

**University of Alberta**

**Metabolite Identification and Quantification in Biofluids Using Liquid  
Chromatography Mass Spectrometry**

by

**Avalyn Elsbeth Stanislaus**

A thesis submitted to the Faculty of Graduate Studies and Research  
in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy**

**Department of Chemistry**

©Avalyn Elsbeth Stanislaus  
Fall 2012  
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

*To my husband, family and friends who believed I could fly*

## **Abstract**

Metabolomics, one of the branches of systems biology, is the comprehensive measurement of all the endogenous metabolites in a biological system. It can be applied in the areas of biomarker discovery and diagnosis of critical diseases and disorders, and can significantly increase our understanding of the pathophysiology involved in development of acute diseases. For metabolomics to reach its true potential, there are many key areas that need to be addressed. This thesis focuses on the development and application of new mass spectrometric techniques aimed at improving several areas in metabolomics research: 1) detection and identification of compound classes, 2) enhancing the electrospray response and improving the chromatographic behavior of metabolites, 3) increasing sample throughput, 4) accurate quantification of endogenous metabolites and 5) expanding current databases used in metabolite identification. The initial work, the comprehensive analysis of acylglycines in human urine, illustrates the analytical challenges of detection and identification of trace levels of metabolites in urine. Mass spectrometric methods were optimized using fragmentation patterns and breakdown graphs, and putative identifications were made using a combination of diagnostic neutral losses. This method was successful in detecting and putatively identifying 43 new acylglycines. Quantification of these metabolites in human urine and plasma was performed using a new labeling strategy. Chemical modification was advantageous as it enhanced the electrospray response of these polar metabolites and provided a way to incorporate a stable isotope onto the molecule. Quantification using a new

“surrogate matrix” strategy was also developed. A new fast LC method was developed for the analysis of several analytes in the vitamin B<sub>12</sub> pathway in less than 36 seconds. This method was very useful in improving high-throughput and in the discovery of the role of ABC protein transporters in B<sub>12</sub> metabolism. Finally, a web-based software called MyCompoundID was developed to expand the metabolome coverage of an existing database (HMDB) by generating metabolic products derived from 76 common biotransformations. This served to increase the number of entries from 8000 to more than 10 million.

## Acknowledgements

This thesis would not have been possible without the help of many people whose contribution deserves special mention and it is my pleasure to convey my gratitude to them all.

I am especially indebted to my supervisor, Dr. Liang Li, for his invaluable advice, supervision, guidance and encouragement throughout my research and the opportunity to contribute to his laboratory. Above all, I appreciate all his contributions of time, ideas and funding that made my PhD experience productive and enjoyable.

I would like to gratefully acknowledge the other members of my supervisory committee, Dr. James Harynuk, Dr. Dennis Hall and the members of my examining committee, Dr. John Klassen and Dr. Jonathan Curtis for their time, interest, insightful questions and helpful comments on my research. I am also grateful to them for taking the time to thoroughly review this thesis and offer knowledgeable and critical comments. I have also benefited from the advice and guidance offered by Dr. Harynuk, Dr. Klassen and Dr. Charles Lucy, who kindly granted me their time to answer my many questions about scientific topics related to my research. I would also like to thank Dr. Karen Waldron for agreeing to participate in my PhD defense and I look forward to her invaluable counsel.

It is a pleasure to pay tribute to the people with whom I collaborated. To Dr. Roy Gravel and Megan McDonald in the Department of Biochemistry and Molecular Biology at the University of Calgary, I would like to thank you for the collaborative work on Vitamin B12 deficiency (Chapter 5). To Dr. Guohui Lin and his students Jianjun Zhou and Ronghong Li in the Department of Computer Science at the University of Alberta, I extend my appreciation for the collaborative work on My Compound ID (Chapter 6). Many thanks, in particular, to Dr. David Wishart for allowing me the opportunity to be a part of the Human

Metabolome Database (HMDB) project and for providing funding and acylglycine standards used in my research (Chapters 2-4).

The members of the Li group have contributed greatly to my experience here at the University of Alberta. The group has been a source of friendships as well as helpful discussions and assistance. I especially thank Dr. J. Bryce Young and Melissa Brown for their training on many of the mass spectrometers in our lab. Many thanks to Kevin Guo for his hard work and great contribution to the carboxylic acid labeling strategy used in Chapters 3 and 4 of this thesis. I am eternally grateful for two very special friends and colleagues, Azeret Zuniga and Mingguo Xu, who stuck it out in graduate school with me and for all the pleasant times working together. They have offered many opinions and advice about my scientific research and life in general, which I really appreciate. We have had many great laughs and I truly enjoyed our times together, especially our lunches. My time here at U of A has been greatly enriched because of them. Collective and individual acknowledgements also go to past and present group members that I had the pleasure to work: Dr. Nan Wang, Dr. Peng Wang, Dr. Fang Wu, Dr. Andrea De Souza, Dr. Andy Lo, Lu Chen, Yanan Tang, Difei Sun, Jun Peng, Ruokun, Zhou, Jared Curle, Chad Iverson, Feifei Fu, Chiao-Li Tseng, Xiaoxia Ye, Tao Huan, Zhendong Li, Yiman Wu, Jiamin Zheng, Tran Tran and Lara Ebert. I am grateful for their contributions.

To Dr. Randy Whittal in the Mass Spectrometry Facility, thank you for giving me the opportunity to work on several of the mass spectrometers in the facility. I am very grateful to him for sharing some of his expert mass spectrometry knowledge and accurate mass calculation tables with me. Special thanks to Bela R  iz for his technical assistance in analyzing samples on the oa-TOF (Chapter 6). I would also like to thank Jing Zheng for her technical advice.

I must also acknowledge Dr. Sandra Marcus and Dr. Garreth Lambkin for the training and the use of their biological laboratory and the equipment therein.

To Allan Chilton in the Electronics Shop, for his endless patience as he came to the rescue to revive several of our instruments when we needed it most, I am eternally grateful. Special acknowledgements also go to Randy Benson (the pump doctor) in the Machine Shop as he has promptly fixed many leaky vacuum pumps and for doing it with such a great sense of humor. I am very thankful to Paul Crothers and Dieter Starke for patiently offering technical advice.

I would also like to thank the departmental staff in the general office, purchasing office and mail room, especially Laura Pham and Sarah McGee for their kind assistance. I would be remiss if I did not thank the staff in Chemistry Chemical Stores and Receiving, especially Ryan Lewis, Bernie Hippel, Andrew Yeung and Marcel Munroe, who readily provided willing and efficient service with a ready smile. Thanks especially to Ryan who was always willing to go above and beyond to assist in whatever way he can.

Much of my research would not have been possible without the kind assistance of Raymond Lemieux, Glen Thomas and Dr. Eric Flaim of the Integrated Nanosystems Research Facility (INRF) facility at the National Institute of Nanotechnology, who allowed me the opportunity to use the QTRAP™ 4000 mass spectrometer. This instrument was used in Chapters 2, 3, 4, 5 and 6 of this thesis. Special acknowledgements also to Raymond Lemieux for our many animated discussions and for offering advice in scientific research and life in general. I am indebted to Jonathan Clarke, the IT System Administrator at the INRF for patiently reviving my laptop when it fell victim to a virus, while I went into full panic. I am eternally grateful. Thanks also to lab managers Dr. Nancy Zheng and Shiau-Yin Wu for all their assistance.

To Dr. Charles Iden, my employer and mentor at State University of New York, Stony Brook, I would like to thank you for giving me a start in research and for encouraging me to pursue my dreams. Thank you also for the training and lessons in electronics and mass spectrometry.

I am very grateful to the committee at the Association for Mass Spectrometry Applications to the Clinical Lab (MSACL) for awarding me a travel grant and for giving me the opportunity to give an oral presentation at the conference. The experience has been invaluable.

I would also like to acknowledge Genome Alberta and Genome Canada for providing the funding that made this research possible. I would also like to thank the Department of Chemistry at the University of Alberta for providing financial support during my graduate studies. Special thanks to the Graduate Students Association for awarding me a travel grant to assist me in attending the American Society of Mass Spectrometry Conference.

Where would I be without my family? I would like to thank my parents, my grandparents and my stepmother and my siblings. I thank my dad for his constant love, support and confidence in me and instilling the love of learning in me as a child and my mom for her love, support and her prayers. My grandmother has always been mentor, my support, my rock, and my counselor and I will always miss her. May she rest in peace. I appreciate all the members of my family, too many to mention, who contributed in many ways to the successful realization of my thesis.

Many thanks also to my friends in Edmonton, New York, and St. Lucia for their friendships, love, prayers and encouragement.

Nobody has been more important to me in the pursuit of my doctorate than my husband, Elisha Stanislaus. Words cannot express my deep appreciation to him for his love, dedication, support, patience and confidence in me. I am truly grateful for the endless happiness and unending inspiration he has brought me.

Finally, I would like to express my appreciation for all those who were important in making this dream a reality, as well as expressing my apologies that I could not mention you all individually. Thank you.

## Table of Contents

Chapter 1. Introduction .....	1
1.1 Metabolomics .....	1
1.1.1 Terminology.....	1
1.1.2 Sampling .....	4
1.1.3 Measuring the Metabolome .....	4
1.2 Challenges in MS-based Metabolomics.....	11
1.2.1 Sample Preparation .....	11
1.2.2 Matrix Effects .....	12
1.2.3 Method Validation/Biomarker Validation .....	12
1.2.4 Metabolite Identification.....	13
1.3 Liquid Chromatography .....	14
1.3.1 High Performance Liquid Chromatography (HPLC) .....	14
1.3.2 Ultra Performance Liquid Chromatography (UPLC) .....	17
1.3.3 Ultra-High Performance Liquid Chromatography (UHPLC).....	18
1.4 Mass Spectrometry.....	21
1.4.1 Electrospray Ionization .....	21
1.4.2 Matrix Effects .....	25
1.4.3 Triple-Quadrupole Linear Ion Trap (QTRAP®) .....	29
1.4.4 Scan Modes .....	35
1.4.5 Information-Dependent Acquisition (IDA) Scans .....	38
1.4.6 Orthogonal Acceleration – Time-of-Flight (oa-TOF).....	38
1.5 Chemical Derivatization in LC-ESI-MS .....	41
1.6 Metabolite Identification by LC-MS .....	43

1.6.1	Ion Annotation and Accurate Mass Measurements .....	44
1.6.2	Mass Spectral Interpretation and Spectral Matching .....	45
1.7	Quantification of the Metabolome by LC-MS .....	47
1.7.1	Calibration Techniques .....	47
1.7.2	Validation of Quantitative Methods.....	50
1.8	Inborn Errors of Metabolism (IEM) – Study of Acylglycines .....	54
1.9	Overview of Thesis .....	60
1.10	Literature Cited.....	61
 Chapter 2. Ultra-Performance Liquid Chromatography Tandem Mass Spectrometry Detection of Acylglycines .....		
		72
2.1	Introduction .....	72
2.2	Experimental .....	74
2.2.1	Materials and Reagents .....	74
2.2.2	Samples .....	75
2.2.3	Microsome Incubation .....	75
2.2.4	Solid Phase Extraction .....	76
2.2.5	Esterification .....	76
2.2.6	UPLC Separation .....	77
2.2.7	Mass Spectrometry.....	77
2.2.8	Breakdown Graphs and Scan Modes .....	78
2.3	Results and Discussion.....	78
2.3.1	Sample Handling Issues .....	78
2.3.2	MS Fragmentation Pattern of Acylglycine Standards .....	82
2.3.3	MS Scan Modes for Sample Analysis .....	88
2.3.4	Human Urine Analysis.....	92

2.4	Conclusions .....	103
2.5	Literature Cited .....	103
Chapter 3. Quantification of Acylglycines in Human Urine using Derivatization and Liquid Chromatography Tandem Mass Spectrometry .....		
107		
3.1	Introduction .....	107
3.2	Experimental .....	111
3.2.1	Materials and Reagents .....	111
3.2.2	Synthesis of <sup>13</sup> C <sub>2</sub> - p-dimethylaminophenacyl bromide.....	112
3.2.3	Derivatization Reaction Optimization .....	113
3.2.4	Urine Samples.....	113
3.2.5	Creatinine Determination.....	114
3.2.6	Preparation and Derivatization of Standards and Samples .....	114
3.2.7	Ultra-high Performance Liquid Chromatography (UHPLC).....	115
3.2.8	Mass Spectrometry.....	115
3.2.9	Method Validation .....	116
3.2.10	Measurement of Endogenous Levels of Acylglycines in Human Urine	121
3.3	Results and Discussion.....	122
3.3.1	Optimization of Derivatization Reaction Conditions .....	122
3.3.2	Effects of Derivatization on Chromatography .....	130
3.3.3	Mass Analysis .....	130
3.3.4	Enhancement of the ESI Signal .....	136
3.3.5	Selection of Surrogate Matrix.....	138
3.3.6	Method Validation .....	138
3.3.7	Measurement of Endogenous Levels of Acylglycines.....	155
3.4	Conclusion.....	161

3.5	Literature Cited .....	162
Chapter 4. Quantification of Acylglycines in Human Plasma using Derivatization and Liquid Chromatography Tandem Mass Spectrometry.....		
4.1	Introduction .....	165
4.2	Experimental .....	167
4.2.1	Chemical and Reagents.....	167
4.2.2	Plasma Samples .....	168
4.2.3	Preparation and Derivatization of Standards and Samples .....	168
4.2.4	Ultra-high Performance Liquid Chromatography.....	169
4.2.5	Mass Spectrometry.....	170
4.2.6	Method Validation .....	170
4.3	Results and Discussion.....	175
4.3.1	Surrogate Matrix Strategy.....	175
4.3.2	Optimization of the LC-MS/MS Method.....	175
4.3.3	Method Validation .....	179
4.3.4	Measurement of endogenous acylglycines .....	191
4.4	Conclusion.....	195
4.5	Literature Cited .....	197
Chapter 5. Development of a Fast Liquid Chromatography for the Analysis of Homocysteine, Succinic Acid and Methylmalonic Acid in <i>C. elegans</i> Worm Media .....		
5.1	Introduction .....	200
5.2	Experimental .....	206
5.2.1	Chemicals and Materials.....	206
5.2.2	Preparation of Samples .....	206
5.2.3	Preparation of <i>C. elegans</i> RNAi-fed Strains.....	207

5.2.4	Preparation of L1 larvae for <i>C. elegans</i> Deletion Strains .....	207
5.2.5	Methylmalonic Acid Accumulation Assay .....	208
5.2.6	<sup>14</sup> C-propionate Incorporation Assay .....	208
5.2.7	Preparation of Standards .....	208
5.2.8	Ultra-Performance Liquid Chromatography (UPLC).....	209
5.2.9	Mass Spectrometry.....	209
5.3	Results and Discussion.....	210
5.3.1	Mass Analysis .....	210
5.3.2	Chromatographic Separation .....	212
5.3.3	Flow Rate Optimization for Fast LC .....	212
5.3.4	Effect of MRM Dwell Time .....	216
5.3.5	Succinic Acid and Methylmalonic Acid Analysis .....	217
5.3.6	Calibration and Validation.....	219
5.3.7	Analysis of Worm Incubation Media.....	221
5.3.8	Analysis of <i>C. Elegans</i> RNAi Samples.....	225
5.3.9	Analysis of ABC Transporter Deletion Strains .....	226
5.4	Conclusion.....	229
5.5	Literature Cited .....	229
Chapter 6. MyCompoundID: Using an Evidenced-Based Metabolome Library for Metabolite Identification.....		234
6.1	Introduction .....	234
6.2	Experimental .....	235
6.2.1	Chemicals and Materials.....	235
6.2.2	Construction of the EML .....	235
6.2.3	Web-service Design and Implementation.....	238

6.2.4	Web Interface .....	239
6.2.5	Samples .....	241
6.2.6	Solid Phase Extraction (SPE).....	241
6.2.7	Liquid Chromatography Mass Spectrometry (LC-MS).....	242
6.2.8	Data Processing.....	243
6.3	Results and Discussion.....	244
6.4	Conclusions .....	254
6.5	Literature Cited .....	255
Chapter 7. Conclusions and Future Work.....		257
7.1	Conclusions .....	257
7.2	Future Work .....	260
7.2.1	Biomarker Discovery .....	260
7.2.2	Dried Urine Spots .....	261
7.2.3	Microwave-assisted Extraction and Derivatization .....	262
7.3	Literature Cited .....	266
Appendix.....		267

## List of Tables

Table 1-1	Definition of terminology used in metabolite studies.....	3
Table 1-2	Chemical diversity of metabolites in biological fluids .....	5
Table 1-3	A list of the advantages and disadvantages of UPLC over HPLC.....	18
Table 1-4	Scan modes used in the QTRAP® 4000 (Ref. 119).....	34
Table 1-5	Acylglycine profile in metabolic disorders. Deficient enzyme in brackets. .....	58
Table 2-1	Comparison of HLB and MAX solid phase extraction cartridges .....	81
Table 2-2	MRM scanning conditions (positive ion mode).....	91
Table 2-3	A list of confirmed and putatively identified acylglycines present in urine of six healthy volunteers .....	97
Table 3-1	List of conditions used in the optimization experiments .....	124
Table 3-2	Masses, abbreviations and retention times of protonated DmPA-labeled acylglycines .....	131
Table 3-3	MRM masses and parameters .....	134
Table 3-4	Internal Standard suitability. Comparison of linearity of <sup>12</sup> C and <sup>13</sup> C- labeled analytes.....	142
Table 3-5	Comparison of slopes in surrogate and authentic matrices and calibration parameters of calibration curve in surrogate urine.....	145
Table 3-6	Lower limits of quantification (LLOQ) measurements in surrogate urine (n=5) .....	146
Table 3-7	Comparison of the standard addition method and the surrogate matrix approach.....	147
Table 3-8	Dilution linearity.....	149
Table 3-9	Matrix Effects study. Comparison of slopes in solvent to slopes in authentic urine using a modified t-test.....	150
Table 3-10	Intra-day, inter-day and derivatization precision of the proposed method .....	152
Table 3-11	Accuracy data obtained using the proposed method.....	153
Table 3-12	Stability results of various storage conditions .....	154

Table 3-13	Concentration ranges of acylglycines in healthy urine .....	156
Table 3-14	Concentration ranges for putatively identified acylglycines in healthy urine .....	161
Table 4-1	Masses, abbreviations and retention times of DmPA-labeled acylglycines .....	177
Table 4-2	Internal Standard Suitability. Comparison of <sup>12</sup> C and <sup>13</sup> C linearity in solvent and surrogate matrix.....	180
Table 4-3	Comparison of slopes in surrogate and authentic matrices and calibration parameters of calibration curve in surrogate plasma .....	183
Table 4-4	Lower limits of quantification (LLOQ) measurements of the proposed method .....	184
Table 4-5	Comparison of standard addition method and surrogate matrix approach .....	185
Table 4-6	Matrix Effects study. Comparison of slopes in solvent to slopes in authentic plasma using a modified t-test.....	187
Table 4-7	Extraction recovery of deuterated acylglycines in plasma.....	188
Table 4-8	Intra-day and inter-day precision measurements of the proposed method .....	189
Table 4-9	Accuracy measurements for acylglycines in plasma using the proposed method .....	190
Table 4-10	Stability results for various storage conditions .....	192
Table 4-11	Endogenous concentrations of acylglycines in plasma. Concentration expressed as μmol/mol creatinine.....	193
Table 5-1	Instrument parameters and MRM transitions used in the polarity switching method.....	210
Table 6-1	List of the 76 most common metabolic reactions (biotransformations)	236
Table 6-2	Results from database searching of urine metabolite features using reaction = 0. MS/MS spectra of unknowns was matched with the HMDB spectra .....	246
Table 6-3	Results from database searching of plasma metabolite features using reaction = 0. MS/MS spectra of unknowns was matched with the HMDB spectra .....	248

## List of Figures

Figure 1-1	Overview of the “omics” cascade and systems biology ..... 2
Figure 1-2	A typical LC-MS/MS workflow of a metabolomics study. This figure is adapted from ref. 37..... 7
Figure 1-3	van Deemter plots (A) showing the relationship between theoretical plate height and linear velocity of the mobile phase, (B) showing the curves for columns packed with different particle sizes. .... 16
Figure 1-4	Typical LC analysis workflow: (A) Sequential workflow using a single binary pump and a single column, (B) Alternating column regeneration workflow using two binary pumps and two identical columns. .... 20
Figure 1-5	Basic components of an electrospray ionization source ..... 22
Figure 1-6	Schematic of the electrospray ionization process in positive ion mode.. 24
Figure 1-7	Schematic of possible mechanisms of ion suppression in electrospray ionization. Adapted from ref.109..... 27
Figure 1-8	Schematic of a quadrupole mass analyzer ..... 30
Figure 1-9	Schematic of the LIT mass analyzer, with axial ion ejection, used in the QTRAP® 4000 ..... 31
Figure 1-10	Schematic of the AB Sciex QTRAP® 4000 hybrid triple quadrupole linear ion trap instrument..... 33
Figure 1-11	Triple quadrupole scan modes ..... 36
Figure 1-12	Schematics of an orthogonal acceleration time-of-flight mass analyzer 39
Figure 1-13	Standard addition curve ..... 48
Figure 1-14	Fatty acid oxidation by the beta-oxidation pathway ..... 56
Figure 1-15	Pathways of branched-chain amino acids. Enzymes: (1) isovaleryl-coenzyme A-CoA dehydrogenase, (2) 3-methylcrotonyl-CoA carboxylase, (3) 3-hydroxy-3-methylglutaryl-CoA lyase, (4) 2-methylbutyryl-CoA-dehydrogenase, (5) 2-methyl3-hydroxybutyryl-CoA-dehydrogenase, (6) 2-methylacetoacetyl-CoA thiolase, (7) propionyl-CoA carboxylase. Black bars indicate enzyme deficiencies and the corresponding disorders can be found in Table 1-5..... 57
Figure 2-1	Structures of acylglycines: (A) dimethylglycine; (B) acetylglycine; (C) propionylglycine; (D) isobutyrylglycine; (E) butyrylglycine; (F) tiglylglycine; (G) 3-methylcrotonylglycine; (H) 2-methylbutyrylglycine; (I) isovalerylglycine; (J) valerylglycine; (K) phenylglycine; (L) hexanoylglycine; (M) octanoylglycine; (N) phenylpropionylglycine; (O)

	phenylacetyl-glycine; (P) 4-hydroxyphenylacetyl-glycine; (Q) hippuric acid (R) glutaryl-glycine and (S) suberyl-glycine .....	80
Figure 2-2	Cartridge reproducibility of the MAX cartridges .....	81
Figure 2-3	Comparison of HPLC and UPLC separation of C <sub>5</sub> -glycine isomers. Identity of peaks are (1) 2-methylbutyryl-glycine; (2) isovaleryl-glycine and (3) valeryl-glycine. ....	83
Figure 2-4	Proposed fragmentation pathway of acyl-glycines. ....	84
Figure 2-5	Proposed fragmentation pathway of glycine-conjugated di-carboxylic acids .....	85
Figure 2-6	Breakdown graphs of (A) valeryl-glycine; (B) tiglyl-glycine; (C) phenylacetyl-glycine and (D) phenylpropionyl-glycine.....	87
Figure 2-7	Ion chromatograms from the UPLC-MS/MS analysis of urine from a healthy volunteer: (A) TIC of precursor ion scanning of <i>m/z</i> 74 (negative mode); (B) TIC of precursor ion scanning of <i>m/z</i> 76 (positive mode); and (C) TIC of neutral loss scanning of 75 Da.....	90
Figure 2-8	Ion chromatograms from UPLC-MS/MS analysis of six healthy human urine samples. Urine samples were collected as first morning void.....	92
Figure 2-9	Comparison of XIC of <i>m/z</i> 224 (hydroxyphenylpropionyl-glycine) in (A) microsome extract and (C) urine, and corresponding product ion spectra, (B) microsomes and (D) urine. The inset in (D) shows the fragmentation pattern of hydroxyphenylpropionyl-glycine.....	94
Figure 2-10	Putative identification of C <sub>6</sub> :1-glycine (position of double bond unknown), (A) XIC of C <sub>6</sub> :1-glycine isomer, <i>m/z</i> 172, (B) MS/MS spectrum of peak at 22.78 minutes, (C) XIC of C <sub>6</sub> :1-glycine methyl ester, <i>m/z</i> 186, (D) MS/MS spectrum of peak at 29.00 minutes.....	95
Figure 2-11	An overlay of total ion chromatograms of urine samples collected as first morning void from a healthy individual for five consecutive days. ....	101
Figure 2-12	An overlay of total ion chromatograms of urine samples collected from a healthy individual 4 hours after a meal and collected over a period of 12 hours. ....	102
Figure 3-1	Derivatization reaction of DmPA and acyl-glycines.....	109
Figure 3-2	Effect of reagent catalyst on the reaction.....	125
Figure 3-3	Effect of reagent concentration on the reaction .....	126
Figure 3-4	Effect of solvent conditions on the reaction .....	127
Figure 3-5	Effect of temperature on the reaction.....	128

Figure 3-6	Effect of reaction time on the reaction.....	129
Figure 3-7	(A) Extracted ion chromatograms of analyte $^{12}\text{C}_2$ -DmPA-labeled hexanoylglycine (HG) and internal standard $^{13}\text{C}_2$ -DmPA-labeled hexanoylglycine (HG). (B) Product ion spectrum of $^{12}\text{C}_2$ -DmPA-HG showing the structure and fragmentation. (C) Product ion spectrum of $^{13}\text{C}_2$ -DmPA-HG showing the structure and fragmentation. Fragments in bold italics are the fragments of the DmPA tag. Fragments $m/z$ 134 and $m/z$ 136 are shown as insets in B and C respectively.....	132
Figure 3-8	Signal enhancement using the DmPA tag. Acylglycines are grouped by peak area ranges. Numbers in bold above the bars are the fold increase of derivatized to underivatized.....	137
Figure 3-9	Representative LC-MS/MS chromatograms of (A) solvent blank, (B) derivatization blank, (C) blank after running ULOQ calibration standard ( $n=5$ ), (D) surrogate blank, (E) internal standards (100 nM) in authentic matrix, (F) analytes (1 nM) in surrogate matrix and (G) internal standards in surrogate matrix. Compounds: 1, acetylglycine; 2, propionylglycine; 3, isobutyrylglycine; 4, butyrylglycine; 5, 4-hydroxyphenylacetylglycine; 6, 2-furoylglycine; 7, tiglylglycine; 8, 2-methylbutyrylglycine; 9, 3-methylcrotonylglycine; 10, isovalerylglycine; 11, valerylglycine; 12, hexanoylglycine; 13, phenylacetylglycine; 14, phenylpropionylglycine; 15, glutarylglycine; 16, heptanoylglycine; 17, octanoylglycine; 18, suberylglycine.....	140
Figure 3-10	Influence on gender on the acylglycine excretion pattern ( $n=10$ for both males and females).....	158
Figure 3-11	Influence of BMI on the acylglycine excretion profile.....	159
Figure 4-1	Positive-ion ESI-MS/MS product ion spectrum of DmPA-labeled-IVG: (A) Overlay of $^{12}\text{C}_2$ -DmPA- and $^{13}\text{C}_2$ -DmPA-labeled-IVG; (B) Product ion spectrum of $^{12}\text{C}_2$ -DmPA-labeled-IVG; (C) Product ion spectrum of $^{13}\text{C}_2$ -DmPA-labeled-IVG. ....	178
Figure 4-2	Selected reaction monitoring chromatograms of acylglycines (left panel) and corresponding internal standards (right panel). (A) surrogate plasma, (B) surrogate plasma spiked with 1.0 nM acylglycines and internal standards, (C) authentic plasma spiked with 100 nM internal standards and showing endogenous levels of acylglycines. ....	182
Figure 4-3	Influence of gender on the acylglycine plasma profile. Red dotted line indicates the lower limit of quantification of the analyte in plasma. ....	196
Figure 5-1	Structures of cobalamin and the enzyme cofactors.....	201
Figure 5-2	Homocysteine pathway. Homocysteine is methylated to form methionine by a folate-dependent reaction. 5-Methyltetrahydrofolate is the methyl donor for the methylation of homocysteine. The enzyme methionine	

	synthase using methylcobalamin as a co-factor catalyzes this conversion. .....	203
Figure 5-3	Methylmalonic pathway. Propionate is a product in the catabolic pathway of odd number fatty acids, cholesterol and the amino acids isoleucine, valine, methionine and threonine. Propionate is then carboxylated to methylmalonic acid, which is isomerized to succinic acid. The isomerization is catalyzed by the enzyme methylmalonyl-CoA mutase and the co-factor adenosylcobalamin. ....	204
Figure 5-4	Fragmentation and product ion spectra of (A) Homocysteine, (B) Succinic acid and (C) Methylmalonic acid. (Inset in (B) shows the structure of fragment ion $m/z$ 99).....	211
Figure 5-5	Total ion chromatogram showing the retention of (1) Homocysteine, (2) Succinic acid and (3) Methylmalonic acid.....	213
Figure 5-6	Effect of flow rate on the chromatographic separation of peaks (1) Hcy, (2) SA and (3) MMA. The flow rates studied are (A) 0.150 mL/min, (B) 0.200 mL/min, (C) 0.250 mL/min, (D) 0.300 mL/min, (E) 0.400 mL/min, and (F) 0.500 mL/min. ....	215
Figure 5-7	Effect of flow rate on resolution between SA and MMA (peaks 1 and 2). Dotted line at $R_s=1$ .....	216
Figure 5-8	Effect of dwell times on the number of points across a chromatographic peak. Chromatogram (A) 50 ms, (B) 100 ms, (C) 150 ms and (D) 200 ms.....	218
Figure 5-9	TIC showing the retention of (1) SA and SA-d <sub>4</sub> , (2) MMA and MMA-d <sub>3</sub> .....	219
Figure 5-10	Calibration curve of MMA .....	220
Figure 5-11	Chromatograms showing the analysis of MMA and MMA-d <sub>3</sub> in PBS media.....	222
Figure 5-12	Chromatograms showing the analysis of MMA and MMA-d <sub>3</sub> in Incubation media.....	223
Figure 5-13	Chromatograms showing the analysis of MMA and MMA-d <sub>3</sub> in CeHR media.....	224
Figure 5-14	Methylmalonic acid detected in the supernatant after RNAi <i>mmcm-1</i> in the presence of 200 mM propionate.....	225
Figure 5-15	Methylmalonic acid accumulation of <i>C. elegans</i> ABC transporter deletion strains .....	227
Figure 6-1	Experimental Workflow .....	240

Figure 6-2	Waters Oasis <sup>®</sup> HLB chemistry.....	250
Figure 6-3	Structures of diagnostic fragment ions for (A) PEG derivatives, (B) monomethoxy-PEG derivatives and (C) cocodiethanolamides .....	253
Figure 7-1	Experimental workflow of urine analysis .....	264
Figure 7-2	Comparison of the liquid-liquid extraction and dried urine spot extraction from (A) Spiked urine sample, final concentration 50 $\mu$ M and (B) Unspiked urine sample.....	265

## List of Abbreviations

ABC	ATP-binding cassette
ACN	Acetonitrile
AdoCbl	Adenosylcobalamin
APCI	Atmospheric pressure chemical ionization
ATP	Adenosine triphosphate
BALF	Broncho-alveolar lavage fluid
BEH	Bridged ethylene hybrid
BMI	Body mass index
CAD	Collision activated dissociation
Cbl	Cobalamin
CE	Capillary electrophoresis
CeHR	<i>C. elegans</i> habitation and reproduction media
CID	Collision induced dissociation
CoA	Coenzyme A
CRM	Charged residue model
CSF	Cerebral spinal fluid
CV	Coefficient of variation
DBS	Dried blood spot
DC	Direct current
DDW	Double distilled water

DmPA	p-Dimethylaminophenacyl
DNA	Deoxyribonucleic acid
DUS	Dried Urine Spot
EMS	Enhanced mass scan
EPI	Enhanced product ion
ER	Enhanced resolution
ESI	Electrospray ionization
ETF	Electron transferring flavoprotein
ETF-DH	Electron transferring flavoprotein dehydrogenase
FAB	Fast atom bombardment
FDA	Food and Drug Administration
FTICR	Fourier transform ion cyclotron resonance
GC-MS	Gas chromatography mass spectrometry
GS1	Gas Source 1
GS2	Gas Source 2
H	Plate height
HB101	Hybrid K12 x B strain bacterium
Hcy	Homocysteine
HILIC	Hydrophilic interaction liquid chromatography
HLB	Hydrophilic-lipophilic balanced
HLM	Human liver microsome

HMDB	Human metabolome database
HPLC	High performance liquid chromatography
IDA	Information dependent acquisition
IEM	Inborn errors of metabolism
IEM	Ion evaporation model
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
IQ	Interquad lens
IR	Infrared
IS	Internal standard
KOH	Potassium hydroxide
LB	Lysogeny broth
LC	Liquid chromatography
LIT	Linear ion trap
LLOQ	Lower limit of quantification
LOD	Limit of detection
MALDI	Matrix-assisted laser desorption ionization
MAX	Mixed mode anion exchange
MCI	MyCompoundID
ME	Matrix Effects
MMA	Methylmalonic acid
MRM	Multiple reaction monitoring

mRNA	messenger RNA
MRP	Multidrug resistance protein
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NADPH	Nicotinamide adenine dinucleotide phosphate
NBS	Newborn Screening
NGM	Nematode growth medium
NIST	National Institute of Standards and Technology
NL	Constant neutral loss scan
NMR	Nuclear magnetic resonance
OA	Organic acidurias/organic acidemias
PBS	Phosphate buffered saline
PC	Precursor ion scan
PGP	P-glycoprotein
Q1	Quadrupole 1
Q3	Quadrupole 3
QC	Quality control
QTRAP	Quadrupole linear ion trap mass spectrometer
R <sup>2</sup>	Coefficient of determination
RE	Relative error
RF	Radio frequency

RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
ROD	Reversed osmosis deionized
RP	Reversed-phase
S/N	Signal to noise ratio
SA	Succinic acid
SIL-IS	Stable-isotope labeled internal standard
SIM	Single ion monitoring
S <sub>N</sub> 2	Bimolecular nucleophilic substitution
SPE	Solid phase extraction
SRM	Selected reaction monitoring
TC	Transcobalamin
TCA	Trichloroacetic acid
TEA	Triethylamine
TEOA	Triethanolamine
TFA	Trifluoroacetic acid
TIC	Total ion chromatogram
TMS	Trimethylsilyl
TOF	Time-of-flight
UHPLC	Ultra-high Performance Liquid Chromatography
ULOQ	Upper limit of quantification

UPLC      Ultra Performance Liquid Chromatography

UV        Ultraviolet

XIC        Extracted ion chromatogram

# Chapter 1. Introduction

## 1.1 Metabolomics

Metabolomics, defined as the study of all metabolites in a biological system, has been emerging as a new field of study in analytical biochemistry<sup>1, 2</sup> and is considered to be the final stage of the “omics” cascade (see Figure 1-1). In genomics, the complete genome is sequenced in order to understand the function of single genes. In the study of transcriptomics, the analysis of gene expression is studied through mRNA, and comprehensive protein analysis or proteomics is studied in an attempt to understand the function of all proteins. Metabolites are the end products of enzymes and cellular processes. When combined with genomics, transcriptomics and proteomics, metabolomics can provide additional understanding and elucidation of many complex biological processes. Together, these four “omics” form a collective approach known as systems biology. Thus the information flows from gene to transcript to protein to metabolite. Metabolomics is therefore an excellent tool for diagnosing diseases, disorder or for studying the effects of drugs or toxicants.

### 1.1.1 Terminology

Metabolomics, as we know it, started in the early 1970s when Pauling *et al.* used gas-liquid chromatography to analyze the content of urine and breath.<sup>3</sup> The name metabolome was first coined in the late 1990s by Oliver *et al.* in the context of functional genomics<sup>4</sup> and later used by Tweeddale *et al.* in studying *E. coli* metabolism.<sup>5</sup> The term “metabolome” was defined by these two groups as the total complement of metabolites in a cell in a particular physiological or developmental state. Since then, there has been some debate in the research community over the exact definitions of the terms used in the field of metabolomics and these terms are constantly evolving.

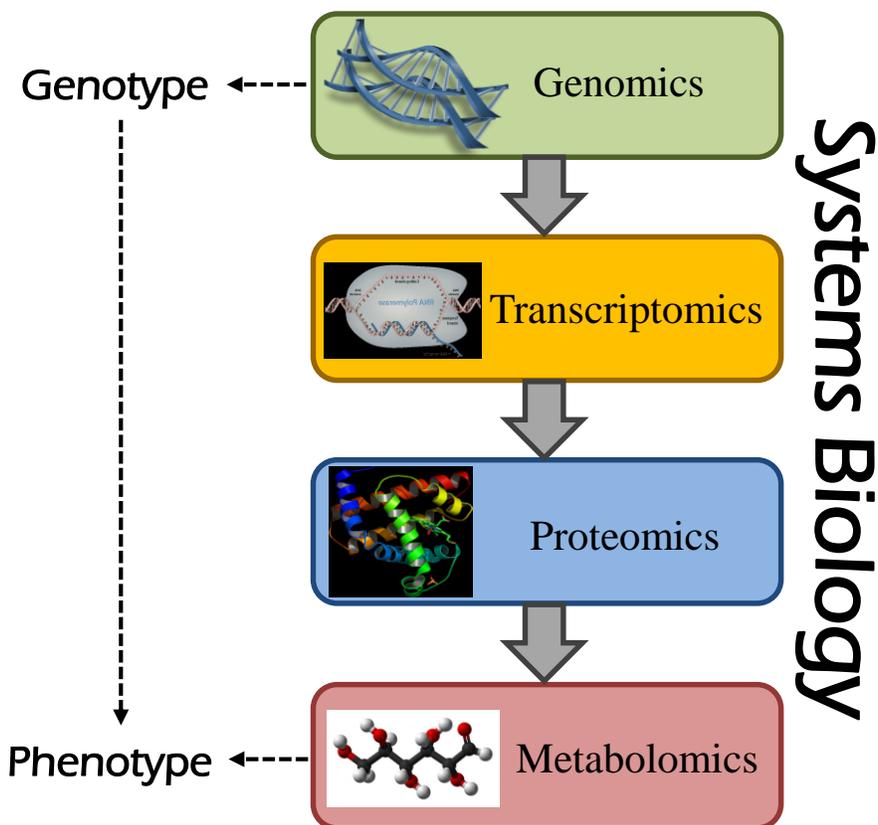


Figure 1-1 Overview of the “omics” cascade and systems biology

One of the major questions is whether to use the term “metabolomics” or “metabonomics” when describing the global measurement of small molecules in any biological system. Fiehn<sup>6</sup> first gave a more detailed definition of the term metabolomics and Nicholson<sup>7</sup> first introduced the term metabonomics. Both terms overlap to a great degree, especially when studying human diseases.<sup>8</sup> Table 1-1 lists the definitions of the terms used in metabolomics studies, initially by Fiehn and Nicholson, and adapted by other authors. Because of the various definitions found in the literature today, describing approaches used in small molecule research can be quite confusing. Despite the confusion, the term metabolomics is used most commonly and two distinct analytical approaches are used most often in many research laboratories: non-targeted and targeted approaches. The non-targeted approach focuses on the qualitative and quantitative analysis of all small

molecules in a system within a certain mass range. It is, in general, a step above metabolite fingerprinting and is often referred to as metabolite profiling.<sup>9</sup> In the targeted approach, efforts tend to center on the analysis of selected or individual classes of compounds,<sup>9</sup> ranging in number from dozens to hundreds of metabolites.

Table 1-1 Definition of terminology used in metabolite studies

<b>Term</b>	<b>Definition</b>
Metabolite	Bio-active small molecule, with molecular mass less than 1000 Da, which is involved in biological networks and pathways. <sup>10</sup>
Metabolome	Total complement of all metabolites present in a biological system. <sup>10-12</sup>
Metabolomics	Non-biased identification, quantitative and comprehensive analysis of the metabolome under certain conditions. <sup>6, 13, 14</sup>
Metabonomics	Quantitative measurement of the dynamic multi-parametric metabolic response of living systems to genetic modification, physiological stimuli or biochemical perturbations caused by disease, drugs or toxins. <sup>7</sup>
Metabolite profiling	Quantitative analysis of groups of metabolites that are associated with a particular biological pathway or belong to a certain class of compounds, e.g. carbohydrates, lipids, organic acids. <sup>6, 11</sup>
Metabolite fingerprinting	A global, high-throughput analysis of the metabolome to provide sample discrimination and classification based on their biological status or origin. <sup>11, 15</sup>
Targeted analysis	Quantitative analysis restricted to one or several metabolites related to a specific reaction or a particular enzyme system. <sup>6, 11</sup>

### 1.1.2 Sampling

The most popular types of samples used in metabolomics studies are urine, plasma, serum, saliva, tissue and cerebral spinal fluid (CSF). Breath and broncho-alveolar lavage fluid (BALF) analyses are also fairly common. The most common tissue samples used are liver, brain and muscle samples. Urine, plasma, serum and saliva are non-invasively obtained and are therefore more readily available, making them the sample matrices of choice.

### 1.1.3 Measuring the Metabolome

The metabolome, being the downstream product of the genome, reflects the functional level of the cell and changes to the metabolome are expected to be greatly amplified relative to the genome, transcriptome and the proteome. Measuring the metabolome is also reflective of environmental stresses on the system. The metabolome consists of compounds that vary widely in chemical structures and properties<sup>16</sup> (summarized in Table 1-2). The heterogeneity of metabolites makes it virtually impossible to determine the complete metabolome using one analytical technique. To add to that, many of the metabolites are present in a wide range of concentrations and there are very high fluxes throughout the metabolic pools. The main challenges are therefore the chemical complexity, the dynamic range of the instrument used and the extraction techniques. Ideally, the analysis technique should be unbiased, able to efficiently separate and detect each individual metabolite and capable of identifying and quantifying the metabolites.

The expanding field of metabolomics has benefited from advances in the analytical techniques of high resolution nuclear magnetic resonance (NMR)<sup>17-21</sup> and mass spectrometry (MS).<sup>22-26</sup> NMR spectroscopy generates valuable structural information of metabolites, generally from biological fluids requiring little or no sample preparation. Compared with MS, NMR methods are lower in sensitivity, generally with limits of detection in the nanomolar to micromolar range. However, NMR is highly quantitative, reproducible, and its sensitivity does

not depend on metabolite chemical properties such as hydrophobicity or  $pK_a$ . These qualities are important when doing comprehensive metabolome measurements using multivariate statistical methods.

Table 1-2 Chemical diversity of metabolites in biological fluids

<b>Chemical Classes of Metabolites in Biological Fluids</b>			
Acyl phosphates	Biotin & derivatives	Indoles & derivatives	Porphyrins
Acylglycines	Carbohydrates	Inorganic ions	Prostanoids
Alcohol phosphates	Carnitines	Ketoacids	Pterins
Alcohols & polyols	Catecholamines & derivatives	Ketones	Purines & derivatives
Aldehydes	Cobalamin & derivatives	Leukotrienes	Pyridoxals & derivatives
Alkanes/Alkenes	Coenzyme A & derivatives	Lipoamides & derivatives	Pyrimidines & derivatives
Amino acid phosphates	Cyclic amines	Minerals & elements	Quinones & derivatives
Amino acids	Dicarboxylic acids	Nucleosides	Retinoids
Amino alcohols	Fatty acids	Nucleotides	Sphingolipids
Amino ketones	Glycerolipids	Peptides	Steroids & derivatives
Aromatic acids	Glycolipids	Polyamines	Sugar phosphates

The high sensitivity of MS detection, with detection limits in the high femtogram to low picogram range, makes it a very useful tool for measuring metabolites, especially in complex biological systems. Combining MS with a separation technique can offer further information on the chemical properties of metabolites and provide separation of isobaric species whilst reducing the

complexity of the mass spectra. Current popular methods include mass spectrometry coupled to gas chromatography (GC)<sup>27, 28</sup>, liquid chromatography (LC)<sup>29-33</sup> or capillary electrophoresis (CE)<sup>34-36</sup>. Since GC analysis is limited to volatile compounds, non-volatile, polar metabolites often need to be derivatized before they can be separated on a GC column. LC-MS plays an important role in quantitative metabolome analysis,<sup>37</sup> with LC evolving to include sub 2  $\mu\text{m}$  particle columns in ultra-performance liquid chromatography (UPLC). CE is particularly suited to separating polar, charged metabolites on the basis of charge-to-mass ratio. Because the separation mechanism of CE is different from LC, it can provide complementary information on the metabolome of a biological sample. Based on the coupling technique and ionization methods used, there may be a bias towards or against certain compound classes. Therefore, analytical platforms, for example GC-MS and LC-MS should be used in parallel in order to comprehensively study the metabolome. Whether GC-MS, LC-MS or CE-MS is used, comprehensive metabolomic analyses comprise many steps. For both the non-targeted and targeted approaches, the workflow is similar: an extraction step, followed by derivatization if needed, subsequent separation of the analytes and then analysis by mass spectrometry and data analysis. However, in the non-targeted more global approach, in order to quantify or even detect all the compounds in an unbiased way, integration of data from various techniques and instrumentation is needed. Merging all the data can be quite a challenge, because although many data analysis packages exist for organizing meta-data on genes and proteins, these are still lacking for metabolites. Another difficult task is the identification of unknown metabolites that can be detected using this methodology. Although there are disadvantages with this approach, one major advantage is that this method enables the discovery of novel biomarkers. A typical LC-MS-based metabolomics workflow is represented in Figure 1-2. It begins with a non-targeted approach to screen potential metabolites of interest that were significantly altered when comparing biological groups. A targeted

approach is then applied to these metabolites for identification, verification, quantification and biological interpretation.

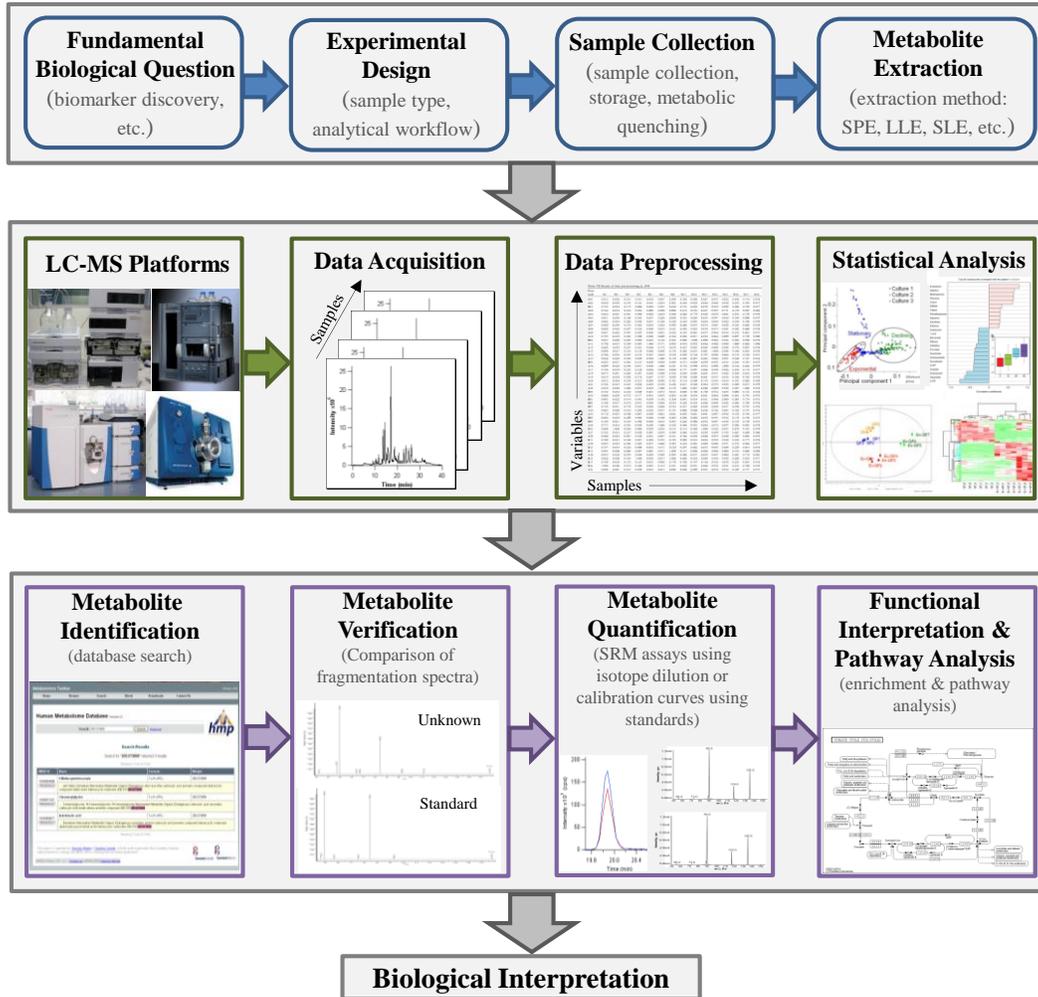


Figure 1-2 A typical LC-MS/MS workflow of a metabolomics study. This figure is adapted from ref. 37

### 1.1.3.1 NMR-based Metabolomics

NMR spectroscopy is based on the net absorption of energy in the radiofrequency region of the spectrum by the nuclei of elements that have spin angular momentum and a magnetic moment. The absorption, resonant frequencies and other spectral features for the nuclei of an element provide useful information on the identity and molecular structure. Chemical shifts in the NMR spectrum depend on the effect of shielding by the electrons orbiting the nucleus and signal intensity depends on the number of identical nuclei. Active nuclei used for the study of biological samples include  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^{31}\text{P}$ .<sup>38</sup> The most widely used is  $^1\text{H}$  NMR, and as the majority of metabolites possess hydrogen atoms, the technique is non-biased. Sensitivity is based on the natural abundance of the atom studied, the strength of the magnetic field, analysis time and use of cryogenic probes.

In metabolomics, NMR spectroscopy is often used as a high throughput fingerprinting technique.<sup>39, 40</sup> Crude samples can be simply dissolved in solvent ( $\text{D}_2\text{O}$  for  $^1\text{H}$  NMR), placed into a probe, inserted into the instrument and analyzed. An alternative method is flow injection NMR, which sequentially loads each sample into the NMR spectrometer and analyzes each sample.<sup>41</sup> NMR has been used extensively in many applications of small molecule analysis, namely clinical<sup>42, 43</sup> and pharmaceutical analysis,<sup>44-47</sup> toxicology studies,<sup>48, 49</sup> plant chemistry,<sup>50-52</sup> disease biomarkers, and even the study of intact cells and tissues.<sup>53, 54</sup> The main obstacles in biological studies are water suppression and signal contribution from protein and other high molecular weight compounds. A range of pulse sequences can serve to improve spectral information and minimize the effects of large amounts of other components. This includes the application of water suppression pulse sequences to minimize peak broadening,<sup>55</sup> and the use of spin-echo and diffusion edited sequences to enhance detection of small molecules and macromolecules.<sup>56</sup>

The advantages of NMR in metabolomics are many<sup>38</sup>: it can provide both qualitative and quantitative information, is non-selective, non-destructive, requires little sample preparation, needs no chemical derivatization, is useful for analysis of intact tissues, and is highly amenable to the analysis of compounds not suitable for LC-MS or GC-MS analyses. Despite its disadvantages, namely sensitivity and complex spectra containing thousands of signals from biological samples, NMR can provide complementary information to MS techniques and is an invaluable tool in metabolite analysis.

### **1.1.3.2 GC-MS based Metabolomics**

The application of GC-MS in metabolite analysis has been in use for decades.<sup>57</sup> GC is however limited to the analysis of molecules that are volatile, thermally stable and have a low molecular weight. Volatile, small molecules can be sampled, prepared and subsequently analyzed directly. However, many of the compounds in metabolomic studies are polar, non-volatile compounds (amino acids, organic acids, alcohols, sugars, amines and many others) and require chemical derivatization. Because of the wide range in chemical functional groups of metabolites, derivatization is done in two stages.<sup>57, 58</sup> First, carbonyl groups are converted to oximes using O-alkylhydroxylamine solutions to eliminate a slow and reversible reaction of silylation reagents with the carbonyl group. Second, trimethylsilyl (TMS) esters are formed using silylating reagents to replace the exchangeable protons in the molecules. Therefore the polar active hydrogens are replaced by the less polar TMS group that increases volatility by reducing dipole-dipole interactions. Silylation is a reversible reaction in the presence of water and so samples must be fully dry before reaction. However, extensive sample drying can result in a loss of volatile metabolites, for example some ketones, aldehydes and lactic acid. Other derivatization techniques exist for other functional groups in targeted analysis.<sup>27, 59</sup> Sample stability depends on sample and environmental conditions and care must be taken to ensure no sample degradation. Metabolite

identification is performed by matching retention time or retention index and mass spectrum with that of a pure compound analyzed under the same conditions. Mass spectral searches against commercially available libraries such as NIST are very useful but these databases do not contain all the metabolites that are expected to be found in metabolic pathways. The mass spectrometer employed can depend on the type of application. Quadrupoles are useful for single ion monitoring (SIM), which enhances sensitivity, time-of-flight (TOF) instruments are valuable for full mass scans with high mass accuracy while ion traps are highly sensitive instruments.

As GC-MS is such a mature analytical technique, its application is wide ranging in metabolomics. In the pharmaceutical field, GC-MS was applied before HPLC-MS for targeted drug and metabolite profiling.<sup>60, 61</sup> More recent applications include profiling urine and blood samples for organic acid<sup>62, 63</sup> and acylglycines<sup>64, 65</sup> in detecting the inborn errors of metabolism of organic acidemias and fatty acid oxidation disorders. Other applications include plant metabolomics,<sup>6, 66-68</sup> clinical<sup>69</sup> and microbial metabolomics,<sup>70, 71</sup> and many others.

Biological samples, being very complex, can contain hundreds to thousands of metabolites and no single GC or HPLC experiment can separate all these compounds. Such a separation can be successfully performed using 2D-chromatography in which orthogonal mechanisms of separation are employed either by using different column chemistries or different mobile phases. Comprehensive 2D-GC (GC x GC), first introduced in 1991 by Liu and Phillips,<sup>72</sup> is an online method where, described in simple terms, compounds elute from the first column, are trapped and sampled periodically by a modulator, which then injects the eluates of the first column onto the second column at regular intervals. The two columns must be orthogonal in selectivity and fractions are separated quickly on the second column before the elution of subsequent fractions from the first column (typically 2 – 6 seconds). Such a technique is an invaluable tool in metabolomics and has application in plasma metabolome<sup>73</sup>, biomedical studies,<sup>28</sup> plant metabolomics<sup>74</sup> and yeast metabolomics.<sup>75, 76</sup>

### **1.1.3.3 CE-MS-based Metabolomics**

CE-MS is a sensitive analytical technique that is suitable for analyzing charged, neutral, and polar compounds; therefore it is a useful tool in both targeted and untargeted metabolomics.<sup>35</sup> However, it has not been used as extensively in metabolomics studies as HPLC and GC techniques. CE offers several advantages over HPLC, namely higher resolution, faster speeds, smaller sample volumes in the nanoliter range and very little organic solvent usage.<sup>35</sup> Because of the low sample volumes used, sensitivity is low when using UV/Vis as the mode of detection. However, with the coupling of CE-MS, nanospray ionization can be used for enhanced sensitivity.<sup>77</sup> Recently, there has been an increase in the use of CE-MS in targeted and non-targeted approaches for clinical and biomedical applications,<sup>78-80</sup> plant<sup>81</sup> and bacterial<sup>82</sup> metabolomics and metabolic fingerprinting.<sup>83</sup>

## **1.2 Challenges in MS-based Metabolomics**

Current challenges being faced in MS-based metabolomics include optimized sample handling and preparation, reducing matrix effects that account for irreproducible signals, certain compound classes that are more amenable to different ionization techniques, identification of unknown metabolites, validated quantification methods and standardization of the methodologies.

### **1.2.1 Sample Preparation**

Due to heterogeneity in the physical and chemical nature of metabolites, there is no one extraction technique that can extract “all” metabolites. Therefore extraction conditions need to be optimized to extract as many metabolites as possible.<sup>84</sup> Concentration ranges also make sample preparation difficult. With the advent of more sensitive mass spectrometers, some samples have to be diluted

because of the high concentration of some of the metabolites, for example hippuric acid and citric acid in urine. On the other hand, most metabolites of interest are present in much lower concentrations, so finding an appropriate sample volume/concentration often poses quite a challenge.

### **1.2.2 Matrix Effects**

Matrix effects (ME) play a major role in atmospheric pressure ionization mass spectrometry. ME are defined as the influence exerted by components of the matrix in which the analyte is present which affect the signal being measured.<sup>85</sup> More on the subject of ME is discussed in Section 1.4.2. ME become a problem in metabolomics if two or more co-eluting compounds have different surface activities or proton affinities, especially if one of the compounds is in high concentration. The analysis is, therefore, biased, with each analyte having the potential to influence the response of other co-eluting analytes. In order to achieve accurate and precise results, ME must be eliminated or minimized. This can be accomplished by improving sample clean-up, optimizing sample preparation, chromatographic conditions, mass spectrometry parameters, and using internal standards, in particular, stable-isotope labeled internal standards (SIL-IS).<sup>86</sup> The use of SIL-IS is considered ideal because the standard possesses identical physical and chemical properties with the analyte and can therefore co-elute with the analyte in LC-MS. It is assumed that the SIL-IS and the analyte are subject to identical ME, thus eliminating the effects of ME on the accuracy and precision of the measurements.

### **1.2.3 Method Validation/Biomarker Validation**

While method validation experiments for exogenous compounds in drug analysis, for example, are regulated by FDA guidelines,<sup>87</sup> there are no such regulations for methods measuring endogenous compounds. There is a need for

an adopted procedure guideline for regulation of quantification of endogenous metabolites. The main challenge is that there are no “true” blanks where the analytes of interest are absent, making validation more difficult. Regulation should include definition of the sample blanks to be used for calibrators and quality control. More on this topic will be discussed in Section 1.7.2.

#### 1.2.4 Metabolite Identification

One of the major bottlenecks in metabolomics currently is the identification of metabolites of interest. The general approach for metabolite identification is as follows: 1) statistical methods are used to choose  $m/z$  values that can distinguish diseased samples from control samples; 2) the  $m/z$  value of the analyte of interest is searched against an accurate mass database, within a specific mass range and these query  $m/z$  values are retrieved as putative identifications; 3) tandem MS scans are performed to generate fragmentation spectra; 4) the MS/MS spectrum is compared to hypothetical spectra (through *in silico* methods) or to previously obtained spectra of compounds compiled into a spectral library; 5) retention times,  $m/z$  values and fragmentation data are used to elucidate structural information.<sup>88</sup> Challenges exist at each step of the identification process. Searches of just one  $m/z$  value can result in hundreds of putative identifications and matching the retention time of MS/MS spectra of each of the searches can be very time-consuming and arduous. Successful identification of metabolites requires high quality MS/MS data. Differences in instrumentation type and experimental conditions such as collision energies, often affect the quality of the MS/MS spectra. *In silico* fragmentation prediction can be based on fragmentation rules or by a combinatorial disconnect of chemical bonds based on bond energies. Missing fragmentation rules in the program can cause missed or incorrect matches, thus generated matches need to be checked with caution.<sup>88</sup> Matching spectra using spectral libraries can also be problematic because the data are usually generated from different instruments and there is

generally limited coverage of the existing spectral libraries. Information on previously identified metabolites are often scattered in many different literature sources and analyzed under many different conditions, making comparison and identification very difficult. More information on metabolite identification can be found in Section 1.6.

### **1.3 Liquid Chromatography**

Most analytical platforms in metabolomics studies involve the use of a separation technique before performing mass spectrometric analyses. The most common method used is the coupling of high performance liquid chromatography (HPLC) to mass spectrometry. HPLC allows the separation of a wide range of compounds ranging in polarity and molecular masses. Good chromatographic separation can reduce sample complexity, lower background noise, minimize matrix effects during ionization and improve sensitivity in detection.

#### **1.3.1 High Performance Liquid Chromatography (HPLC)**

In HPLC-MS, reversed-phase columns (most commonly C<sub>18</sub> columns) are often the columns of choice in small molecule analysis. Reversed-phase columns are very efficient at separating a range of non-polar to semi-polar compounds such as alkaloids, steroids, etc. More recently, the emergence of more polar columns like hydrophilic interaction liquid chromatography (HILIC) allows the separation of polar compounds, for example amino acids, sugars, carboxylic acids, etc. HILIC separation, although quite similar to normal-phase separation, uses polar organic mobile phases that are more compatible with ESI-MS.<sup>89</sup>

With the sheer number and complexity of biological samples used in metabolomics studies, many researchers are searching for more efficient approaches to the separation as well as “fast” LC separation to speed up the

analyses. Measures to improve efficiency as well as speed include using shorter columns, employing faster flow rates and increased temperatures.<sup>90</sup> However, in HPLC, columns with particle sizes of 3-5  $\mu\text{m}$  are generally used and limitations in efficiency, resolution and pressure are soon reached. This is illustrated in the van Deemter equation,<sup>91, 92</sup> which is an empirical formula describing the relationship between linear velocity (flow rate) and plate height (H).

$$H = A + B/u + Cu$$

where A, B and C are constant for a particular solute, column and experimental conditions and u is the linear velocity of the mobile phase. A plot of the van Deemter curve (plate height vs. linear velocity) can be observed in Figure 1-3 (A). The “A” term represents eddy diffusion and it does not vary with flow rate. It is dependent on the arrangement and size of the particles in the column. The “B” term represents longitudinal diffusion, which is time dependent. This effect decreases at higher flow rates because the analyte spends less time on the column. The “C” term which is mass transfer between and within the mobile phase and stationary phase is flow dependent. As the flow rate of the mobile phase increases, the more an analyte molecule interacting with the stationary phase will tend to lag behind the analyte molecules in the faster moving mobile phase. This results in incomplete transfer and causes band broadening. A plot of the combination of all three terms is known as the van Deemter plot (Figure 1-3 (A)). Since the C term is proportional to particle size, a van Deemter curve can be used to examine and compare chromatographic performance between different columns. However, as illustrated by Figure 1-3 (B), as the column particle size decreases to below 2  $\mu\text{m}$ , there is a significant increase in efficiency (and thus resolution), which is not affected by flow rate. By decreasing particle size, efficient fast LC separations can be achieved, without sacrificing resolution. This technique became known as Ultra Performance Liquid Chromatography (UPLC).

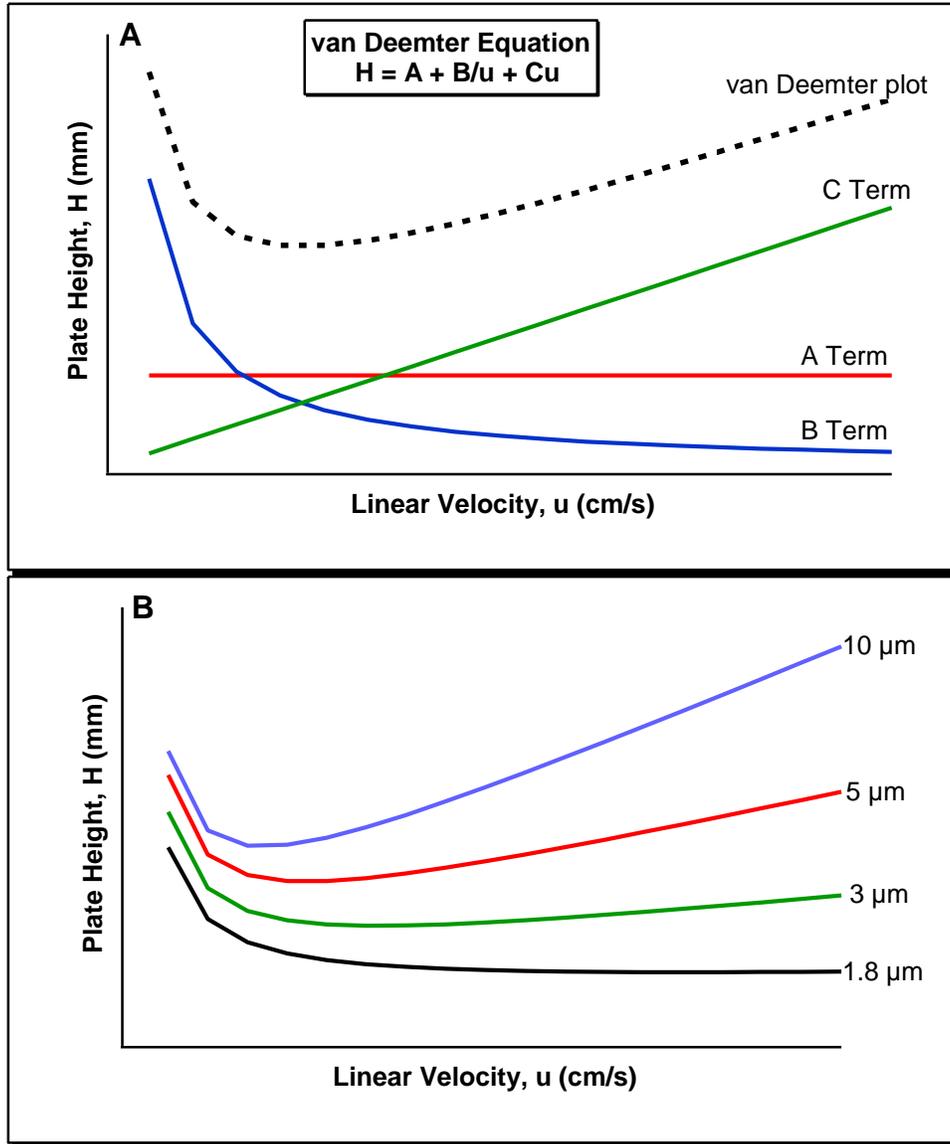


Figure 1-3 van Deemter plots (A) showing the relationship between theoretical plate height and linear velocity of the mobile phase, (B) showing the curves for columns packed with different particle sizes.

### 1.3.2 Ultra Performance Liquid Chromatography (UPLC)

In order for UPLC to be successful and fully take advantage of the speed, resolution and sensitivity offered by sub 2  $\mu\text{m}$  particles, the instrumentation must be capable of handling the speed and high pressure requirements. To minimize the effects of frictional heating, which occurs at high pressures, columns with smaller diameters, typically in the range of 1.0 – 2.1 mm are used for UPLC. With particle sizes of 1.7  $\mu\text{m}$ , peak widths (at half height) of 1 second or even less can be achieved. Therefore, the detector sampling rate must be high enough to obtain enough data points across the peak to integrate the peak accurately and reproducibly. One of the first commercially available UPLC system to successfully meet these requirements was the Waters Acquity UPLC™ System, which became available in 2004.

As mentioned above, mass spectrometry coupled to liquid chromatography is a commonly accepted analytical tool for qualitative and quantitative analysis of metabolites. In order to be compatible with the very narrow peaks produced by UPLC, MS instrumentation with faster scan rates for improved peak definition has been developed. The use of UPLC-MS as a tool for the analysis of metabolites has been fully investigated and many papers have been published on the topic.<sup>93-95</sup> In many cases, the technique has been particularly useful in the separation of metabolite isomers,<sup>96</sup> that are otherwise undifferentiated in HPLC-MS. Using UPLC, one can take full advantage of the principles of chromatography to do analyses using shorter column and higher flow rates and still maintain superior efficiency, resolution and sensitivity. The advantages of using UPLC are many and are listed in Table 1-3, along with some disadvantages.<sup>92</sup> In analyzing the complex mixtures found in metabolomics, UPLC can provide a great improvement over HPLC in that many more analytes can be detected.

Table 1-3 A list of the advantages and disadvantages of UPLC over HPLC

<b>Advantages</b>	<b>Disadvantages</b>
Better resolution	Higher cost of instruments, parts and columns
Higher separation efficiency	Detector and data collection system may not be able to cope with sharper peaks
Faster chromatography	Only binary pump systems available
Better sensitivity or signal to noise ratio (sharper and taller peaks)	Number of stationary phases still limited compared to HPLC.
Less solvent usage	
Increased sample throughput	
Lower ion suppression	

### 1.3.3 Ultra-High Performance Liquid Chromatography (UHPLC)

The terms UPLC and UHPLC are often confused and used interchangeably. UPLC is a trademark of Waters Corporation and is used to describe their instrumentation, but the generic term used to describe the technique is UHPLC (Ultra High Performance Liquid Chromatography). In 2003, one year before the Waters Acquity UPLC™ system debuted, Agilent produced the 1200 series UHPLC system that used 1.8 µm particle columns. This section briefly describes the UHPLC system that was used for metabolite quantification studies in this thesis work. The Agilent 1290 UHPLC system was equipped with two binary pumps, along with an autosampler and a thermostatted column compartment.<sup>97</sup> Usually, in a typical LC run, after injection, the analytes are separated in a gradient run (typically 10 – 30 minutes in metabolomics), followed by a column wash to remove strongly bound hydrophobic compounds and then the column is equilibrated to initial mobile phase conditions. These steps are performed sequentially before the next analysis can begin. Generally the column

wash and equilibration steps can often take up to 50% of the analysis time (and sometimes can be even longer than the gradient run in “fast” LC analyses). Figure 1-4 (A) illustrates a sequential workflow in a typical LC run using one binary pump system. By using two identical columns and a second binary gradient pump, a procedure known as alternating column regeneration can be employed. This workflow is demonstrated in Figure 1-4 (B). In this procedure, the two columns can be switched between the eluent (or analytical) pump and the regeneration pump, using a 2-position valve. The column wash step and the equilibration step can be performed in column 2 by the regeneration pump simultaneously, while the analytical run is carried out in column 1 by the eluent pump. This procedure can reduce the cycle time by as much as 50%, improving sample through-put significantly. The 1-minute rinse procedure is used to rinse the volume between the eluent pump and the switching valve port. The mobile phase at the end of the previous analytical run is usually a high percentage organic solvent and that remains in the volume between the eluent pump and the switching valve port, and if pumped onto the equilibrated column in the subsequent run it can cause unpredictable chromatographic results.<sup>98</sup> It is therefore advantageous to rinse out this volume of high organic solvent with the mobile phase composition used in the equilibration step to ensure reproducible chromatographic separation. The alternating column regeneration procedure was used to develop the UHPLC methods in Chapters 3 and 4 of this thesis.

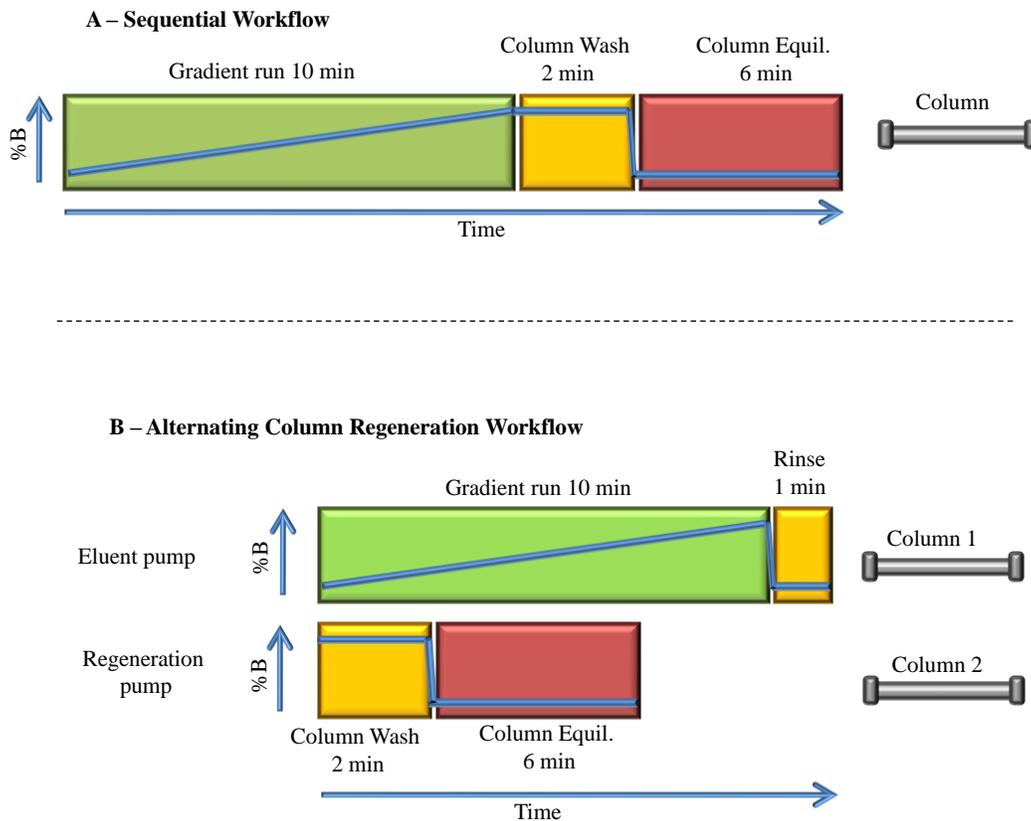


Figure 1-4 Typical LC analysis workflow: (A) Sequential workflow using a single binary pump and a single column, (B) Alternating column regeneration workflow using two binary pumps and two identical columns.

## 1.4 Mass Spectrometry

There has been great progress in mass spectrometry for metabolomics in recent years, with the number of annual publications exceeding NMR-based metabolomics publications. In coupling LC and CE to a mass spectrometer, ionization may be achieved using electrospray (ESI) or atmospheric pressure chemical ionization (APCI). Choices in mass spectrometry analyzers include quadrupoles and ion traps, which offer excellent sensitivity but lower resolution, time-of-flight (TOF) instruments, which offer higher resolution and high speeds, orbitrap and Fourier transform ion cyclotron resonance (FTICR), which offer very high resolution and mass accuracy, typically  $< 1$  ppm. Similar analyzers can be arranged in a tandem-in-space configuration for example, triple quadrupole mass spectrometer or different analyzers can be combined to form a hybrid instrument for e.g. quadrupole-TOF, ion trap-orbitrap and quadrupole-linear ion trap. Current platforms that involve the use of high resolution or newer hybrid mass spectrometers, capable of tandem MS measurements, are becoming more popular. These instruments allow for improved structural elucidations and rapid quantification of many metabolites in a wide dynamic concentration range.

The focus of this thesis is the study of metabolites using liquid chromatography coupled to electrospray ionization mass spectrometry. The mass spectrometers used extensively in this study were a hybrid triple quadrupole-linear ion trap and a time-of flight instrument. The following sections will focus primarily on the aforementioned ionization technique and mass spectrometers.

### 1.4.1 Electrospray Ionization

As the name suggests, electrospray is a process whereby a fine spray of charged droplets is produced under the influence of an electric field. A simplified view of the components of an electrospray source is illustrated in Figure 1-5

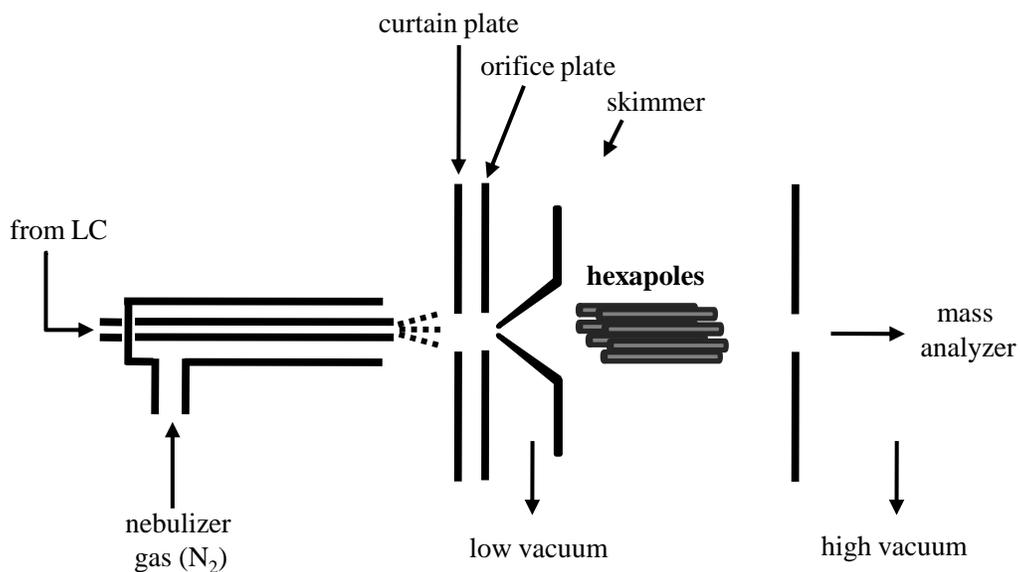


Figure 1-5 Basic components of an electrospray ionization source

A dilute sample solution, usually the effluent of an LC run, is pumped continuously through a stainless steel capillary at a low flow rate (0.1 – 10  $\mu\text{L}/\text{min}$ ). For stable operation at higher flow rates, a heated gas is used to assist in the fine droplet formation and the desolvation process. The tip of the capillary is held at a high voltage, generally 2-5 kV, with respect to a metal plate or counter-electrode. This potential is sufficiently high to generate a fine mist of charged droplets from the flowing solution. The solvent in these droplets then evaporate converting the charged droplet to gas-phase ions. These ions are then transported from the atmospheric pressure region of the instrument to the high vacuum region of the analyzers via a series of pressure-reduction stages.

Despite the widespread use of ESI-MS, the mechanism of ESI is still not fully understood and is a highly debated topic. It is proposed that there are four main steps in the production of gas-phase ions from solution.<sup>99-101</sup>

1) Charged droplet formation at the tip of the capillary. A detailed diagram of the droplet formation is given in Figure 1-6. The voltage applied to the capillary can be either positive or negative, depending on the application and the analytes of interest. For the purpose of simplicity, only positive ionization will be explained further and illustrated in Figure 1-6. As a voltage of about 2-5 kV is applied to the capillary and the capillary is about 1 mm in diameter and located 1 – 3 cm from the counter-electrode, the resulting electric field is very high (about  $10^6$  V/m). The electric field, which is essentially inversely proportional to the capillary outer radius, will penetrate the solution and will be highest near the capillary tip. In the positive mode, cations concentrate at the capillary tip and migrate towards the counter-electrode. Anions migrate away from the tip. This will cause an enrichment of cations at the surface of the liquid meniscus. The migration of ions resisted by the surface tension of the liquid causes a distortion of the meniscus forming a Taylor cone, which is shown in Figure 1-6. Under a sufficiently high electric field, the tip of the cone becomes unstable and breaks into a fine jet, which emerges at the tip. The jet breaks into small charged droplets due to repulsive charges at the surface. As the positively charged droplets are carried off, there is a requirement for charge balance. In the positive ion mode, oxidation occurs in the solution at the metal surface and reduction occurs at the counter-electrode.

2) Evaporation of charged droplets. The initial charge and size of the droplets are influenced by a number of factors, namely applied potential, solvent flow rate, capillary diameter, and solvent characteristics. Evaporation of the solvent is carried out using a flow of a dry gas, usually nitrogen, which may be heated.

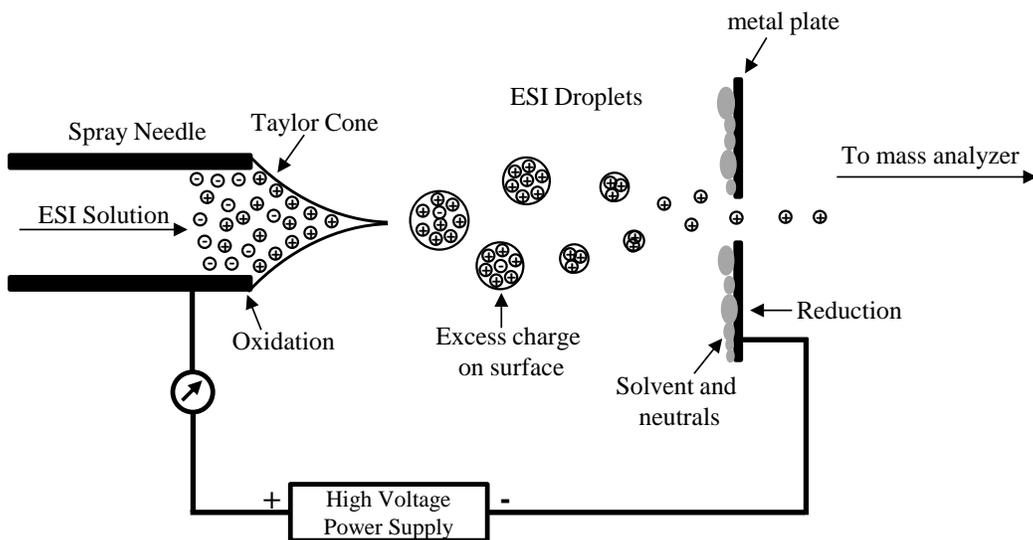


Figure 1-6 Schematic of the electrospray ionization process in positive ion mode

3) Droplet shrinkage and fission. The charged droplets shrink due to solvent evaporation, while the charge remains constant. This leads to an increase in the charge density on the surface. At a certain radius, the Rayleigh instability limit is reached where the repulsive forces of the charges exceed the droplet surface tension. Further evaporation leads to further instability, called Coulomb fission, which leads to fission of the droplet into small charged progeny droplets. The resulting fission is uneven with the offspring droplets possessing 20% of the parents' mass and 15% of the charge. About 20 progeny droplets are created with each repeated fission.<sup>99</sup>

4) Formation of gas-phase ions. Two mechanisms have been proposed to explain the formation of gas-phase ions from the charged droplets (now very small and very highly charged): the charged residue model (CRM) and the ion evaporation model (IEM). CRM was first proposed by Dole in 1968.<sup>102</sup> According

to this model, which was used for the analysis of macromolecules, solvent evaporation and droplet fission occurs repeatedly until the droplet size is so small that it only contains one analyte molecule. As the solvent completely evaporates, it leads to a gas-phase analyte ion, which retains the charge of the droplet. It is believed that the model applies to proteins and other macromolecules. IEM was first proposed by Iribarne and Thomson in 1976.<sup>103</sup> The IEM proposal relies on the repeated evaporation and fission process. Ion expulsion from the droplet occurs at a droplet radius of about 10 nm, when the electric field is sufficiently high but less than the Rayleigh limit. The expelled ion is repelled by the other charges on the droplet (Coulomb repulsion) but can still be attracted to the droplet by polarization at a close distance. At further distances from the droplet, the repulsive force overcomes the attractive forces and will facilitate the escape of the ion into the gas phase. The process is called ion evaporation and replaces Coulomb fission. It is assumed that the evaporating ion is one of the ions with significant surface activity. IEM applies to small organic and inorganic ions.

The ESI response of an analyte can depend on the chemical nature of the analyte, presence and concentration of electrolytes in the liquid,<sup>104</sup> volatility of the solvent<sup>99</sup>, surface activity in the droplet,<sup>105</sup> presence of non-volatile components,<sup>104</sup> flow rate of the LC eluent,<sup>106</sup> concentration of other ionizable species<sup>105</sup> and competition of gas-phase ion transfer reactions between analytes and other ions.<sup>99, 107</sup>

#### **1.4.2 Matrix Effects**

Matrix effects are considered the “Achilles heel” of electrospray ionization. According to the FDA guidelines for bioanalytical validation<sup>87</sup>, matrix effects can be defined as “the direct or indirect alteration or interference in response due to the presence of unintended analytes (for analysis) or other interfering substances in the sample.” This definition includes effects due to ion enhancements or ion suppression. Matrix effects can cause significant errors in

reproducibility, accuracy and precision when there is a differential suppression or enhancement between and within calibration samples and biological samples. The mechanism involved in matrix effects is not fully understood, simply because it is related to the mechanism of the electrospray process, which is also not fully established. It is well known that one of the most important factors that influences ESI response is the physical and chemical nature of the analyte. In general the basicity, charge density and hydrophobicity of the analyte affect the response – basicity guarantees protonation, while charge density and non-polarity determine how likely the analyte ions are to stay on the surface of the droplet.<sup>108</sup> Surface activity is closely related to the hydrophobicity of the analyte; higher hydrophobicity generally leads to higher surface activity.

Since ion suppression is more common and problematic than ion enhancement, only suppression will be discussed further in this section. There are several causes and mechanisms for ion suppression, as shown schematically in Figure 1-7.

