University of Alberta

Evaluation of Derivatization Reaction Conditions for Metabolite Labelling Using Differential Isotopic Labelling for Biological Samples Analyzed by Liquid Chromatography Mass Spectrometry

by

Jared Patrick Curle

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

Master of Science

Department of Chemistry

©Jared Patrick Curle
Fall 2013
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.

Abstract

The objective of this work was to assess the impact of reaction conditions for isotopic labeling of metabolites prior to analysis using liquid chromatography mass spectrometry (LC-MS). Differential isotope $^{12}C/^{13}C$ dansyl chloride was used to label metabolites containing either primary amines, secondary amines, or phenolic functional groups and those were determined by both liquid chromatography electrospray ionization Fourier transform ion cyclotron resonance MS (LC-ESI-FTICR-MS) and ultra-performance liquid chromatography with ultraviolet absorbance detection (UPLC-UV). Using UPLC-UV, metabolite standards labeled under varying reaction conditions were quantitatively determined in terms of reaction efficiency comparison. Compared with conventional heating, microwave assisted heating (MAH) methods demonstrated better labeled product yields. In addition ambient temperature reactions were found provide the best labelling without using MAH. The same reaction conditions were assessed for metabolite profiling using LC-MS. It was found that a reaction temperature of 40°C provided the best results based on the number of peak pairs found. However other reaction methods provided unique peak pairs not found in the other methods.

Acknowledgment

I would like to thank Dr. Liang Li for his support and guidance during my research and learning. I would like to thank my committee for the time spent reviewing my thesis and the comments that the graciously provided. The staff of the University of Alberta Mass Spectrometry Lab was very helpful during my research and I appreciate their tutelage and help troubleshooting. Thanks go to all members of the Li group that I had the pleasure of working with and learning from: Azerat Zuniga, Avalyn Stanislaus, Ming guo Xu, Ruokun Zhou, Tammy Zheng, Tran Tran, Tao Huan, Chiaoli Tseng, and Yiman Wu.

Finally I would like to thank all of my family and friends for their enduring support in all my endeavours.

Table of Contents

Chapter 1 Introduction 1

1.1 General Introduction 1

1.1.1 Targeted Metabolomics 2

1.1.2 Profiling Metabolomics 5

1.1.3 Derivatization and Differential Stable Isotope Labeling for Metabolomics 11

1.2 Liquid Chromatography 14

1.2.1 High Performance Liquid Chromatography (HPLC) 15

1.2.2 Ultra Performance Liquid Chromatography (UPLC) 17

1.3 Mass Spectrometry 18

1.3.1 Electrospray Ionization (ESI) 18

1.3.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry 21

1.4 Overview of Thesis 25

1.5 Literature Cited 26

Chapter 2 :UPLC-UV Quantification for the Determination of Optimal Derivatization Reaction Conditions for Isotope Labelling Metabolomics Studies 32

2.1 Introduction 32

2.2 Experimental 35

2.2.1 Chemicals and Reagents 35

2.2.2 Sample Preparation 35

2.2.3 Derivatization Reactions 35

2.2.4 UPLC-UV Separation and Detection 36

2.2.5 Method Validation 37

2.3 Results and Discussion 39

2.3.1 LC Gradient Method Development 39

2.3.2 Quantification of Labeled Metabolites 54

2.3.3 Method Validation 72

2.4 Conclusions 84

2.5 Literature Cited 85

Chapter 3 : Evaluation of Reaction Conditions for Differential Isotope Labeling for Metabolomic Profiling of Biological Samples by Liquid Chromatography Mass Spectrometry 88

3.1 Introduction 88

3.2 Experimental 89

3.2.1 Chemicals and Reagents 89

3.2.2 Sample Collection and Processing 90

3.2.3 Derivatization Reactions 91

3.2.4 UPLC-UV 92

3.2.5 LC-FTICR 92

3.2.6 Method Validation 93

3.2.7 Data Processing and Statistical Analysis 94

3.3 Results and Discussion 94

3.3.1 Assessment of Derivatization Conditions for Dansyl Labelling of Biofluids .. 94

3.3.2 Urine and Plasma Samples with Varying Reaction Temperatures 111

3.3.3 Comparison of Conventional and Microwave Assisted Heating 129

3.3.4 Method Validation 134

3.4 Conclusion 136

3.5 Literature Cited 136

Chapter 4 : Conclusions and Future Work 139

List of Tables

Table 2-1 List of compounds in standard mixture used for method development............ 39

Table 2-2 Experimental Conditions that affect α* in Reverse Phase Chromatography... 48

Table 2-3 Mixture of dansylated compounds for comparison of reaction conditions and retention times from UPLC-UV analysis. 55

Table 2-4 Repeatability (RSD), R^2 and relative response vs. internal standard (Dns-Gly) at four different concentrations of amino acid standard (n=3). Total metabolite concentrations are 28, 10, 1, and 0.1 mM. Reaction conditions are: Conventional heat; 60 minutes, 60°C. 73

Table 2-5 Repeatability (RSD), R^2 and relative response vs. internal standard (Dns-Gly) at four different concentrations of amino acid standard (n=3). Total metabolite concentrations are 28, 10, 1, and 0.1 mM. Reaction conditions are: CEM Discover, 200 s, and 100 watts. 75

Table 2-6 Results of external calibration absolute quantification using purchased predansylated standards. Initial concentration is the measured amount of metabolite added initially to the reaction; final concentration is the amount calculated using the calibration curve. Microwave reaction is 200 s at 100 w, conventional reaction is 60 minutes at 60°C. 80

Table 2-7 Intra-day, inter-day and method reproducibility of the analytical method. 82

Table 2-8 LLOD and LLOQ of internal standard and standards used for external calibration. 84

Table 3-1 Peak area results before and after mixing of labeled urine samples with different reaction conditions. Volumes given are prior to mixing and further UPLC-UV analysis shown in the after mixing section. 102

Table 3-2 Labeled metabolites that appear in 9 of 9 urine samples analyzed and are unique to the reaction temperature used. Those with more than one name have multiple HMDB matches. 121

Table 3-3 Labeled metabolites that appear in 9 of 9 plasma samples analyzed and are unique to the reaction temperature used. Those with more than one name have multiple HMDB matches. 124

List of Figures

Figure 1-1 Omics cascade and systems biology. 3

Figure 1-2 Representation for a typical workflow in LC-MS based profiling metabolomics. 9

Figure 1-3 Dansyl labeling reaction scheme (A). Dansyl molecule with chemical specialties highlighted (B). Workflow for Dansyl labeling (C). 13

Figure 1-4 Representation of the electro spray ion process which leads to ionization of analyte molecules and transfer to a mass analyzer. 21

Figure 1-5 Sequence of events involved in FTICR-MS excitation and detection. 23

Figure 1-6 Ion cyclotron motion where an ion of mass m and charge q is constrained to a circular orbit by the magnetic field, B. This orbit has a characteristic angular frequency ωc or the cyclotron frequency (A). Ion excitation and detection (B). Depiction of ICR cell (C). 24

Figure 2-1 Representative UV spectra of four selected Dansylated metabolites obtained with a UV-Vis PDA detector (A). Full PDA spectrum of dansylated amino acid standard mixture (B). 42

Figure 2-2 Representative UPLC-UV Chromatograms of the purchase Sigma® standard (1.25 mM each) (A) and the standard created from individual compounds (3 mM each) (B). Dansylated compounds are identified by their one letter amino acid code and a list with corresponding retention times can be found in table 2-1. Highlighted boxes in both A and B show groups of co-eluting compounds. 44

Figure 2-3 Gradient method development UPLC separation chromatograms with varying gradient times. k*=5 with an 8 minute gradient (A). k*=15 with a 19 minute gradient (B). k*=20 with a 26.5 minute gradient (C). Conditions: All methods had 10% B initial and

80% B final with a linear gradient, H_2O:ACN w/ FA; flow rate, 0.65 mL/min; column temperature = 30°C... 47

Figure 2-4 UPLC-UV chromatograms comparing separation temperatures – 30°C (A) and 50°C (B). The arrows denote where a reduction of resolution between peaks occurred at the higher temperature.. 49

Figure 2-5 UPLC-UV chromatograms comparing separation of standards on a 15 cm column (A), and a 5 cm column (B & C). k* values are also given for comparison. 50

Figure 2-6 UPLC-UV chromatograms comparing acetonitrile (A) and methanol (B) used as different B phase solvent types.. 52

Figure 2-7 UPLC-UV chromatograms of labeled metabolites and the final UPLC method showing segmented chromatogram... 53

Figure 2-8 UPLC-UV chromatogram of the mixture of dansylated compounds used for reaction condition assessment in this project. Each dansylated metabolite is labeled with a three letter amino acid code. 54

Figure 2-9 UPLC-UV quantification of labeled amino acids under different reaction temperatures and reaction times. Relative response is the peak area ratio of each labeled peak to that of the internal standard Dns-glycine. Reaction temperatures are given in the legend with ambient temperature (AT, 21-24°C), 40°C, 60°C, or 80°C. 58

Figure 2-10 UPLC-UV quantification of labeled amino acids labeled using a CEM Discover microwave reactor. Relative response is the peak area ratio of each labeled peak to that of the internal standard Dns-Glycine. The amount of power applied (w) to each reaction is given in the legend.. 63