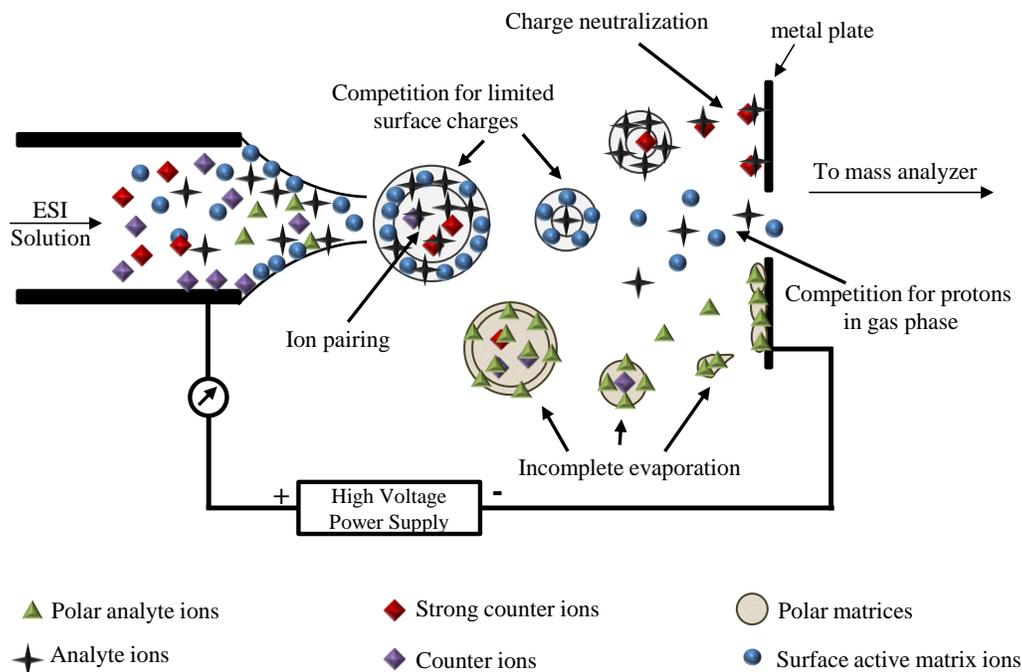


Figure 1-7 Schematic of possible mechanisms of ion suppression in electrospray ionization. Adapted from ref.109

The first possible mechanism to explain ion suppression of the analyte can be competition for limited surface excess charge.<sup>108</sup> The excess charge on an ESI droplet is fixed and controlled by the applied voltage and the flow rate. In addition, because excess charges all reside on the droplet surface, competition for charge and surface space both occur. Matrix components with higher surface activity and charge density (for example surfactants<sup>105</sup>, polymers<sup>109</sup> and lipophilic components) are expected to out-compete analytes with lower surface activity and charge. In the fission process, these matrix components will occupy the droplet surface in each subsequent offspring droplet. Analytes will then be left in the interior of the preceding droplet and remain undetected.

Another possible mechanism is incomplete evaporation of the droplet.<sup>108</sup> Ions generated in the droplet can only be detected once they are ejected into the gas phase so droplet evaporation is essential. Nonvolatile components in the

droplet can alter the viscosity, volatility and conductivity of the solution. This can cause incomplete evaporation and affect Taylor cone formation, resulting in a hindrance to the fission and lowering the efficiency of gas phase ion formation.<sup>104</sup> The number of analytes converted to the gas phase is therefore reduced and thus the ESI response is decreased.

Many ions can undergo reactions in the gas phase after they are emitted from the droplet due to the fact that their gas phase basicities can be different from their solution phase basicities. Gas phase reactions such as charge neutralization, charge stripping and charge transfer can occur and have a significant effect on the ESI response.<sup>107</sup> If the matrix components are stronger gas phase bases, they will compete with the analyte gas phase ions for protons and suppress the ESI response of the analytes.<sup>110</sup> Ion suppression can also be caused by strong acids such as trifluoroacetic acid (TFA), which causes strong ion pairing with basic analytes.<sup>111</sup> This ion pairing process keeps the analytes in the more neutral interior of the droplet and prevents the analyte from partitioning into the surface phase; therefore the analyte cannot be released into the gas phase.

Matrix effects can come from exogenous sources, for example, plasticizers from plastic vials and anticoagulants such as Li-heparin,<sup>112</sup> or from endogenous sources, and these can range from polar components that elute early on a reversed-phase column to hydrophobic components such as lipids. There are several strategies for overcoming matrix effects and they include: (a) decreasing the injection volume, which decreases the amount of matrix that enters the ion source at the same time as the analyte,<sup>113</sup> (b) using a divert valve, which diverts the unwanted early portion of the LC run, (c) avoiding exogenous matrices by using the same brand of tubes for processing and storing standards and samples, (d) use of stable isotope-labeled internal standards, which is based on the rationale that the internal standard will experience the same matrix effects as the analyte,<sup>113</sup> (e) by using the same matrix as the samples to prepare calibration standards (blank matrix) to eliminate the matrix difference between the calibration standards and the samples,<sup>113</sup> (f) by using extensive sample clean-up and purification to

separate the analytes from the matrices,<sup>114, 115</sup> (g) by employing the standard addition method to correct for matrix effects, if sample preparation and using an appropriate internal standard is too time-consuming or difficult,<sup>116</sup> and finally (h) by reducing the ESI flow rate to nL/min, which is well known for increasing desolvation, ionization and ion-transfer efficiency.<sup>117</sup>

### 1.4.3 Triple-Quadrupole Linear Ion Trap (QTRAP®)

The most important role of a quadrupole, historically, has been as a mass analyzer. The quadrupole<sup>101</sup> is constructed from four parallel rods in which a two-dimensional hyperbolic field is established (see Figure 1-8). The rods are generally 4-10 mm in diameter and 10-30 cm long. An applied potential of

$$(U + V \cos (\omega t))$$

is applied to two opposite rods and a potential of

$$- (U + V \cos (\omega t))$$

is applied to the other two opposite rods, with the RF components of the rods being 180° out of phase. Simply put, the filtering role of the quadrupole is achieved by the application of a combination of time-independent DC and a time-dependent AC voltage. For a given DC and AC voltage, only ions of a certain  $m/z$  value have stable paths, all other ions are not transmitted. In practice, the DC potential is held at a fixed fraction of the RF potential. A mass spectrum is obtained by sweeping the RF and DC voltages in a systematic way to allow ions of increasing or decreasing  $m/z$  values to reach the detector. In the RF-only mode, the DC component is set to zero and ions of a wide range of  $m/z$  values have stable trajectories and are transmitted through the quadrupole.

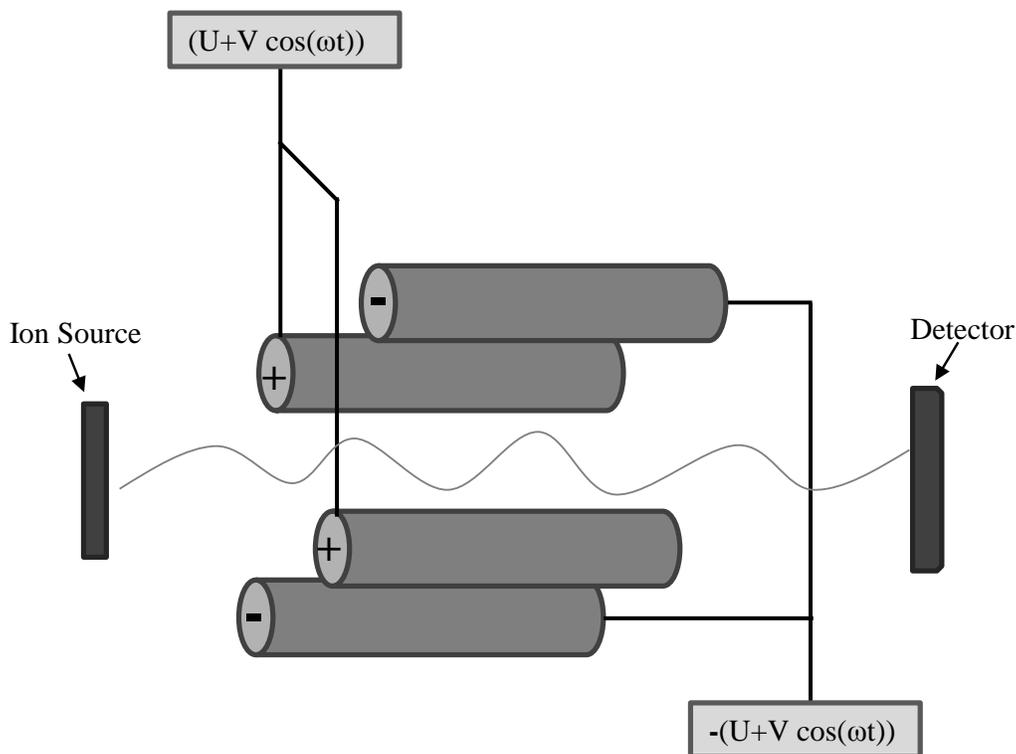


Figure 1-8 Schematic of a quadrupole mass analyzer

The advantages of a quadrupole mass filter are many. It is mechanically simple, can produce spectra with a linear mass range, has high scan speeds and transmission, is not dependent on the initial energy distribution of ions and is well suited for selected reaction monitoring. It is also more tolerant of high pressures than a TOF instrument, which is an advantage in the hybrid QTRAP<sup>®</sup> instrument.

One of the newer additions to the family of mass spectrometers is the quadrupole linear ion trap (LIT).<sup>101</sup> The construction of the LIT analyzer is fundamentally identical to the quadrupole mass analyzer. It consists of four parallel rods and uses RF and DC voltages for mass analysis. The rods are arranged in the x/y plane and two entrance and exit aperture lenses in the z-axis. A schematic of a LIT analyzer is represented in Figure 1-9.

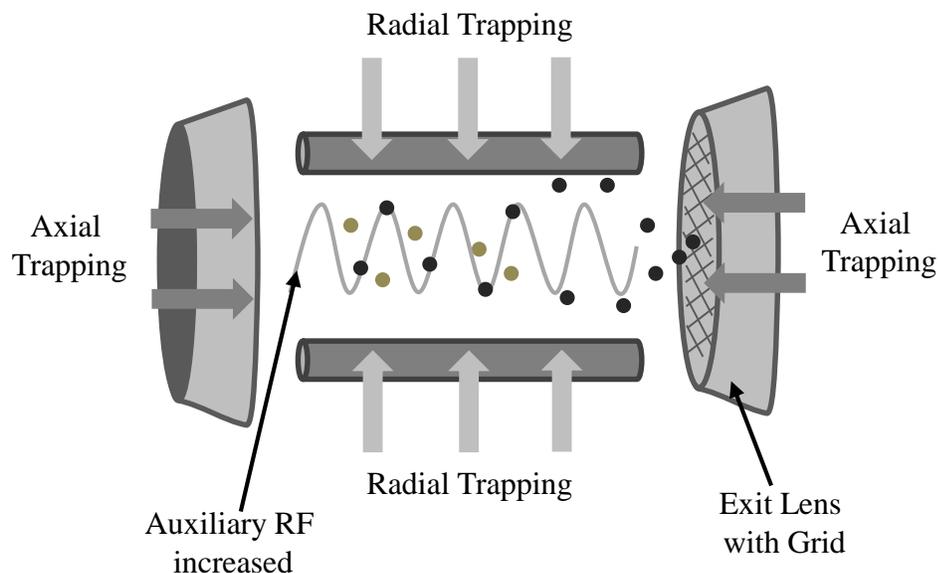


Figure 1-9 Schematic of the LIT mass analyzer, with axial ion ejection, used in the QTRAP<sup>®</sup> 4000

Ions are trapped in the LIT by the RF voltage applied to the rods and by the DC voltage applied to the aperture plates. Ions can be ejected from the LIT either radially (through slits in the x-rods) or axially (through the exit lens). In this section, only axial ejection will be further explained, as this is the physical design of the QTRAP<sup>®</sup> instrument. In the LIT, ions move in a radial (x/y plane) direction and in an axial direction (z plane). The potential in the hyperbolic field is the same as that of the quadrupole and is only valid well within the rods, sufficiently far away from the end lens. At the ends of the rods, the field terminates due to the presence of the lenses. This fringing field distorts the ideal quadrupole field at the entrance and exit of the LIT and the effects can couple the radial and axial motion of the ion. Prior to ejection, the ions can be excited by application of a supplementary AC voltage to the quadrupoles (or in a few cases to the exit lens). When the frequency of the ion matches that of the auxiliary AC, an increase in the

radial displacement occurs. If the ions are near the exit lens, the fringing field effect is highest and the ions can have sufficient kinetic energy gained from the fringing field to be ejected. By changing the frequency and the amplitude of the auxiliary AC voltage, ions of different  $m/z$  can be ejected to produce a mass spectrum.

Advantages of the LIT over a 3D ion trap include improved trapping efficiency, increased ion storage capacity due to the larger volume of the LIT and reduced space charge effects due to radial confinements of ions in a line rather than in a point as in 3D ion traps.

The QTRAP<sup>®</sup> 4000 instrument used in the thesis work is based on a triple quadrupole platform, where Q3 can be operated in either the normal quadrupole mode or the LIT mode.<sup>118</sup> Changes in Q3 from a conventional quadrupole to a LIT analyzer can occur in less than 1 ms. A schematic of the instrument is shown in Figure 1-10. Ions are generated by an electrospray ion source (pneumatically assisted) and they enter into an RF-only quadrupole ion guide (Q0) through an orifice plate and a skimmer. The Q0 vacuum chamber is maintained at about  $6 \times 10^{-3}$  Torr. The Q0 chamber is separated from the analyzer chamber by a differential pumping aperture called IQ1. The collision cell (q2) is an enclosed LINAC quadrupole and has two lenses, IQ2 and IQ3, located at either end. Nitrogen is used as the collision gas and pressure is maintained at approximately  $5 \times 10^{-3}$  Torr. The quadrupoles Q1 and Q3 are mechanically identical and the typical pressure in Q3 is approximately  $3 \times 10^{-5}$  Torr. Two aperture lenses are located at the exit of Q3 and ions are detected using a dynode electron multiplier. Auxiliary AC voltage can be applied to Q3 and is ramped proportionate to mass. The trapped ions are then mass selectively axially ejected. The LIT analyzer can be operated in three scan speeds: 250, 1000 and 4000 amu/s. The scan functions of the instrument are represented in Table 1-4. All the typical scan modes of the triple quadrupole are maintained, with the addition of the trap scan modes.

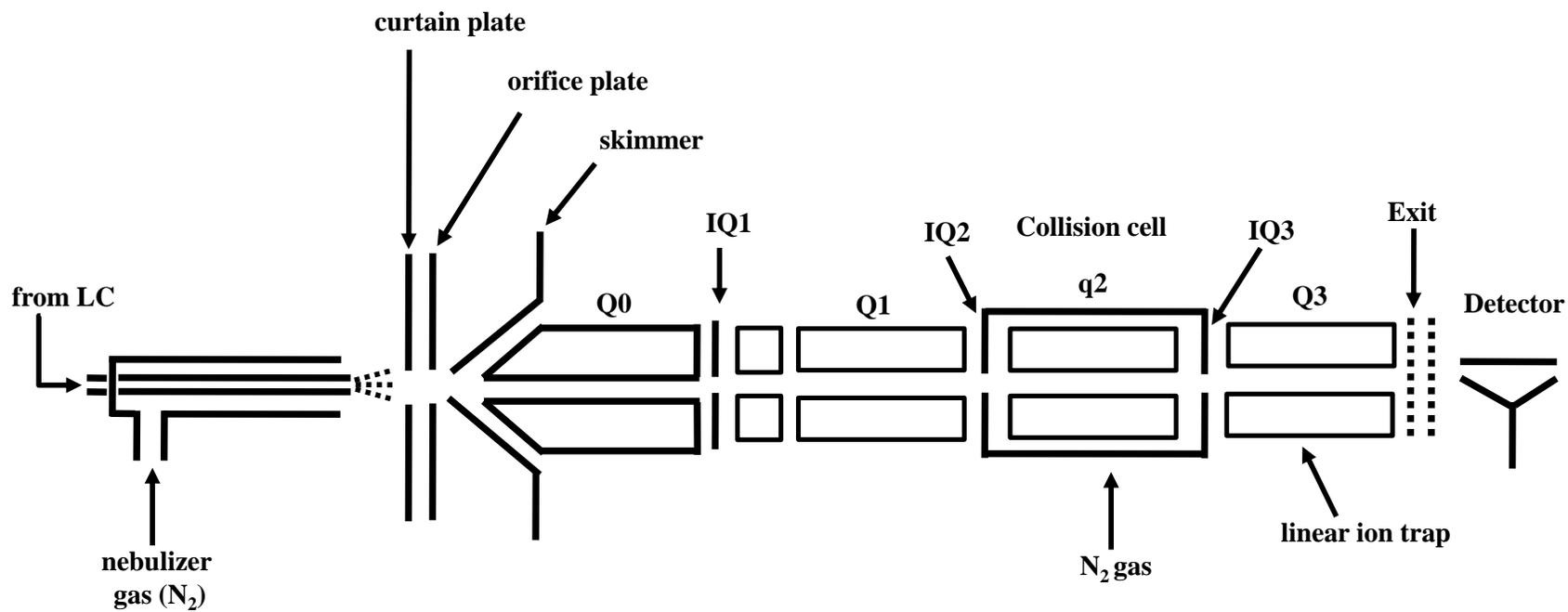


Figure 1-10 Schematic of the AB Sciex QTRAP<sup>®</sup> 4000 hybrid triple quadrupole linear ion trap instrument

Table 1-4 Scan modes used in the QTRAP<sup>®</sup> 4000 (Ref. 119)

<b>Mode of Operation</b>	<b>Q1</b>	<b>Q2</b>	<b>Q3</b>
<b>Triple Quad Scans</b>			
Q1 Scan	Resolving (Scan)	RF-only	RF-only
Q3 Scan	RF-only	RF-only	Resolving (Scan)
Product Ion Scan	Resolving (Fixed)	Fragment	Resolving (Scan)
Precursor Ion Scan	Resolving (Scan)	Fragment	Resolving (Fixed)
Neutral Loss Scan	Resolving (Scan)	Fragment	Resolving(Scan offset)
Selected Reaction Monitoring Mode	Resolving (Fixed)	Fragment	Resolving (Fixed)
<b>Triple Quad in combination with LIT scans</b>			
Enhanced Q3 Single MS	RF-only	No Fragment	Trap/scan
Enhanced Product Ion scan	Resolving (Fixed)	Fragment	Trap/scan
MS3	Resolving (Fixed)	Fragment	Isolation/frag Trap/scan
Time delayed fragmentation	Resolving (Fixed)	Trap/No frag	Frag/trap/scan
Enhanced Resolution Q3 Single MS	RF-only	No Fragment	Trap/scan
Enhanced Multiply Charged	RF-only	No Fragment	Trap/scan

#### 1.4.4 Scan Modes

For the purpose of brevity, only the scan modes discussed in later chapters of this thesis will be explained here. Figure 1-11 shows a graphical representation of these scan modes. Fragmentation in the QTRAP<sup>®</sup> can be performed either in the LIT analyzer by resonance excitation or in the collision cell, q2 as in triple quadrupole instruments.<sup>118</sup> In the enhanced product ion (EPI) mode, selection of a precursor ion is done in Q1 using RF/DC mode isolation at any of three different resolutions: low, high or open resolution. The ions are fragmented via collision induced dissociation (CID) in the collision cell, q2 and the fragment ions are trapped in Q3 operating in the LIT mode. The ions are then mass-selectively axially scanned out to create a mass spectrum. The term “enhanced” is used whenever Q3 is operated in the LIT mode because of the enhanced sensitivity when compared to Q3 operated in the conventional quadrupole mode. The advantage to generating fragments in q2 rather than in the LIT mode is that in q2, ions can undergo multiple collisions. However, in a LIT analyzer, fragmentation occurs by resonance excitation of the precursor ion and in many cases the product ions are too cool to fragment further. As a result, CID conducted outside the LIT generates triple quadrupole fragmentation patterns, which typically yield a richer product ion spectrum than those obtained via resonance excitation. Another reason is that the duty cycle is enhanced because the precursor ion isolation and fragmentation generation steps are spatially separated. Also, there is no inherent low mass cut-off as in ion trap MS/MS spectra.

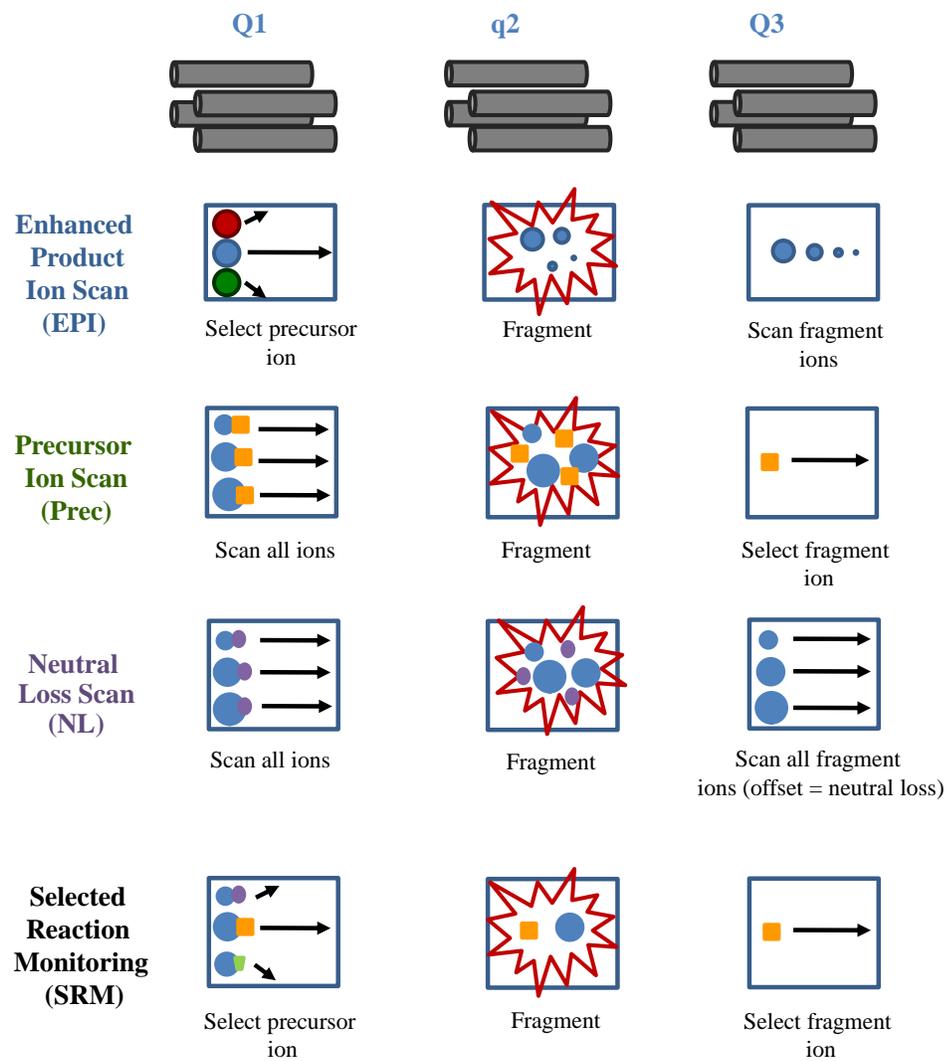


Figure 1-11 Triple quadrupole scan modes

The triple quadrupole scans, precursor ion (PC), neutral loss (NL) and selected reaction monitoring (SRM) scans are typically the same as for a conventional triple quadrupole mass spectrometer.<sup>101</sup> When an ion loses a diagnostic fragment as a charged fragment or a neutral fragment, it can be detected using PC or NL scans respectively. In PC scans, Q1 is used as a resolving RF/DC transmission quadrupole to scan across a mass range. Ions are fragmented via CID in q2 and Q3 is set as an RF/DC quadrupole mass filter to only transmit the fragment ion of interest. Only ions that are transmitted through Q1 that fragment to generate the diagnostic fragment will be detected.

In NL scans, Q1 is used in the same way as in PC scans, in a resolving RF/DC mode across a mass range and the ions fragment in q2. Q3 is set in the resolving mode as well, scanning across a similar mass range that is offset by the mass of the neutral fragment. Only the ions passing through Q1, which lose a neutral fragment of the defined mass will then be transmitted through Q3 and be detected.

In SRM mode, Q1 is set in the resolving RF/DC mode to select the precursor ion of interest and Q3 is set to select a specific fragment of interest. The collision energy used in q2 to fragment the precursor ion is optimized to generate an abundance of the diagnostic fragment ion. Only ions with the exact transition of precursor to product ion will be detected. This type of scan has the highest duty cycle of the scans. Many SRM scans can be combined together (known as multiple reaction monitoring or MRM) in one single LC-MS run to detect the presence of as many known metabolites as possible in a complex biological matrix. PC and NL scans are very useful in identifying metabolites, especially those belonging to a family or class of compounds that share diagnostic fragment ions or neutral losses in their fragmentation patterns. Despite the relatively low duty cycles of these scans, they are very selective and much more sensitive than full MS scans, although not as sensitive as MRM scans.

#### **1.4.5 Information-Dependent Acquisition (IDA) Scans**

The use of IDA scans has become very useful in increasing throughput and generally requires no foreknowledge of the analytes. This procedure combines two or more different scan modes sequentially in a single LC-MS or LC-MS/MS run. The first scan is known as the survey scan and the data is processed “on the fly” to determine the masses of interest based on pre-defined selection criteria. Once the criteria are met, then a second data-dependent scan is performed. Depending on the cycle time of the experiment, several data-dependent scans can be performed. In most LC-MS experiments, a typical IDA experiment would include a full MS scan as a survey scan followed by an EPI scan as a dependent scan. In this thesis work, the most common IDA experiments used were an MRM survey scan followed by one to four dependent EPI scans. This IDA scan also offered a way to perform quantitative and qualitative analyses in one run. PC and NL scans were also used as survey scans. IDA has the potential to offer a wealth of information in just one run.

#### **1.4.6 Orthogonal Acceleration – Time-of-Flight (oa-TOF)**

The TOF mass analyzer consist of a long field-free tube about 100 cm in length. A schematic of an oa-TOF mass spectrometer is represented in Figure 1-12. The flight tube is a vacuum enclosure between the ion source and detector. The basic principle of the TOF is as follows: ions are separated on the basis of their differences in velocities. A short pulse of ions having defined kinetic energy is dispersed in time when it travels down the long flight tube.

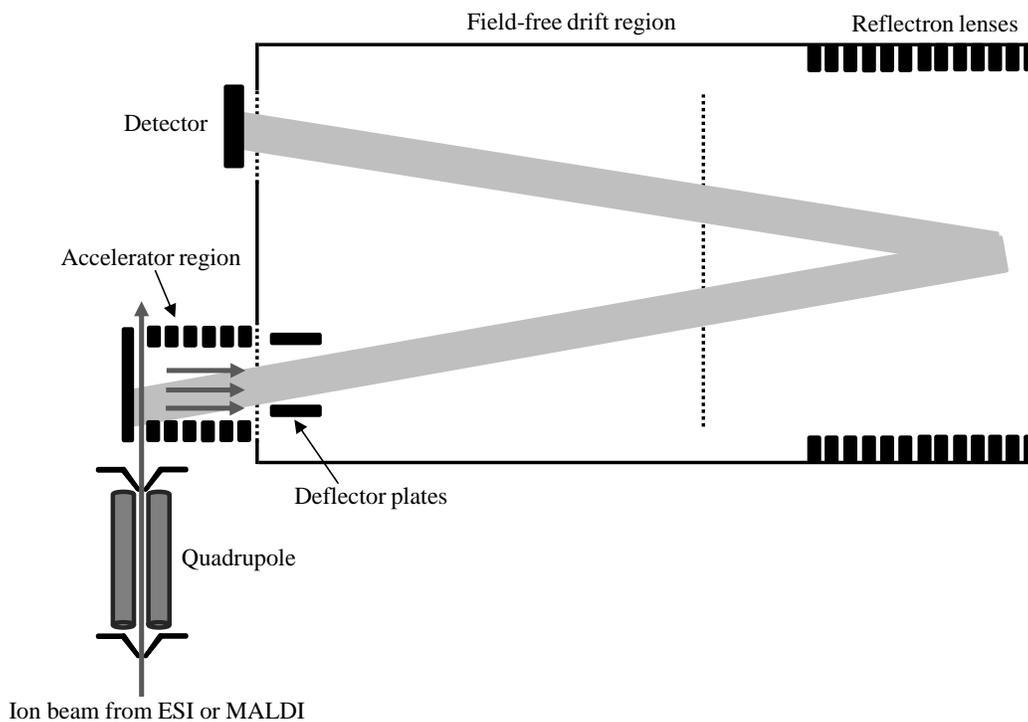


Figure 1-12 Schematics of an orthogonal acceleration time-of-flight mass analyzer

The velocities of the ions can be defined as:

$$v = \sqrt{\frac{2eV}{m}}$$

where  $V$  is voltage,  $m$  is the mass of the ion and  $e$  is the charge of the ion. The velocity of an ion is inversely proportional to its  $m/z$  value, so lower  $m/z$  ions travel faster and reach the detector earlier. The time an ion takes to get to the detector is given by:

$$t = \frac{L}{v} = L \sqrt{\frac{m}{2eV}}$$

where  $L$  is the length of the flight tube. Timing circuits are used to coordinate the start time with the initial pulse of the ions or other events in the experiment. The measured arrival times of the ions can be converted to a mass spectrum by the calibration of the instrument. In a TOF mass spectrometer, resolution is related to the temporal width of the ion packet when it reaches the detector. The main factors contributing to resolution are the initial kinetic energy spread of the ions as well as the spatial distribution, owing to ions of the same mass being ionized at different locations in the source. A device used to correct for the spatial and energy spread is a reflectron. The reflectron is an electrostatic mirror with a series of grids and electrodes in a decelerating or retarding field. The reflectron is placed at the end of the flight tube. Ions from the flight tube enter the grid of the reflectron, which is generally at the flight tube potential and as they pass through the series of electrodes, they are slowed down by the retarding field until they stop. The same electrical field that slowed the ions down on entering the reflectron will now accelerate them as their direction is reversed and they leave the reflectron. The principle is that when ions with the same mass enter the reflectron, those with more energy will travel faster and arrive sooner but will penetrate deeper and spend more time in the field than a slower ion. As a result, ions with the same  $m/z$  value will arrive simultaneously at the detector, which is

placed at the other end of the tube. The added length in the flight path also contributes to resolution.

Because of the start-stop timing that is essential to TOF operation, the instruments are optimally combined with pulsed-mode ion sources such as matrix-assisted laser desorption ionization (MALDI). Alternatively, using an oa-TOF mass spectrometer can accommodate a pulsed-ion beam from a continuous ion beam source such as ESI. The ion beam enters an ion acceleration region at a perpendicular direction to the main axis of the TOF instrument, illustrated in Figure 1-12. At regular time intervals, a short pulse of an orthogonal accelerating field is applied to ions in a section of the beam, and their kinetic energy is increased. The beam fills the acceleration region in about the same time it takes for a TOF scan to be completed after orthogonal acceleration. In the “fill-up mode” the first stage of the acceleration region is field-free. Once the acceleration region is filled, a field is generated rapidly by applying a voltage pulse to one or more of the electrodes. This is referred to as the “push-out” mode and the ions are accelerated orthogonally relative to their velocities in the ion beam. The pulse is generally used to start the timing electronics. This approach reduces spatial and energy spreads and provides high efficiencies to gate ions from an external continuous source. Coupling the ion source with the oa-TOF is generally achieved using an RF-only quadrupole, or a LIT. These have the added advantage of focusing the ion beam, minimizing divergence and increasing ion-transport efficiency.

## **1.5 Chemical Derivatization in LC-ESI-MS**

There are many derivatization reactions used in LC-MS applications.<sup>119, 120</sup> There are a range of compounds that are weakly acidic or basic and display poor sensitivity with ESI-MS, such as sugars, alcohols, amino acids, and vitamins. To address this shortcoming, several derivatization strategies have been developed to

improve the sensitivity of these analytes. Most published approaches focus on making small modifications to the analyte to improve ESI response. No matter what the functional group is or the strategy used, the ideal reagents and the reaction should have the following characteristics: inexpensive, simple synthesis steps, high yield, non-toxic reagents, stable over long periods of time, reaction should proceed under mild conditions, with minimal side products, reaction mixture should be compatible with the LC-MS systems and not require too much cleanup.

Chemical derivatization of the analyte is often used to enhance the ESI response by either adding a chargeable functional group<sup>121</sup> or by increasing its surface activity.<sup>122</sup> Chargeable functional groups facilitate ion formation through protonation or sodium adduct formation. Another method is to facilitate charging by adding electrochemically reactive functional groups. Ionization can then be accomplished through electrochemical oxidation or reduction.<sup>123</sup> Derivatization can be used to modify hydrophilic compounds to improve their surface activity as they tend to reside in the droplet interior. Not only are their ESI responses dramatically improved but also their retention on a reversed-phase HPLC column. Increasing the retention time also avoids the pitfalls associated with suppression due to salts and other early-eluting compounds. In addition, derivatization of the analyte may shift the mass of the derivatized product out of the low-mass region of the mass spectrum, which is usually complicated by the presence of solvent clusters and contaminants. Derivatization can also be used to generate fragments efficiently during collision induced dissociation (CID) in tandem MS and produce intense fragment ions that can be valuable in selected reaction monitoring (SRM) quantification methods. It is important to note that the ESI response can be increased exponentially by using a derivatizing agent that increases both the proton affinity and the hydrophobicity of the analyte.<sup>124</sup> Derivatization can also improve quantitative analysis by using stable isotope labeled reagents.<sup>120</sup> The “light” isotope reagent can be used to label the analyte and the “heavy” isotope reagent can be used to label the standard to be used as the internal standard. The

advantage to this method is that an internal standard can be produced for each analyte. The purchase or synthesis of an internal standard for each analyte is therefore not required, lowering the cost of the experiment.

## **1.6 Metabolite Identification by LC-MS**

One of the major bottlenecks in metabolomic studies is the accurate identification of small molecules or metabolites. Unlike such biopolymers as DNA and proteins, the diversity in the chemical and physical properties of metabolites makes the task very difficult. Because of the vast number of publications in the field of metabolite identification, each employing very different levels of evidence, the Metabolomics Standard Initiative (MSI) published several guidelines for the publication of metabolomics experiments. One of these publications discusses the reporting standards for chemical analysis and defines the confidence levels for identifying non-novel metabolites, ranging from level 1 to 4.<sup>125</sup> It states that the minimum standards for identifying a non-novel metabolite is by analyzing the metabolite using two independent and orthogonal data sets relative to an authentic compound analyzed under identical experimental conditions, for example retention time and mass spectra, accurate mass and tandem mass spectrometry. Identification of novel compounds usually requires extraction, purification, accurate mass measurement, elemental analysis, fragmentation patterns and the use of other analytical techniques beyond mass spectrometry: NMR, IR or chemical derivatization. Typically, MS-based metabolite identification in untargeted metabolomics is mainly achieved through integration of LC-MS and LC-MS/MS data. This includes ion annotation, an accurate mass search to obtain elemental composition, isotopic pattern evaluation, mass spectral interpretation and spectral matching.

### 1.6.1 Ion Annotation and Accurate Mass Measurements

Ion annotation is the procedure whereby a group of ions likely to originate from the same compound are identified. In LC-MS, one compound can be represented by multiple peaks with different  $m/z$  values at similar retention times. These peaks are due to the presence of isotopes, adducts, charge states, and neutral loss fragments. The solvent and buffer composition, pH, pKa, organic modifier and analyte concentration can influence the formation of adducts and the occurrence of multiply charged species. The correct adduct ion must be determined in order to obtain the accurate mass of the neutral molecule. Also different adducts can alter the fragmentation pathways. Usually,  $[M+H]^+$  ions are preferable for fragmentation because they yield more informative spectra at common collision energies of 10 – 50 eV. Mass spectrometry software from many vendors that has charge state determination included can help detect adduct ions.

After successfully annotating the ions, their monoisotopic exact masses can be deduced based on the difference in the mass of the adduct and charge state deconvolution. These masses can then be queried against a metabolite database such as HMDB,<sup>16</sup> METLIN<sup>126</sup> or MassBank.<sup>127</sup> Accurate mass is generally measured on an FTICR-MS (less than 1 ppm accuracy), magnetic sector (less than 1 ppm accuracy), TOF or quadrupole-TOF (5-10 ppm accuracy) or orbitrap (less than 1 ppm accuracy). The aim of the library search is to obtain a hit containing the correct structure of a compound in the library or partial insights on structure based on compounds that almost match. Identification based on accurate mass seldom results in unique identification of these compounds. The monoisotopic mass of a compound is not usually sufficient to determine its elemental composition, even for mass accuracies in the sub ppm range.<sup>128</sup> Therefore both accurate mass and isotope patterns are required and many algorithms depend on both.

## 1.6.2 Mass Spectral Interpretation and Spectral Matching

Spectral interpretation of tandem mass spectra obtained using LC-MS derived its beginnings from fragmentation interpretation of electron ionization spectra. It involves deducing the possible structure of a compound through predicting fragmentation by *in silico* methods or by diagnostic cleavages. Because the fragmentation is a gas-phase reaction, cleavage sites on a molecule can be estimated and described with fragmentation rules. A list of common neutral losses that occur during CID fragmentation and common fragmentations typical in atmospheric pressure ionization sources have been published to assist analysts in interpreting mass spectra.<sup>129</sup> There are also many commercially available software sources that differ in their approaches that can be employed by the mass spectrometrists to help in spectral interpretation. The Advanced Chemistry Development (ACD) Fragmenter software is such a tool that uses a database of fragmentation rules. HighChem MassFrontier ([www.highchem.com](http://www.highchem.com)) possesses a large library of fragmentation schemes extracted from the literature and in-house spectral libraries. Other *in silico* tools use a different rationale to predict structure, where a list of fragments is generated through combinatorial disconnection of the chemical bonds in the compound. Such examples are Fragment Identifier (FiD)<sup>130</sup> and MetFrag<sup>131</sup>. The bond energies are calculated and low-energy cleavages are preferred over high-energy cleavages. After hypothetical fragments are predicted, they are then assigned a similarity score when compared against the experimental spectrum.

Another approach uses diagnostic neutral losses and fragment ions to predict and interpret the fragmentation of compounds belonging to the same chemical class. Bourcier and Hoppilliard used characteristic neutral losses to identify tyrosine derivatives.<sup>132</sup> It is well known that compounds with similar structures can share similar spectral characteristics. The fragmentation of compounds belonging to the same chemical class can be studied and fragmentation reactions and pathways deduced using existing fragmentation rules. Under identical experimental conditions, the MS/MS spectra of unknown samples

can be generated and compared for putative identification of the compound class. Using a well known list of common neutral losses can also assist in interpretation of unknown mass spectra. If logical losses can be assigned, then parts of the unknown metabolite can be identified. Intelligent deduction of the putative identity of the unknown metabolite can be made based on LC retention time, fragmentation of similar standards, known fragmentation patterns and literature searches. This form of interpretation is level 3 of the MSI and is termed the putatively characterized compound class.<sup>125</sup>

Spectral matching is the verification of metabolite identification by comparing the MS/MS spectrum of the unknown with the MS/MS spectrum of authentic standards assembled in a spectral library. There are several spectral libraries that are available to the general public.<sup>126, 127</sup> To measure the similarity between the unknown and authentic spectra, an appropriate scoring function is important for any spectral matching algorithm. The most common scoring function calculates the similarity of the query spectrum and the library spectrum by treating the two spectra as vectors and determining their dot product, as in, for example, MassBank.<sup>127</sup> Spectral libraries can be limited in their coverage and one way of expanding the coverage is by allowing metabolites of similar structure or substructure to be matched, as in MassFrontier.

The above computational methods assist in reducing the search space and prioritizing putative identifications, especially since the number of available standards is limited. It is important to remember that for confirmation of the identity of an unknown, the authentic standards still need to be obtained and analyzed under identical conditions of the unknown sample to compare retention times and MS/MS spectra.

## 1.7 Quantification of the Metabolome by LC-MS

### 1.7.1 Calibration Techniques

As in other analytical techniques, quantification using LC-MS is based on the comparison of the intensity of analyte response (measuring either peak height or area) in a sample with the intensity of known amounts of the authentic analyte or a chemical analog, measured under identical conditions. There are three common procedures for comparison with these standards, namely external calibration method, standard addition method and the internal calibration method.<sup>133</sup> External calibration is the simplest method of the three and involves an external standard or known material, which is prepared separate or external to the unknown material or sample. In this method, a series of such external standards containing known amounts of the analytes of interest is prepared. Ideally, at least five of these solutions are needed to create a reliable calibration. The intensity of the analyte signal is plotted as a function of the known analyte concentration and a calibration curve is created by fitting the plot using the method of least squares. It is important to fit the correct form of the curve to the calibration data. Samples are then prepared and analyzed in an identical fashion and the concentration of the analyte in a sample is calculated from the calibration curve. It should be noted that external calibration does not take into account any matrix effects or sample losses due to storage or preparation.

Standard addition method can be used to address matrix effects and is often used to determine the concentration of analytes in complex matrices such as biological fluids or tissues. The rationale is to add analyte in known concentrations to the unknown sample matrix and the increase in response after analysis is thought to be due only to the change in analyte concentration and not to any other interfering species. The procedure of standard addition is to add the analyte in increasing concentrations to equal aliquots of the sample. Ideally, at least five added concentrations should be used. The detector response of the standard addition solutions is then plotted against the added concentrations. The

calibration curve is fitted using linear regression and the concentration of the analyte in the original solution calculated by extrapolation of this curve to intercept the x-axis. This is illustrated in Figure 1-13. Although this method takes into account matrix effects, it does not account for sample losses due to sample handling and preparation. This approach is also time-consuming and impractical if the number of samples is high or sample volumes are limited as a series of “spiked” samples must be prepared and analyzed for each unknown sample. For these reasons, this method has limited use in metabolomics.

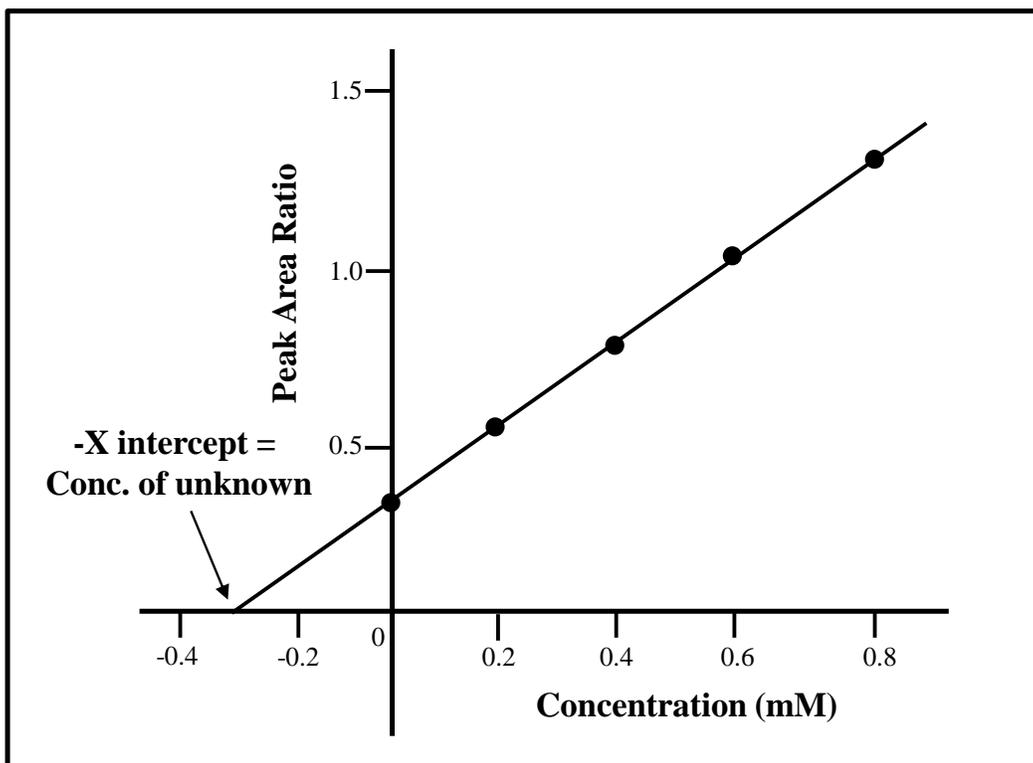


Figure 1-13 Standard addition curve

The internal standard method can be employed to overcome both matrix effects and losses due to sample handling. The procedure involves the comparison of the instrument response from the analyte in the sample to the response of an internal standard added to the sample early in the analytical process (preferably before sample preparation). This internal standard is a suitably selected compound that closely resembles the analyte in physical and chemical properties. A constant and equal amount of the internal standard is added to all unknown samples. The same amount is added to calibration standards, prepared with increasing concentrations of the analyte. Intensity signals from both the analyte and the internal standard are measured in the standards and the samples. The ratio of the two signals is then used to generate the calibration curve and determine the concentration of analyte in the samples. If the choice of internal standard is appropriate, losses in the analyte will be mirrored by the losses in the internal standard and the matrix effects experienced by both would be, in theory, identical. In this way, it also compensates for variation in sample preparation, injection volumes and chromatographic separation. An ideal choice in internal standard is the use of the stable isotope-labeled analog of the analyte. A stable isotope-labeled internal standard (SIL-IS) has the same chemical formula and structure as the analyte but differs in the isotopes,  $^1\text{H}$ ,  $^{12}\text{C}$ ,  $^{14}\text{N}$  or  $^{16}\text{O}$ , which are usually replaced by  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  or  $^{18}\text{O}$ . The SIL-IS, therefore exhibits identical behavior with the target analytes. Using SIL-IS is therefore preferred over using a structural analog, but they can be quite expensive and in the study of metabolomics, a limited number of SIL-IS are available commercially. Deuterated SIL-IS are the least expensive and are most commonly used; however, they tend to suffer from isotope effects in the chromatographic separation, when using a reversed-phase column. Isotope effect is generally observed as deuterated compounds elute earlier than their protiated counterparts. The zero point energy, lowest energy of a system, is dependent on the reduced mass of the atom; the heavier the atom the lower the frequency of vibration and the smaller the zero point energy. The reverse is true and lighter atoms have a greater frequency of vibration and a higher zero point energy. Since deuterium is heavier than

hydrogen, it has a lower zero point energy, resulting in different bond dissociation energies for C-H and C-D. The nature of isotope effects is due to a large extent on vibrational frequencies, which is dependent on the mass of the nuclei. The smaller amplitude of vibrations for the C-D bond results in a more compact electron distribution and a decrease in polarizability and molecular volume for the C-D bond.<sup>134, 135</sup> Therefore there will be less interaction between the deuterated compound and the hydrophobic stationary phase than for the protiated isotopomer and the deuterated compound elutes earlier. This effect becomes more pronounced as the number of deuterium atoms in the molecule increases and the effect on polarity is altered depending on the proximity of a heteroatom to the substituted deuterium.<sup>135</sup> For accurate quantification, co-elution of the analyte and SIL-IS is preferred. In short, the use of deuterated internal standard is not always the optimal choice. SIL-IS containing isotopes <sup>13</sup>C, <sup>15</sup>N or <sup>18</sup>O are often a better choice as they do not exhibit isotope effects; however these isotopomers tend to be very expensive and even fewer compounds are commercially available. Of these isotopomers, <sup>13</sup>C-SIL-IS are most often used.

### **1.7.2 Validation of Quantitative Methods**

In small molecule analysis, after choosing an appropriate calibration technique and creating an LC-MS method, optimizing separation and detection of the metabolites, it is important to validate the quantitative method. Well-characterized and fully validated analytical methods are essential to generate reproducible and reliable data that can be used to accurately evaluate and interpret the biological findings. The U.S. Food and Drug Administration (FDA)<sup>87</sup> provides guidance for bioanalytical method validation of drugs and their metabolites in biological fluids and these procedures are adopted in quantitative metabolomics. This guide defines the essential parameters used in method validation – namely accuracy, precision, selectivity, reproducibility, limits of detection and quantification and stability – and addresses how to assess and determine these

parameters. Evaluation of these parameters in drug analysis is very straightforward and relies on the use of calibration standards, which are known amounts of the analytes spiked into the biological sample matrix and quality control samples, which are spiked sample matrix (with known quantities of analyte) used to assess the integrity and validity of the results of the unknown samples. These investigations are straightforward because the drug and metabolites to be tested are exogenous to the sample matrix.

However, as biomarkers have become increasingly important as indicators of the pharmacodynamic effect of drug treatment, safety and mechanism of drug action, the application of method validation has become complicated. One of the major reasons for this is because biomarkers tend to be proteins (used as surrogate endpoints) and small molecules (which aid in understanding the biology of the disease), which are endogenous in nature. Metabolomics is often used in biomarker discovery and an integration of proteomic and genomic approaches can provide a more comprehensive understanding of the biology of the disease. Validation of an analytical method used in the measurement of endogenous compounds is more complex than those used conventionally to measure xenobiotic drugs and challenges the existing guidelines which are still in use by many researchers doing metabolomics studies. Lee *et al* proposed a fit-for purpose method for the validation of a bioanalytical method for biomarker measurement, which can be adapted for metabolomic studies.<sup>136</sup>

A challenge exists because of the presence of the analyte in the control or blank matrix, which is used to prepare calibration and quality control standards. The presence of the analyte also presents added difficulties when evaluating selectivity, recovery (the extraction efficiency of a known amount of analyte from the sample matrix), matrix effects, and limits of detection and quantification. It is often difficult, if not impossible, to obtain analyte-free authentic biological matrix or matrix with known analyte concentration for preparation of the reference standards. Occasionally, sample matrix with negligible concentration of the analytes can be used. It must be noted that it is essential to consider a proper

approach for constructing the calibration curve as it is the key to the success of the method. Ideally, the use of standard addition methods can be used to circumvent this problem. The advantage of using this approach is that matrix effects are minimized and an accurate determination of the analyte in the unknown sample can be determined. This approach, however, requires more sample volume and analysis time than other conventional methods, which may not be practical in metabolomics where sample volumes can be small (microliter amounts) and sample numbers high (hundreds). Another disadvantage is that if the endogenous concentration of the analyte is high in the control matrix, addition of relatively small concentrations of analyte may not be distinguishable or alternatively, result in saturation of the detector. Extrapolation of the calibration curve may lead to erroneous results. Furthermore, it is more difficult to determine the limits of detection and quantification using this approach.<sup>137, 138</sup>

Another approach is to use authentic analytes but use a “surrogate” matrix in an attempt to perform matrix-matching.<sup>137</sup> A surrogate matrix could be an artificial or synthetic matrix, a modified matrix or matrix from another species. In its simplest form, surrogates can be water or a buffer such as phosphate buffered saline at pH 7.4. Synthetic urine and cerebral spinal fluid can also be prepared and used.<sup>139</sup> Synthetic urine, plasma and serum are also available commercially. An alternative to using the authentic biological matrix is to treat the matrix to remove the analyte. Examples include oxidation of the analyte,<sup>140</sup> stripping by activated carbon,<sup>141</sup> or by affinity chromatography. Disadvantages to this approach include irreproducibility in preparing the synthetic matrices, and the fact that the analyte solubility and extraction from the synthetic matrix may differ in the authentic matrix.<sup>141</sup> Alternatively, in some cases, another species can be used as a surrogate matrix.<sup>142</sup> Some analytes may occur in higher concentrations in human plasma, than in rat, rabbit, pig or other mammalian species or are absent. Thus the use of these other biological matrices as a surrogate may be appropriate. The disadvantage of using the surrogate matrix approach is that it may not correct for matrix effects. One way to determine the suitability of this approach is to prepare

the calibration standards in the surrogate and the authentic matrices and determine the slopes of both calibration curves. Similar slopes indicate a similarity of the matrices and suggest that this matrix can be used. If there is significant deviation in the slopes, an alternative matrix should be used.<sup>138</sup>

A third and interesting approach is the use of authentic matrix and “surrogate” analyte.<sup>138, 143</sup> A surrogate analyte is an analyte that does not occur naturally in the matrix but is similar to the authentic analyte in its physical and chemical properties. For LC-MS analysis, an isotope-labeled analyte standard that differs only in mass is the best choice for the surrogate analyte. The surrogate analyte is used to prepare calibration standards and a third compound, either a non-endogenous chemical analog or preferably another isotope-labeled form of the analyte is used as the internal standard. To test suitability of this approach, the response of the authentic and surrogate analytes must be determined and theoretically be equal. Their response must also be constant over time and over the concentration range tested.<sup>138</sup> The disadvantages to this approach are that it can only be used for methods involving MS detection and the use of many isotope-labeled compounds can be very expensive.

It is important to remember that method validation is performed to ensure the reliability and quality of the results of the authentic analyte in the authentic matrix. For that purpose, surrogate analyte or matrix should only be used when preparing calibration standards and for determinations of limits of detection and quantification. All validation parameters should be determined in authentic matrix using authentic analytes. Precision, accuracy, and stability studies are performed using spiked or un-spiked authentic matrix calculated from the surrogate calibration curve. Variability limits, often measured as coefficient of variation, is the same as stated in the FDA guidelines,<sup>87</sup> less than 15% or 20% at the LLOQ.

## 1.8 Inborn Errors of Metabolism (IEM) – Study of Acylglycines

Analysis of glycine conjugates (acylglycines) has become important in the investigation of several inborn errors of metabolism (IEM) known collectively as mitochondrial fatty acid oxidation disorders and organic acidurias. Fatty acids are an important source of energy in humans and may come from different sources: diet, released from adipose tissue or synthesized *de novo*. For energy production, fatty acids undergo  $\beta$ -oxidation in the mitochondria,<sup>144, 145</sup> although a small portion is oxidized in the peroxisomes as well. Once released, fatty acids are rapidly activated to form a Coenzyme-A (CoA) ester by a variety of different acyl-CoA synthases. The oxidation, which occurs in both the mitochondria and peroxisomes, consists of four enzymatic steps, in which an acyl-CoA ester undergoes subsequent steps of dehydrogenation, hydration, another dehydrogenation and finally a thiolitic cleavage.<sup>144</sup> Each dehydrogenation step yields electrons, which are transported via electron transferring flavoprotein (ETF) and electron transferring flavoprotein dehydrogenase (ETF-DH) to the respiratory chain. Each cycle of  $\beta$ -oxidation produces one acetyl-CoA molecule, which is converted to acetoacetate in the hydroxymethylglutaryl cycle. This oxidation is illustrated in Figure 1-14. Acyl-CoA esters can react with glycine in the mitochondrion and leave the cell in the form of acylglycines. A similar process can occur with carnitine to form acylcarnitines. The reaction of acyl-CoA with glycine is catalyzed by the enzyme glycine-N-acylase and this enzyme is found to be specific for certain acyl groups.<sup>146</sup> Defects in the mitochondrial  $\beta$ -oxidation system can result in an elevation of acyl-CoA, which is then conjugated to either glycine or carnitine. In fact, most of the currently identified mitochondrial  $\beta$ -oxidation deficiencies are associated with distinct acylcarnitine or acylglycine profiles. Excess acyl-CoA may also be subjected to  $\omega$ -oxidation to yield dicarboxylic acids, which can then be conjugated to glycine<sup>147</sup> or carnitine. Therefore acylglycines are used to diagnose defects in the fatty acid oxidation process.

Acylglycines are also biomarkers for organic acidurias (OA), also known as organic acidemias.<sup>148, 149</sup> The term “organic acidurias” or “organic acidemias” is used to describe a group of disorders that result from a dysfunction, usually an enzyme deficiency, in a specific step in amino acid catabolism. These disorders are usually characterized by elevated levels of non-amino organic acids, acylglycines and acylcarnitine in the urine. The majority of OA are caused by abnormal catabolism of the branched chain amino acids and lysine. The organic acids that are elevated due to a block in that pathway give the name to that condition. Figure 1-15 demonstrates the catabolism of amino acids and the enzyme defects that can result. Table 1-5 summarizes the fatty acid disorders, organic acidurias and the acylglycines used in the diagnosis.<sup>149</sup> There are many other organic acidurias and fatty acid oxidation disorders that are characterized by organic acids and acylcarnitines in plasma but for relevancy are not discussed here.

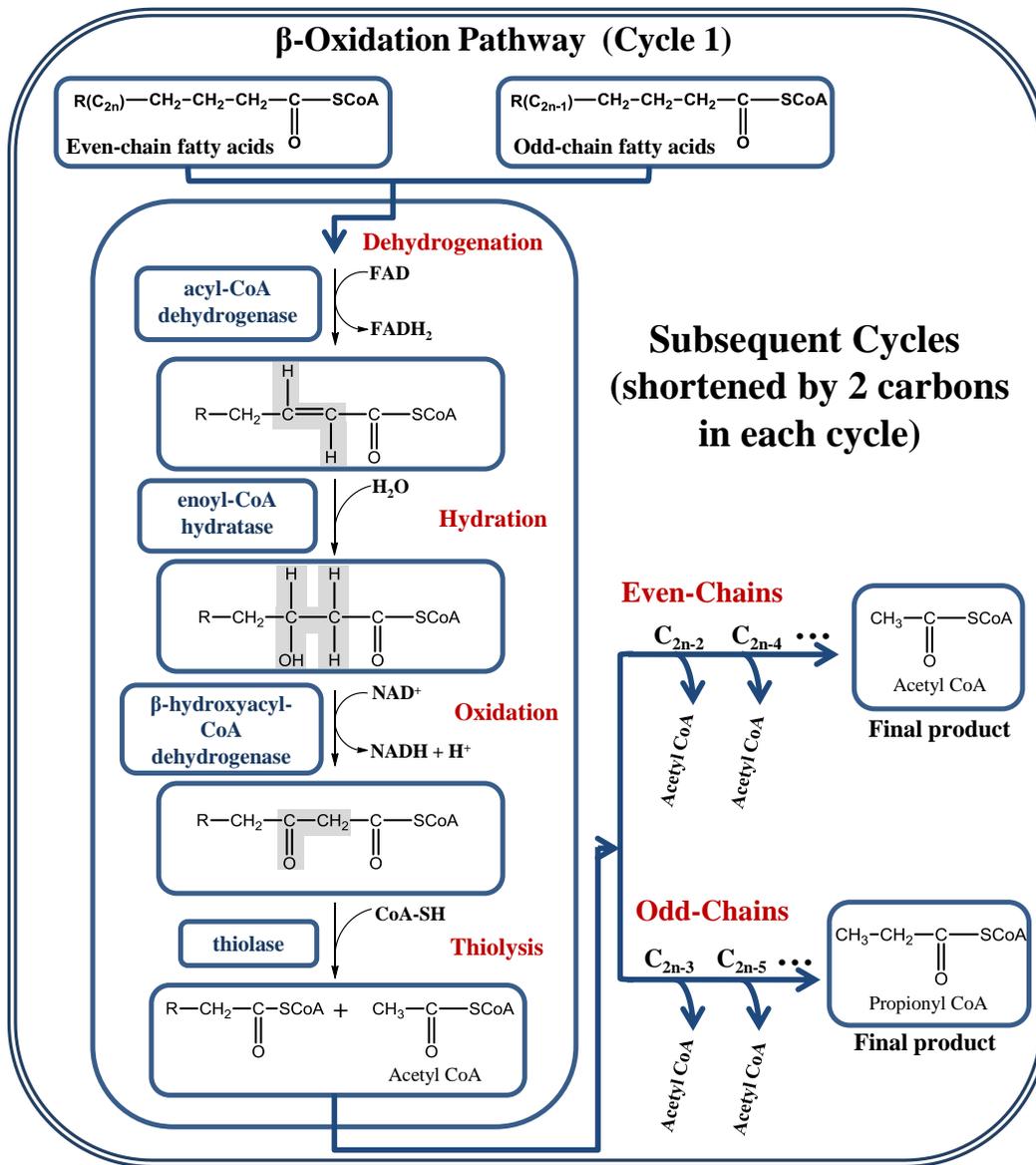


Figure 1-14 Fatty acid oxidation by the beta-oxidation pathway

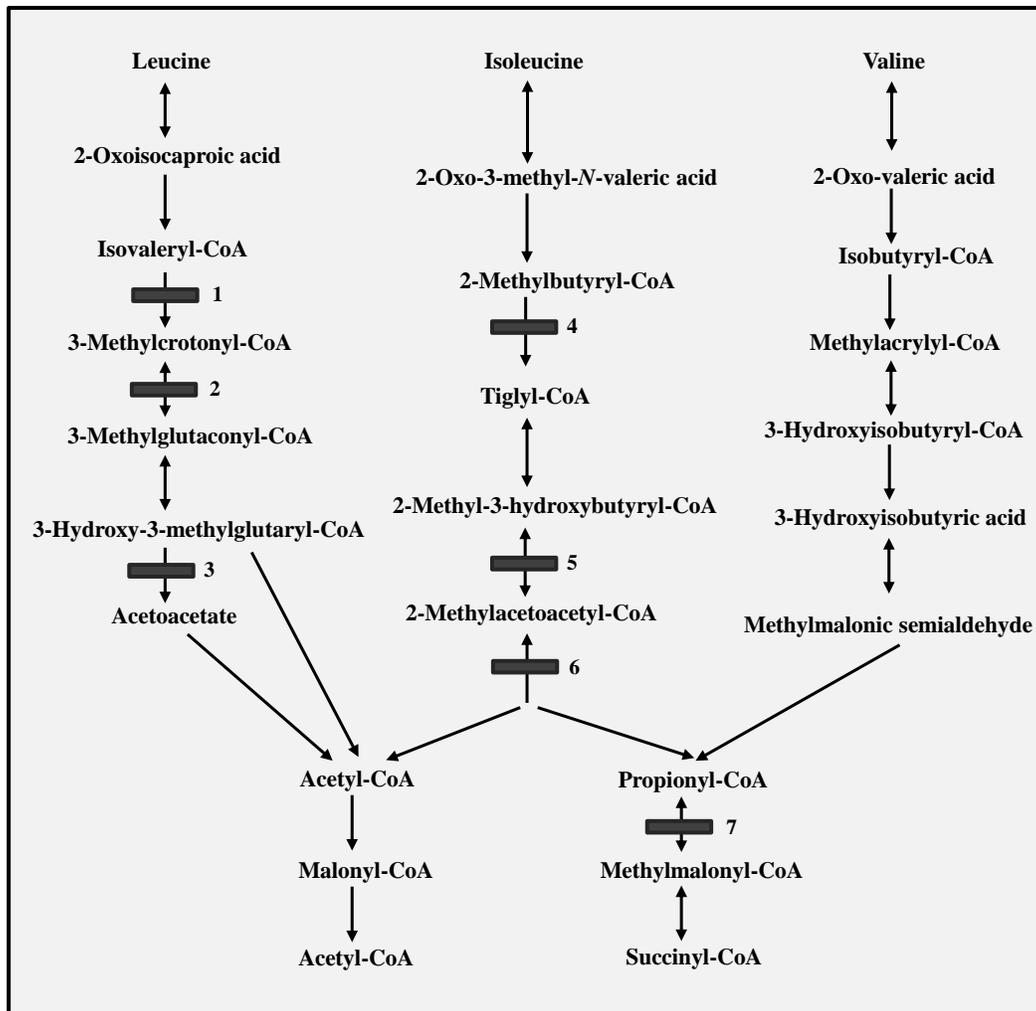


Figure 1-15 Pathways of branched-chain amino acids. Enzymes: (1) isovaleryl-coenzyme A-CoA dehydrogenase, (2) 3-methylcrotonyl-CoA carboxylase, (3) 3-hydroxy-3-methylglutaryl-CoA lyase, (4) 2-methylbutyryl-CoA-dehydrogenase, (5) 2-methyl-3-hydroxybutyryl-CoA-dehydrogenase, (6) 2-methylacetoacetyl-CoA thiolase, (7) propionyl-CoA carboxylase. Black bars indicate enzyme deficiencies and the corresponding disorders can be found in Table 1-5.

Table 1-5 Acylglycine profile in metabolic disorders. Deficient enzyme in brackets.

Disorders	Urinary Acylglycines
<b>Organic Acidurias</b>	
Isovaleric acidemia ( <i>Enzyme 1</i> )	Isovalerylglycine
3-Methylcrotonyl-CoA carboxylase deficiency ( <i>Enzyme 2</i> )	3-Methylcrotonylglycine
Multiple carboxylase deficiency ( <i>Enzyme 2, deficiency of all biotin-dependent carboxylases</i> )	3-Methylcrotonylglycine
3-Hydroxy-3-methylglutaric aciduria ( <i>Enzyme 3</i> )	3-Methylcrotonylglycine
3-Ketothiolase deficiency ( <i>Enzyme 6</i> )	Tiglylglycine
2-Methylbutyryl-CoA dehydrogenase deficiency ( <i>Enzyme 4</i> )	2-Methylbutyrylglycine
2-Methyl-3-hydroxybutyryl-CoA-dehydrogenase deficiency ( <i>Enzyme 5</i> )	2-Methylbutyrylglycine
Propionic acidemia ( <i>Enzyme 7</i> )	Propionylglycine Tiglylglycine
<b><math>\beta</math>-oxidation fatty acid defects</b>	
Glutaric aciduria type II ( <i>Deficiency in enzyme called electron transfer flavoprotein</i> )	Isovalerylglycine Isobutyrylglycine 2-Methylbutyrylglycine
Short chain acyl-CoA dehydrogenase deficiency ( <i>Deficiency in acyl-CoA dehydrogenase for short chains</i> )	Butyrylglycine
Medium chain acyl-CoA dehydrogenase deficiency ( <i>Deficiency in acyl-CoA dehydrogenase for medium chains</i> )	Hexanoylglycine Phenylpropionylglycine Suberylglycine

Collectively, fatty acid disorders and organic acidurias are part of the newborn screening program in all provinces across Canada.<sup>150</sup> Newborn Screening (NBS) is a process in which infants are screened shortly after birth for a list of metabolic disorders that are treatable but not easily clinically detected. The interval between birth and first sign of symptoms may range from hours to weeks, dependent on the defect, and it is believed to coincide with the catabolism of carbohydrate, proteins and fats. The signs of the disorder are poor feeding, drowsiness, recurrent hypoglycemia, cardiomyopathy, cardiac arrhythmias, easily fatigued, muscle weakness and pain, after which the newborn sinks into an unexplained progressive coma. The goals of treatment for fatty acid oxidation disorders are to (1) control endogenous lipolysis, (2) enhance gluconeogenesis, (3) ensure availability of Co-enzyme A within the mitochondrion, and (4) provide adequate and appropriate substrate for ATP synthesis.<sup>151</sup> Therefore, limiting dietary lipids and administering glucose intravenously are necessary. The buildup of excessive toxic CoA-esters limits the availability of free CoA for other pathways. Therefore, providing supplementary L-carnitine can cause the conversion of the toxic CoA-esters to acylcarnitine, releasing the CoA. Providing a substrate depends on the disorder. If the disorder is a long chain fatty acid oxidation defect, providing medium chain triglyceride formula would be appropriate since there would be no dependence on the long chain enzymes. In the case of OA, early removal of the toxic acids is required and this is generally done by administering a blood exchange transfusion.<sup>152</sup> This is followed by a stringent protein-restricted diet consisting of an amino acid mixture, free of the amino acid for which there is a disorder. The absent amino acid is then gradually introduced into the diet. Oral glycine and intravenous L-carnitine administration are also effective treatments.

Organic acids and acylglycines are often analyzed together in NBS laboratories.<sup>149</sup> To date, there are about 250 compounds (mainly acids and about 10-15 acylglycines) that can be identified in the urine of a normal individual. There are also 70 IEMs that are known to give characteristic organic acid and

acylglycine profiles, although most of those disorders are currently characterized by organic acids. Urine is the main biofluid used to test for acylglycines and in many laboratories today they are analyzed by GC-MS.<sup>64, 149</sup> The main steps used are (1) extraction using liquid-liquid extraction, solid-phase extraction or ion exchange chromatography, (2) derivatization, usually silylation and (3) analysis by GC-MS. Although urine is almost always used, dried blood spots have also been analyzed for acylglycines.<sup>153</sup> Dried blood spots (DBS) are obtained by pricking the heel of the newborn and spotting the blood on special filter papers known as Guthrie cards. DBS are conveniently used in the NBS programs because they are easy to obtain, do not require much storage space, can be stored at room temperature and ship easily. More recently, tandem MS methods are increasingly being used for acylglycine detection, using stable isotope dilution.<sup>148, 154</sup> Although these methods are attractive because they are fast, requiring only a few minutes for the analysis, they lack the specificity to distinguish between isomers and are therefore limited in diagnosing some IEMs. To that end, several LC-MS methods were published, using HPLC and UHPLC alike to attempt to separate the isomers and employ tandem MS to perform accurate quantification.<sup>149, 155, 156</sup>

## **1.9 Overview of Thesis**

The main objective of this work is to address some of the challenges faced in MS-based metabolomics, namely metabolite identification and quantification in biological fluids. This thesis focuses on the development of liquid chromatography tandem mass spectrometric methods for these qualitative and quantitative studies. Chapter 2 explains a strategy, developed for the putative identification of novel acylglycines in human urine, based on fragmentation patterns and diagnostic neutral losses. Microsome incubations were also used to generate hydroxylated standards to assist in the identification process and breakdown graphs were constructed to determine trends in the fragmentation.

Chapters 3 and 4 focus on the development of LC-MS methods for the quantification of endogenous acylglycines in human urine and plasma, respectively, using a surrogate matrix approach. Chapter 5 describes the development of a fast LC-MS method for the analysis of homocysteine, succinic acid and methylmalonic acid, which are products of the vitamin B<sub>12</sub> pathway. Chapter 6 describes the creation of software design for metabolite identification, which expands an existing database. Finally, Chapter 7 concludes this thesis work and also discusses future perspectives of this research in metabolomics studies.

### 1.10 Literature Cited

- (1) Ryan, D.; Robards, K. *Anal. Chem.* **2006**, *78*, 7954-7958.
- (2) Schmidt, C. *J. Natl. Cancer Inst.* **2004**, *96*, 732-734.
- (3) Pauling, L.; Robinson, A. B.; Teranishi, R.; Cary, P. *Proc. Nat. Acad. Sci. U. S.* **1971**, *68*, 2374-2376.
- (4) Oliver, S. G.; Winson, M. K.; Kell, D. B.; Baganz, F. *Trends Biotechnol.* **1998**, *16*, 373-378.
- (5) Tweeddale, H.; Notley-McRobb, L.; Ferenci, T. *J. Bacteriol.* **1998**, *180*, 5109-5116.
- (6) Fiehn, O. *Plant Mol. Biol.* **2002**, *48*, 155-171.
- (7) Nicholson, J. K.; Lindon, J. C.; Holmes, E. *Xenobiotica* **1999**, *29*, 1181-1189.
- (8) Robertson, D. G. *Toxicol. Sci.* **2005**, *85*, 809-822.
- (9) Robertson, D. G.; Watkins, P. B.; Reily, M. D. *Toxicol. Sci.*, **2011**, *120*, S146-S170.
- (10) Oldiges, M.; Luetz, S.; Pflug, S.; Schroer, K.; Stein, N.; Wiendahl, C. *Appl. Microbiol. Biotechnol.* **2007**, *76*, 495-511.
- (11) Nielsen, J.; Oliver, S. *Trends Biotechnol.* **2005**, *23*, 544-546.
- (12) Nicholson, J. K.; Wilson, I. D. *Nat. Rev. Drug Discovery* **2003**, *2*, 668-676.

- (13) Villas-Boas, S. G.; Hoejer-Pedersen, J.; Aakesson, M.; Smedsgaard, J.; Nielsen, J. *Yeast* **2005**, *22*, 1155-1169.
- (14) Goodacre, R.; Vaidyanathan, S.; Dunn, W. B.; Harrigan, G. G.; Kell, D. B. *Trends Biotechnol.* **2004**, *22*, 245-252.
- (15) Sumner, L. W.; Mendes, P.; Dixon, R. A. *Phytochemistry* **2003**, *62*, 817-836.
- (16) Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S.; Mandal, R.; Sinelnikov, I.; Xia, J.; Jia, L.; Cruz, J. A.; Lim, E.; Sobsey, C. A.; Shrivastava, S.; Huang, P.; Liu, P.; Fang, L.; Peng, J.; Fradette, R.; Cheng, D.; Tzur, D.; Clements, M.; Lewis, A.; De, S. A.; Zuniga, A.; Dawe, M.; Xiong, Y.; Clive, D.; Greiner, R.; Nazzyrova, A.; Shaykhutdinov, R.; Li, L.; Vogel, H. J.; Forsythe, I. *Nucleic Acids Res.* **2009**, *37*, D603-610.
- (17) Jung, J. Y.; Lee, H.-S.; Kang, D.-G.; Kim, N. S.; Cha, M. H.; Bang, O.-S.; Ryu, D. H.; Hwang, G.-S. *Stroke* **2011**, *42*, 1282-1288.
- (18) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat. Protoc.* **2010**, *5*, 536-549.
- (19) Aranibar, N.; Borys, M.; Mackin, N. A.; Ly, V.; Abu-Absi, N.; Abu-Absi, S.; Niemitz, M.; Schilling, B.; Li, Z. J.; Brock, B.; Russell, R. J., II; Tymiak, A.; Reily, M. D. *J. Biomol. NMR* **2011**, *49*, 195-206.
- (20) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Trends Biotechnol.* **2011**, *29*, 267-275.
- (21) Keun, H. C.; Athersuch, T. J. *Methods Mol. Biol.* **2011**, *708*, 321-334.
- (22) Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrom. Rev.* **2007**, *26*, 51-78.
- (23) Werner, E.; Croixmarie, V.; Umbdenstock, T.; Ezan, E.; Chaminade, P.; Tabet, J.-C.; Junot, C. *Anal. Chem.* **2008**, *80*, 4918-4932.
- (24) Mishur, R. J.; Rea, S. L. *Mass Spectrom. Rev.* **2012**, *31*, 70-95.
- (25) Brown, M.; Dunn, W. B.; Dobson, P.; Patel, Y.; Winder, C. L.; Francis-McIntyre, S.; Begley, P.; Carroll, K.; Broadhurst, D.; Tseng, A.; Swainston, N.; Spasic, I.; Goodacre, R.; Kell, D. B. *Analyst* **2009**, *134*, 1322-1332.
- (26) Dudley, E.; Yousef, M.; Wang, Y.; Griffiths, W. J. *Adv. Protein Chem. Struct. Biol.* **2010**, *80*, 45-83.

- (27) McNaney, C.; Hnatyshyn, S.; Reily, M.; Drexler, D. 2011; American Chemical Society; TOXI-81.
- (28) Koek, M. M.; Kloet, F. M.; Kleemann, R.; Kooistra, T.; Verheij, E. R.; Hankemeier, T. *Metabolomics* **2011**, 7, 1-14.
- (29) Chen, C.; Gonzalez, F. J.; Idle, J. R. *Drug Metab. Rev.* **2007**, 39, 581-597.
- (30) Yoshida, H.; Yamazaki, J.; Ozawa, S.; Mizukoshi, T.; Miyano, H. *J. Agric. Food Chem.* **2009**, 57, 1119-1126.
- (31) Sumner, L. W. *Biotechnol. Agric. For.* **2006**, 57, 21-32.
- (32) t'Kindt, R.; Storme, M.; Deforce, D.; Van, B. J. *J. Sep. Sci.* **2008**, 31, 1609-1614.
- (33) t'Kindt, R.; van, B. J. *Handbook on Mass Spectrometry: Instrumentation, Data and Analysis, and Applications* **2009**, Nova Science Publishers, Inc.; 39-73.
- (34) Britz-McKibbin, P. *Methods Mol. Biol.* **2011**, 708, 229-246.
- (35) Ramautar, R.; Mayboroda, O. A.; Somsen, G. W.; de, J. G. J. *Electrophoresis* **2011**, 32, 52-65.
- (36) Ramautar, R.; Somsen, G. W.; de, J. G. J. *Electrophoresis* **2009**, 30, 276-291.
- (37) Zhou, B.; Xiao, J. F.; Tuli, L.; Ransom, H. W. *Mol. BioSyst.* **2012**, 8, 470-481.
- (38) Reo, N. V. *Drug Chem. Toxicol.* **2002**, 25, 375-382.
- (39) Georgiev, M. I.; Ali, K.; Alipieva, K.; Verpoorte, R.; Choi, Y. H. *Phytochemistry* **2011**, 72, 2045-2051.
- (40) Ali, K.; Maltese, F.; Zyprian, E.; Rex, M.; Choi, Y. H.; Verpoorte, R. *J. Agric. Food Chem.* **2009**, 57, 9599-9606.
- (41) Spraul, M.; Hofmann, M.; Ackermann, M.; Shockcor, J. P.; Lindon, J. C.; Nicholls, A. W.; Nicholson, J. K.; Damment, S. J. P.; Haselden, J. N. *Anal. Commun.* **1997**, 34, 339-341.
- (42) Vinaixa, M.; Rodriguez, M. A.; Rull, A.; Beltran, R.; Blade, C.; Brezmes, J.; Canellas, N.; Joven, J.; Correig, X. *J. Proteome Res.* **2010**, 9, 2527-2538.

- (43) McClay, J. L.; Adkins, D. E.; Isern, N. G.; O'Connell, T. M.; Wooten, J. B.; Zedler, B. K.; Dasika, M. S.; Webb, B. T.; Webb-Robertson, B.-J.; Pounds, J. G.; Murrelle, E. L.; Leppert, M. F.; van, d. O. E. J. C. G. *J. Proteome Res.* **2010**, *9*, 3083-3090.
- (44) Halouska, S.; Fenton, R. J.; Barletta, R. G.; Powers, R. *ACS Chem. Biol.* **2012**, *7*, 166-171.
- (45) Powers, R. *Expert Opin. Drug Discovery* **2009**, *4*, 1077-1098.
- (46) Rosenblum, E. S.; Tjeerdema, R. S.; Viant, M. R. *Environ. Sci. Technol.* **2006**, *40*, 7077-7084.
- (47) Forgue, P.; Halouska, S.; Werth, M.; Xu, K.; Harris, S.; Powers, R. *J. Proteome Res.* **2006**, *5*, 1916-1923.
- (48) Fukuhara, K.; Ohno, A.; Ando, Y.; Yamoto, T.; Okuda, H. *Drug Metab. Pharmacokinet.* **2011**, *26*, 399-406.
- (49) McKelvie, J. R.; Wolfe, D. M.; Celejewski, M. A.; Alae, M.; Simpson, A. J.; Simpson, M. J. *Environ. Pollut.* **2011**, *159*, 3620-3626.
- (50) Yuliana, N. D.; Khatib, A.; Verpoorte, R.; Choi, Y. H. *Anal. Chem.* **2011**, *83*, 6902-6906.
- (51) Lopez-Gresa, M. P.; Maltese, F.; Belles, J. M.; Conejero, V.; Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Phytochem. Anal.* **2010**, *21*, 89-94.
- (52) Lima, M. R. M.; Felgueiras, M. L.; Graca, G.; Rodrigues, J. E. A.; Barros, A.; Gil, A. M.; Dias, A. C. P. *J. Exp. Bot.* **2010**, *61*, 4033-4042.
- (53) Hu, J. Z. *Methods Mol. Biol.* **2011**, *708*, 335-364.
- (54) Bankefors, J.; Kaszowska, M.; Schlechtriem, C.; Pickova, J.; Braennaes, E.; Edebo, L.; Kiessling, A.; Sandstroem, C. *Food Chem.* **2011**, *129*, 1397-1405.
- (55) Ding, G.; Hu, H.; Li, L.; Mao, X. a.; Ye, C. *Prog. Nat. Sci.* **1997**, *7*, 41-45.
- (56) Wang, Y.; Bollard, M. E.; Keun, H.; Antti, H.; Beckonert, O.; Ebbels, T. M.; Lindon, J. C.; Holmes, E.; Tang, H.; Nicholson, J. K. *Anal. Biochem.* **2003**, *323*, 26-32.
- (57) Miller, J. M.; McNair, H. *Gas chromatography* **2000**; John Wiley & Sons, Inc.; 217-254.

- (58) Tanaka, K.; West-Dull, A.; Hine, D. G.; Lynn, T. B.; Lowe, T. *Clin. Chem.* **1980**, *26*, 1847-1853.
- (59) Drozd, J. *Chemical Derivatizations in Gas Chromatography*; Elsevier: Amsterdam, **1981**.
- (60) Lehrer, M. *Clin. Lab. Med.* **1990**, *10*, 271-288.
- (61) Brooks, C. J. W.; Anthony, G. M.; Rocher, P.; Middleditch, B. S.; Stillwell, W. G. *J. Chromatogr. Sci.* **1971**, *9*, 35-43.
- (62) Lo, S. F.; Young, V.; Rhead, W. J. *Methods Mol. Biol.* **2010**, *603*, 433-443.
- (63) Jones, P. M.; Bennett, M. J. *Methods Mol. Biol.* **2010**, *603*, 423-431.
- (64) Costa, C. G.; Guerand, W. S.; Struys, E. A.; Holwerda, U.; ten, B. H. J.; Tavares, d. A. I.; Duran, M.; Jakobs, C. *J. Pharm. Biomed. Anal.* **2000**, *21*, 1215-1224.
- (65) Kimura, M.; Yamaguchi, S. *J. Chromatogr., B: Biomed. Sci. Appl.* **1999**, *731*, 105-110.
- (66) Berkov, S.; Bastida, J.; Viladomat, F.; Codina, C. *Talanta* **2011**, *83*, 1455-1465.
- (67) Qualley, A. V.; Dudareva, N. *Methods Mol. Biol.* **2009**, *553*, 329-343.
- (68) Shuman, J. L.; Cortes, D. F.; Armenta, J. M.; Pokrzywa, R. M.; Mendes, P.; Shulaev, V. *Methods Mol. Biol.* **2011**, *678*, 229-246.
- (69) Kuhara, T. *Rinsho Kagaku (Nippon Rinsho Kagakkai)* **2010**, *39*, 166-174.
- (70) Lu, H.; Gan, D.; Zhang, Z.; Liang, Y. *Metabolomics* **2011**, *7*, 191-205.
- (71) Aura, A.-M.; Mattila, I.; Seppanen-Laakso, T.; Miettinen, J.; Oksman-Caldentey, K.-M.; Oresic, M. *Phytochem. Lett.* **2008**, *1*, 18-22.
- (72) Liu, Z.; Phillips, J. B. *J. Chromatogr. Sci.* **1991**, *29*, 227-231.
- (73) Culbertson, A. W.; Williams, W. B.; McKee, A. G.; Zhang, X.; Marchs, K. L.; Naylor, S.; Valentine, S. J. *LCGC North Am.* **2008**, 74-81.
- (74) Hope, J. L.; Prazen, B. J.; Nilsson, E. J.; Lidstrom, M. E.; Synovec, R. E. *Talanta* **2005**, *65*, 380-388.
- (75) Mohler, R. E.; Dombek, K. M.; Hoggard, J. C.; Pierce, K. M.; Young, E. T.; Synovec, R. E. *Analyst* **2007**, *132*, 756-767.

- (76) Mohler, R. E.; Tu, B. P.; Dombek, K. M.; Hoggard, J. C.; Young, E. T.; Synovec, R. E. *J Chromatogr A* **2008**, *1186*, 401-411.
- (77) Lapainis, T.; Rubakhin, S. S.; Sweedler, J. V. *Anal. Chem.* **2009**, *81*, 5858-5864.
- (78) Ramautar, R.; van der Plas, A. A.; Nevedomskaya, E.; Derks, R. J. E.; Somsen, G. W.; de Jong, G. J.; van Hilten, J. J.; Deelder, A. M.; Mayboroda, O. A. *J. Proteome Res.* **2009**, *8*, 5559-5567.
- (79) Sirén, Heli; Seppänen-Laakso, T.; Orešič, M. *J. Chromatogr. B* **2008**, *871*, 375-382.
- (80) Soga, T.; Igarashi, K.; Ito, C.; Mizobuchi, K.; Zimmermann, H.-P.; Tomita, M. *Anal. Chem.* **2009**, *81*, 6165-6174.
- (81) Urakami, K.; Zangiaccomi, V.; Yamaguchi, K.; Kusuhara, M. *Biomed. Res.* **2010**, *31*, 161-163.
- (82) Baidoo, E. E. K.; Benke, P. I.; Neususs, C.; Pelzing, M.; Kruppa, G.; Leary, J. A.; Keasling, J. D. *Anal. Chem.* **2008**, *80*, 3112-3122.
- (83) Allard, E.; Bałckstrołm, D.; Danielsson, R.; Sjolberg, P. J. R.; Bergquist, J. *Anal. Chem.* **2008**, *80*, 8946-8955.
- (84) t'Kindt, R.; De, V. L.; Storme, M.; Deforce, D.; Van, B. J. *J. Chromatogr., B: Analyt. Technol. Biomed. Life Sci.* **2008**, *871*, 37-43.
- (85) Mallet, A. I.; Down, S. *Dictionary of Mass Spectrometry*; John Wiley & Sons: West Sussex, 2009.
- (86) Chambers, E.; Wagrowski-Diehl, D. M.; Lu, Z.; Mazzeo, J. R. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2007**, *852*, 22-34.
- (87) U.S.FDA *Bioanalytical Method Validation* **May 2001**.
- (88) Xiao, J. F.; Zhou, B.; Resson, H. W. *TrAC, Trends Anal. Chem.* **2012**, *32*, 1-14.
- (89) Cubbon, S.; Antonio, C.; Wilson, J.; Thomas-Oates, J. *Mass Spectrom. Rev.* **2010**, *29*, 671-684.
- (90) Tiller, P. R.; Romanyshyn, L. A.; Neue, U. D. *Anal. Bioanal. Chem.* **2003**, *377*, 788-802.
- (91) van Deemter, J. J.; Zuiderweg, F. J.; Klinkenberg, A. *Chem. Eng. Sci.* **1956**, *5*, 271-289.

- (92) Miller, J. M. *Chromatography Concepts and Contrasts*, 2nd ed.; Wiley: Hoboken, 2005.
- (93) Wang, X.; Wang, H.; Zhang, A.; Lu, X.; Sun, H.; Dong, H.; Wang, P. *J. Proteome Res.* **2012**, *11*, 1284-1301.
- (94) Wang, X.; Sun, H.; Zhang, A.; Wang, P.; Han, Y. *J. Sep. Sci.* **2011**, *34*, 3451-3459.
- (95) Lee, S. H.; An, J. H.; Park, H.-M.; Jung, B. H. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2012**, 887-888, 8-18.
- (96) Li, H.; Deng, Z.; Liu, R.; Loewen, S.; Tsao, R. *Food Chem.* **2012**, *132*, 508-517.
- (97) Huesgen, A. G.; Naegele, E. *Agilent Technologies publication 5990-5069EN December 2009*.
- (98) Huber, U. *Agilent Technologies publication 5988-7831EN May 2007*.
- (99) Kebarle, P.; Peschke, M. *Anal. Chim. Acta* **2000**, *406*, 11-35.
- (100) Cech, N. B.; Enke, C. G. *Mass Spectrom. Rev.* **2002**, *20*, 362-387.
- (101) Dass, C. *Fundamentals of Contemporary Mass Spectrometry*; Wiley Hoboken, New Jersey, 2007.
- (102) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240-2249.
- (103) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287-2294.
- (104) Constantopoulos, T. L.; Jackson, G. S.; Enke, C. G. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 625-634.
- (105) Cech, N. B.; Enke, C. G. *Anal. Chem.* **2001**, *73*, 4632-4639.
- (106) Asperger, A.; Efer, J.; Koal, T.; Engewald, W. *J. Chromatogr., A* **2001**, *937*, 65-72.
- (107) Stephenson, J. L., Jr.; McLuckey, S. A. *Anal. Chem.* **1996**, *68*, 4026-4032.
- (108) Mei, H. In *Using Mass Spectrometry for Drug Metabolism Studies*; Korfmacher, W. A., Ed.; CRC Press: Boca Raton, 2005, pp 103 - 150.
- (109) Tong, X. S.; Wang, J.; Zheng, S.; Pivnichny, J. V.; Griffin, P. R.; Shen, X.; Donnelly, M.; Vakerich, K.; Nunes, C.; Fenyk-Melody, J. *Anal. Chem.* **2002**, *74*, 6305-6313.

- (110) Amad, M. a. H.; Cech, N. B.; Jackson, G. S.; Enke, C. G. *J. Mass Spectrom.* **2000**, *35*, 784-789.
- (111) Kuhlmann, F. E.; Apffel, A.; Fischer, S. M.; Goldberg, G.; Goodley, P. C. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1221-1225.
- (112) Mei, H.; Hsieh, Y.; Nardo, C.; Xu, X.; Wang, S.; Ng, K.; Korfmacher, W. *A. Rapid Commun. Mass Spectrom.* **2003**, *17*, 97-103.
- (113) Trufelli, H.; Palma, P.; Famigliani, G.; Cappiello, A. *Mass Spectrom. Rev.* **2011**, *30*, 491-509.
- (114) Rossi, D. T.; Zhang, N. *J. Chromatogr., A* **2000**, *885*, 97-113.
- (115) Ding, J.; Neue, U. *D. Rapid Commun. Mass Spectrom.* **1999**, *13*, 2151-2159.
- (116) Ito, S.; Tsukada, K. *J. Chromatogr., A* **2002**, *943*, 39-46.
- (117) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.
- (118) Hopfgartner, G.; Varesio, E.; Tschaepaet, V.; Grivet, C.; Bourgoigne, E.; Leuthold, L. A. *J. Mass Spectrom.* **2004**, *39*, 845-855.
- (119) Halket, J. M.; Waterman, D.; Przyborowska, A. M.; Patel, R. K. P.; Fraser, P. D.; Bramley, P. M. *J. Exp. Bot.* **2005**, *56*, 219-243.
- (120) Deng, P.; Zhan, Y.; Chen, X.; Zhong, D. *Bioanalysis* **2012**, *4*, 49-69.
- (121) Van, B. G. J.; Asano, K. G. *Anal. Chem.* **1994**, *66*, 2096-2102.
- (122) Cech, N. B.; Krone, J. R.; Enke, C. G. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1040-1044.
- (123) Van, B. G. J.; Quirke, J. M. E.; Tagani, R. A.; Dilley, A. S.; Covey, T. R. *Anal. Chem.* **1998**, *70*, 1544-1554.
- (124) Okamoto, M.; Takahashi, K.-I.; Doi, T. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 641-643.
- (125) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. *Metabolomics* **2007**, *3*, 211-221.

- (126) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. *Ther. Drug Monit.* **2005**, *27*, 747-751.
- (127) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T. *J. Mass Spectrom.* **2010**, *45*, 703-714.
- (128) Kind, T.; Fiehn, O. *BMC Bioinf.* **2006**, *7*, 234-243.
- (129) Levsen, K.; Schiebel, H.-M.; Terlouw, J. K.; Jobst, K. J.; Elend, M.; Preiss, A.; Thiele, H.; Ingendoh, A. *J. Mass Spectrom.* **2007**, *42*, 1024-1044.
- (130) Heinonen, M.; Rantanen, A.; Mielikainen, T.; Kokkonen, J.; Kiuru, J.; Ketola, R. A.; Rousu, J. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 3043-3052.
- (131) Wolf, S.; Schmidt, S.; Mueller-Hannemann, M.; Neumann, S. *BMC Bioinf.* **2010**, *11*, 148-159.
- (132) Bourcier, S.; Hoppilliard, Y. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 93-103.
- (133) Ardrey, R. E.; *Liquid Chromatography Mass Spectrometry: An Introduction* Wiley, **2003**, pp 7-33.
- (134) Turowski, M.; Yamakawa, N.; Meller, J.; Kimata, K.; Ikegami, T.; Hosoya, K.; Tanaka, N.; Thornton, E. R. *J. Am. Chem. Soc.* **2003**, *125*, 13836-13849.
- (135) Wade, D. *Chem.-Biol. Interact.* **1999**, *117*, 191-217.
- (136) Lee, J. W.; Devanarayan, V.; Barrett, Y. C.; Weiner, R.; Allinson, J.; Fountain, S.; Keller, S.; Weinryb, I.; Green, M.; Duan, L.; Rogers, J. A.; Millham, R.; O'Brien, P. J.; Sailstad, J.; Khan, M.; Ray, C.; Wagner, J. A. *Pharm. Res.* **2006**, *23*, 312-328.
- (137) Houghton, R.; Pita, C. H.; Ward, I.; Macarthur, R. *Bioanalysis* **2009**, *1*, 1365-1374.
- (138) van, d. M. N. C. *TrAC, Trends Anal. Chem.* **2008**, *27*, 924-933.

- (139) Wilson, S. F.; James, C. A.; Zhu, X. C.; Davis, M. T.; Rose, M. J. *J. Pharm. Biomed. Anal.* **2011**, *56*, 315-323.
- (140) Boomsma, F.; Alberts, G.; Vaneijk, L.; Manintveld, A. J.; Schalekamp, M. *Clin. Chem.* **1993**, *39*, 2503-2508.
- (141) Zhang, Y. H.; Dufield, D.; Klover, J.; Li, W. L.; Szekely-Klepser, G.; Lepsy, C.; Sadagopan, N. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2009**, *877*, 513-520.
- (142) Zhang, G.; Zhang, Y.; Fast, D. M.; Lin, Z.; Steenwyk, R. *Anal. Biochem.* **2011**, *416*, 45-52.
- (143) Li, W. L.; Cohen, L. H. *Anal. Chem.* **2003**, *75*, 5854-5859.
- (144) Rinaldo, P.; Matern, D. *Annu. Rev. Physiol.* **2002**, *64*, 477-502.
- (145) Gregersen, N.; Andresen, B. S.; Pedersen, C. B.; Olsen, R. K. J.; Corydon, T. J.; Bross, P. *J. Inherited Metab. Dis.* **2008**, *31*, 643-657.
- (146) Bartlett, K.; Gompertz, D. *Biochem. Med.* **1974**, *10*, 15-23.
- (147) Gron, I.; Gregersen, N.; Kolvraa, S.; Rasmussen, K. *Biochem. Med.* **1978**, *19*, 133-140.
- (148) Bonafe, L.; Troxler, H.; Kuster, T.; Heizmann, C. W.; Chamoles, N. A.; Burlina, A. B.; Blau, N. *Mol. Genet. Metab.* **2000**, *69*, 302-311.
- (149) la Marca, G.; Rizzo, C. In *Metabolic Profiling: Methods and Protocols*; Metz, T. O., Ed., 2010; Vol. 708, pp 73-98.
- (150) Therrell, B. L.; Adams, J. *J. Inherited Metab. Dis.* **2007**, *30*, 447-465.
- (151) Charles R, R. *Seminars in Neonatology* **2002**, *7*, 37-47.
- (152) Ogier de Baulny, H.; Saudubray, J. M. *Seminars in Neonatology* **2002**, *7*, 65-74.
- (153) Carter, S. M. B.; Watson, D. G.; Midgley, J. M.; Logan, R. W. *J. Chromatogr., B: Biomed. Appl.* **1996**, *677*, 29-35.
- (154) Grant, W. D.; Bonham Carter, S. M.; McKechnie, J. M.; Pierotti, A.; Galloway, P. *J. Inherited Metab. Dis.* **2007**, *30*, 31-31.
- (155) Lewis-Stanislaus, A. E.; Li, L. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 2105-2116.

- (156) Ombrone, D.; Salvatore, F.; Ruoppolo, M. *Anal. Biochem.* **2011**, *417*, 122-128.

## Chapter 2. Ultra-Performance Liquid Chromatography Tandem Mass Spectrometry Detection of Acylglycines\*

### 2.1 Introduction

Metabolomics is an emerging field that is poised to play a significant role in many disciplines of biosciences. One of the current challenges in metabolomics is to profile the metabolome in a comprehensive manner, ideally covering the entire set of metabolites present in a biological system. However, due to technical limitations, only a small portion of the metabolome is analyzed. There is a great need to expand the metabolome coverage to reveal subtle changes of the metabolome in biological studies or disease biomarker discovery. Liquid chromatography mass spectrometry (LC-MS) is a sensitive technique that can detect many metabolites in a metabolome sample.<sup>1</sup> However, metabolite identification from the mass spectral data alone is often difficult.<sup>2-6</sup> There are only a limited number of metabolite standards available for spectral comparison to identify unknowns. For example, in the Human Metabolome Database (HMDB),<sup>7</sup> product ion spectra of about 900 known standards obtained by tandem mass spectrometry (MS/MS) are included. To create this MS/MS spectral library, almost all the possible commercial sources were utilized to acquire the available metabolite standards. This number is still quite small compared to the size of the human metabolome; the total number of human metabolites is unknown, but the HMDB contains over 8000 entries of endogenous human metabolites. It is therefore necessary to expand this spectral library to include many more metabolites found in biological sources, such as human biofluids. A comprehensive spectral library will enable many metabolomics researchers to

---

\* A version of this chapter has been published as Lewis-Stanislaus, A.E., Li, L., **2010**, "A Method for the Comprehensive Analysis of Urinary Acylglycines by Using Ultra-Performance Liquid Chromatography Quadrupole Linear Ion Trap Mass Spectrometry", *J. Am. Soc. Mass Spectrom.* **21**, 2105 – 2116.

take advantage of the database resource for identifying metabolites for biological studies or biomarker discovery.

A strategy to expand metabolomic coverage and the spectral library systematically is currently being practiced. This includes detecting and identifying metabolites sharing one or more similar structural moieties, such as glycine. Because some standards from this group are available, analyzing the unknown metabolites of the same group becomes more manageable. This work demonstrates the development and application of this strategy to detect and identify unknown acylglycines. Acylglycines are an important class of metabolites that have been used in the diagnosis of several inborn errors of metabolism (IEM).<sup>8</sup>

IEMs are inherited disorders that are due to gene defects coding specific enzymes involved in the metabolism of amino acids or organic acids.<sup>9</sup> One of the main classes of IEMs consists of disorders of fatty acid oxidation and mitochondrial metabolism. In individuals with fatty acid oxidation disorder, a wide array of symptoms is due to the toxic fatty acid acyl-coA esters that accumulate in the mitochondria. Glycine conjugation with these acyl-coA esters has been of clinical interest due to their role in the detoxification process. Increased concentration of urinary acylglycines is generally indicative of IEMs.<sup>9-</sup>

11

The most widely used analytical methods for determining acylglycines in urine include solvent extraction, derivatization, followed by separation and detection using gas chromatography mass spectrometry (GC-MS).<sup>12-17</sup> Although this established method offers many advantages, it requires specific derivatization reactions for analyzing the polar acylglycines. Also diagnoses of some IEMs can be challenging due to the relatively low sensitivity of the analytical process.<sup>18</sup> The application of tandem mass spectrometry (MS/MS) has been used as an alternative to GC-MS and as a high-throughput method. Analysis by fast atom bombardment (FAB) and lately electrospray ionization (ESI) have been well

documented.<sup>19-22</sup> Although these methods can detect a larger number of disorders in a single run, they lack the ability to distinguish between isomers and detect relatively low abundance acylglycines, thus may not be able to distinguish between certain disorders.<sup>22</sup>

In this work, an LC-MS method is reported based on the use of triple quadrupole linear ion trap (QTRAP<sup>®</sup>) mass spectrometer that offers higher sensitivity and more data-dependent scanning modes than a conventional triple quadrupole tandem mass spectrometer.<sup>23</sup> Separation is carried out by using ultra performance liquid chromatography (UPLC), which allows for high resolution separation of isomers and closely related acylglycines. The aim of this work is to identify as many unknown acylglycines as possible in order to expand the overall molecular coverage of this important class of metabolites. The described method provides a sensitive analytical procedure to detect previously undetected acylglycines in the urine of healthy individuals; 65 acylglycines were found using the new method, compared to the currently known 18 compounds.

## **2.2 Experimental**

### **2.2.1 Materials and Reagents**

All chemicals except those noted were purchased from Sigma-Aldrich Canada (Oakville, ON, Canada). Human pooled liver microsomes and nicotinamide adenine dinucleotide phosphate (NADPH) regeneration solutions A and B were purchased from BD Gentest (Franklin Lakes, NJ, USA). Optima grade methanol and water, Optima LCMS grade acetonitrile and acetic anhydride and pyridine were purchased from Fisher Scientific (Ottawa, ON, Canada). HPLC-grade formic acid and ammonium hydroxide were obtained from Fluka (Milwaukee, OH, USA). The standard acylglycines used were dimethylglycine, phenylglycine, acetylglycine, propionylglycine, isobutyrylglycine, butyrylglycine, 4-hydroxyphenylacetylglycine, 2-methylbutyrylglycine, isovalerylglycine,

valerylglycine, tiglylglycine, 3-methylglycine, suberylglycine, glutarylglycine, phenylacetylglycine, phenylpropionylglycine, hexanoylglycine, octanoylglycine and hippuric acid.

### **2.2.2 Samples**

This study was conducted in accordance with the Arts, Science & Law Research Ethics Board policy at the University of Alberta. Urine was collected from six healthy volunteers who were not on any special diet. The volunteers were all adults, ranging in age from 24 to 38 years old. The urine samples were collected as first morning void samples for all six volunteers and urine was also collected over 12 hours for two of those individuals. The samples were centrifuged, aliquoted and stored at -20 °C without any added preservatives until further analysis. For long-term storage, the samples were kept at -80 °C.

### **2.2.3 Microsome Incubation**

Acylglycines standards (50 µM) were individually incubated with human liver microsomes (2 mg/mL) in 100 mM potassium phosphate, pH 7.4. The standards were first pre-incubated at 37 °C for five min, followed by the addition of NADPH regenerating solution, containing components A (26.1 mM NADP<sup>+</sup>, 66 mM glucose-6-phosphate and 66 mM MgCl<sub>2</sub> in H<sub>2</sub>O) and B (40 U/mL glucose-6-phosphate dehydrogenase in 5 mM sodium citrate) in 100 mM potassium phosphate, pH 7.4. The solutions were incubated, with shaking, in a 37 °C incubator for periods of 30 minutes, 1, 3, 6 and 24 hrs. Control incubations omitting either the standards, NADPH regenerating solution or microsomes were also performed by substituting an equal volume of potassium phosphate buffer. The reactions were then terminated by the addition of 5% acetic acid in acetonitrile (v/v). The samples were centrifuged at 14,000 g for 5 minutes and the supernatants were subjected to solid phase extraction (SPE).

#### 2.2.4 Solid Phase Extraction

Oasis mixed-mode anion exchange (MAX) cartridges (Waters, Mississauga, ON) were operated at a flow rate of 1 mL/min using a vacuum manifold (Alltech from Fisher Scientific, Ottawa, ON). The 30-mg or 60-mg cartridges were preconditioned with 1 mL acetonitrile followed by equilibration with 1 mL water. Urine (1 mL) was loaded onto the columns and the columns were washed with 1 mL 5% ammonium hydroxide solution. Acylglycines were eluted in two fractions, 1 mL 2% formic acid in 40% acetonitrile/60% water followed by 1 mL 2% formic acid in acetonitrile. The eluents were evaporated to dryness, using a vacuum centrifuge concentrator, and reconstituted in 100  $\mu$ L 4% acetonitrile, 0.1% formic acid in water. Samples extracted were urine, urine spiked with standards, and microsomal incubations.

#### 2.2.5 Esterification

In order to confirm the identification of acylglycines found, methylation and acetylation were performed. These reactions were optimized using standards and confirmed to be useful in assisting in compound identification. For the urine samples, after extraction of the samples by SPE, the eluents were evaporated to dryness in a vacuum centrifuge concentrator. Three hundred microliters of 3 M methanolic HCl was added to the dried extract in a reaction vial and allowed to react at 65 °C for 15 minutes.<sup>20</sup> The reaction mixture was then divided into two equal portions; one portion was evaporated under a stream of nitrogen and reconstituted in 100  $\mu$ L 10/90 ACN/H<sub>2</sub>O (v/v), 0.1% formic acid. The other portion was also evaporated under a stream of nitrogen and treated with 100  $\mu$ L of 50:50 acetic anhydride/pyridine (v/v) and allowed to react at room temperature for 1 hour. The solvents were evaporated under a stream of nitrogen and the residue was reconstituted in 100  $\mu$ L 10/90 ACN/H<sub>2</sub>O (v/v), 0.1% formic acid.

### 2.2.6 UPLC Separation

Chromatographic separation was done on a Waters ACQUITY UPLC system with a 1.7- $\mu\text{m}$  bridged ethylene hybrid (BEH), 150 mm x 1.0 mm C<sub>18</sub> column. The column was maintained at ambient temperature. Elution was done according to the following method: 100% A for 11 minutes, then a linear gradient of 0-35% B over 50 minutes, 35-100% B over 5 minutes and held at 100% B for 9 minutes, where mobile phase A consisted of 4% acetonitrile, 0.1% formic acid in water and B consisted of 0.1% formic acid acetonitrile. The flow rate of the method was 0.050 mL/min and 5.0  $\mu\text{L}$  of each sample was injected onto the column.

### 2.2.7 Mass Spectrometry

Mass analysis was carried out in an ABI 4000 QTRAP<sup>®</sup> mass spectrometer equipped with a TurboIonSpray source (Applied Biosystems, Foster City, CA). The UPLC and mass spectrometer were both controlled by Analyst software v.1.5 from Applied Biosystems. The mass spectrometer was operated in both positive and negative ESI mode. General mass spectrometric conditions were: for positive mode, spray voltage, 4800 V; temperature, 200 °C; GS1, 40; GS2, 10; curtain gas, 10; CAD, high; declustering potential, 35; collision energy, 22 eV; for negative mode: spray voltage, -3100 V; temperature, 200 °C; GS1, 40; GS2, 10; curtain gas, 10; CAD, high; declustering potential, -40; collision energy, -20 eV. The acquisition method consisted of several information dependent acquisition (IDA) scan cycles including precursor ion, neutral loss or multiple reaction monitoring (MRM) as the survey scan and four dependent enhanced product ion (EPI) scans. IDA criteria were set to allow the four most intense peaks to trigger the EPI scans. The EPI scans were in the range of 50 - 500 Da, scanned at 250, 1000 and 4000 amu/s and source parameters were the same as mentioned above. The total duty cycle was approximately 1.2 s. Data was collected in profile mode and analyzed using Analyst v. 1.5 and LightSight v. 2.0.

## **2.2.8 Breakdown Graphs and Scan Modes**

To obtain optimal conditions for the detection of acylglycines, breakdown graphs from collision-induced dissociation (CID) of precursor ions were constructed. Solutions of each acylglycine standard were dissolved in water at a concentration of 0.5 mM and stored at -20 °C. These solutions were serially diluted to 10 µM in a mixture of 50:50 ACN/H<sub>2</sub>O (v/v), 0.1% formic acid and infused into the ion source with a syringe pump at a rate of 10 µL/min. Product ion spectra were obtained from the precursor ion in the positive mode by averaging 30 cycles. Scans were performed at different collision energies ranging from 5 to 40 eV, with a step size of 5 eV. Mass spectrometric conditions used are described above.

Based on the information obtained from the breakdown graphs, appropriate neutral loss and precursor ion scans were performed as survey scans in IDA experiments, which triggered four dependent EPI scans. MRM transitions were set up based on masses of acylglycines that were detected in the neutral loss and precursor ion scans and those acylglycines likely to be found in urine but whose signals were too low to be detected by the constant neutral loss and precursor ion scans. IDA experiments were created with these MRM transitions as a survey scan, in an experiment similar to the precursor and neutral loss methods.

## **2.3 Results and Discussion**

### **2.3.1 Sample Handling Issues**

The structures of the nineteen acylglycine standards used in this study are shown in Figure 2-1. The main purpose of this work was to develop a means of detecting as many acylglycines as possible. Each step of the analysis procedure was therefore optimized to maximize the performance of the technique. For example, current methods of extracting acylglycines from urine include liquid-

liquid extraction followed by derivatization.<sup>12, 13</sup> Human urine, however, is a complex mixture of salts, hydrophilic and hydrophobic compounds, peptides and proteins. In this work, a more selective sample extraction method was developed. In order to optimize the solid phase extraction procedure, two cartridges, hydrophilic-lipophilic balanced (HLB) and mixed-mode anion-exchange (MAX) cartridges, were tested. Table 2-1 shows a comparison of the extraction of seven acylglycines using both cartridges. The comparison is made using the peak areas of triplicate experiments.

Factors to optimize the conditions include composition, volumes and pH of the wash and eluting solutions. MAX cartridge was chosen based on its high selectivity in retaining the acylglycines. The optimal conditions are listed in the experimental section. This method was found to be effective in isolating acylglycines and organic acids from human urine; for acylglycine standards, recoveries of greater than 88% were obtained. The extraction method was also reproducible. As an example, four urine samples were extracted using four different SPE cartridges and were injected into the UPLC to compare the performance of the different MAX cartridges. The results shown in Figure 2-2 indicate excellent cartridge-to-cartridge reproducibility, signifying the robustness of the extraction procedure.

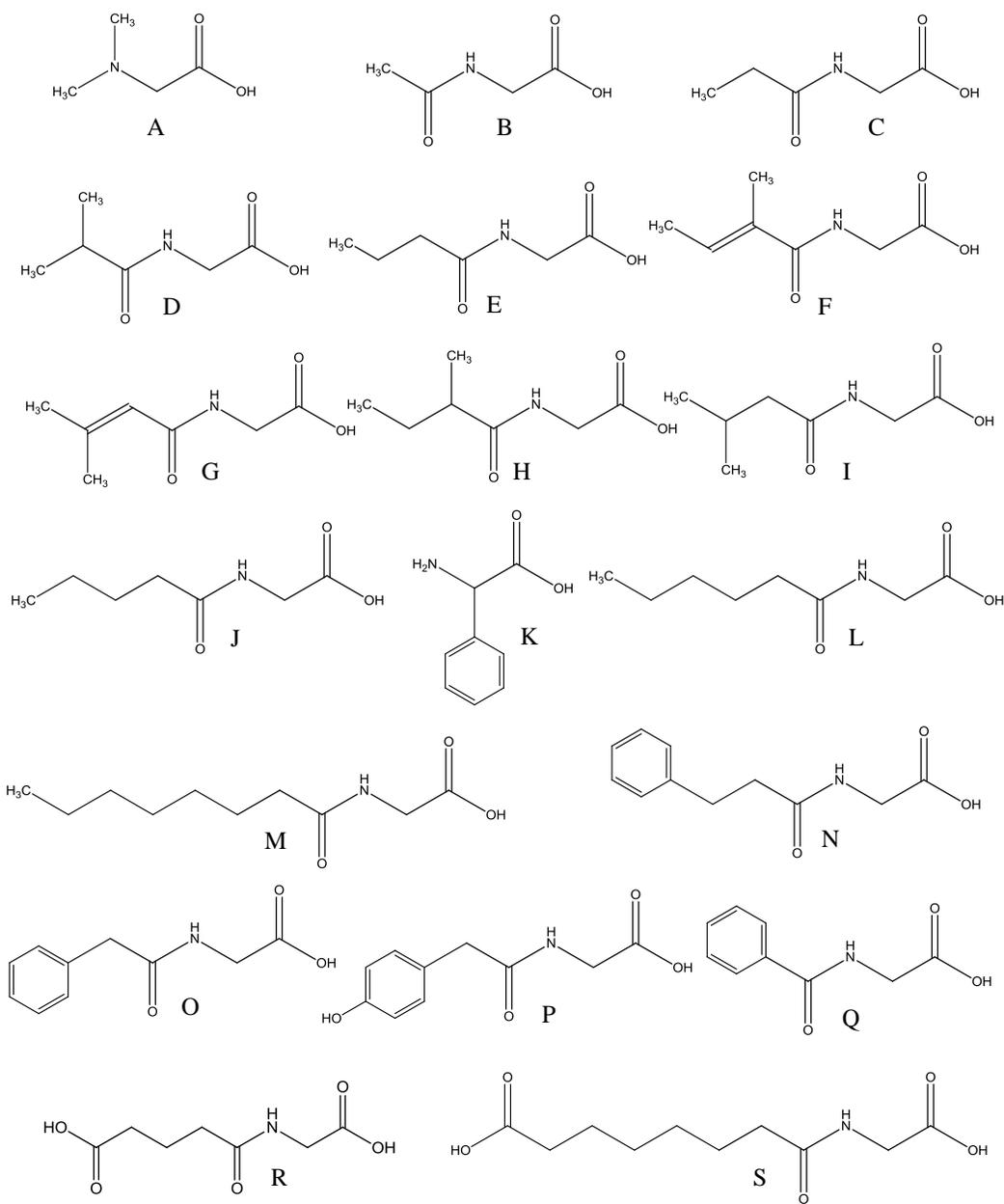


Figure 2-1 Structures of acylglycines: (A) dimethylglycine; (B) acetylglycine; (C) propionylglycine; (D) isobutyrylglycine; (E) butyrylglycine; (F) tiglylglycine; (G) 3-methylcrotonylglycine; (H) 2-methylbutyrylglycine; (I) isovalerylglycine; (J) valerylglycine; (K) phenylglycine; (L) hexanoylglycine; (M) octanoylglycine; (N) phenylpropionylglycine; (O) phenylacetylglycine; (P) 4-hydroxyphenylacetylglycine; (Q) hippuric acid (R) glutarylglycine and (S) suberylglycine

Table 2-1 Comparison of HLB and MAX solid phase extraction cartridges

Acylglycines	HLB		MAX	
	Peak Areas	% CV	Peak Areas	% CV
Propionylglycine	7.33E+05	11.01	6.42E+06	10.23
Isobutyrylglycine	8.55E+05	9.56	2.69E+06	8.66
3-Methylcrotonylglycine	6.10E+06	8.75	3.77E+07	9.17
Isovalerylglycine	5.59E+06	8.2	7.41E+07	11.88
Hexanoylglycine	4.78E+07	3.55	7.41E+07	5.79
Octanoylglycine	1.54E+06	9.11	3.67E+07	13.9
Phenylpropionylglycine	2.62E+08	6.23	7.03E+07	7.89

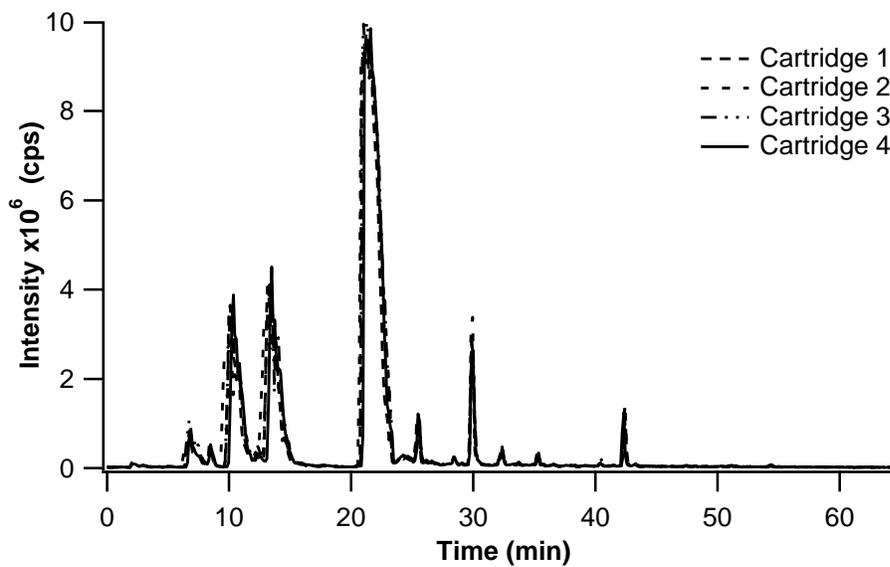


Figure 2-2 Cartridge reproducibility of the MAX cartridges

Another improvement is in the area of chromatographic separation. Instead of HPLC, UPLC was used in this work. UPLC capitalizes on the use of sub-2- $\mu\text{m}$  particles to offer superior efficiency and resolution compared to HPLC.<sup>24</sup> The higher efficiency translates into better detection sensitivity with narrower peaks which enhances the detection of acylglycines present in low concentrations in urine. The higher resolution is significant in separating isomeric and isobaric species. It was found that the UPLC method used was efficient in separating the isomers of acylglycines, which is very important in aiding the assignment of potential structures to unknown acylglycines. An example of the high resolution separation is shown in Figure 2-3. Baseline separation can be observed for the isomers of C<sub>5</sub>-glycine in pre-concentrated urine. In contrast, HPLC could not resolve these peaks well and some of the low intensity peaks were not observed. The signal to noise was also increased.

### 2.3.2 MS Fragmentation Pattern of Acylglycine Standards

Analyzing the MS fragmentation patterns of available standards can facilitate the development and optimization of a MS/MS method for sensitive and selective detection of a group of similar compounds including unknowns. In this work, nineteen acylglycines were used to generate the MS/MS fragmentation information. The fragmentation pathways of acylglycines and of the glycine-conjugated di-carboxylic acids are shown in Figure 2-4 and Figure 2-5 respectively. As these schemes show, common neutral losses of masses 18 (H<sub>2</sub>O), 46 (H<sub>2</sub>O + CO), 75 (NH<sub>2</sub>CH<sub>2</sub>COOH), and 103 (NH<sub>2</sub>CH<sub>2</sub>COOH + CO) are observed for most of the acylglycines. Less common fragments observed are the neutral losses of 60 (H<sub>2</sub>O + CH<sub>2</sub>=C=O), 93 (NH<sub>2</sub>CH<sub>2</sub>COOH + H<sub>2</sub>O), observed in straight chain acylglycines longer than 6 carbons and 117 (CH<sub>2</sub>=C(OH)NHCH<sub>2</sub>COOH), which is observed as a major fragment in phenylpropionylglycine, resulting in the stable tropyllium ion.

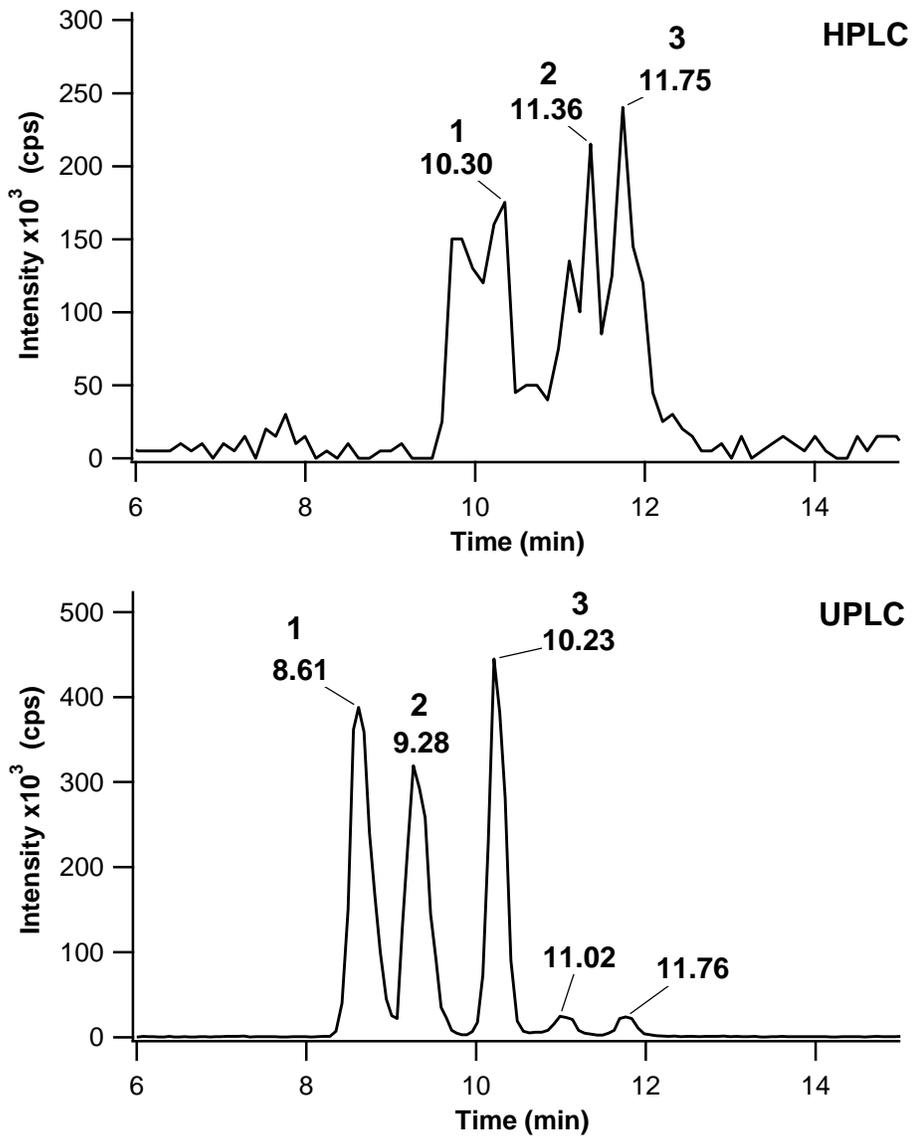


Figure 2-3 Comparison of HPLC and UPLC separation of C<sub>5</sub>-glycine isomers. Identity of peaks are (1) 2-methylbutyrylglycine; (2) isovalerylglycine and (3) valerylglycine.

Fragmentations of many acylglycines yield a major fragment ion at  $m/z$  76, which corresponds to the protonated glycine ( $\text{H}_3^+\text{NCH}_2\text{COOH}$ ). This fragmentation pathway occurs because aliphatic acylglycines lose the acyl moiety as a ketene. There must be an available proton on carbon 2 of the acyl group in order for this reaction to occur.<sup>19</sup>

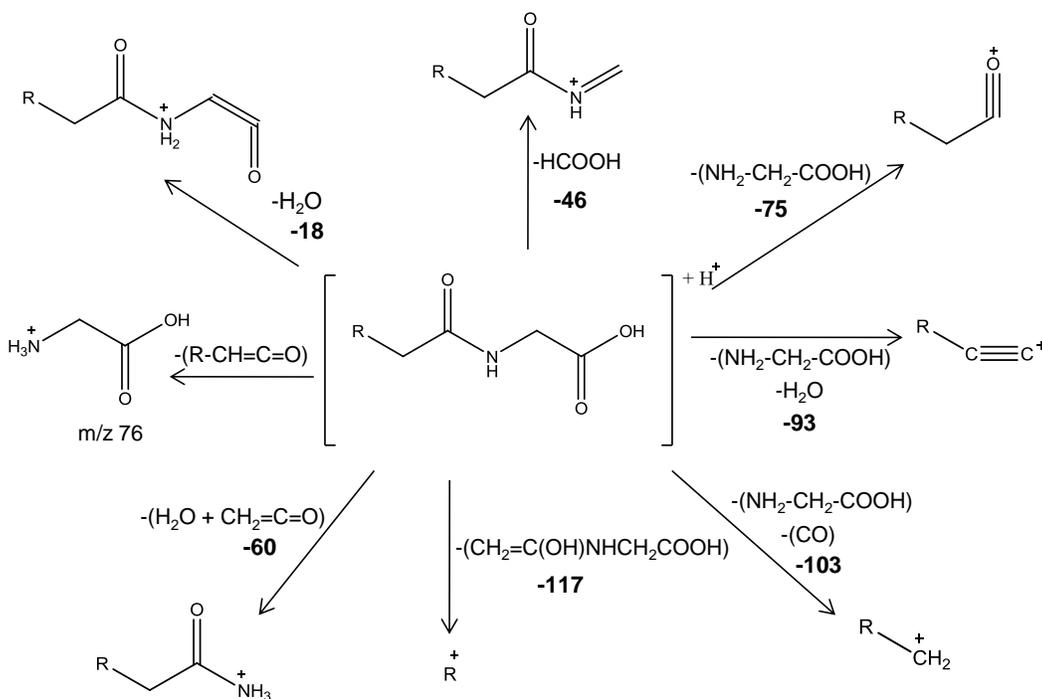


Figure 2-4 Proposed fragmentation pathway of acylglycines.

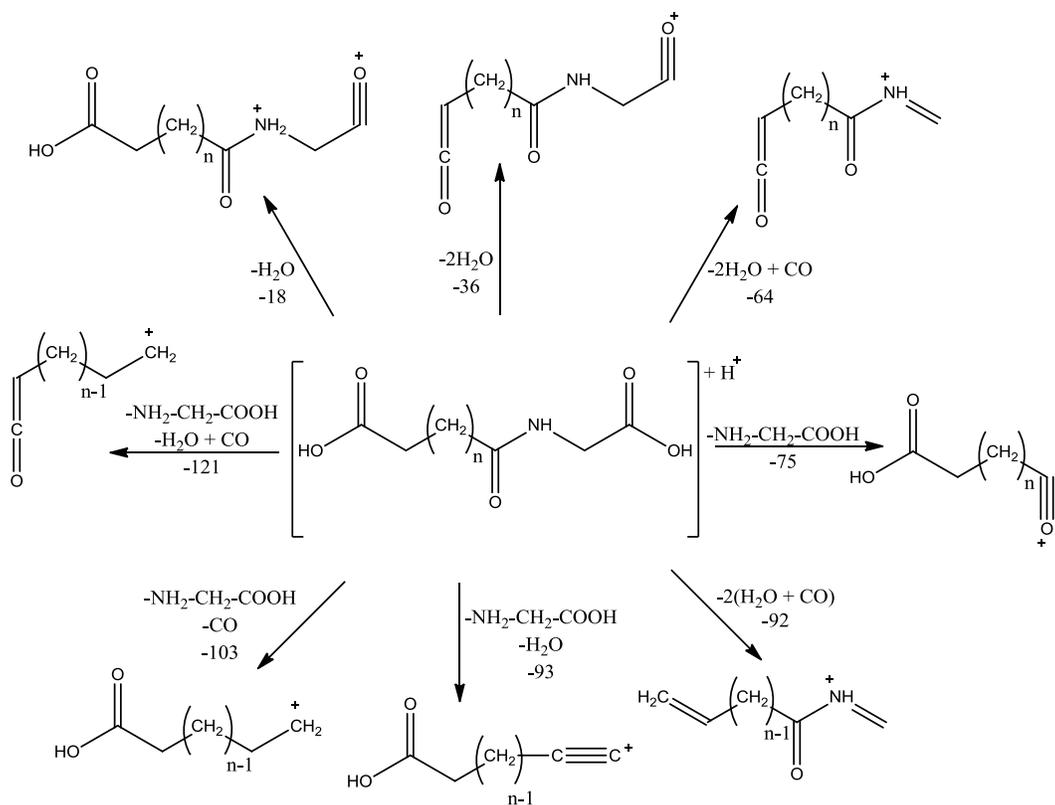


Figure 2-5 Proposed fragmentation pathway of glycine-conjugated di-carboxylic acids

Fragmentation in QTRAP<sup>®</sup> can be induced either inside the linear ion trap or in the collision cell in the triple quadrupole mode. The instrument is capable of all the scans common to the triple quadrupole instrument, such as precursor scanning (PC), constant neutral loss scanning (NL) and multiple-reaction monitoring (MRM), as well as trap scan modes enhanced product ion (EPI), enhanced mass (EMS) and enhanced resolution (ER) scans. The instrument can also perform both quadrupole and trap scans in a single run and switching from one to the other only takes a few milliseconds. This can be done in an information-dependent acquisition (IDA), which combines two or more scan modes in a single LC-MS/MS run. The mass spectrometer switches from the first scan or a survey scan usually done as an MS scan to a second data dependent MS/MS scan when an eluting peak rises above a predetermined threshold level.

In this work, MRM, constant neutral loss scanning and precursor scanning were used as survey scans with the enhanced product ion scans as the dependent MS/MS scans.

Currently, all acylglycines detected by direct infusion ESI-MS/MS use a precursor of  $m/z$  76 or if methylation is done, a precursor of  $m/z$  90 method.<sup>20</sup> In our work, the optimal energetic conditions for the detection of acylglycines were obtained from the breakdown graphs, several examples of which can be observed in Figure 2-6. The best choice for collision energy and MRM transitions can be readily obtained from these graphs. A typical breakdown graph in this work is shown in Figure 2-6(A) for valerylglycine. Most of the typical losses (as shown in Figure 2-4) are observed. At low to medium collision energy ( $15 < CE < 30$ ), the most abundant fragment ion is  $m/z$  76 and at high collision energy ( $CE > 30$ ), the fragment ion corresponding to the loss of 103 is the most abundant.

Although most acylglycines yield a fragment at  $m/z$  76, not all of them do. The breakdown graph of tiglylglycine, shown in Figure 2-6(B), demonstrates this. The major loss observed is the loss of 75 Da, the neutral glycine moiety. No fragment ion at  $m/z$  76 is generated because there is no available proton on carbon 2 of the acyl group.<sup>19, 22</sup> An IDA method, which uses a survey scan of a neutral loss of 75, should be used to detect acylglycines, including tiglylglycine, 3-methycrotonylglycine, hippuric acid, and hydroxyhippuric acid isomers, among others. These acylglycines would not be detected using a precursor of  $m/z$  76 method. Therefore the optimal transition for MRM is precursor ion fragmentation corresponding to the loss of 75 Da.

It is important to note that the fragment ion at  $m/z$  76 may not be the most intense fragment in the product ion spectra and if that transition is used in detection, sensitivity is lost. Examples of this are shown in Figure 2-6(C) and (D). Breakdown graphs of aromatic acylglycines reveal that the neutral loss of 103 is the most abundant fragment and not the fragment at  $m/z$  76. The optimal conditions for the detection of aromatic acylglycines at a CE of 20 is an MRM

transition from the precursor ion to the fragment ion produced with the loss of 103 (e.g., for phenylacetyl-glycine,  $m/z$  194  $\rightarrow$  91). Other trends observed in the fragmentation of acylglycines are the following: losses of 18 and 46 are usually low in intensity and sometimes not observed, especially in aromatic glycines and glycines conjugated to unsaturated fatty acids. Fragment ion  $m/z$  76 is the major fragment for straight and branched chain acylglycines. Loss of 103 is more intense in branched chains than in straight chains, especially when it is branched at the 2-position of the fatty acid chain. For the glycine conjugates of dicarboxylic acid, the fragment ion at  $m/z$  76 is not the major product ion. In this case, loss of 64 fragment ion is the most abundant fragment ion at a CE of 20 eV. Losses of 18 and 46 are higher in intensities than in the other acylglycines.

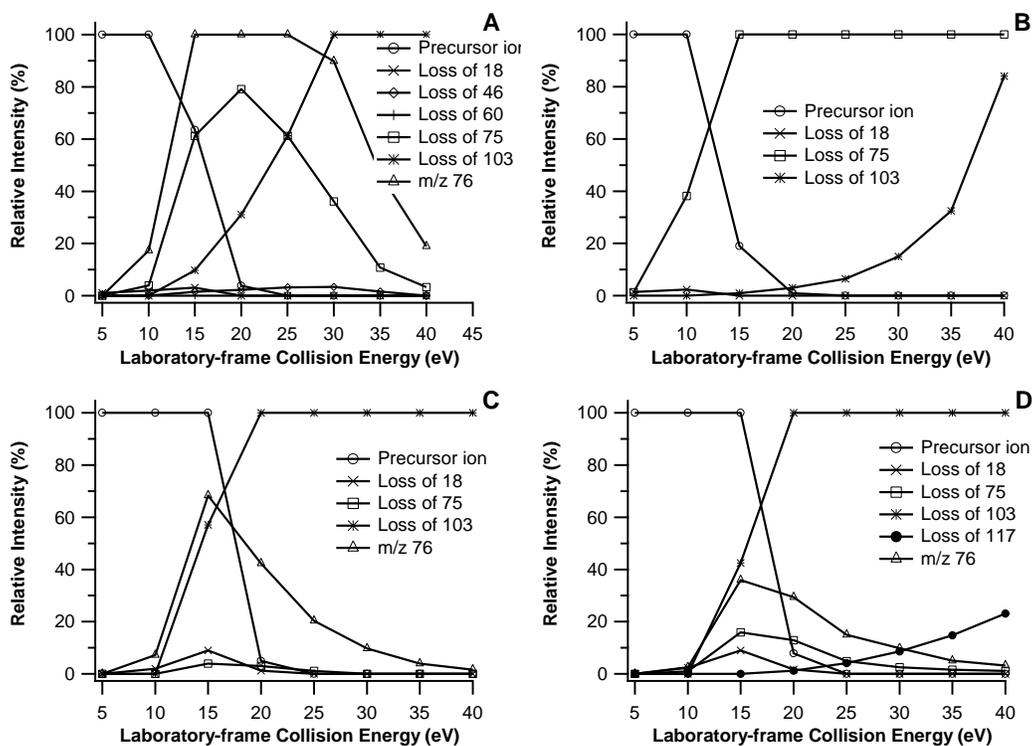


Figure 2-6 Breakdown graphs of (A) valeryl-glycine; (B) tiglylglycine; (C) phenylacetyl-glycine and (D) phenylpropionyl-glycine

### 2.3.3 MS Scan Modes for Sample Analysis

Because most acylglycines yield an intense fragment ion at  $m/z$  76 in the positive mode, an IDA experiment was performed using precursor scanning as a survey scan followed by an enhanced product ion scan as the data dependent MS/MS scan. Acylglycines also possess a carboxylic acid group, so they can be analyzed in the negative ionization mode. In the negative ion mode they lose the deprotonated glycine moiety ( $\text{H}_2\text{NCH}_2\text{COO}^-$ ,  $m/z$  74). A similar IDA experiment was performed with a precursor of  $m/z$  74 scan as the survey scan. A comparison of the positive and negative precursor scans is shown in Figure 2-7 (A) and (B). Negative ion scan did not significantly improve the selectivity and sensitivity in comparison to the positive mode. Most acylglycines were detected in both modes and the overall intensities were quite similar. Some acylglycines that do not generate a fragment ion at  $m/z$  76 will still generate the fragment ion at  $m/z$  74, such as tiglylglycine and 3-methylcrotonylglycine. Several earlier eluting peaks like acetylglycine, isobutyrylglycine, butyrylglycine and hydroxyphenylacetylglycine were detected at considerably higher intensities using the positive ionization mode. Due to this and the fact that positive MS/MS scans give much more structural information than those acquired in negative mode, all subsequent scans were done in the positive ion mode.

As mentioned earlier, some acylglycines do not fragment to yield a product ion at  $m/z$  76. However, all acylglycines commonly lose the neutral glycine fragment (a neutral loss of 75 Da) and the remaining acyl fragment retains the positive charge. An IDA experiment in the positive ionization mode, similar to that described above, was performed except the survey scan used was a neutral loss of 75 Da scan. Figure 2-7 (C) shows a total ion chromatogram of an IDA experiment using constant neutral loss of 75 Da as the survey scan. Acylglycines such as hippuric acid, 3-methylcrotonylglycine, tiglylglycine, phenylglycine, hydroxyhippuric acid, and methylhippuric acid were not detected in the precursor method and exhibited major fragment ions at  $m/z$  105, 83, 83, 77, 121 and 119, respectively, corresponding to the neutral loss of 75 Da. Comparison of the TICs

shown in Figure 2-7 (B) and (C) shows that most of the acylglycines observed in the precursor method were also found using the neutral loss method. However, the neutral loss method was not as selective, because other classes of compounds that gave the neutral loss of 75 Da (the identities of these compounds were unknown) were also detected and many acylglycines were detected at lower intensities.

When targeting certain compound families in biological matrices, high selectivity and sensitivity become essential. Precursor and constant neutral loss scans are highly selective, but are only moderately sensitive. A much better alternative is to use MRM as the survey scan to trigger the enhanced product ion data collection. MRM is much more selective and sensitive than either precursor scanning or constant neutral loss scanning. Due to the differences in the fragmentation of certain acylglycines, specific MRM transitions were chosen for each acylglycine using fragmentation information gleaned from the precursor scanning, constant neutral loss scanning methods and the breakdown curves. The precursor scanning and constant neutral loss scanning survey scans provided information about the masses of all possible acylglycines that can be observed and the breakdown curves were used to provide the best possible transitions. These optimized MRM transitions as well as dwell times and collision energy are shown in Table 2-2. The survey scan contained 62 MRM transitions and each transition was performed with a dwell time of 15 ms and a CE of 20 eV. The scan time was 1.2 s for all transitions. The IDA intensity threshold was set to 100 counts per second (cps) and the dependent scans were performed in enhanced product ion mode with a CE of 20 eV. Four enhanced product ion scans were performed on the four most intense peaks before switching back to MRM scan mode.

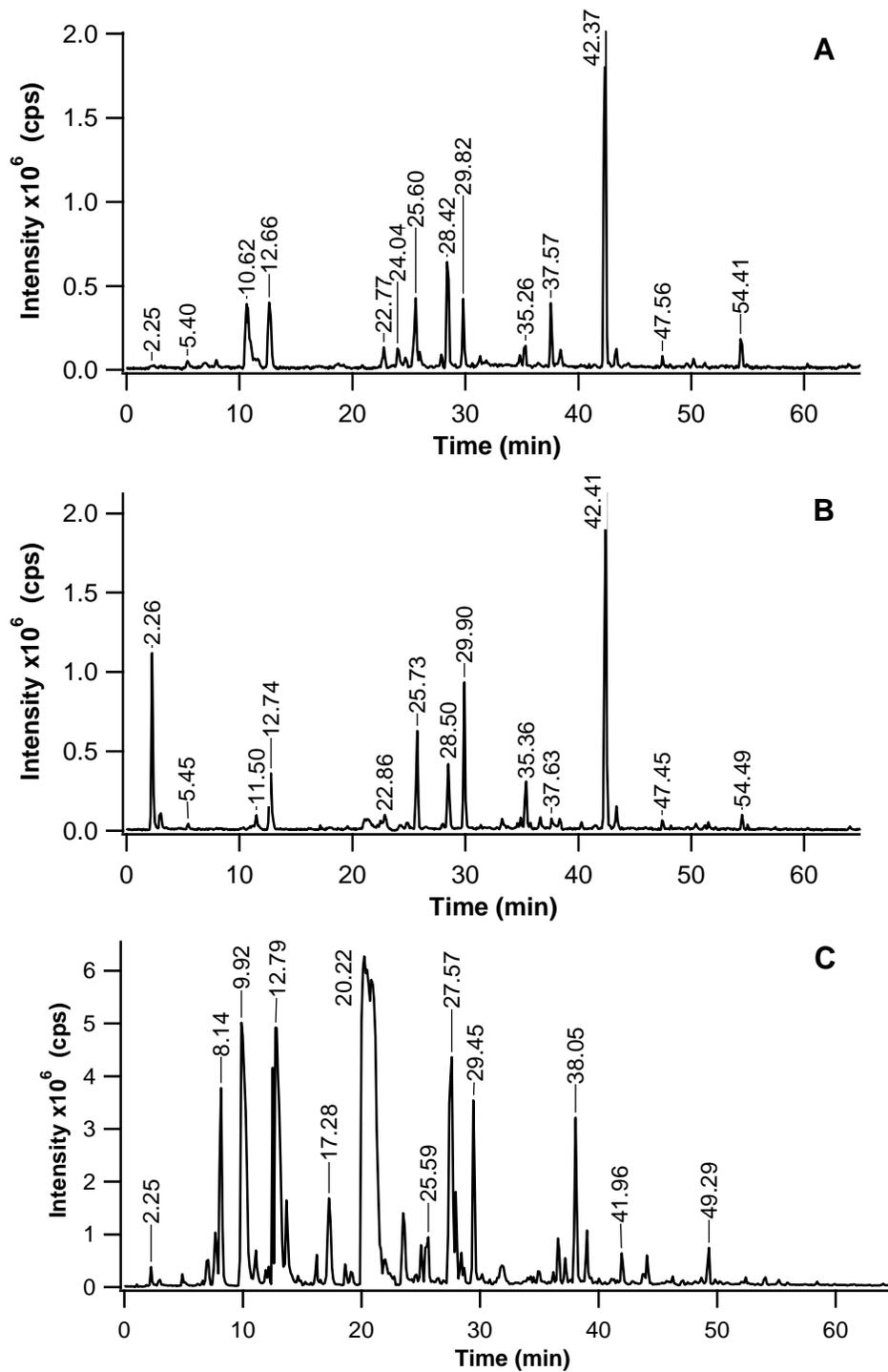


Figure 2-7 Ion chromatograms from the UPLC-MS/MS analysis of urine from a healthy volunteer: (A) TIC of precursor ion scanning of  $m/z$  74 (negative mode); (B) TIC of precursor ion scanning of  $m/z$  76 (positive mode); and (C) TIC of neutral loss scanning of 75 Da.

Table 2-2 MRM scanning conditions (positive ion mode)

Q1 ( <i>m/z</i> )	Q3 ( <i>m/z</i> )	Dwell time (ms)	CE (eV)	Q1 ( <i>m/z</i> )	Q3 ( <i>m/z</i> )	Dwell time (ms)	CE (eV)
104	58	15	20	204	76	15	20
118	76	15	20	204	140	15	20
146	76	15	20	206	76	15	20
152	77	15	20	206	103	15	20
158	83	15	20	208	105	15	20
160	76	15	20	210	107	15	20
162	76	15	20	212	137	15	20
170	95	15	20	214	76	15	20
172	75	15	20	214	111	15	20
172	69	15	20	216	76	15	20
174	76	15	20	216	123	15	20
176	76	15	20	218	76	15	20
176	64	15	20	218	125	15	20
180	105	15	20	218	154	15	20
181	106	15	20	222	76	15	20
184	109	15	20	222	119	15	20
186	76	15	20	224	76	15	20
186	83	15	20	224	121	15	20
188	76	15	20	226	76	15	20
188	95	15	20	232	168	15	20
190	76	15	20	242	149	15	20
190	126	15	20	242	167	15	20
194	91	15	20	244	169	15	20
194	119	15	20	246	76	15	20
196	121	15	20	246	182	15	20
198	76	15	20	250	147	15	20
198	95	15	20	256	76	15	20
200	76	15	20	256	163	15	20
200	97	15	20	258	194	15	20
202	76	15	20	260	196	15	20
202	109	15	20	276	201	15	20

### 2.3.4 Human Urine Analysis

After the method was developed, it was applied to the analysis of human urine samples using the 62 MRM transitions to detect acylglycines. Figure 2-8 shows the ion chromatograms obtained from UPLC MS/MS of urine samples of six healthy volunteers.

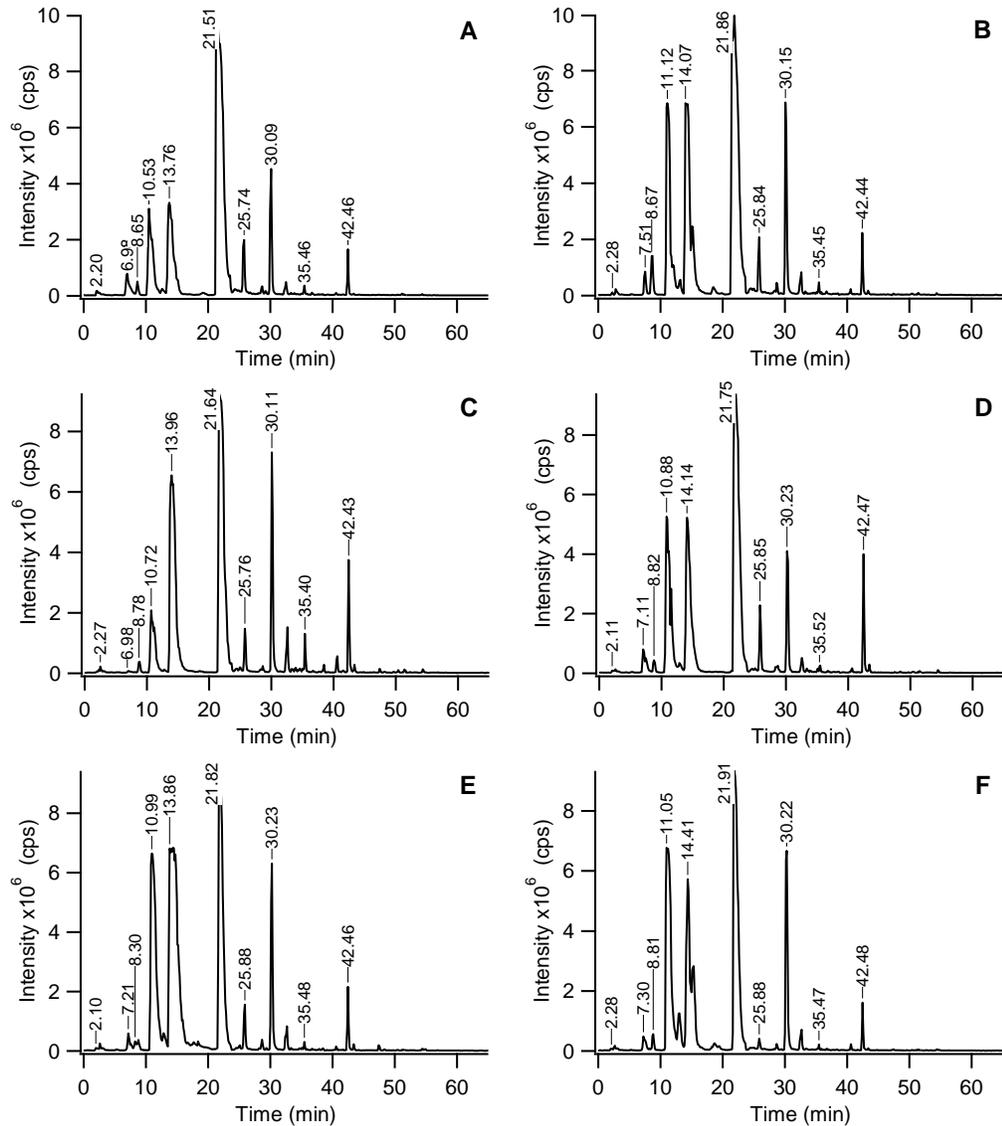


Figure 2-8 Ion chromatograms from UPLC-MS/MS analysis of six healthy human urine samples. Urine samples were collected as first morning void.

Relative differences in intensities of acylglycines can be observed. Most acylglycines are common to all six individuals with the exception of dimethylglycine and 4-hydroxyphenylacetyl-glycine, which are only detected in 4 and 5 individuals, respectively. Acylglycine identification was confirmed using standards and putative identification of other acylglycines was done using retention time and class-specific fragmentation patterns. In order to identify an unknown acylglycine, at least four of seven neutral losses must be observed. Water loss is very common and not considered to be structurally valuable. Due to the lack of standards, positive identification of many unknown acylglycines cannot be made. The use of human liver microsome (HLM) metabolite production and chemical derivatization was employed to provide putative identification for some of the unknowns. In the HLM method, individual acylglycine standards were incubated with human liver microsomes to generate the metabolites of acylglycines potentially found in human urine. The hydroxylated and carbonyl-substituted metabolites were the most common metabolites observed in the microsome incubations. A comparison of the retention times and fragmentation patterns of the metabolites was used to aid in the identification of some of the unknown acylglycines. An example of such a comparison is shown in Figure 2-9. The ion chromatograms from microsome incubation and urine are shown in Figure 2-9 (A) and (C), respectively. Figure 2-9 (B) shows the MS/MS spectrum from the chromatographic peak labeled with a diamond at the retention time of 21.11 minutes (Figure 2-9 (A)). Figure 2-9 (D) shows the MS/MS spectrum from the peak with a diamond at the retention time of 21.22 minutes (Figure 2-9 (C)). The inset in Figure 2-9 (D) shows the fragmentation of hydroxyphenylpropionylglycine. Comparison of the MS/MS spectra shows losses of 18, 46, 75, 103 and 117 in the spectrum of the urine sample. Losses of 18 and 46 (not diagnostic) and the precursor ion are not observed in the MS/MS spectrum of the microsome extract (possibly too high CE), but losses of 75, 103 and 117 are common to both, with the major fragment ion being the loss of 117. Even though not all the fragment ions are observed, the diagnostic losses are common. Comparison of retention times and MS/MS

spectra allows for the identification of hydroxyphenylpropionylglycine. It is important to note that the isomers generated by the microsomes are not necessarily the same isomers excreted in urine but the isomeric information is nonetheless valuable.

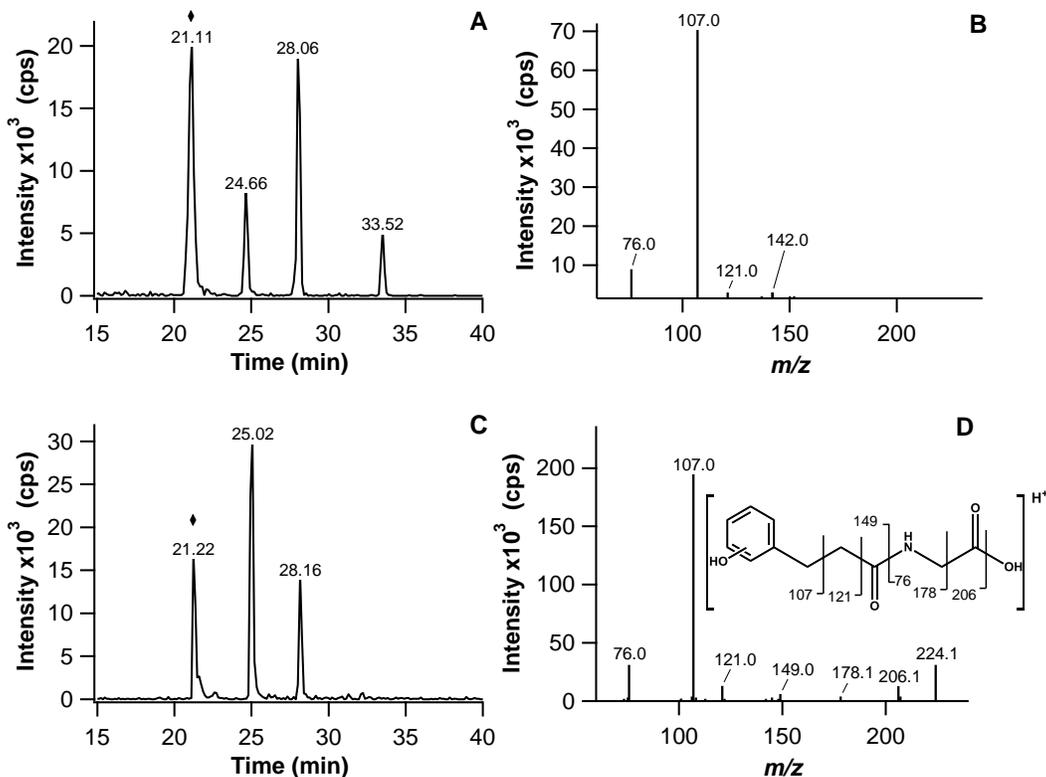


Figure 2-9 Comparison of XIC of  $m/z$  224 (hydroxyphenylpropionylglycine) in (A) microsome extract and (C) urine, and corresponding product ion spectra, (B) microsomes and (D) urine. The inset in (D) shows the fragmentation pattern of hydroxyphenylpropionylglycine.

Chemical derivatization is also useful for acylglycine identification. It is important to note that compounds other than acylglycines can also produce a fragment ion of  $m/z$  76 in the product ion scan. However, comparison between underivatized urine and derivatized urine can assist in identifying acylglycines and distinguish them from other compounds. A shift of 14 Da in the precursor

mass (28 Da in the case of di-carboxylic acids), an appropriate shift in retention time and corresponding losses in the product ion scans should be observed for acylglycines. For example, C<sub>6</sub>:1 acylglycine, previously unreported in the urine of healthy individuals, was confirmed as an acylglycine by methylation and the results are shown in Figure 2-10.

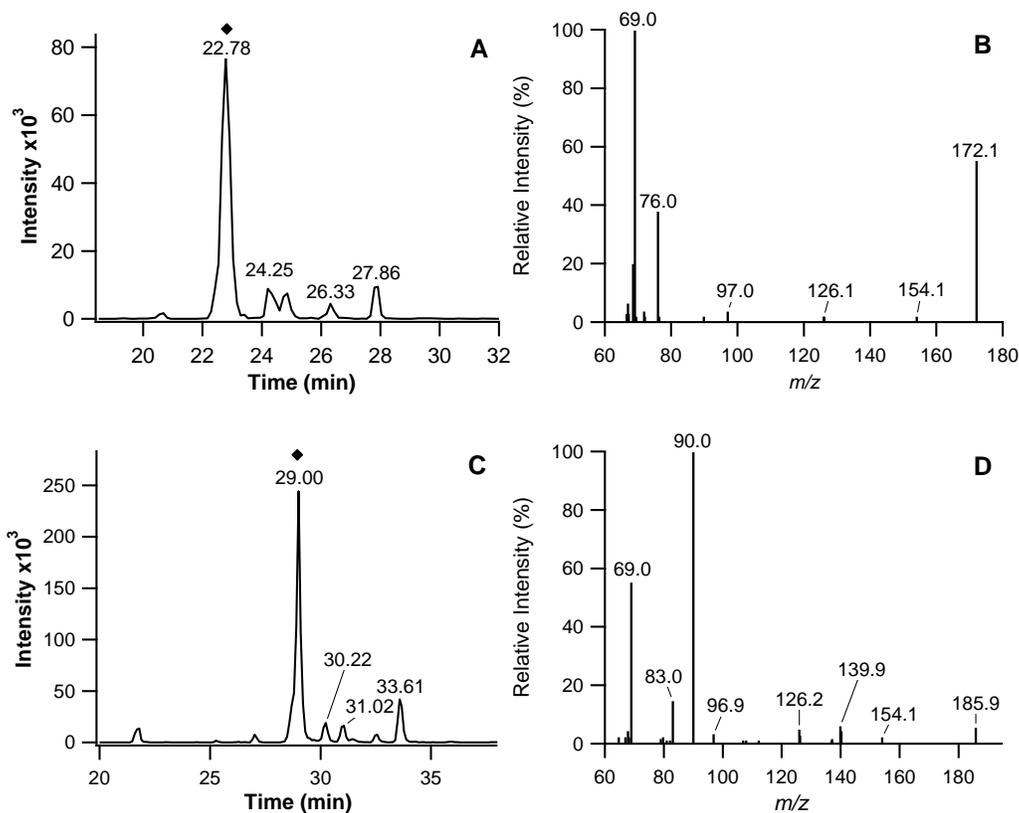


Figure 2-10 Putative identification of C<sub>6</sub>:1-glycine (position of double bond unknown), (A) XIC of C<sub>6</sub>:1-glycine isomer, *m/z* 172, (B) MS/MS spectrum of peak at 22.78 minutes, (C) XIC of C<sub>6</sub>:1-glycine methyl ester, *m/z* 186, (D) MS/MS spectrum of peak at 29.00 minutes.

A total of 65 acylglycines was detected and are shown in Table 2-3 in three groups, positively identified, putatively identified and unknown structures. Only 18 of them have been previously reported in the literature. All of the

acylglycines used in screening for fatty acid disorders were detected with the exception of propionylglycine. Propionylglycine was not observed in any of the individuals but was detected in low levels in a spiked urine sample (the concentration of the spiked propionylglycine was 10  $\mu\text{M}$ ); this compound appears to be present in very low concentrations in urine of healthy individuals. The standard deviations of the chromatographic retention times of acylglycines detected from different individuals were calculated and also presented in Table 2-3. This information is useful for other users who use similar separation conditions for analyzing these metabolites. Examples of the fragmentation and spectral interpretation for these compounds are provided in Figures A2-1 to Figures A2-4 in the Appendix. A full list of the spectral library of acylglycines is presented in the Acylglycines MS/MS library in the Electronic Appendix. Please contact Professor Liang Li for a copy of the Electronic Appendix. The MS/MS spectra of these compounds will be deposited to a free, public-accessible database ([www.hmdb.ca](http://www.hmdb.ca)).<sup>7</sup> This database has over 8000 entries of mainly endogenous human metabolites including MS/MS spectra of about 900 metabolites.

Table 2-3 A list of confirmed and putatively identified acylglycines present in urine of six healthy volunteers\*

<b>Confirmed Acylglycines</b>			
<b>Acylglycine</b>	<b><i>m/z</i></b>	<b>RT (min)</b>	<b>References</b>
Dimethylglycine <sup>‡</sup>	104	2.12 ± 0.0	25, 26
Phenylglycine	152	2.26 ± 0.06	
Acetylglycine	118	2.37 ± 0.06	19, 20
Isobutyrylglycine	146	5.44 ± 0.13	13, 27
4-Hydroxyphenylacetylglycine <sup>†</sup>	210	10.22 ± 0.18	
2-Methylbutyrylglycine	160	11.12 ± 0.12	13, 28, 29
Tiglylglycine	158	11.89 ± 0.19	14, 30-32
Isovalerylglycine	160	12.86 ± 0.16	13, 14, 17, 33, 34
3-Methylcrotonylglycine	158	12.97 ± 0.19	14, 30, 35
Valerylglycine	160	16.38 ± 0.18	
Hippuric acid	180	22.21 ± 0.04	15, 30, 36-38
Suberylglycine	232	25.43 ± 0.06	13, 17
Phenylacetylglycine (PAG)	194	25.88 ± 0.06	20
Hexanoylglycine	174	30.07 ± 0.04	13, 14, 17, 19
Phenylpropionylglycine (PPG)	208	33.88 ± 0.04	13
Octanoylglycine	202	51.16 ± 0.03	
<b>Putatively Identified Acylglycines</b>			
<b>Acylglycine</b>	<b><i>m/z</i></b>	<b>RT (min)</b>	<b>References</b>
Malonylglycine/Hydroxybutyrylglycine (or isomers) <sup>†</sup>	162	2.59 ± 0.05	
Succinylglycine/Hydroxyvalerylglycine	176	4.42 ± 0.14	

(or isomers)			
Dihydroxyhippuric acid (hydroxyl positions unknown)	212	5.27 ± 0.06	
Dihydroxyhippuric acid (hydroxyl positions unknown)	212	6.98 ± 0.22	
2-Furoylglycine	170	7.20 ± 0.19	30, 38, 39
Dihydroxyhippuric acid (hydroxyl positions unknown)	212	8.24 ± 0.27	
Hydroxyhippuric acid (hydroxyl position unknown)	196	8.77 ± 0.06	22, 30, 32, 40
Dihydroxyhippuric acid (hydroxyl positions unknown)	212	10.85 ± 0.23	
Hydroxyhippuric acid (hydroxyl position unknown)	196	10.91 ± 0.26	22, 30, 32, 40
Hydroxyhexanoylglycine (hydroxyl position unknown)	190	11.61 ± 0.16	
2-Pentenoylglycine	158	13.85 ± 0.17	
Hydroxyhippuric acid (hydroxyl position unknown)	196	14.19 ± 0.27	22, 30, 32, 40
Nicotinuric acid	181	15.03 ± 0.21	41, 42
Hydroxyphenylacetyl <sup>†</sup> glycine (hydroxyl position unknown)	210	15.74 ± 0.26	
Hydroxyphenylpropionylglycine (hydroxyl position unknown)	224	21.37 ± 0.16	
Hydroxyphenylpropionylglycine (hydroxyl position unknown)	224	22.68 ± 0.09	
Hexenoylglycine (double bond position unknown)	172	23.06 ± 0.07	

Hexenoylglycine (double bond position unknown)	172	24.53 ± 0.05	
Hydroxyphenylpropionylglycine (hydroxyl position unknown)	224	28.16 ± 0.04	
Methylvalerylglycine (methyl position unknown)	174	28.65 ± 0.03	
Heptenoylglycine (double bond position unknown)	186	28.71 ± 0.03	
Hydroxyoctenoylglycine (hydroxyl position unknown)	216	28.93 ± 0.03	
Methylhippuric acid (methyl position unknown)	194	32.58 ± 0.05	30, 43, 44
Hydroxyoctanoylglycine (hydroxyl position unknown)	218	33.31 ± 0.04	
Octadienoylglycine (double bond positions unknown)	198	35.07 ± 0.03	
Hydroxyoctanoylglycine (hydroxyl position unknown)	218	35.46 ± 0.03	
Octadienoylglycine (double bond positions unknown)	198	36.75 ± 0.03	
Sebacylglycine	260	38.12 ± 0.06	
Methylhexanoylglycine (methyl position unknown)	188	38.47 ± 0.02	
Heptanoylglycine	188	41.07 ± 0.04	
Octadienoylglycine (double bond positions unknown)	198	41.40 ± 0.01	
Phenylbutyrylglycine	222	41.61 ± 0.05	

Cis-3,4-Methylene-heptanoylglycine	200	42.46 ± 0.02	
Octenoylglycine (double bond position unknown)	200	43.38 ± 0.02	
Nonenoylglycine (double bond position unknown)	214	47.44 ± 0.01	
Nonanoylglycine	216	54.95 ± 0.02	
<b>Unknown Acylglycines</b>			
<b>Acylglycine</b>	<b><i>m/z</i></b>	<b>RT (min)</b>	<b>References</b>
Unknown #1	210	30.41 ± 0.03	
Unknown #2	242	32.90 ± 0.03	
Unknown #3	188	35.22 ± 0.03	
Unknown #4 <sup>‡</sup>	258	35.29 ± 0.06	
Unknown #5	200	35.45 ± 0.02	
Unknown #6	244	36.28 ± 0.03	
Unknown #7	256	38.06 ± 0.03	
Unknown #8	242	40.27 ± 0.04	
Unknown #9	222	50.27 ± 0.04	
Unknown #10	226	52.60 ± 0.03	
Unknown #11	224	54.39 ± 0.03	
Unknown #12	250	56.84 ± 0.03	
Unknown #13	256	66.54 ± 0.02	

\*Compounds marked with <sup>†</sup> are only found in five individuals and those marked with <sup>‡</sup> are only found in four individuals. References quoted are for acylglycines detected in healthy individuals by other methods. The MS/MS spectra of all the listed compounds can be found in the Electronic Appendix.

While Figure 2-8 illustrates that many acylglycines can be detected from urine samples of different individuals, a similar number of acylglycines can also be detected from urine samples of an individual collected at different time points. This is illustrated in a preliminary work where the effect of the time of sampling on the acylglycine profile of urine was examined. For urine sample collection, the most common type of collection is the first morning voiding. Often a sample is collected during fasting because analytes of interest may be excreted in higher concentrations. However, other conditions that occur in fasting can lead to an abnormal excretion of other metabolites. In this study, urine was collected from a healthy individual at the first morning voiding, for five consecutive days (to assess the effect of diet), at four hours after a heavy meal and also was collected over 12 hrs. Figure 2-11 shows the ion chromatograms of urine collected at the first morning voiding from an individual from five consecutive days. An overlay of total ion chromatograms of multiple reaction monitoring scans using 62 diagnostic transitions are shown.

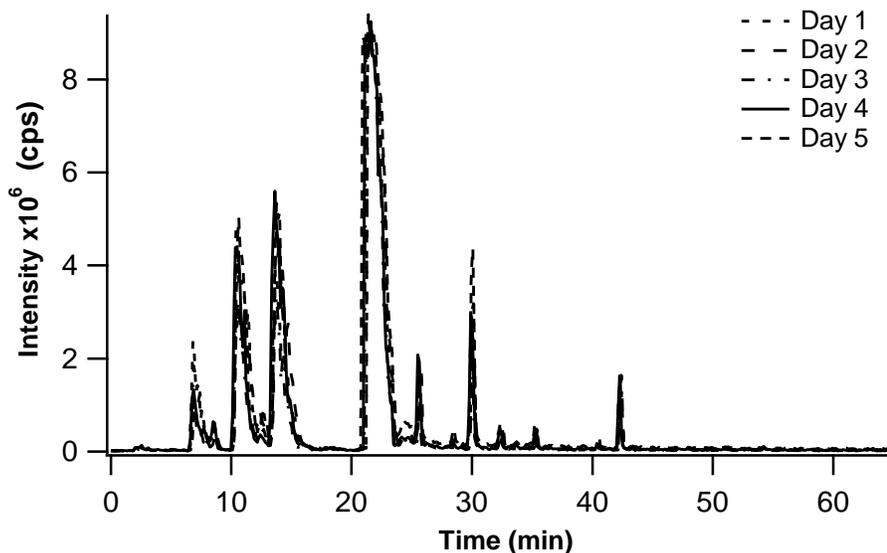


Figure 2-11 An overlay of total ion chromatograms of urine samples collected as first morning void from a healthy individual for five consecutive days.

The ion chromatograms of other urine samples collected, four hours after a meal and collected over 12 hours, are shown in Figure 2-12. Results of this comparison study revealed very few differences in the acylglycine profile. Relative intensities differed between time collections but the same acylglycines were detected in all the time collections.

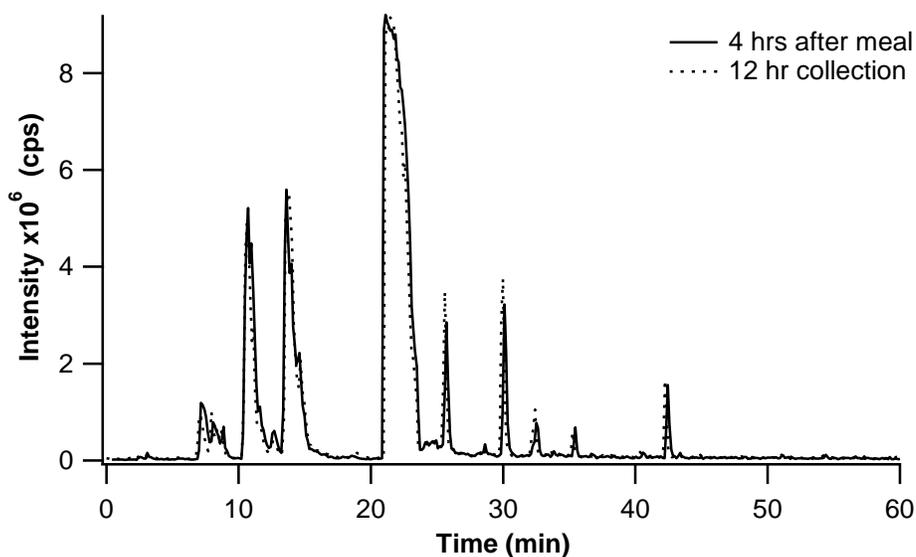


Figure 2-12 An overlay of total ion chromatograms of urine samples collected from a healthy individual 4 hours after a meal and collected over a period of 12 hours.

The above examples demonstrate that a large number of acylglycines can be consistently detected from urine samples collected from different individuals or from the same individual at different times. Thus, the developed method should be suitable for profiling many acylglycines for potential discovery of new biomarkers for IEMs and other diseases by using a large pool of samples with clinical information on the individuals.

## 2.4 Conclusions

In this study, a method for the comprehensive analysis of acylglycines in the urine of healthy subjects was developed. Putative identification of acylglycines was achieved using retention time information and characteristic fragmentation patterns. Derivatization or esterification was not necessary to enhance detection, when using selective extraction methods like SPE and sensitive detection instruments like the hybrid QTRAP<sup>®</sup> mass spectrometer. Greater selectivity and sensitivity were achieved using MRM scans and simultaneous product ion scans provide structural elucidation. The use of optimal acquisition modes is very important. Constant neutral loss and precursor methods are valuable for detection of wide range of acylglycines and for the discovery of expected as well as unexpected, or previously undetected, acylglycines. Breakdown graphs provided optimal transitions and conditions, as well as important trends to note for detection of classes of acylglycines. With this strategy, detection of expected acylglycines was enhanced and novel acylglycines were discovered. A total of 65 acylglycines were detected and of those, only 15 are currently used for diagnosis of metabolic diseases. Additional studies need to be done using clinical samples of subjects with known IEMs or other diseases to determine the effectiveness of the method in the diagnoses of these disorders. The data set generated in this study will be significant in future research to further understand, discover and improve the diagnosis of new inborn errors of metabolism and other metabolic diseases related to the excretion of acylglycines.

## 2.5 Literature Cited

- (1) Scalbert, A.; Brennan, L.; Fiehn, O.; Hankemeier, T.; Kristal, B. S.; van Ommen, B.; Pujos-Guillot, E.; Verheij, E.; Wishart, D.; Wopereis, S. *Metabolomics* **2009**, *5*, 435-458.
- (2) Bowen, B. P.; Northen, T. R. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1471-1476.

- (3) Bullinger, D.; Fux, R.; Nicholson, G.; Plontke, S.; Belka, C.; Laufer, S.; Gleiter, C. H.; Kammerer, B. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1500-1513.
- (4) Zhou, M.; McDonald, J. F.; Fernández, F. M. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 68-75.
- (5) Lin, L.-C.; Wu, H.-Y.; Tseng, V. S.-M.; Chen, L.-C.; Chang, Y.-C.; Liao, P.-C. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 232-241.
- (6) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2010**, *5*, 1005-1018.
- (7) Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S.; Mandal, R.; Sinelnikov, I.; Xia, J. G.; Jia, L.; Cruz, J. A.; Lim, E.; Sobsey, C. A.; Shrivastava, S.; Huang, P.; Liu, P.; Fang, L.; Peng, J.; Fradette, R.; Cheng, D.; Tzur, D.; Clements, M.; Lewis, A.; De Souza, A.; Zuniga, A.; Dawe, M.; Xiong, Y. P.; Clive, D.; Greiner, R.; Nazzyrova, A.; Shaykhtudinov, R.; Li, L.; Vogel, H. J.; Forsythe, I. *Nucleic Acids Res.* **2009**, *37*, D603-D610.
- (8) Sahai, I.; Marsden, D. *Crit. Rev. Clin. Lab. Sci.* **2009**, *46*, 55-82.
- (9) Pasquali, M.; Monsen, G.; Richardson, L.; Alston, M.; Longo, N. *American Journal of Medical Genetics Part C-Seminars in Medical Genetics* **2006**, *142C*, 64-76.
- (10) Sim, K. G.; Hammond, J.; Wilcken, B. *Clin. Chim. Acta* **2002**, *323*, 37-58.
- (11) Hale, D. E.; Bennett, M. J. *J. Pediatr.* **1992**, *121*, 1-11.
- (12) Kouremenos, K. A.; Pitt, J.; Marriott, P. J. *J. Chromatogr., A* **2010**, *1217*, 104-111.
- (13) Costa, C. G.; Guerand, W. S.; Struys, E. A.; Holwerda, U.; ten Brink, H. J.; de Almeida, I. T.; Duran, M.; Jakobs, C. *J. Pharm. Biomed. Anal.* **2000**, *21*, 1215-1224.
- (14) Hagen, T.; Korson, M. S.; Sakamoto, M.; Evans, J. E. *Clin. Chim. Acta* **1999**, *283*, 77-88.
- (15) Suh, J. W.; Lee, S. H.; Chung, B. C. *Clin. Chem.* **1997**, *43*, 2256-2261.
- (16) Carter, S. M. B.; Midgley, J. M.; Watson, D. G.; Logan, R. W. *J. Pharm. Biomed. Anal.* **1991**, *9*, 969-975.