Figure 2-11 The effect of microwave power and time on the amount of dansyl labeled amino acid using a CEM Discover microwave. Power comparison was done for a reaction time of 120 s (A). Time comparison is done with 100 watts applied power (B). 64

Figure 2-12 UPLC-UV quantification of labeled amino acids labeled using a Panasonic consumer microwave. Relative response is the peak area ratio of each labeled peak to that of the internal standard Dns-Glycine. The amount of power applied (power level) is given in the legend. 68

Figure 2-13 The effect of microwave power and time on the amount of dansyl labeled amino acid using a Panasonic consumer microwave. Power comparison was done for a reaction time of 120 s (A). Time comparison is done with 200 w applied microwave power (B). 69

Figure 2-14 Comparison of quantification results for various derivatization methods. The conventional heating methods were both reacted for 60 minutes while the microwave reactions took place over 2 minutes. 71

Figure 2-15 Quantification of amino acid standards using UPLC-UV over the range of 0.1 mM- 28 mM total metabolite concentration. Samples derivatized using conventional heating at 60°C for 60 minutes. 77

Figure 2-16 Quantification of amino acid standards using UPLC-UV over the range of 0.1 mM- 28 mM total metabolite concentration. Samples derivatized using CEM discover microwave heating for 2 minutes at 100 watts of power. 78

Figure 2-17 External calibration curves for Dns-aspartic acid and Dns-threonine with concentrations ranging from 0.01 mM to 2 mM using a 5 μL injection volume. 79
Figure 3-1 Workflow for LC-MS metabolomic analysis of biofluids using dansyl chloride for differential labeling of biological samples. 96

Figure 3-2 Representative UPLC-UV spectrum for a ^{12}C dansyl labeled urine sample (A) and a labeled plasma sample (B). Calibration curves for urine (C) and plasma (D)......... 98

Figure 3-3 UPLC-UV total peak area of dansyl metabolites for urine samples before ^{12}C/^{13}C mixing (A) and after mixing (B). Reaction conditions used for each sample are given as time-temperature-reaction method. (c = conventional oven, at = ambient temperature, mic = microwave assisted heating) 101

Figure 3-4 Total ion chromatogram from LC-FTICR analysis of dansyl labeled urine. 104

Figure 3-5 Mass spectrum of dansyl labeled glycine from analysis of labeled urine sample (A). Extracted ion chromatogram of both ^{12}C and ^{13}C labeled glycine (B). 105

Figure 3-6 Number of differentially labeled peak pairs obtained using (A). Extracted ion chromatograms for each different reaction type (B). 108

Figure 3-7 Partial least squares discriminant analysis (PLSDA) analysis of urine samples using variable reaction conditions.. 110

Figure 3-8 UPLC-UV quantitation of urine (A) and plasma (B) samples after derivatization with $^{12}C/^{13}C$ DnsCl and mixing of the separate reactions. 112

Figure 3-9 LC-MS total ion chromatograms comparing replicate samples of urine (A) and plasma (B). 114

Figure 3-10 Spectra of peak pair classification level 1 (A), level 2 (B), and level 3 (C). Mass labels define which peaks are being discussed. 116

Figure 3-11 Peak pair counts from samples of urine (A) and plasma (B). The data presented is separated into peak pair level and the color coding is explained in the legend. 119

Figure 3-12 Number of metabolites found in each of the replicates on the left and those found in at least 6 on the right. Urine (A) and plasma (B) samples. 120

Figure 3-13 PCA score plots for urine samples (A) and plasma samples (B). 128

Figure 3-14 Peak pair amount and level comparison for a samples derivatized using both conventional and MAH heating. Reaction times for microwave reactions are given in seconds not minutes (i.e. 120 and 300 s) 130

Figure 3-15 PCA plot of sample labeled using conventional heating in red and MAH in green (A). PLS-DA plot of each separate reaction (B). 132

Figure 3-16 Comparison of unique peak pairs found in MAH reactions and conventionally heated reactions. The number in the middle is common peak pairs found in the two samples. These peak pairs were found in at least 2 of the 3 replicates assessed for each reaction type. 133

Figure 3-17 Representative total ion chromatograms for acetonitrile blank (A), reaction mixture blank (B), labeled amino acid standard (C), raw urine (D), and ACN blank after three successive labeled urine blank samples to assess carryover (E). 135

List of Abbreviations

%RSD	% Relative Standard Deviation
ACN	Acetonitrile
AC	Alternating current
APCI	Atmospheric pressure chemical ionization
API	Atmospheric pressure ionization
$(\Delta tr)_{avg}$	Average Difference in Retention Time
$(\Delta t_{r)avg}$	Average Retention Time
CE	Capillary electrophoresis
CSF	Cerebrospinal fluid
CRM	Charge residue model
CI	Chemical ionization
R_s	Chromatographic Resolution
CV	Coefficient of variation
CE	Collision energy
CID	Collision-induced dissociation
CMO	Consumer Microwave Oven
Da	Dalton
DnsBP	Dansyl byproduct
Dns	Dansyl
DnsCl	Dansyl Chloride or 5-(dimethylamino)naphthalene-1-sulfonyl chloride
DP	Declustering potential
Δt_R	Difference in Retention Time
DIL	Differential isotope labeling
DC	Direct current
EI	Electron impact ionization
EI-MS	Electron impact ionization mass spectrometry
ESI	Electrospray ionization
EMS	Enhanced MS
EPI	Enhanced product ion
FAB	Fast atom bombardment
FFT	Fast Fourier transformation
FA	Formic Acid
FT-ICR-MS	Fourier Transform ion cyclotron resonance mass spectrometry
GC-MS	Gas chromatography mass spectrometry
k*	gradient retention factor
α*	gradient selectivity
t_g	gradient time
HPLC	High performance liquid chromatography
B	Homogeneous magnetic field
HMDB	Human Metabolome Database

HILIC	Hydrophilic interaction chromatography
IEM	Ion evaporation model
LOQ	Limit of Quantitation
LIT	Linear ion trap
LC	Liquid chromatography
LC-MS	Liquid chromatography mass spectrometry
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LLOQ	lower limit of quantitation
MS	Mass Spectrometry
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionization
MeOH	Methanol
PC	Microwave power control
TTC	Microwave time temperature control
MAH	Microwave-assisted heating
MCI	Mild cognitive impairment
M	Molar
MRM	Multiple reaction monitoring
NMR	Nuclear magnetic resonance
PLS-DA	Partial least squares-discriminant analysis
DmPA	*p*-dimethylaminophenacyl bromide
PDA	Photo Diode Array
PCA	Principal component analysis
QIT	Quadrupole ion trap
Q-TOF-MS	Quadrupole-time of flight-mass spectrometry
QC	quality control
rf	Radio frequency
t_R	Retention time
RPLC	Reverse phase liquid chromatography
s	seconds
SIL	Stable isotope labeled
MS/MS	Tandem mass spectrometry
TLC	Thin layer chromatography
TOF	Time of flight
QqQ	Triple quadrupole
2D	Two-dimensional
UV	ultraviolet
UV-Vis	Ultraviolet-visible
W	Watts

Chapter 1 Introduction

1.1 General Introduction

Metabolomics is a relatively new field of study that is based on the detection and quantification of endogenous small molecules in any number of biological samples. This type of research is being applied in diverse fields such as nutrition,[1] drug development,[2] agriculture,[3] clinical diagnostics,[4] and environmental chemistry and toxicology.[5] The common definition of a metabolomics study is that it determines the molecular make up of a biological sample under a certain size (<1500 Da) in samples of biological origin. This is the final link of the 'omics' cascade that include genomics, transcriptomics, and proteomics. Metabolomics studies commonly use either nuclear magnetic resonance (NMR) or mass spectrometry (MS) for detection and classification of molecules based on either spectral interpretation or mass to charge ratio, respectively. Both of these instruments have been used with for number of studies publishing articles based on these techniques (see Figure 1 from [6]). NMR is a useful technique that is useful for determination of compound identification by structural interpretation.[7] Despite this characteristic it is challenging to couple NMR detection with separation techniques such as liquid chromatography (LC), gas chromatography (GC) and capillary electrophoresis (CE).[8-9] In addition NMR generally is less sensitive than MS techniques. On the other hand MS has been successfully coupled with these techniques and they are now in widespread use.[10-12] GC requires that metabolites are volatile in nature and therefore limits the metabolome available for analysis by this technique.[13] Although CE has been linked with MS, robust techniques required for metabolomics are still not in place.[14] Often the technique of choice for robust metabolomics study is liquid chromatography mass spectrometry (LC-MS) coupled with an electrospray ionization (ESI) interface. In

fact LC-MS has become so routine that it is used as a clinical diagnostic technique.[15] As mentioned MS type detection makes up about fifty percent of all metabolomics study.