- (17) Kimura, M.; Yamaguchi, S. *J. Chromatogr., B.* **1999**, *731*, 105-110.
- (18) Downing, M.; Allen, J. C.; Bonham, J. R.; Edwards, R. G.; Manning, N. J.; Olpin, S. E.; Pollitt, R. J. *J. Inherited Metab. Dis.* **1999**, *22*, 289-292.
- (19) Millington, D. S.; Kodo, N.; Terada, N.; Roe, D.; Chace, D. H. *Int. J. Mass Spectrom.* **1991**, *111*, 211-228.
- (20) Bonafe, L.; Troxler, H.; Kuster, T.; Heizmann, C. W.; Chamoles, N. A.; Burlina, A. B.; Blau, N. *Mol. Genet. Metab.* **2000**, *69*, 302-311.
- (21) Shigematsu, Y.; Hata, I.; Tanaka, Y. *Clin. Chim. Acta* **2007**, *386*, 82-86.
- (22) Rashed, M. S. *J. Chromatogr., B: Biomed. Sci. Appl.* **2001**, *758*, 27-48.
- (23) Hopfgartner, G.; Varesio, E.; Tschappat, V.; Grivet, C.; Bourgoigne, E.; Leuthold, L. A. *J. Mass Spectrom.* **2004**, *39*, 845-855.
- (24) Guillarme, D.; Schappler, J.; Rudaz, S.; Veuthey, J. L. *Trac-Trends Anal. Chem.* **2010**, *29*, 15-27.
- (25) Laryea, M. D.; Steinhagen, F.; Pawliczek, S.; Wendel, U. *Clin. Chem.* **1998**, *44*, 1937-1941.
- (26) Moolenaar, S. H.; Poggi-Bach, J.; Engelke, U. F. H.; Corstiaensen, J. M. B.; Heerschap, A.; de Jong, J. G. N.; Binzak, B. A.; Vockley, J.; Wevers, R. A. *Clin. Chem.* **1999**, *45*, 459-464.
- (27) Sass, J. O.; Sander, S.; Zschocke, J. *J. Inherited Metab. Dis.* **2004**, *27*, 741-745.
- (28) Gibson, K. M.; Burlingame, T. G.; Hogema, B.; Jakobs, C.; Schutgens, R. B. H.; Millington, D.; Roe, C. R.; Roe, D. S.; Sweetman, L.; Steiner, R. D.; Linck, L.; Pohowalla, P.; Sacks, M.; Kiss, D.; Rinaldo, P.; Vockley, J. *Pediatric Res.* **2000**, *47*, 830-833.
- (29) Tein, I.; Haslam, R. H. A.; Rhead, W. J.; Bennett, M. J.; Becker, L. E.; Vockley, J. *Neurology* **1999**, *52*, 366-372.
- (30) Liebich, H. M.; Forst, C. *J. Chromatogr., B: Biomed. Appl.* **1990**, *525*, 1-14.
- (31) Bennett, M. J.; Powell, S.; Swartling, D. J.; Gibson, K. M. *Clin. Chem.* **1994**, *40*, 1879-1883.
- (32) Garcia-Villoria, J.; Navarro-Sastre, A.; Fons, C.; Perez-Cerda, C.; Baldellou, A.; Fuentes-Castello, M. A.; Gonzalez, I.; Hernandez-

- Gonzalez, A.; Fernandez, C.; Campistol, J.; Delpiccolo, C.; Cortes, N.; Messeguer, A.; Briones, P.; Ribes, A. *Clin. Biochem.* **2009**, *42*, 27-33.
- (33) Rinaldo, P.; Welch, R. D.; Previs, S. F.; Schmidtsommerfeld, E.; Gargus, J. J.; Oshea, J. J.; Zinn, A. B. *Pediatric Res.* **1991**, *30*, 216-221.
- (34) Fries, M. H.; Rinaldo, P.; SchmidtSommerfeld, E.; Jurecki, E.; Packman, S. *J. Pediatr.* **1996**, *129*, 449-452.
- (35) Eminoglu, F. T.; Ozcelik, A. A.; Okur, I.; Tumer, L.; Biberoglu, G.; Demir, E.; Hasanoglu, A.; Baumgartner, M. R. *J. Child Neurology* **2009**, *24*, 478-481.
- (36) Roowi, S.; Stalmach, A.; Mullen, W.; Lean, M. E. J.; Edwards, C. A.; Crozier, A. *J. Agric. Food Chem.* **2010**, *58*, 1296-1304.
- (37) Ahmadi, F.; Asgharloo, H.; Sadeghi, S.; Gharehbagh-Aghababa, V.; Adibi, H. *J. Chromatogr., B: Anal. Tech. Biomed. Life Sci.* **2009**, *877*, 2945-2951.
- (38) Liebich, H. M.; Gesele, E.; Woll, J. *J. Chromatogr., B* **1998**, *713*, 427-432.
- (39) Pettersen, J. E.; Jellum, E. *Clin. Chim. Acta* **1972**, *41*, 199-207.
- (40) Chen, Y. F.; Sullards, M. C.; Hoang, T. T.; May, S. W.; Orlando, T. M. *Anal. Chem.* **2006**, *78*, 8386-8394.
- (41) Li, A. C.; Chen, Y. L.; Junga, H.; Shou, W. Z.; Jiang, X.; Naidong, W. *Chromatographia* **2003**, *58*, 723-731.
- (42) Stratford, M. R. L.; Dennis, M. F. *J. Chromatogr., B: Biomed. Appl.* **1992**, *582*, 145-151.
- (43) Pacenti, M.; Dugheri, S.; Villanelli, F.; Bartolucci, G.; Calamai, L.; Boccalon, P.; Arcangeli, G.; Vecchione, F.; Alessi, P.; Kikic, I.; Cupelli, V. *Biomed. Chromatogr.* **2008**, *22*, 1155-1163.
- (44) Ohashi, Y.; Mamiya, T.; Mitani, K.; Wang, B. L.; Takigawa, T.; Kira, S.; Kataoka, H. *Anal. Chim. Acta* **2006**, *566*, 167-171.

# Chapter 3. Quantification of Acylglycines in Human Urine using Derivatization and Liquid Chromatography Tandem Mass Spectrometry\*

## 3.1 Introduction

The analysis of acylglycines in biofluids is of major interest in diagnosing inborn errors of metabolism (IEMs) because of their prominent role in metabolic pathways.<sup>1-3</sup> IEMs change the way the body processes nutrients and uses them for energy production. These disorders are characterized by a defect in a single enzyme, resulting in reduced or nonexistent catalytic activity, or by the impaired activity of transporters or cofactors along the metabolic pathway. This can result in an accumulation of metabolites that can become substrates for enzymes in alternative metabolic pathways. In disorders of amino acid and fatty acid metabolisms, the intermediary metabolism of amino acids or fatty acids is blocked and the intermediary product formed is in excess and must be conjugated to carnitine or glycine to facilitate excretion or a balance of coenzyme A. Glycine conjugation serves an important detoxification role and occurs mainly in the liver. Therefore acylglycines in urine is an indicator of the accumulation of the corresponding acyl CoA esters and can be important in the diagnoses of IEMs like fatty acid disorders and amino acid disorders. In order to diagnose as many IEMs as possible, a wide range of acylglycines should be quantified in urine.

Traditionally acylglycines in urine have been analyzed by GC/MS<sup>4, 5</sup> and are usually derivatized for analysis. GC/MS provides the combination of a high resolution separation in gas chromatography and mass spectrometric detection, which reduces the number of false positives. Metabolic profiling can also be done

---

\* A version of this chapter has been submitted as Stanislaus, A.E., Li, L., **2012**, Development of an Isotope Labeling Ultra-High Performance Liquid Chromatography Mass Spectrometric Method for Quantification of Acylglycines in Human Urine to *Analytica Chimica Acta*

by GC/MS analysis. Tandem mass spectrometry<sup>6</sup> has been used to analyze acylglycines, and it is a fast and powerful diagnostic tool, but it lacks the capability to distinguish between isomeric and isobaric species and is limited in diagnosing certain IEMs. In a recent study,<sup>7</sup> acylglycines were analyzed using UPLC-MS/MS and 44 acylglycines were detected that were previously undetected in normal human urine. The UPLC has the advantage of high resolution separation like that of gas chromatography and high sensitivity. Although it would seem that their extraction and analysis from urine would be simple, differences in their concentration and complexity resulted in difficulties resolving and detecting all of the analytes. In that study, trace quantities of acylglycines were detected by extracting and concentrating the analytes.

To my knowledge, only three LC-MS methods have been reported to date for quantification of several acylglycines.<sup>8-10</sup> Quantification in LC-MS is related to the ionization efficiencies of the analytes. Matrix effect and suppression can reduce sensitivity and can cause complication in quantification. The best strategy is to use an internal standard (IS), which differs from the analyte of interest by its isotopic form. This IS also helps to overcome matrix effects by compensating for ion suppression or enhancement. Unfortunately, it is not always possible or practical to synthesize an IS for every metabolite when performing a more comprehensive profiling study.

An alternate strategy is to use chemical labeling to attach a stable isotope-coded derivatizing agent to all analytes possessing a common reactive moiety. In acylglycines, the target functional group is the terminal carboxyl group. Several derivatization procedures have been developed for the positive-ion detection of acids. One of the common derivatization reagents used to label carboxyl-containing compounds is phenacyl bromide.<sup>11-13</sup> Although the main use of the tag was to allow carboxylic acids to be detectable by UV/Vis, the benzene group is quite amenable to ESI. The derivatization agent used in this study is a new isotope label, p-dimethylaminophenacyl (DmPA) bromide, used previously in our laboratory to label carboxylic acids.<sup>14</sup> This label is a derivative of phenacyl

bromide, in which a dimethyl amino group is added in the para-position. This group is an easily ionizable group to which a proton could be easily added. Also an isotope of carbon can be coded on the dimethyl groups (see Figure 3-1). Absolute and relative quantification of all samples can be easily achieved with this method. There are several main advantages of using this method: first, the reaction is a one-step reaction under relatively mild conditions; second, ionization efficiencies are increased through the introduction of a positive charge on the amino moiety; third, a stable isotope labeled standard can be generated for each standard obtained; fourth, using an isotope of carbon instead of deuterium means that these analytes do not suffer from isotope effects in chromatography; and lastly, chromatographic retention is improved.

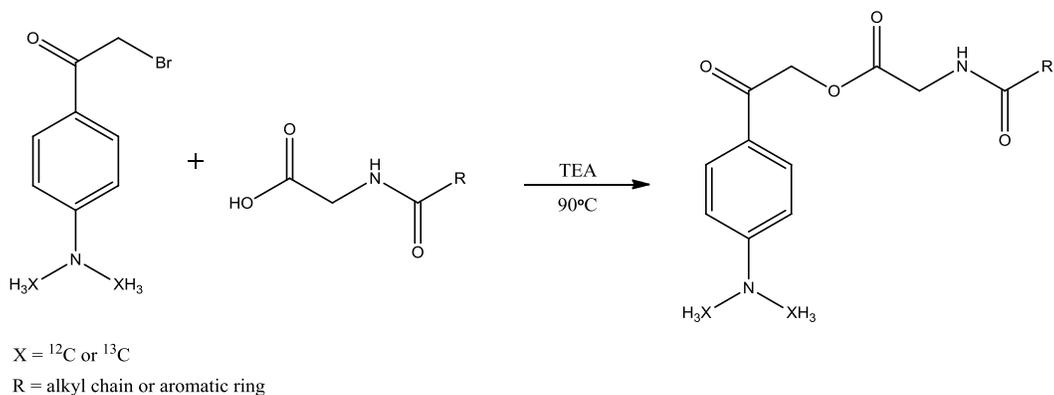


Figure 3-1 Derivatization reaction of DmPA and acylglycines

One of the major challenges faced in quantification experiments is the complication of quantifying naturally occurring (endogenous) analytes in biological matrices, due to background interferences from the analytes being present in the blank matrix. Method validation is primarily done in the authentic blank matrix, which is often analyte-free and preparation of calibrators and reference samples in the matrix are quite straightforward. In the presence of endogenous analytes; however, calibrators and reference samples have to be prepared in a different way, using alternate analytes and matrices, and validation

methods have to be adjusted to fit the approach used. Another major challenge is the lack of regulatory guidelines on accurate method validation. Many researchers apply the same validation procedures as those used for exogenous compounds and these are not accurate or applicable in many cases. An important part of method development and validation is to construct calibration curves in the matrix of interest for quantification and assessment of matrix effects. This is usually problematic because the matrix contains unknown quantities of the analytes of interest. The use of standard addition method is a suitable and accurate approach for endogenous analyte quantification, which eliminates any matrix effects issues. This method consists of the addition of a series of standard solutions in increasing concentration to each individual sample and sample concentration is determined by the x-intercept of the calibration curve. The problem with this approach however, is that it can be impractical when there are many samples or limited sample volumes and the method is only accurate when the sample concentration is greater than matrix baseline concentration. One strategy is to use a “surrogate analyte” approach,<sup>15-18</sup> where quantification of the endogenous analyte is done using a calibration curve constructed using an alternate analyte. For this approach, a suitable surrogate analyte would be an isotopic analogue of the analyte. The IS could then be a chemical analogue or more ideally another isotopic analogue of the analyte. One factor to be considered with this method is that the purchase of several isotopic analogues of several analytes could be very expensive and synthesis could be very time consuming. In this study, another strategy called the “surrogate matrix” approach was chosen to address the issue of endogenous interference by using a calibration curve constructed in an alternate matrix. Pure aqueous or organic solutions,<sup>19, 20</sup> synthetic matrices,<sup>21-23</sup> or matrices from alternate species,<sup>24-26</sup> have been previously used but these may not account for recovery in sample preparation and matrix effects in LC-MS analyses.

In this work, I attempt to validate the procedure for the quantification of derivatized acylglycines in urine by using the underivatized urine as the “surrogate matrix.” The analytical approach used in this study is described as

follows: reference standards of acylglycines were obtained and derivatized with both  $^{12}\text{C}$ -DmPA reagent (referred to as analytes) and  $^{13}\text{C}$ -DmPA reagent (referred to as IS). Using the analytes and IS, calibration curves were constructed and evaluated in three different matrices: solvent mix (water/ acetonitrile, 50/50, v/v, 0.1% formic acid), derivatized urine and underivatized urine. Regression lines of each of the three curves were compared statistically to establish similarities in the matrices and determine the choice of surrogate matrix, which is underivatized urine in this study. The standard addition method was also used to verify the validity of this approach. Lastly, method validation and quantification of the endogenous analytes were performed using the regression equations of the calibration curve in surrogate matrix.

This work summarizes the development and validation of a novel and sensitive LC-MS/MS method for the accurate determination of acylglycines in urine of healthy subjects. To the best of my knowledge, this is the first LC-MS assay that uses the DmPA tag for enhancing sensitivity of the acylglycines and the “surrogate matrix” strategy for accurate quantification of these analytes.

## **3.2 Experimental**

### **3.2.1 Materials and Reagents**

Optima grade acetonitrile, methanol and water were purchased from Fisher Scientific (Ottawa, ON). Triethylamine and formic acid were obtained from Sigma-Aldrich (Oakville, ON). Acetylglycine, propionylglycine, isobutyrylglycine, butyrylglycine, 2-methylbutyrylglycine, isovalerylglycine, valerylglycine, hexanoylglycine, heptanoylglycine, tiglylglycine, glutarylglucine and suberylglucine were purchased from Dr. Herman J. ten Brink (Amsterdam, Netherlands). 3-Methylcrotonylglycine, 2-furoylglycine, phenylacetylglucine, phenylpropionylglucine, and 4-hydroxyphenylacetylglucine standards were obtained from HMBD. Octanoylglycine was purchased from Sigma-Aldrich. 2-

Bromo-1-(4-dimethylaminophenyl)ethanone was purchased from Combi-Blocks Inc. (San Diego, CA) and also synthesized in our laboratory, according to the method described below.

### 3.2.2 Synthesis of $^{13}\text{C}_2$ - p-dimethylaminophenacyl bromide

The synthesis of  $^{13}\text{C}$ -DmPA derivatizing reagent was based on a two-step procedure. In the first step, in a 25-mL round-bottom flask, 1.1 g of p-aminoacetophenone was added to 2.55 g of  $\text{NaHCO}_3$  in 4.2 mL of  $\text{H}_2\text{O}$ . The flask was placed in a 10 °C water bath for 5 minutes. About 1.5 mL of  $^{13}\text{C}$ -dimethyl sulphate was added drop-wise over 90 minutes to the solution in an ice-water bath at 10 °C. The resulting solution was warmed to 50 °C for 10 minutes. About 1.6 mL of saturated KOH solution was added to hydrolyze un-reacted dimethyl sulfate. The resulting solution was subjected to liquid/liquid extraction by water, then chloroform. The precipitate of tertiary amine formed by the above treatment was washed a few times with water, then with heptane, and then dissolved in chloroform and dried with magnesium sulfate. Subsequently, chloroform was removed and the tertiary amine was dissolved in hot heptane under reflux and filtered through a hot Schott funnel. After cooling, the crystallized amine was filtered, washed a few times with heptane. The residue was purified by flash chromatography (silica gel, 40 × 3 cm, 20 mL ethylacetate), and further purified by a semi-preparative Grace Apollo silica normal-phase HPLC column (10 × 150 mm, 5 µm particle size).

In the second step, 0.65 g of  $^{13}\text{C}$ -N,N-Dimethylamino-p-acetophenone synthesized from the first step was dissolved in 4 mL of concentrated  $\text{H}_2\text{SO}_4$  and then cooled to 0 °C in a water-ice bath, which resulted in a positively charged quaternary amine. To the solution, 0.2 mL of bromine was added drop-wise at 0 °C and was then gradually warmed to room temperature and stirred for 6 hours. The reaction mixture was poured into ice/water. The yellow precipitate was extracted by liquid/liquid extraction with chloroform, and then dried, then re-

dissolved in 5 mL of tetrahydrofuran, and cooled to 0 °C in ice-water bath. About 0.45 mL of diethylphosphite and 0.47 mL of triethylamine in 2.5 mL of tetrahydrofuran at 0 °C were added drop-wise to the solution. The resulting mixture was gradually warmed to room temperature and stirred for 6 hours. The solution then was extracted by liquid/liquid extraction with chloroform/cold water, and dried. This material was further purified by the semi-prep RP column (Agilent Zorbax Rx-C18, 9.4 x 250mm, 5 µm particle size). The purity was tested using LC-FTICR-MS and LC/UV. NMR was also used to characterize the reaction products and confirm the identity and purity of the final product. The purity of the labeling reagent was >99.5% by HPLC, UV, MS and NMR analysis.

### **3.2.3 Derivatization Reaction Optimization**

In order to optimize the reaction conditions of the derivatization reaction, pooled urine (obtained by mixing equal volumes of all urine samples) was used. A volume of 25 µL of pooled urine, equivalent to approximately 200 nmol creatinine, was dried and the following reaction conditions were tested. Sample solvent: dried urine (no solvent), 5% water in acetonitrile (final concentration), 100% acetonitrile; reagent: TEA, TEOA and K<sup>+</sup> salts/18 crown 6; reagent concentration: 10 mg/mL, 20 mg/mL, 30 mg/mL; reaction temperature: 60 °C, 75 °C, 90 °C; reaction time: 30 minutes, 60 minutes 90 minutes. IS, <sup>13</sup>C<sub>2</sub>-DmPA-labeled analytes, was added for a final concentration of 100 nM.

### **3.2.4 Urine Samples**

This study was conducted in accordance with the Arts, Science & Law Research Ethics Board policy at the University of Alberta. Urine samples were collected from 10 males and 10 females, who were not known to be on any special diet or medications or have any inborn errors of metabolism. The volunteers were all adults, ages 25-30, with body mass indices (BMIs) ranging

from 18 to 34.3. The urine samples were all collected as second morning void for 3 consecutive days (a total of 60 samples). The urine was centrifuged, filtered, aliquoted and stored at -80 °C without any added preservatives, until further analysis.

### **3.2.5 Creatinine Determination**

Determination of creatinine in urine was carried out using an assay based on a kinetic Jaffe reaction using the QuantiChrom™ Creatinine Assay kit (Gentaur, Montreal, QC) with a linear detection range of 8.0 µM to 4.4 mM. Acylglycine levels were normalized to a creatinine content of 200 nmol.

### **3.2.6 Preparation and Derivatization of Standards and Samples**

#### **3.2.6.1 Standard Solutions**

A standard stock solution was prepared by dissolving each accurately weighed acylglycine into acetonitrile to give a concentration of 1.0 mM. A mixture of all acylglycines at a concentration of 10 µM in 20% water in acetonitrile was mixed with 100 µL of either <sup>12</sup>C<sub>2</sub>-DmPA or <sup>13</sup>C<sub>2</sub>-DmPA (30 mg/mL) and 100 µL TEA (30 mg/mL) in acetonitrile. The mixture was vortexed and heated in a heating block at 90 °C for 1 hour. The reaction was then terminated with 100 µL acetic acid dissolved in acetonitrile (10% v/v) reacted at 90 °C for 15 minutes. The reaction scheme is presented in Figure 3-1. The resulting mixture was evaporated to dryness and reconstituted in 100 µL acetonitrile/water (50/50, v/v), 0.1% formic acid before analysis by LC-MS/MS.

### 3.2.6.2 Urine Samples

Volumes of urine equivalent to 200 nmol creatinine were aliquoted into vials and evaporated to dryness. The dried urine was then reconstituted in 100  $\mu$ L of 20% water in acetonitrile and vortexed. To the mixture was added 100  $\mu$ L triethylamine (30 mg/mL) and 100  $\mu$ L DmPA (30 mg/mL) and the mixture heated in a heating block at 90 °C for 1 hour. The reaction was terminated with 100  $\mu$ L acetic acid dissolved in acetonitrile (10% v/v) reacted at 90 °C for 15 minutes. The resulting mixture was evaporated to dryness and reconstituted in 100  $\mu$ L acetonitrile/water (50/50, v/v), 0.1% formic acid before it was subjected to LC-MS/MS analysis.

### 3.2.7 Ultra-high Performance Liquid Chromatography (UHPLC)

Liquid chromatography was performed on an Agilent 1290 Series LC (Mississauga, ON). Chromatographic separation was done on two Phenomenex Kinetex 1.7  $\mu$ m minibore 50 x 2.1 mm C<sub>18</sub> columns. Elution conditions were as follows: linear gradient of 18% - 28% mobile phase B over 17 minutes, 28% - 43% B over 18 minutes, and 43% - 90% B over 10 minutes, where mobile phase A consisted of 2% acetonitrile, 0.1% formic acid in water and mobile phase B consisted of 2% water, 0.1% formic acid in acetonitrile. The columns were maintained at 25 °C and autosampler set at 4 °C. Flow rate was set at 0.250 mL/min and 5.0  $\mu$ L of each sample was injected. The regeneration pump performed column wash and equilibration steps in parallel on the second C<sub>18</sub> column using 100% mobile phase B and 18% mobile phase B respectively.

### 3.2.8 Mass Spectrometry

Mass analysis was carried out on an AB Sciex 4000 QTRAP<sup>®</sup> hybrid triple quadrupole linear ion trap mass spectrometer (Concord, ON), equipped with an electrospray ionization (ESI) interface. The ESI source was set to perform in

the positive ion mode. The optimized parameters were as follows: spray voltage, 4800; curtain gas, 10; CAD, high; temperature, 250 °C; GS1, 40; GS2, 30; declustering potential, 45; collision energy, 25eV. The acquisition method consisted of an information dependent acquisition (IDA) scan using multiple reaction monitoring (MRM) as a survey scan and two dependent enhanced product ion (EPI) scans. The EPI scans were in the range 50 – 800 Da, scanned at 4000 Da/s. Data was acquired and processed using Analyst,<sup>®</sup> version 1.5.1 software (Concord, ON).

### **3.2.9 Method Validation**

The method was validated to prove its efficiency by examining the following: selectivity, linearity of IS and analytes, accuracy and precision. Carry-over and matrix effects were also evaluated. All required statistical analyses were performed using R script 2.11.1 and Igor Pro 6.01.

#### **3.2.9.1 Selectivity and Carryover**

Selectivity of the method was determined by analyzing the following: an instrument blank (injecting only mobile phase A), a solvent blank (consisting of acetonitrile/water (50/50, v/v), 0.1% formic acid), a reaction blank (reaction mixture without analytes), blank surrogate matrix and authentic matrix. These were all analyzed using the LC-MS/MS method described below and were evaluated for the presence of any interfering signals. Carry-over was assessed by analyzing instrument blanks after analysis of 5 replicates of upper limit of quantification (ULOQ) standards, 5 replicates of high QC standards and 3 replicates of high concentration samples.

### 3.2.9.2 Linearity

#### 3.2.9.2.1 Comparison of Linearity of Internal Standard and Analytes

To increase the efficiency of the surrogate matrix method, a stable isotope labeled IS should be used whenever possible. In order to determine the suitability of the  $^{13}\text{C}$ -DmPA-labeled analytes as the IS, two sets of separate calibration curves were constructed: one for each analyte and one for each corresponding IS. Two sets of standards, one containing eighteen  $^{12}\text{C}$ -DmPA-labeled acylglycines at concentrations of 1, 5, 10, 50, 100, 500, 1000 nM and another containing the  $^{13}\text{C}$ -labeled acylglycines at the same concentrations were made. The calibration standards were prepared in the solvent mix and surrogate matrix separately and analyzed in triplicates. Similarities in detector response for the analytes and their IS were evaluated by comparing the slopes (sensitivity) of the regression curves obtained in surrogate and authentic matrix using a modified student t-test, at the 95% confidence level<sup>27</sup> using equation 3-1.

$$t = \frac{b_1 - b_2}{s_{b_1-b_2}}, \text{ where } s_{b_1-b_2} = \sqrt{\frac{(s^2_{Y.X})_p}{(\sum x^2)_1} + \frac{(s^2_{Y.X})_p}{(\sum x^2)_2}}$$
$$\text{and } (s^2_{Y.X})_p = \frac{(\text{residual SS})_1 + (\text{residual SS})_2}{(\text{residual DF})_1 + (\text{residual DF})_2}$$

Equation 3-1 Test statistic t from the modified t test. Parameters used: b, slope of regression line; 's<sub>b<sub>1</sub>-b<sub>2</sub></sub>', the standard error of the difference between regression coefficients; (s<sup>2</sup><sub>Y.X</sub>)<sub>p</sub>, the pooled residual mean square; subscripts 1 and 2 refer to the two regression lines being compared. Critical value has (n<sub>1</sub> + n<sub>2</sub> - 4) degrees of freedom.

#### 3.2.9.2.2 Calibration Curve

The linearity of the ESI response for the DmPA-labeled acylglycines was examined by analyzing 5 replicates at 7 different concentration levels. Serial dilutions of a mixture of eighteen  $^{12}\text{C}_2$ -DmPA-labeled acylglycines (10  $\mu\text{M}$ ) were

performed to obtain concentrations of 1, 5, 10, 50, 100, 500, 1000 nM in surrogate urine. Here, the surrogate matrix used is pooled underivatized urine prepared by mixing equal volumes of the urine samples. A fixed volume of  $^{13}\text{C}_2$ -DmPA-labeled acylglycines (final concentration 100 nM) was added to the urine as an internal standard. Peak area ratios were calculated by dividing the peak area of each analyte by the peak area of the corresponding IS. Calibration curves were constructed by plotting peak area ratios against concentration of each acylglycine. Different weighting factors were evaluated and the curves were fitted to linear regression analysis with a weighting of  $1/\text{response}$  ( $1/y$ ). The limit of detection (LOD) and the lower limit of quantification (LLOQ) were determined for each acylglycine using the equations  $3 \times (\sigma_B)/m$  and  $10 \times (\sigma_B)/m$  respectively, where  $\sigma_B$  is the standard deviation of the response and  $m$  is the slope derived from the calibration curve. The standard deviation of the blank response was estimated from the regression parameter, standard deviation of the y-intercept.<sup>28, 29</sup> The LLOQ of each analyte was confirmed experimentally by analyzing five replicates of surrogate matrix spiked with the analytes at a concentration equal to the LLOQ.

### **3.2.9.3 Authentication of the Surrogate Matrix**

#### **3.2.9.3.1 Comparison of Calibration Slopes in Surrogate and Authentic Matrices**

Calibration standards were prepared as described above by spiking the  $^{12}\text{C}$ -DmPA-labeled analytes in derivatized pooled urine (known in this study as authentic matrix) at 7 different concentration levels and analyzed by the same LC-MS/MS method. The suitability of the surrogate matrix as the calibrator matrix was evaluated by comparing the slopes (sensitivity) of the regression curves obtained in surrogate and authentic matrix using the modified student t-test described above, at the 95% confidence level.

### 3.2.9.3.2 Validation of Surrogate Matrix by Standard Addition

A standard addition experiment was performed to evaluate the suitability of underivatized urine as a surrogate matrix. The endogenous concentrations of acylglycines in a derivatized pooled urine sample were determined by analyzing three replicates of the samples by LC-MS/MS and using the surrogate matrix calibration curve as described above. A pooled urine sample was used to prepare a standard addition curve by adding a series of derivatized standards in increasing concentration. The concentration of each acylglycine in the pooled urine was determined by extrapolating to determine the magnitude of the negative x-intercept. This concentration was compared to the concentration of acylglycines in pooled urine derived from the surrogate matrix calibration curve. Percent relative error (% RE) was calculated using the following equation:

$$\% \text{ RE} = \frac{\text{surrogate matrix concentration} - \text{standard addition concentration}}{\text{standard addition concentration}} \times 100$$

Equation 3-2 Equation used to calculate percent relative error.

### 3.2.9.4 Dilution Linearity

Three volumes of pooled derivatized urine were fortified with 100 nM derivatized standards. Dilutional linearity was determined by serially diluting pooled derivatized urine 2-, 5-, 10-, and 20-fold with surrogate urine. The measured concentrations of the diluted samples were plotted against expected concentrations for each analyte and a linear regression was performed.

### 3.2.9.5 Matrix Effects

Matrix effects were assessed by comparing the slopes of the calibration curves prepared in solvent mix to those prepared in surrogate (underivatized) and authentic (derivatized) urine. The modified t-test, at the 95% confidence level,

was used to compare the slopes to evaluate if the differences in the slopes are statistically significant and indicate matrix effects. In addition ion suppression/enhancement was also calculated according to Equation 3-2.

$$\% \text{ Suppression/enhancement} = \frac{\text{Slope of curve in matrix}}{\text{Slope of curve in solvent}} \times 100 - 100$$

Equation 3-3 Equation to calculate percent suppression or enhancement.

### **3.2.9.6 Inter-day and Intra-day Precision**

The between-day and within-day precision were determined by analyzing a series of 10 injections of a derivatized pooled urine sample over 3 consecutive days under the same operating conditions. Intra-day precision was evaluated for the 10 injections (total of 10 replicates per day) and inter-day precision was evaluated for a period of three days (total of 30 replicates). Precision was evaluated by calculating coefficients of variations (% CV) from the measured concentrations.

### **3.2.9.7 Method Reproducibility**

To measure the reproducibility of the derivatization method, five replicates of pooled urine were derivatized by the method described above and analyzed by LC-MS/MS. Method precision was evaluated by calculating coefficients of variations (% CV) from the measured concentrations.

### **3.2.9.8 Accuracy**

To demonstrate the applicability of the method to accurately determine the concentrations of acylglycines in urine, three sets of pooled derivatized urine were fortified with derivatized analytes at three concentrations (three replicates per

level): 10 nM (low QC), 150 nM (medium QC) and 400 nM (high QC). The recoveries were calculated by subtracting the amount of endogenous acylglycine in an un-spiked sample from the amount found in fortified sample and comparing to the added amounts. Percent relative error (% RE) or deviation from the added concentration was calculated as a measure of accuracy.

### **3.2.9.9 Stability**

Three pooled derivatized urine samples were diluted 1:10 and fortified with derivatized analytes at a concentration of 100 nM to study the stability of the analytes. The urine stability samples were analyzed immediately after spiking (labeled “fresh” samples) and at selected times over the storage period. Stability was assessed after each freeze-thaw cycle for three cycles, after 24 hours in the autosampler at 4 °C, and after 5 hours on a benchtop at room temperature. Freezer stability was assessed by analyzing the sample after storage at -20 °C for two and eight weeks. The percent recovery of each analyte was determined by calculating the concentration at each condition and expressed as a percentage of the fresh sample.

## **3.2.10 Measurement of Endogenous Levels of Acylglycines in Human Urine**

### **3.2.10.1 Absolute Quantification**

Three volumes of urine from each individual, equivalent to 200 nmol creatinine, were dried and derivatized with  $^{12}\text{C}_2$ -DmPA as explained above and analyzed by LC-MS/MS. Peak area ratios for each acylglycine were calculated and quantification was performed using the equations derived from the calibration curves in surrogate matrix.

### 3.2.10.2 Relative Quantification

Relative quantification was done on putatively identified acylglycines, for which there were no standards available. Putative identification was done using fragmentation patterns based on a previous study.<sup>7</sup> Urine samples were prepared as described above for absolute quantification and analyzed using identical conditions. Peak area ratios were calculated using the peak area of each acylglycine and peak area of the closest eluting IS. Quantification was done using the equations derived from the calibration curves in surrogate matrix.

## 3.3 Results and Discussion

### 3.3.1 Optimization of Derivatization Reaction Conditions

The strategy of experimentation used in this study was the “one-factor-at-a-time” approach. The factors optimized were reagents (catalysts), solvents, reagents (DmPA/TEA) concentration, reaction temperature and time.

Table 3-1 shows how the factors were controlled. Water was added to acetonitrile (final 5%, v/v) as a solvent to dissolve the more polar acylglycines. Each factor level was changed over its range (amounting to 3 different levels each), while the other factors were kept constant at starting point conditions. Only thirteen of the analytes were monitored; the other five were at low levels in the pooled urine samples. A series of graphs were plotted as shown in Figure 3-2, Figure 3-3, Figure 3-4, Figure 3-5, and Figure 3-6, describing how the peak area ratios of the analytes were affected by varying each factor. Each data point plotted represents an average of three individual experiments using three individual samples. Errors bars indicate standard deviation.

The derivatization reaction, represented in Figure 3-1, between the DmPA reagent and the urinary acylglycines is an  $S_N2$  reaction and requires a solvent system in which all the reagents are soluble. The polar acylglycines found in urine

tend to be soluble in methanol, acetonitrile, and water. Acetonitrile is the best solvent because it is a polar aprotic solvent. A solvent mixture of 5% water seems to be better than using 100% acetonitrile or no solvent, especially for the acylglycines with two carboxylic acid groups because of solubility issues. It has been reported that protic or aprotic solvents can be used in the reaction without a decrease in the yield, making total exclusion of water unnecessary.<sup>30</sup> Pooled urine was used because of the averaging effect and it was thought that it would make a good representative of all 60 samples. The concentration of TEA was held at an equivalent concentration to the DmPA reagent in each experiment and different temperatures and reaction times were tested. The optimal conditions chosen for this study were: reagents, DmPA and TEA; concentration, 30 mg/mL; solvent, 5% water in acetonitrile; temperature, 90 °C and time, 60 minutes. For times longer than 1 hour, no significant increase was observed.

Table 3-1 List of conditions used in the optimization experiments

Exp.	Solvent	Reagents			Temp (°C)	Time (min)
		DmPA/TEA <sup>a</sup> (mg/mL)	DmPA/TEOA (mg/mL)/(mM)	DmPA/18-crown-6 (mg/mL)		
A	5% H <sub>2</sub> O		20/750		90	60
	5% H <sub>2</sub> O			20/1	90	60
	5% H <sub>2</sub> O	20			90	60
B	5% H <sub>2</sub> O	10			90	60
	5% H <sub>2</sub> O	20			90	60
	5% H <sub>2</sub> O	30			90	60
C	No Solvent	30			90	60
	5% H <sub>2</sub> O	30			90	60
	ACN	30			90	60
D	5% H <sub>2</sub> O	30			60	60
	5% H <sub>2</sub> O	30			75	60
	5% H <sub>2</sub> O	30			90	60
E	5% H <sub>2</sub> O	30			90	30
	5% H <sub>2</sub> O	30			90	60
	5% H <sub>2</sub> O	30			90	90

<sup>a</sup> The concentrations of DmPA and TEA are equal

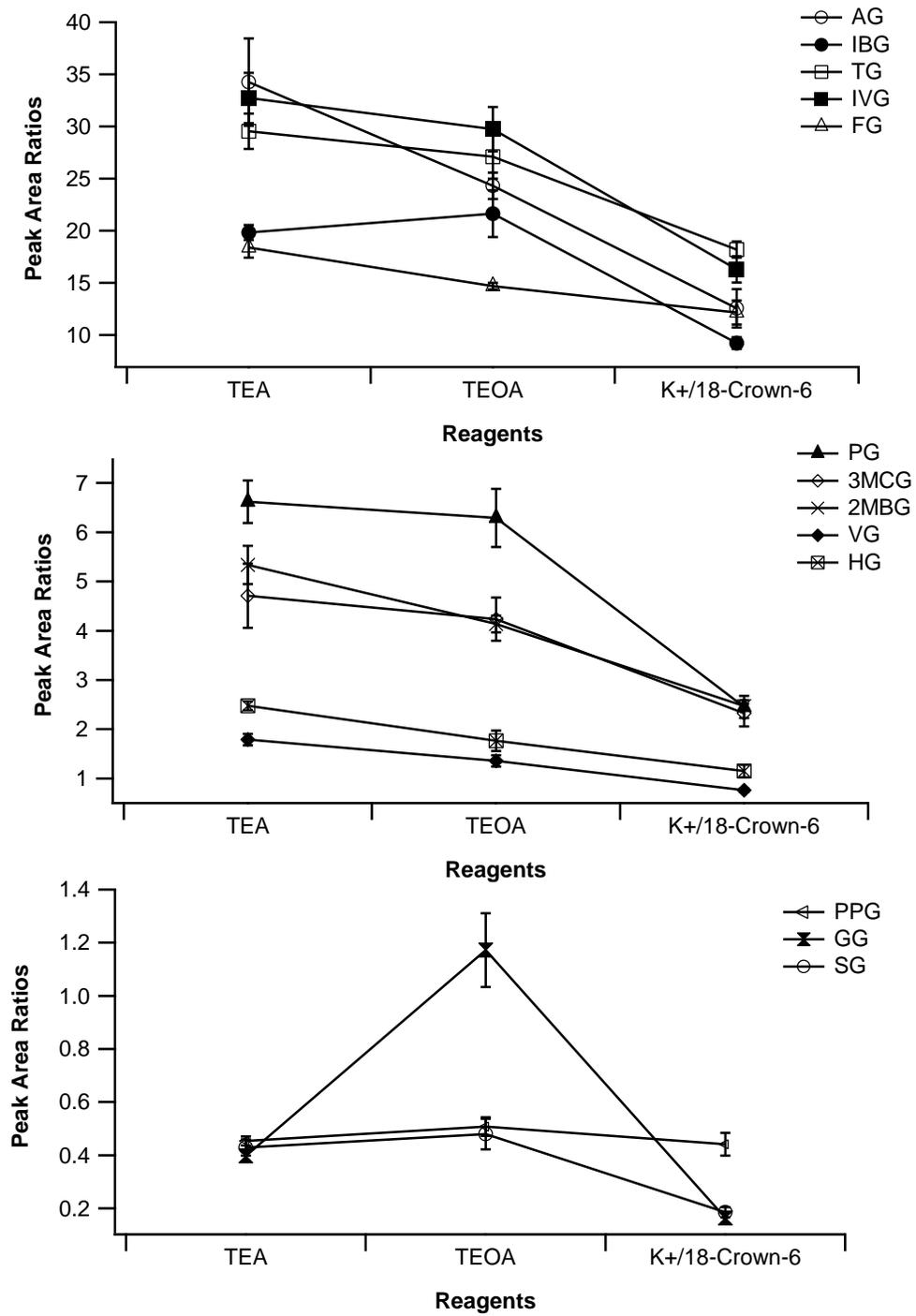


Figure 3-2 Effect of reagent catalyst on the reaction

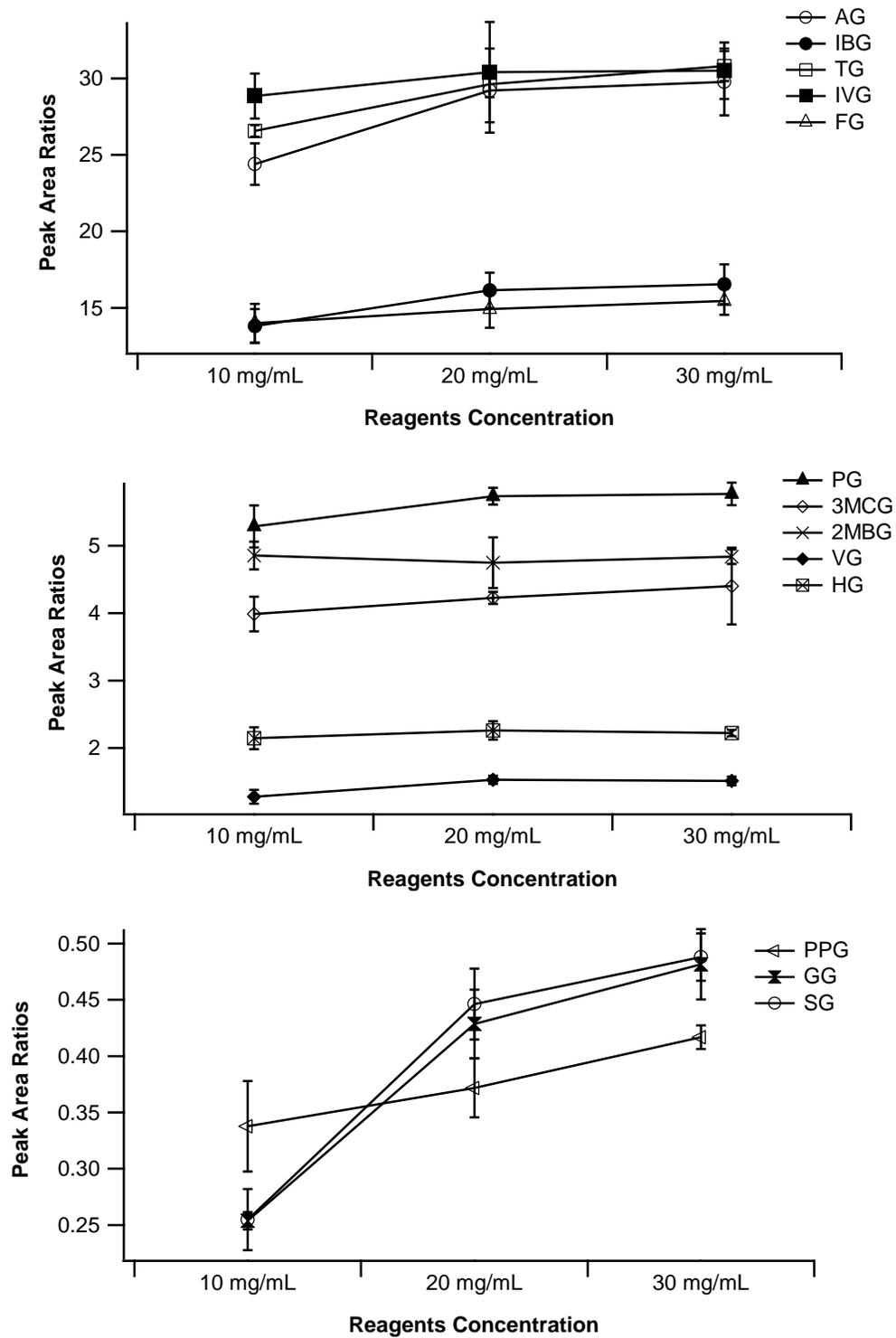


Figure 3-3 Effect of reagent concentration on the reaction

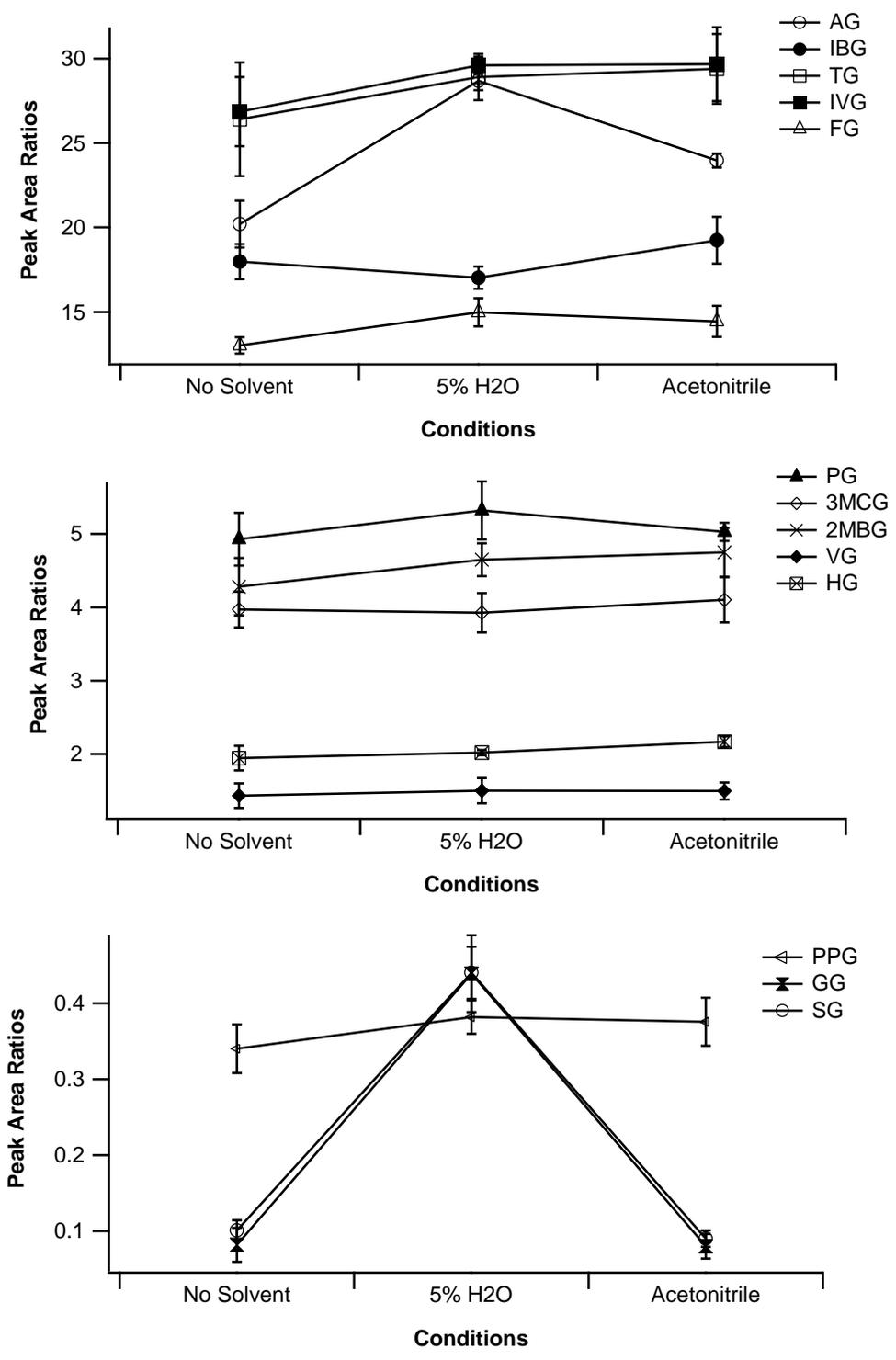


Figure 3-4 Effect of solvent conditions on the reaction

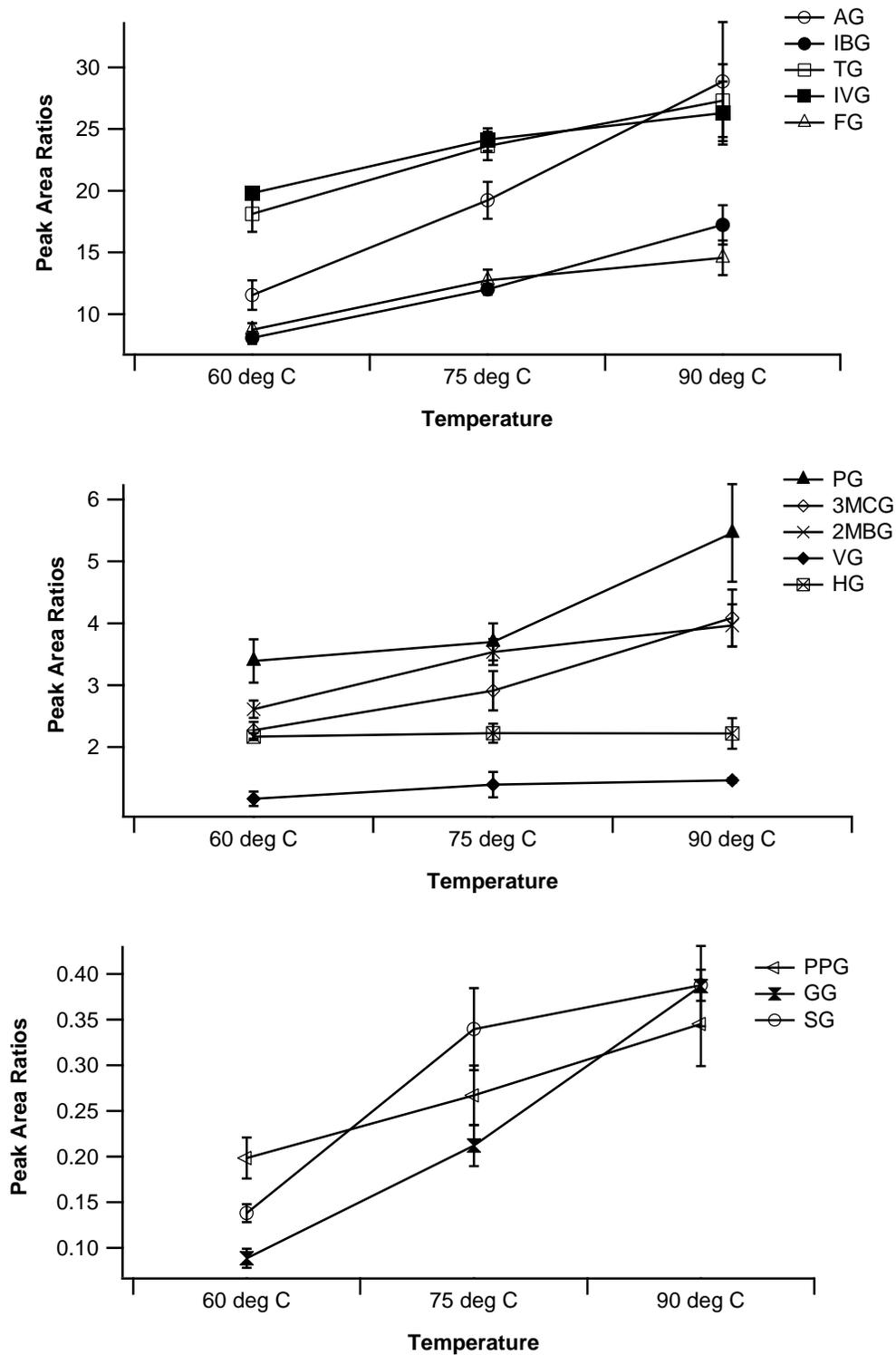


Figure 3-5 Effect of temperature on the reaction

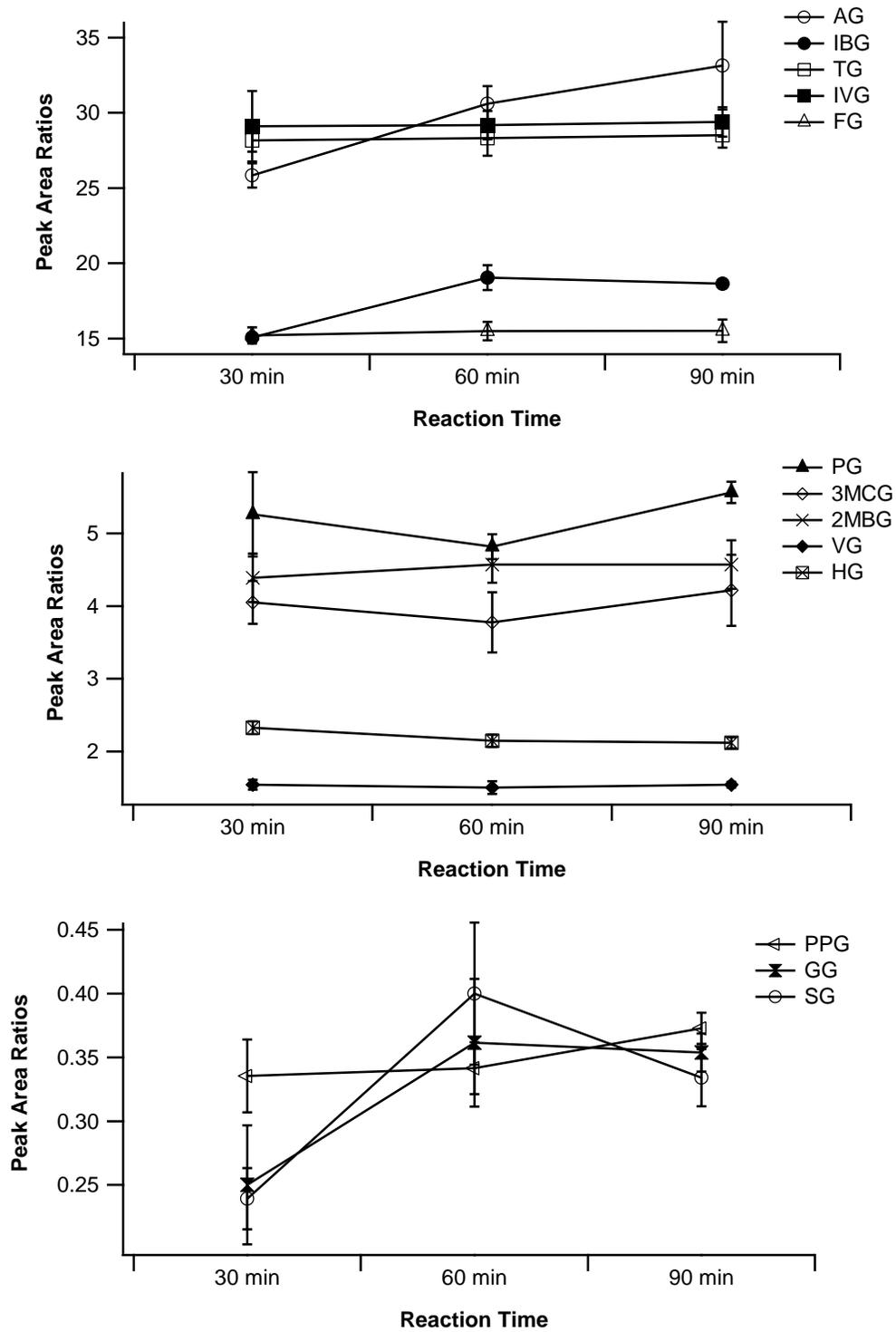


Figure 3-6 Effect of reaction time on the reaction

### 3.3.2 Effects of Derivatization on Chromatography

Chromatographic separation of polar analytes before mass spectrometric detection enhances their sensitivity, partly because it reduces ion suppression. Acylglycines are polar molecules and they eluted in a higher percentage of aqueous solvent when using a reversed phase (RP) column, especially the smaller and more polar acylglycines like acetylglycine, glutarylglycine and propionylglycine. The addition of the tag, which contains an aromatic ring increases the hydrophobicity of the acylglycines and therefore increases their retention on a RP column. This results in a higher resolution in the chromatographic separation and isomers are better separated when compared to the unlabeled species. Under the experimental conditions of the present method, the analysis run time was 37 minutes and the retention times of the eighteen derivatized acylglycines and their abbreviations are represented in Table 3-2. The incorporation of a stable-isotope labeling reagent does not result in an isotope effect in the RP separation. Co-elution of the light and heavy labeled analyte and IS respectively are essential for accurate quantification by LC-MS. Differences in retention time of analyte and IS can cause any ion suppression to have significant variability in the detection of both and can result in quantification errors. Figure 3-7 shows that the  $^{12}\text{C}/^{13}\text{C}$  -labeled analytes perfectly co-elute using the example of  $^{12}\text{C}/^{13}\text{C}$ -DmPA-labeled hexanoylglycine (HG).

### 3.3.3 Mass Analysis

The structure of a representative derivatized acylglycine, hexanoylglycine and its internal standard are shown in Figure 3-7. For these two compounds the protonated precursor molecules were observed at  $m/z$  335 and  $m/z$  337 for hexanoylglycine and IS respectively. During the IDA experiment, the precursors are subjected to collision induced dissociation to yield product ion spectra. The product ion spectra of hexanoylglycine and the IS are also shown in Figure 3-7, as well as a proposed fragmentation scheme.

Table 3-2 Masses, abbreviations and retention times of protonated DmPA-labeled acylglycines

<b>DmPA-labeled Acylglycine</b>	<b>Abbreviation</b>	<b>Protonated Mass (Da)</b>	<b>RT (min)</b>
Acetylglycine	AG	279	3.46
Propionylglycine	PG	293	5.12
Isobutyrylglycine	IBG	307	7.86
Butyrylglycine	BG	307	8.12
4-Hydroxyphenylacetylglycine	HPAG	371	8.57
2-Furoylglycine	FG	331	9.44
Tiglylglycine	TG	319	11.36
2-Methylbutyrylglycine	2MBG	321	11.86
3-Methylcrotonylglycine	3MCG	319	12.19
Isovalerylglycine	IVG	321	12.35
Valerylglycine	VG	321	13.34
Hexanoylglycine	HG	335	19.65
Phenylacetylglycine	PAG	355	20.65
Phenylpropionylglycine	PPG	369	20.82
Glutarylglycine	GG	512	22.88
Heptanoylglycine	HpG	349	23.69
Octanoylglycine	OG	363	26.60
Suberylglycine	SG	554	26.68

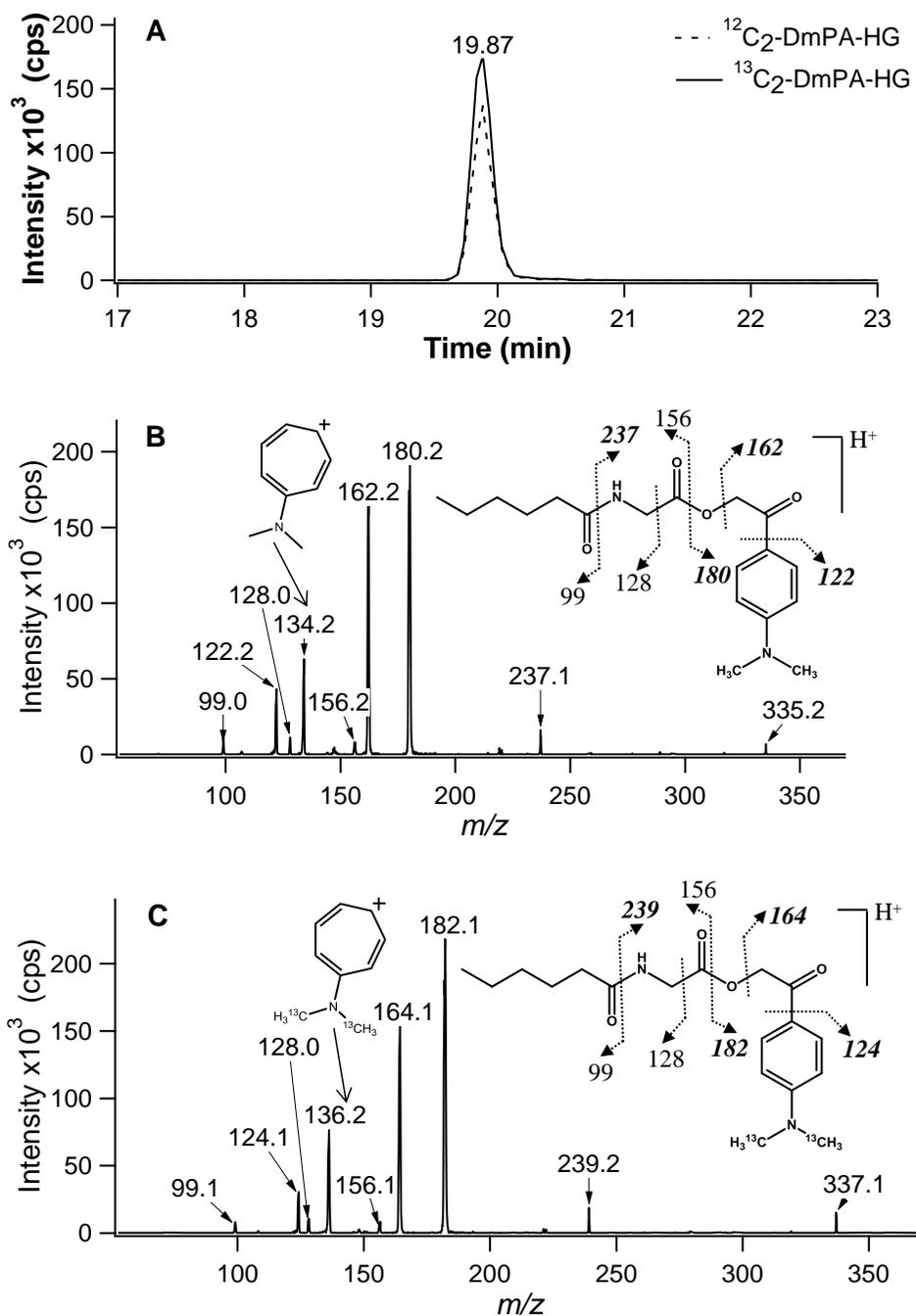


Figure 3-7 (A) Extracted ion chromatograms of analyte  $^{12}\text{C}_2\text{-DmPA}$ -labeled hexanoylglycine (HG) and internal standard  $^{13}\text{C}_2\text{-DmPA}$ -labeled hexanoylglycine (HG). (B) Product ion spectrum of  $^{12}\text{C}_2\text{-DmPA}$ -HG showing the structure and fragmentation. (C) Product ion spectrum of  $^{13}\text{C}_2\text{-DmPA}$ -HG showing the structure and fragmentation. Fragments in bold italics are the fragments of the DmPA tag. Fragments  $m/z$  134 and  $m/z$  136 are shown as insets in B and C respectively.

Under optimized conditions to yield highest sensitivity of the analytes, the most abundant fragment ions produced are  $m/z$  180 and  $m/z$  182 for hexanoylglycine and IS respectively. These two fragments are generated from the loss of a neutral ketene (on the side of the acylglycine) resulting in a protonated DmPA tag. These fragments are generated in the product ion spectra of all the analytes and their corresponding internal standards and were used in the MRM survey scans, with the Q1 ions being the precursor ions of either the analyte or IS and Q3 ion being either  $m/z$  180 or  $m/z$  182, respectively. A list of Q1 and Q3 ions used for the MRM survey scan and instrument parameters are represented in Table 3-3. Fragment ions  $m/z$  162,  $m/z$  134 and  $m/z$  122 are produced from the further fragmentation of the DmPA tag and are present in all product ion spectra of the analytes. Fragment ions  $m/z$  99,  $m/z$  128 and  $m/z$  156 are produced from dissociation of the acylglycine portion of the molecule and are more diagnostic in nature. These result from neutral losses of 236, 207, and 179 respectively and these losses were used as diagnostic losses to detect other acylglycines. Explanation of the fragmentation can be found in a previous study.<sup>7</sup> The fragment ion at  $m/z$  237 is also diagnostic because it corresponds to the protonated mass of (DmPA + glycine).

An IDA experiment was employed to generate simultaneous product ion spectra, which were used to confirm the identity of the acylglycine detected. With 63 precursor ions scanned in the MRM survey scan and a dwell time of 15 milliseconds (see Table 3-3), peak shape integrity was maintained and a minimum of 11 points across a peak could still be obtained. No cross-talk was detected in the mass spectrometric analysis. This proposed method enabled the determination of low-level concentrations of derivatized acylglycines in human urine.

Table 3-3 MRM masses and parameters

Analyte		Internal Standard		DP	CE	Dwell time (ms)
Q1	Q3	Q1	Q3			
265.1	180.1	267.1	180.1	45	18	15
279.1	180.1	281.1	182.1	45	18	15
293.1	180.1	295.1	182.1	45	18	15
307.2	180.1	309.2	182.1	45	18	15
313.2	180.1			45	18	15
319.2	180.1	321.2	182.1	45	18	15
321.2	180.1	323.2	182.1	45	20	15
323.2	180.1			45	20	15
331.2	180.1	333.2	182.1	45	20	15
333.2	180.1			45	20	15
335.2	180.1	337.2	182.1	45	20	15
337.2	180.1			45	20	15
345.2	180.1			45	20	15
347.2	180.1			45	20	15
349.2	180.1	351.2	182.1	45	20	15
351.2	180.1			45	20	15
355.2	180.1	357.2	182.1	45	20	15
356.2	180.1	358.2	182.1	45	20	15
357.2	180.1			45	20	15
359.2	180.1			45	20	15
361.2	180.1			45	20	15
363.2	180.1	365.2	182.1	45	20	15
365.2	180.1			45	20	15
367.2	180.1			45	20	15

Analyte		Internal Standard		DP	CE	Dwell time (ms)
Q1	Q3	Q1	Q3			
369.2	180.1	371.2	182.1	45	20	15
371.2	180.1	373.2	182.1	45	20	15
373.2	180.1			45	20	15
375.2	180.1			45	20	15
377.2	180.1			45	20	15
379.2	180.1			45	20	15
383.2	180.1			45	20	15
385.2	180.1			55	20	15
387.2	180.1			55	20	15
393.3	180.1			55	20	15
403.3	180.1			55	20	15
405.3	180.1			55	20	15
407.3	180.1			55	20	15
411.3	180.1			55	20	15
417.3	180.1			55	20	15
419.3	180.1			55	20	15
421.3	180.1			55	20	15
437.3	180.1			55	20	15
498.4	319.2			55	20	15
512.4	333.2	516.4	335.2	55	20	15
526.4	347.2			55	20	15
540.4	361.2			55	22	15
554.4	375.2			55	22	15
558.4	377.2			55	22	15

### 3.3.4 Enhancement of the ESI Signal

Three volumes of a urine sample (normalized to creatinine) were analyzed before and after derivatization. The ESI signals of DmPA-labeled acylglycines were compared to their unlabeled counterparts and can be represented as a bar chart in Figure 3-8 (same amount injected). The signals for the labeled acylglycines were approximately 1-3 orders of magnitude higher than the signal for the corresponding unlabeled acylglycines.

Also, despite the differences in the chain lengths and aromaticity of the fatty acid portion of the glycine conjugates, similar responses are observed for all the acylglycines in the ESI (note the similarity of their calibration slopes). Equimolar concentrations of standards produce similar responses in ESI. The signal enhancement can be due to several factors. First, the increased pKa of the labeled acylglycines, due to the presence of the dimethylamine moiety, leads to preferential protonation in solution, which is advantageous for the generation of gas-phase ions in ESI. Secondly, the increase in hydrophobicity of the labeled acylglycines allows them to be less solvated and increases their chances of staying on the surface layer of the droplets formed in electrospray. In comparison with the native acylglycines, they are more likely to have a higher surface activity and are therefore more likely to be significantly enriched in the daughter droplets, which are eventually desolvated and form gas-phase ions. Ionization efficiency is also enhanced because the labeled analyte now elutes in a higher percentage organic solvent rather than in a solvent with higher water content as with the corresponding native analyte. As the retention time increases in reversed-phase chromatography, the eluting analytes are less hydrated and therefore more hydrophobic.

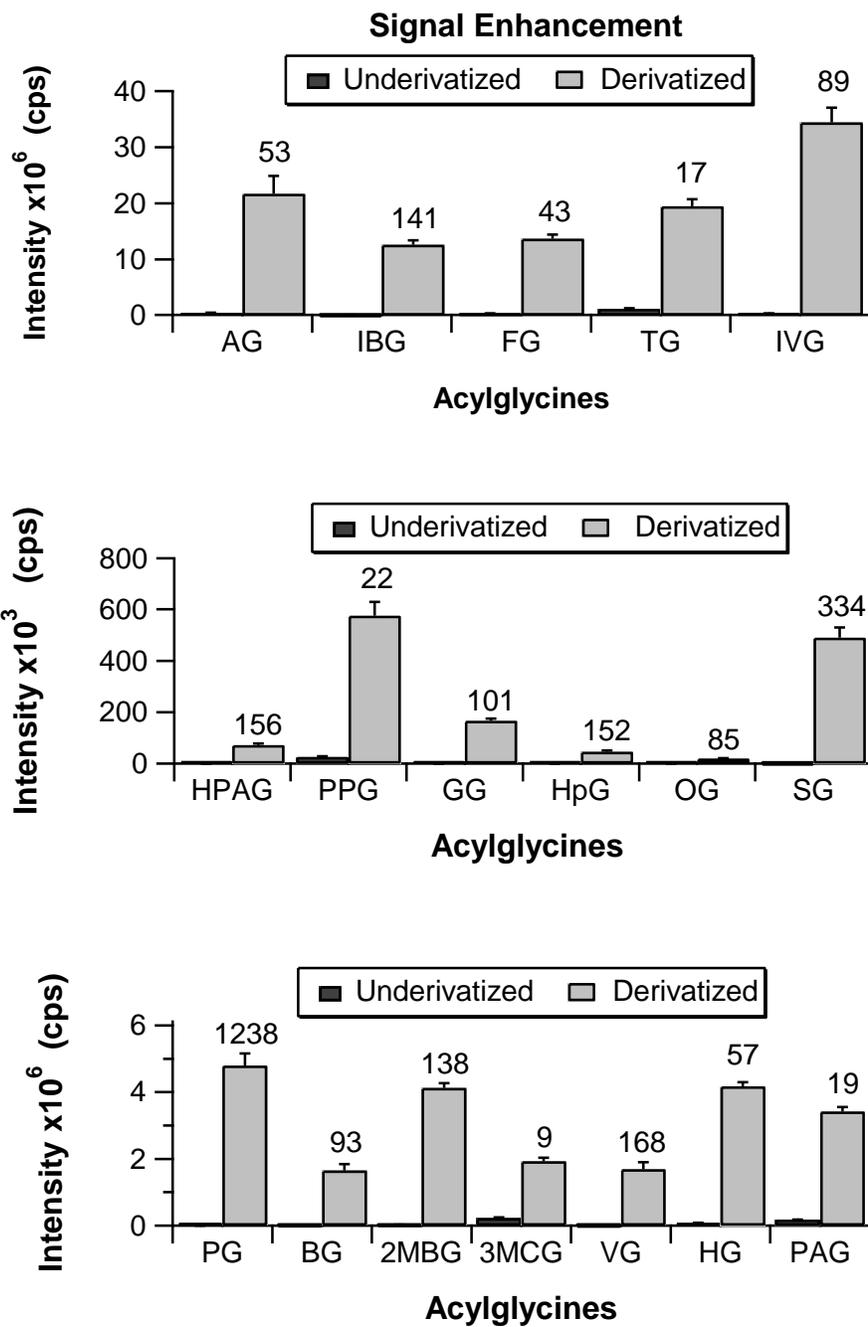


Figure 3-8 Signal enhancement using the DmPA tag. Acylglycines are grouped by peak area ranges. Numbers in bold above the bars are the fold increase of derivatized to underivatized.

### **3.3.5 Selection of Surrogate Matrix**

It has been established that one of the major challenges in validating methods for analysis of endogenous compound analysis is obtaining a true blank matrix without the presence of analytes in order to construct a calibration curve. Initially, a solvent mix of water/ acetonitrile, (50/50, v/v) 0.1% formic acid was used to prepare the calibration curve. Water alone could not be used because of the poor solubility of the derivatized acylglycines. However, when comparing the acylglycine concentrations using standard addition and using a calibration equation in solvent mix, the relative differences were higher than the recommended guideline of 15%. Also, as will be discussed below, matrix effects were observed in the urine. Using diluted derivatized urine was not an option because several of the analytes' concentrations were too high and the dilution would have to be greater than 10-fold. For the measurement of derivatized acylglycines in urine, pooled underivatized urine (equal volumes from all samples) was used as the surrogate matrix because the latter does not contain the derivatized acylglycines. Not only does it not contain the target analytes, it is quite similar to the authentic matrix; more so than the urine obtained from other mammalian species. It is also likely that most mammalian species would have endogenous levels of acylglycines.

### **3.3.6 Method Validation**

#### **3.3.6.1 Selectivity and Carryover**

In this study of endogenous analytes, selectivity was more difficult to evaluate because an analyte-free control matrix was difficult to obtain. Selectivity assessment included the analysis of an instrument blank, a solvent blank, a derivatized blank and blank surrogate matrix to determine interferences with the analyte and internal standard signals and analysis of the authentic matrix to determine interferences with the internal standard signal. Carryover was also

assessed by analyzing instrument blanks after analysis of several replicates of high concentration standards and samples. Representative chromatograms of blanks, surrogate and authentic matrices are shown in Figure 3-9. The chromatograms of the solvent blank, derivatized blank and surrogate blank in Figure 3-9 (A), (B) and (D) respectively revealed that there are no interfering peaks at the retention times where each analyte and IS are expected to elute, indicating sufficient selectivity of the method. No significant carryover ( $\leq 5\%$  of LLOQ) originating from the method was observed in Figure 3-9 (C). In Figure 3-9 (E), the MRM chromatogram with the transitions corresponding to the internal standards in authentic matrix is shown. Only two of the peaks at retention times 3.44 minutes and 11.41 minutes are observed, corresponding to the isotopic contribution from acetylglycine and tiglylglycine (two of the most abundant acylglycines in urine). The isotopic contribution of the  $^{13}\text{C}$  of the analytes can be accounted for and subtracted from the internal standard signal. Figure 3-9 (F) and (G) display a standard mix (1nM) and an internal standard mix (100 nM) in surrogate matrix respectively. No interferences were observed.

### **3.3.6.2 Linearity of Internal Standard and Analytes**

The concept behind the use of an internal standard methodology is that it is chosen to mimic the analyte of interest so that any changes in extraction or instrumental techniques would be reflected in a change in the IS as well. In most methods, a calibration curve is set up by analyzing a set of standards containing various concentrations of the analyte and one concentration of IS. The assumption here is that the IS has the same detector response as the analyte and that a calibration curve of the IS will have the same linearity as that of the analyte. Choosing the concentration of the IS then becomes important because there could be a chance that the linearity of the analyte is different from the IS and the chosen concentration of the IS might lie in a non-linear region of the curve.

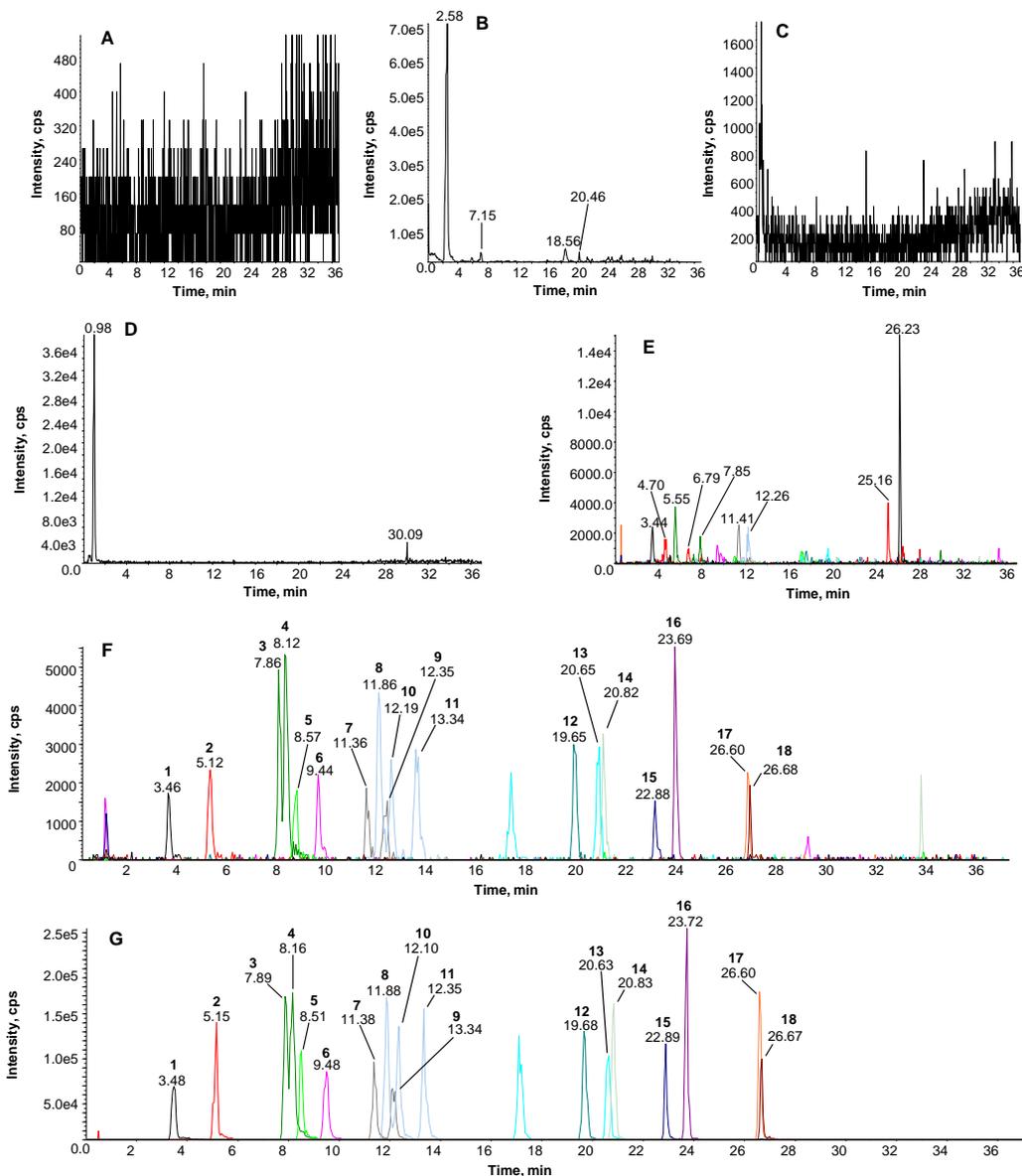


Figure 3-9 Representative LC-MS/MS chromatograms of (A) solvent blank , (B) derivatization blank, (C) blank after running ULOQ calibration standard (n=5), (D) surrogate blank, (E) internal standards (100 nM) in authentic matrix, (F) analytes (1 nM) in surrogate matrix and (G) internal standards in surrogate matrix. Compounds: 1, acetylglycine; 2, propionylglycine; 3, isobutyrylglycine; 4, butyrylglycine; 5, 4-hydroxyphenylacetylglycine; 6, 2-furoylglycine; 7, tiglylglycine; 8, 2-methylbutyrylglycine; 9, 3-methylcrotonylglycine; 10, isovalerylglycine; 11, valerylglycine; 12, hexanoylglycine; 13, phenylacetylglycine; 14, phenylpropionylglycine; 15, glutarylglycine; 16, heptanoylglycine; 17, octanoylglycine; 18, suberylglycine.

As suggested by A. K. Hewavitharana,<sup>31</sup> the more accurate way to calculate the concentration of the analyte when using a single IS calibration point is to ensure linearity around the chosen IS concentration. The linearity studies comparing <sup>12</sup>C-DmPA and <sup>13</sup>C-DmPA-labeled acylglycines are represented in Table 3-4. For each of the eighteen analytes, calibration curves were constructed and the slopes recorded as a comparison of detector response. The slopes were then compared statistically and calculated t-values compared to the critical values of t at the 95% confidence level. As is observed in Table 3-4, the t-values are all less than the critical values indicating that there are no significant differences between the slopes. A concentration of 100 nM was chosen for the IS and linearity was ensured for higher as well as lower concentration around this point. Because of ion suppression and enhancement that can be observed in electrospray ionization, the concentration of the IS can be lower or higher than expected and this linearity test ensures the accuracy of the method.

Table 3-4 Internal Standard suitability. Comparison of linearity of  $^{12}\text{C}$  and  $^{13}\text{C}$ -labeled analytes

Acyl glycines	Calibration Curve in Solvent			Calibration Curve in Surrogate Matrix		
	$^{12}\text{C}$ Slopes	$^{13}\text{C}$ Slopes	t-value <sup>a</sup>	$^{12}\text{C}$ Slopes	$^{13}\text{C}$ Slopes	t-value <sup>a</sup>
AG	1.18E+07	1.18E+07	0.09	9.36E+06	9.38E+06	0.13
PG	1.59E+07	1.57E+07	0.11	1.64E+07	1.74E+07	1.34
IBG	2.32E+07	2.41E+07	0.89	2.51E+07	2.52E+07	0.51
BG	3.76E+07	3.80E+07	0.28	3.87E+07	3.89E+07	0.33
HPAG	1.54E+07	1.56E+07	0.42	1.34E+07	1.33E+07	0.47
FG	1.70E+07	1.70E+07	0.10	1.59E+07	1.60E+07	0.07
TG	1.48E+07	1.44E+07	2.00	1.31E+07	1.31E+07	0.23
2MBG	2.88E+07	2.84E+07	0.39	2.54E+07	2.53E+07	0.31
3MCG	1.27E+07	1.36E+07	1.92	1.49E+07	1.48E+07	0.72
IVG	1.90E+07	1.89E+07	0.12	1.90E+07	1.91E+07	1.68
VG	2.25E+07	2.21E+07	0.56	2.20E+07	2.22E+07	1.18
HG	2.25E+07	2.28E+07	0.60	2.53E+07	2.50E+07	0.89
PAG	1.98E+07	2.12E+07	1.66	2.00E+07	2.03E+07	2.01
PPG	2.15E+07	2.19E+07	0.65	2.06E+07	2.06E+07	0.08
GG	8.28E+06	8.63E+06	0.88	5.18E+06	5.13E+06	0.80
HpG	2.85E+07	2.89E+07	0.74	2.72E+07	2.71E+07	0.48
OG	2.10E+07	2.17E+07	1.72	1.92E+07	1.92E+07	0.96
SG	8.93E+06	8.35E+06	1.92	5.21E+06	5.03E+06	1.94

<sup>a</sup> t-values do not exceed critical values (critical value = 2.306, except for PG which is 2.228)

### 3.3.6.3 Calibration, Authentication of Surrogate, LOD and LLOQ

The linearity of the method was investigated in solvent, surrogate matrix and authentic matrix, covering the range of 1.0 – 1000 nM. Slopes, intercepts, coefficients of determination ( $R^2$ ) and standard errors were determined by the least-squares linear regression model. Due to the heteroscedasticity of the data, the best linear fit and residuals of the model was achieved using a weighting factor of  $1/y$ . The calibration curve was linear over the range 1.0 – 1000 nM for the solvent and 1.0 – 500 nM for the surrogate and authentic matrices, except for propionylglycine (linear range 1.0 – 1000 nM). The endogenous levels of eight acylglycines, acetylglycine, isobutyrylglycine, tiglylglycine, 3-methylcrotonylglycine, isovalerylglycine, 2-furoylglycine, hexanoylglycine, and phenylacetylglycine, were on the upper end of the linear dynamic range and standard addition calibration curves of these eight analytes were non-linear. Surrogate and authentic matrices were therefore diluted 1:9 with solvent and fortified with increasing concentrations of these eight acylglycines and calibration curves were then constructed.

In order to determine if the responses of the analytes were similar in surrogate and authentic matrices, a comparison of slopes was performed using the modified t-test. Similarities in the detector responses in both matrices would authenticate the use of the surrogate matrix in the validation process. Table 3-5 summarizes the calibration parameters for both surrogate and authentic matrices. Comparison of the sensitivity for acylglycines in surrogate and authentic matrices using the modified t-test indicates that the two slopes were not significantly different at the 95% confidence level. This indicates that the surrogate matrix can be used as a substitute for the authentic matrix and the quantitation method will be accurate. The  $R^2$  values of the weighted calibration curves were all higher than 0.9980 for each acylglycine. Overall CV % values, calculated as an average of the CV% at each level, are all less than 15%. Examples of 4 calibration plots are represented in Figure A3-1 in the Appendix. Plots of all 18 calibration curves are

presented in Figure S3-1 in the Electronic Appendix. A statistical summary of the calibration curves is represented in Table A3-1 in the Appendix.

LOD and LLOQ values were calculated using the standard deviation of the y-intercept and the slope, derived from the linear regression model. The LOD ranges in value from 1.0 – 2.0 nM and LLOQ ranges from 2.0 – 5.0 nM. This information can also be found in Table 3-5. The % RE and % CV derived from the replicate analysis at the LLOQ concentration of each acylglycine ranges from -17.47 and 7.34 and 3.67 – 15.21 respectively, indicating good accuracy and precision. This information is summed up in Table 3-6. It is important to note that the signal-to-noise ratios of the LLOQ at the respective concentrations were significantly higher than 10 for all analytes, except for glutaryl-glycine. This indicates that these values of LLOQ would be considerably lower if LLOQ was measured using the signal-to-noise method and this is better than current methods.<sup>8, 10</sup>

Table 3-5 Comparison of slopes in surrogate and authentic matrices and calibration parameters of calibration curve in surrogate urine

Acylglycines	Urine Dilution	Slopes (Surrogate)	Slopes (Authentic)	t-value <sup>a</sup>	Linearity (R <sup>2</sup> )	LOD (nM)	LLOQ (nM)	Linear Dynamic Range	Overall %CV
AG	10x	8.6	9.1	0.48	0.9993	1	3	3 - 500	12.0
PG	1x	9.15	8.8	1.25	0.9996	1	3	3 - 1000	8.2
IBG	10x	10.4	10.9	1.93	0.9996	1	2	2 - 500	14.9
BG	1x	10.0	10.3	1.97	0.9992	1	4	4 - 500	12.8
HPAG	1x	8.7	8.5	1.43	0.9985	1	5	5 - 500	10.8
FG	10x	8.90	8.9	0.01	0.9997	1	2	2 - 500	10.0
TG	10x	8.6	8.8	0.52	0.9986	1	5	5 - 500	11.7
2MBG	1x	9.2	9.1	0.75	0.9992	1	3	3 - 500	8.5
3MCG	1x	9.4	9.69	1.95	0.9987	1	4	4 - 500	9.1
IVG	10x	10.1	9.4	1.94	0.9994	1	3	3 - 500	11.9
VG	1x	9.2	9.1	0.50	0.9983	2	5	5 - 500	7.1
HG	10x	9.25	8.9	1.91	0.9997	1	2	2 - 500	10.8
PAG	10x	9.3	9.62	1.93	0.9995	1	3	3 - 500	10.8
PPG	1x	10.0	9.8	1.37	0.9996	1	2	2 - 500	10.1
GG	1x	8.8	8.69	1.61	0.9995	1	2	2 - 500	12.5
HpG	1x	9.2	9.0	0.95	0.9987	1	4	4 - 500	10.5
OG	1x	9.5	9.4	0.20	0.9994	1	3	3 - 500	8.1
SG	1x	8.3	8.61	2.14	0.9987	1	4	4 - 500	14.9

<sup>a</sup> t-values do not exceed critical values (critical value = 2.228, except for PG which is 2.201)

Table 3-6 Lower limits of quantification (LLOQ) measurements in surrogate urine (n=5)

Acylglycines	LLOQ (nM)	Measured Concentration (nM)	RE (%)	CV (%)	S/N Ratio
AG	3	2.7 ± 0.3	-11.0	9.7	40 ± 2
PG	3	2.7 ± 0.4	-9.2	14.2	67 ± 8
IBG	2	1.9 ± 0.3	-5.2	14.8	41 ± 3
BG	4	4.2 ± 0.5	5.5	12.3	115 ± 13
HPAG	5	5.3 ± 0.4	5.9	7.6	77 ± 19
FG	2	1.7 ± 0.1	-16.9	6.8	17 ± 2
TG	5	4.6 ± 0.4	-7.4	7.5	59 ± 7
2MBG	3	2.8 ± 0.3	-7.7	10.3	79 ± 27
3MCG	4	4.1 ± 0.6	3.0	14.6	73 ± 6
IVG	3	2.8 ± 0.3	-6.9	10.3	52 ± 20
VG	5	5.1 ± 0.3	1.1	6.3	92 ± 15
HG	2	1.7 ± 0.1	-17.5	6.6	32 ± 7
PAG	3	2.5 ± 0.1	-17.1	3.7	50 ± 5
PPG	2	1.7 ± 0.2	-14.7	13.8	32 ± 4
GG	2	2.2 ± 0.2	8.1	11.2	13 ± 2
HpG	4	3.7 ± 0.4	-7.0	11.2	111 ± 19
OG	3	2.9 ± 0.4	-4.6	15.2	157 ± 21
SG	4	4.3 ± 0.5	7.3	12.1	40 ± 7

### 3.3.6.4 Standard Addition Results

Due to the presence of endogenous acylglycines in derivatized urine, the use of underivatized urine as a blank matrix was validated using a standard addition experiment. Comparison of the acylglycine concentrations determined

using a standard addition method and calibration in surrogate matrix can be observed in Table 3-7. Percent relative differences calculated were all less than 15%, except for HpG, OG and HPAG, whose endogenous levels are below the LLOQ. The LLOQ values presented here is represented in units of ( $\mu\text{mol/mol}$  creatinine). These values of % RE support the use of underivatized urine as the surrogate urine.

Table 3-7 Comparison of the standard addition method and the surrogate matrix approach

Acylglycines	Standard Addition Conc ( $\mu\text{mol/mol}$ creatinine)	Surrogate Calibration Conc. ( $\mu\text{mol/mol}$ creatinine)	RE (%)
AG	2746 $\pm$ 190	2624 $\pm$ 3	-4.5
PG	163 $\pm$ 13	164 $\pm$ 2	0.4
IBG	530 $\pm$ 25	516 $\pm$ 1	-2.7
BG	26 $\pm$ 2	23 $\pm$ 2	-13.8
HPAG	3.3 $\pm$ 0.5	< 2.5	N/A
FG	682 $\pm$ 29	594 $\pm$ 2	-12.9
TG	1489 $\pm$ 75	1494 $\pm$ 3	0.3
2MBG	9 $\pm$ 3	93 $\pm$ 2	-5.8
3MCG	165 $\pm$ 15	163 $\pm$ 2	-0.7
IVG	1312 $\pm$ 48	1125 $\pm$ 1	-14.2
VG	28.7 $\pm$ 0.7	27 $\pm$ 3	-5.5
HG	78 $\pm$ 24	70 $\pm$ 1	-10.0
PAG	65 $\pm$ 16	63 $\pm$ 2	-4.4
PPG	9.5 $\pm$ 0.9	10 $\pm$ 1	1.1
GG	6.9 $\pm$ 0.3	7.0 $\pm$ 3	1.6
HpG	< 2.0	< 2.0	N/A
OG	< 1.5	< 1.5	N/A
SG	14 $\pm$ 1	12 $\pm$ 3	-9.3

### 3.3.6.5 Dilution Linearity

The premise of the parallelism and dilution linearity is that if the diluting matrix is suitable and there are no matrix interferences from either matrix then the concentration of the undiluted sample and the back calculated concentrations should be the same. The dilution linearity was assessed for all analytes by plotting observed results vs. expected results and performing linear regression. Results are shown in Table 3-8. The slope of the resulting line and the correlation coefficient ( $R^2$ ) is given for each analyte. Parallelism is proven when the slopes are nearly 1. Acylglycine concentrations at different dilutions showed good agreement indicating that the method had good dilution linearity. These results also indicate that there are no matrix interferences when surrogate urine is the diluent and that the surrogate is a suitable matrix. Surrogate matrix would be used to dilute samples with concentration higher than the ULOQ.

### 3.3.6.6 Evaluation of Matrix Effects

The most common methods of evaluating matrix effects are post-column infusion and post-extraction addition. These methods are only appropriate when there are blank matrices available. Another option available when there are endogenous levels of analytes is a comparison of slopes in matrix or matrix-match solutions with the slopes obtained in matrix-free standard solution.<sup>32, 33</sup> One indicator that the sample matrix affects the analysis is a deviation in the slopes. The slopes of the calibration curves in solvent were compared to the slopes in both surrogate and authentic urine using the modified t-test defined above and the results are reported in Table 3-9, as well as calculations of matrix effects according to the equation defined in the experimental section. Negative values signify suppression, whereas positive values indicate enhancement. All analytes, with the exception of isobutyrylglycine, experienced suppression in the urine samples, though the suppression was rather modest (less than 20%). The modified t-test demonstrated that for most of the analytes, the slopes in solvent were significantly different than

those in either of the urine samples. On the other hand, Table 3-5 shows that the slopes in surrogate and authentic matrices are not significantly different at the 95% confidence level. To minimize matrix effects and for a more accurate validation and quantification, all calibrations and validations were done using either surrogate or authentic urine samples.

Table 3-8 Dilution linearity

Acylglycines	Slope	R <sup>2</sup>
AG	0.9680	0.9978
PG	0.9873	0.9978
IBG	0.9955	0.9997
BG	1.0078	0.9971
HPAG	1.0489	0.9914
FG	0.9977	0.9995
TG	0.9811	0.9975
2MBG	1.0024	0.9970
3MCG	0.9938	0.9992
IVG	0.9828	0.9987
VG	0.9964	0.9986
HG	0.9966	0.9990
PAG	1.0021	0.9943
PPG	1.0091	0.9967
GG	1.0199	0.9988
HpG	1.1082	0.9983
OG	1.0023	0.9593
SG	0.9506	0.9983

Table 3-9 Matrix Effects study. Comparison of slopes in solvent to slopes in authentic urine using a modified t-test

Acyl glycine	Solvent	Surrogate			Authentic		
	Slope	Slope	t- value <sup>a</sup>	Matrix Effects	Slope	t- value <sup>b</sup>	Matrix Effects
AG	10.4	8.6	<b>4.25</b>	-17.0	9.1	1.30	-12.6
PG	10.3	9.15	<b>8.53</b>	-11.1	8.8	<b>6.12</b>	-14.8
IBG	10.2	10.4	0.18	1.8	10.9	0.67	6.5
BG	10.3	10.0	1.32	-3.1	10.3	0.06	-0.1
HPAG	10.1	8.7	<b>8.87</b>	-13.6	8.5	<b>15.12</b>	-15.7
FG	10.2	8.90	<b>15.52</b>	-12.7	8.9	<b>9.00</b>	-12.7
TG	10.0	8.6	<b>8.24</b>	-14.2	8.8	<b>3.67</b>	-11.5
2MBG	10.3	9.2	<b>9.48</b>	-10.0	9.1	<b>9.34</b>	-10.9
3MCG	10.2	9.4	<b>4.99</b>	-7.7	9.69	<b>4.22</b>	-4.6
IVG	10.1	10.1	0.19	-0.4	9.4	<b>3.03</b>	-7.4
VG	10.4	9.2	<b>4.33</b>	-11.1	9.1	<b>5.13</b>	-11.8
HG	9.9	9.25	<b>2.68</b>	-6.3	8.9	<b>5.85</b>	-10.2
PAG	10.2	9.3	<b>6.28</b>	-8.3	9.62	<b>4.94</b>	-5.4
PPG	10.4	10.0	<b>2.77</b>	-4.4	9.8	<b>5.45</b>	-6.4
GG	10.3	8.8	<b>11.33</b>	-14.2	8.69	<b>21.89</b>	-15.3
HpG	10.1	9.2	<b>4.38</b>	-8.8	9.0	<b>7.64</b>	-10.5
OG	10.5	9.5	<b>5.71</b>	-9.6	9.4	<b>5.64</b>	-10.1
SG	10.4	8.3	<b>11.00</b>	-19.7	8.61	<b>15.59</b>	-17.0

<sup>a</sup> t-values in bold exceeded critical values at 95% confidence level (critical value = 2.262; PG, 2.228) ; <sup>b</sup> t-values in bold exceeded critical values at 95% confidence level (critical value = 2.228; PG, 2.201)

### **3.3.6.7 Method Reproducibility**

Method reproducibility was defined through the analysis of five experimental replicates under the same reaction conditions. The obtained results are presented in Table 3-10. Results are expressed as % CV and the values are all under 10% for all acylglycines, except for heptanoylglycine and octanoylglycine, whose concentrations were at or below the LLOQ. The reported % CV values are 12.54 and 14.76% respectively, which are still under 15% signifying that the current derivatization method demonstrated adequate precision.

### **3.3.6.8 Precision and Accuracy**

Intra- and inter-day precision of the proposed method were also determined and presented in Table 3-10. Results are expressed as % CV and intra-day values varied between 2.50% and 11.40% and inter-day values varied between 3.11% and 11.70%. Accuracy was investigated for three concentrations, 10.0, 150.0, 400.0 nM and the results are summarized in Table 3-11. The results are expressed as % RE and % CV and both sets of values are lower than 15% for all analytes at all three concentrations. These results indicate that the proposed method demonstrated satisfactory accuracy and precision.

### **3.3.6.9 Stability**

Samples were considered to be stable if the concentration of analytes were within  $\pm 15\%$  (i.e. 85-115%) of initial concentration in the fresh samples. Results of the stability study conducted under various conditions are expressed as percent recovery and are summarized in Table 3-12. All eighteen acylglycines were within the acceptable limits and were shown to be stable in human urine samples after 3 freeze/thaw cycles and up to 8 weeks storage at  $-20\text{ }^{\circ}\text{C}$ . It was also

demonstrated that they were stable for up to 5 hours at room temperature and 24 hours at 4 °C in the autosampler.

Table 3-10 Intra-day, inter-day and derivatization precision of the proposed method

<b>Acylglycines</b>	<b>Intra-day CV % (n=10)</b>	<b>Inter-day CV % (n=30)</b>	<b>Method CV % (n=5)</b>
AG	2.9	3.1	3.7
PG	8.3	11.7	6.9
IBG	6.7	7.9	6.0
BG	7.8	8.3	9.7
HPAG	2.6	6.2	7.3
FG	5.5	6.7	6.1
TG	2.5	3.1	6.4
2MBG	7.5	6.9	7.3
3MCG	6.4	6.4	9.5
IVG	2.8	4.4	9.5
VG	8.0	7.2	7.5
HG	3.5	3.9	5.1
PAG	3.3	4.8	5.2
PPG	11.4	10.9	9.2
GG	9.2	11.3	4.1
HpG	4.1	3.8	12.5
OG	4.2	5.7	14.8
SG	5.7	5.4	4.5

Table 3-11 Accuracy data obtained using the proposed method

Acyl glycines	QC Low (10.00 nM)			QC Mid (150.00 nM)			QC Mid (400 nM)		
	Measured value	RE (%)	CV (%)	Measured value	RE (%)	CV (%)	Measured value	RE (%)	CV (%)
AG	10 ± 1	4.2	11.3	155 ± 21	3.6	13.6	380 ± 34	-5.0	9.1
PG	9.6 ± 0.6	-4.3	6.6	148 ± 5	-1.0	3.6	362 ± 21	-9.6	5.9
IBG	10 ± 1	-0.9	11.0	150 ± 10	-0.1	6.7	379 ± 26	-5.3	7.0
BG	10 ± 1	-0.2	9.7	147 ± 3	-2.0	2.3	361 ± 12	-9.8	3.4
HPAG	11 ± 1	7.7	11.8	149 ± 2	-0.4	1.4	367 ± 22	-8.3	5.9
FG	10 ± 1	3.2	10.1	147 ± 12	-1.7	8.2	395 ± 25	-1.3	6.2
TG	10.6 ± 0.4	5.9	4.1	153 ± 6	2.0	4.2	358 ± 46	-10.6	13.0
2MBG	10.9 ± 0.8	8.5	7.6	152 ± 1	1.2	0.8	407 ± 14	1.8	3.4
3MCG	10.8 ± 0.2	7.7	2.0	150 ± 2.0	-0.3	1.1	403 ± 16	0.7	4.1
IVG	10 ± 1	-4.8	10.8	155 ± 16	3.4	10.3	395 ± 23	-1.3	5.9
VG	9.7 ± 0.2	-2.6	2.2	150.3 ± 0.5	0.2	0.4	405 ± 8	1.4	2.0
HG	11 ± 1	6.4	12.7	136 ± 12	-9.1	8.7	349 ± 30	-12.9	8.7
PAG	9.5 ± 0.3	-5.1	3.5	155 ± 12	3.1	7.8	367 ± 17	-8.3	4.7
PPG	10.2 ± 0.2	2.3	2.1	147 ± 7	-2.0	4.9	368 ± 11	-8.0	3.0
GG	10.5 ± 0.9	4.6	8.8	155 ± 3	3.1	1.7	400 ± 8	0.1	2.0
HpG	10 ± 1	-1.9	12.3	147 ± 4	-2.0	2.8	341 ± 13	-14.9	3.9
OG	10.5 ± 0.5	4.9	4.9	141 ± 5	-6.0	3.4	360 ± 24	-9.9	6.7
SG	9.8 ± 0.8	-2.1	8.0	148 ± 3	-1.1	1.9	402 ± 8	0.5	1.9

Table 3-12 Stability results of various storage conditions

Acyl glycine	Room Temp 5 hours		Auto sampler 24 hours		Freeze/Thaw Cycle 1		Freeze/Thaw Cycle 2		Freeze/Thaw Cycle 3		-20 °C 2 weeks		-20 °C 8 weeks	
	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)
AG	98.1	4.7	96.0	3.7	97.0	8.8	96.3	4.1	98.3	5.7	94.1	10.1	94.0	3.1
PG	99.7	9.7	98.1	10.4	97.7	9.6	97.2	11.0	100.6	9.4	95.0	10.7	103.0	11.8
IBG	93.5	7.4	97.2	6.3	93.5	6.8	100.1	8.9	94.8	8.4	91.2	9.7	95.2	8.1
BG	105.1	12.7	104.9	6.6	95.8	8.9	93.9	8.7	104.2	5.5	99.9	11.1	105.1	7.8
HPAH	90.4	4.4	98.3	6.7	99.0	5.1	95.9	10.4	102.2	4.6	101.3	6.6	95.1	8.7
FG	104.7	4.1	107.3	9.1	108.0	5.6	104.5	3.3	101.1	5.7	98.5	8.4	103.6	3.4
TG	95.3	3.4	100.0	3.2	96.8	7.8	98.2	5.5	100.4	5.9	98.8	10.7	92.6	6.1
2MBG	97.5	6.8	100.9	5.5	97.1	4.2	95.3	6.6	97.2	3.5	99.3	12.2	96.2	3.5
3MCG	102.8	6.5	99.1	6.3	101.8	7.2	99.0	5.9	100.6	5.8	95.9	9.4	104.2	6.8
IVG	97.4	2.2	98.8	6.5	101.5	6.2	98.3	2.9	98.0	4.8	95.7	10.6	98.4	2.8
VG	104.5	4.4	102.5	2.2	104.9	4.0	104.6	2.8	101.5	2.5	96.2	8.4	102.9	4.1
HG	99.9	6.1	98.5	4.3	98.2	4.4	96.7	4.3	94.7	4.5	93.2	9.2	99.2	5.1
PAG	95.7	11.3	95.4	10.3	93.7	10.5	97.5	10.2	99.6	13.1	99.8	13.3	96.2	11.4
PPG	97.2	7.8	96.4	5.8	95.4	7.9	97.2	5.6	100.5	5.8	95.7	9.8	98.5	5.9
GG	95.3	5.2	95.6	6.0	95.1	9.2	96.5	6.3	94.3	5.6	98.2	5.9	95.8	5.1
HpG	94.5	6.0	92.5	5.9	95.1	7.5	94.2	6.3	95.0	6.4	91.9	6.4	94.9	5.5
OG	99.4	5.3	100.0	4.3	100.9	5.7	102.1	5.2	102.3	5.0	106.0	7.4	106.4	4.6
SG	102.5	9.4	106.4	7.9	96.0	6.8	100.9	6.8	102.3	8.4	94.1	9.9	103.7	3.7

### **3.3.7 Measurement of Endogenous Levels of Acylglycines**

The validated LC-MS/MS method was then used in the analysis of urine samples from 20 healthy individuals, collected over 3 consecutive days (total of 60 samples). Each sample was derivatized and analyzed in triplicates. Concentrations were back-calculated using the peak area ratios of analyte/IS from a calibration curve constructed in surrogate urine.

#### **3.3.7.1 Absolute Quantification**

The ranges of concentrations of each acylglycine over three days are shown in Table 3-13. The concentrations are represented in  $\mu\text{mol/mol}$  creatinine. The values obtained in this study are comparable with published data.<sup>8, 10</sup> An expanded table listing the daily concentrations of each of the 20 individual can be found in Table S3-1 in the Electronic Appendix. For most of the individuals, day-to-day variations in acylglycine concentrations were relatively small. The identities of the eighteen analytes were confirmed by the co-elution of the internal standard and the interpretation of the product ion spectra.

Table 3-13 Concentration ranges of acylglycines in healthy urine

Acylglycines	Concentration ( $\mu\text{mol/mol}$ creatinine)		
	Day 1	Day 2	Day 3
AG	597 - 9789	494 - 1848	418 - 3466
PG	86 - 378	59 - 402	61 - 630
IBG	216 - 1622	174 - 1249	164 - 839
BG	8.1 - 135	7 - 170	12 - 237
HPAG	< 2.5 - 7.9	< 2.5 - 10	< 2.5 - 5
FG	154 - 2183	81 - 3279	168 - 5192
TG	569 - 5034	560 - 3519	470 - 3190
2MBG	55 - 262	35 - 364	64 - 225
3MCG	56 - 646	40 - 624	65 - 444
IVG	593 - 7151	542 - 5656	477 - 4772
VG	7 - 215	8 - 147	12 - 90
HG	21 - 199	27 - 375	34 - 198
PAG	36 - 136	39 - 170	37 - 131
PPG	< 1.0 - 66	< 1.0 - 66	2 - 57
GG	4 - 57	3 - 52	1 - 59
HpG	< 2.0 - 16	< 2.0 - 9	< 2.0 - 7
OG	< 1.5 - 35	< 1.5 - 41	< 1.5 - 49
SG	8 - 68	4 - 52	< 2.0 - 39

### 3.3.7.2 Acylglycine Excretion Pattern

The acylglycine profile of the individuals was investigated to determine if there was any dependence of the acylglycine excretion pattern on parameters such as gender or BMI. The influence of gender on the acylglycine pattern is seen in Figure 3-10. The box plots were generated using the full range of values. Fifteen acylglycines were plotted (only acylglycines with values greater than LLOQ were plotted) and each value plotted is an average of three days. There are differences between males and females but there does not seem to be a consistent pattern. There is no discernible pattern between short- and medium-chain acylglycines. Thus, it appears that gender does not play a role in the excretion pattern of acylglycines in urine.

Similar results were observed for the dependence of the profile on BMI. BMI was calculated by dividing the individual body mass in kilograms by the square of his/her height in meters and the units are expressed in  $\text{kg/m}^2$ . The BMIs were divided into four categories, according to the Canadian guidelines for body weight classification: (a) underweight – BMI less than 18.5; (b) normal weight – BMI between 18.5 and 24.9; (c) overweight – BMI between 25 and 29.9 and (d) obese – BMI of 30 and more. The number of individuals that fit into each category was n=1 (underweight), n=15 (normal weight), n=2 (overweight) and n=2 (obese). The mean values of BMI for males and females were very similar, 23.5 in males and 22.2 in females. The influence of BMI can be seen in the bar charts in Figure 3-11.

Values plotted are an average of acylglycine values over the three days. Acylglycines are arranged based on similar concentration ranges. As can be observed, the available data does not indicate any influence of BMI on the excretion pattern. Unfortunately, due to the small number of samples investigated, the interpretation of the data remains limited.

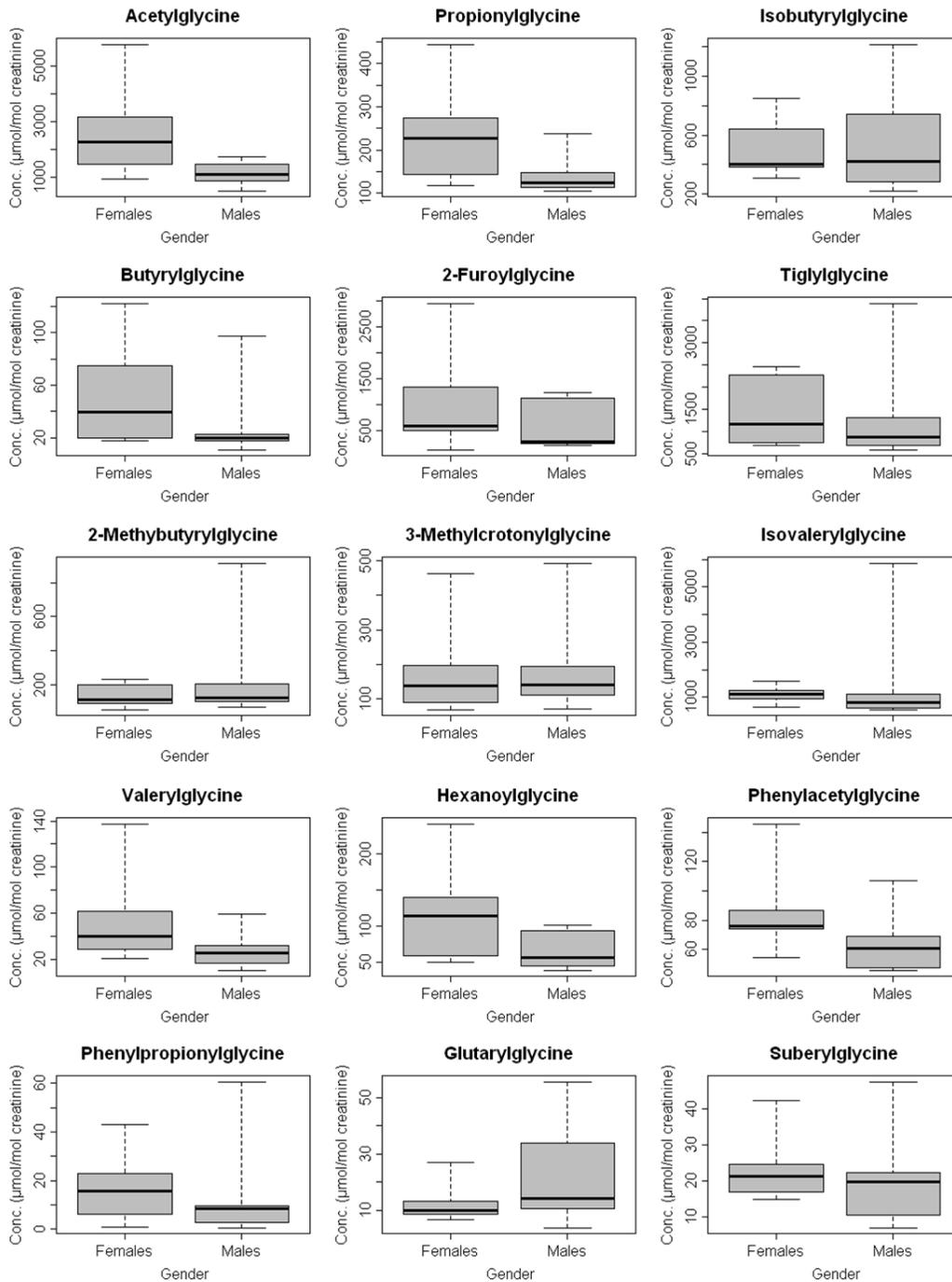


Figure 3-10 Influence on gender on the acylglycine excretion pattern (n=10 for both males and females)

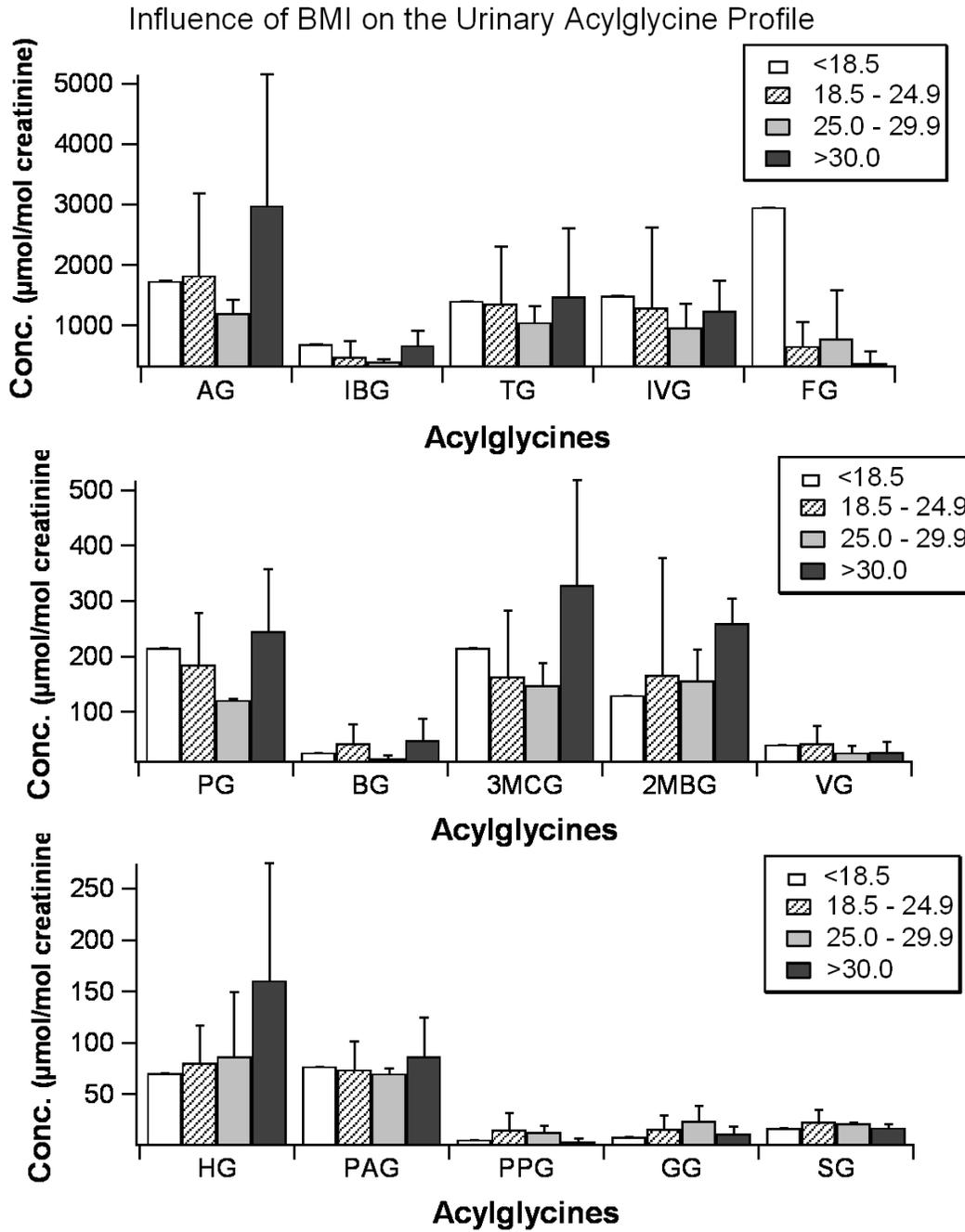


Figure 3-11 Influence of BMI on the acylglycine excretion profile

### 3.3.7.3 Relative Quantification

In the relative quantification of the derivatized acylglycines, for which there were no standards, putative identification was done by interpretation of the product ion spectra generated. In the spectra of the unknown acylglycine the following factors were investigated: (a) Fragments corresponding to the tag, (b) Fragments corresponding to the dissociation of the acylglycine, i.e. neutral losses of 236, 207, and 179, (c) fragment ion  $m/z$  237 and (d) logical retention time. The fragmentation of the acylglycine portion of the molecule resulted in low abundance fragments that are unobservable. It is believed that there are many more acylglycines present in the samples based on a previous study,<sup>7</sup> but with low abundance fragment ions it was difficult and often impossible to putatively identify these compounds. Consequently, only six compounds could be putatively identified. These are C<sub>5</sub>:1-G (pentenoylglycine or isomer), retention time, 12.47 minutes; C<sub>6</sub>:1-G (hexanoylglycine or isomer), retention time, 16.50 minutes; C<sub>6</sub>-G (methyl valerylglycine or isomer), retention time, 19.07 minutes; C<sub>8</sub>:1-G (cis-3,4-methylene heptanoylglycine or isomer), retention time, 23.89 minutes; C<sub>8</sub>:1-G (octenoylglycine or isomer), retention time, 24.22 minutes; and C<sub>9</sub>:1-G (cis-3,4-methylene octanoylglycine or isomer), retention time, 25.94 minutes. The ranges of concentrations for these acylglycines over three days are shown in Table 3-14. The concentrations are represented in  $\mu\text{mol/mol}$  creatinine. An expanded table presenting the daily concentrations of each of the 20 individuals can be found in Table S3-2 in the Electronic Appendix. The concentration for each day is shown in the columns and is represented as the mean  $\pm$  standard deviation of the three experimental replicates. The concentrations are represented in  $\mu\text{mol/mol}$  creatinine. EPI product ion spectra used in the putative identification of each acylglycine is represented in Figure A3-2 – Figure A3-4 in the Appendix.

Table 3-14 Concentration ranges for putatively identified acylglycines in healthy urine

Acylglycines	RT (min)	Concentration ( $\mu\text{mol/mol creatinine}$ )		
		Day 1	Day 2	Day 3
C <sub>5</sub> :1-G	12.47	13 - 647	9 - 652	23 - 525
C <sub>6</sub> :1-G	16.50	4 - 94	4 - 48	6 - 51
C <sub>6</sub> -G	19.07	18 - 317	13 - 134	22 - 118
C <sub>8</sub> :1-G	23.89	104 - 792	78 - 650	113 - 532
C <sub>8</sub> :1-G (2)	24.22	4 - 121	4 - 83	7 - 67
C <sub>9</sub> :1	25.94	< 1.5 - 35	< 1.5 - 71	< 1.5 - 62

### 3.4 Conclusion

This chapter describes a sensitive LC-MS/MS method for the quantification of derivatized acylglycines in human urine. The method was validated using underivatized urine as the surrogate matrix and proved to be reliable in terms of selectivity, linearity, accuracy and precision. The complications associated with performing validation for endogenous compounds have been satisfactorily addressed and method performance was evaluated based on a fit-for-purpose approach. The surrogate matrix calibration method was successfully applied and demonstrated an advantage over the use of solvent (water) as a suitable surrogate matrix. Parallelism experiments were used during validation to evaluate any bias originating from the differences between authentic and surrogate matrices. The typical acylglycine excretion profile shows no dependence on BMI or gender. Except for a few individuals, day-to-day variations are relatively minor. It is important to note however, that no biological significance can be drawn from such a limited number of samples.

Application of this method offers many merits over previously reported methods, namely a more accurate method validation approach, potentially lower

limits of detection, signal enhancement and the ability to create internal standards for each standard obtained. This method can be adopted for the accurate quantification of many endogenous biomarkers that can be readily derivatized, with little or no modifications. Although this chapter mainly focuses on the methodology of quantifying acylglycines, this method can easily be applied to clinical samples where the concentrations of the analytes are expected to increase for IEMs and similar disorders.

### 3.5 Literature Cited

- (1) Pasquali, M.; Monsen, G.; Richardson, L.; Alston, M.; Longo, N. *Am. J. Med. Genet. C* **2006**, *142C*, 64-76.
- (2) Kompare, M.; Rizzo, W. B. *Semin. Pediatr. Neurol.* **2008**, *15*, 140-149.
- (3) Waddell, L.; Wiley, V.; Carpenter, K.; Bennetts, B.; Angel, L.; Andresen, B. S.; Wilcken, B. *Mol. Genet. Metab.* **2006**, *87*, 32-39.
- (4) Costa, C. G.; Guerand, W. S.; Struys, E. A.; Holwerda, U.; ten Brink, H. J.; de Almeida, I. T.; Duran, M.; Jakobs, C. *J. Pharm. Biomed. Anal.* **2000**, *21*, 1215-1224.
- (5) Hagen, T.; Korson, M. S.; Sakamoto, M.; Evans, J. E. *Clin. Chim. Acta* **1999**, *283*, 77-88.
- (6) Bonafe, L.; Troxler, H.; Kuster, T.; Heizmann, C. W.; Chamoles, N. A.; Burlina, A. B.; Blau, N. *Mol. Genet. Metab.* **2000**, *69*, 302-311.
- (7) Lewis-Stanislaus, A. E.; Li, L. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 2105-2116.
- (8) Ombrone, D.; Salvatore, F.; Ruoppolo, M. *Anal. Biochem.* **2011**, *417*, 122-128.
- (9) Ito, T.; Kidouchi, K.; Sugiyama, N.; Kajita, M.; Chiba, T.; Niwa, T.; Wada, Y. *J. Chromatogr., B: Biomed. Appl.* **1995**, *670*, 317-322.
- (10) Fong, B. M. W.; Tam, S.; Leung, K. S. Y. *Talanta* **2011**, *88*, 193-200.
- (11) Bodoprost, J.; Rosemeyer, H. *Int. J. Mol. Sci.* **2007**, *8*, 1111-1124.

- (12) Obert, J. C.; Hughes, D.; Sorenson, W. R.; McCann, M.; Ridley, W. P. *J. Agric. Food Chem.* **2007**, *55*, 2062-2067.
- (13) Francois, I.; Sandra, P. *J. Chromatogr. A* **2009**, *1216*, 4005-4012.
- (14) Guo, K.; Li, L. *Anal. Chem.* **2010**, *82*, 8789-8793.
- (15) van de Merbel, N. C. *Trac-Trends in Anal. Chem.* **2008**, *27*, 924-933.
- (16) Li, W. L.; Cohen, L. H. *Anal. Chem.* **2003**, *75*, 5854-5859.
- (17) Jemal, M.; Schuster, A.; Whigan, D. B. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1723-1734.
- (18) Shi, J. X.; Liu, H. F.; Wong, J. M.; Huang, R. N.; Jones, E.; Carlson, T. J. *J. Pharm. Biomed. Anal.* **2011**, *56*, 778-784.
- (19) Kindt, E.; Shum, Y.; Badura, L.; Snyder, P. J.; Brant, A.; Fountain, S.; Szekely-Klepser, G. *Anal. Chem.* **2004**, *76*, 4901-4908.
- (20) Zhang, Y. H.; Dufield, D.; Klover, J.; Li, W. L.; Szekely-Klepser, G.; Lepsy, C.; Sadagopan, N. *J. Chromatogr. B* **2009**, *877*, 513-520.
- (21) Gilibili, R. R.; Kandaswamy, M.; Sharma, K.; Giri, S.; Rajagopal, S.; Mullangi, R. *Biomed. Chromatogr.* **2011**, *25*, 1352-1359.
- (22) Wilson, S. F.; James, C. A.; Zhu, X. C.; Davis, M. T.; Rose, M. J. *J. Pharm. Biomed. Anal.* **2011**, *56*, 315-323.
- (23) Houghton, R.; Pita, C. H.; Ward, I.; Macarthur, R. *Bioanalysis* **2009**, *1*, 1365-1374.
- (24) Zhang, G. D.; Zhang, Y. Z.; Fast, D. M.; Lin, Z. S.; Steenwyk, R. *Anal. Biochem.* **2011**, *416*, 45-52.
- (25) Li, H.; Rose, M. J.; Tran, L.; Zhang, J.; Miranda, L. P.; James, C. A.; Sasu, B. J. *J. Pharm. Toxicol. Methods* **2009**, *59*, 171-180.
- (26) Flaherty, J. M.; Connolly, P. D.; Decker, E. R.; Kennedy, S. M.; Ellefson, M. E.; Reagen, W. K.; Szostek, B. *J. Chromatogr. B* **2005**, *819*, 329-338.
- (27) Zar, J. H. In *Biostatistical Analysis*, 5 ed.; Snavey, S. L., Ed.; Pearson Prentice Hall Inc.: Upper Saddle River, 2010, pp 363-378.
- (28) ICH, Bioanalytical Method Validation, 1996.
- (29) Ermer, J. *J. Pharm. Biomed. Anal.* **2001**, *24*, 755-767.

- (30) Mukherjee, P. S.; Karnes, H. T. *Biomed. Chromatogr.* **1996**, *10*, 193-204.
- (31) Hewavitharana, A. K. *Crit. Rev. Anal. Chem.* **2009**, *39*, 272-275.
- (32) Beltrán, E.; Ibáñez, M.; Sancho, J. V.; Hernández, F. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1801-1809.
- (33) Villagrasa, M.; Guillamon, M.; Eljarrat, E.; Barcelo, D. *J. Chromatogr. A* **2007**, *1157*, 108-114.

## **Chapter 4. Quantification of Acylglycines in Human Plasma using Derivatization and Liquid Chromatography Tandem Mass Spectrometry\***

### **4.1 Introduction**

Acylglycines have long been studied as biomarkers in inborn errors of metabolism (IEM) in newborns and are used in the diagnoses of these disorders.<sup>1, 2</sup> IEMs are genetic diseases or disorders of the metabolic system. These disorders are characterized by a defect in a single enzyme, resulting in reduced catalytic activity, or by impaired activity of transporters or cofactors along metabolic pathways. This can result in an accumulation of metabolites that can become toxic or interfere with the normal function of systems. One class of IEMs is fatty acid disorders, which is the results of enzyme defects affecting the body's ability to oxidize fatty acids to produce energy. Fatty acid oxidation occurs in the mitochondria of cells and is a series of four reactions, each catalyzed by a different enzyme, depending on the length of the fatty acid chain.<sup>1</sup> Deficiencies of any of these enzymes results in different disorders named for the deficient enzyme, for e.g. medium-chain acyl-coenzyme A dehydrogenase deficiency. In individuals where there is an impairment of fatty acid oxidation, fatty acid CoA esters may build up when there is an increase in energy requirements or fasting and must be conjugated to carnitine or glycine to facilitate excretion or a balance of coenzyme A. Glycine conjugation then serves an important detoxification role and the products are known as acylglycines.

Acylglycines are excreted by healthy individuals as well as those with IEMs and they can be detected in biological matrices (for e.g. blood, plasma, serum, urine) but traditionally have been analyzed in urine.<sup>3-5</sup> This is because

---

\* A version of this chapter has been prepared for submission as Stanislaus, A.E., Li, L., Development of a Method for the Quantitation of Acylglycines in Human Plasma using Derivatization and Liquid Chromatography Tandem Mass Spectrometry.

urine is generally easier to obtain and is relatively free of proteinaceous material, making sample handling and analysis less complicated. In recent times, dried blood spots have become more popular in newborn screening for IEMs.<sup>6-9</sup> Because of their lower levels, acylglycines are not generally analyzed in dried blood spots or plasma. Very few papers are published with reference values for acylglycines in plasma.<sup>10</sup>

This chapter focuses on increasing the electrospray signal of acylglycines in biofluids so they can be analyzed in plasma by LC-MS. The derivatizing tag used in this work was used previously in Chapter 3 and targets the carboxylic acid moiety in acylglycines. The tag, p-dimethylaminophenacyl (DmPA) bromide, is a derivative of phenacyl bromide, in which a dimethyl amino group is added in the para-position. There are several advantages of using this labeling technique: first, the amino group is easily ionizable, increasing ionization efficiencies; second, an isotope of carbon can be coded onto the dimethyl group creating an isotope-labeled standard for each analyte purchased; third, the increase in hydrophobicity of the labeled acylglycines allows them to be less solvated and increases their surface activity and lastly, the retention time increases in reversed-phase chromatography and causes an improvement in the separation.

In this work, quantification was performed using mass spectrometry coupled with liquid chromatography. The quantification is based on the comparison of the signal generated from the analytes of interest in matrix versus the signal from standards in the blank matrix. Isotope dilution, which involves the use of stable-isotope labeled standards, is a widely accepted method for quantification using mass spectrometry. The method involves the preparation of a series of standards over a range of concentrations which are added to the blank matrix and a known volume of the isotope-labeled analyte is added to the calibration standards, quality control standards and unknown samples. This internal standard serves to compensate for sample preparation and instrument variability. When the analytes of interest exist naturally in the matrix and exhibit high background levels, it is often difficult to obtain a blank matrix. A new

approach, a “surrogate matrix” strategy,<sup>11, 12</sup> has been developed and has been previously applied to the quantification of endogenous acylglycines in urine (Chapter 3). The fundamental concept of this approach is to generate a calibration curve in a “blank surrogate” matrix, which in this study is the underivatized plasma, using peak area ratios of standard to internal standard. The concentration of the analytes in authentic matrix, that is the derivatized plasma, is calculated based on the regression equations from the calibration curves. The belief is that the underivatized plasma would be a better substitute and would better mimic the authentic matrix than water or synthetic matrix. Compared to current methods that use water as a surrogate matrix, this approach can offer matrix-matching, which improves accuracy. An LC-MS/MS method for the quantification of endogenous acylglycines in human plasma is presented.

## **4.2 Experimental**

### **4.2.1 Chemical and Reagents**

Optima grade acetonitrile, methanol and water were purchased from Fisher Scientific (Ottawa, ON). Triethylamine and formic acid were obtained from Sigma-Aldrich (Oakville, ON). Acetylglycine, propionylglycine, isobutyrylglycine, butyrylglycine, 2-methylbutyrylglycine, isovalerylglycine, valerylglycine, hexanoylglycine, heptanoylglycine, tiglylglycine, glutarylglycine and suberylglycine were purchased from Dr. Herman J. ten Brink (Amsterdam, Netherlands). 3-Methylcrotonylglycine, 2-furoylglycine, phenylacetylglucine, phenylpropionylglycine, and 4-hydroxyphenylacetylglucine standards were obtained from HMBD. Octanoylglycine was purchased from Sigma-Aldrich. 2-Bromo-1-(4-dimethylaminophenyl)ethanone was purchased from Combi-Blocks Inc. (San Diego, CA) and also synthesized in our laboratory, according to the method described in Chapter 3, Section 3.2.2.

## 4.2.2 Plasma Samples

This study was conducted in accordance with the Arts, Science & Law Research Ethics Board policy at the University of Alberta. Blood samples were collected from 5 males and 5 females, who were not known to be on any special diet or medications or have any inborn errors of metabolism. The volunteers were all adults ranging in age from 25 to 35 years. The plasma was generated by centrifuging whole blood at 4000 g at 4 °C for 20 minutes. The plasma supernatant was collected and aliquoted and stored at -80 °C until further analysis.

## 4.2.3 Preparation and Derivatization of Standards and Samples

### 4.2.4.1 Standard Solutions

Standard stock solutions containing a mixture of all eighteen acylglycines were prepared by dissolving accurately weighed acylglycines in acetonitrile to yield a final concentration of 1.0 mM. A working stock solution of 100 µM was made in 20% water in acetonitrile. To 100 µL of the acylglycine mixture was added 100 µL of either  $^{12}\text{C}_2$ -DmPA or  $^{13}\text{C}_2$ -DmPA (30 mg/mL) and 100 µL TEA (30 mg/mL) in acetonitrile. The mixture was vortexed and heated in a heating block at 90 °C for 1 hour. The reaction was then terminated with 100 µL acetic acid dissolved in acetonitrile (10% v/v) and reacted at 90 °C for 15 minutes. The reaction scheme is presented in Figure 3-1. The resulting mixture was evaporated to dryness and reconstituted in 100 µL acetonitrile/water (50/50, v/v), 0.1% formic acid. The standards were then spiked into solvent and surrogate plasma (calibration standards) and authentic plasma (quality control standards) and analyzed by LC-MS/MS. Internal standard (IS), that is  $^{13}\text{C}_2$ -DmPA-labeled standards, were spiked into calibration standards and samples to give a final concentration of 100 nM before analysis by LC-MS/MS.

#### 4.2.4.2 Plasma samples

Frozen plasma samples were thawed at 4 °C to prevent deterioration of the sample. For protein precipitation, 400 µL methanol was added to 100 µL plasma. The plasma mixture was vigorously vortexed for 10 seconds and incubated at 4 °C for 30 minutes. The mixture was then vortexed again for another 10 seconds and centrifuged at 14000 g for 15 minutes at 4 °C. The supernatant was decanted, evaporated and the residue was reconstituted in 100 µL of 20% water in acetonitrile. The extracted plasma was then derivatized in an identical method as the standard solutions. To 100 µL plasma extract was added 100 µL of either <sup>12</sup>C<sub>2</sub>-DmPA (30 mg/mL) and 100 µL TEA (30 mg/mL) in acetonitrile. The mixture was vortexed and heated in a heating block at 90 °C for 1 hour. The reaction was then terminated with 100 µL acetic acid dissolved in acetonitrile (10% v/v) and reacted at 90 °C for 15 minutes. The resulting mixture was evaporated to dryness and reconstituted in 100 µL acetonitrile/water (50/50, v/v), 0.1% formic acid.

#### 4.2.4 Ultra-high Performance Liquid Chromatography

Liquid chromatography was performed on an Agilent 1290 Series LC (Mississauga, ON) equipped with dual pumps, an eluting pump and a regeneration pump. Chromatographic separation was done on two Phenomenex Kinetex 1.7 µm minibore 50 x 2.1 mm C<sub>18</sub> columns. Elution conditions were as follows: linear gradient of 18% - 28% mobile phase B over 17 minutes, 28% - 43% B over 8 minutes, and 43% - 100% B over 15 minutes and returning initial conditions for 2 minutes. Mobile phase A consisted of 2% acetonitrile, 0.1% formic acid in water and mobile phase B consisted of 2% water, 0.1% formic acid in acetonitrile. The column was maintained at 25 °C and autosampler set at 4 °C. Flow rate was set at 0.250 mL/min and 5.0 µL of each sample was injected. The regeneration pump performed column wash and equilibration steps in parallel on the second C<sub>18</sub> column using 100% mobile phase B and 18% mobile phase B respectively.

#### **4.2.5 Mass Spectrometry**

Mass analysis was carried out on an AB Sciex 4000 QTRAP<sup>®</sup> hybrid triple quadrupole linear ion trap mass spectrometer (Concord, ON), equipped with an electrospray ionization (ESI) interface. The ESI source was set to perform in the positive ion mode. The optimized parameters were as follows: spray voltage, 4800; curtain gas, 10; CAD, high; temperature, 250 °C; GS1, 40; GS2, 30; declustering potential, 45; collision energy, 25 (parameters are unit-less). The acquisition method consisted of an information dependent acquisition (IDA) scan using multiple reaction monitoring (MRM) as a survey scan and two dependent enhanced product ion (EPI) scans. The EPI scans were in the range 50 – 800 Da, scanned at 4000 Da/s. Data was acquired and processed using Analyst<sup>®</sup> version 1.5.1 software (Concord, ON).

#### **4.2.6 Method Validation**

The method was validated for the following: selectivity, linearity of IS and analytes, matrix effects, extraction efficiency, stability, accuracy and precision. All required statistical analyses were performed using R script 2.11.1 and Igor Pro 6.01.

##### **4.2.7.1 Selectivity**

Method selectivity was evaluated by analyzing an instrument blank, a solvent blank, surrogate plasma, surrogate plasma containing only IS at 100 nM and surrogate plasma containing analytes at 1.0 nM. These samples were all analyzed using the LC-MS/MS method described above and were evaluated for the presence of any interfering signals in the MRM channels of the analytes and IS.

#### 4.2.7.2 Linearity

The linearity of the ESI response for the DmPA-labeled acylglycines was examined by analyzing 5 replicates at 7 different concentration levels. Calibration standards were prepared by serial dilutions of a mixture of eighteen  $^{12}\text{C}_2$ -DmPA-labeled acylglycines (10  $\mu\text{M}$ ) to obtain concentrations of 1, 5, 10, 50, 100, 500, 1000 nM. Due to the endogenous presence of acylglycines in the derivatized plasma, underivatized pooled plasma (referred to as surrogate plasma) was used to prepare the calibration standards. Pooled plasma was prepared by mixing equal volumes of all the plasma samples. A fixed volume of  $^{13}\text{C}_2$ -DmPA-labeled acylglycines (final concentration 100 nM) was added to the surrogate plasma as an internal standard. Calibration curves were constructed by plotting peak area ratios (analyte peak area/IS peak area) against concentration of each acylglycine and were fitted to linear regression analysis with a weighting of  $1/\text{response}$  ( $1/y$ ), after evaluation of different weighing factors. Determination of the limit of detection (LOD) and the lower limit of quantification (LLOQ) for each acylglycine was done using the equations  $3 \times (\sigma_B)/m$  and  $10 \times (\sigma_B)/m$  respectively, where  $\sigma_B$  is the standard deviation of the response and  $m$  is the slope derived from the calibration curve. The standard deviation of the blank response was estimated from the standard deviation of the y-intercept, obtained from the regression analysis. The LLOQ of each analyte was confirmed experimentally by analyzing five replicates of surrogate matrix spiked with the analytes at a concentration equal to the LLOQ.

#### 4.2.7.3 Extraction Procedure

The extraction recovery of native acylglycines from plasma was determined by spiking in known amounts of deuterated acylglycine standards into surrogate plasma. The three deuterated standards used were isobutyrylglycine- $\text{d}_2$ , hexanoylglycine- $\text{d}_2$  and suberylglycine- $\text{d}_2$  at two concentrations, 25 nM and 400 nM. Six different samples of pooled plasma were thawed; three plasma samples

were spiked with deuterated standards, final concentration of 25 nM and three plasma samples were spiked with deuterated standards, final concentration 400 nM. Protein precipitation and derivatization were performed as described above for the plasma samples. IS was added in the same manner and the replicate samples were analyzed by LC-MS/MS. Extraction efficiency (as percent recovery) was assessed by comparing the peak area ratios of the extracted standards with that of deuterated standards of the same concentration spiked in extracted plasma.

#### **4.2.7.4 Authentication of Surrogate Plasma**

Two sets of calibration standards were prepared as described above in underivatized pooled plasma and derivatized pooled plasma. The slopes of the calibration curves obtained in each matrix were compared using a modified t-test<sup>13</sup> at the 95% confidence level using Equation 3-1 (Chapter 3).

Another way of evaluating the suitability of underivatized plasma as a surrogate matrix is to compare the concentration of acylglycines in pooled plasma derived from the calibrations curves in surrogate plasma and authentic plasma. Using the surrogate calibration curve, quantification was performed using the derived calibration regression equations. Using the calibration curves in authentic plasma (standard addition experiment), the concentration of each acylglycine in the pooled plasma was calculated by extrapolating to determine the magnitude of the x-intercept. This concentration was compared to the concentration of acylglycines derived from the surrogate plasma calibration curve. Percent relative error (% RE) was calculated using Equation 3-2:

$$\% \text{ RE} = \frac{\text{surrogate matrix concentration} - \text{standard addition concentration}}{\text{standard addition concentration}} \times 100$$

#### **4.2.7.5 Matrix Effects**

Matrix effects were assessed by comparing the slopes of the calibration curves prepared in solvent mix to those prepared in surrogate and authentic plasma. The modified t-test, at the 95% confidence level, was used to compare the slopes to evaluate if the slope differences were statistically significant and indicative of matrix effects. In addition ion suppression/enhancement was also calculated according to Equation 3-2 (Chapter 3).

#### **4.2.7.6 Accuracy and Precision**

To determine the accuracy of the method, four sets of pooled derivatized plasma were fortified with derivatized analytes at four different concentrations (three replicates per concentration level): 10 nM or 25 nM (low QC), 150 nM (medium QC) and 400 nM (high QC). The concentration of the low QC was determined as 3 x LLOQ for each analyte. The recoveries were calculated by subtracting the amount of endogenous acylglycine in an un-spiked sample from the amount found in fortified sample and comparing the value to the added amounts. Deviation from the added concentration or percent relative error (% RE) was calculated as a measure of accuracy.

Intra-day and inter-day precision were determined by analyzing a series of 10 injections of a derivatized pooled plasma sample over 3 consecutive days under the same LC-MS/MS conditions. Intra-day precision was evaluated for the 10 injections (total of 10 replicates per day) and inter-day precision was evaluated for a period of three days (total of 30 replicates). Precision was evaluated by calculating coefficients of variations (CV %) from the measured concentrations.

#### 4.2.7.7 Stability

Pooled derivatized plasma samples, fortified with derivatized analytes at a concentration of 25 nM, was used to study the stability of the analytes. One set of plasma samples were analyzed immediately after spiking (referred to as “fresh” samples) and at selected times over the storage period. Stability was assessed after each freeze-thaw cycle for three cycles, after 24 hours in the autosampler at 4 °C, and after 5 hours on a benchtop at room temperature. Freezer stability was assessed by analyzing the plasma sample after storage at -20 °C for two and eight weeks. The percent recovery of each analyte was determined by calculating the concentration at each condition and expressed as a percentage of the fresh plasma sample.

#### 4.2.7.8 Measurement of endogenous levels of acylglycines

Chromatographic data and peak integrations were performed using Analyst<sup>®</sup> version 1.5.1 software. Applying the surrogate matrix approach, QC concentrations and endogenous concentrations of acylglycines were determined using the regression equations derived from the calibration curves in surrogate plasma. The equation used in this study is shown below where m is the slope of the regression line and b is the intercept.

$$\frac{\text{Area } (^{12}\text{C}_2 - \text{DmPA analytes})}{\text{Area } (^{13}\text{C}_2 - \text{DmPA analytes})} = m [\text{Conc. } (^{12}\text{C}_2 - \text{DmPA analytes})] + b$$

Equation 4-1 Equation used for calibration curves

## **4.3 Results and Discussion**

### **4.3.1 Surrogate Matrix Strategy**

Assays based on electrospray ionization mass spectrometry (ESI-MS) can be affected by matrix-induced ion suppression when analyzing biological samples. Matrix effects occur when endogenous compounds co-eluting with the analyte of interest alters the ionization efficiency of the analyte. The competition between these matrix components and analytes may cause the decrease (ion suppression) or the increase (ion enhancement) in the efficiency of analyte ion formation. Thus the efficiency of analyte ion formation and evaporation depends on the matrix in which the analyte exists. It is important to take matrix effects into consideration when developing mass spectrometric assays for endogenous compounds because the matrix used in the preparation of standard curves and that of the unknown samples are sometimes different. A common practice to minimize matrix effects is to match the matrix of the calibration standards with that of the sample. In such cases, the matrix used is a “blank matrix” in which the analyte is absent or in very low concentration. In the case of endogenous analytes, no true blank matrix is available and an alternate matrix must be used. Also, it must be considered that it is not possible to match a matrix exactly because each sample matrix is different. Based on this, underivatized extracted pooled plasma was selected to be used as the surrogate matrix to perform the matrix-matched calibration.

### **4.3.2 Optimization of the LC-MS/MS Method**

Acylglycines, being relatively small and polar, are not easily separated using reversed-phase chromatography. Method development involves initial mobile phase conditions that are almost 100% aqueous to achieve some retention, but those conditions are not desirable for ESI. As expected from previous work, adding a hydrophobic tag to the acylglycines increased their retention on a

reversed-phase column and improved the separation of the analytes, especially separation of the isomers. The hydrophobic tag also enhances the ESI response of the labeled acylglycines compared to the native acylglycines. The optimized analysis time was 41 minutes and the retention times of the eighteen derivatized acylglycines and their abbreviations are represented in Table 4-1. The incorporation of a  $^{13}\text{C}$  stable-isotope labeling reagent does not display any isotope effect in the reversed-phase separation as shown in Figure 4-1. Perfect co-elution of the light and heavy labeled analyte and IS respectively are essential for accurate quantification by LC-MS.

For each acylglycine, the precursor ion  $[\text{M}+\text{H}]^+$  was used to trigger the MS/MS product ion spectra. The positive-ion ESI mass spectrum and MS/MS product ion spectrum of DmPA-labeled isovalerylglycine are shown in Figure 4-1. Explanations of the fragmentation patterns of the labeled acylglycines have been given in Chapter 3. Both analyte and IS yielded different product ions in the MRM channel resulting from the loss of a portion of the tag, preventing cross talk.

Table 4-1 Masses, abbreviations and retention times of DmPA-labeled acylglycines

<b>Acylglycine</b>	<b>Abbreviation</b>	<b>Protonated Mass</b>	<b>RT (min)</b>
Acetylglycine	AG	279	3.25
Propionylglycine	PG	293	4.73
Isobutyrylglycine	IBG	307	7.42
Butyrylglycine	BG	307	7.65
4-Hydroxyphenylacetylglycine	HPAG	371	8.02
2-Furoylglycine	FG	331	8.92
Tiglylglycine	TG	319	10.85
2-Methylbutyrylglycine	2MBG	321	11.32
3-Methylcrotonylglycine	3MCG	319	11.57
Isovalerylglycine	IVG	321	11.80
Valerylglycine	VG	321	12.80
Hexanoylglycine	HG	335	19.20
Phenylacetylglycine	PAG	355	20.17
Phenylpropionylglycine	PPG	369	20.42
Glutarylglycine	GG	512	22.57
Heptanoylglycine	HpG	349	23.41
Octanoylglycine	OG	363	26.36
Suberylglycine	SG	554	26.40

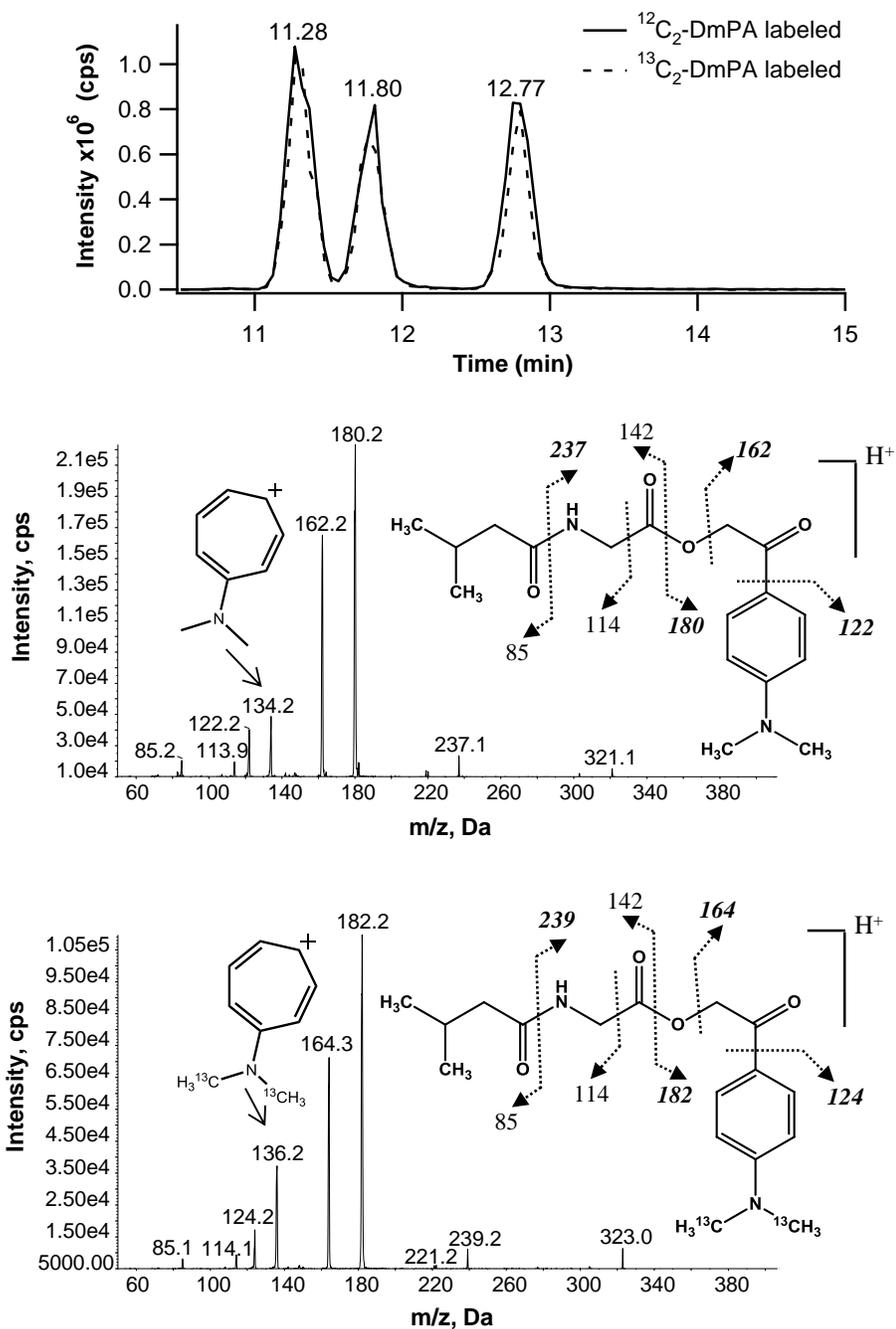


Figure 4-1 Positive-ion ESI-MS/MS product ion spectrum of DmPA-labeled-IVG: (A) Overlay of  $^{12}\text{C}_2$ -DmPA- and  $^{13}\text{C}_2$ -DmPA-labeled-IVG; (B) Product ion spectrum of  $^{12}\text{C}_2$ -DmPA-labeled-IVG; (C) Product ion spectrum of  $^{13}\text{C}_2$ -DmPA-labeled-IVG.

### 4.3.3 Method Validation

#### 4.3.3.1 Internal Standard Selection

Another approach to compensate for matrix effects is the use of internal standards to ensure accuracy and precision of the method. Using this procedure, a stable-isotope labeled (SIL) analog with identical chemical and physical properties is most commonly used. In choosing a SIL internal standard (SIL-IS), using a  $^{13}\text{C}$  label is preferred to using a  $^2\text{H}$  label because the latter may not co-elute sufficiently with the analyte to compensate for differences in matrix effects. For the IS to fully compensate for any ion suppression or enhancement present, it must perfectly co-elute with the analyte and have a similar response in the mass spectrometer. Thus, one of the first steps in developing a quantification method is the selection of a suitable SIL-IS. An appropriate concentration of the IS should be considered in order to ensure linearity of the calibration plot using peak area ratios. Choosing the concentration of the IS then becomes important because there is a chance that the linearity of the analyte is different from the IS and the chosen concentration of the IS might lie in a non-linear region of the curve. The results of the linearity studies of  $^{12}\text{C}$ -DmPA and  $^{13}\text{C}$ -DmPA-labeled acylglycines can be seen in Table 4-2. For each of the eighteen analytes, a statistical comparison of the slopes was then performed using the modified t-test and the calculated t-values compared to the critical values at the 95% confidence level. As is observed in Table 4-2, the t-values are all less than the critical values indicating that there are no significant differences between the two slopes. A concentration of 100 nM was chosen for the IS and linearity was ensured for concentrations lower and higher than that concentration.

Table 4-2 Internal Standard Suitability. Comparison of <sup>12</sup>C and <sup>13</sup>C linearity in solvent and surrogate matrix

Acylglycines	Calibration Curve in Solvent			Calibration Curve in Surrogate Matrix		
	<sup>12</sup> C Slopes	<sup>13</sup> C Slopes	t-value	<sup>12</sup> C Slopes	<sup>13</sup> C Slopes	t-value
AG	1.18E+07	1.18E+07	0.09	9.49E+06	9.22E+06	1.76
PG	1.59E+07	1.57E+07	0.11	1.44E+07	1.44E+07	0.15
IBG	2.32E+07	2.41E+07	0.89	2.80E+07	2.88E+07	1.90
BG	3.76E+07	3.80E+07	0.28	3.38E+07	3.35E+07	0.37
HPAG	1.54E+07	1.56E+07	0.42	1.36E+07	1.38E+07	1.67
FG	1.70E+07	1.70E+07	0.10	1.69E+07	1.68E+07	0.20
TG	1.48E+07	1.44E+07	2.00	1.30E+07	1.32E+07	1.82
2MBG	2.88E+07	2.84E+07	0.39	2.76E+07	2.72E+07	2.10
3MCG	1.27E+07	1.36E+07	1.92	1.44E+07	1.44E+07	0.14
IVG	1.90E+07	1.89E+07	0.12	2.02E+07	2.00E+07	1.06
VG	2.25E+07	2.21E+07	0.56	2.44E+07	2.40E+07	2.10
HG	2.25E+07	2.28E+07	0.60	2.42E+07	2.42E+07	0.02
PAG	1.98E+07	2.12E+07	1.66	2.03E+07	2.04E+07	0.35
PPG	2.15E+07	2.19E+07	0.65	1.97E+07	1.98E+07	0.12
GG	8.28E+06	8.63E+06	0.88	6.43E+06	6.39E+06	0.35
HpG	2.85E+07	2.89E+07	0.74	2.75E+07	2.70E+07	1.81
OG	2.10E+07	2.17E+07	1.72	1.81E+07	1.81E+07	0.50
SG	8.93E+06	8.35E+06	1.92	6.57E+06	6.59E+06	0.33

<sup>a</sup> t-values do not exceed critical values (critical value = 2.306, except for AG and PG which is 2.228)

#### 4.3.3.2 Selectivity

To determine the selectivity of the plasma method, several different chromatograms were evaluated. The degree of interference was assessed by examination of the MRM chromatograms of blanks and surrogate matrix. As can be observed in Figure 4-2, no interfering signals from the instrument or solvent blank or from the surrogate matrix were found in the MRM channels or at the retention times of the analytes. There were also no interfering signals from the plasma in the MRM channels and at the retention times of the internal standards.

#### 4.3.3.3 Calibration Curves, LOD and LLOQ

Calibration curves for all 18 analytes were obtained by analyzing seven standards, concentration range 1.0 – 1000 nM, in surrogate plasma. The slopes, intercepts, standard errors and coefficients of determination ( $R^2$ ) were all determined using the least-squares linear regression model. Different weighting factors were evaluated and the best linear fit and residuals of the model was achieved using a weighting factor of  $1/y$ , due to the heteroscedasticity of the data. Plots of 4 calibration curves and a statistical summary are shown in Figure A4-1 and Table A4-1 respectively in the Appendix. Plots of all 18 calibration curves are presented in Figure S4-1 in the Electronic Appendix. Calibration parameters are summarized and can be found in Table 4-3. The  $R^2$  values of the weighted calibration curves were all higher than 0.990 for each acylglycine. Overall CV % values, calculates as an average of the CV% at each level, were all less than 15%.

LOD and LLOQ values were calculated using the standard deviation of the y-intercept and the slope, derived from the linear regression model. The LOD ranges in value from 0.3 – 3.0 nM and LLOQ ranges from 1.0 – 10.0 nM. This information can also be found in Table 4-3. The % RE and % CV obtained from the replicate analysis at the LLOQ concentration of each acylglycine ranges from

-6.1 – 17.4% and 3.4 – 18.2% respectively, indicating good accuracy and precision. This information is summed up in Table 4-4.

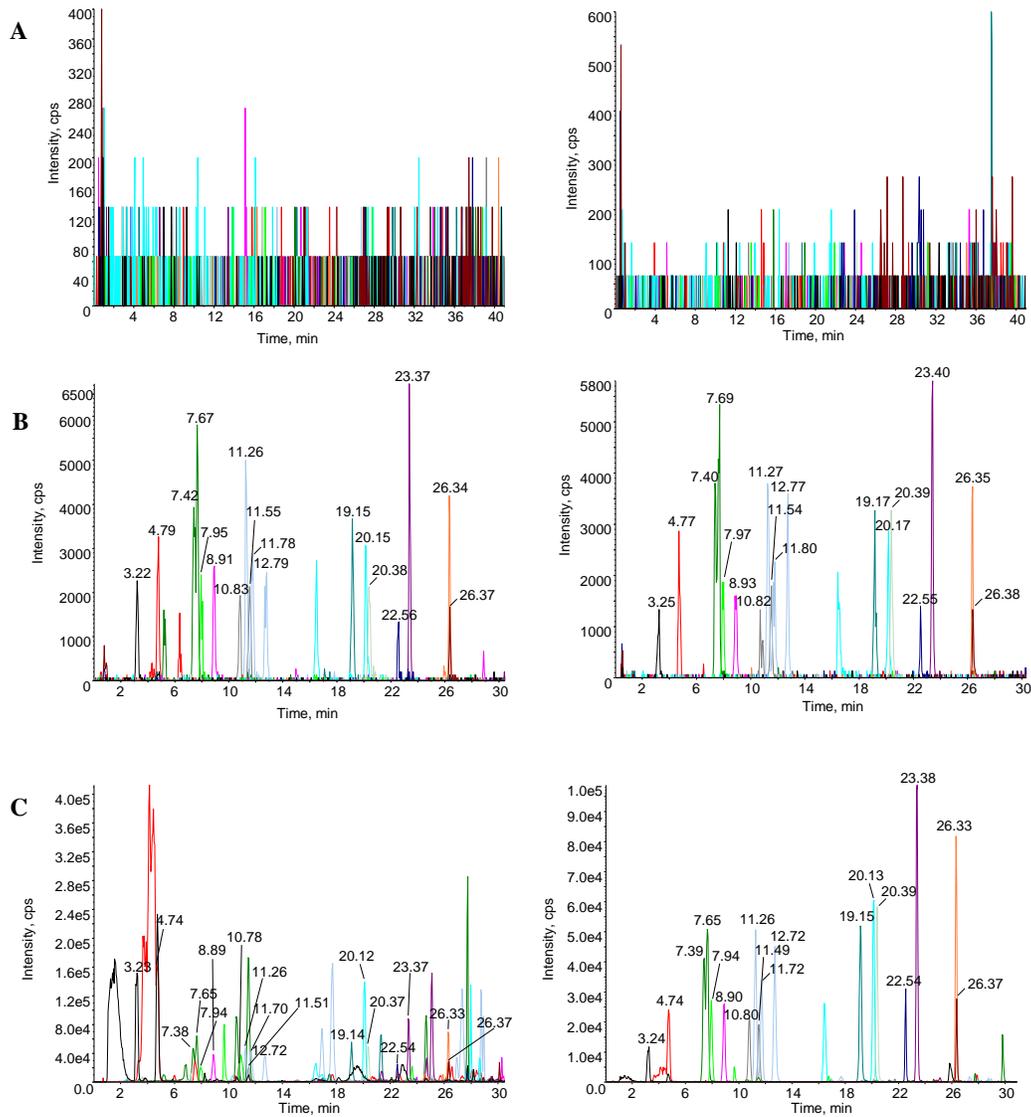


Figure 4-2 Selected reaction monitoring chromatograms of acylglycines (left panel) and corresponding internal standards (right panel). (A) surrogate plasma, (B) surrogate plasma spiked with 1.0 nM acylglycines and internal standards, (C) authentic plasma spiked with 100 nM internal standards and showing endogenous levels of acylglycines.

Table 4-3 Comparison of slopes in surrogate and authentic matrices and calibration parameters of calibration curve in surrogate plasma

Acylglycines	Slopes (Surrogate)	Slopes (Authentic)	t-value <sup>a</sup>	Linearity (R <sup>2</sup> )	LOD (nM)	LLOQ (nM)	Linear Dynamic Range	Overall CV % <sup>b</sup>
AG	9.0	8.9	1.19	0.9989	2	6	6 -1000	9.0
PG	8.2	8.2	1.33	0.9992	2	5	5 -1000	5.6
IBG	9.9	10.1	1.34	0.9974	2	7	7 - 500	11.7
BG	10.2	10.3	0.63	0.9994	1	3	3 - 500	9.6
HPAH	8.8	8.8	0.38	0.9984	2	5	5 - 500	8.4
FG	9.5	9.3	1.32	0.9988	1	4	4 - 500	8.9
TG	9.1	9.0	0.26	0.9944	3	10	10 - 500	10.7
2MBG	8.8	8.6	1.94	0.9995	1	3	3 - 500	8.9
3MCG	9.5	9.5	0.50	0.9989	1	4	4 - 500	9.0
IVG	8.9	8.9	0.28	0.9991	1	4	4 - 500	9.5
VG	9.1	8.8	1.32	0.9994	1	3	3 - 500	7.9
HG	9.28	9.18	2.08	0.9999	0.3	1	1 - 500	8.7
PAG	9.77	9.9	0.57	0.9998	1	2	2 - 500	8.6
PPG	9.0	8.91	1.11	0.9994	1	3	3- 500	11.4
GG	8.71	8.7	0.34	0.9996	1	2	2 - 500	9.6
HpG	9.1	9.1	0.48	0.9995	1	3	3 - 500	8.1
OG	8.6	8.5	0.86	0.9993	1	3	3 - 500	6.6
SG	10.4	10.3	1.29	0.9987	1	4	4 - 500	8.5

<sup>a</sup> t-values do not exceed critical values (critical value = 2.262; AG and PG, 2.201) ; <sup>b</sup> Overall CV % calculated as the average of all CV % at all concentration levels

Table 4-4 Lower limits of quantification (LLOQ) measurements of the proposed method

Acylglycines	LLOQ (nM)	Measured Concentration (nM)	RE %	% CV
AG	6.0	5.5	8.6	9.0
PG	5.0	5.3	-6.1	13.4
IBG	7.0	7.1	-1.4	7.9
BG	3.0	3.0	-0.3	9.5
HPAG	5.0	4.9	1.4	3.4
FG	4.0	3.7	6.5	9.3
TG	10.0	10.6	-5.8	9.4
2MBG	3.0	2.7	11.2	12.7
3MCG	4.0	3.9	2.4	9.3
IVG	4.0	4.0	0.5	13.5
VG	3.0	2.8	8.2	18.2
HG	1.0	1.0	0.2	15.5
PAG	2.0	1.7	17.4	12.9
PPG	3.0	2.9	4.1	16.8
GG	2.0	1.8	10.1	7.5
HpG	3.0	2.5	16.7	16.7
OG	3.0	2.7	9.9	3.7
SG	4.0	4.4	-9.3	12.5

#### 4.3.3.4 Authentication of Surrogate Matrix

In order to authenticate the use of surrogate plasma as a substitute for authentic plasma, calibration curves prepared in both surrogate and authentic plasma were compared using a modified t-test. Table 4-3 also summarizes the results of the comparison. These results show that the t values calculated were all less than the critical values of t at the 95% confidence level, indicating that the two slopes were not significantly different at that confidence level. This indicates

that the surrogate plasma can be used as a substitute for the authentic plasma and application of the calibration curve in surrogate plasma will yield accurate quantification. Another way of authenticating the use of surrogate plasma was by comparing the concentrations of acylglycines in a pooled plasma sample obtained using the calibration curve in surrogate plasma and the concentrations obtained using a standard addition experiment in the pooled plasma sample. The results are summarized in Table 4-5. For 10 acylglycines, the endogenous concentrations were less than the LLOQ; for the other 8 acylglycines, the % RE between the two calculated concentrations were all less than 15%, supporting the appropriate use of the surrogate plasma.

Table 4-5 Comparison of standard addition method and surrogate matrix approach

<b>Acylglycines</b>	<b>Standard Addition Conc. (nM)</b>	<b>Surrogate Calibration Conc. (nM)</b>	<b>RE %</b>
AG	1446 ± 30	1417 ± 51	-2.0
PG	568 ± 8	553 ± 55	-2.7
IBG	8.0 ± 0.9	<7.00	N/A
BG	8.3 ± 0.7	7.0 ± 0.8	-14.9
HPAG	< 5.00	< 5.00	N/A
FG	55 ± 2	49 ± 2	-11.3
TG	16.2 ± 0.9	15.3 ± 0.8	-5.4
2MBG	< 3.0	< 3.0	N/A
3MCG	< 4.0	< 4.0	N/A
IVG	69 ± 3	64 ± 1	-8.3
VG	< 3.0	< 3.0	N/A
HG	7.9 ± 0.3	7.9 ± 0.7	-0.01
PAG	137 ± 8	123 ± 4	-10.2
PPG	< 3.0	< 3.0	N/A
GG	< 2.0	< 2.0	N/A
HpG	< 3.0	< 3.0	N/A
OG	< 3.0	< 3.0	N/A
SG	< 4.0	< 4.0	N/A

#### 4.3.3.5 Evaluation of Matrix Effects

Matrix effects were evaluated using a comparison of slopes method, used in assays where no blank matrix is available.<sup>14</sup> A comparison of the slopes is made between matrix-free standard solutions and matrix or matrix-matched solutions and any significant deviation in the slopes indicate matrix effects of the sample matrix. The slopes of the calibration curves prepared in solvent were compared to the slopes in both surrogate and authentic plasma using the modified t-test defined above and the results are reported in Table 4-6, as well as calculations of matrix effects according to the equation defined in the experimental section. Negative values signify suppression, whereas positive values indicate enhancement. All analytes, with the exception of suberylglycine in surrogate plasma, experienced suppression, though the suppression was modest (less than 21%). The modified t-test demonstrated that for most of the analytes, the calculated t-values were higher than the critical values of t (at 95% confidence level) indicating that the slopes in solvent were significantly different than those in either of the plasma matrices. This indicates that matrix effects were observed in both plasma matrices. On the other hand, Table 4-3 shows that the slopes in surrogate and authentic matrices are not significantly different at the 95% confidence level. To minimize matrix effects and for more accurate validation and quantification, all calibrations and validations were done using either surrogate or authentic plasma.

Table 4-6 Matrix Effects study. Comparison of slopes in solvent to slopes in authentic plasma using a modified t-test

Acylglycines	Slopes in Solvent	Slopes in Surrogate	t-values <sup>a</sup>	Matrix Effect in Surrogate	Slopes in Authentic	t-values <sup>a</sup>	Matrix Effect in Authentic
AG	10.4	9.0	4.64	-12.95	8.9	5.66	-13.75
PG	10.3	8.2	14.78	-20.01	8.2	17.05	-20.77
IBG	10.24	9.9	2.66	-3.38	10.1	<b>0.73</b>	-0.93
BG	10.3	10.2	<b>0.52</b>	-1.20	10.3	<b>0.21</b>	-0.46
HPAG	10.1	8.8	8.24	-12.91	8.8	9.56	-12.44
FG	10.2	9.5	6.18	-6.58	9.3	8.16	-8.34
TG	10.0	9.1	3.89	-8.71	9.0	5.88	-9.27
2MBG	10.3	8.8	14.56	-14.28	8.6	16.23	-16.16
3MCG	10.2	9.5	4.00	-6.02	9.5	4.98	-6.56
IVG	10.1	8.9	5.16	-11.76	8.9	5.47	-12.16
VG	10.4	9.1	5.06	-12.49	8.8	5.81	-14.66
HG	9.9	9.28	12.35	-6.00	9.18	3.41	-7.00
PAG	10.2	9.77	3.10	-3.98	9.9	<b>1.32</b>	-2.71
PPG	10.4	9.0	8.54	-13.54	8.91	9.98	-14.38
GG	10.3	8.71	12.14	-15.14	8.7	11.82	-15.50
HpG	10.1	9.1	5.02	-9.64	9.1	5.41	-10.20
OG	10.5	8.6	10.76	-18.07	8.5	11.92	-18.98
SG	10.4	10.4	<b>0.31</b>	0.60	10.3	<b>0.71</b>	-1.13

<sup>a</sup> t-values in bold do not exceed critical values at 95% confidence level (critical value = 2.262; AG and PG, 2.228)

#### 4.3.3.6 Extraction Recovery

The extraction recovery was calculated by comparing the peak area ratios of human plasma samples spiked with deuterated acylglycines prior to extraction with those spiked post-extraction. The percent recovery was determined at 25 nM and 400 nM in human plasma. The recovery of the three analytes from plasma is represented in Table 4-7 and ranges from 78.5 – 96.6% at 25 nM and 7.49 – 108.5% at 400 nM.

Table 4-7 Extraction recovery of deuterated acylglycines in plasma

Acylglycines	25 nM		400 nM	
	Recovery (%)	CV %	Recovery (%)	CV %
Isobutyrylglycine-d <sub>2</sub>	96.6	12.9	101.0	3.4
Hexanoylglycine-d <sub>2</sub>	95.2	9.9	108.5	12.1
Suberylglycine-d <sub>2</sub>	78.5	14.3	79.4	12.7

#### 4.3.3.7 Precision and Accuracy

Intra-day (days 1 thru 3) and inter-day precision of the proposed method were determined and presented in Table 4-8. Results are expressed as % CV and are all less than 15%. Accuracy was investigated for four concentrations, 10.0, 25.0, 150.0, 400.0 nM and the results are summarized in Table 4-9. QC low was determined as at least two times the LLOQ for each analyte. For acetylglycine, propionylglycine, isobutyrylglycine, hydroxyphenylacetylglycine and tiglylglycine, whose LLOQ are all higher than 4.0 nM, the QC low was determined at 25.0 nM. The results are expressed as % RE and % CV and both sets of values are lower than 15% for all analytes at all concentrations. These results indicate that the proposed method demonstrated satisfactory accuracy and precision.

Table 4-8 Intra-day and inter-day precision measurements of the proposed method

Acylglycines	Intra-day Precision						Inter-day Precision	
	Day 1		Day 2		Day 3		Conc. (nM)	CV %
	Conc. (nM)	CV %	Conc. (nM)	CV %	Conc. (nM)	CV %		
AG	1733	3.7	1647	4.7	1626	4.6	1669	5.0
PG	541	10.3	567	11.6	589	8.2	566	10.3
IBG	114	6.1	117	7.8	116	8.8	116	7.5
BG	115	5.9	110	4.3	112	5.9	112	5.6
HPAG	113	3.8	111	5.5	107	3.4	110	4.9
FG	158	4.7	158	5.5	159	3.6	158	4.5
TG	124	3.7	120	5.8	120	3.5	122	4.6
2MBG	107	4.7	105	5.1	108	2.4	107	4.2
3MCG	105	4.1	107	4.4	104	4.3	105	4.3
IVG	170	2.5	172	2.9	171	4.1	171	3.1
VG	105	2.2	106	4.6	105	2.5	105	3.2
HG	104	4.6	105	3.0	105	2.8	105	3.5
PAG	224	8.3	233	5.5	233	9.5	230	7.9
PPG	104	4.1	101	5.6	100	4.2	102	4.7
GG	105	5.2	103	4.6	104	4.0	104	4.5
HpG	106	3.2	109	4.7	105	3.9	106	4.1
OG	113	5.2	110	4.5	113	2.1	112	4.2
SG	96	4.9	97	7.0	97	5.7	96	5.7

Table 4-9 Accuracy measurements for acylglycines in plasma using the proposed method

Acylglycines	QC Low (10.00 nM/25.00 nM)			QC Mid (150.00 nM)			QC Mid (400.00 nM)		
	Measured value	RE %	CV %	Measured value	RE %	CV %	Measured value	RE %	CV %
AG	26 ± 2	2.1	3.1	148 ± 16	1.6	10.9	404 ± 13	0.9	3.2
PG	25 ± 1	1.3	4.1	151 ± 5	1.0	3.3	404 ± 4	0.9	0.9
IBG	25 ± 2	1.1	6.8	149 ± 8	0.5	5.1	400 ± 18	0.1	4.5
BG	10 ± 1	2.8	9.7	142 ± 8	5.4	6.0	398 ± 9	0.6	2.3
HPAG	26 ± 1	2.9	5.8	147 ± 5	1.8	3.5	403 ± 20	0.8	4.9
FG	10 ± 1	0.9	14.2	145 ± 9	3.0	5.9	411 ± 7	2.7	1.8
TG	24.7 ± 0.8	1.2	3.0	144 ± 6	3.8	4.3	407 ± 30	1.9	7.3
2MBG	10.3 ± 0.6	3.2	6.1	146 ± 4	2.7	2.9	400 ± 10	0.1	2.5
3MCG	10.4 ± 0.5	3.6	4.9	149 ± 6	0.6	4.2	397 ± 10	0.8	2.4
IVG	11 ± 1	6.8	12.6	152 ± 7	1.4	4.3	419 ± 16	4.8	3.9
VG	10.4 ± 0.7	4.3	6.4	143 ± 7	4.9	4.8	397 ± 11	0.7	2.7
HG	10.6 ± 0.5	5.7	4.4	144 ± 3	4.3	2.3	399 ± 11	0.2	2.8
PAG	10.0 ± 0.4	0.2	3.9	143 ± 8	4.7	5.9	402 ± 16	0.4	4.0
PPG	9.9 ± 0.9	0.8	9.2	153 ± 2	2.0	1.0	394 ± 4	1.6	1.1
GG	10 ± 1	2.6	9.5	146 ± 5	2.9	3.4	424 ± 10	6.1	2.3
HpG	10.5 ± 0.4	5.3	3.5	150 ± 5	0.3	3.0	401 ± 10	0.2	2.6
OG	10.6 ± 0.8	6.4	7.2	155 ± 5	3.5	3.5	406 ± 7	1.6	1.6
SG	10.0 ± 0.7	0.2	7.3	147 ± 7	1.8	4.7	401 ± 15	0.3	3.7

#### **4.3.3.8 Stability**

Samples were considered to be stable if the concentrations of analytes were within  $\pm 15\%$  (i.e. 85-115%) of initial concentration in the fresh samples. Results of the stability study conducted under various conditions are expressed as percent recovery and are summarized in Table 4-10. Recovery of all eighteen acylglycines were within the acceptable limits and the analytes were shown to be stable in human plasma samples after 3 freeze/thaw cycles and up to 8 weeks storage at  $-20\text{ }^{\circ}\text{C}$ . It was also demonstrated that they were stable for up to 5 hours at room temperature and 24 hours at  $4\text{ }^{\circ}\text{C}$  in the autosampler.