Even though MS is powerful and widely applicable it can still not determine each and every metabolite contained in complex biological samples. This is due to the fact that metabolites vary widely in physiochemical properties unlike genomics and proteomics, that have much simpler molecular make ups with four and twenty repeating units respectively. To qualitatively and quantitatively determine this molecular diversity our research group employs a strategy that uses chemical derivatization for metabolite labeling. This labeling achieves a number of things, improved chromatography, enhanced ESI sensitivity, and differential isotope labels for relative and absolute detection and quantification. Derivatization labels like dansyl chloride, used in this work, are chemically targeted for specific functional groups. Dansyl chloride reacts with primary/secondary amines and phenols, thereby acting as a means of data reduction. This thesis is based on the application dansyl chloride for differential isotope labeling of metabolites in different biofluids with separation and detection by LC-MS under variable derivatization reaction conditions.

1.1.1 Targeted Metabolomics

Metabolomics as a field strives to detect and characterize the small molecules population of any number of biological samples. The sub-field of targeted metabolomics is a less ambitious approach that focuses on small groups of analytes, as little as one or two and as many as hundreds. Like other metabolomics approach this is achieved through detection and identification of analytes with instruments like NMR and MS. NMR has the advantage of little to no sample preparation required for analysis. Despite this, NMR has less sensitivity than MS, and targeted analysis often requires the sensitivity to detect metabolites in low abundance. When MS is coupled with a separation technique and

proper sample preparation techniques are developed, highly sensitive and accurate methods with very good dynamic ranges can be realized.[16-17] There are a number of separation techniques,

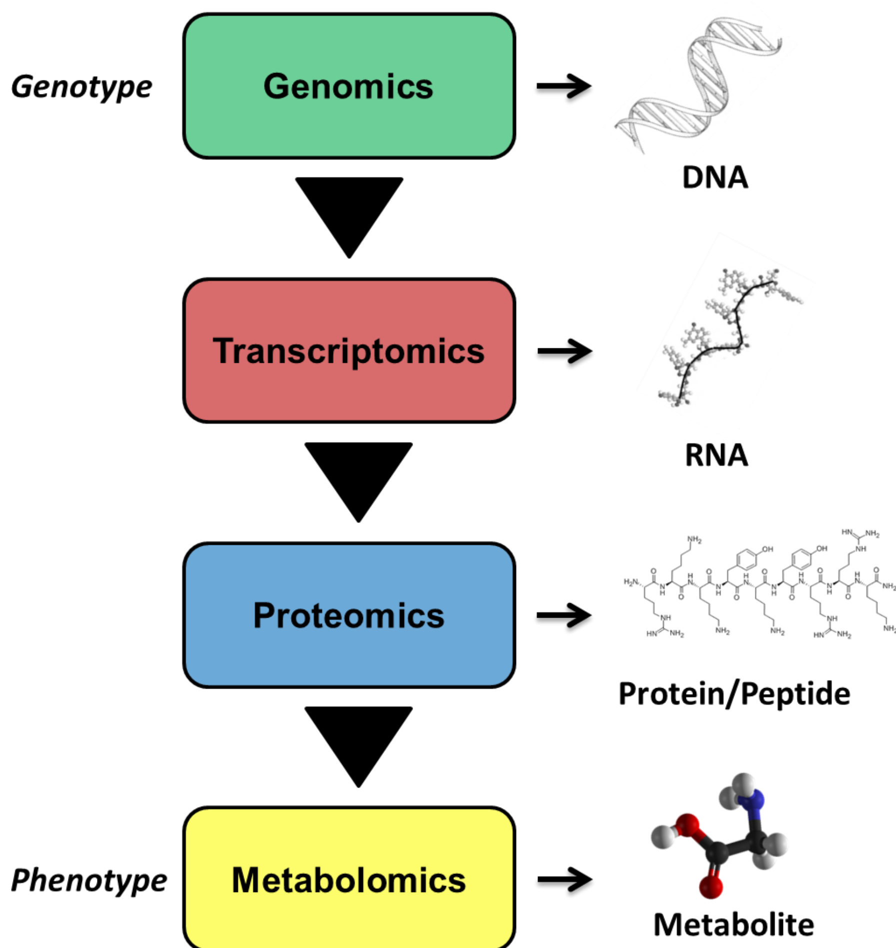


Figure 1-1 Omics cascade and systems biology.

ionization sources, and mass spectrometers that have been employed in targeted metabolic analysis. As discussed GC[18], CE[19], and most notably HPLC[20] (and UPLC) have been used extensively for the reduction of sample complexity through separation by different fundamentals. LC-MS analysis requires a specific interface to ionize samples,

examples are; electrospray ionization (ESI), atmospheric pressure ionization (APCI), or atmospheric pressure photo ionization (APPI).[21] There are a number of different mass spectrometers employed such as quadrupole (QP), ion traps (IT), and linear ion traps (LIT) which all have very good sensitivity but low resolution. For higher mass resolution analysis users can turn to quadrupole time of flight (QTOF), FT-ICR, or Orbitrap instruments.[22-24]

The basic workflow of targeted metabolomics analysis like most other analytical disciplines starts with sample preparation. After which the sample is separated, ionized, and finally analyzed on an MS. The sample preparation process depends on the type of sample being analyzed. Common sample types include urine, blood (plasma or serum), cerebral spinal fluid (CSF), saliva, cells, tissues, or environmental samples. After collection an extraction step is carried out in order to reduce the complexity of the sample. Some samples such as plasma, serum, or saliva need a protein precipitation with organic solvent. Others, such as tissues and cells, require quenching with rapid cooling or rapid heating to denature proteins if analytes are thermally labile. Cells and tissues also require a mechanical method to release metabolites into solution. Following sample preparation most samples are stable if stored at low temperature (-80°C). Chromatography will be discussed in greater detail in a later section but briefly; samples are injected into a chromatographic system and separated on a column that works by causing chemicals to retain on a stationary phase. This separates components before they are ionized. ESI is the most popular choice and acts by charging molecules so they can be manipulated by electric and magnetic fields in the mass spectrometer.

The choice of mass spectrometer for a study depends on the needs of the analyst. Often low resolution is acceptable and sensitivity is important so QT, IT, or LIT are used. This is often the case in targeted metabolomics and these are the instruments of choice.

Specifically the triple quadrupole (QqQ) instrument is used for this type of analysis. There are two modes most often used; multiple reaction monitoring (MRM) and high resolution full scan. MRM mode is the most often used and it takes advantage of multiple quadrupole' s by selecting a mass in the first quadrupole, fragmenting, and then finally determining the specific fragments of each molecule for an accurate identification of analytes. This mode takes some time for each metabolite of interest and therefore when an analysis becomes more complex it is impossible to quantify these all accurately. Future analysis looks towards using full scan modes with better resolution than that of the QqQ in order to perhaps realize the goal of combining targeted and profiling metabolomics studies.

Targeted metabolomics are currently important for relative and absolute quantitation of metabolites in biological samples. For these to be valid each of the steps mentioned above must be robust and validated. It is the future goal of many to try and take targeted type quantification to a larger stage and apply it whole metabolomes.

1.1.2 Profiling Metabolomics

Other 'omics' studies such as genomics and proteomics have met the challenge of determining their respective targets using specific methods. Metabolomics however has not yet reached its potential of being an approach than can quantitatively and comprehensively analyze all of the metabolites found in biosystems thereby revealing their 'metabolome', as was first defined by Fiehn.[25] There have been successes within the field as signalled by the commercial availability of kits that provide metabolomic data for embryonic *in vitro* fertilization selection.[26] There are still a number of major hurdles for field to become successful. These mainly lie in the field of MS based metabolomics as they pertain to this study. As has been discussed there, are many separation techniques, interfaces, and mass analyzers. This diversity in instrumentation throughout the field is a

large reason that there is a lack of cohesiveness between separate studies. Major issues involved with this are standardization of methods, robustness, and reproducibility. There is a current attempt to try and address this through the development of field wide standards, as has been seen in other fields. [27] This evolution is reasonable as during the initial years of metabolomics study many would ‘pick the low hanging fruit’ and publish easy and variable studies. There was little in the way of standardization to preclude certain work from being published. There are also a number of initiatives to produce databases of determined metabolites; examples of these include the human metabolome database (HMDB), METLIN, and My Compound ID (MCID).[28-31] This is not perfect and many groups are often doing replicate analysis that otherwise could be documented. Notwithstanding these problems the field has had results and will continue to in the future.

It is hard to draw a clear line between targeted and profiling metabolomics. Both are analyses that look at multiple small molecule analytes using the same type of instrumentation. Some targeted studies look at hundreds of metabolites which makes it hard to define what targeted really means.[32] To define profiling metabolomics we must remove the defined hypothesis, often this type of analysis is looking to determine only what is in a sample without specific pre-knowledge. Another defining factor for these types of analysis is that they employ advanced and multivariate statistical analysis. Targeted analysis can often can be easily be done as as it relies heavily on MRM and other MS/MS modes as discussed.

The workflow of a profiling metabolomics analysis is much like that of a targeted analysis with the major steps being experimental design, sample collection and preparation, analysis (LC-MS), and data processing and statistical analysis. Often overlooked but very important is the role of experimental planning and sample collection.