#### **4.3.4 Measurement of endogenous acylglycines**

The validated LC-MS/MS method was used to analyze healthy plasma from 10 volunteers. Each sample was derivatized and analyzed in triplicates. Concentrations were back-calculated using peak area ratios and the calibration curve prepared in surrogate plasma and was expressed as mean  $\pm$  standard deviation. The endogenous concentration of each acylglycines is shown in Table 4-11. Only seven acylglycines, acetylglycine, propionylglycine, butyrylglycine, tiglylglycine, isovalerylglycine, hexanoylglycine, and hydroxyphenylacetylglycine were detected in the 10 individuals. The other acylglycines were detected in quantities less than LLOQ in most of the individuals. Confirmation of acylglycine identity was performed using co-elution of the internal standard and the interpretation of the product ion spectra.

Table 4-10 Stability results for various storage conditions

Acyl glycine	Room Temp 5 hours		Autosampler 24 hours		Freeze/Thaw Cycle 1		Freeze/Thaw Cycle 2		Freeze/Thaw Cycle 3		-20 °C 2 weeks		-20 °C 8 weeks	
	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)	Accuracy (%)	CV (%)
AG	99.4	8.0	98.3	9.1	99.2	6.5	102.9	7.8	99.9	6.8	94.3	11.2	94.1	6.5
PG	99.5	9.8	100.9	3.8	97.8	1.8	97.3	5.3	99.4	5.6	97.2	9.3	99.4	1.1
IBG	100.6	5.4	97.5	2.3	97.1	2.7	97.8	4.0	101.3	9.4	99.4	13.4	94.4	1.8
BG	98.6	6.8	100.5	8.7	99.7	8.0	96.2	5.9	96.0	6.2	94.2	7.0	98.2	6.1
HPAG	98.8	6.4	99.5	6.0	98.7	4.7	100.6	5.6	101.5	10.1	97.0	5.8	97.3	4.8
FG	101.1	6.0	96.8	6.3	99.9	6.3	103.7	5.9	101.5	7.8	98.8	8.5	94.2	6.0
TG	96.2	2.3	99.2	3.1	100.2	4.4	98.0	5.0	98.8	1.4	100.0	3.4	97.0	0.8
2MBG	100.8	3.5	96.6	6.6	96.5	4.8	99.5	5.2	97.1	4.1	99.7	12.0	98.6	6.7
3MCG	98.2	7.1	98.9	5.3	97.1	6.6	94.0	9.3	98.1	5.6	99.1	8.5	94.2	5.4
IVG	99.7	4.5	102.8	4.2	100.2	3.8	102.7	3.3	102.8	3.5	100.8	10.2	98.8	3.8
VG	99.2	3.1	99.1	4.8	96.0	5.6	98.4	6.1	98.6	5.1	98.1	3.5	98.2	6.6
HG	100.1	4.8	97.8	4.2	98.2	5.5	100.4	5.9	98.1	6.9	100.0	5.0	98.6	8.7
PAG	98.9	2.2	101.1	0.4	96.9	1.4	102.8	1.9	99.6	2.8	101.3	4.8	103.0	2.4
PPG	95.1	3.9	93.8	5.1	93.1	3.5	93.1	4.7	95.8	4.7	97.3	9.1	98.1	4.3
GG	98.9	6.4	100.6	6.0	93.2	7.4	93.3	11.3	93.2	8.0	93.1	7.7	94.1	6.1
HpG	95.6	5.5	100.4	7.9	98.8	5.2	97.5	6.3	98.0	6.1	96.0	8.5	98.8	5.4
OG	93.9	5.8	96.7	7.6	96.4	5.8	98.6	8.0	101.3	6.5	98.0	5.8	100.4	6.0
SG	101.6	9.0	98.5	12.0	98.0	12.0	95.9	10.5	94.6	10.7	97.0	8.9	94.5	10.9

Table 4-11 Endogenous concentrations of acylglycines in plasma. Concentration expressed as  $\mu\text{mol/mol}$  creatinine.

<b>Sample Name</b>	<b>AG</b>	<b>PG</b>	<b>IBG</b>
Individual #1	1013 $\pm$ 153	331 $\pm$ 2	< 7.0
Individual #2	1344 $\pm$ 69	268 $\pm$ 44	< 7.0
Individual #3	1482 $\pm$ 73	599 $\pm$ 81	< 7.0
Individual #4	1073 $\pm$ 148	388 $\pm$ 51	< 7.0
Individual #5	1693 $\pm$ 244	331 $\pm$ 47	< 7.0
Individual #6	977 $\pm$ 69	309 $\pm$ 20	< 7.0
Individual #7	1659 $\pm$ 256	497 $\pm$ 75	< 7.0
Individual #8	1908 $\pm$ 118	502 $\pm$ 47	7.0 $\pm$ 0.8
Individual #9	1753 $\pm$ 173	417 $\pm$ 43	< 7.0
Individual #10	1535 $\pm$ 182	493 $\pm$ 40	< 7.0
<b>Sample Name</b>	<b>BG</b>	<b>HPAG</b>	<b>FG</b>
Individual #1	3.5 $\pm$ 0.5	< 5.0	< 4.0
Individual #2	3.4 $\pm$ 0.2	< 5.0	24.3 $\pm$ 0.1
Individual #3	5.1 $\pm$ 0.3	< 5.0	< 4.0
Individual #4	5.2 $\pm$ 0.7	< 5.0	38.5 $\pm$ 2.8
Individual #5	4.0 $\pm$ 0.6	< 5.0	< 4.0
Individual #6	3.66 $\pm$ 0.02	< 5.0	232.5 $\pm$ 49.3
Individual #7	6.8 $\pm$ 0.5	< 5.0	< 4.0
Individual #8	4.6 $\pm$ 0.7	< 5.0	5.0 $\pm$ 0.6
Individual #9	6.3 $\pm$ 0.8	< 5.0	< 4.0
Individual #10	4.4 $\pm$ 0.2	< 5.0	5.6 $\pm$ 0.2
<b>Sample Name</b>	<b>TG</b>	<b>2MBG</b>	<b>3MCG</b>
Individual #1	11 $\pm$ 2	< 3.0	< 4.0
Individual #2	20.9 $\pm$ 0.1	< 3.0	< 4.0
Individual #3	13 $\pm$ 2	< 3.0	< 4.0
Individual #4	13 $\pm$ 1	< 3.0	< 4.0
Individual #5	10 $\pm$ 1	< 3.0	< 4.0
Individual #6	22 $\pm$ 4	< 3.0	< 4.0
Individual #7	11 $\pm$ 1	< 3.0	< 4.0
Individual #8	10.0 $\pm$ 0.8	< 3.0	< 4.0
Individual #9	12 $\pm$ 2	< 3.0	< 4.0
Individual #10	< 10.0	< 3.0	< 4.0

<b>Sample Name</b>	<b>IVG</b>	<b>VG</b>	<b>HG</b>
Individual #1	40 ± 2	< 3.0	6.04 ± 0.01
Individual #2	45 ± 3	< 3.0	< 1.0
Individual #3	39 ± 3	< 3.0	1.1 ± 0.1
Individual #4	64 ± 7	< 3.0	1.5 ± 0.1
Individual #5	46 ± 6	< 3.0	4.9 ± 0.1
Individual #6	76 ± 11	< 3.0	1.2 ± 0.2
Individual #7	54 ± 2	< 3.0	4.4 ± 0.2
Individual #8	62.4 ± 0.9	< 3.0	5.1 ± 0.7
Individual #9	62 ± 2	< 3.0	1.8 ± 0.2
Individual #10	67.06 ± 0.02	< 3.0	1.90 ± 0.03
<b>Sample Name</b>	<b>PAG</b>	<b>PPG</b>	<b>GG</b>
Individual #1	106 ± 18	< 3.0	< 2.0
Individual #2	88 ± 15	< 3.0	< 2.0
Individual #3	95 ± 16	< 3.0	< 2.0
Individual #4	121 ± 17	< 3.0	< 2.0
Individual #5	81 ± 8	< 3.0	< 2.0
Individual #6	59 ± 9	< 3.0	< 2.0
Individual #7	133 ± 11	< 3.0	< 2.0
Individual #8	127 ± 5	< 3.0	< 2.0
Individual #9	113 ± 9	< 3.0	< 2.0
Individual #10	111 ± 2	< 3.0	2.57 ± 0.06
<b>Sample Name</b>	<b>HpG</b>	<b>OG</b>	<b>SG</b>
Individual #1	< 3.0	< 3.0	< 4.0
Individual #2	< 3.0	< 3.0	< 4.0
Individual #3	< 3.0	< 3.0	< 4.0
Individual #4	< 3.0	< 3.0	< 4.0
Individual #5	< 3.0	< 3.0	< 4.0
Individual #6	< 3.0	< 3.0	< 4.0
Individual #7	< 3.0	< 3.0	< 4.0
Individual #8	< 3.0	< 3.0	< 4.0
Individual #9	< 3.0	< 3.0	< 4.0
Individual #10	< 3.0	< 3.0	< 4.0

The acylglycine concentration profile of the 10 individuals was examined to determine if there was any dependence of the acylglycine excretion pattern on gender. The influence on gender is shown in Figure 4-3. The box plots were generated using the full range of values. Eight acylglycines, whose values were above LLOQ for at least 5 of the individuals, were plotted using the mean values. From the box plots, there is no discernible pattern observed in the acylglycine profile, except for 2-furoylglycine, where the values for females are below LLOQ and the values for males are above LLOQ. The source of 2-furoylglycines is exogenous, originating from the metabolism via oxidation of furfural to furoic acid. Furoic acid is then conjugated to glycine and excreted in urine. Furfural can be found in processed foods and beverages, chocolate and as a solvent in lubricating oils.<sup>15, 16</sup> It is unclear at this time why 2-furoylglycine is higher in males than in females.

#### **4.4 Conclusion**

A validated and sensitive LC-MS/MS method is described for the determination of acylglycines in human plasma samples. The surrogate matrix approach was able to overcome the analytical challenges incurred when measuring the endogenous acylglycines. The method was successfully applied in determining the endogenous levels of acylglycines in the plasma samples of 10 healthy individuals. This strategy can be adopted in developing an assay for the determination of acylglycines in the plasma of patients or newborns with IEMs.

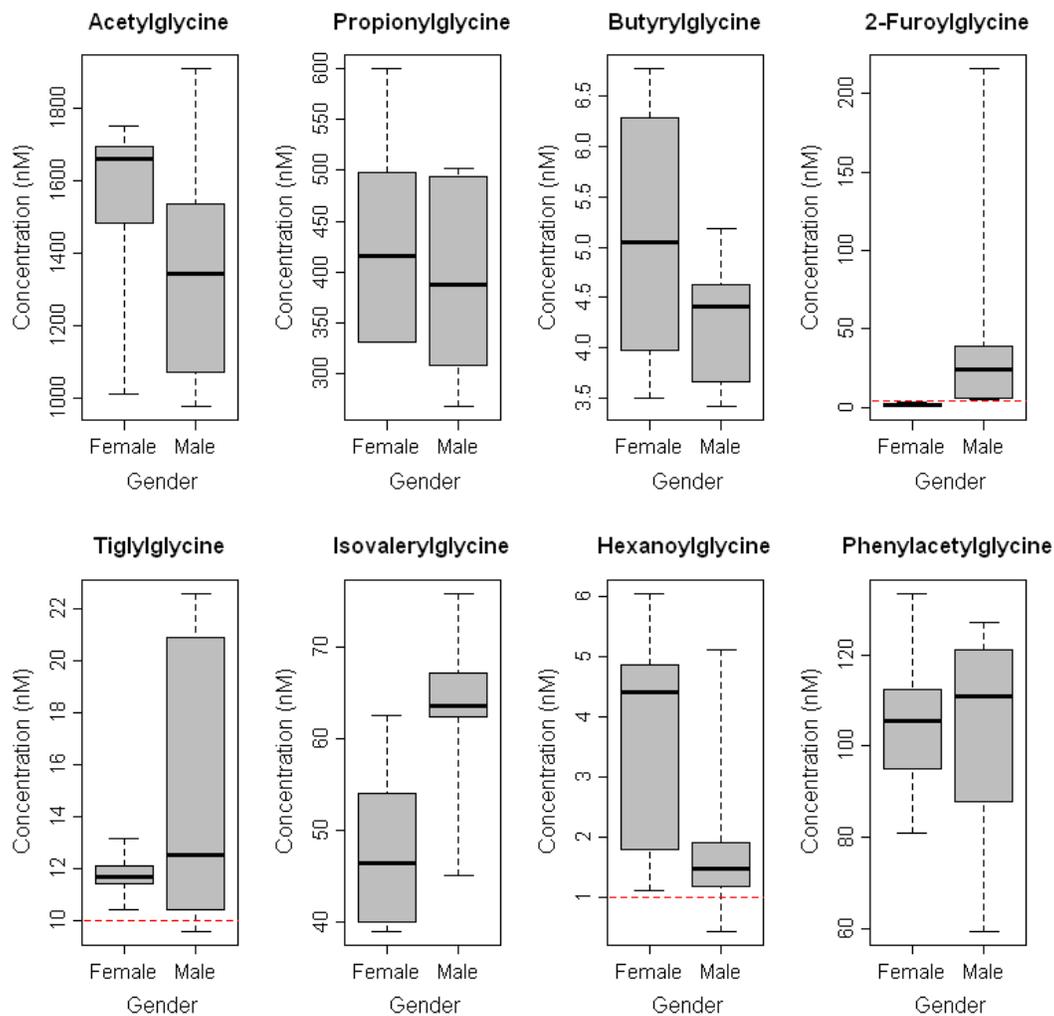


Figure 4-3 Influence of gender on the acylglycine plasma profile. Red dotted line indicates the lower limit of quantification of the analyte in plasma.

## 4.5 Literature Cited

- (1) Kompare, M.; Rizzo, W. B. *Seminars in Pediatric Neurology* **2008**, *15*, 140-149.
- (2) la Marca, G.; Rizzo, C. In *Metabolic Profiling: Methods and Protocols*; Metz, T. O., Ed.; Humana Press Inc: Totowa, 2011; Vol. 708, pp 73-98.
- (3) Lewis-Stanislaus, A. E.; Li, L. *J. Am. Soc. Mass Spectrom.* **2011**, *21*, 2105-2116.
- (4) Ombrone, D.; Salvatore, F.; Ruoppolo, M. *Anal. Biochem.* **2011**, *417*, 122-128.
- (5) Bonafe, L.; Troxler, H.; Kuster, T.; Heizmann, C. W.; Chamoles, N. A.; Burlina, A. B.; Blau, N. *Mol. Genet. Metab.* **2000**, *69*, 302-311.
- (6) Ensenauer, R.; Fingerhut, R.; Maier, E. M.; Polanetz, R.; Olgemoller, B.; Roschinger, W.; Muntau, A. C. *Clin. Chem.* **2011**, *57*, 623-626.
- (7) Kasper, D. C.; Ratschmann, R.; Metz, T. F.; Mechtler, T. P.; Moslinger, D.; Konstantopoulou, V.; Item, C. B.; Pollak, A.; Herkner, K. R. *Wien. Klin. Wochen.* **2010**, *122*, 607-613.
- (8) Kennedy, S.; Potter, B. K.; Wilson, K.; Fisher, L.; Geraghty, M.; Milburn, J.; Chakraborty, P. *BMC Pediatr.* **2010**, *10*, 8.
- (9) McHugh, D. M. S.; Cameron, C. A.; Abdenur, J. E.; Abdulrahman, M.; Adair, O.; Al Nuaimi, S. A.; Ahlman, H.; Allen, J. J.; Antonozzi, I.; Archer, S.; Au, S.; Auray-Blais, C.; Baker, M.; Bamforth, F.; Beckmann, K.; Pino, G. B.; Berberich, S. L.; Binard, R.; Boemer, F.; Bonham, J.; Breen, N. N.; Bryant, S. C.; Caggana, M.; Caldwell, S. G.; Camilot, M.; Campbell, C.; Carducci, C.; Cariappa, R.; Carlisle, C.; Caruso, U.; Cassanello, M.; Castilla, A. M.; Ramos, D. E. C.; Chakraborty, P.; Chandrasekar, R.; Ramos, A. C.; Cheillan, D.; Chien, Y. H.; Childs, T. A.; Chrastina, P.; Sica, Y. C.; de Juan, J. A. C.; Colandre, M. E.; Espinoza, V. C.; Corso, G.; Currier, R.; Cyr, D.; Czuczy, N.; D'Apolito, O.; Davis, T.; de Sain-Van der Velden, M. G.; Pecellin, C. D.; Di Gangi, I. M.; Di Stefano, C. M.; Dotsikas, Y.; Downing, M.; Downs, S. M.; Dy, B.; Dymerski, M.; Rueda, I.; Elvers, B.; Eaton, R.; Eckerd, B. M.; El Mougy, F.; Eroh, S.; Espada, M.; Evans, C.; Fawbush, S.; Fijolek, K. F.; Fisher, L.; Franzson, L.; Frazier, D. M.; Garcia, L. R. C.; Bermejo, M.; Gavrillov, D.; Gerace, R.; Giordano, G.; Irazabal, Y. G.; Greed, L. C.; Grier, R.; Grycki, E.; Gu, X. F.; Gulamali-Majid, F.; Hagar, A. F.; Han, L. S.; Hannon, W. H.; Haslip, C.; Hassan, F. A.; He, M. A.; Hietala, A.; Himstedt, L.; Hoffman, G. L.; Hoffman, W.; Hoggatt, P.; Hopkins, P. V.;

Hougaard, D. M.; Hughes, K.; Hunt, P. R.; Hwu, W. L.; Hynes, J.; Ibarra-Gonzalez, I.; Ingham, C. A.; Ivanova, M.; Jacox, W. B.; John, C.; Johnson, J. P.; Jonsson, J. J.; Karg, E.; Kasper, D.; Klopper, B.; Katakouzinou, D.; Khneisser, I.; Knoll, D.; Kobayashi, H.; Koneski, R.; Kozich, V.; Kouapei, R.; Kohlmüller, D.; Kremensky, I.; la Marca, G.; Lavochkin, M.; Lee, S. Y.; Lehotay, D. C.; Lemes, A.; Lepage, J.; Lesko, B.; Lewis, B.; Lim, C.; Linard, S.; Lindner, M.; Lloyd-Puryear, M. A.; Lorey, F.; Loukas, Y. L.; Luedtke, J.; Maffitt, N.; Magee, J. F.; Manning, A.; Manos, S.; Marie, S.; Hadachi, S. M.; Marquardt, G.; Martin, S. J.; Matern, D.; Gibson, S. K. M.; Mayne, P.; McCallister, T. D.; McCann, M.; McClure, J.; McGill, J. J.; McKeever, C. D.; McNeilly, B.; Morrissey, M. A.; Moutsatsou, P.; Mulcahy, E. A.; Nikoloudis, D.; Norgaard-Pedersen, B.; Oglesbee, D.; Oltarzewski, M.; Ombrone, D.; Ojodu, J.; Papakonstantinou, V.; Reoyo, S. P.; Park, H. D.; Pasquali, M.; Pasquini, E.; Patel, P.; Pass, K. A.; Peterson, C.; Pettersen, R. D.; Pitt, J. J.; Poh, S.; Pollak, A.; Porter, C.; Poston, P. A.; Price, R. W.; Queijo, C.; Quesada, J.; Randell, E.; Ranieri, E.; Raymond, K.; Reddic, J. E.; Reuben, A.; Ricciardi, C.; Rinaldo, P.; Rivera, J. D.; Roberts, A.; Rocha, H.; Roche, G.; Greenberg, C. R.; Mellado, J. M. E.; Juan-Fita, M. J.; Ruiz, C.; Ruoppolo, M.; Rutledge, S. L.; Ryu, E. J.; Saban, C.; Sahai, I.; Garcia-Blanco, M. I. S.; Santiago-Borrero, P.; Schenone, A.; Schoos, R.; Schweitzer, B.; Scott, P.; Seashore, M. R.; Seeterlin, M. A.; Sesser, D. E.; Sevier, D. W.; Shone, S. M.; Sinclair, G.; Skrinska, V. A.; Stanley, E. L.; Strovel, E. T.; Jones, A. L. S.; Sunny, S.; Takats, Z.; Tanyalcin, T.; Teofoli, F.; Thompson, J. R.; Tomashitis, K.; Domingos, M. T.; Torres, J.; Torres, R.; Tortorelli, S.; Turi, S.; Turner, K.; Tzanakos, N.; Valiente, A. G.; Vallance, H.; Vela-Amieva, M.; Vilarinho, L.; von Döbeln, U.; Vincent, M. F.; Vorster, B. C.; Watson, M. S.; Webster, D.; Weiss, S.; Wilcken, B.; Wiley, V.; Williams, S. K.; Willis, S. A.; Woontner, M.; Wright, K.; Yahyaoui, R.; Yamaguchi, S.; Yssel, M.; Zakowicz, W. M. *Genet. Med.* **2011**, *13*, 230-254.

- (10) Carter, S. M. B.; Watson, D. G.; Midgley, J. M.; Logan, R. W. J. *Chromatogr., B: Biomed. Sci. Appl.* **1996**, *677*, 29-35.
- (11) van de Merbel, N. C. *Trac-Trends Anal. Chem.* **2008**, *27*, 924-933.
- (12) Houghton, R.; Pita, C. H.; Ward, I.; Macarthur, R. *Bioanalysis* **2009**, *1*, 1365-1374.
- (13) Zar, J. H. In *Biostatistical Analysis*, 5 ed.; Snavely, S. L., Ed.; Pearson Prentice Hall Inc.: Upper Saddle River, 2010, pp 363-378.
- (14) Beltrán, E.; Ibáñez, M.; Sancho, J. V.; Hernández, F. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1801-1809.

- (15) Kumps, A.; Duez, P.; Mardens, Y. *Clin. Chem.* **2002**, *48*, 708-717.
- (16) Tsai, S.; Huang, M. *J. Occupational Safety Health* **2008**, *16*, 370-382.

## Chapter 5. Development of a Fast Liquid Chromatography for the Analysis of Homocysteine, Succinic Acid and Methylmalonic Acid in *C. elegans* Worm Media\*

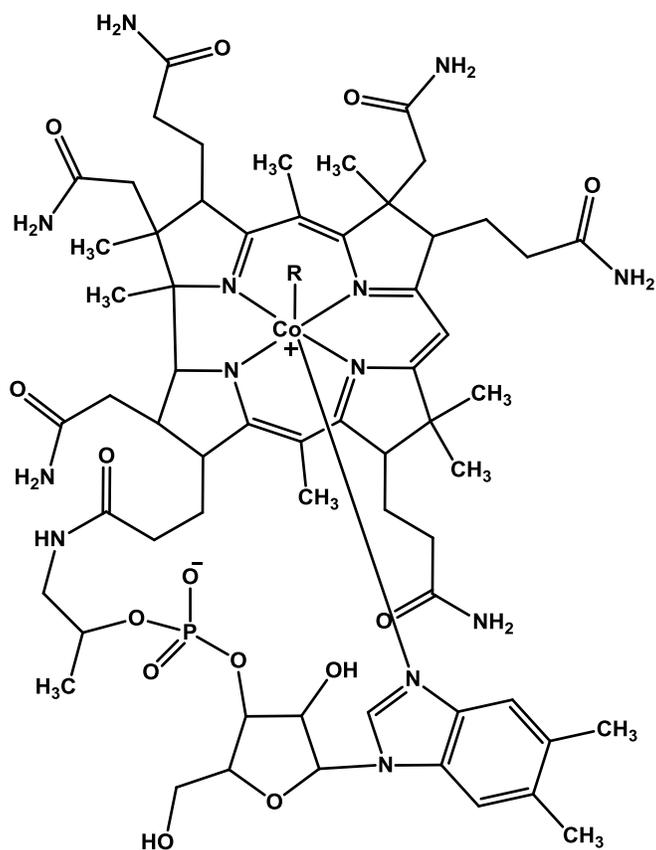
### 5.1 Introduction

Vitamin B<sub>12</sub> (cobalamin) is a complex cobalt-containing molecule that is an essential nutrient for humans with a key role in normal functioning of the nervous system and blood cell production.<sup>1-3</sup> The structure of vitamin B<sub>12</sub> and its common coenzyme forms can be seen in Figure 5-1. Both the cyano- and hydroxycobalamin must be converted to methyl and adenosylcobalamin, which function as two enzyme co-factors.<sup>2,3</sup> Cobalamin is not manufactured in the body and must be absorbed through dietary or supplementary intake. The vitamin must be modified through intracellular metabolism to generate the different physiological forms.<sup>2-4</sup> The metabolism of cobalamin is quite complex, requiring many steps in the process. If any of these steps are absent or blocked, cobalamin deficiency can occur. Common causes of cobalamin deficiency include dietary deficiency, impaired intestinal uptake, pernicious anemia and malabsorption. Less common are genetic diseases that are known to affect newborns and young children that block the metabolism of the vitamin B<sub>12</sub>. These diseases may cause defects in the transcobalamins (cobalamin transport proteins) or in the intracellular enzymes.<sup>3-5</sup>

In mammals, cobalamin plays a significant role as a cofactor in two physiological reactions: (A) the synthesis of methionine from homocysteine (Hcy) and (B) the conversion of methylmalonic acid (MMA) to succinic acid (SA).<sup>3</sup>

---

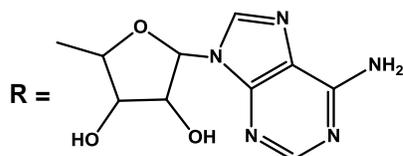
\* This work was done in collaboration with Dr. Roy Gravel's laboratory at the University of Calgary. Avalyn Stanislaus contributed by developing a fast LC-MS/MS method for the analysis of the analytes.



R =  $\text{C}\equiv\text{N}$       **Cyanocobalamin**

R = OH      **Hydroxycobalamin**

R =  $\text{CH}_3$       **Methylcobalamin**



**Adenosylcobalamin**

Figure 5-1 Structures of cobalamin and the enzyme cofactors

Both of these reactions are represented in Figure 5-2 and Figure 5-3 respectively. In these reactions, cobalamin is an essential cofactor for the enzymes methionine synthase, which is a cytosolic enzyme and methylmalonyl-CoA mutase, which is a mitochondrial enzyme. For each of these enzymes, the cofactor is in a different physiological form: methylcobalamin for methionine synthase and adenosylcobalamin for methylmalonyl-CoA.

In the case of a deficiency of the vitamin, the two reactions are impaired and concentrations of Hcy and MMA are elevated. A series of analytical procedures and chromatographic techniques have been employed for the diagnosis of cobalamin deficiency: total vitamin B<sub>12</sub>,<sup>3, 6-8</sup> holotranscobalamin (transcobalamin bound to cobalamin) determination,<sup>8-13</sup> Hcy measurements,<sup>14-17</sup> and MMA concentration.<sup>15, 17-19</sup> Determination of Hcy and MMA is preferable to measuring cobalamin for the following reasons: (1) cobalamin is not as stable as Hcy and MMA<sup>20</sup>, (2) cobalamin concentrations fall after the concentrations of MMA and Hcy increase and may not adequately reflect tissue concentrations,<sup>21, 22</sup> (3) measuring an increase in cobalamin concentration is less difficult than measuring a decrease due to the low sensitivity of the cobalamin assay.

Chromatographic methods for Hcy, SA and MMA include gas chromatography,<sup>23-29</sup> capillary electrophoresis<sup>30-34</sup> and high performance liquid chromatography<sup>35-45</sup> have been previously used. Gas chromatographic methods can be time consuming and labor intensive as there is an additional sample derivatization step involved in the analysis. Liquid chromatography has significant advantages over the other two methods in that it yields shorter analysis times and no derivatization is necessary.

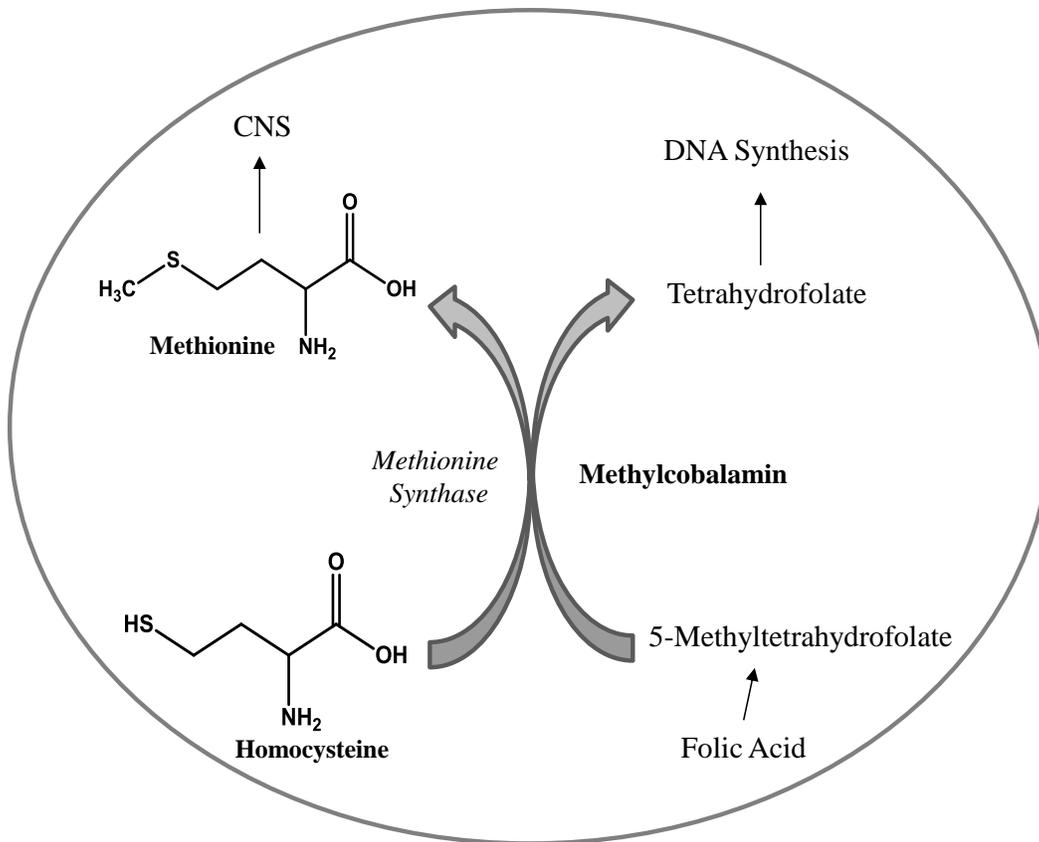


Figure 5-2 Homocysteine pathway. Homocysteine is methylated to form methionine by a folate-dependent reaction. 5-Methyltetrahydrofolate is the methyl donor for the methylation of homocysteine. The enzyme methionine synthase using methylcobalamin as a co-factor catalyzes this conversion.

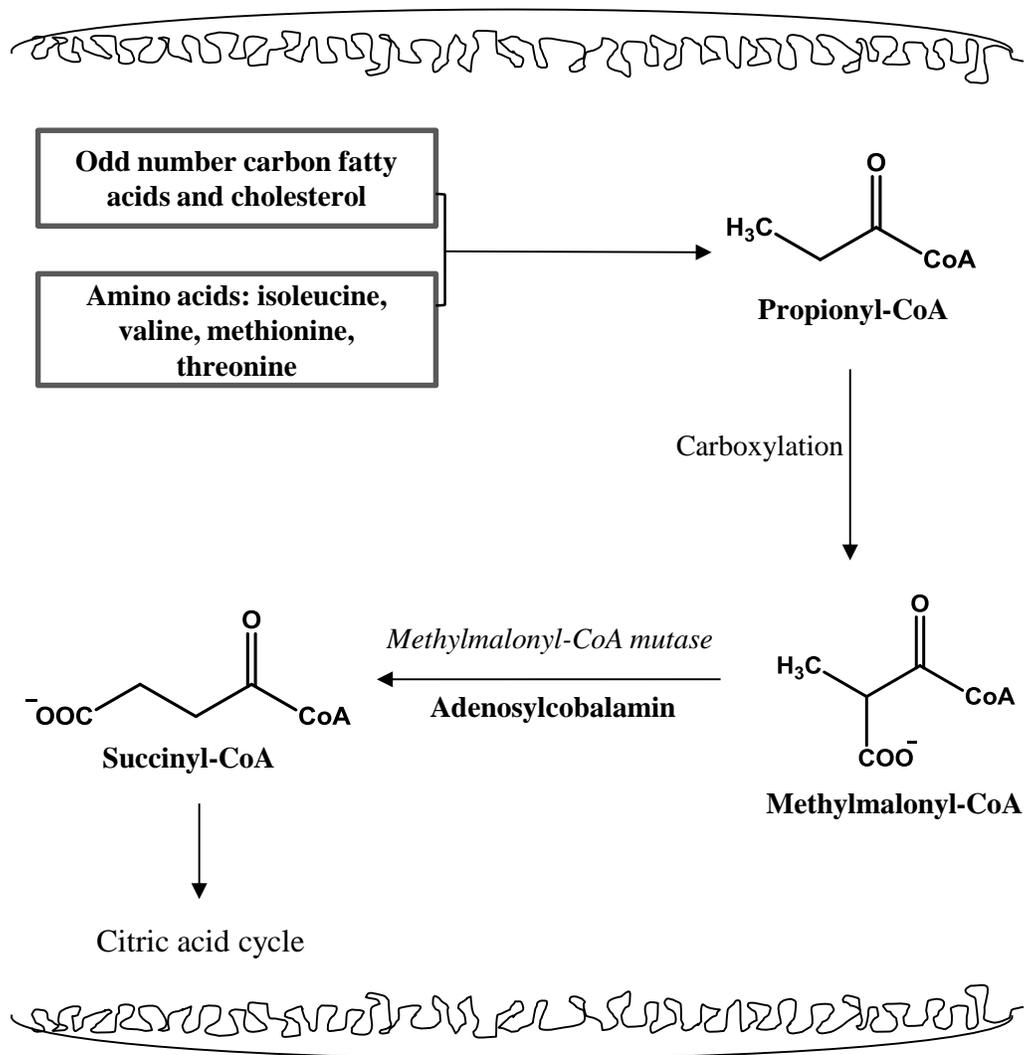


Figure 5-3 Methylmalonic pathway. Propionate is a product in the catabolic pathway of odd number fatty acids, cholesterol and the amino acids isoleucine, valine, methionine and threonine. Propionate is then carboxylated to methylmalonic acid, which is isomerized to succinic acid. The isomerization is catalyzed by the enzyme methylmalonyl-CoA mutase and the co-factor adenosylcobalamin.

In this work, a fast LC-MS/MS method was developed for the simultaneous determination of Hcy, MMA and SA excreted from *Caenorhabditis elegans*. The nematode *C. elegans* has been utilized to study cobalamin metabolism and deficiency in humans.<sup>46-52</sup> This organism was chosen as a model because i) they are small, transparent, feed on bacteria and have a short life span, ii) as micro-organisms, they can be easily grown, iii) they share cellular structures and biological pathways with those of higher organisms and iv) its genome has been sequenced and their DNA sequences have cobalamin pathway genes that are quite similar to humans. Because these organisms have similar metabolism, they can be a powerful tool in helping to unravel human diseases.<sup>53-55</sup> Another useful characteristic of *C.elegans* is that it is a simple task to disrupt the function of certain genes using RNA interference (RNAi).<sup>56-59</sup> In silencing genes, it is possible to ascertain the function of that gene. *C.elegans* were used in an effort to identify unknown genes and metabolic processes in the cobalamin metabolism. *C. Elegans* worms were grown on bacteria expressing RNAi for individual genes coding for relevant enzymes in the Vitamin B<sub>12</sub> pathway. The worms feed on the bacteria and incorporate the double-stranded RNA, which will knock out expression of a particular gene. In the case of these experiments, if expression of a vitamin B<sub>12</sub> gene is knocked out, then synthesis of one of the cofactors or enzymes for e.g. methylmalonyl-CoA mutase will be blocked. The worms will then have elevated concentrations of methylmalonic acid, which will be excreted.

In plasma and interstitial fluids, cobalamin travels bound to transcobalamin (TC) protein complex, is internalized by cells and is filtered through the kidneys. It is commonly believed that cobalamin is secreted from enterocytes complexed to TC, but it is not known how intracellular cobalamin is actually exported from cells, including the enterocytes and exchanged between organs and cells. Whatever the means of cobalamin excretion from a cell, whether in free form or as a complex, it must involve transport across the plasma membrane. In this study, *C. elegans* is used as a model organism to define the role of ATP-binding cassette (ABC) transporters in cobalamin metabolism. ABC

transporters are transmembrane proteins that use energy in the form of ATP to carry out certain biological functions, including transport of various substrates across the plasma membrane and DNA repair. They are also involved in inherited human diseases as well as resistance to multiple drugs. *C. elegans* genome contains 60 ABC transporters and these include the multidrug resistance protein (MRP) and P-glycoprotein (PGP), each of which have several homologs.<sup>60</sup> Recently, the MRP-1 homolog was identified as a cobalamin intestinal efflux transporter.<sup>61</sup> Currently, 42 available ABC transporter mutant strains of *C.elegans* are expected to be screened using the current LC-MS/MS assay.

## **5.2 Experimental**

### **5.2.1 Chemicals and Materials**

Optima grade acetonitrile, methanol and water were purchased from Fisher Scientific (Ottawa, ON). MMA, SA, DL-Hcy, MMA-d<sub>3</sub> and SA-d<sub>4</sub> were purchased from Sigma Aldrich (Oakville, ON). Formic acid used in the LC-MS/MS analyses was also obtained from Sigma-Aldrich.

### **5.2.2 Preparation of Samples**

*C. elegans* were recovered after growth on the bacterial feeder strains, washed and placed in an incubation medium for one to fourteen days, during which time, they may or may not excrete MMA. Three different incubation media were tested: i) phosphate buffered saline (PBS), ii) Puck's saline mix also known in this study as incubation mix, consisting of phosphate buffered saline, glucose, alanine and 15% fetal bovine serum) and iii) *C. elegans* habitation and reproduction (CeHR) worm media, consisting of salts, minerals, vitamins, growth factors, nucleic acids, amino acids and 20% skim milk. An antibiotic (50 µg/mL kanamycin) was added to all the media to prevent microbial growth. All three

types of media may be fortified with 10 mM propionate as a precursor molecule in the MMA pathway. The media were then analyzed for the presence of methylmalonic acid, excreted by the worms. Method development was done by spiking the analytes at different concentrations of 0.1  $\mu$ M, 1  $\mu$ M, 10  $\mu$ M, and 100  $\mu$ M into these media. This was done to evaluate each media for interferences and matrix effects and to select the best media for the analysis.

### **5.2.3 Preparation of *C. elegans* RNAi-fed Strains**

RNAi bacterial strains were streaked on LB agar plates with 100  $\mu$ g/mL ampicillin and 20  $\mu$ g/mL tetracycline and grown overnight at 37 °C. The strains were grown overnight from individual colonies in LB with 100  $\mu$ g/mL of ampicillin at 37 °C and 250 rpm. One hundred microliters of an overnight culture was spread on NGM plates containing 100  $\mu$ g/mL ampicillin and 1 mM IPTG. Plates were allowed to dry and placed at 37 °C overnight and stored at 4 °C. Single N2 worms were picked to each and placed at 20 °C until bacterial clearance. The plates were harvested using Multiprobe II Plus Program. Worms were washed once with sterile ROD Water in 96 deep-well plate and spun at 800 g for 5 minutes at 4C. They were then once with 1 mL sterile media for assay, using PBS + kanamycin for <sup>14</sup>C propionate assay and LC-MS/MS assay.

### **5.2.4 Preparation of L1 larvae for *C. elegans* Deletion Strains**

One gravid *C. elegans* worms was picked per 14 cm diameter NGM agar plates containing HB101 bacteria and incubated at 20 °C for 7 or more days. Gravid worms from a single plate were washed three times with 0.1 M sodium chloride and combined. The tubes were centrifuged at 800 g at 4 °C. Pelleted worms were resuspended in 0.1 M sodium chloride and washed with bleach/sodium hydroxide

(2:1 v,v). Eggs were washed with cold sterile water, resuspended in M9 buffer, transferred to a tissue culture flask and rocked at room temperature overnight.

### **5.2.5 Methylmalonic Acid Accumulation Assay**

Worms prepared as described above were added to well of a 24-well plate topped PBS + Kanamycin. Sodium propionate (200 mM) was added and the plate sealed. The plate was rotated at 175 rpm at 20 °C for 7 days and the contents of the wells transferred to tubes and centrifuged at 800 g at 4 °C. Supernatants were transferred to fresh tubes and stored at -20 °C until analysis by LC-MS/MS.

### **5.2.6 <sup>14</sup>C-propionate Incorporation Assay**

Worms were added to wells topped up with PBS + kanamycin and sodium propionate. <sup>14</sup>C-propionate was then added to the wells. The plate was covered and shaken at 200 rpm at room temperature for 24 hours. Contents of the cells were transferred and centrifuged at 800 g and the pellets resuspended by vortexing in 5% trichloroacetic acid. The tubes were heated to 85 °C and placed on ice for 20 minutes. The tubes were then centrifuged at 18,000 g and resuspended in acetone and centrifuged again. Pellets were air-dried and resuspended by vortexing in 5% SDS. Sterile DDW was added and 460 µL was mixed with 3 mL CytoScint fluid and read on a Beckman Scintillation counter.

### **5.2.7 Preparation of Standards**

Stock solutions of MMA, SA and Hcy, each at 1.0 mM, were prepared in Optima grade water. Internal standards, MMA-d<sub>3</sub> and SA-d<sub>4</sub>, each at 0.5mM were also prepared in Optima grade water. Working solutions were made by diluting the stock solutions in either PBS solutions fortified with antibiotic and propionic

acid, worm incubator mix and CeHR incubation mix to the following concentrations: 0.001  $\mu\text{M}$ , 0.005  $\mu\text{M}$ , 0.010  $\mu\text{M}$ , 0.050  $\mu\text{M}$ , 0.100  $\mu\text{M}$ , 0.500  $\mu\text{M}$ , 1.0  $\mu\text{M}$ , 5.0  $\mu\text{M}$ , 10.0  $\mu\text{M}$ , 50.0  $\mu\text{M}$ , 100.0  $\mu\text{M}$ .

### **5.2.8 Ultra-Performance Liquid Chromatography (UPLC)**

Chromatographic separation was performed on a Waters ACQUITY UPLC system with a 1.7- $\mu\text{m}$  bridged ethylene hybrid (BEH), 2.1 mm x 50 mm  $\text{C}_{18}$  column. The separation was carried out by isocratic elution using a mobile phase consisting of 4% acetonitrile in water/ 0.1% formic acid for 2 minutes, followed by a ramp to 100% acetonitrile/0.1% formic acid in 0.2 minutes, and held at 100% for 1.5 minutes. The column was maintained at a temperature of 25  $^{\circ}\text{C}$ . The flow rate of the method was varied (0.150 mL/min – 0.500 mL/min) and 5.0  $\mu\text{L}$  of each sample was injected onto the column.

### **5.2.9 Mass Spectrometry**

The LC system was coupled to an AB Sciex 4000 QTRAP<sup>®</sup> hybrid triple quadrupole linear ion trap mass spectrometer (Concord, ON), equipped with a Turbo Ionspray ionization interface. The ESI source was set to perform in polarity switching mode, operating in both positive and negative modes using multiple reaction monitoring (MRM) in a single run. The optimized source parameters were as follows: ion spray voltage, 4800 V (positive), -4200 V (negative); curtain gas, 10 arbitrary units; collision gas, high; temperature, 350  $^{\circ}\text{C}$ ; GS1, 30; GS2, 20 (gas measured in arbitrary units); dwell time for each ion pair, 100 ms; nitrogen was used as the nebulizer, curtain and collision gas. The compound parameters were optimized and are represented in Table 5-1. These parameters were tuned to produce the most intense ion signals of the analytes. Data was acquired and processed using Analyst<sup>®</sup> version 1.4.2 software.

Table 5-1 Instrument parameters and MRM transitions used in the polarity switching method

ESI mode	Analyte	MRM Transition	Compound-dependent Parameters			
			DP	EP	CE	CXP
Positive	Hcy	136.0/90.0	20	12	20	10
Negative	SA, MMA	117.0/73.0	-20	-14	-15	-18

DP, declustering potential; EP, entrance potential; CE, collision energy; CXP, collision cell exit potential

## 5.3 Results and Discussion

### 5.3.1 Mass Analysis

Whereas Hcy was better ionized in the positive ion mode, succinic acid and methylmalonic acid had higher sensitivity in the negative mode. Because of the different ionization conditions, both positive and negative ion modes were alternated in order to detect all the analytes. The MRM transitions used are represented in Table 5-1. Product ion spectra of Hcy, MMA, and SA are represented in Figure 5-4. The product ion spectrum of Hcy in Figure 5-4 (A) shows the protonated molecular ion  $[M+H]^+$  at  $m/z$  136.0. Fragmentation produced two major fragments at  $m/z$  118.0 and  $m/z$  90.0, which are the loss of  $H_2O$  and  $(H_2O + CO)$  respectively. Minor fragment ions observed were  $m/z$  73 and  $m/z$  56, which may possibly be the losses of  $(H_2O + CO + NH_3)$  and  $(H_2O + CO + H_2S)$ .

In the negative mode, the MRM transition of  $m/z$  117 to  $m/z$  73.0 was used for MMA and SA, which is based on the loss of the carboxyl group ( $COO^-$ ). Fragmentation of SA and MMA can be observed in Figure 5-4 (B) and (C). An additional fragment can be observed in the product ion spectrum of SA (Figure 5-4 (B)),  $m/z$  99.0, which corresponds to the loss of  $H_2O$ . The structure of the fragment ion  $m/z$  99 is included as an inset.<sup>20</sup>

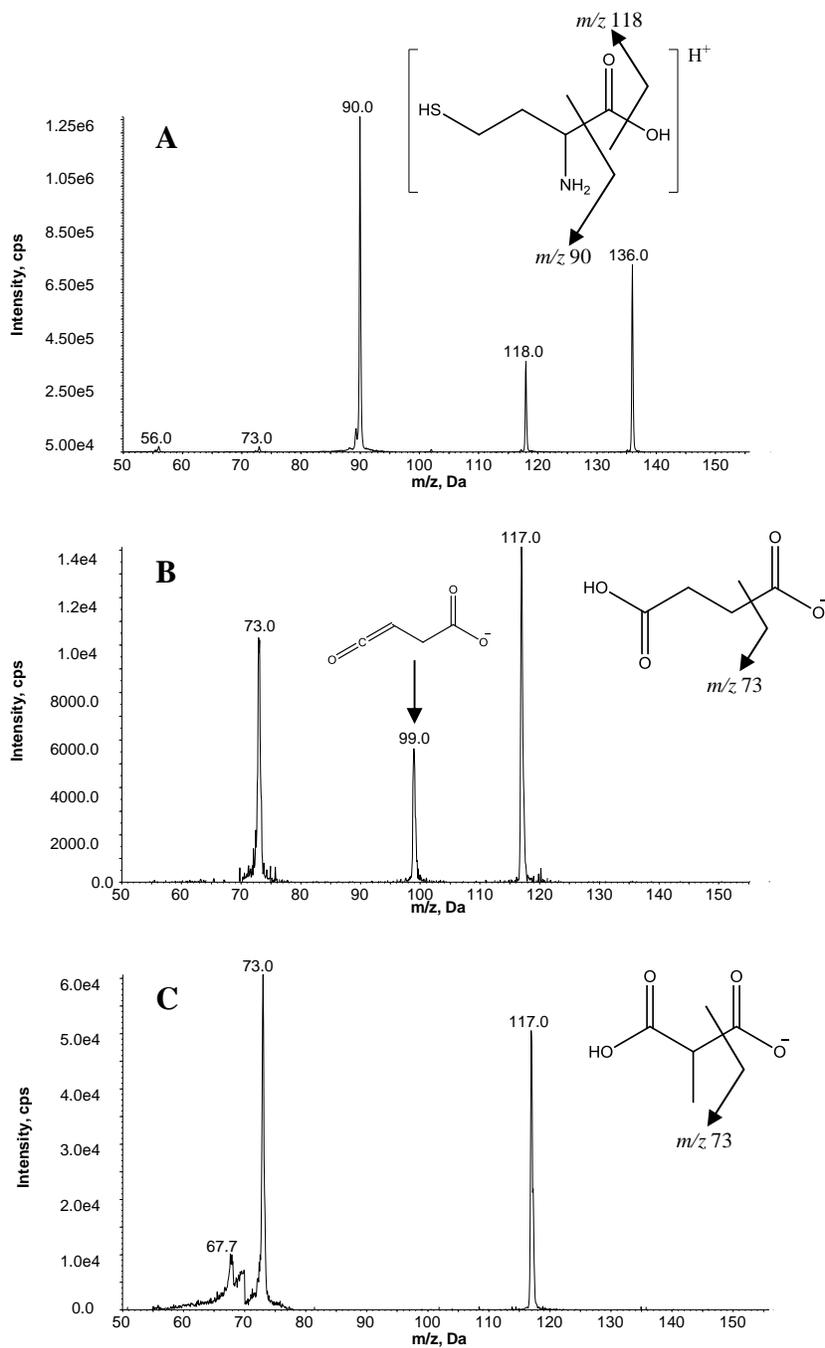


Figure 5-4 Fragmentation and product ion spectra of (A) Homocysteine, (B) Succinic acid and (C) Methylmalonic acid. (Inset in (B) shows the structure of fragment ion  $m/z$  99)

### 5.3.2 Chromatographic Separation

Hcy, MMA and SA are all small polar molecules, which make their retention on a reversed phase column very difficult. Therefore, method development, involving choice of columns and mobile phase optimization remains an important task, in order to achieve the desired separation. Satisfactory retention of the analytes was achieved when using a sub 2  $\mu\text{m}$  particle C18 column and mobile phase 4% acetonitrile in water/ 0.1% formic acid. This column also showed high stability under highly aqueous conditions. All the analytes were eluted in less than 1 minute and the retention times were 0.37 minutes (Hcy), 0.60 minutes (SA) and 0.70 minutes (MMA). Figure 5-5 shows a typical chromatogram of all three analytes at a concentration of 1.0  $\mu\text{M}$ .

### 5.3.3 Flow Rate Optimization for Fast LC

Flow rate affects the LC system pressure, separation quality and analysis time. For small molecules, flow rate is expected to be an important factor for resolution. The optimum flow rates for columns with sub-micron particles are higher than for columns with larger particles (van Deemter equation). From the van Deemter equation, the C-term (and to some extent the A-term) are influenced by particle size. Smaller particles have shorter diffusion path lengths, allowing the analyte to spend less time travelling in and out of the particle, where peak diffusion can occur. Thus the analyte elutes as a narrow peak. As the particle size increases, there is less of an effect of higher flow rates on the efficiency. The ability to use a wider range of higher flow rates translates to a much faster analysis time. With the use of a shorter column, fast LC can be performed without compromising separation quality (or significantly lowering the resolution between peaks of interest). It is important to note that in order to reduce the system volume, all connecting tubing was changed to the lowest diameter available, which is 0.12 mm id tubing. Reducing the volume also reduces the time the solvent takes to reach the head of the column.

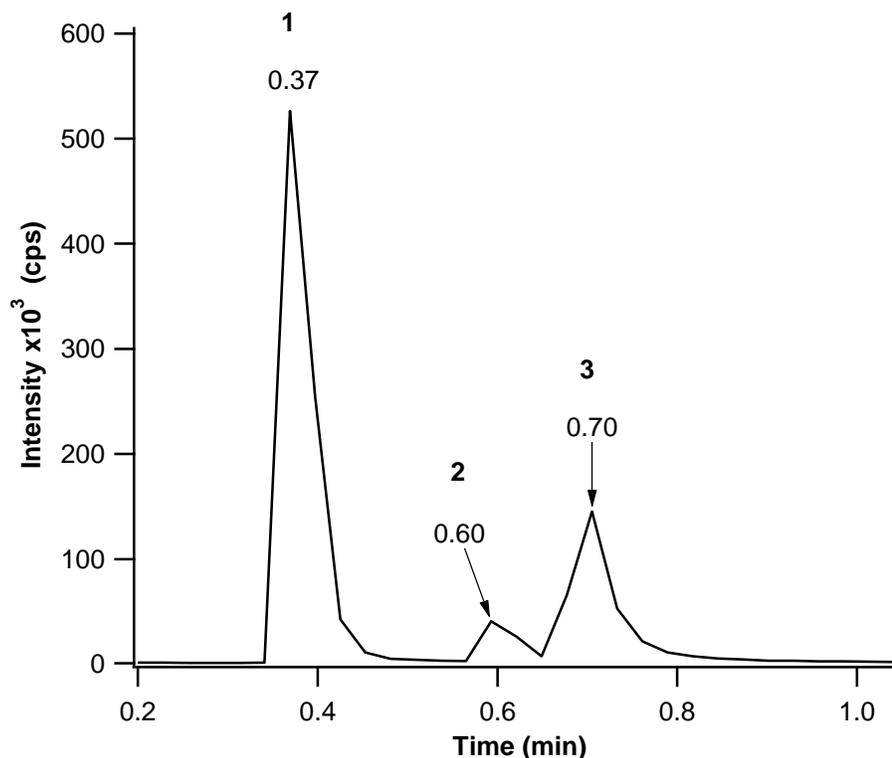


Figure 5-5 Total ion chromatogram showing the retention of (1) Homocysteine, (2) Succinic acid and (3) Methylmalonic acid.

Because the analytes are small and polar, there are not many changes that can be made using a gradient separation. Most reported methods use an isocratic separation.<sup>39</sup> In this work, the flow rate was varied between 0.150 mL/min and 0.500 mL/min and the chromatographic analysis was performed keeping all other conditions constant. Figure 5-6 shows the results of the effects of flow rate on the chromatographic separation, with (A) thru (F) being flow rates of 0.15, 0.20, 0.25, 0.30, 0.4, 0.50 mL/min respectively.

Resolution was calculated for SA and MMA as they were the critical pairs in each of the chromatograms. Resolution was calculated as the difference in retention times of MMA and SA divided by the average peak width of the two

peaks. The resolution as a function of flow rates are plotted and presented in Figure 5-7.

It can be observed that increasing the flow rate from 0.150 mL/min to 0.500 mL/min does not significantly affect the resolution of the chromatographic separation. At the higher flow rates, the resolution of SA and MMA is still close to 1. The peak width also does not change significantly with flow rate. The highest flow rate with a resolution of 1, corresponding to a flow rate of 0.4 mL/min, was chosen as the optimum flow rate for all subsequent runs. In order to keep the method as short as possible, the flow rate for the column wash and equilibration segments was increased to 0.8 mL/min. The optimized method consist of isocratic elution using a mobile phase consisting of 4% acetonitrile in water/ 0.1% formic acid for 1 minute, followed by a quick ramp to 100% acetonitrile/0.1% formic acid (in 0.2 minutes), which was held for 1.5 minutes and equilibration time of 2.5 minutes. The total run time is 5 minutes from injection to injection.

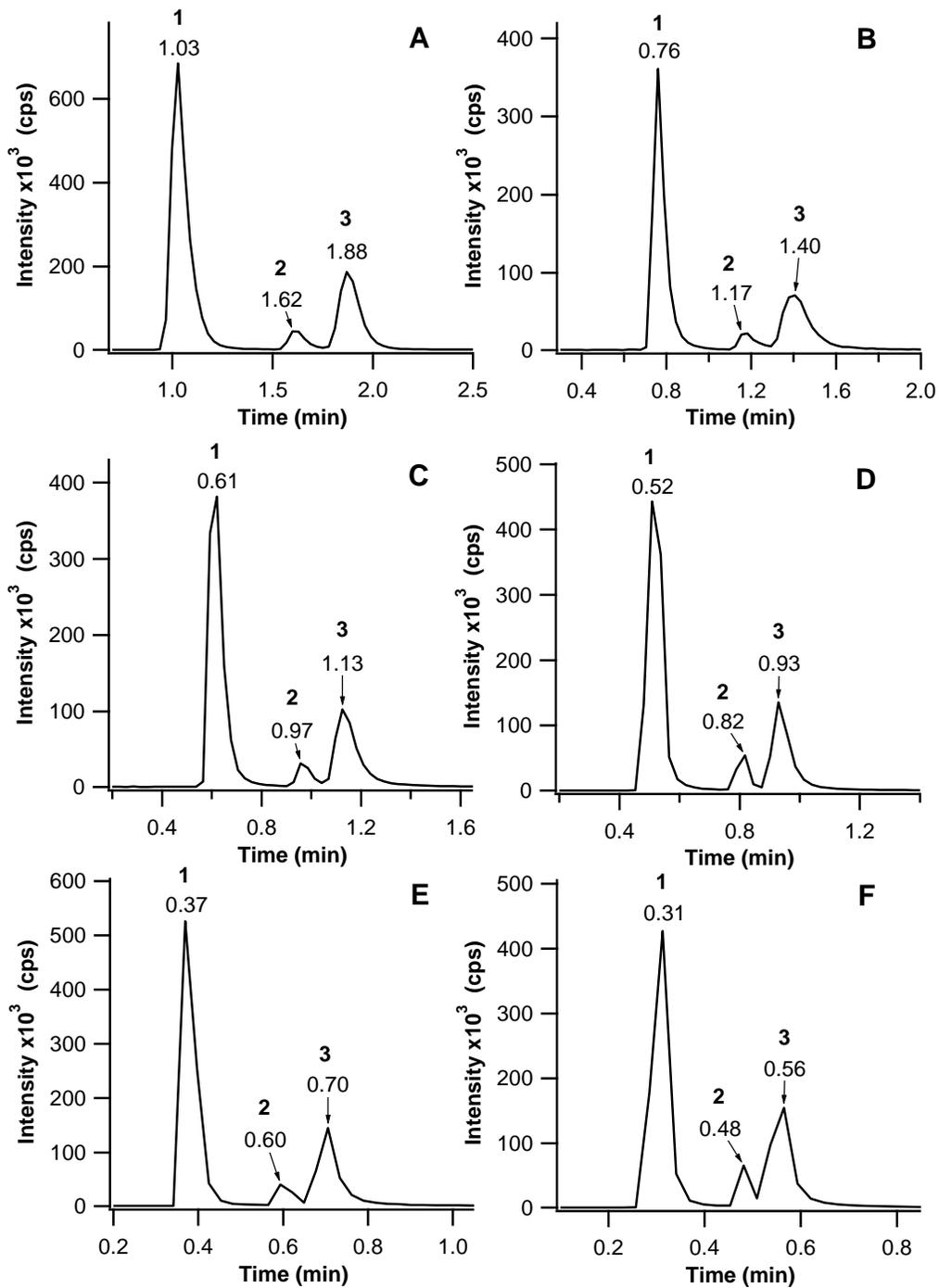


Figure 5-6 Effect of flow rate on the chromatographic separation of peaks (1) Hcy, (2) SA and (3) MMA. The flow rates studied are (A) 0.150 mL/min, (B) 0.200 mL/min, (C) 0.250 mL/min, (D) 0.300 mL/min, (E) 0.400 mL/min, and (F) 0.500 mL/min.

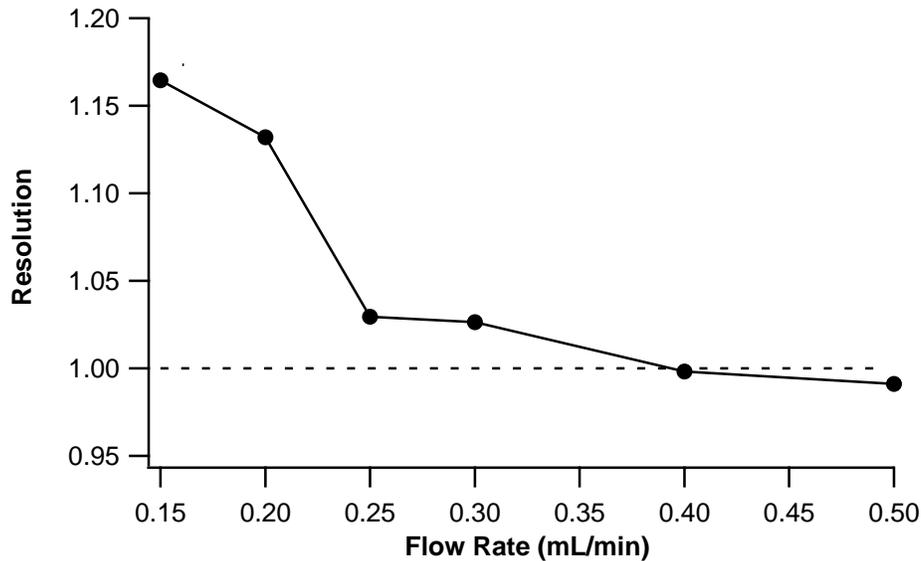


Figure 5-7 Effect of flow rate on resolution between SA and MMA (peaks 1 and 2). Dotted line at  $R_s=1$

### 5.3.4 Effect of MRM Dwell Time

In an MRM experiment, dwell time can be defined as the length of time in which each discrete ion signal is measured.<sup>62</sup> Increasing the dwell time will cause the signal to be averaged over a longer period of time and result in fewer data points per chromatographic peak. In the same way, decreasing the dwell time will increase the sampling rate and result in more data points per peak. As the number of ions being monitored increases, the dwell time should be decreased in order to properly characterize the chromatographic peak. However, decreasing the dwell time can negatively affect signal to noise (S/N) and data quality is affected. In theory, S/N increases with the square root of increased dwell time. A common accepted rule is to obtain 10-15 points per peak and optimizing dwell time should be done to obtain at least 10 data points per chromatographic peak. For analyses

in which only a few analytes are monitored, dwell time is generally not a limitation.

Figure 5-8 shows the effect of dwell time on the number of points across a chromatographic peak. The range was chosen based on the dwell times in methods reported in literature.<sup>39</sup> A dwell time of 50 ms caused a significant decrease in the S/N of the SA and MMA peaks. It may be important to note that this may have been influenced by the polarity switching which significantly increased the duty cycle. As mentioned above, however, there was not much difference in the peaks obtained with dwell times of 100 ms and 200 ms, i.e. dwell time was not a limitation. A dwell time of 100 ms was chosen for all subsequent methods.

### **5.3.5 Succinic Acid and Methylmalonic Acid Analysis**

Since this was a preliminary study, only the analysis of methylmalonic acid was necessary at the initial stage. The determination of Hcy is not as specific as its level is influenced by many other factors, like folate and vitamin B<sub>6</sub> deficiencies and creatinine levels.<sup>5, 63</sup> Moreover, total Hcy determination requires an additional reduction step, reducing homocystine to homocysteine. A representative chromatogram of MMA and SA is shown in Figure 5-9. The chromatographic conditions are same as before; the main difference is that the ESI source was set to perform in negative mode instead of polarity switching mode, using MRM transition for MMA and SA. Dwell time was set at 150 ms. Succinic acid (SA), which is an isomer of MMA was also monitored to guarantee the chromatographic separation of these two isomers in the analysis.

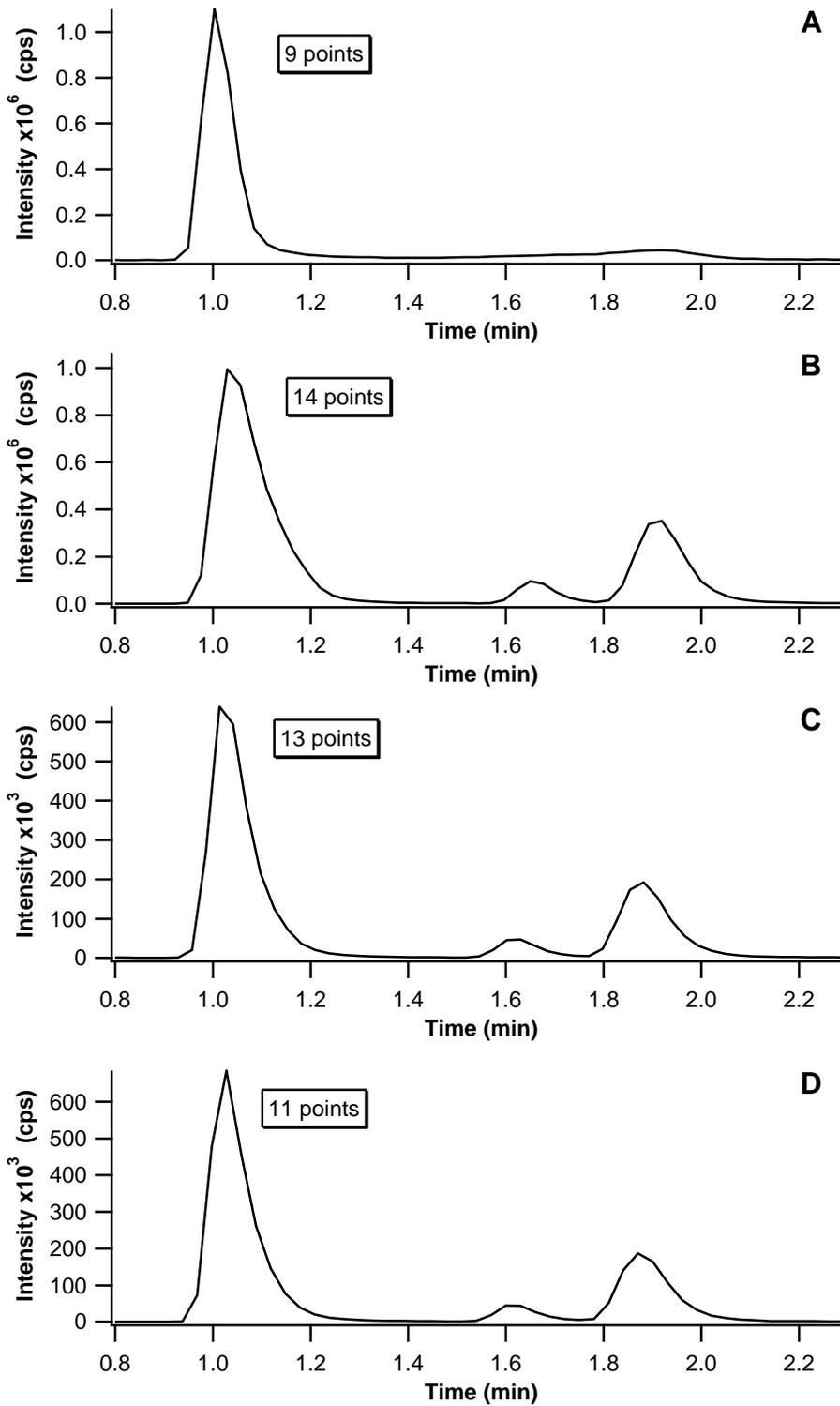


Figure 5-8 Effect of dwell times on the number of points across a chromatographic peak. Chromatogram (A) 50 ms, (B) 100 ms, (C) 150 ms and (D) 200 ms.

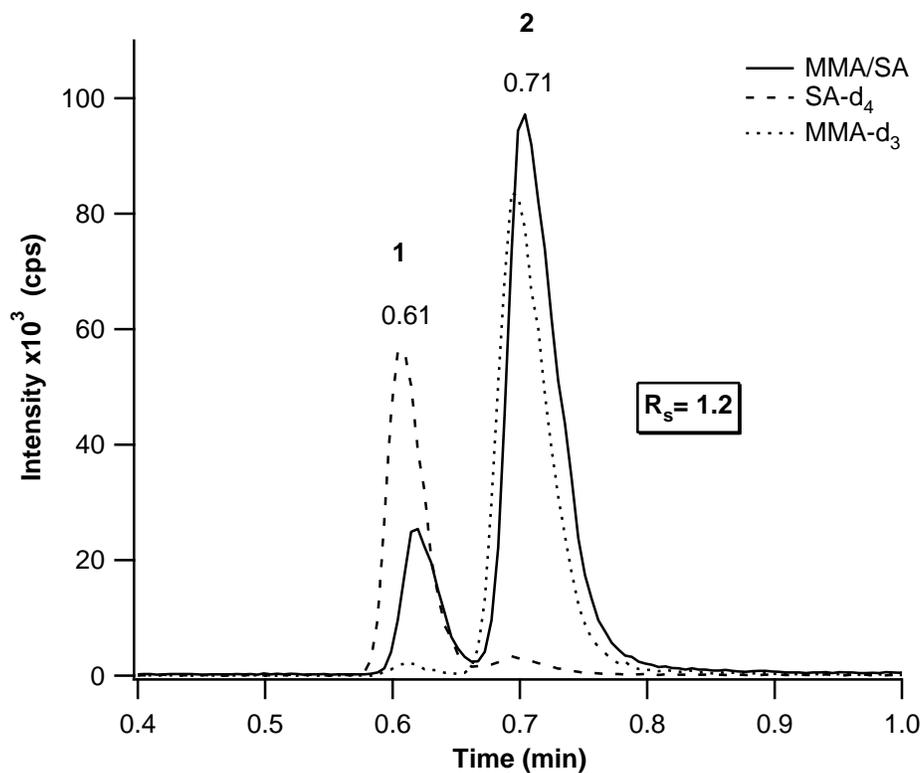


Figure 5-9 TIC showing the retention of (1) SA and SA-d<sub>4</sub>, (2) MMA and MMA-d<sub>3</sub>

### 5.3.6 Calibration and Validation

Calibration standards were prepared by spiking MMA into PBS medium at the following concentrations: 0.050  $\mu\text{M}$ , 0.100  $\mu\text{M}$ , 0.500  $\mu\text{M}$ , 1.0  $\mu\text{M}$ , 5.0  $\mu\text{M}$ , 10.0  $\mu\text{M}$ , 50.0  $\mu\text{M}$ . The linearity of the ESI response for MMA was examined by analyzing triplicates at these 7 different concentration levels. Limit of detection (LOD) and lower limit of quantification (LLOQ) were calculated using the equations  $3 \times (\sigma_B)/m$  and  $10 \times (\sigma_B)/m$  respectively, where  $\sigma_B$  is the standard deviation of the response and  $m$  is the slope derived from the calibration curve. The standard deviation of the blank response was estimated from the regression

parameter, standard deviation of the y-intercept. Figure 5-10 represents the calibration curve of MMA in PBS/propionate medium. The LOD and LLOQ values calculated were 0.016  $\mu\text{M}$  and 0.051  $\mu\text{M}$  respectively. Precision and accuracy were performed by analyzing five replicates of 5 $\mu\text{M}$  (expected concentration in PBS experiments) throughout a sequence. The average concentration measured is 5.008  $\mu\text{M}$ , with accuracy and precision values of 100.16% and 4.89%, respectively.

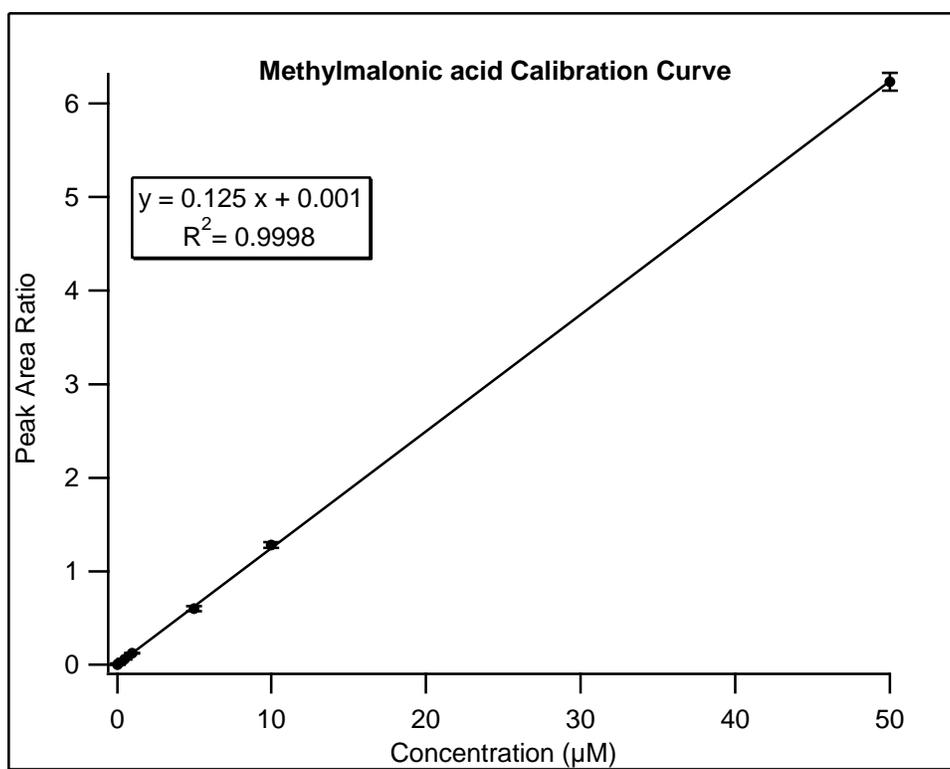


Figure 5-10 Calibration curve of MMA

### 5.3.7 Analysis of Worm Incubation Media

The analysis of MMA in different incubation media was examined. CeHR and the incubation mix are usually the best liquid worm media to use, however, with the presence of salts, nutrients and proteins, detection of MMA may be problematic. Propionate (10 mM) is added to induce elevated production of MMA, which can be easily measured. The expected concentration of excreted MMA in these media is generally in the range of 10-100  $\mu$ M. PBS plus propionate can also be used but this generates a lower concentration of MMA, generally 1-10  $\mu$ M, which is still about a 50-fold increase in the level of MMA in control worms.

Figure 5-11 to Figure 5-13 show the chromatograms of the three different media. Blank media with the absence and presence of propionate are shown in chromatograms (A) and (C) of Figure 5-11 to Figure 5-13. MMA and MMA-d<sub>3</sub> were spiked at concentration 10  $\mu$ M and 5  $\mu$ M respectively into each medium and the chromatograms can be seen in chromatograms (B) and (D) of Figure 5-11 to Figure 5-13. In both incubation medium and CeHR medium, there is an interference in the SA channel at the same retention time of 0.67 minutes. This can be seen in Figure 5-12 and Figure 5-13 (A) and (C) respectively. At this time, the identity of this peak is unknown. The medium chosen for the analysis is PBS because of the much lower background and the absence of any interference in the MMA/SA and MMA-d<sub>3</sub> channels. Since the goal is to develop high-throughput fast LC methods, a medium that does not require extraction or derivatization of the analytes to enhance detection would be desirable.

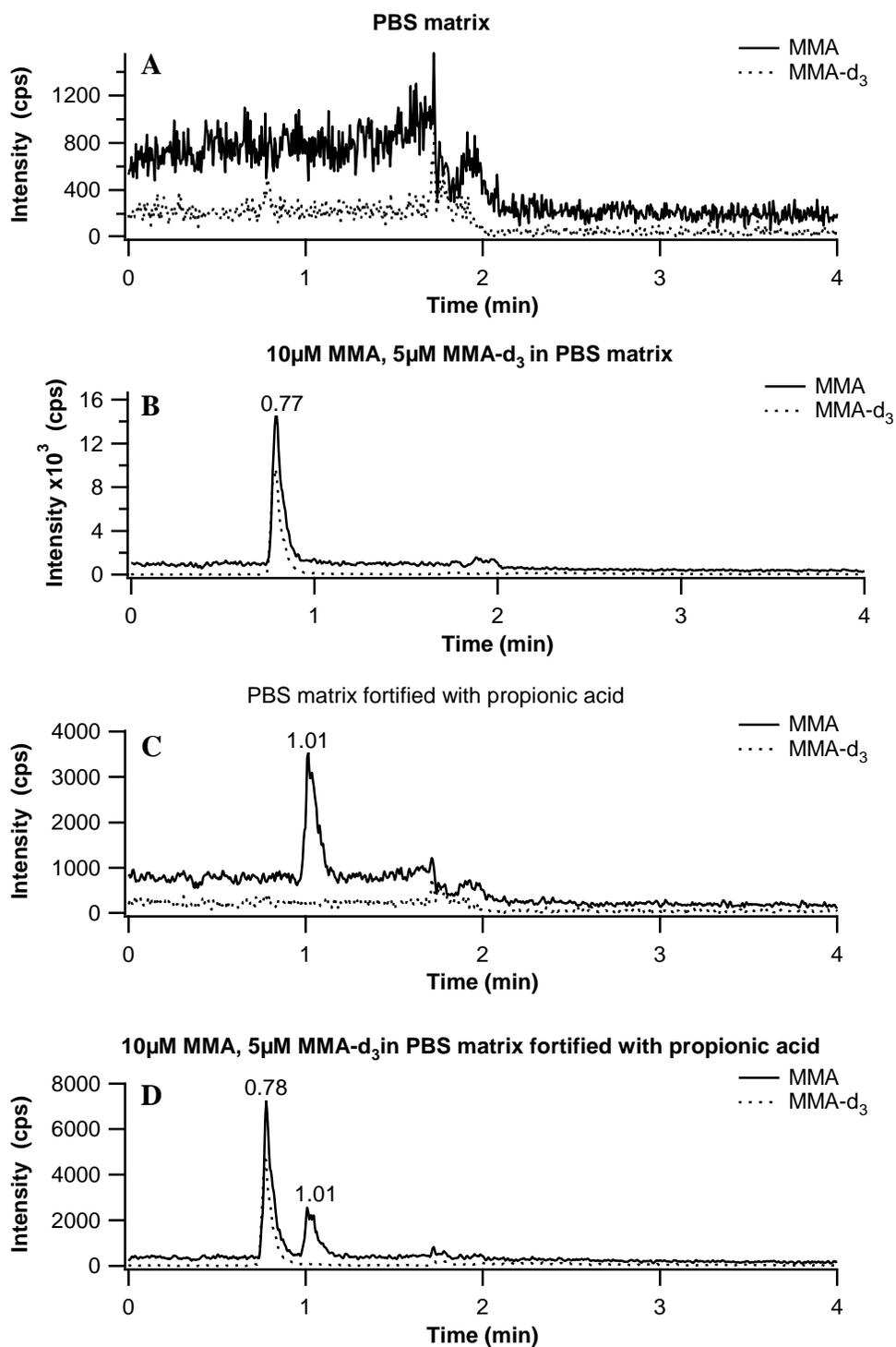


Figure 5-11 Chromatograms showing the analysis of MMA and MMA-d<sub>3</sub> in PBS media

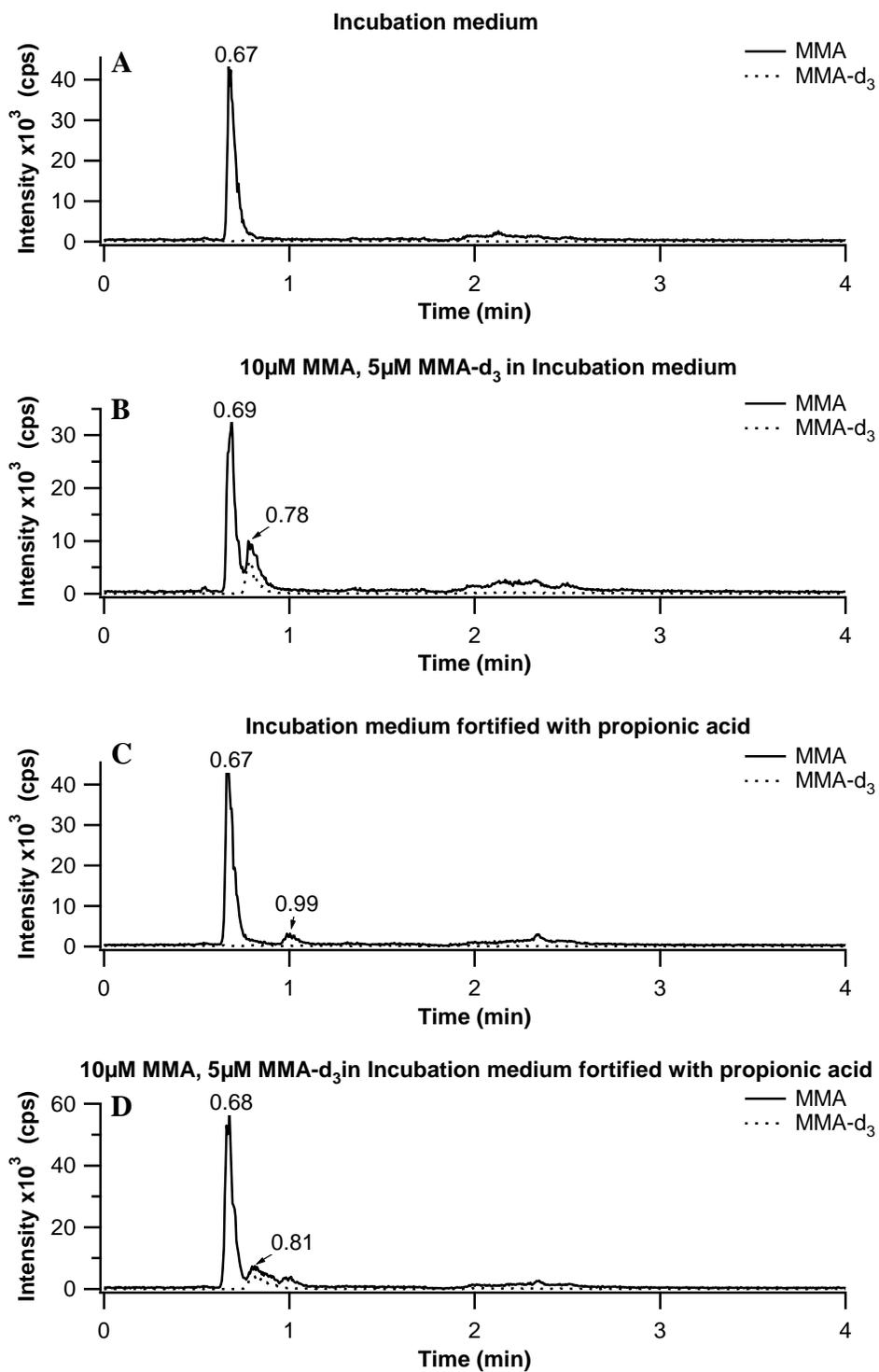


Figure 5-12 Chromatograms showing the analysis of MMA and MMA-d<sub>3</sub> in Incubation media

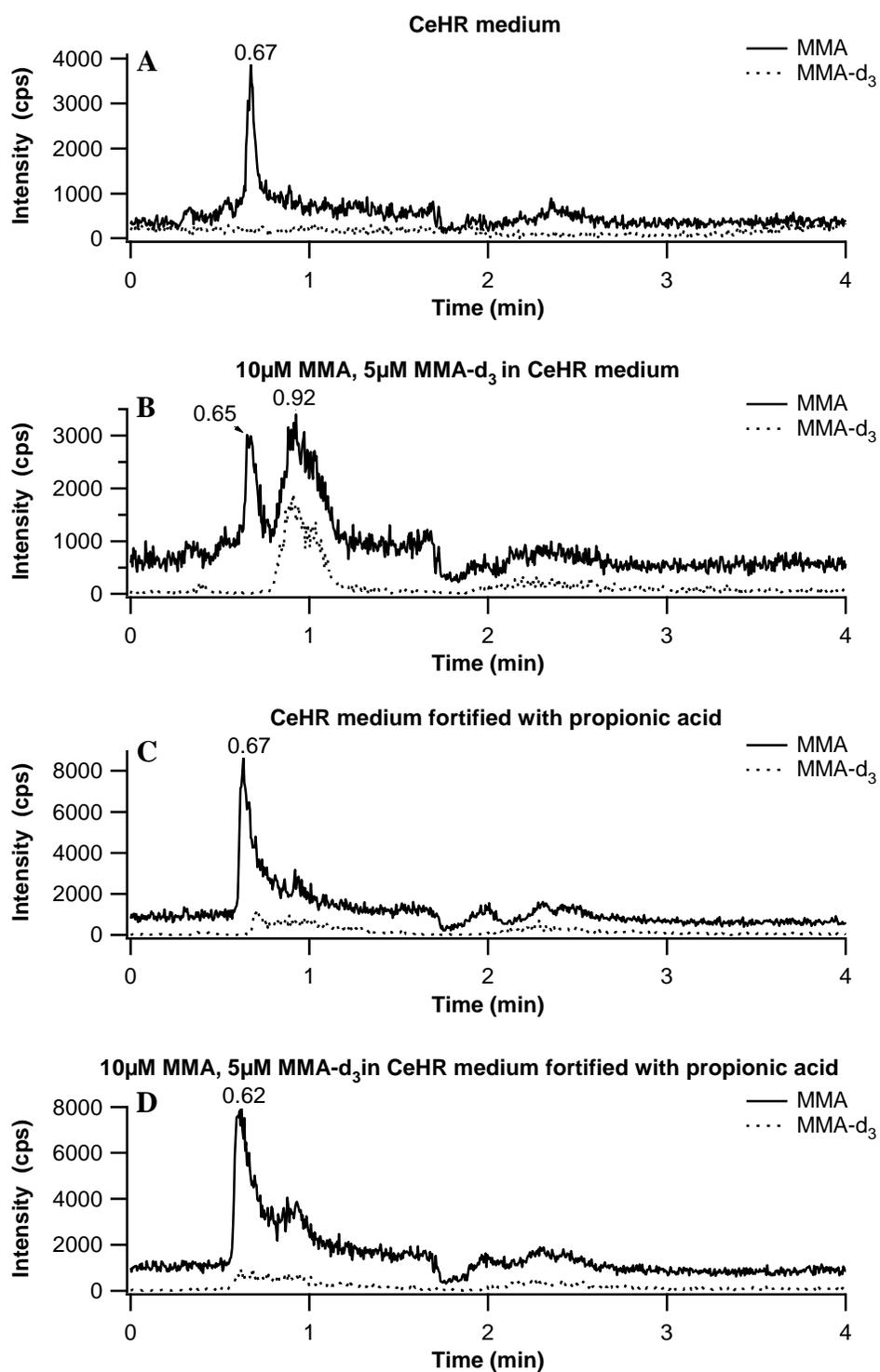


Figure 5-13 Chromatograms showing the analysis of MMA and MMA-d<sub>3</sub> in CeHR media

### 5.3.8 Analysis of *C. Elegans* RNAi Samples

In the preliminary test of RNAi against *mmcm-1*, the concentration of methylmalonic acid increases relative to the wild type. This is illustrated in Figure 5-14. The accumulation of MMA in the supernatant was about four times the levels observed in the wild type nematodes. *mmcm-1* is a gene that codes for the production of methylmalonyl CoA mutase and the RNAi creates a block prior to the formation of succinyl-CoA causing an increase in the concentration and excretion of methylmalonic acid. The values are the average from three independent experiments with the standard deviation indicated by the error bars.