In order for any statistical results at the end of study to be meaningful, the initial experimental plan must be sound. Important are the steps of sample preparation which are often overlooked due their assumed simplicity. If standardized methods of preparation and storage are not employed, large variation in results are often observed. It is important in profiling studies that there be the least amount of steps possible and that any chemically or otherwise harsh treatments be minimized. Due to complex samples that have many different physiochemical properties any perturbations can cause true biological results to be lost. To outline this, a number of studies have examined the effect of altering sample preparation techniques on profiling results. [33-34] Also some studies use derivatization for the improvement of analysis by improved chromatography, ionization, detection, data handling etc.[35-36]

Profiling metabolomics presents many challenges to the actual analysis of biological samples. Studies often involve many samples that may need to be analyzed with different conditions and at least 2 replicates. This large sample load drives the need for efficient but still comprehensive analysis. The major time concern is the LC (or other separation method) needed to reduce the sample complexity. LC runs need to be of a certain length in order to provide adequate peak capacity and peak width for proper MS analysis. There are two types of chromatography commonly used; they are reverse phase (RP) and hydrophilic interaction chromatography (HILIC).[37] By using different, orthogonal separations it is often possible to increase the metabolome coverage. In addition 2-dimensional techniques with orthogonal columns or column switching.[38] Furthermore ultra-performance liquid chromatography (UPLC) or ultrahigh pressure liquid chromatography (UHPLC) offer faster separation for profiling.[39]

Ionization is important for this type of analysis, as already mentioned there are a number of available choices each often giving a different result. ESI in both positive and

negative modes and APCI are the options available. Although the use of these is routine in many labs, the difference in sample molecules can often lead to challenging optimization of condition. There are studies that assess how changes in ionization can affect the final data of an experiment.[40] To analyze the ionized molecules there are a number of different MS instruments used. Most studies rely on QTOF instruments, because of their sensitivity, data acquisition speed (important with fast chromatography), high mass accuracy, and acceptable resolution. High resolution instruments such as FTICR-MS and Orbitrap are useful especially for exact mass determination and applications that require high resolution. The downside of these instruments is their prohibitive cost and slow scan speed. Instrument choice is not all important however as there is often further analysis (e.g. NMR, retention comparison) for true conformation of metabolite identification.

Data from LC-MS experiments is often very complicated and requires the application of unique statistical strategies and software for it to be useful. Raw data is m/z ratio for each scan based on retention time. For high or medium resolution instruments used in profiling this can add up to hundreds of thousands of data point for each scan. If an analysis is twenty minutes with 4 seconds for a type FTICR-MS scan there will be 300 individual scans, this results in tens of millions of data points. All of this is obviously not useful metabolite data. In addition to useful signals there is noise, adduct ions, multiply charged species, fragments, or poor chromatographic peaks to consider. To make this useful there are a number of steps usually followed; noise filtering, baseline correction, centering, normalization, peak picking, peak integration, and retention time alignment. The data analyzed is generally in the form of matrix with axis represented by m/z and retention and peak intensity used for each corresponding point. There are a number of

peak picking software options available. For example vendor software (Masshunter, Markerlynx), independent developer software,

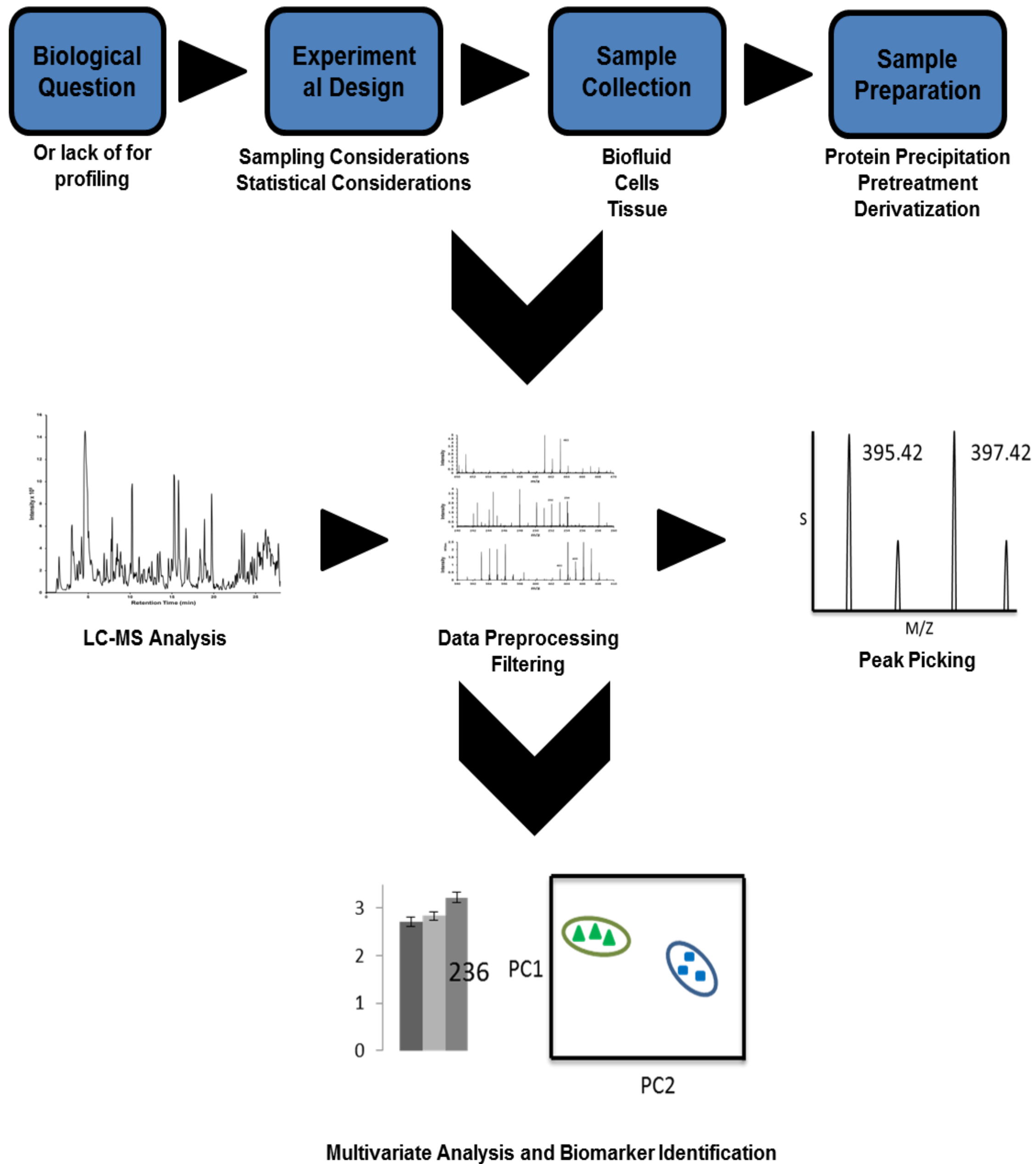


Figure 1-2 Representation for a typical workflow in LC-MS based profiling metabolomics.

open access software (XCMS, Metlalign), in-house script development (using R, Java, Excel etc.). Choosing which software the analyst should use often depends on what the goal of the analysis is, the control they want over data handling, and the experience or capability of the analyst to create their own software. Most studies will compare multiple replicate samples and chromatographic runs and therefore alignment software is important. The two factors to consider in this are retention tolerance and mass tolerance. Although retention times should be relatively stable analysis over long periods of time can experience large shifts. There is also the matter of what type of chromatography is being used (UPLC or HPLC) when thinking about the peak width. Also mass accuracy should be high but small variations make the choice of mass tolerance an important one. After peak lists have been extracted it is important to manually inspect the data and use proper quality control samples to ensure the data qualifies. Finally profiling data is analyzed by multivariate analysis that will be discussed in further detail in Section 1.4 of this thesis.

The overall goal of profiling is the identification of actual metabolites, also known as biomarker identification. This is often the most challenging aspect of the analysis because: comparison of data from MS to MS is hard[41] and adducts formed in LC-MS through ionization cannot be safely predicted or managed. Although a high resolution mass spectrometer is necessary it is not the only step required for marker validation. By searching metabolite or other chemical databases one can create a short list of possible matches. From these isotopic patterns, MS/MS experiments, retention comparison, or using complimentary techniques (CE-MS, GC-MS, NMR) can provide the structure of the biomarker of interest. It is important to note that this process is by no means trivial. Further the validation of any method for which there are many considerations is often complex. Proper control and QC samples can increase confidence

in the results. Replicate samples provide a means of reducing false positives. A current recommendation is the "80% rule" in which a metabolite must be found in 8 of 10 samples in order to be further analyzed.[42]

The area of profiling metabolomics employing LC-MS is an already large and growing facet of metabolomics studies. This interest has and will continue to drive improvements in technology and methods within the field. This improvement will perhaps reveal as yet undiscovered metabolites or help to qualify those which have already been detected. Creation of standard practices will be important, so that data can be shared between investigators.

1.1.3 Derivatization and Differential Stable Isotope Labeling for Metabolomics

Metabolites in their native state often pose challenges for metabolomics analysis with LC-MS. There are three main areas for which this technique suffers the most; first is the ionization of analytes, second is the separation of analytes, and third is the accurate detection and identification of analytes. To improve all of these problems a strategy of chemical derivatization can be employed. Derivatization has been used extensively in other analysis types such as GC,[43] GC-MS,[36, 44] LC with UV/VIS detection,[45] and non-metabolomics LC-MS analysis.[46] To reduce this complexity it is possible to apply a chemical derivatization strategy that will effectively fraction the metabolome based on functional groups. This may seem like added work but it has consistently been mentioned that because of the chemical diversity present in the metabolome, no one technique or method can be used for total coverage.[47] The advantages outlined above do not come without drawbacks. Derivatization adds another sample preparation step, and along with it the increased variability in samples due to reaction efficiency differences, byproduct formation, and possible metabolite degradation. Despite these problems this approach is believed to have enough positive payback to warrant application and further

study. Derivatization strategies target different functional groups found widely on metabolites in order to tailor chemical tags. These chemical tags commonly contain an ionizable charged group, added hydrophobic bulk, and finally the ability to easily attach isotopic labels. These tags will often increase the ionization of different groups such as lipids or to improve chromatography of highly polar molecules.[48-50]

Using high resolution spectrometers yields exact m/z that allows for putative metabolite identification. This putative match is only the first step in determining the true identity of a metabolite. In addition isotopic profile, chemical rules, elemental ratios, and chemical standards can be used to help with putative formula determination. There have been a number of metabolomics studies that employ stable isotope labeling (SIL) of metabolites for LC-MS. SIL compounds are very common in LC-MS analysis, especially as internal standard for absolute and relative quantification analysis. This allows for discrimination of metabolites based on increased mass and a change in isotopic pattern.[51-52] Often this analysis uses cells or other extracts that grown or soaked in media with isotopic labeled nutrients or other molecules.[53] This approach takes advantage of the characteristic shift in isotopic pattern or the appearance of a light and a heavy labeled peak when samples are differentially labeled. By creating isotopically labeled derivatization reagents it is possible to take advantage of both the chemical improvements in retention and ionization and also the spectral differences for data analysis and peak detection. Because of the specificity of the chemical derivatization reaction, this approach can effectively fractionate the metabolome based on functional group. A number of studies have successfully used this approach to determine metabolic profiles of amine and phenols and carboxylic acid containing metabolites.[54-58] The application of this technique can be used of both relative and absolute quantification, easier peak detection, and improved chromatography.

This thesis will use Dansyl chloride as a derivatization reagent for isotopic

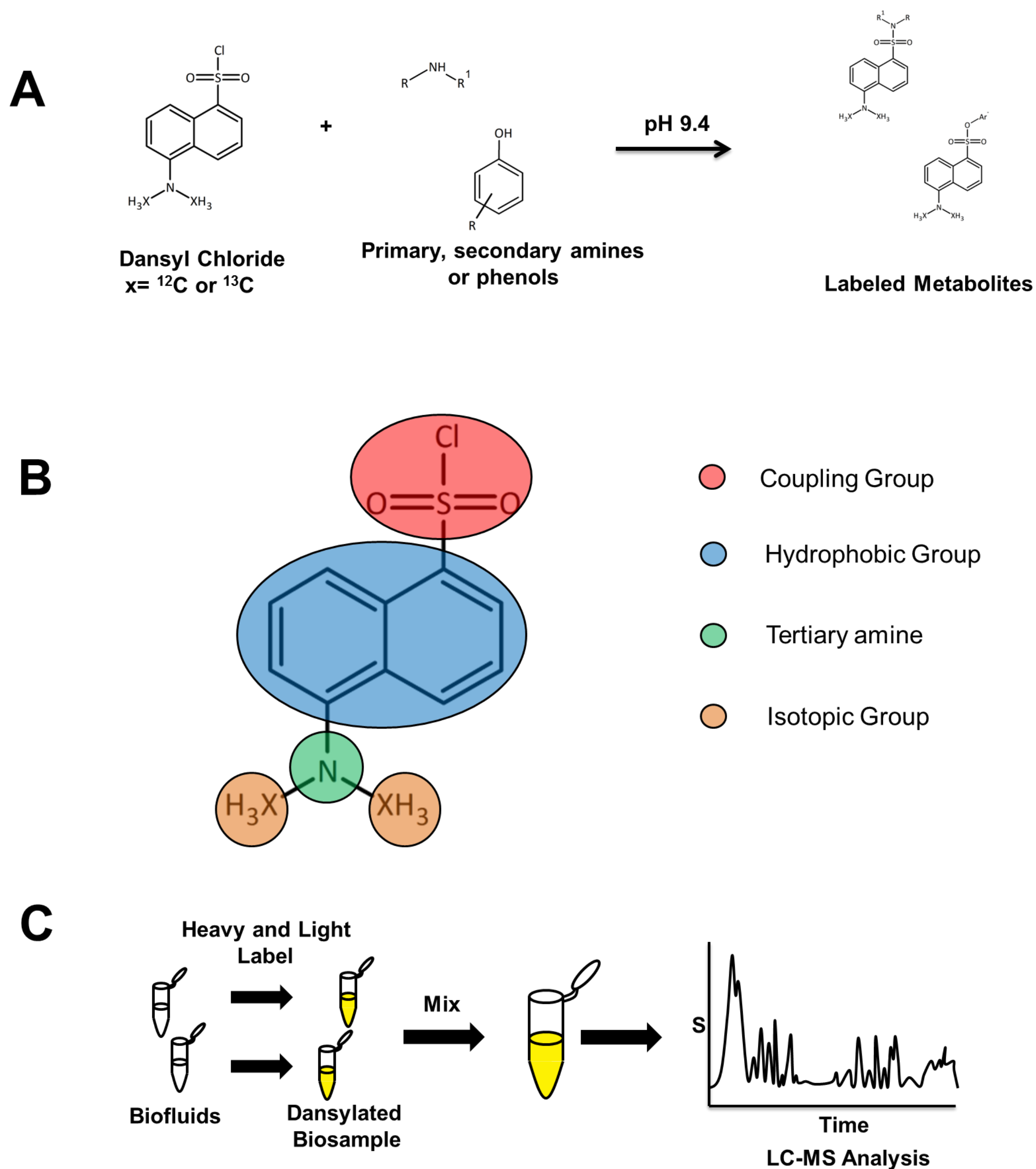


Figure 1-3 Dansyl labeling reaction scheme (A). Dansyl molecule with chemical specialties highlighted (B). Workflow for dansyl labeling (C).

labeling. This compound is well studied, simple and robust and been used for numerous other applications including pre column LC derivatization for signal enhancement in UV and fluorescence detection.[59-61] Studies using this compound focused on targeted type analysis of many metabolic compounds that include biogenic amines,[62] ß-estradiol and estrone[62], and phenol containing drug metabolites[63]. The dansyl label contains a hydrophobic naphthalene group (Figure 1-3 B) that makes relatively polar metabolites to elute much later in the chromatogram and therefore in a higher proportion of the less polar organic mobile phase. This increases the ionization desolvation rate and stabilizes the electrospray in ESI thereby improving analyte response. Also because of the increased hydrophobicity the analyte has increased surface affinity in the ESI source. Also of importance is the tertiary amine group that increases the chargeability of the molecule. This is important as molecules compete for a limited amount of charge within the ESI spray droplet. As well as these enhancements Dansyl labeling shifts most low mass metabolites (50-150 Da) into a higher mass region which contains less chemical noise due to solvent and other factors.[64] These enhancements increase the confidence and coverage of metabolomic analysis of the phenol and amine containing metabolome.

1.2 Liquid Chromatography

Chromatography as a technique has its roots in early separations of mixed color compounds on paper. This led to the further development of chromatographic separation techniques for analysis and industrial purposes. The specific mechanism by which this separation occurs is variable but in general it involves the equilibration of solute (analyte) between two different phases. The two separate phases are known as the stationary and the mobile phase. Stationary phases vary greatly and have many different properties depending on the application and often consist of a liquid or a solid attached to some sort of support. The mobile phase is either gas or liquid. The movement of solutes through the

column is based on the affinity of the solute for the stationary and the mobile phase. That is a solute with a higher affinity for the stationary in comparison with others would spend more time not moving down the column and therefore would retain longer. The types of chromatography are adsorption, partition, ion-exchange, molecular exclusion, and affinity. Modern liquid and gas chromatography mostly rely on partition chromatography and that is what will be discussed in the following sections and what is used for this study.