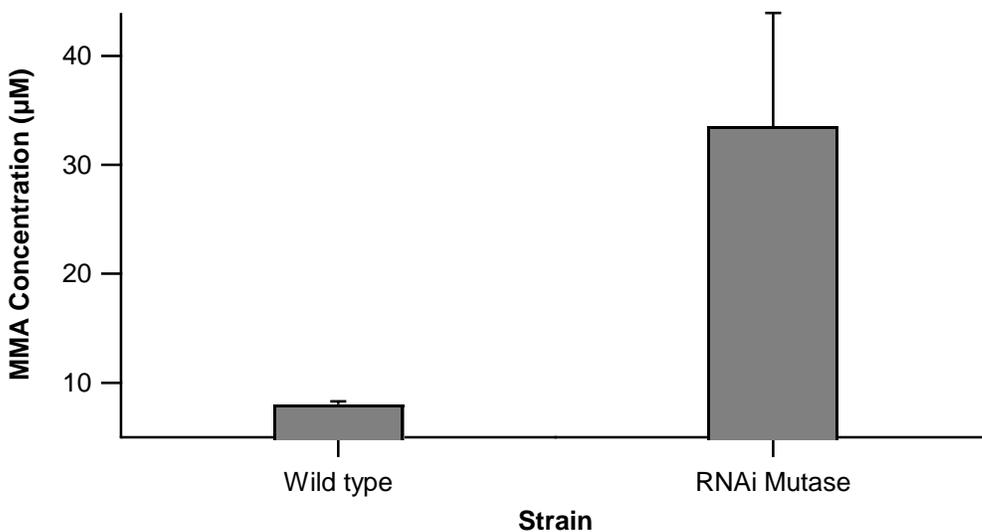


Figure 5-14 Methylmalonic acid detected in the supernatant after RNAi *mmcm-1* in the presence of 200 mM propionate.

### 5.3.9 Analysis of ABC Transporter Deletion Strains

ABC transporters are one of the largest families of transport proteins and are grouped into several subfamilies based on the amino acid sequence and domain organization. They are collectively able to accommodate a large variety of different substrates and this diversity of function is often manifest at the family level as well as in individual members of the subfamily. In *C. elegans*, 61 ABC transporters have been reported grouped into eight subfamilies A-H. In this work, knockouts for the following ABC transporters have been studied: Subfamily A – ABT (homolog ABT-1), Subfamily B – PGP (homologs PGP-4, PGP-5, PGP-6, PGP-7, PGP-8, PGP-10, PGP-12 and PGP-13), Subfamily B – HAF (homolog HAF-2), Subfamily C – MRP (homologs MRP-1, MRP-2, MRP-3, MRP-4, and MRP-6) and Subfamily G – C16C10.12 and WHT-6. All the nematode P-glucoprotein (PGP) examined thus far seem to be expressed in the intestines. The functions of some of these transporter proteins are still not yet known in *C. elegans* and it is the goal of this study to identify those that may be involved in cobalamin metabolism. Figure 5-15 represent the changes in MMA excretion in response to mutations in the ABC transporter strain. MMA excretion is represented as changes in MMA concentration relative to the wild-type worms and is expressed as a log scale. The absent or defective protein in the deletion strain is labeled in italics. Mutations in the gene that codes for methylmalonyl CoA mutase (represented as *mutase*) and the gene that codes for adenosylcobalamin formation (represented as *AdoCbl*) are also shown. These are used to compare the MMA concentration against the MMA concentration of the deletion strains because the effects of these mutations are known. Most of the mutations in PGP transporters studied resulted in a reduced excretion of MMA, except for PGP-10 and PGP-13. Mutations in MRP-1, MRP-2, MRP-3 and MRP-6 show reduced levels of MMA whereas mutations in MRP-4, ABT-2 and WHT-6 resulted in higher levels of MMA excreted.

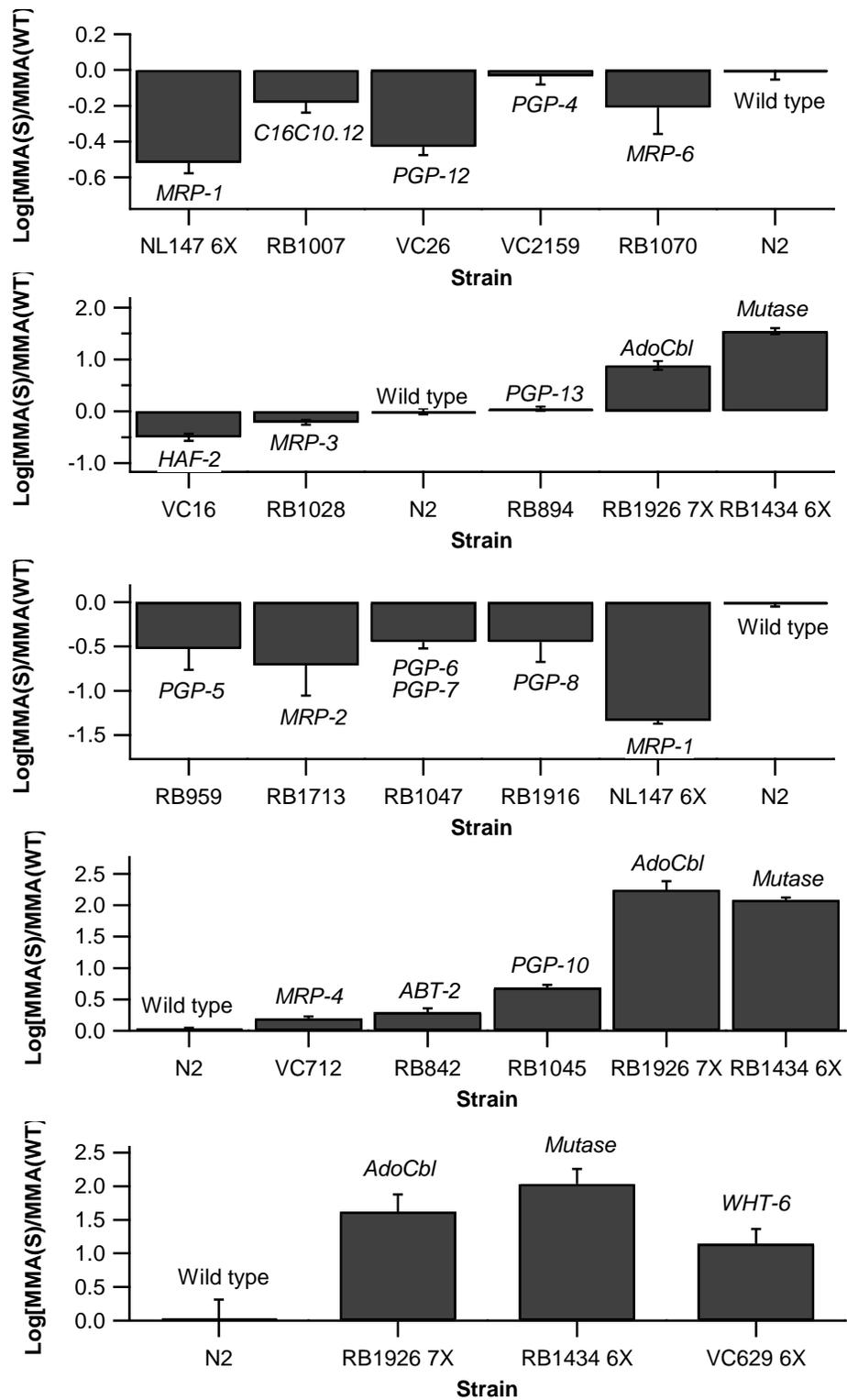


Figure 5-15 Methylmalonic acid accumulation of *C. elegans* ABC transporter deletion strains

In parallel with the LC-MS/MS assay, a  $^{14}\text{C}$ -propionate incorporation assay was also done in Dr. Gravel's laboratory at the University of Calgary. Using this assay, the incorporation of  $^{14}\text{C}$ -propionate into fibroblast cell protein is tested and this provides a measurement of total propionate pathway, involving the conversion of methylmalonyl CoA to succinyl CoA.<sup>64</sup> Incorporation of radioactive propionate into trichloroacetic acid (TCA)-precipitable materials is measured. Defects in the propionate pathway will result in significantly reduced levels of radioactive propionate incorporated into TCA-precipitable materials and the propionate incorporation increases in the presence of cobalamin. The results of the  $^{14}\text{C}$ -propionate incorporation test revealed that the proteins MRP-1, MRP-2 and WHT-6 are involved in cobalamin transport.

MRP-1 and MRP-2 are localized in the *C. elegans* intestinal region and is believed to act redundantly as cobalamin intestinal efflux transporters. Mutations in the genes coding for either MRP-1 or MRP-2 resulted in diminished excretion of MMA in comparison to the wild type worms. It is believed that worms existing on the PBS medium with very little food are quite lethargic and have slowed metabolism. In such a situation, most of the MMA excretion will come from the more active intestines and thus MMA excretion is observed in the wild type worms. On addition of cobalamin to the medium the excretion of MMA is reduced. In the case of MRP-1 and MRP-2 mutations, cobalamin concentration in the intestine is increased and thus the concentration of MMA is reduced. In contrast, mutations in WHT-6, which are also localized in the intestines of *C. elegans*, resulted in an increase in the MMA excretion relative to the wild type nematodes. It is believed that WHT-6 is a lysosomal or mitochondrial transporter but at this stage, more assays are required in order understand the action of these proteins.

## 5.4 Conclusion

A sensitive and high-throughput fast LC method has been developed for the analysis of Hcy, MMA and SA in worm medium. To the best of my knowledge, this is the shortest elution time recorded for the analysis of these three analytes. This method has been successfully applied to the analysis of MMA in PBS worm media. Concerning the RNAi experiments, as indicated by University of Calgary, RNAi against other genes involved in cobalamin metabolism did not produce reliable MMA accumulation or <sup>14</sup>C-propionate incorporation. Preliminarily, the transporter deletion strains experiments suggest that there are multiple ABC transporters involve in cobalamin metabolism. MRP-1 and MRP-2 are thought to act redundantly as intestinal efflux transporters, whereas WHT-6 may be either a lysosomal or mitochondrial transporter. Further studies are being done at the University of Calgary to determine the location, function and mechanism of action of these proteins.

## 5.5 Literature Cited

- (1) Buchanan, J. M. *Medicine* **1964**, *43*, 697-709.
- (2) Bor, M. V.; Nexo, E. In *Vitamins in the Prevention of Human Diseases*; Hermann, W., Obeid, R., Eds.; Walter de Gruyter: Berlin, 2011, pp 187-271.
- (3) Miller, J.; Green, R. In *Handbook of Vitamins, Fourth Edition*; CRC Press, 2007, pp 414-459.
- (4) Andrès, E.; Loukili, N. H.; Noel, E.; Kaltenbach, G.; Abdelgheni, M. B.; Perrin, A. E.; Noblet-Dick, M.; Maloisel, F.; Schlienger, J.-L.; Blicklé, J.-F. In *CMAJ: Canadian Medical Association Journal*; Canadian Medical Association, 2004; Vol. 171, pp 251-259.
- (5) Carmel, R. *Annu. Rev. Med.* **2000**, *51*, 357.
- (6) Karmi, O.; Zayed, A.; Baragethi, S.; Qadi, M.; Ghanem, R. *IIOAB J.* **2011**, *2*, 23-32.

- (7) Matchar, D. B.; McCrory, D. C.; Millington, D. S.; Feussner, J. R. *Am J Med Sci* **1994**, *308*, 276-283.
- (8) Al, A. F.; Al-Hashmi, H.; Mula-Abed, W.-A. *Oman Med J* **2010**, *25*, 9-12.
- (9) Herrmann, W.; Obeid, R.; Schorr, H.; Geisel, J. *Clin. Chem. Lab. Med.* **2003**, *41*, 1478-1488.
- (10) Lee, Y. K.; Kim, H.-S.; Kang, H. J. *Ann. Clin. Lab.Sci.* **2009**, *39*, 361-366.
- (11) Obeid, R.; Herrmann, W. *Clin. Chem. Lab. Med.* **2007**, *45*, 1746-1750.
- (12) Nexo, E.; Hoffmann-Lucke, E. *Am. J. Clin. Nutr.* **2011**, *94*, 359S-365S.
- (13) Hvas, A. M.; Nexo, E. *J Intern Med* **2005**, *257*, 289-298.
- (14) Bhatt, S. P.; Gamit, T. V. *Indian Pharm.* **2007**, *6*, 32-36.
- (15) Lamers, Y. *Curr. Opin. Clin. Nutr. Metab. Care* **2011**, *14*, 445-454.
- (16) Liang, D.; Liang, X.; Wang, Y. *Linchuang Jianyan Zazhi (Nanjing, China)* **2003**, *21*, 122-124.
- (17) Ueland, P. M.; Schneede, J. *Tidsskr Nor Laegeforen* **2008**, *128*, 690-693.
- (18) Pfeiffer, C. M.; Smith, S. J.; Miller, D. T.; Gunter, E. W. *Clin. Chem.* **1999**, *45*, 2236-2242.
- (19) Magera, M. J.; Helgeson, J. K.; Matern, D.; Rinaldo, P. *Clin. Chem.* **2000**, *46*, 1804-1810.
- (20) Kushnir, M. M.; Komaromy-Hiller, G.; Shushan, B.; Urry, F. M.; Roberts, W. L. *Clin. Chem.* **2001**, *47*, 1993-2002.
- (21) Klee, G. G. *Clin. Chem.* **2000**, *46*, 1277-1283.
- (22) Norman, E. J. *J Am Geriatr Soc* **1999**, *47*, 1158-1159.
- (23) Shinohara, Y.; Hasegawa, H.; Kaneko, T.; Tamura, Y.; Hashimoto, T.; Ichida, K. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2010**, *878*, 417-422.
- (24) Kaluzna-Czaplinska, J.; Michalska, M.; Rynkowski, J. *Acta Biochim. Pol.* **2011**, *58*, 31-34.
- (25) Yazdanpanah, M.; Chan, P.-C.; Evrovski, J.; Romaschin, A.; Cole, D. E. *C. Clin. Biochem.* **2003**, *36*, 617-620.

- (26) Lee, J. A.; Jung, B. C., *Method for detecting concentration of homocysteine in blood plasma with gas chromatography-mass spectrometry*. KR773878B1, **2007**, pp 7.
- (27) Ye, Q.; Zheng, D.; Zhu, F. *Anal. Methods* **2010**, 2, 354-358.
- (28) Kushnir, M. M.; Komaromy-Hiller, G. *J. Chromatogr., B: Biomed. Sci. Appl.* **2000**, 741, 231-241.
- (29) Inoue, Y.; Ohse, M. *Anal. Bioanal. Chem.* **2011**, 400, 1953-1958.
- (30) Schneede, J.; Ueland, P. M. *Anal. Chem.* **1995**, 67, 812-819.
- (31) Ivanov, A. V.; Luzyanin, B. P.; Moskovtsev, A. A.; Rotkina, A. S.; Kubatiev, A. A. *J. Anal. Chem.* **2011**, 66, 317-321.
- (32) Marsh, D. B.; Nuttall, K. L. *J. Capillary Electrophor.* **1995**, 2, 63-67.
- (33) Causse, E.; Siri, N.; Bellet, H.; Champagne, S.; Bayle, C.; Valdiguie, P.; Salvayre, R.; Couderc, F. *Clin. Chem.* **1999**, 45, 412-414.
- (34) Franke, D. R.; Marsh, D. B.; Nuttall, K. L. *J. Capillary Electrophor.* **1996**, 3, 125-129.
- (35) Carvalho, V. M.; Kok, F. *Anal. Biochem.* **2008**, 381, 67-73.
- (36) Casetta, B. *Methods Mol. Biol.* **2010**, 603, 253-260.
- (37) Turgeon, C. T.; Magera, M. J.; Cuthbert, C. D.; Loken, P. R.; Gavrilov, D. K.; Tortorelli, S.; Raymond, K. M.; Oglesbee, D.; Rinaldo, P.; Matern, D. *Clin. Chem.* **2010**, 56, 1686-1695.
- (38) Hu, L.-F.; Li, J.-W.; Wang, X.-Q.; Xu, R.-A.; Xu, X.-G.; Jiang, H.-Y.; Zhang, X.-H. *Pharmazie* **2010**, 65, 720-722.
- (39) Hempen, C.; Wanschers, H.; Sluijs, V. G. *Anal. Bioanal. Chem.* **2008**, 391, 263-270.
- (40) Zhang, X.; Ye, X.; Hu, L.; Zhang, N. *Yaowu Fenxi Zazhi* **2010**, 30, 2267-2269.
- (41) Rafii, M.; Elango, R.; House, J. D.; Courtney-Martin, G.; Darling, P.; Fisher, L.; Pencharz, P. B. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2009**, 877, 3282-3291.
- (42) Kirchhoff, F.; Lorenzl, S.; Vogeser, M. *Clin. Chem. Lab. Med.* **2010**, 48, 1647-1650.

- (43) Lakso, H.-A.; Appelblad, P.; Schneede, J. *Clin. Chem.* **2008**, *54*, 2028-2035.
- (44) Persichilli, S.; Gervasoni, J.; Iavarone, F.; Zuppi, C.; Zappacosta, B. *J. Sep. Sci.* **2010**, *33*, 3119-3124.
- (45) Tomaiuolo, M.; Vecchione, G.; Margaglione, M.; Pisanelli, D.; Grandone, E. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2009**, *877*, 3292-3299.
- (46) Li, Y. N.; Gulati, S.; Baker, P. J.; Brody, L. C.; Banerjee, R.; Kruger, W. D. *Human Mol. Genet.* **1996**, *5*, 1851-1858.
- (47) Yamada, K.; Tobimatsu, T.; Toraya, T. *Biosci. Biotechnol. Biochem.* **1998**, *62*, 2155-2160.
- (48) Kasai, S.; Yamazaki, T. *Gene* **2001**, *264*, 281-288.
- (49) Chandler, R. J.; Aswani, V.; Tsai, M. S.; Falk, M.; Wehrli, N.; Stabler, S.; Allen, R.; Sedensky, M.; Kazazian, H. H.; Venditti, C. P. *Mol. Genet. Metab.* **2006**, *89*, 64-73.
- (50) Lerner-Ellis, J. P.; Tirone, J. C.; Pawelek, P. D.; Dore, C.; Atkinson, J. L.; Watkins, D.; Morel, C. F.; Fujiwara, T. M.; Moras, E.; Hosack, A. R.; Dunbar, G. V.; Antonicka, H.; Forgetta, V.; Dobson, C. M.; Leclerc, D.; Gravel, R. A.; Shoubridge, E. A.; Coulton, J. W.; Lepage, P.; Rommens, J. M.; Morgan, K.; Rosenblatt, D. S. *Nature Genet.* **2006**, *38*, 93-100.
- (51) Froese, D. S.; Gravel, R. A. *Expert Rev. Mol. Med.* **2010**, *12*, 1-20.
- (52) Dobson, C. M.; Wai, T.; Leclerc, D.; Kadir, H.; Narang, M.; Lerner-Ellis, J. P.; Hudson, T. J.; Rosenblatt, D. S.; Gravel, R. A. *Human Mol. Genet.* **2002**, *11*, 3361-3369.
- (53) Culetto, E.; Sattelle, D. B. *Human Mol. Genet.* **2000**, *9*, 869-877.
- (54) Schmid, A.; DiDonato, C. J. *J. Child Neurology* **2007**, *22*, 1004-1012.
- (55) Barr, M. M. *Physiol. Genomics* **2003**, *13*, 15-24.
- (56) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. *Nature* **1998**, *391*, 806-811.
- (57) Fire, A.; Albertson, D.; Harrison, S. W.; Moerman, D. G. *Development* **1991**, *113*, 503-514.

- (58) Montgomery, M. K.; Xu, S.; Fire, A. *Proc. Natl. Acad. Sci.* **1998**, *95*, 15502-15507.
- (59) Kamath, R. S.; Martinez-Campos, M.; Zipperlen, P.; Fraser, A. G.; Ahringer, J. *Genome Biol.* **2001**, *2*.
- (60) Sheps, J. A.; Ralph, S.; Zhao, Z.; Baillie, D. L.; Ling, V. *Genome Biol.* **2004**, *5*, R15-R31.
- (61) Beedholm-Ebsen, R.; van, d. W. K.; Hardlei, T.; Nexoe, E.; Borst, P.; Moestrup, S. K. *Blood* **2010**, *115*, 1632-1639.
- (62) Mallet, A.; Down, S. *Dictionary of Mass Spectrometry*; John Wiley & Sons: West Sussex, 2009.
- (63) Monsen, A. L. B.; Ueland, P. M. *Am. J. Clin. Nutr.* **2003**, *78*, 7-21.
- (64) Kakinuma, H.; Kobayashi, A.; Takahashi, H. *Clin. Chim. Acta* **2004**, *343*, 209-212.

## Chapter 6. MyCompoundID: Using an Evidenced-Based Metabolome Library for Metabolite Identification\*

### 6.1 Introduction

Metabolomics is a rapidly evolving field for studying biological systems and discovering potential disease biomarkers.<sup>1, 2</sup> Advance in metabolomics is largely driven by the development of new analytical techniques, such as liquid chromatography mass spectrometry (LC-MS). However, metabolite identification remains a major analytical challenge.<sup>3, 4</sup> The vast majority of spectral features observed in LC-MS cannot be assigned to known compounds.<sup>5-7</sup> This serious deficiency hinders the development of sophisticated bioinformatics tools for integrating the metabolome data with the proteome and transcriptome information for studies in systems biology and medicine. Clearly new tools for metabolite identification are urgently needed.

An accurate mass-MS/MS approach is reported for metabolite identification based on compound library searching. An evidence-based metabolome library (EML) is constructed that is composed of known published metabolites, as well as their possible metabolic products that are predicted by biotransformation reactions commonly encountered in metabolism. The predicted metabolites have indirect evidence of their potential existences in a given species as they are derived from known metabolites and metabolic reactions. The rationale is that a known metabolite can be involved in various metabolic reactions in biological systems, producing different metabolic products. Some of them have been identified and documented with assigned chemical structures, while many others have not been identified. The hypothesis is that, by including

---

\* A version of this chapter has been submitted as Li, Liang; Li, Ronghong; Zhou, Jianjun; Zuniga, Azeret; Stanislaus, Avalyn; Shi, Yi; Wishart, David; Lin, Guohui, 2012, MyCompoundID: Using an Evidence-based Metabolome Library for Metabolite Identification to *Analytical Chemistry*. Avalyn Stanislaus contributed by extracting and analyzing the urine sample and performed data analysis and spectral interpretation of the urine metabolites.

all of the possible metabolic products in the library, many unknowns that are structurally related to the known metabolites can potentially be identified using the accurate mass-MS/MS approach. This approach is illustrated by human metabolite identification. The 8,021 entries found in the Human Metabolome Database (HMDB)<sup>8</sup> is used to create the EML. This EML is then applied for identification of metabolites present in simple extracts of human urine and plasma and demonstrated the possibility of identifying many more metabolites than the conventional approach of using the standard HMDB.

## **6.2 Experimental**

### **6.2.1 Chemicals and Materials**

Optima grade methanol and water, Optima LCMS grade acetonitrile were purchased from Fisher Scientific (Ottawa, ON, Canada). HPLC-grade formic acid was obtained from Fluka (Milwaukee, OH, USA).

### **6.2.2 Construction of the EML**

There are currently 8,021 entries in the Human Metabolome Database (HMDB).<sup>8</sup> These entries were used to create the evidence-based metabolome library. By examining the literature information, 76 commonly encountered metabolic reactions were identified and are represented in Table 6-1. Based on these reactions, *in silico* biotransformation of the 8,021 known metabolites were performed. Each reaction generates a product with the addition or subtraction of an expected group (e.g., +O in oxidation or -O in de-oxidation) from the reactant, a known metabolite. Several possible structures of the product (isomers) could exist, but all with a characteristic mass shift from the added or subtracted group. The number of the new entries in the EML with one metabolic reaction is 375,809; some of the impossible transformations (e.g., -O from a metabolite

containing no oxygen) were excluded during the construction of the library. Currently there is also an option of generating the library with two metabolic reactions [e.g., a metabolite undergoes methylation (+CH<sub>2</sub>) and then oxidation (+O), or a metabolite undergoes demethylation (-CH<sub>2</sub>) and then oxidation (+O)], which produced a library with 10,583,901 entries.

Table 6-1 List of the 76 most common metabolic reactions (biotransformations)

#	Reaction	Mass Difference (Da)	Description
1	-H <sub>2</sub>	-2.015650	dehydrogenation
2	+H <sub>2</sub>	2.015650	hydrogenation
3	-CH <sub>2</sub>	-14.015650	demethylation
4	+CH <sub>2</sub>	14.015650	methylation
5	-NH	-15.010899	loss of NH
6	+NH	15.010899	addition of NH
7	-O	-15.994915	loss of oxygen
8	+O	15.994915	Oxidation
9	-NH <sub>3</sub>	-17.026549	loss of ammonia
10	+NH <sub>3</sub>	17.026549	addition of ammonia
11	-H <sub>2</sub> O	-18.010565	loss of water
12	+H <sub>2</sub> O	18.010565	addition of water
13	-CO	-27.994915	loss of CO
14	+CO	27.994915	addition of CO
15	-C <sub>2</sub> H <sub>4</sub>	-28.031300	loss of C <sub>2</sub> H <sub>4</sub>
16	+C <sub>2</sub> H <sub>4</sub>	28.031300	addition of C <sub>2</sub> H <sub>4</sub>
17	-C <sub>2</sub> H <sub>2</sub> O	-42.010565	Deacetylation
18	+C <sub>2</sub> H <sub>2</sub> O	42.010565	Acetylation
19	-CO <sub>2</sub>	-43.989830	loss of CO <sub>2</sub>
20	+CO <sub>2</sub>	43.989830	addition of CO <sub>2</sub>
21	SO <sub>3</sub> H->SH	-47.984745	Sulfonic acid to Thiol
22	SH->SO <sub>3</sub> H	47.984745	Thiol to Sulfonic acid
23	-C <sub>2</sub> H <sub>3</sub> NO	-57.021464	loss of glycine
24	+C <sub>2</sub> H <sub>3</sub> NO	57.021464	glycine conjugation
25	-SO <sub>3</sub>	-79.956817	loss of sulfate
26	+SO <sub>3</sub>	79.956817	sulfate conjugation

27	-HPO <sub>3</sub>	-79.966333	loss of Phosphate
28	+HPO <sub>3</sub>	79.966333	addition of Phosphate
29	-C <sub>4</sub> H <sub>3</sub> N <sub>3</sub>	-93.032697	loss of Cytosine
30	+C <sub>4</sub> H <sub>3</sub> N <sub>3</sub>	93.032697	addition of Cytosine
31	-C <sub>4</sub> H <sub>2</sub> N <sub>2</sub> O	-94.016713	loss of Uracil
32	+C <sub>4</sub> H <sub>2</sub> N <sub>2</sub> O	94.016713	addition of Uracil
33	-C <sub>3</sub> H <sub>5</sub> NOS	-103.009186	loss of cysteine
34	+C <sub>3</sub> H <sub>5</sub> NOS	103.009186	cysteine conjugation
35	-C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub> S	-107.004101	loss of taurine
36	+C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub> S	107.004101	taurine conjugation
37	-C <sub>5</sub> H <sub>4</sub> N <sub>2</sub> O	-108.032363	loss of Thymine
38	+C <sub>5</sub> H <sub>4</sub> N <sub>2</sub> O	108.032363	addition of Thymine
39	-(C <sub>5</sub> H <sub>5</sub> N <sub>5</sub> - H <sub>2</sub> O)	-117.043930	loss of Adenine
40	+(C <sub>5</sub> H <sub>5</sub> N <sub>5</sub> - H <sub>2</sub> O)	117.043930	addition of Adenine
41	-C <sub>3</sub> H <sub>5</sub> NO <sub>2</sub> S	-119.004101	loss of S-cysteine
42	+C <sub>3</sub> H <sub>5</sub> NO <sub>2</sub> S	119.004101	S-cysteine conjugation
43	-C <sub>5</sub> H <sub>8</sub> O <sub>4</sub>	-132.042260	loss of D-ribose
44	+C <sub>5</sub> H <sub>8</sub> O <sub>4</sub>	132.042260	addition of D-ribose
45	-C <sub>5</sub> H <sub>3</sub> N <sub>5</sub>	-133.038845	loss of Guanine
46	+C <sub>5</sub> H <sub>3</sub> N <sub>5</sub>	133.038845	addition of Guanine
47	-C <sub>7</sub> H <sub>13</sub> NO <sub>2</sub>	-143.094629	loss of Carnitine
48	+C <sub>7</sub> H <sub>13</sub> NO <sub>2</sub>	143.094629	addition of Carnitine
49	-C <sub>5</sub> H <sub>7</sub> NO <sub>3</sub> S	-161.014666	loss of N-acetyl-S-cysteine
50	+C <sub>5</sub> H <sub>7</sub> NO <sub>3</sub> S	161.014666	addition of N-acetyl-S-cysteine
51	-C <sub>6</sub> H <sub>10</sub> O <sub>5</sub>	-162.052825	loss of Hexose
52	+C <sub>6</sub> H <sub>10</sub> O <sub>5</sub>	162.052825	addition of Hexose
53	-C <sub>6</sub> H <sub>8</sub> O <sub>6</sub>	-176.032090	loss of glucuronic acid
54	+C <sub>6</sub> H <sub>8</sub> O <sub>6</sub>	176.032090	addition of glucuronic acid
55	-C <sub>10</sub> H <sub>12</sub> N <sub>2</sub> O <sub>4</sub>	-224.079708	loss of Thymidine
56	+C <sub>10</sub> H <sub>12</sub> N <sub>2</sub> O <sub>4</sub>	224.079708	addition of Thymidine
57	-C <sub>9</sub> H <sub>11</sub> N <sub>3</sub> O <sub>4</sub>	-225.074957	loss of Cytidine
58	+C <sub>9</sub> H <sub>11</sub> N <sub>3</sub> O <sub>4</sub>	225.074957	addition of Cytidine
59	-C <sub>9</sub> H <sub>10</sub> N <sub>2</sub> O <sub>5</sub>	-226.058973	loss of Uridine
60	+C <sub>9</sub> H <sub>10</sub> N <sub>2</sub> O <sub>5</sub>	226.058973	addition of Uridine
61	-C <sub>16</sub> H <sub>30</sub> O	-238.229665	loss of Palmitic acid
62	+C <sub>16</sub> H <sub>30</sub> O	238.229665	addition of Palmitic acid
63	-C <sub>6</sub> H <sub>11</sub> O <sub>8</sub> P	-242.019158	loss of Glucose-6-phosphate

64	+C <sub>6</sub> H <sub>11</sub> O <sub>8</sub> P	242.019158	addition of Glucose-6-phosphate
65	-C <sub>10</sub> H <sub>11</sub> N <sub>5</sub> O <sub>3</sub>	-249.086190	loss of Adenosine
66	+C <sub>10</sub> H <sub>11</sub> N <sub>5</sub> O <sub>3</sub>	249.086190	addition of Adenosine
67	-C <sub>10</sub> H <sub>11</sub> N <sub>5</sub> O <sub>4</sub>	-265.081105	loss of Guanosine
68	+C <sub>10</sub> H <sub>11</sub> N <sub>5</sub> O <sub>4</sub>	265.081105	addition of Guanosine
69	-C <sub>10</sub> H <sub>15</sub> N <sub>3</sub> O <sub>5</sub> S	-289.073244	loss of Glutathione
70	+C <sub>10</sub> H <sub>15</sub> N <sub>3</sub> O <sub>5</sub> S	289.073244	addition of Glutathione
71	-C <sub>10</sub> H <sub>15</sub> N <sub>3</sub> O <sub>6</sub> S	-305.068159	loss of S-Glutathione
72	+C <sub>10</sub> H <sub>15</sub> N <sub>3</sub> O <sub>6</sub> S	305.068159	addition of S-Glutathione
73	-C <sub>12</sub> H <sub>20</sub> O <sub>10</sub>	-324.105650	loss of di-hexose
74	+C <sub>12</sub> H <sub>20</sub> O <sub>10</sub>	324.105650	addition of di-hexose
75	-C <sub>18</sub> H <sub>30</sub> O <sub>15</sub>	-486.158475	loss of tri-hexose
76	+C <sub>18</sub> H <sub>30</sub> O <sub>15</sub>	486.158475	addition of tri-hexose

### 6.2.3 Web-service Design and Implementation

In MyCompoundID (MCI), all known human endogenous metabolites are imported from the Human Metabolome Database and stored in a local MySQL database. These metabolites and their one- or two-reaction products are indexed using the molecular masses up to the millionth precision. The web server for MCI was constructed within Apache using Java and JavaScript to ensure the most efficiency and the largest platform compatibility. The 76 commonly encountered metabolic reactions are implemented in the web server, which accepts single and batch queries with 0, 1 and 2 allowed metabolic reactions. The subtraction reactions are logically validated using the compound's MOL files. All query results are prepared for easier manual inspection that includes a ChemDraw Plugin. The web server interacts with the local computer to allow the users to exclude any output entry and to associate an output entry to any experimental evidence. Such post-curated query results can then be exported to a local archive. All these functions are enabled and efficiently executed in Java and JavaScript, with extendibility for further development.

#### 6.2.4 Web Interface

To use the EML library for metabolite identification, a web-based search and data interpretation program called MyCompoundID ([www.mycompoundid.org](http://www.mycompoundid.org)) was developed. In the workflow (see Figure 6-1), both MS and MS/MS spectra of a metabolome sample are generated using one or more high performance mass spectrometers, such as Fourier transform (FT) MS, time-of-flight (TOF) MS and quadrupole linear trap (QTrap) tandem MS. In MyCompoundID, the user enters a mass (either a single value or multiple values in batch mode) and a mass tolerance value determined by the mass accuracy of the instrument used, and then selects the reaction number (0, 1, or 2). The program searches the EML to find any matches between library entries and the query mass within the defined mass tolerance. The search result is displayed in an interactive table and the matched entries can be sorted (e.g., based on the order of mass error).

One important functionality of the program is that the user can upload the chemical structure of the parent metabolite into ChemDraw or a free-ware ChemDraw Plugin. Both ChemDraw and ChemDraw Plugin allow the user to add or subtract a reaction group in the uploaded structure to create a new structure. Furthermore, the user can use the Mass Fragmentation tool therein to break the chemical bond(s) to generate fragment ion structures and masses. Using the experimental MS/MS spectrum produced from the precursor ion of the query mass, the user can examine the spectral fragmentation pattern and compare it to the fragment ions generated by the Mass Fragmentation tool. If the pattern matches, putative metabolite identification can be made on the query mass. To document the identification process, all metadata, including the structure of the proposed match, the experimental MS/MS spectrum, fragment ion structures, fragmentation pathways, and any other documents (e.g., a word file to describe the process), can be saved to the matched entry. Finally, the results can be exported to a spreadsheet for presentation and other uses. A tutorial for the use of

the program and an example of the process described above are given in the Electronic Appendix.

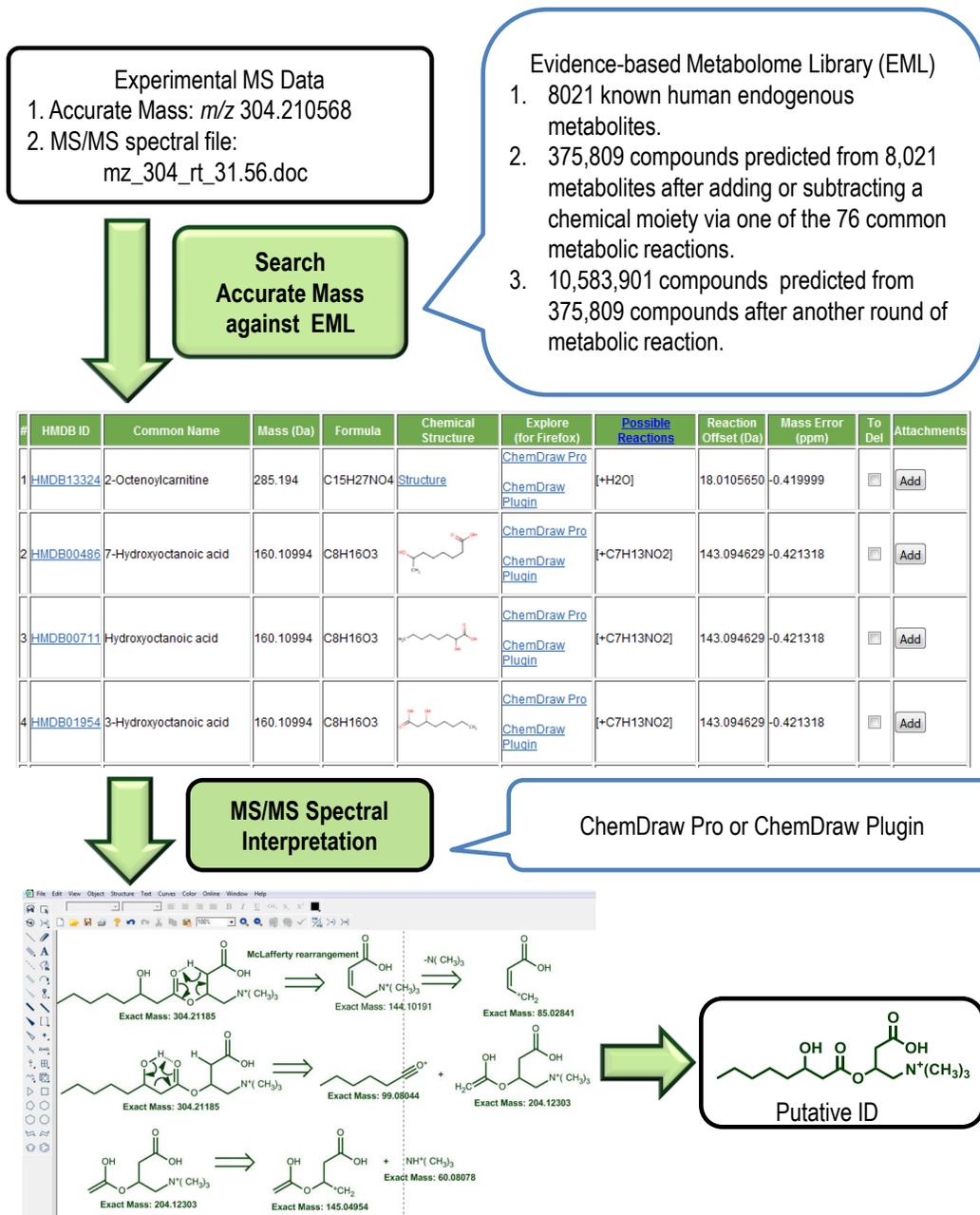


Figure 6-1 Experimental Workflow

### **6.2.5 Samples**

This study was conducted in accordance with the Arts, Science & Law Research Ethics Board policy at the University of Alberta. Urine and plasma samples were obtained from a healthy adult female. The urine sample was collected as first morning void samples and was subsequently centrifuged, aliquoted and stored at -20 °C without any added preservatives until further analysis. The blood sample was collected using K<sub>2</sub>EDTA as the anticoagulant and plasma produced by centrifuging the whole blood samples at 4000g at 4 °C for 20 minutes. The plasma supernatant was collected and aliquoted and stored at -80 °C until further analysis.

### **6.2.6 Solid Phase Extraction (SPE)**

Plasma and urine samples were extracted by SPE using Waters Oasis HLB cartridges (volume, 3 mL; sorbent weight, 60 mg; particle size, 3 µm), operated at a flow rate of 1 mL/min using a vacuum manifold (Alltech from Fisher Scientific, Ottawa, ON). The cartridges were preconditioned with 1mL methanol followed by equilibration with 1 mL water. A 1-mL volume of sample (either urine or plasma) was then passed through the cartridges and the cartridges were washed with 1mL water. The compounds in the samples were eluted from the cartridges with 1 mL methanol. The extracts were then evaporated to dryness in a Savant SpeedVac concentrator system (Global Medical Instrumentation or GMI, Ramsey, Minnesota). The dried urine extract was reconstituted in 100 µL mobile phase A (4% acetonitrile, 0.1% formic acid in water) and the plasma extract was reconstituted in 40% acetonitrile, 0.1% formic acid in water.

### 6.2.7 Liquid Chromatography Mass Spectrometry (LC-MS)

LC-MS and LC-MS/MS analyses were performed on an QTRAP<sup>®</sup> 4000 hybrid triple quadrupole linear ion trap mass spectrometer (Applied Biosystems, Foster City, CA) as well as a 6220 oa-TOF (Agilent Technologies, Santa Clara, CA), each equipped with a 1200 series High Performance Liquid Chromatography system (Agilent Technologies, Santa Clara, CA). The Agilent 1200 series was equipped with an autosampler with refrigerated sample compartment, a binary pump with online degasser and a heated column compartment. Chromatographic separation was achieved using mobile phase A, consisting of 4% acetonitrile, 0.1% formic acid in water and mobile phase B, consisting of 0.1% formic acid in acetonitrile. The method consists of an initial 0% B maintained for 10 minutes, followed by a linear increase to 80% B in 40 minutes, then increased to 100% B in 5 minutes, kept constant at 100% B for 5 minutes and switched back to the initial conditions in 20 minutes. Five microliters of each sample was injected and flow rate was 100  $\mu$ L/min. A Waters BEH (ethylene bridged hybrid) 2.1 x 50mm, 1.7 $\mu$ m C<sub>18</sub> column was used for the separation at room temperature.

The QTRAP<sup>®</sup> 4000 was equipped with a turbo ion spray and operated in positive electrospray mode. The instrument was calibrated weekly with polypropylene glycol (PEG) and ESI tuning mix (AB Sciex) for mass accuracy and resolution. The software Analyst, version 1.5, was used for data acquisition and processing. Optimizing of the mass spectrometric conditions were performed by infusing a 0.1  $\mu$ M of acylcarnitine standard solution dissolved in 50% acetonitrile in water, 0.1 % formic acid at a flow rate of 10  $\mu$ L/min. MS/MS spectra of the metabolites were obtained in IDA mode, using enhanced mass (EMS) survey scan and 4 dependent enhanced product ion (EPI) scans. The EMS survey scan was conducted at mass range 50 – 1000 amu at a scan rate of 1000 amu/s. The following optimized source parameters were used: curtain gas, 10; CAD gas, high; ion source, 4800V; temperature, 250; gas 1, 25; gas 2, 15; declustering

potential (DP), 50; collision energy (CE), 5; (parameters are unit-less). The IDA threshold was set at 500 counts per second (cps). The EPI scan rate was 4000 amu/s and the scan range set at 50 – 1000 amu. Source parameters were identical to the EMS scan and CE was set at 27 eV with a CE spread of 5 eV.

The TOF-MS analysis was performed using full scan mode in the mass range  $m/z$  50 – 1000 in positive ion mode. The conditions of the ESI source are as follows: drying gas ( $N_2$ ) flow rate, 9.9 L/min; drying gas temperature, 325 °C; capillary voltage, 3200 V; nebulizer gas, 20 psig; sheath gas temperature, 40 °C; fragmentor, 4 V, skimmer voltage, 63 V; and octopole RF, 250 V. Prior to the analysis of the samples, the instrument was calibrated daily. Data acquisition and processing were performed using the Mass Hunter software version B.03.01.

### **6.2.8 Data Processing**

The QTRAP<sup>®</sup> IDA mass list from each sample, containing retention time information, was manually extracted to an Excel spreadsheet and compared to the retention time data obtained from the TOFMS instrument. Data files from the TOFMS were converted to mzdata files and processed using publicly available XCMS and R software. XCMS was downloaded from Scripps (<http://metlin.scripps.edu/download/>). The mzdata files were processed using the following optimized XCMS parameters: fwhm=30, step=0.005 and sn=2. These parameters yielded highest mass accuracy of the analytes in the standard solution. After XCMS processing, the data was exported into a Microsoft Excel spreadsheet. In this way, accurate mass and fragmentation information could be obtained for each feature detected.

### 6.3 Results and Discussion

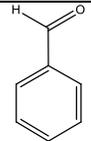
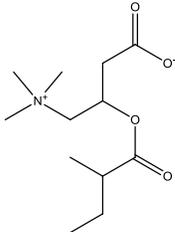
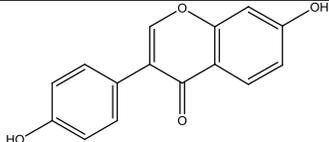
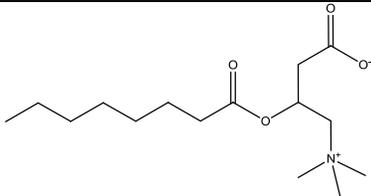
To demonstrate the utility of MyCompoundID combined with our EML for metabolite identification, LC-TOF-MS and LC-QTRAP-MS/MS data were acquired from human urine and plasma samples. Using a simple extraction to capture a small fraction of the metabolome, LC-TOF-MS and LC-QTRAP-MS/MS detected 17,969 and 2,316 features, respectively, in urine and 5,761 and 2,247 features, respectively, in plasma. Out of these features, 325 peaks were extracted in urine and 116 in plasma that were commonly detected by TOF-MS, and QTRAP-MS/MS. The common individual peaks detected by both methods had similar retention times. These metabolite peaks are listed in Supplementary Tables S6-1 and S6-2, available in the Electronic Appendix. To identify these metabolites, the HMDB was first searched using the accurate masses (<5 ppm) and MS/MS spectra against a library of about 900 metabolite standards.<sup>8</sup> Only 8 metabolites were matched in urine and 7 in plasma (see Table 6-2 and Table 6-3). This low rate of success reflects the current status of metabolite identification by LC-MS, i.e., many peaks detected cannot be readily identified using the current database resources.<sup>3, 8-11</sup>

Next, MyCompoundID was used to search the accurate masses of the remaining features against the 8021 known metabolites (i.e., with reaction = 0) to generate a list of mass-matches, followed by MS/MS spectral interpretation of individual matches. Fourteen metabolites were putatively identified in urine and 34 in plasma (see Tables A6-1 and A6-2 and their corresponding Evidence Folders detailing the spectral interpretations of the matches found in the Appendix and Electronic Appendix respectively).

MyCompoundID was then used to search the accurate masses of the remaining features against EML with one reaction. In conjunction with MS/MS spectral interpretation, 41 metabolites were putatively identified in urine and 14 in plasma (see Tables A6-3 and A6-4 and their corresponding Evidence Folders detailing the spectral interpretations of the matches found in the Appendix and

Electronic Appendix respectively). The use of EML with two reactions only led to the tentative identification of 3 more metabolites in urine and none in plasma (see Tables A6-5 and the corresponding Evidence Folders detailing the spectral interpretations of the matches found in the Appendix and Electronic Appendix respectively). This low rate of identification was mainly due to the presence of many possible structures for each matched mass, resulting in difficulty in manual spectral interpretation for structure assignment. Development of an automated spectral interpretation program in the future will likely facilitate metabolite identification using EML with two or more reactions. Nevertheless, using MyCompoundID, a total of 58 additional metabolites in urine and 48 in plasma were putatively identified, compared to 8 and 7 metabolites identified using the standard compound library, respectively. These examples illustrate that MyCompoundID can significantly increase the number of metabolites identified in biofluids.

Table 6-2 Results from database searching of urine metabolite features using reaction = 0. MS/MS spectra of unknowns was matched with the HMDB spectra

Feature ID #	Accurate $m/z$ TOF	RT (min) TOF	$m/z$ QTRAP	RT (min) QTRAP	Ion Type	Putative ID	Error (ppm)	Structure	HMDB Link
1	107.04928	17.20	107.0	17.92	$[M + H]^+$	Benzaldehyde	1.29		HMDB06115
2	246.16966	21.50	246.2	21.30	$[M + H]^+$	2-Methylbutyrylcarnitine or isomers	-1.34		HMDB00378
3	255.06550	34.30	255.1	34.79	$[M + H]^+$	Daidzein	1.24		HMDB03312
4	288.21704	34.90	288.2	35.18	$[M + H]^+$	Octanoylcarnitine or isomers	0.35		HMDB00791

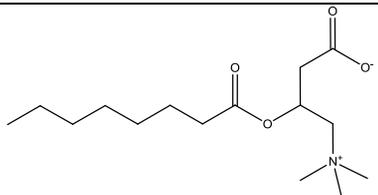
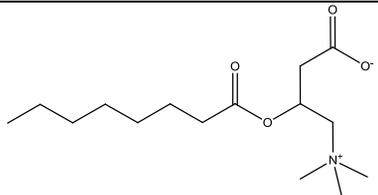
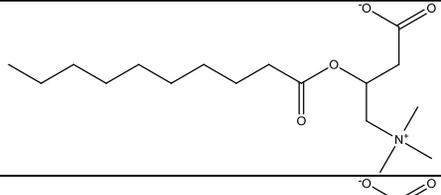
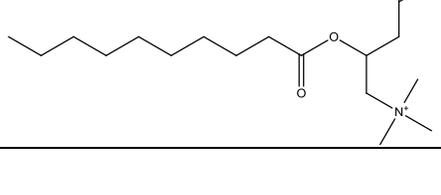
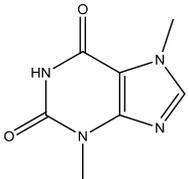
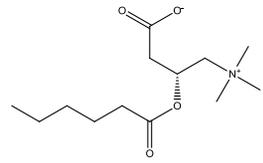
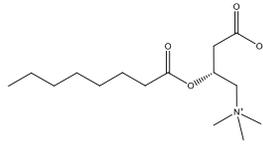
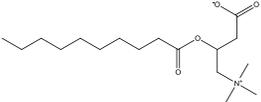
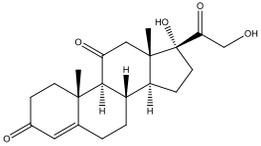
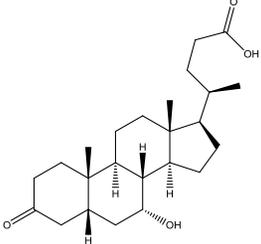
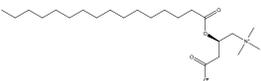
5	288.21667	35.70	288.2	35.87	$[M + H]^+$	Octanoylcarnitine or isomers	-0.92		HMDB00791
6	288.21659	36.80	288.1	37.08	$[M + H]^+$	Octanoylcarnitine or isomers	-1.19		HMDB00791
7	316.24837	40.40	316.2	40.53	$[M + H]^+$	Decanoylcarnitine or isomers	0.41		HMDB00651
8	316.24852	41.40	316.1	41.57	$[M + H]^+$	Decanoylcarnitine or isomers	0.91		HMDB00651

Table 6-3 Results from database searching of plasma metabolite features using reaction = 0. MS/MS spectra of unknowns was matched with the HMDB spectra

Feature ID #	Accurate $m/z$ TOF	RT range (min) TOF	$m/z$ QTRAP	RT (min) QTRAP	Ion type	Putative ID	error (ppm)	Structure	HMDB link
1	181.07217	3.60 - 4.10	181.0	3.40	$[M+H]^+$	Theobromine	0.91		HMDB02825
2	260.18547	28.20 - 28.80	260.2	30.80	$[M+H]^+$	Hexanoylcarnitine	-0.63		HMDB00756
3	288.21723	36.30 - 36.80	288.2	37.00	$[M+H]^+$	Octanoylcarnitine	1.03		HMDB00791

4	316.24876	40.70 - 41.20	316.2	41.50	[M+H] <sup>+</sup>	Decanoylcarnitine	1.65		HMDB00651
5	361.20086	35.10 - 35.60	361.2	35.90	[M+H] <sup>+</sup>	Cortisone	-0.24		HMDB02802
6	391.28379	66.30 - 66.80	391.3	66.30	[M+H] <sup>+</sup>	7a-Hydroxy-3-oxo-5b-cholanoic acid	-1.27		HMDB00503
7	400.34178	52.80 - 53.30	400.4	53.70	[M+H] <sup>+</sup>	Palmitoylcarnitine	-0.90		HMDB00222

Many of the compounds putatively identified using this method possess an aromatic ring. This is due to the chemistry of the hydrophilic-lipophilic balanced reversed-phase (HLB) cartridge. The HLB cartridge is suitable for the capture of a wide range of acids, bases and neutrals. The structure of the sorbent chemistry is shown in Figure 6-2. The most common retention mechanism in HLB is based on Van der Waals forces and polar interactions. Because of the benzene rings on the sorbent, aromatic compounds are attracted due to the non-covalent,  $\pi$ - $\pi$  interaction between the aromatic rings. For this reason, many of the long chain and very long chain fatty acids in plasma were not captured using this cartridge chemistry.

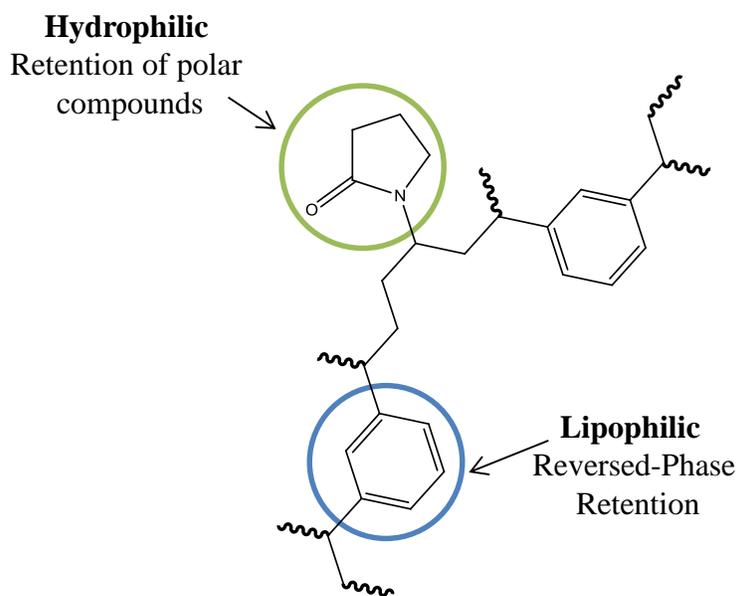


Figure 6-2 Waters Oasis<sup>®</sup> HLB chemistry

It was observed in plasma that many features less than  $m/z$  250 could not be assigned and it is possible that these compounds were not found in the HMDB. There were many metabolites whose structural isomers could not be distinguished based on the fragmentation pattern. Position of functional groups, such as hydroxyls and double bonds, could not be assigned based solely on fragmentation.

In many instances, the queries returned by the software were not consistent with the fragmentation patterns. In contrast, certain compounds were putatively assigned based on diagnostic neutral losses, even though the software returned no compounds hits, either because the compounds were not in the database or the error was higher than 5 ppm.

As can be expected, there were many more phase II metabolites in urine than in plasma and more lipid species detected in plasma. In this work, the samples were only analyzed in the positive ion mode, which limits the number of features found. Therefore very few glucuronide conjugates, fatty acids, organic acids, taurine conjugates and sulphate conjugates were detected because they have better ionization efficiency in the negative mode. LysoPC species were detected both as the protonated and sodiated ion. The CID fragmentation pattern differs significantly dependent on the ion type, so both ion type and their respective fragmentation pathways are included.

For the urine and plasma samples analyzed, MyCompoundID also allowed the identification of several interesting fragmentation patterns from metabolite peaks that are likely from exogenous metabolites. Eighteen exogenous compounds were putatively identified based on their accurate masses, diagnostic fragment ions, molecular weight distributions and literature searches.<sup>12, 13</sup> These compounds were not found in the HMDB due to their exogenous nature. A list of these compounds is presented in Supplementary Tables S6-1 and S6-2 in the Electronic Appendix. These compounds are identified as poly(ethylene glycol) or PEG derivatives which are common additives in processed food, drug formulation, toothpaste, eye drops, etc. There were 10 polyethylene glycol (PEG) and 3 monomethoxy-PEG (MMPEG) homologues found in urine and four cocodiethanolamides (CDEA) homologues found in plasma. These polymers are used extensively in pharmaceuticals (for example, slow-release drugs and beauty supplies) due their lack of toxicity, antigenicity and their rapid elimination from the body. Molecular weight distributions could be observed in the mass spectra of PEG with the difference between each homologue being 44 amu. Protonated

molecular ion clusters  $[M+H]^+$  and ammonium adducts  $[M+NH_4]^+$  were present from  $m/z$  371 to  $m/z$  740. Diagnostic fragments ions  $m/z$  89, 133 and 177 originating from the molecular ions of PEG homologues can be explained and the structures proposed.<sup>12</sup> The structures of these fragments are represented in Figure 6-3(A). For MMPEG homologues, fragments ions present in the spectra are  $m/z$  89, 103 and 147 and can be explained by the structures<sup>12</sup> in Figure 6-3(B). In the spectra of CDEA homologues, two main fragment ions are observed,  $m/z$  88 and 106 and their structures are shown in Figure 6-3(C).<sup>13</sup> The use of diagnostic ions and “fingerprint spectra” facilitated the rapid identification of homologues between the same families of exogenous compounds.

In several cases, trends in the fragmentation patterns, mostly common fragments and neutral losses were used to facilitate identification. These trends are very useful in eliminating certain compound classes generated by the accurate mass queries of the database and can lead to the identification of other compound classes. Some of these diagnostic ions were determined and the metabolites putatively identified. Others could not be identified but the diagnostic ions were still recorded, as shown in Supplementary Tables S6-3 and S6-4 in the Electronic Appendix.

There were 16 unknown metabolites in urine with fragmentation patterns similar to that of glucuronides (shown in Supplementary Table S6-5 in the Electronic Appendix); exogenous metabolites often form this type of derivatives. In total, 45 exogenous metabolites in urine and 2 in plasma were found; more metabolites found in urine than plasma is consistent with the notion that many more metabolites are excreted in urine.

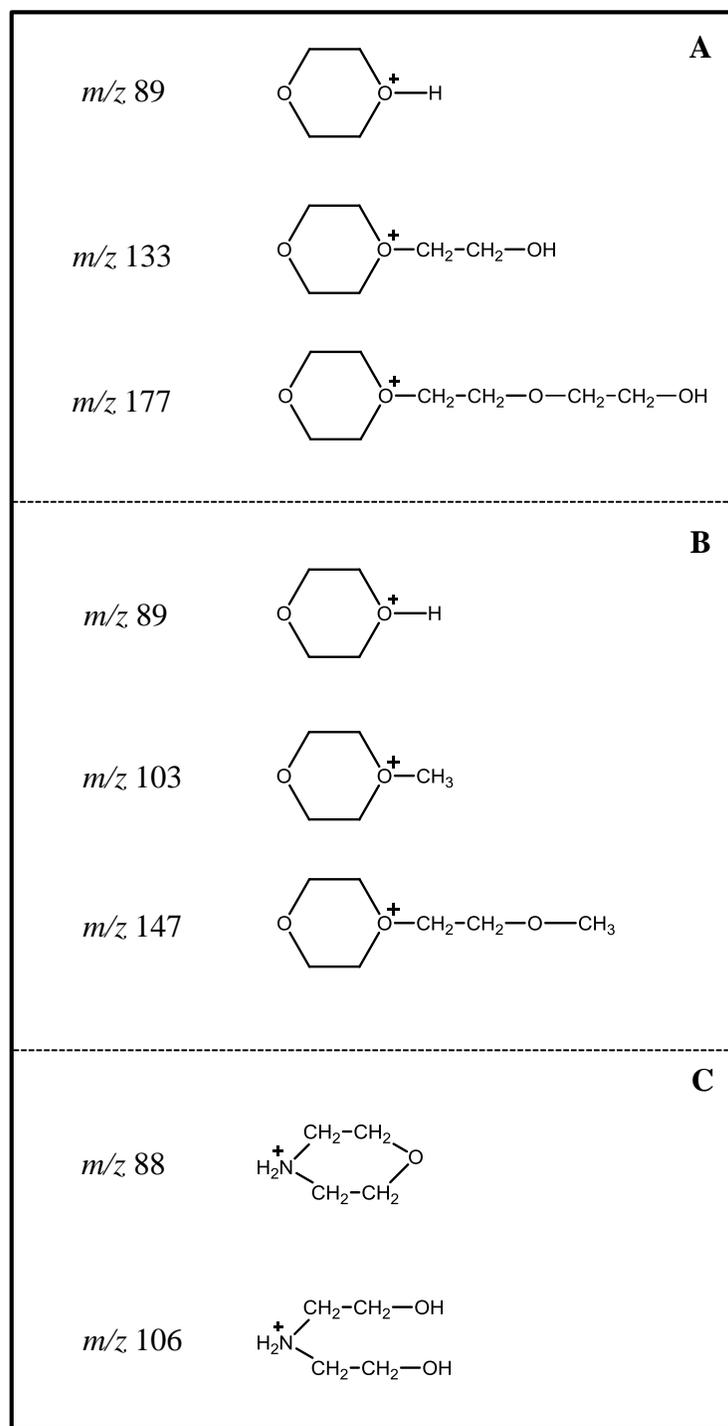


Figure 6-3 Structures of diagnostic fragment ions for (A) PEG derivatives, (B) monomethoxy-PEG derivatives and (C) cocodiethanolamides

It should be noted that MyCompoundID only allows the user to putatively identify a metabolite based on the match of accurate molecular mass and matches of fragment ions detected in MS/MS to the proposed structure. However, this exercise can narrow down the list of metabolite candidates into one or a few unique structures. If positive identification is required (e.g., a potentially useful biomarker of a disease after comparative metabolome profiling of diseased group and healthy controls), authentic standard(s) may be synthesized for comparison. Reducing the number of possible metabolite candidates by this combination of mass search and MS/MS interpretation with EML would save time and efforts, as only a few standards need to be made. In cases where the standards of putatively identified metabolites are difficult to synthesize, the use of microsome- or other cell/tissue-based biotransformation of structurally related standards<sup>14</sup> may be explored to produce the needed standards for metabolite validation.

## **6.4 Conclusions**

In summary, a publicly accessible web-based tool was developed that can facilitate the identification of unknown metabolites in metabolome profiling. Combining with LC-MS, it is shown to be useful for identifying many more metabolites in human urine and blood samples than using a standard library. This MyCompoundID tool features a dynamic compound library that can be expanded in the future by inclusion of the metabolites and their predicted metabolic products from different origins, including human, microbe, plant, food, etc. It is believed that an expanded compound library will increase the number of metabolites identifiable from human biofluids and open the possibility of using MyCompoundID for analyzing the metabolomes of other species. Future plans include adding a functionality for data sharing among the researchers who are interested in chemical identification (e.g., deposition of MS/MS spectra and their

interpretation and spectral assignment for newly identified compounds). This MyCompoundID tool features a dynamic compound library that can be expanded in the future by inclusion of the metabolites and their predicted metabolic products from different origins, including human, microbe, plant, food, etc.

## 6.5 Literature Cited

- (1) Sreekumar, A.; Poisson, L. M.; Rajendiran, T. M.; Khan, A. P.; Cao, Q.; Yu, J.; Laxman, B.; Mehra, R.; Lonigro, R. J.; Li, Y.; Nyati, M. K.; Ahsan, A.; Kalyana-Sundaram, S.; Han, B.; Cao, X.; Byun, J.; Omenn, G. S.; Ghosh, D.; Pennathur, S.; Alexander, D. C.; Berger, A.; Shuster, J. R.; Wei, J. T.; Varambally, S.; Beecher, C.; Chinnaiyan, A. M. *Nature* **2009**, *457*, 910-914.
- (2) Jenkins, H.; Hardy, N.; Beckmann, M.; Draper, J.; Smith, A. R.; Taylor, J.; Fiehn, O.; Goodacre, R.; Bino, R. J.; Hall, R.; Kopka, J.; Lane, G. A.; Lange, B. M.; Liu, J. R.; Mendes, P.; Nikolau, B. J.; Oliver, S. G.; Paton, N. W.; Rhee, S.; Roessner-Tunali, U.; Saito, K.; Smedsgaard, J.; Sumner, L. W.; Wang, T.; Walsh, S.; Wurtele, E. S.; Kell, D. B. *Nat. Biotechnol.* **2004**, *22*, 1601-1606.
- (3) Cui, Q.; Lewis, I. A.; Hegeman, A. D.; Anderson, M. E.; Li, J.; Schulte, C. F.; Westler, W. M.; Eghbalnia, H. R.; Sussman, M. R.; Markley, J. L. *Nat. Biotechnol.* **2008**, *26*, 162-164.
- (4) Han, X.; Yang, K.; Gross, R. W. *Mass Spectrom. Rev.* **2012**, *31*, 134-178.
- (5) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2010**, *5*, 1005-1018.
- (6) Zehethofer, N.; Pinto, D. M. *Anal. Chim. Acta* **2008**, *627*, 62-70.
- (7) Guo, K.; Li, L. *Anal. Chem.* **2010**, *82*, 8789-8793.
- (8) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M.-A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; MacInnis, G. D.; Weljie, A. M.; Dowlatabadi, R.;

- Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. *Nucleic Acids Res.* **2007**, *35*, D521-D526.
- (9) Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; Yamanishi, Y. *Nucleic Acids Res.* **2008**, *36*, D480-D484.
- (10) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. *Ther. Drug Monit.* **2005**, *27*, 747-751.
- (11) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T. *J. Mass Spectrom.* **2010**, *45*, 703-714.
- (12) Seraglia, R.; Traldi, P.; Mendichi, R.; Sartore, L.; Schiavon, O.; Veronese, F. M. *Anal. Chim. Acta* **1992**, *262*, 277-283.
- (13) Ferrer, I.; Furlong, E. T.; Thurman, E. M. *Prepr. Ext. Abstr. ACS Natl. Meet., Am. Chem. Soc., Div. Environ. Chem.* **2002**, *42*, 352-358.
- (14) Clements, M.; Li, L. *Anal. Chim. Acta* **2012**, *685*, 36-44.

## Chapter 7. Conclusions and Future Work

### 7.1 Conclusions

LC-MS-based metabolomics has gained growing popularity as the platform of choice for metabolomics studies. This is due to its potential for high throughput, soft ionization techniques and coverage of metabolites detected. This platform aims at identification and quantification of metabolites in both targeted and non-targeted studies. An ideal LC-MS platform should offer confident identification of metabolites through high quality MS and MS/MS spectra and accurate quantification of the endogenous compounds.

In Chapter 1, key terms are explained and instrumentation, methodologies, chemical derivatization, qualitative and quantitative methods are discussed as related to LC-MS based metabolomics. As these qualitative and quantitative methods are applied to the study of acylglycine biomarkers, a brief introduction to the metabolism and role of these compounds are included. Challenges relating to both identification and quantification of metabolite biomarkers are also briefly examined. Addressing several of these challenges will be focus of the projects in this thesis research work.

In Chapter 2, a targeted LC-MS/MS approach is developed and applied to the analysis of acylglycines in human urine. This work is significant because currently more than 70 inborn errors of metabolisms are known but only about 10-15 acylglycines are used in the diagnosis of a very small fraction of these disorders. Because acylglycines are small polar molecules, and they are present in lower concentrations in urine than organic acids, their detection is often a challenge. The goal of this work is to improve the detection and identification of novel acylglycines detected in urine. Based on the fragmentation of standards, sensitive MRM methods are developed. Using LC retention time, known fragmentation rules, fragmentation patterns of the standards, and literature

searches, identifications are made on the putatively characterized compound class. In an article published in the *Journal of American Society for Mass Spectrometry* in 2010 and adapted for this chapter, forty seven novel acylglycines were putatively identified and reported for the first time. Current methods can only detect 10-15 acylglycines.<sup>1-3</sup> With the discovery of novel acylglycines, a better understanding of the disease biology can be determined and potentially more disorders can be more confidently diagnosed and treated.

In Chapters 3 and 4, a novel quantitative surrogate matrix approach is applied to quantify acylglycines in urine and plasma respectively. The acylglycines are chemically derivatized using a labeling strategy developed in our laboratory. The labeling reagent, <sup>12</sup>C/<sup>13</sup>C- p-dimethylaminophenacyl (DmPA) bromide, is a derivative of phenacyl bromide which is commonly used to label carboxylic acids and fatty acids. Derivatization with the label serves to enhance the ESI response through increased surface activity (addition of the phenyl group) and increased chargeability (addition of the dimethylamine group). Propionylglycine, which was undetected in urine in the previous project (Chapter 2), is enhanced when labeled and its signal is magnified up to 3 orders of magnitude. Other acylglycine signals are enhanced in urine by 1-3 orders of magnitude. In plasma, the acylglycines concentrations are much lower than in urine and are undetected unless they are labeled. The surrogate matrix approach is applied, where the underivatized biofluid is used as the surrogate matrix, while the derivatized biofluid is the authentic matrix. The concept used is one of matrix-matching using the underivatized matrix to prepare calibration standards in the absence of a “true” blank. Slopes in water, surrogate and authentic matrix are compared and it is found that water is not a good surrogate matrix due to differences in matrix effects. Although several LC-MS/MS methods have been reported for sensitive detection of acylglycines, these methods provide a simple and straightforward sample preparation and validation procedures, which offers better LLOQ values.<sup>4, 5</sup> This approach is an attempt to address the challenge of accurate quantification of endogenous biomarkers.

In Chapter 5, a fast UPLC-MS/MS method is developed for the analysis of homocysteine, methylmalonic acid and succinic acid. For high throughput analysis, the separation of the analytes is performed in about 36 seconds with a resolution of 1.0 or higher. To the best of my knowledge, this is the first reported polarity switching method that can perform this separation in less than 1 minute; current methods are about 3 minutes long.<sup>6, 7</sup> The method is applied successfully in determining the role of ABC transporters in vitamin B<sub>12</sub> metabolism. *C. Elegans* is chosen as a model organism and the concentration of methylmalonic acid is determined in mutant worms as compared to wild-type worms. Two new ABC transporters were reported for the first time as participating in cobalamin metabolism and potentially new cobalamin deficiencies could be identified. Currently only the role of MRP-1 transporter protein has been linked to Vitamin B<sub>12</sub> metabolism.<sup>8</sup> This work is done in collaboration with Dr. Roy Gravel's laboratory at the University of Calgary. Previously, methylmalonic acid concentration was determined in his laboratory using extraction, chemical derivatization and GC-MS. The procedure was laborious and time consuming as hundreds of samples are generated. The fast LC method developed addresses the issue of high throughput and reports a lower limit of detection and quantification.

In Chapter 6, a novel web-based metabolite identification tool called "My Compound ID" (MCI) is described that allows searching and interpreting mass spectrometry data against a newly constructed metabolome library. This program uses the database HMDB entries and 76 common biotransformations reported in literature to create an evidence-based metabolite library. By generating *in silico* biotransformations of the 8021 entries found in the HMDB, the library can be expanded to include the 8,021 known human endogenous metabolites and their predicted metabolic products (375,809 compounds from one metabolic reaction and 10,583,901 from two reactions). Putative identifications are made *de novo* for compounds in urine and plasma obtained from the same individual. Queries returned from searches of accurate masses are eliminated based on the common neutral losses obtained from fragmentation spectra. LC retention time,

fragmentation rules, diagnostic fragment ions and literature searches are also used to assist in the structural elucidation of the unknown compounds and aid in putative identification. This project seeks to address the challenge of limited coverage in database searches, by increasing metabolic coverage using known biotransformations.

The focus of this research was to develop novel approaches to overcome several of the limitations encountered in identifying and quantifying metabolites. Several tools have been developed that facilitate the identification of putative compound classes, in the absence of standards. The quantification approach has the potential to overcome problems associated with matrix effects, namely ion suppression, reversed-phase separation of polar metabolites and bioanalytical validation using endogenous compounds.

## **7.2 Future Work**

### **7.2.1 Biomarker Discovery**

Simply put, biomarkers are molecules, whose concentrations change with disease. Biomarker discovery is therefore a set of experiments used to find biomarkers. The biomarker discovery process consists of the following stages: choosing the correct metabolomics strategy, collecting the right clinical samples, optimizing metabolite separations, performing the best mass spectrometry using the most appropriate instrumentation and accurately analyzing the data.

In Chapter 2, sixty three acylglycines were detected, with over forty putatively identified in the urine of healthy individuals. In Chapters 3 and 4, a small fraction of these acylglycines were quantified in urine and plasma. In this thesis, the strategy, separations and mass spectrometry stages of the biomarker process have already been optimized. The next stage would be to obtain valid clinical samples to determine if any of the putatively identified acylglycines undergo changes in their profile or concentrations and are viable biomarker

candidates. Valid clinical samples could include individuals with known and unknown inborn errors of metabolism. These biomarker candidates then go through a verification process to make sure that they change with disease states and a validation process to determine whether they are useful in a clinical setting. In order to make a positive and confirmed identification of the biomarkers, a combination of NMR, accurate mass determination, isotope deconvolution and fragmentation information need to be acquired. Standards are then purchased or synthesized and analyzed to confirm the identity of the acylglycine biomarkers. The discovery of any novel biomarkers can be used in further understanding and improving the diagnosis of novel disorders.

### **7.2.2 Dried Urine Spots**

Dried blood spots (DBS) have been used in newborn screening programs and are mainly used in the screening of compounds like organic acids and acylcarnitines. Because of the low concentration of acylglycines in blood, DBS are seldom used in the analysis of acylglycines. Collecting urine from newborns can be quite problematic, however, and an alternative collection method could be performed by blotting filter paper on a used pamper or diaper of the infected newborn and dried before extraction. In other words, dried urine spots (DUS) can be collected in this manner. This would simplify sample collection and address many sample storage issues. The acylglycines can then be extracted from the DUS, derivatized and then subjected to LC-MS/MS analysis.

In a preliminary study, a urine sample was obtained from a healthy individual and analyzed. Two sets of prepared urine were used: unspiked urine and urine spiked with 50  $\mu\text{M}$  mixture of acylglycine standards. Sample preparation was identical for both sets of urine samples. The filter paper used in this study was Whatman™ 903 Protein Saver Cards. The filter paper was saturated with 100  $\mu\text{L}$  urine and allowed to dry overnight in a dessicator. For elution of the

metabolites from the paper, 3mm holes were punched and placed in a vial containing 300  $\mu\text{L}$  methanol and extracted by sonication for 30 minutes. The methanol extract was then dried and derivatized according to the procedure recorded in Chapter 3. For comparison, an equivalent volume of urine was extracted by liquid-liquid extraction using 300  $\mu\text{L}$  ethylacetate. The extract was dried and derivatized using an identical method as the DUS. The samples were then analyzed by LC-MS/MS. The experimental workflow is illustrated in Figure 7-1. Comparisons of the extraction techniques used for both sets of urine samples are presented in Figure 7-2. Additional experiments are needed to optimize the extraction procedures and determine recovery and stability of the acylglycines in the filter paper. If the dried urine filter paper is a feasible and reliable method, then screening could be more practical and the method could be adapted to diagnosing IEMs using clinical samples. Samples could be sent from remote places even distant countries to be analyzed.

### **7.2.3 Microwave-assisted Extraction and Derivatization**

Microwave-assisted extraction (MAE) has received increasing attention recently as a potential alternative to solid-liquid extraction techniques. Microwave irradiation produces efficient and rapid internal heating by direct coupling of microwave energy with the molecules. MAE allows for simple, rapid and low solvent-consuming extraction methods and is particularly useful in extraction of small molecules collected on filters and absorbents.<sup>9</sup>

The derivatization step can become a tedious and time-consuming process in any workflow. It has been shown that microwave irradiation could remarkably accelerate the derivatization of carbonyls<sup>10</sup> and carboxylic acids,<sup>11</sup> amphetamines and diuretics<sup>12</sup> and has been successfully applied to many other compounds.<sup>13</sup> Using sealed vessel technology, the temperature of the reaction mixture can be raised far above the boiling point of the solutions compared to conventional methods. As a result, reaction times can be reduced from several hours to several

minutes. Coupling extraction and derivatization using microwave-assisted techniques could be an alternative route to analyzing acylglycines in dried urine spots. This one-step strategy would create a simple and rapid method for acylglycine analysis.

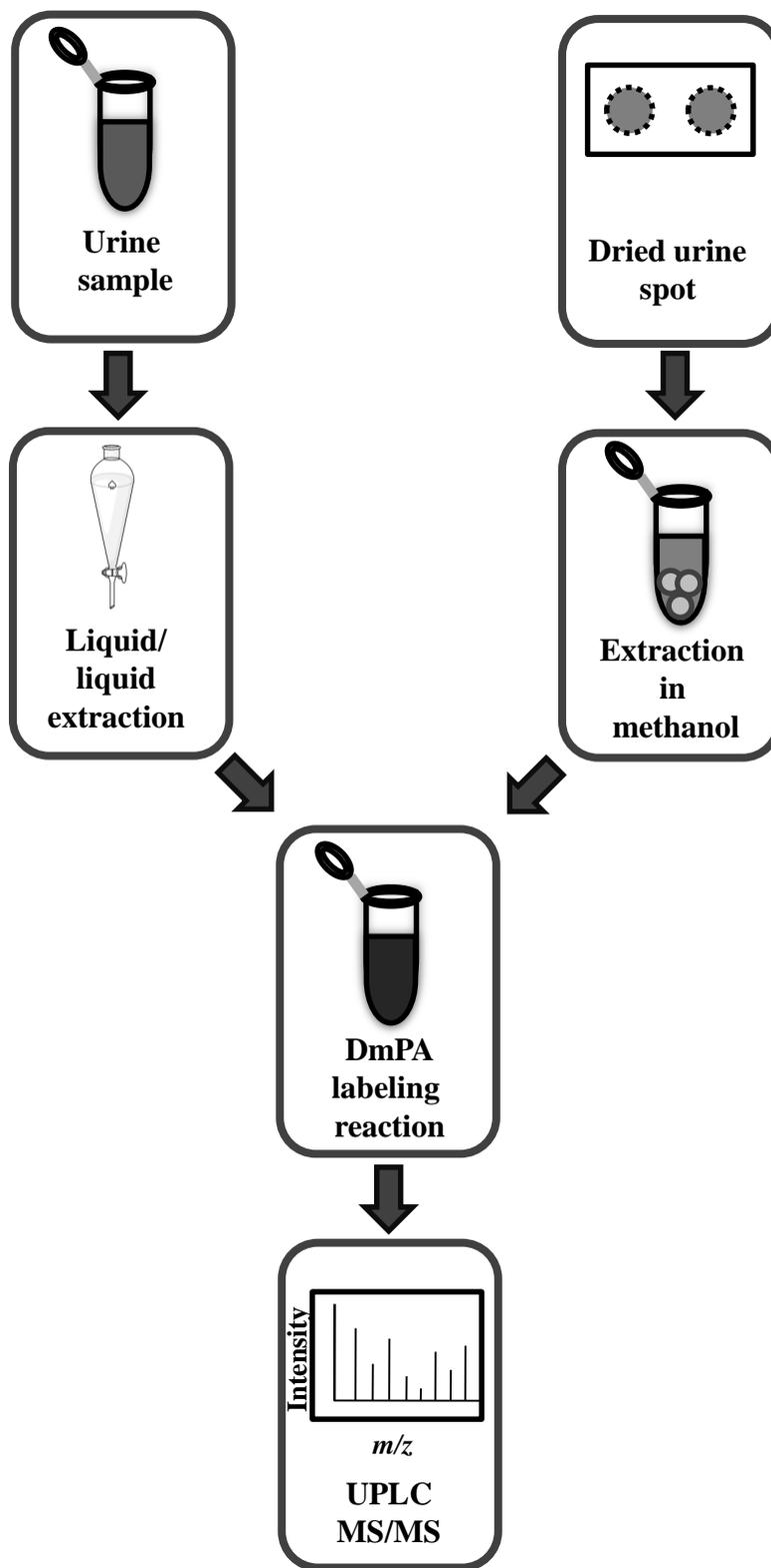


Figure 7-1 Experimental workflow of urine analysis

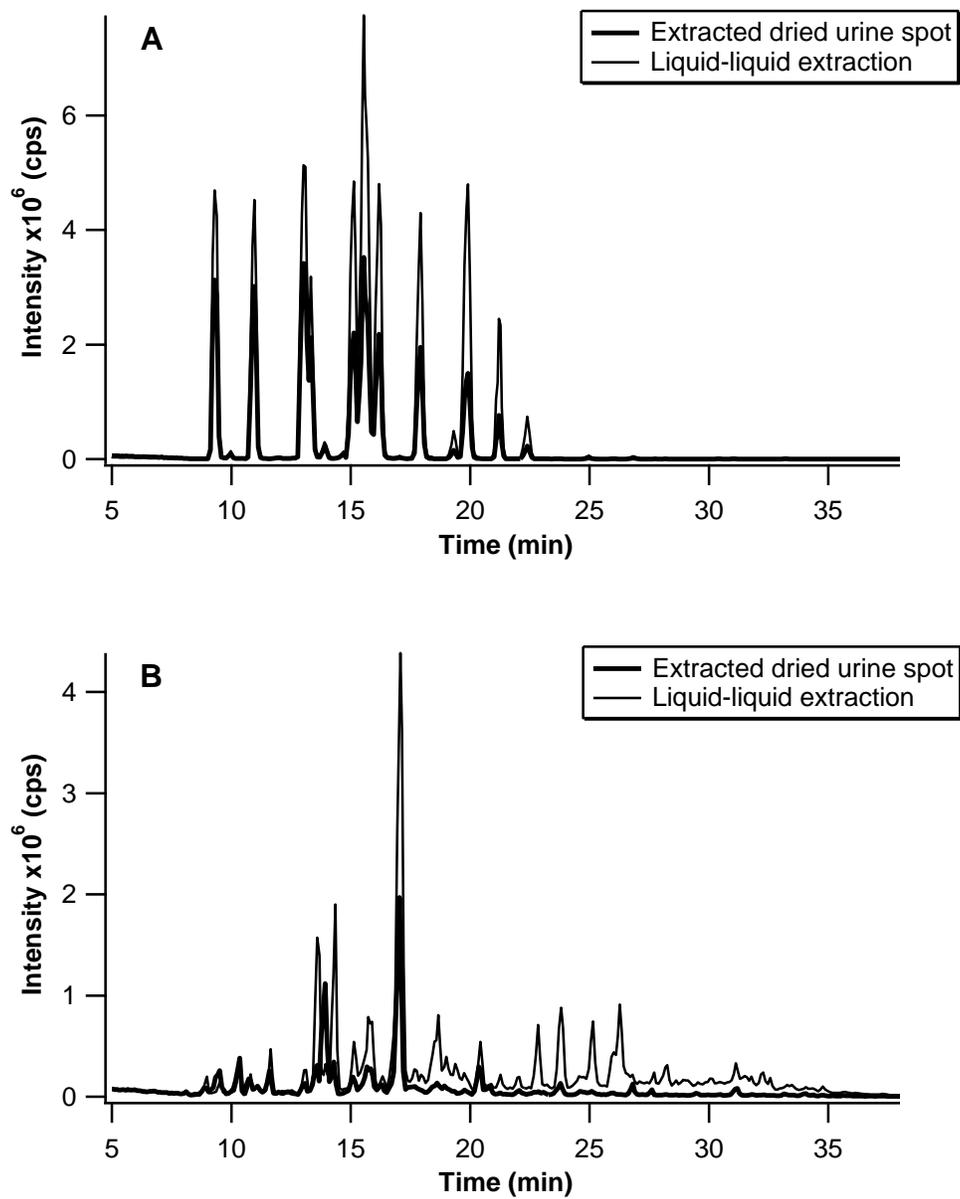


Figure 7-2 Comparison of the liquid-liquid extraction and dried urine spot extraction from (A) Spiked urine sample, final concentration  $50 \mu\text{M}$  and (B) Unspiked urine sample.