1.2.1 High Performance Liquid Chromatography (HPLC)

Liquid chromatography is one specific type of modern chromatography routinely used in almost every facet of chemistry, biology, and many other fields. Liquid chromatography fills the large gap left by gas chromatography for molecular analysis of non-volatile and thermally unstable compounds. By reducing the size of the particle in the stationary phase the rate at which the solute transfers between the liquid and stationary phase increases and therefore the efficiency (peak sharpness) of the separation increases. The downside to decreasing particle size is concurrent increase in backpressure. To manage this increase LC systems are built with components that can withstand the more extreme condition, these are known as high-performance liquid chromatography (HPLC) systems. The most commonly used form of HPLC is reversed phase (RP) chromatography. RPLC employs a non-polar stationary phase that is often microporous silica particles with a bonded phase of hydrocarbon chains of variable length with C_{18} being the most commonly used. There are many other types of stationary phases available and chromatography modes available but those will not be discussed here. The mobile phase used in RP is most often a mixture of two miscible solvents with differing polarities. The most common mixture is water with a less polar solvent such as acetonitrile (ACN) or methanol (MeOH) and can be mixed with buffers or other additives

in relatively low concentrations depending on the analysis. Once a sample is injected the different solute components will interact and partition differently with the mobile and stationary phase. Therefore each solute will move through the column at a different rate. Polar metabolites will interact less with the non-polar stationary phase and more with the relatively polar water in the stationary phase and hence elute earlier. On the other hand non polar solutes will interact more strongly with the stationary phase and less with the polar mobile phase. For simple mixtures with only a few components separation is relatively routine but as the complexity of the sample and the variety of chemicals contained within increase many solutes will begin to co-elute with one another. Isocratic separations refer to those who have a fixed ratio of mobile phase A to mobile phase B while gradient separations alter the ratio dynamically through the separation. While both have their uses, gradient separations are much more powerful with complex mixtures as they can separate a much larger diversity of analytes in shorter amount of time. In order to determine when and what type of eluate molecules are eluting from the column a detector is required. There are a number of possibilities but the most commonly used are UV detection or MS detection. UV detection is simpler and less expensive and takes advantage of molecular absorbance at different wavelengths for the detection of analytes. The main problem with this technique is that if compounds co-elute with one another the signal from each is not discernible. However MS solves this problem by adding the detection of ions through m/z. This m/z can be correlated to analyte identity through mass which allows for coeluting compounds to be analyzed. As already briefly discussed the liquid non-charged eluant must have energy in some form applied to it to be converted to a charged gas phase molecule for analysis by MS. This is accomplished by a number of techniques such as ESI, atmospheric pressure photo ionization (APPI), or atmospheric pressure chemical ionization (APCI). ESI is by far the most used technique and is also what was used for this study. It will be discussed in further detail in Section 1.4.

In the field of metabolomics RPLC separations are the most used method of separation before MS analysis. Most metabolomic studies of the targeted nature use conventional HPLC systems with a variety of conventional columns but efficiency is always better and technological improvement always welcome. Profiling studies using HPLC are mostly interested in increased peak capacity, where peak capacity is the number of theoretical peaks that can be separated in a chromatographic run. By increasing this it is possible to more efficiently use the 'separation space' and therefore detect more metabolites. Recent advances in column and chromatography technology over the same span of growth in metabolomics have seen both grow synergistically. The use of packed columns with sub 2 µm porous particles has allowed for much faster and efficient separations.[65] There is interest in the use of monolithic columns that contain a continuous porous silica rod and allow for efficient separations reduced backpressure.[66-67] There is also interest in the reduction of particle size further to increase efficiency.[68] This as mentioned before comes with the trade-off of increased backpressure. Chromatography systems that operate these sub 2 µM particles are known as ultra-pressure liquid chromatography (UPLC) or ultra-high pressure liquid chromatography (UHPLC), with the former being the name trademarked by Waters Corp. of America. These will be discussed further in the next section.

1.2.2 Ultra Performance Liquid Chromatography (UPLC)

By decreasing the size of stationary phase particles below 2 µm it is possible to increase the speed, resolution and sensitivity of the analysis. In order to handle large increases in backpressure special instrumentation is required. Due to frictional heating found in high pressure environments the bore of the columns is kept narrow, between 1 and 2 mm. The peaks obtained on these systems can be as narrow as 1 second; this is close to efficiency normally reserved for GC systems. This improved speed causes a

problem with choosing an MS system to couple with the UPLC. Because peak width is narrow the scan speed of the MS must be fast enough to provide quantitative data (usually 10-12 points/peak) for an analysis. This excludes using this separation technique to its full potential for high resolution MS system such as FT-ICR and Orbitrap instruments. UPLC has been successfully employed for both targeted[69] and profiling metabolomics studies[70-71].

1.3 Mass Spectrometry

Mass spectrometry (MS) is a technique based on measuring the mass of an ion. This measurement is known as the mass to charge ratio (m/z) and its determination depends on the type of MS used. The initial idea of the mass spectrometry dates back to the 19th century with J. J. Thompson inventing a crude MS.[72] Development continued through the 20th century with a major development being the invention of the electrospray ionization source (ESI) this allowed the MS to be couple with LC and the use of both techniques was greatly enhanced. The general schematic of a MS consists of four main parts; sample introduction or inlet, ionization source, a mass analyzer, and a detector. Analyte molecules enter the system via the inlet and are ionized by the source. Once ionized electric or magnetic fields of the mass analyzer will guide and separate ions based on their m/z ratio. The detector is affected in some way by analyte ions to create an analog or digital signal that is then recorded. There are many variations of these four components depending on what samples, analytes, and goals of the analysis are required.

1.3.1 Electrospray Ionization (ESI)

As mentioned above, the mass analyzer works by controlling ions. In order to take molecules and make ions, an ionization source is required. Once ionized the ion is then introduced into the low vacuum environment of the mass analyzer. There are two general classes of ionization techniques, hard and soft. Hard ionization imparts larger

amounts (70 eV) on molecules causing fragmentation. Ionization modes that accomplish this are electron impact (EI) and chemical ionization (CI). Soft ionization on the other hand has minimal fragmentation and modes that accomplish this are fast atom bombardment (FAB), matrix assisted laser desorption ionization (MALDI), atmospheric pressure chemical ionization (APCI), atmospheric photo ionization (APPI), and electrospray ionization (ESI). Not all of these methods are still extensively used but each has its own characteristics that make them valuable, for example EI is one of the oldest techniques and causes extensive fragmentation of analytes that leaves a characteristic 'fingerprint' that can be used as a means of molecular identification secondary to retention time and molecular mass. The technique used for this study, ESI, is also one of the most widely used because of its ease in coupling with LC. The rest of this section will discuss this ionization in more detail.

ESI was first reported in 1914[73], but not until 1978 and two separate groups, Fenn[74] and Aleksandrov[75], *et al,* did the true power of the technique become apparent. ESI is useful because it can ionize relatively large (<20000 Da), non-volatile, and labile molecules. The one main requirement is that molecules be acidic or basic enough in the gas phase to accept or lose a proton and become ionized. Solution containing analytes, often from the LC, is introduced to a capillary needle that has a high positive or negative voltage (2-4 kV) applied to it. The liquid is charged upon exiting the tip and forms what is known as a Taylor cone. The actual method of droplet formation and ionization is still not fully understood and debate continues on the topic.[76-77] The proposed mechanism is broken into four steps that are also shown in Figure 1-4. All discussion will be about positive mode ionization. (1) A charged droplet forms at the tip of capillary, because of the high voltage applied and relatively small area of the capillary tip the electric field at this point is very high. Cations formed from the electric field migrate towards the

negative electrode while anions migrate away from the tip causing enrichment of the cations. The movement of ions is resisted by surface tension of the liquid and the Taylor cone forms. The tip of the cone becomes unstable and droplets break off into a fine jet. This jet breaks into smaller charged droplets due to repulsive surface charges. (2) Evaporation of charged droplets. Dry gas, often heated nitrogen is flowed over the source increasing the rate of solvent evaporation in the droplets. (3) Droplet shrinkage and fission. The charged droplets shrink due to solvent evaporation but the charge remains constant. This results in increased charge density at the surface of the drops. As the droplet shrinks further the repulsive forces of the charges exceed the surface tension and the Rayleigh limit is met. Further evaporation leads to Coulomb fission in which droplets break off into smaller droplets about a fifth of the size of the parent with a sixth of the charge.[76] (4) The formation of gas phase ions. This is the step in which debate is still being waged. There are two possible mechanisms: the charged residue model (CRM);[78-79] and the ion evaporation model (IEM).[80] The CRM proposes continuous evaporation and fission until only the analyte remains; this is believed to be what happens with macromolecules. The IEM differs in that ion are expelled from the charged droplet by coulombic repulsion. This applies to small molecules that are the focus of metabolomics. The response of an analyte is dependent on a variety of factors and is therefore complicated. These include the nature of the analyte and concentration of electrolytes[81], the volatility of the solvent[76], the surface activity of the droplet,[63] flow rate of the inlet,[82] other ionizable species,[63] and the competition for charge with other gas phase ions.[83]

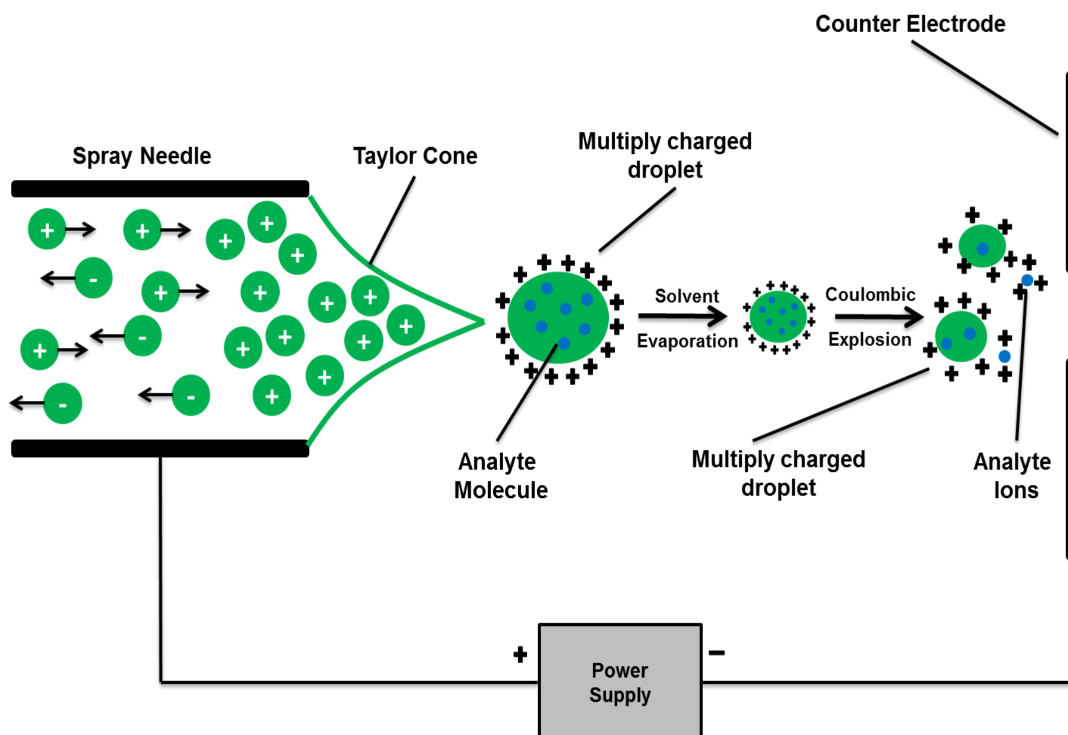


Figure 1-4 Representation of the electro spray ion process which leads to ionization of analyte molecules and transfer to a mass analyzer.

1.3.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

The mass analyzer is the most important part of a mass spectrometer. It performs the actual separation of the charged ions based on their mass. There are a large variety of analyzers available each with their own specific attributes. Magnetic sector instruments were the original workhorse analyzer used until about the 1980's. Since the invention of ESI an explosion in MS technology has been realized. Mass analyzers used in small molecule analysis are quadrupole, ion trap, linear ion trap, time-of-flight (TOF), Fourier transform ion cyclotron resonance (FT-ICR), and Orbitrap. In addition there are combinations of these analyzer types for improved functionality and increased modes of operation. The work of this thesis used an FT-ICR MS which will be discussed further.

The FT-ICR MS is a trapping mass analyzer that provides very high resolution and mass accuracy. It was initially developed by Comisarow and Marshall in 1974.[84] It operates like other trapping based MS instruments. The front end of the analyzer uses a number of hexapole and quadrupole ion guides to help move and focus ions before admission to the Penning trap. All mass analyzers require a high vacuum to minimize ion collisions with residual gas and therefore reduce the amount of available ions for detection. Upon entrance into this trap the ions are excited at their resonant cyclotron frequencies. This is achieved by a spatially uniform static magnetic induction field. This causes the ions to undergo ion cyclotron motion at the characteristic frequency of their given m/z *Eq 1.1*a and b explain the motion of an ion of mass, *m,* and charge, *q*, moving in a spatially uniform magnetic field, *B*, rotates about the magnetic field direction as shown in Figure 1-5.

$$Force = mass \times acceleration = m\frac{d\boldsymbol{v}}{dt} = q(\vec{v} \times \vec{B}) \qquad (1.1)$$

Where, m = ionic mass, q = charge, v = velocity and x B = vector cross product indicating that the Lorentz force moves perpendicular to the plane established by the ion velocity and the magnetic fields.[85-86]

With angular acceleration in the plane perpendicular to B expressed as dv/dt = v_{xy}^2/r, where r is the radius of the cyclotron orbital and angular frequency is expressed as $\omega_c = v_{xy}/r$, which allows Eq 1.1 to be simplified to:

$$\omega_c = \frac{qB}{m} \qquad (1.2)$$

This Eq 1.2 is known as the cyclotron equation and describes the motion of unperturbed ions within the magnetic field. The main point to take from this equation is that ions of m/q or m/z have the same ICR frequency independent of their velocity. This

means that m/z is inversely related to cyclotron frequency and this is the fundamental idea on which FTICR can produce its unique type of data.[87]

The FT-ICR relies on a sequence of events for ion detection that can be broken into four parts: (1) Quench. A large voltage is applied to the trapping plates (see Figure 1-6 C for representation of a trap) in order to remove all remaining ions (from the previous scan) in the trap space.(2) Ionization. An electron beam of laser is used to ionize molecules or as in the case of an LC-MS experiment they are introduced using ion optics on the front end. (3) RF Chirp. A fast RF sweep excites all ions to larger cyclotron orbits. (See figure 1-6 B). (4) Signal. Decay of ions in cyclotron motion produces an image current which is amplified, digitized and stored. Fourier transform is then performed to convert the frequency signal to that of m/z values and obtain the corresponding mass spectrum.

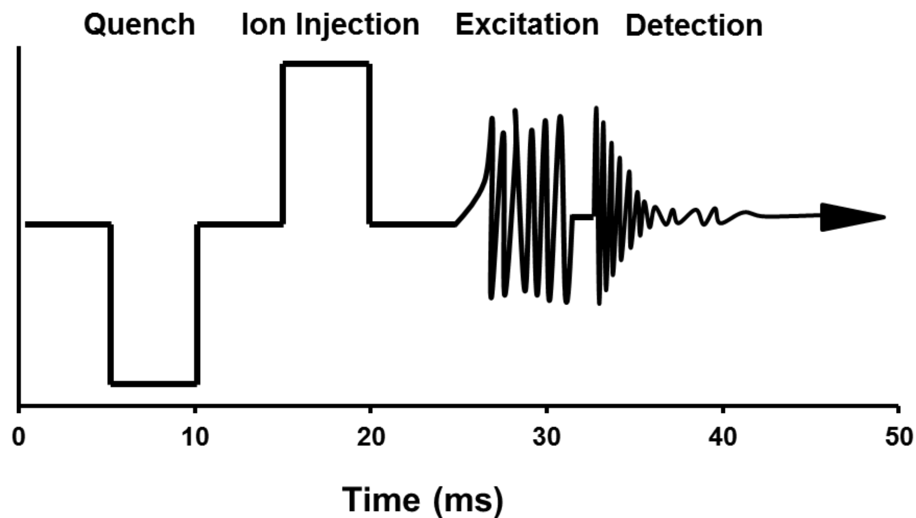


Figure 1-5 Sequence of events involved in FTICR-MS excitation and detection. The y-axis is arbitrary current.

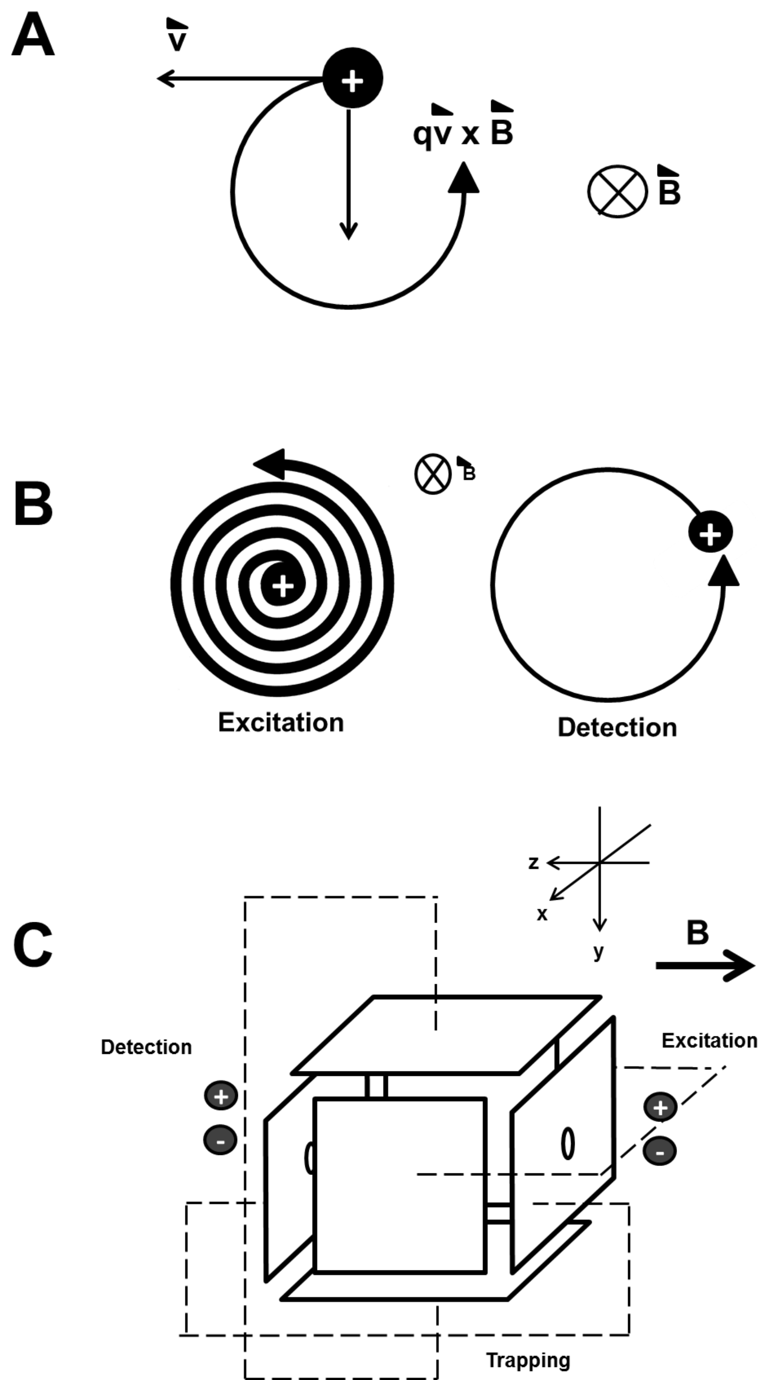


Figure 1-6 Ion cyclotron motion where an ion of mass m and charge q is constrained to a circular orbit by the magnetic field, B. This orbit has a characteristic angular frequency ωc or the cyclotron frequency (A). Ion excitation and detection (B). Depiction of ICR cell (C).