### 7.3 Literature Cited

- (1) Bonafe, L.; Troxler, H.; Kuster, T.; Heizmann, C. W.; Chamoles, N. A.; Burlina, A. B.; Blau, N. *Mol. Genet. Metab.* **2000**, *69*, 302-311.
- (2) Kimura, M.; Yamaguchi, S. *J. Chromatogr. B.* **1999**, *731*, 105-110.
- (3) Costa, C. G.; Guerand, W. S.; Struys, E. A.; Holwerda, U.; ten Brink, H. J.; de Almeida, I. T.; Duran, M.; Jakobs, C. *J. Pharm. Biomed. Anal.* **2000**, *21*, 1215-1224.
- (4) Ombrone, D.; Salvatore, F.; Ruoppolo, M. *Anal. Biochem.* **2011**, *417*, 122-128.
- (5) Fong, B. M. W.; Tam, S.; Leung, K. S. Y. *Talanta* **2011**, *88*, 193-200.
- (6) Hempen, C.; Wanschers, H.; Sluijs, V. G. *Anal. Bioanal. Chem.* **2008**, *391*, 263-270.
- (7) Pedersen, T. L.; Keyes, W. R.; Shahab-Ferdows, S.; Allen, L. H.; Newman, J. W. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2011**, *879*, 1502-1506.
- (8) Beedholm-Ebsen, R.; van, d. W. K.; Hardlei, T.; Nexoe, E.; Borst, P.; Moestrup, S. K. *Blood* **2010**, *115*, 1632-1639.
- (9) Wathey, B.; Tierney, J.; Lidstrom, P.; Westman, J. *Drug Discovery Today* **2002**, *7*, 373-380.
- (10) Strassnig, S.; Wenzl, T.; Lankmayr, E. P. *J. Chromatogr. A.* **2000**, *891*, 267-273.
- (11) Ye, Q.; Zheng, D.; Zhu, F. *Anal. Methods* **2010**, *2*, 354-358.
- (12) Amendola, L.; Colamonici, C.; Mazzarino, M.; Botrè, F. *Anal. Chim. Acta* **2003**, *475*, 125-136.
- (13) Damm, M.; Rechberger, G.; Kollroser, M.; Kappe, C. O. *J. Chromatogr. A.* **2009**, *1216*, 5875-5881.

## Appendix

### Chapter 2

The following figures are examples of structural elucidations which led to putative identification of acylglycines. A full list of the spectral library of acylglycines is presented in the Acylglycines MS/MS library in the Electronic Appendix. Please contact Professor Liang Li for a copy of the Electronic Appendix.

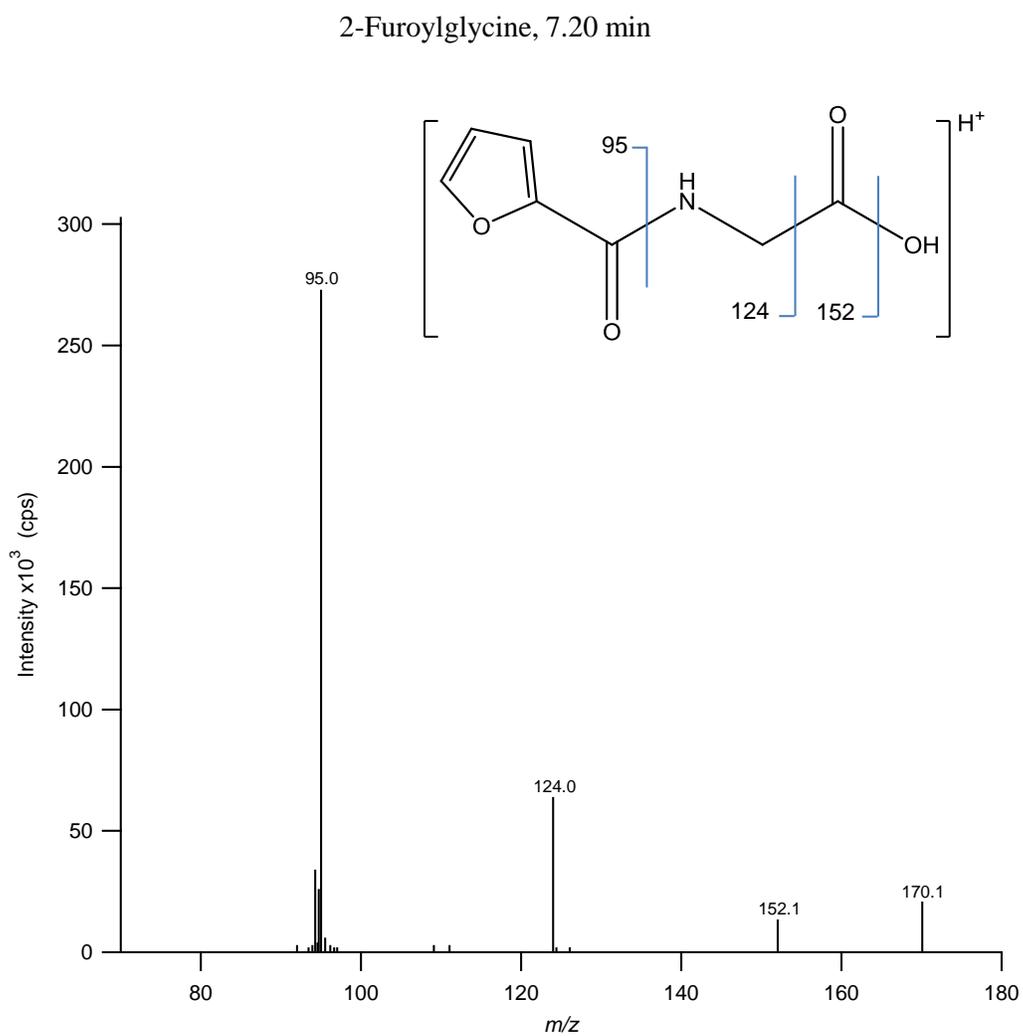


Figure A2-1 Structural elucidation of 2-furoylglycine.

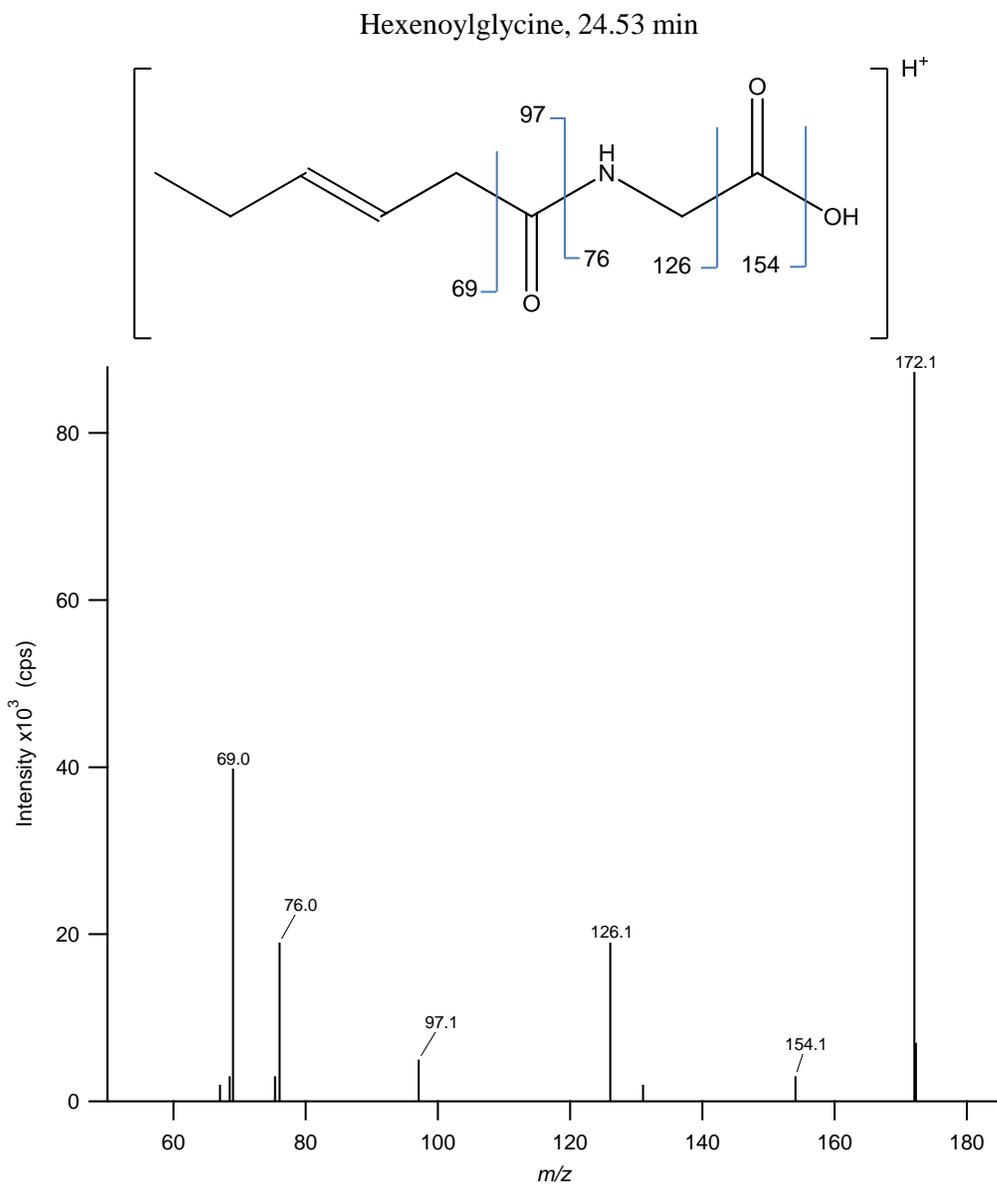


Figure A2-2 Structural elucidation of hexenoylglycine

Cis-3,4-Methylene-heptanoylglycine, 42.46 min

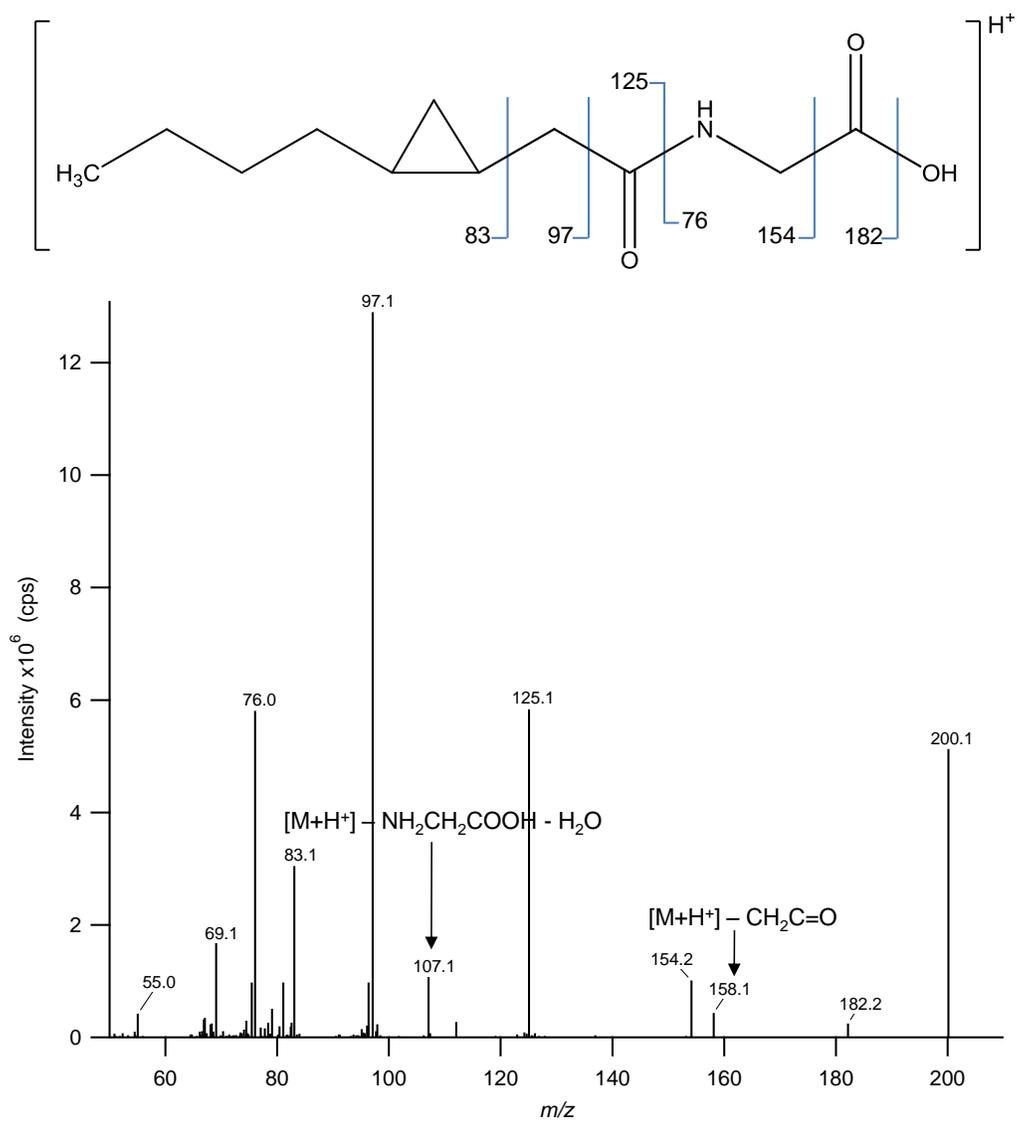


Figure A2-3 Structural elucidation of cis-3,4-methylene-heptanoyl

Hydroxyphenylpropionylglycine, 21.37 min

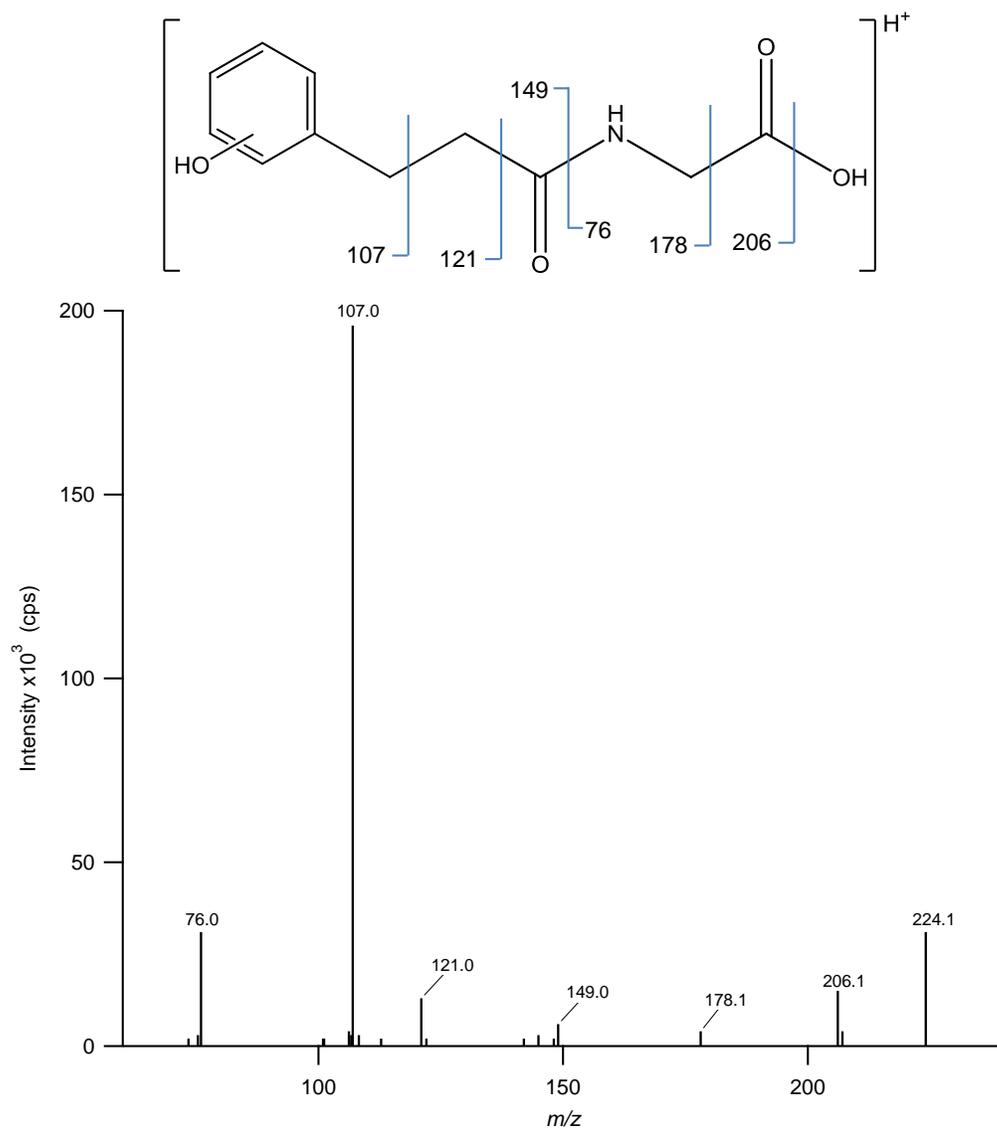


Figure A2-4 Structural elucidation of hydroxyphenylpropionylglycine

### Chapter 3

The following figures are examples of calibration curves. A full list of the calibration curves is presented in Figure S3-1 in the Electronic Appendix.

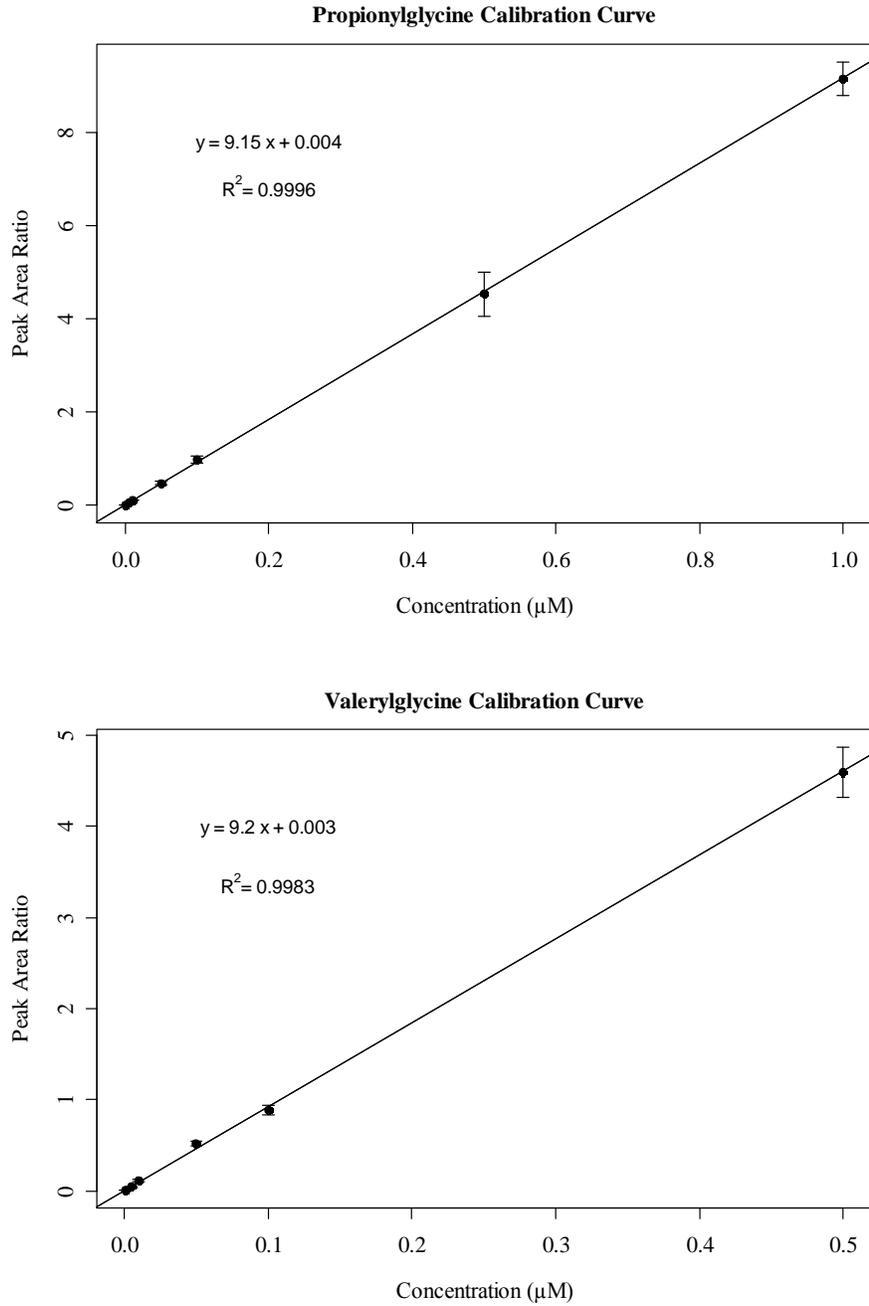


Figure A3-1 Acylglycines calibration curves in surrogate urine

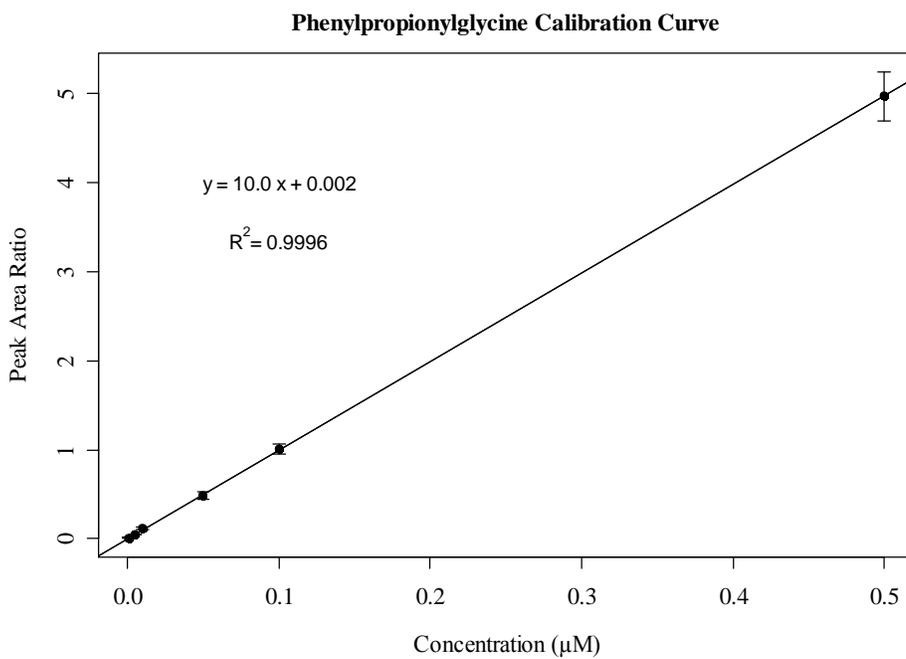
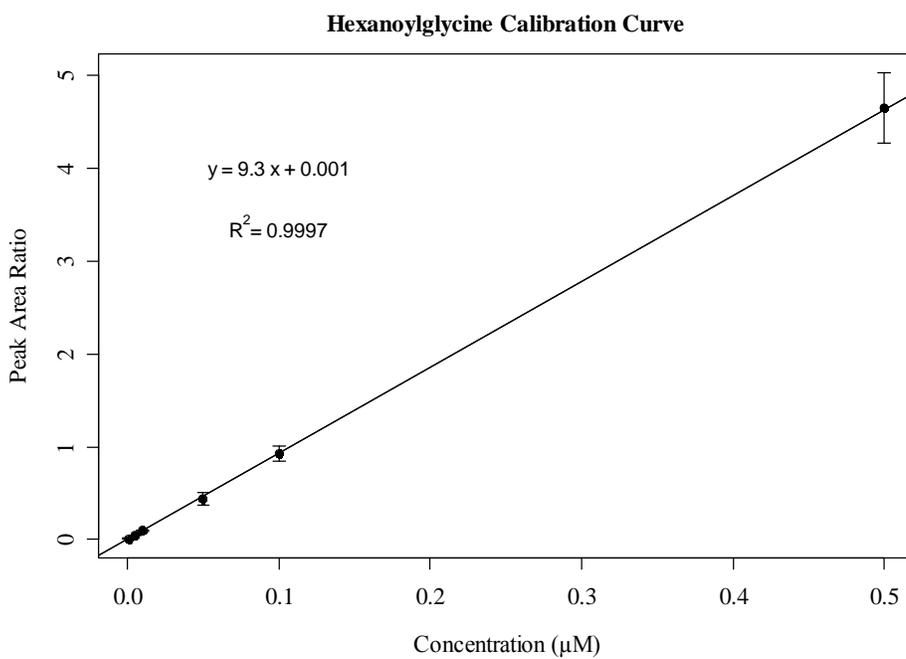


Figure A3-1 Acylglycines calibration curves in surrogate urine

Table A3-1 Surrogate Urine Calibration Curves Statistical Summary

Acylglycines	Slope	Error in slope	Intercept	Error in Intercept	Std. Error	F Value	p value	df
AG	8.6	0.1	0.004	0.003	0.03	5330	2e-7	4
PG	9.15	0.08	0.004	0.003	0.03	12374	1e-9	5
IBG	10.4	0.1	0.000	0.002	0.03	8921	8e-8	4
BG	10.0	0.1	0.004	0.004	0.04	4721	3e-7	4
HPAG	8.7	0.2	0.004	0.004	0.05	2678	8e-7	4
FG	8.90	0.07	0.001	0.002	0.02	15099	3e-8	4
TG	8.6	0.2	0.005	0.004	0.04	2802	7e-7	4
2MBG	9.2	0.1	0.002	0.003	0.03	4872	2e-7	4
3MCG	9.4	0.2	-0.000	0.004	0.04	3121	6e-7	4
IVG	10.1	0.1	0.000	0.003	0.03	6666	1e-7	4
VG	9.2	0.2	0.003	0.005	0.05	2296	1e-6	4
HG	9.25	0.08	0.000	0.001	0.02	13780	3e-8	4
PAG	9.3	0.1	0.004	0.002	0.03	8342	9e-8	4
PPG	10.0	0.1	0.002	0.002	0.03	9486	7e-8	4
GG	8.8	0.1	-0.000	0.002	0.03	8491	8e-8	4
HpG	9.2	0.2	0.002	0.003	0.04	3053	6e-7	4
OG	9.5	0.1	0.004	0.002	0.03	6454	1e-7	4
SG	8.3	0.2	0.003	0.003	0.04	3013	7e-7	4

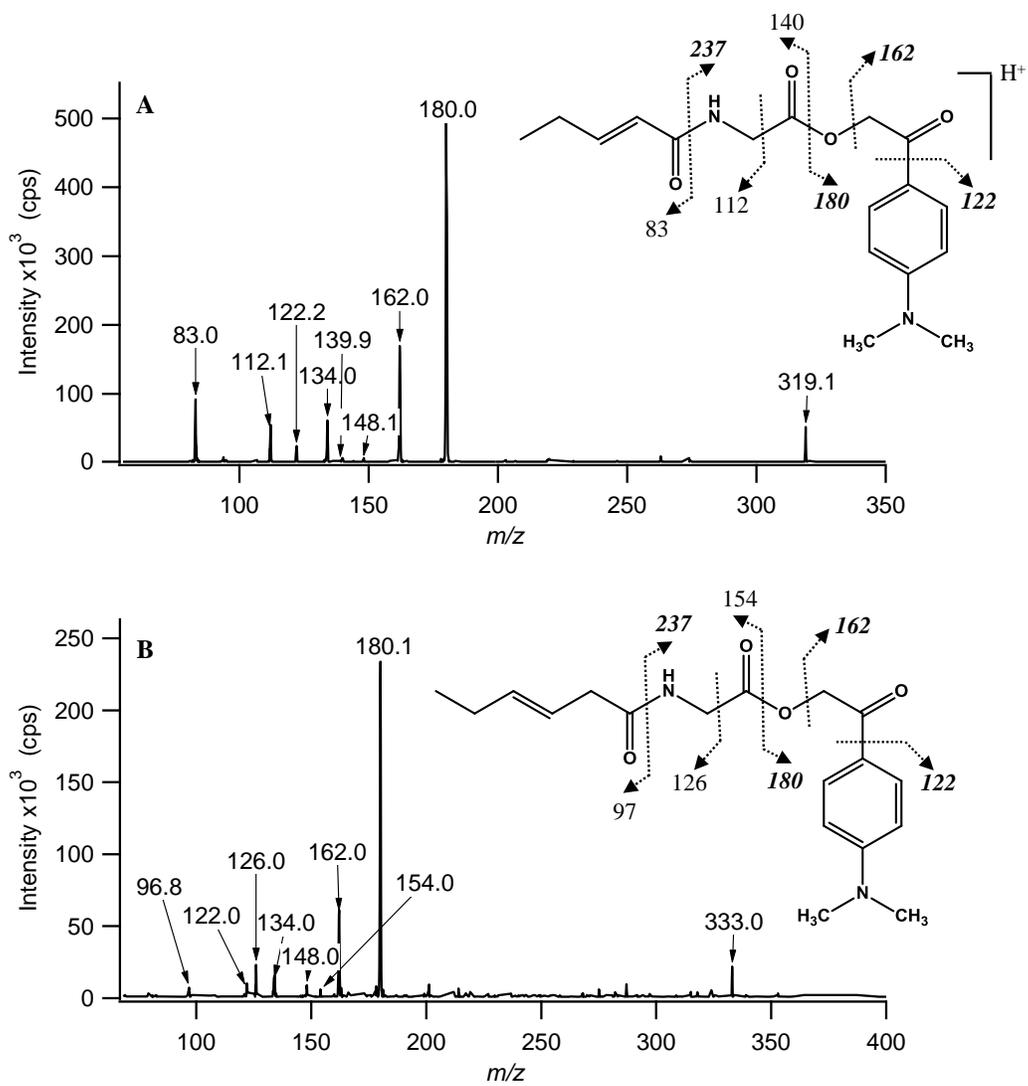


Figure A3-2 Fragmentation and MS/MS spectra of (A) C<sub>5</sub>:1-G; (B) C<sub>6</sub>:1-G

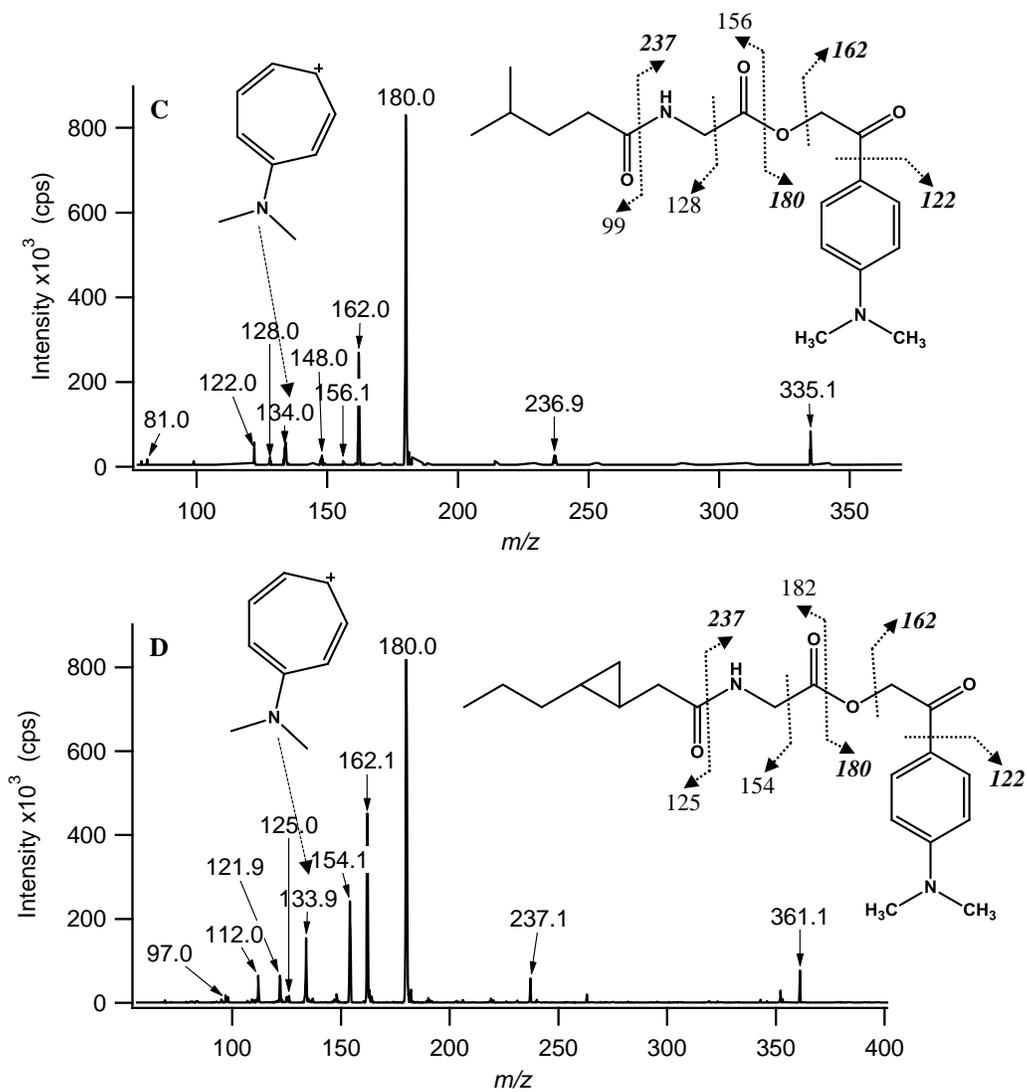


Figure A3-3 Fragmentation and MS/MS spectra of (C)  $C_6$ -G; (D)  $C_8$ :1-G

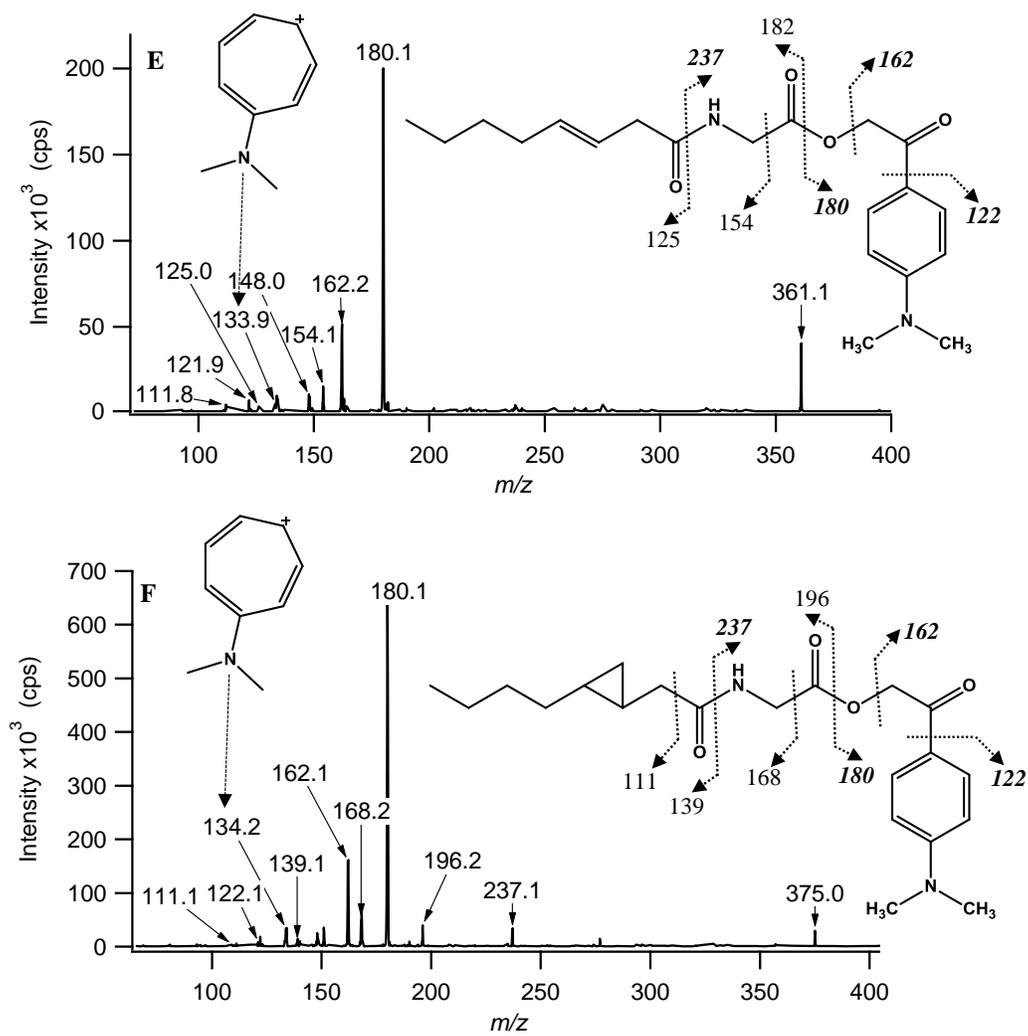


Figure A3-4 Fragmentation and MS/MS spectra of (E) C<sub>8</sub>:1-G; (F) C<sub>9</sub>:1-G

## Chapter 4

The following figures are examples of calibration curves. A full list of the calibration curves is presented in Figure S4-1 in the Electronic Appendix.

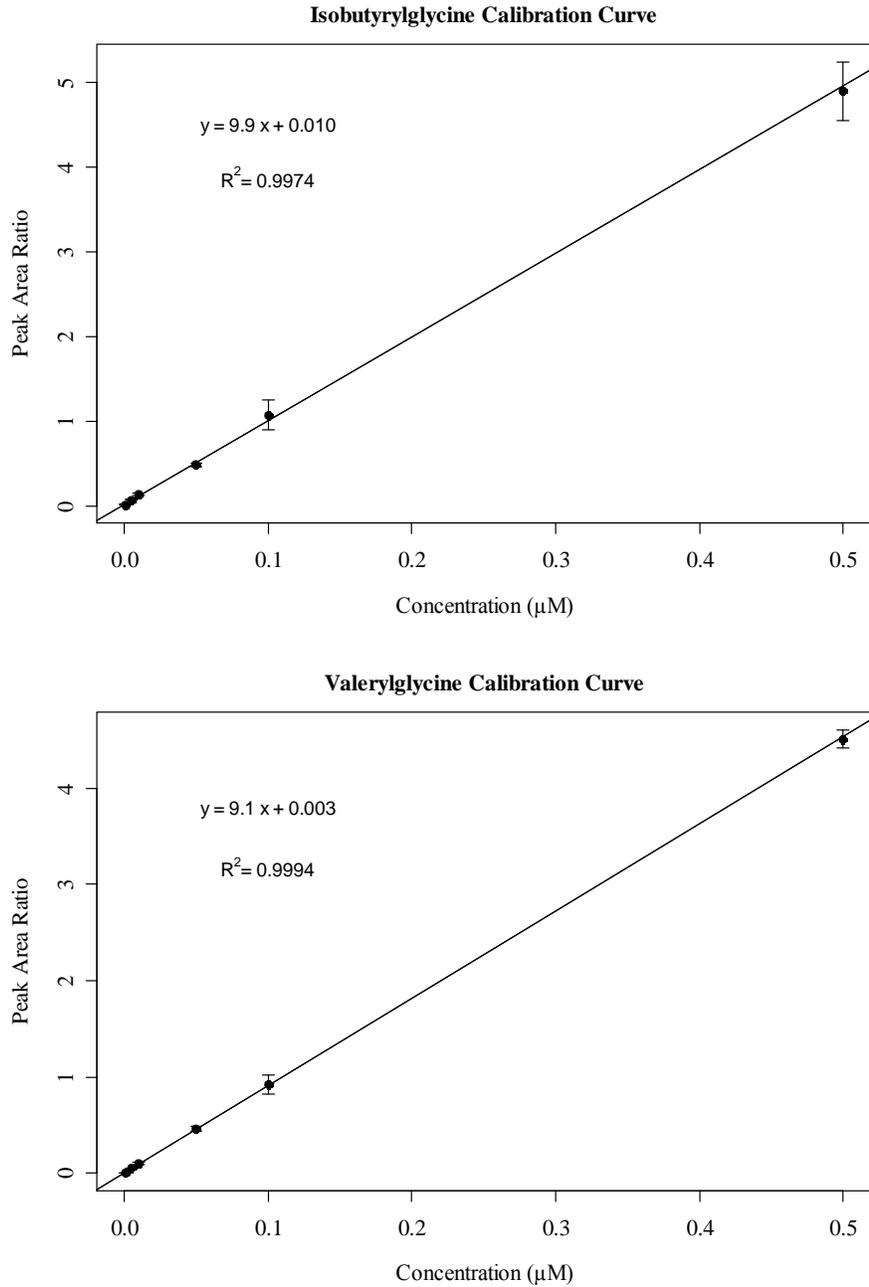


Figure A4-1 Acylglycines Calibration Curves in Surrogate Plasma

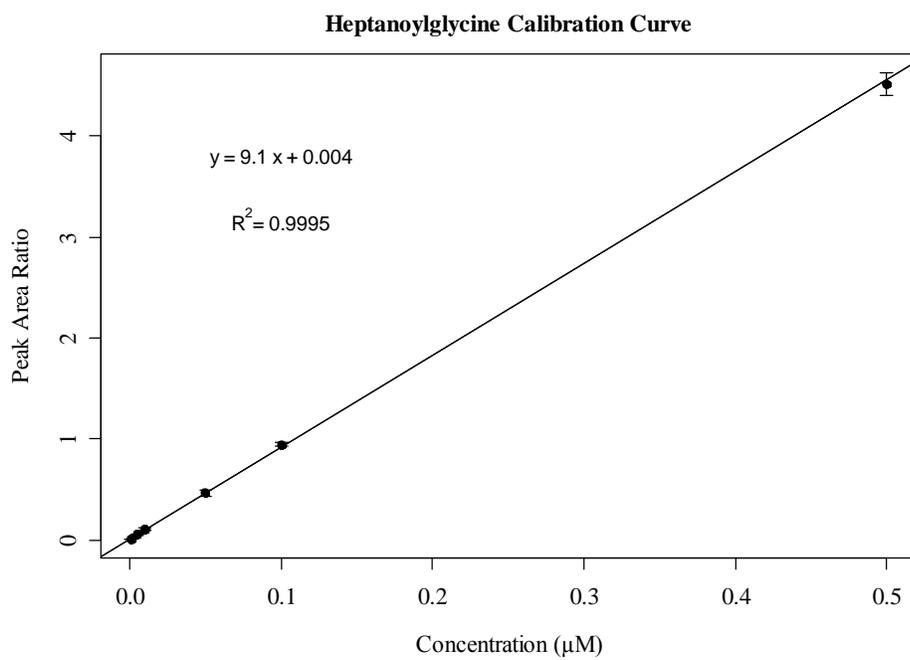
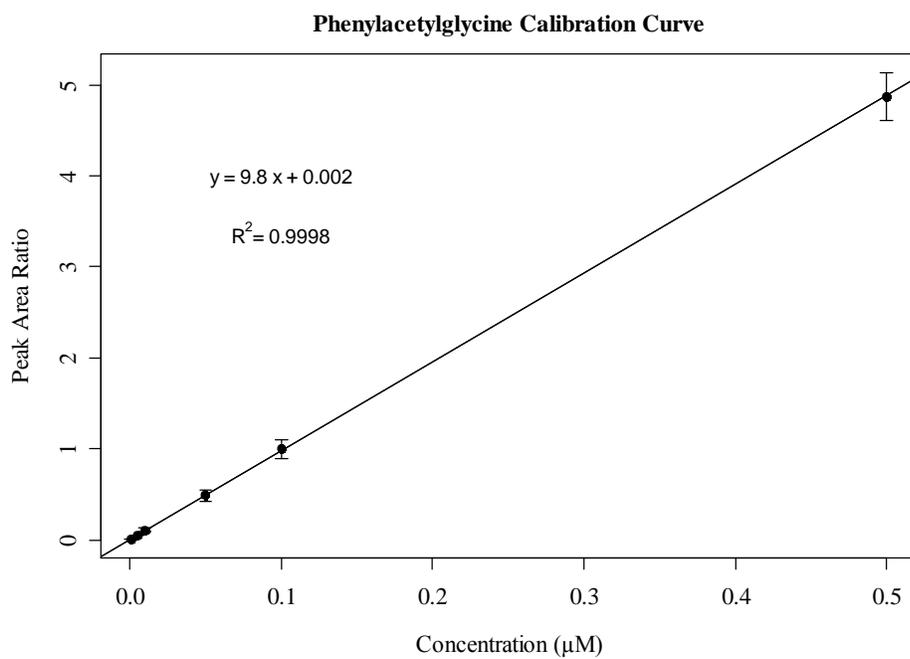


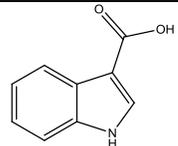
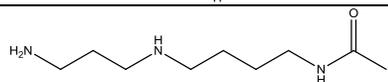
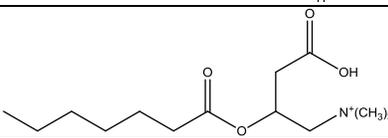
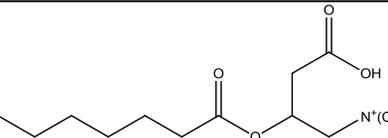
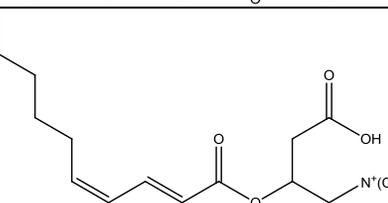
Figure A4-1 Acylglycines Calibration Curves in Surrogate Plasma

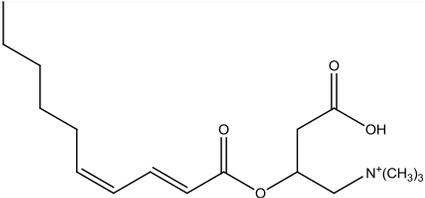
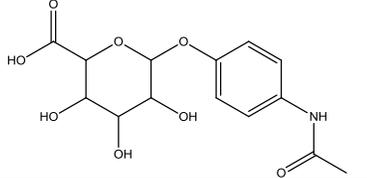
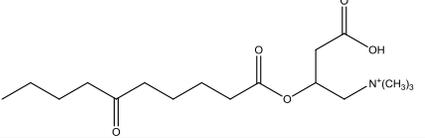
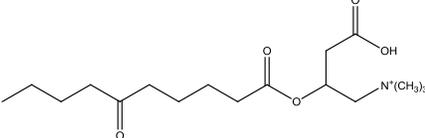
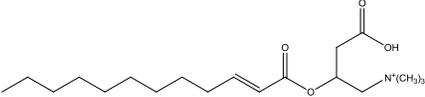
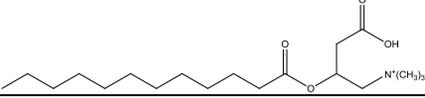
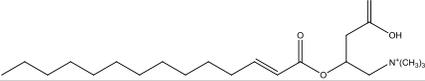
Table A4-1 Surrogate Plasma Calibration Curves Statistical Summary

Acylglycines	Slope	Error in slope	Intercept	Error in Intercept	Std. Error	F Value	p value	df
AG	9.0	0.1	0.009	0.006	0.06	4582	1e-8	5
PG	8.2	0.1	0.007	0.004	0.05	5929	7e-9	5
IBG	9.9	0.3	0.010	0.007	0.06	1506	3e-6	4
BG	10.2	0.1	0.005	0.003	0.03	6971	1e-7	4
HPAG	8.8	0.2	0.005	0.004	0.05	2437	1e-6	4
FG	9.5	0.2	0.004	0.004	0.04	3269	6e-7	4
TG	9.1	0.2	0.005	0.004	0.05	2760	7e-7	4
2MBG	8.8	0.1	0.005	0.002	0.03	7402	1e-7	4
3MCG	9.5	0.2	0.006	0.004	0.04	3639	5e-7	4
IVG	8.9	0.1	0.005	0.003	0.04	4587	3e-7	4
VG	9.1	0.1	0.003	0.002	0.03	6916	1e-7	4
HG	9.28	0.04	0.002	0.001	0.01	49545	2e-9	4
PAG	9.77	0.07	0.002	0.001	0.02	18134	8e-8	4
PPG	9.0	0.1	0.003	0.002	0.03	7061	1e-7	4
GG	8.71	0.09	0.002	0.002	0.02	9799	6e-8	4
HpG	9.1	0.1	0.004	0.002	0.03	7855	1e-7	4
OG	8.6	0.1	0.004	0.002	0.03	5392	2e-7	4
SG	10.4	0.2	0.002	0.004	0.05	3040	6e-7	4

## Chapter 6

Table A6-1 Results from database searching of urine metabolite features using reaction = 0. No MS/MS spectra for these compounds were found in the HMDB database. MS/MS spectral interpretation was done to putatively identify structure.

Feature ID #	Accurate <i>m/z</i> (TOF)	RT (min) TOF	<i>m/z</i> QTRAP	RT (min) QTRAP	Ion Type	Putative ID	Error (ppm)	Structure	HMDB Link
1	162.05493	26.10 - 26.60	162.1	26.10	[M + H] <sup>+</sup>	Indole-3-carboxylic acid or isomers	-0.15		HMDB03320
2	188.17491	1.50 - 2.00	188.2	1.22	[M + H] <sup>+</sup>	N8-Acetylspermidine or isomers	-4.45		HMDB02189
3	274.20019	31.90 - 32.30	274.2	32.68	[M + H] <sup>+</sup>	Heptanoylcarnitine	-4.02		HMDB13238
4	274.20090	32.60 - 33.10	274.1	33.29	[M + H] <sup>+</sup>	Heptanoylcarnitine	-1.43		HMDB13238
5	312.2169	36.50 - 37.30	312.1	37.07	[M + H] <sup>+</sup>	2-trans,4-cis-Decadienoylcarnitine	-0.02		HMDB13325

6	312.21748	38.10 - 38.60	312.2	38.57	[M + H] <sup>+</sup>	2-trans,4-cis-Decadienoylcarnitine	1.75		HMDB13325
7	328.10302	26.30 - 26.90	328.1	26.49	[M + H] <sup>+</sup>	Acetaminophen glucuronide	0.99		HMDB10316
8	330.22735	33.10 - 33.60	330.1	33.79	[M + H] <sup>+</sup>	6-Keto-decanoylcarnitine	-0.47		HMDB13202
9	330.22633	34.60 - 35.10	330.2	35.16	[M + H] <sup>+</sup>	6-Keto-decanoylcarnitine	-3.54		HMDB13202
10	342.26378	43.00 - 43.50	342.2	43.46	[M + H] <sup>+</sup>	trans-2-Dodecenoylcarnitine	-0.30		HMDB13326
11	344.27888	45.20 - 45.70	344.3	45.63	[M + H] <sup>+</sup>	Dodecanoylcarnitine	-1.94		HMDB02250
12	370.29446	46.80 - 47.50	370.3	47.45	[M + H] <sup>+</sup>	Trans-2-Tetradecenoylcarnitine or isomer	-1.98		HMDB13329

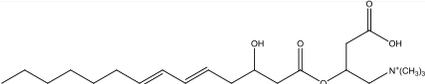
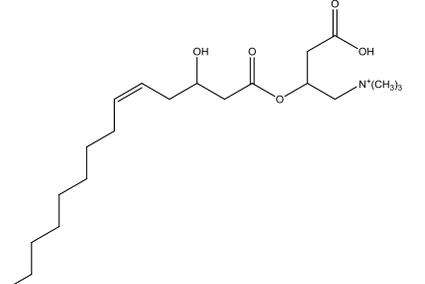
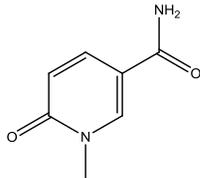
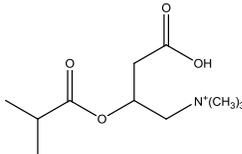
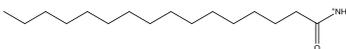
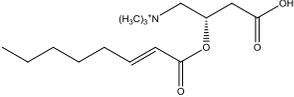
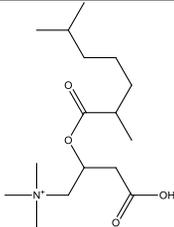
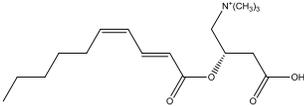
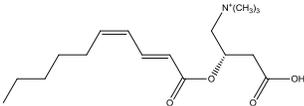
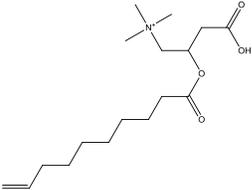
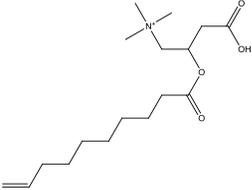
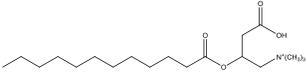
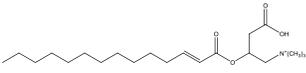
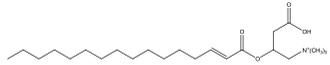
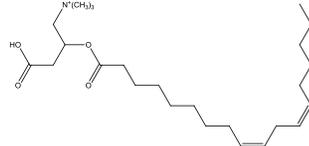
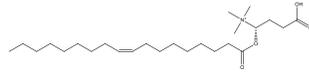
13	384.27459	41.90 - 42.40	384.3	42.33	[M + H] <sup>+</sup>	3-Hydroxy-5, 8- tetradecadiencarnitine	0.36		HMDB13332
14	386.28919	43.60 - 44.20	386.3	44.10	[M + H] <sup>+</sup>	3-Hydroxy-cis-5- tetradecenylcarnitine	-2.37		HMDB13330

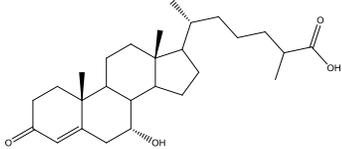
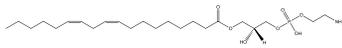
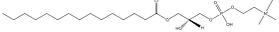
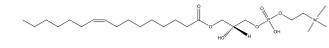
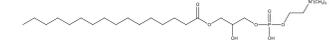
Table A6-2 Results from database searching of plasma metabolite features using reaction = 0. No MS/MS spectra for these compounds were found in the HMDB database. MS/MS spectral interpretation was done to putatively identify the structure.

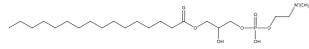
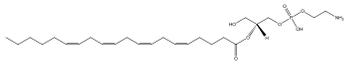
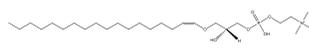
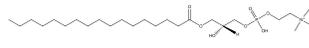
Feature ID #	Accurate <i>m/z</i> TOF	RT range (min) TOF	<i>m/z</i> QTRAP	RT (min) QTRAP	Ion type	Putative ID	error (ppm)	Structure	HMDB link
1	153.06586	2.30 - 2.70	153.1	2.10	[M +H] <sup>+</sup>	N1-Methyl-2-pyridone-5-carboxamide or N1-Methyl-4-pyridone-3-carboxamide	0.01		HMDB04193
2	232.15440	3.60 - 4.20	232.2	4.40	[M +H] <sup>+</sup>	Isobutyryl or butyrylcarnitine	0.29		HMDB00736
3	256.26299	58.90 - 59.50	256.2	59.60	[M +H] <sup>+</sup>	Palmitic amide	-1.96		HMDB12273

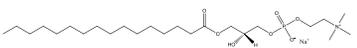
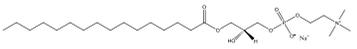
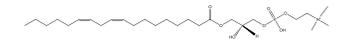
4	286.20142	33.20 - 33.70	286.2	34.30	[M +H] <sup>+</sup>	2-octenoylcarnitine	0.49		HMDB13324
5	302.23220	37.20 - 37.60	302.2	37.90	[M +H] <sup>+</sup>	2,6-dimethylheptanoylcarnitine Nonanoylcarnitine	1.27		HMDB13288
6	312.21678	37.00 - 37.40	312.2	37.80	[M +H] <sup>+</sup>	2-trans,4-cis- Decadienoylcarnitine or isomer	-0.49		HMDB13325
7	312.21655	37.60 - 38.00	312.2	38.30	[M +H] <sup>+</sup>	2-trans,4-cis- Decadienoylcarnitine or isomer	-1.24		HMDB13325

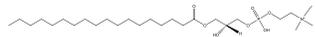
8	314.23284	38.90 - 39.40	314.2	39.80	[M +H] <sup>+</sup>	9-Decenoylcarnitine or isomer	0.80		HMDB13205
9	314.23219	39.80 - 40.20	314.2	40.50	[M +H] <sup>+</sup>	9-Decenoylcarnitine or isomer	0.80		HMDB13205
10	344.27926	44.70 - 45.20	344.3	45.70	[M +H] <sup>+</sup>	Dodecanoylcarnitine	-0.80		HMDB02250
11	370.29507	46.40 - 46.90	370.3	47.50	[M +H] <sup>+</sup>	trans-2- Tetradecenoylcarnitine or cis-5- Tetradecenoylcarnitine	-0.32		HMDB13329

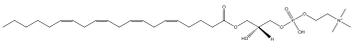
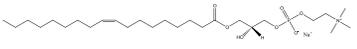
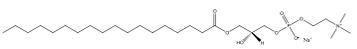
12	398.32598	49.70 - 50.40	398.3	50.80	[M +H] <sup>+</sup>	trans-Hexadec-2-enoyl carnitine	-1.28		HMDB06317
13	424.34187	50.90 - 51.40	424.3	51.90	[M +H] <sup>+</sup>	Linoleylcarnitine or isomers	-0.62		HMDB06461
14	426.35725	53.60 - 54.10	426.4	54.50	[M +H] <sup>+</sup>	Oleoylcarnitine or isomers	-1.25		HMDB05065

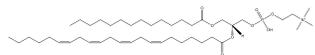
15	431.31472	49.00 - 49.50	431.3	50.10	[M +H] <sup>+</sup>	7 alpha-Hydroxy-3-oxo-4-cholestenoate	-2.02		HMDB12458
16	478.29269	48.50 - 49.00	478.3	49.50	[M +H] <sup>+</sup>	LysoPE(0:0/18:2(9Z,12Z)) or LysoPE(18:2(9Z,12Z)/0:0)	-0.26		HMDB11477
17	482.32383	48.20 - 48.70	482.3	49.30	[M +H] <sup>+</sup>	LysoPC(15:0)	-0.59		HMDB10381
18	494.32440	47.30 - 47.90	494.3	48.30	[M +H] <sup>+</sup>	LysoPC(16:1(9Z))	0.57		HMDB10383
19	496.34060	49.50 - 49.80	496.3	50.60	[M +H] <sup>+</sup>	LysoPC(16:0) isomer (branched)	1.67		HMDB10382

20	496.33767	50.30 - 50.90	496.4	51.40	[M +H] <sup>+</sup>	LysoPC(16:0)	-4.23		HMDB10382
21	502.29305	48.60 - 49.10	502.4	49.00	[M +H] <sup>+</sup>	LysoPE(0:0/20:4(5Z,8Z,11Z,14Z)), LysoPE(0:0/20:4(8Z,11Z,14Z,17Z)), LysoPE(20:4(5Z,8Z,11Z,14Z)/0:0) or LysoPE(20:4(8Z,11Z,14Z,17Z)/0:0)	0.47		HMDB11487
22	508.37562	52.20 - 52.70	508.4	53.30	[M +H] <sup>+</sup>	LysoPC(P-18:0)	-1.05		HMDB13122
23	510.35492	52.60 - 53.10	510.4	53.70	[M +H] <sup>+</sup>	LysoPC(17:0)	-0.98		HMDB12108

24	518.32156	49.50 - 49.80	518.3	50.60	$[M + Na]^+$	LysoPC(16:0) isomer	-0.32		HMDB10382
25	518.32248	50.30 - 50.90	518.3	51.40	$[M + Na]^+$	LysoPC(16:0)	1.55		HMDB10382
26	520.33798	48.60 - 49.20	520.3	49.00	$[M + H]^+$	LysoPC(18:2(9Z,12Z))	-3.45		HMDB10386

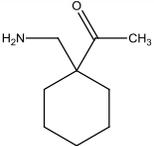
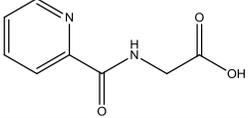
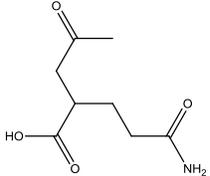
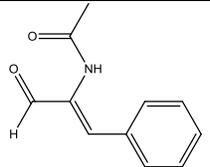
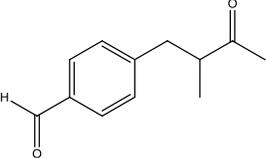
27	522.35552	51.30 - 51.90	522.4	52.40	[M +H] <sup>+</sup>	LysoPC(18:1(9Z)) or LysoPC(18:1(11Z))	0.20		HMDB02815
28	524.37206	55.00 - 55.60	524.4	55.20	[M +H] <sup>+</sup>	LysoPC(18:0) or LysoPC(0:0/18:0)	1.90		HMDB10384
29	542.32191	48.60 - 49.20	542.6	49.00	[M +Na] <sup>+</sup>	LysoPC(18:2(9Z,12Z))	0.37		HMDB10386

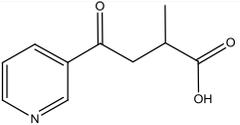
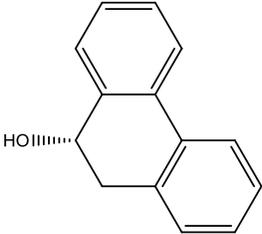
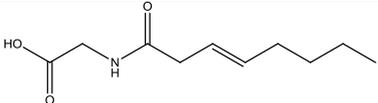
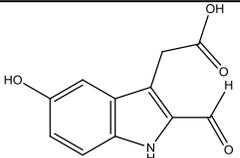
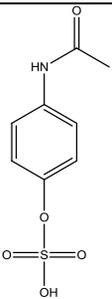
30	544.34029	48.70 - 49.20	544.6	49.74	[M +H] <sup>+</sup>	LysoPC(20:4(5Z,8Z,11Z,14Z)) LysoPC(20:4(8Z,11Z,14Z,17Z))	0.95		HMDB10395
31	544.33744	51.30 - 51.90	544.6	52.40	[M +Na] <sup>+</sup>	LysoPC(18:1(9Z)) or LysoPC(18:1(11Z))	0.14		HMDB02815
32	546.35509	50.20 - 50.70	546.7	51.20	[M +H] <sup>+</sup>	LysoPC(20:3(5Z,8Z,11Z)) or LysoPC(20:3(8Z,11Z,14Z))	-0.60		HMDB10393
33	546.35288	55.00 - 55.60	546.8	56.10	[M +Na] <sup>+</sup>	LysoPC(18:0) or LysoPC(0:0/18:0)	-0.26		HMDB10384

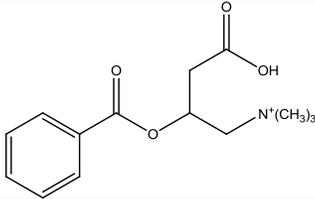
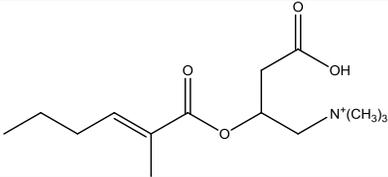
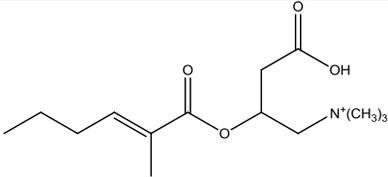
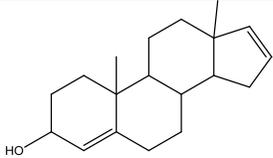
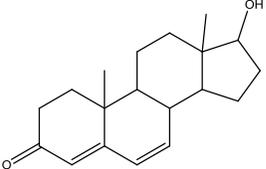
34	782.56595	63.30 - 64.00	782.7	63.30	[M +H] <sup>+</sup>	-4.45	<p>Any of</p> <p>PC(14:0/22:4(7Z,10Z,13Z,16Z)),</p> <p>PC(16:0/20:4(5Z,8Z,11Z,14Z)),</p> <p>PC(16:0/20:4(8Z,11Z,14Z,17Z)),</p> <p>PC(16:1(9Z)/20:3(5Z,8Z,11Z)),</p> <p>PC(16:1(9Z)/20:3(8Z,11Z,14Z)),</p> <p>PC(18:0/18:4(6Z,9Z,12Z,15Z)),</p> <p>PC(18:1(11Z)/18:3(6Z,9Z,12Z)),</p> <p>PC(18:1(11Z)/18:3(9Z,12Z,15Z)),</p> <p>PC(18:1(9Z)/18:3(6Z,9Z,12Z)),</p> <p>PC(18:1(9Z)/18:3(9Z,12Z,15Z)),</p> <p>PC(18:2(9Z,12Z)/18:2(9Z,12Z)),</p> <p>PC(18:3(6Z,9Z,12Z)/18:1(11Z)),</p> <p>PC(18:3(6Z,9Z,12Z)/18:1(9Z)),</p> <p>PC(18:3(9Z,12Z,15Z)/18:1(11Z)),</p> <p>PC(18:3(9Z,12Z,15Z)/18:1(9Z)),</p> <p>PC(18:4(6Z,9Z,12Z,15Z)/18:0),</p> <p>PC(20:3(5Z,8Z,11Z)/16</p>		HMDB07889
----	-----------	------------------	-------	-------	---------------------	-------	--	---	-----------

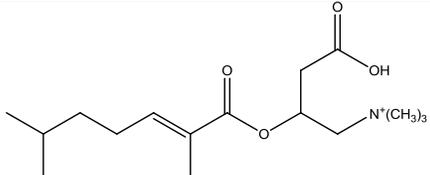
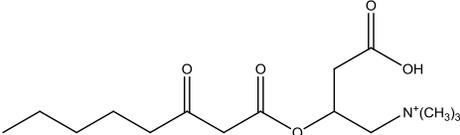
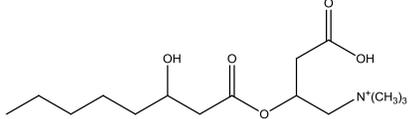
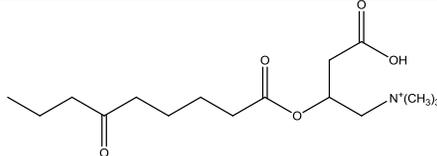
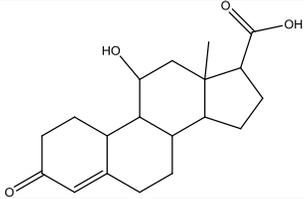
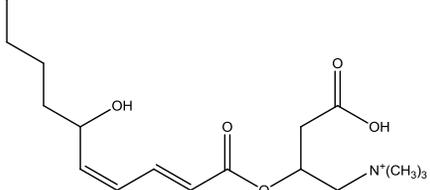
:1(9Z)),  
PC(20:3(8Z,11Z,14Z)/1  
6:1(9Z)),  
PC(20:4(5Z,8Z,11Z,14  
Z)/16:0),  
PC(20:4(8Z,11Z,14Z,1  
7Z)/16:0),  
PC(22:4(7Z,10Z,13Z,1  
6Z)/14:0)

Table A6-3 Results from database searching of urine metabolite features using reaction = 1. MS/MS spectral interpretation was done to putatively identify structure.

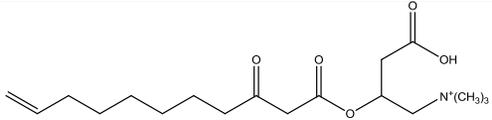
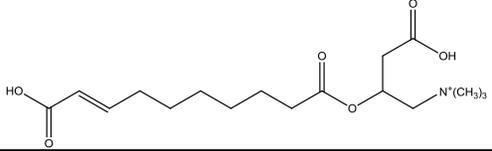
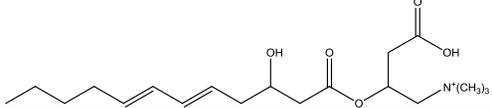
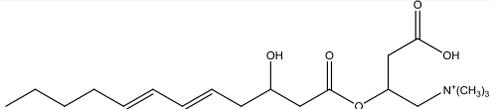
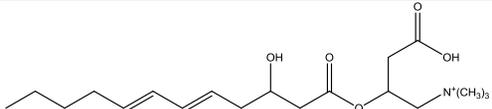
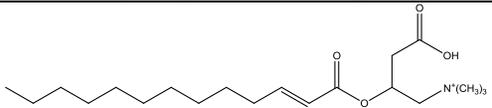
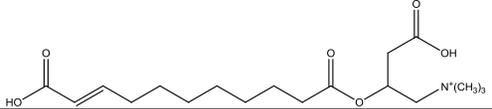
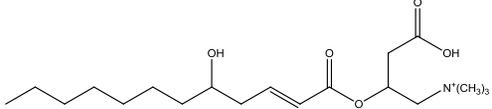
Feature ID #	Accurate $m/z$ TOF	RT (min) TOF	$m/z$ QTRAP	RT (min) QTRAP	Ion Type	Putative ID	Error (ppm)	Structure
1	156.13807	35.80 - 36.30	155.9	36.37	$[M + H]^+$	Gabapentin - O	-1.4	
2	181.06001	5.80 - 6.30	181.2	6.35	$[M + H]^+$	Picolinic acid + C <sub>2</sub> H <sub>3</sub> NO (glycine) or isomer	-4.2	
3	188.12758	10.70 - 11.20	188.2	11.00	$[M + H]^+$	N-Acetyl-leucine + CH <sub>2</sub>	-2.9	
4	190.08581	16.70 - 17.90	190.3	17.45	$[M + H]^+$	N-Acetyl-L-phenylalanine - H <sub>2</sub> O	-2.38	
5	191.10706	30.80 - 31.30	191.1	31.16	$[M + H]^+$	Cuminaldehyde + C <sub>2</sub> H <sub>2</sub> O	2.12	

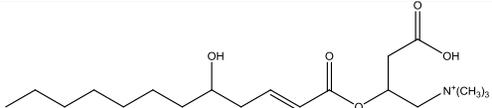
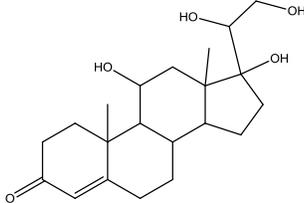
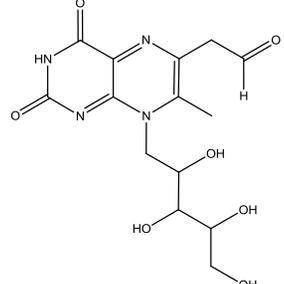
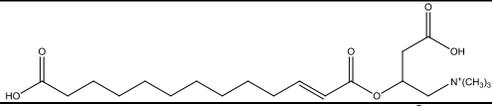
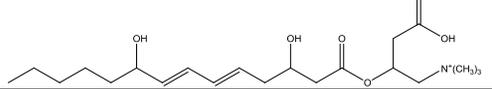
6	194.08110	22.10 - 23.00	194.2	22.49	[M + H] <sup>+</sup>	3-Succinoylpyridine + CH <sub>2</sub>	-0.36	
7	197.09607	47.00 - 47.60	197.2	47.74	[M + H] <sup>+</sup>	(+)-(1R,2R)-1,2-Diphenylethane-1,2-diol - H <sub>2</sub> O	-0.09	
8	200.12830	33.60 - 34.10	200.2	34.21	[M + H] <sup>+</sup>	Capryloylglycine + H <sub>2</sub>	0.88	
9	220.06023	26.30 - 26.80	220.2	26.36	[M + H] <sup>+</sup>	5-Hydroxyindoleacetic acid + CO	-0.94	
10	232.02726	6.40 - 7.00	232.1	7.29	[M + H] <sup>+</sup>	Acetaminophen + SO <sub>3</sub>	-0.72	

11	266.13904	25.80 - 26.40	266.1	26.08	[M + H] <sup>+</sup>	Benzoic acid + C <sub>7</sub> H <sub>13</sub> NO <sub>2</sub> (carnitine)	1.32	
12	272.18492	29.80 - 30.30	272.2	30.58	[M + H] <sup>+</sup>	Tiglylcarnitine + C <sub>2</sub> H <sub>4</sub> or isomer	-2.59	
13	272.18453	30.90 - 31.40	272.1	31.61	[M + H] <sup>+</sup>	Tiglylcarnitine + C <sub>2</sub> H <sub>4</sub> or isomer	-4.05	
14	273.22063	41.90 - 42.40	273.2	42.38	[M + H] <sup>+</sup>	Androstenol - H <sub>2</sub> or isomer	-2.47	
15	287.19942	36.20 - 36.80	287.2	36.82	[M + H] <sup>+</sup>	Testosterone - H <sub>2</sub> or isomer	-2.02	

16	300.21738	36.70 - 37.20	300.1	37.21	[M + H] <sup>+</sup>	2,6 dimethylheptanoyl carnitine - H <sub>2</sub> or isomer	1.50	
17	302.19765	27.80 - 28.40	302.2	28.50	[M + H] <sup>+</sup>	2-Octenoylcarnitine + O or isomers	4.84	
18	304.21057	30.90 - 31.30	304.1	31.51	[M + H] <sup>+</sup>	3-Hydroxyoctanoic acid + C <sub>7</sub> H <sub>13</sub> NO <sub>2</sub> (carnitine) or isomer	-4.21	
19	316.21114	30.30 - 30.80	316.1	31.01	[M + H] <sup>+</sup>	6-Keto-decanoylcarnitine - CH <sub>2</sub> or isomer	-2.23	
20	319.19112	41.70 - 42.20	319.2	42.19	[M + H] <sup>+</sup>	11beta-hydroxyprogesterone - CH <sub>2</sub> or isomer	2.31	
21	328.21054	30.00 - 30.50	328.1	30.26	[M + H] <sup>+</sup>	2-trans,4-cis-Decadienoylcarnitine + O or isomer	-3.99	

22	328.21181	32.60 - 33.10	328.2	33.23	$[M + H]^+$	2-trans,4-cis-Decadienylcarnitine + O or isomer	-0.12	
23	328.24671	39.70 - 40.30	328.2	40.29	$[M + H]^+$	4,8 dimethylnonanoyl carnitine -H <sub>2</sub> or isomer	-4.70	
24	328.24889	40.90 - 41.30	328.2	41.28	$[M + H]^+$	4,8 dimethylnonanoyl carnitine -H <sub>2</sub> or isomer	1.96	
25	328.24842	41.70 - 42.00	328.2	41.95	$[M + H]^+$	4,8 dimethylnonanoyl carnitine -H <sub>2</sub> or isomer	0.51	
26	332.20657	29.40 - 30.00	332.2	30.23	$[M + H]^+$	Nonate + C <sub>7</sub> H <sub>13</sub> NO <sub>2</sub> (carnitine)	-0.57	

27	342.22799	31.90 - 32.40	341.9	32.58	[M + H] <sup>+</sup>	9-Decenoylcarnitine + CO	1.42	
28	344.20615	30.90 - 31.40	344.2	31.37	[M + H] <sup>+</sup>	Decenedioic acid + C <sub>7</sub> H <sub>13</sub> NO <sub>2</sub> (carnitine)	-1.79	
29	356.24373	34.80 - 35.30	356.2	35.35	[M + H] <sup>+</sup>	3-Hydroxy-5, 8-tetradecadiencarnitine - C <sub>2</sub> H <sub>4</sub> or isomer	1.62	
30	356.24265	37.20 - 37.70	356.4	37.62	[M + H] <sup>+</sup>	3-Hydroxy-5, 8-tetradecadiencarnitine - C <sub>2</sub> H <sub>4</sub> or isomer	-1.43	
31	356.24375	38.50 - 39.10	356.1	38.92	[M + H] <sup>+</sup>	3-Hydroxy-5, 8-tetradecadiencarnitine - C <sub>2</sub> H <sub>4</sub> or isomer	1.65	
32	356.27859	44.90 - 45.40	356.3	45.27	[M + H] <sup>+</sup>	trans-2-Dodecenoylcarnitine + CH <sub>2</sub> or isomer	-2.64	
33	358.22209	32.30 - 33.10	358.2	32.97	[M + H] <sup>+</sup>	9-Decenoylcarnitine + CO <sub>2</sub> or isomer	-0.91	
34	358.25828	36.20 - 36.90	358.2	36.71	[M + H] <sup>+</sup>	trans-2-Dodecenoylcarnitine + O or isomer	-1.48	

35	358.25867	38.80 - 39.30	357.8	39.27	$[M + H]^+$	trans-2-Dodecenoylcarnitine + O or isomer	-0.40	
36	365.23186	35.20 - 35.70	365.1	35.72	$[M + H]^+$	Cortisol + H <sub>2</sub> or isomer	-1.05	
37	377.10623	28.10 - 28.50	377.2	28.01	$[M + Na]^+$	6,7-Dimethyl-8-(1-D-ribose)lumazine + CO or isomer	-1.57	
38	386.25346	35.80 - 36.30	386.2	36.36	$[M + H]^+$	trans-2-Dodecenoylcarnitine + CO <sub>2</sub> or isomer	-.070	
39	400.26951	34.70 - 35.20	400.2	35.22	$[M + H]^+$	3-Hydroxy-5, 8-tetradecadiencarnitine + O or isomer	0.34	

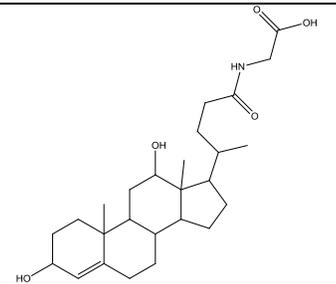
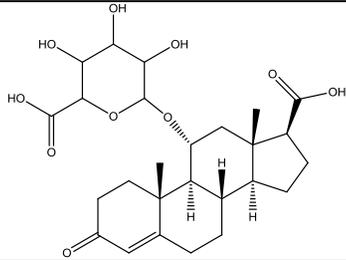
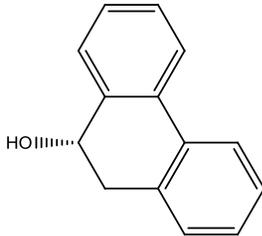
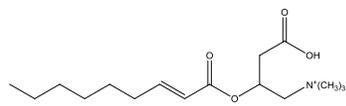
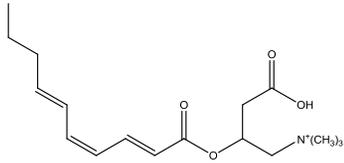
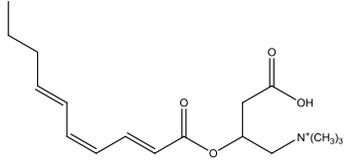
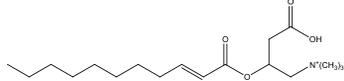
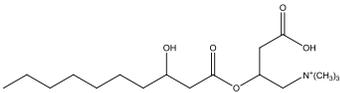
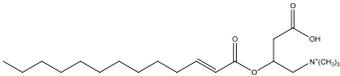
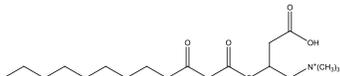
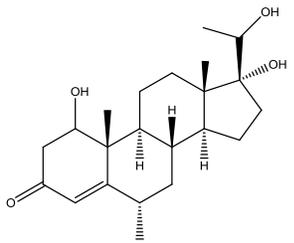
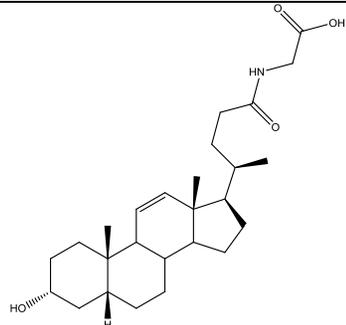
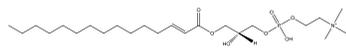
40	448.30556	40.70 - 41.30	448.3	41.16	$[M + H]^+$	Deoxycholic acid glycine conjugate - H <sub>2</sub> or isomer	-0.42	 <p>The structure shows a steroid nucleus with a hydroxyl group at C-3 and a side chain at C-14 consisting of a propyl group and a glycine conjugate (-CH<sub>2</sub>-CH<sub>2</sub>-NH-CH<sub>2</sub>-COOH).</p>
41	531.22010	43.60 - 44.10	531.3	44.05	$[M + Na]^+$	11beta-Hydroxyprogesterone + C <sub>6</sub> H <sub>8</sub> O <sub>6</sub> (glucuronic acid) or isomer	0.05	 <p>The structure shows a steroid nucleus with a ketone at C-3, a double bond at C-4, and a hydroxyl group at C-11. It is conjugated to a glucuronic acid moiety at C-20 via an ester bond. The glucuronic acid moiety has hydroxyl groups at C-2, C-3, and C-6, and a carboxylic acid group at C-5.</p>

Table A6-4 Results from database searching of plasma metabolite features using reaction = 1. MS/MS spectral interpretation was done to confirm structure

Feature ID #	Accurate <i>m/z</i> TOF	RT range (min) TOF	<i>m/z</i> QTRAP	RT (min) QTRAP	Ion type	Putative ID	error (ppm)	Structure
1	197.09619	46.70 - 47.20	197.1	47.90	[M +H] <sup>+</sup>	(+)-(1R,2R)-1,2-Diphenylethane-1,2-diol – H <sub>2</sub> O	0.51	
2	300.21627	35.20 - 35.70	300.1	36.00	[M +H] <sup>+</sup>	2-Octenoylcarnitine + CH <sub>2</sub>	-2.20	

3	310.20077	36.10 - 36.70	310.1	36.90	[M +H] <sup>+</sup>	2-trans,4-cis- Decadienoylcarnitine – H <sub>2</sub>	-1.65	
4	310.20088	35.00 - 35.50	310.2	35.90	[M +H] <sup>+</sup>	2-trans,4-cis- Decadienoylcarnitine (isomer) – H <sub>2</sub>	-1.65	
5	328.24815	40.50 - 40.90	328.2	41.30	[M +H] <sup>+</sup>	trans-2-Dodecenoylcarnitine – CH <sub>2</sub>	-0.31	

6	332.24270	36.30 - 36.80	332.2	37.00	[M +H] <sup>+</sup>	(R)-3-Hydroxydecanoic acid + carnitine	-1.37	
7	356.27946	44.40 - 44.80	356.2	45.30	[M +H] <sup>+</sup>	trans-2-Dodecenoylcarnitine (or isomers) + CH <sub>2</sub>	-0.26	
8	358.25860	39.10 - 39.60	358.2	39.90	[M +H] <sup>+</sup>	3-Oxododecanoic acid (or isomers) + carnitine	-0.54	

9	363.25219	34.80 - 35.40	363.2	35.60	[M +H] <sup>+</sup>	Medroxyprogesterone + H <sub>2</sub> O	-2.19	
10	432.31057	41.80 - 42.30	432.3	42.70	[M +H] <sup>+</sup>	Deoxycholic acid glycine conjugate (or isomer) – H <sub>2</sub> O	-0.59	
11	480.31027	51.10 - 51.60	480.3	52.70	[M +H] <sup>+</sup>	LysoPC(15:0) – H <sub>2</sub>	3.77	

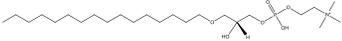
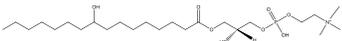
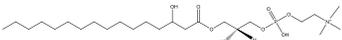
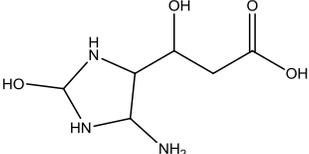
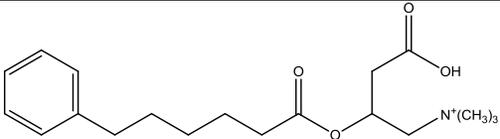
12	482.35911	51.20 - 51.70	482.3	52.30	[M +H] <sup>+</sup>	LysoPC(O-18:0) – C <sub>2</sub> H <sub>4</sub>	-2.87	
13	512.33418	40.40 - 41.00	512.4	41.30	[M +H] <sup>+</sup>	LysoPC(16:1(9Z)) + H <sub>2</sub> O	-0.97	
14	512.33418	47.00 - 47.50	512.3	48.10	[M +H] <sup>+</sup>	LysoPC(16:0) + O	-0.99	

Table A6-5 Results from database searching of urine metabolite features using reaction = 2. MS/MS spectral interpretation was done to putatively identify the structure.

Feature ID #	Accurate $m/z$ TOF	RT (min) TOF	$m/z$ QTRAP	RT (min) QTRAP	Ion Type	Putative ID	Error (ppm)	Structure
1	192.09865	6.10 - 6.60	192.2	6.28	$[M + H]^+$	L-Histidine + $H_2O$ + $H_2O$	4.01	
2	336.21741	36.90 - 37.40	336.1	37.44	$[M + H]^+$	5-Phenylvaleric acid + $CH_2$ + $C_7H_{13}NO_2$ (carnitine)	1.42	
3	338.23302	38.50 - 39.00	338.1	39.03	$[M + H]^+$	Perillyl alcohol + $C_2H_2O$ + $C_7H_{13}NO_2$ (carnitine)	1.28	