The characteristics of an FTICR-MS are that it has a wide mass range of m/z values (up to 5000). It has high resolving power with resolution (R) =1000000 possible and R = 10000 routine. The mass measurement accuracy is between 2 and 5 ppm. Ions in the cell are all detected but there may be some issues with ion transfer into the cell. The downside of this technique is the slow scan speed, shown in Figure 1-5 a single scan is 40 ms or 4 s thereby constraining the analysis when coupled with chromatographic techniques. Another issue is the prohibitive cost and space required to house this type of instrument. Despite this FTICR one of the only instruments that can attain resolution high enough for putative formula assignment through mass and spectral pattern. This has been shown to be possible but still has problems with high false positive ratios.[88-89] This problem becomes worse as the signal to noise ratio gets low (~20 or lower S/N).[90] This is due to decreased mass accuracy below S/N of 10 and because isotopic patterns at these signal levels are undetectable.[91] Even if the putative formula is determined there are often many hits. For a simple metabolite like glycine ($C_2H_5NO_2$) there are 19 matches in Chemspider and if just the mass is used in METLIN, a specific metabolite search engine, there are 5 results. Event thought this process is not perfect it can be improved through proper data filtering, controls, and strategies such as those employed in this thesis.

1.4 Overview of Thesis

The main objective of this thesis is to assess and evaluate different reaction conditions for the labeling of metabolites in biological samples using dansyl chloride stable isotopic reagent. The samples are then analyzed using LC-MS for use in metabolomic profiling and quantitative studies. To determine the effect of changes of temperature, time, and reaction method on specific metabolites a simpler analysis was developed using UPLC-UV. Chapter 2 outlines this approach and the results of quantifying labeled metabolites and comparing the amount to determine optimal reaction

conditions. Rapid labeling of metabolites using microwave assisted heating were assessed in an attempt to reduce the overall sample preparation time required for this analysis and to determine if there were other positive or negative effects. Chapter 3 is focused on determining how varying the dansyl derivatization reaction conditions affect the outcome of LC-MS profiling experiments. A number of factors were evaluated in an attempt to discern what optimal conditions for this type of experiment may be. Peak pair number was extracted and used as a comparison of the number of putative metabolites determined for each different reaction. Also the data was compared using multivariate statistical analysis for separation and classification of peak pair results. Chapter 4 outlines the conclusions of the thesis and directions for future work.

1.5 Literature Cited

1. Hammarqvist, F.; Andersson, K.; Luo, J.-L.; Wernerman, J., *Clinical Nutrition* **2005,** *24* (2), 236-243.

2. Wang, M.; Lamers, R.-J. A. N.; Korthout, H. A. A. J.; van Nesselrooij, J. H. J.; Witkamp, R. F.; van der Heijden, R.; Voshol, P. J.; Havekes, L. M.; Verpoorte, R.; van der Greef, J., *Phytotherapy Research* **2005,** *19* (3), 173-182.

3. De Vos, R. C.; Moco, S.; Lommen, A.; Keurentjes, J. J.; Bino, R. J.; Hall, R. D., *Nature protocols* **2007,** *2* (4), 778-91.

4. Wishart, D. S., *American Journal of Transplantation* **2005,** *5* (12), 2814-2820.

5. Antignac, J. P.; Courant, F.; Pinel, G.; Bichon, E.; Monteau, F.; Elliott, C.; Le Bizec, B., *TrAC Trends in Analytical Chemistry* **2011,** *30* (2), 292-301.

6. Kuehnbaum, N. L.; Britz-McKibbin, P., *Chemical Reviews* **2013,** *113* (4), 2437-2468.

7. Nicholson, J. K.; Buckingham, M. J.; Sadler, P. J., *The Biochemical journal* **1983,** *211* (3), 605-15.

8. Jayawickrama, D. A.; Sweedler, J. V., *Journal of chromatography. A* **2003,** *1000* (1-2), 819-40.

9. Cloarec, O.; Campbell, A.; Tseng, L. H.; Braumann, U.; Spraul, M.; Scarfe, G.; Weaver, R.; Nicholson, J. K., *Anal Chem* **2007,** *79* (9), 3304-11.

10. Kuhara, T.; Shinka, T.; Inoue, Y.; Ohse, M.; Zhen-Wei, X.; Yoshida, I.; Inokuchi, T.; Yamaguchi, S.; Takayanagi, M.; Matsumoto, I., *Journal of Chromatography B: Biomedical Sciences and Applications* **1999,** *731* (1), 141-147.

11. Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P., *Anal Chem* **2009,** *81* (7), 2506-15.

12. Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C., *Science* **1989,** *246* (4926), 64-71.

13. Shushan, B., *Mass spectrometry reviews* **2010,** *29* (6), 930-944.

14. Maxwell, E. J.; Chen, D. D. Y., *Analytica chimica acta* **2008,** *627* (1), 25-33.

15. Chace, D. H.; Kalas, T. A.; Naylor, E. W., *Clinical chemistry* **2003,** *49* (11), 1797-817.

16. Want, E. J.; Nordström, A.; Morita, H.; Siuzdak, G., *Journal of proteome research* **2006,** *6* (2), 459-468.

17. Rabinowitz, J. D.; Kimball, E., *Analytical Chemistry* **2007,** *79* (16), 6167-6173.

18. Lisec, J.; Schauer, N.; Kopka, J.; Willmitzer, L.; Fernie, A. R., *Nature protocols* **2006,** *1* (1), 387-96.

19. Monton, M. R.; Soga, T., *Journal of chromatography. A* **2007,** *1168* (1-2), 237-46; discussion 236.

20. Zelena, E.; Dunn, W. B.; Broadhurst, D.; Francis-McIntyre, S.; Carroll, K. M.; Begley, P.; O'Hagan, S.; Knowles, J. D.; Halsall, A.; Wilson, I. D.; Kell, D. B., *Analytical Chemistry* **2009,** *81* (4), 1357-1364.

21. Lu, W.; Bennett, B. D.; Rabinowitz, J. D., *Journal of Chromatography B* **2008,** *871* (2), 236-242.

22. Marshall, A. G., *International Journal of Mass Spectrometry* **2000,** *200* (1–3), 331-356.

23. Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R., *Journal of Mass Spectrometry* **2005,** *40* (4), 430-443.

24. Stroh, J. G.; Petucci, C. J.; Brecker, S. J.; Huang, N.; Lau, J. M., *J Am Soc Mass Spectrom* **2007,** *18* (9), 1612-6.

25. Fiehn, O., *Comparative and Functional Genomics* **2001,** *2* (3), 155-168.

26. Botros, L.; Sakkas, D.; Seli, E., *Molecular Human Reproduction* **2008,** *14* (12), 679-690.

27. Fiehn, O.; Robertson, D.; Griffin, J.; Werf, M.; Nikolau, B.; Morrison, N.; Sumner, L.; Goodacre, R.; Hardy, N.; Taylor, C.; Fostel, J.; Kristal, B.; Kaddurah-Daouk,

R.; Mendes, P.; Ommen, B.; Lindon, J.; Sansone, S.-A., *Metabolomics* **2007,** *3* (3), 175-178.

28. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., *Therapeutic drug monitoring* **2005,** *27* (6), 747-51.

29. Li, L.; Li, R.; Zhou, J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y.; Huan, T.; Zheng, J.; Shi, Y.; Wishart, D. S.; Lin, G., *Analytical Chemistry* **2013,** *85* (6), 3401-3408.

30. 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 research* **2007,** *35* (Database issue), D521-6.

31. 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 Souza, A.; Zuniga, A.; Dawe, M.; Xiong, Y.; Clive, D.; Greiner, R.; Nazyrova, A.; Shaykhutdinov, R.; Li, L.; Vogel, H. J.; Forsythe, I., *Nucleic acids research* **2009,** *37* (Database issue), D603-10.

32. Zuniga, A.; Li, L., *Analytica chimica acta* **2011,** *689* (1), 77-84.

33. Bruce, S. J.; Jonsson, P.; Antti, H.; Cloarec, O.; Trygg, J.; Marklund, S. L.; Moritz, T., *Analytical Biochemistry* **2008,** *372* (2), 237-249.

34. Want, E. J.; O'Maille, G.; Smith, C. A.; Brandon, T. R.; Uritboonthai, W.; Qin, C.; Trauger, S. A.; Siuzdak, G., *Analytical Chemistry* **2005,** *78* (3), 743-752.

35. Guo, K.; Bamforth, F.; Li, L., *J. Am. Soc. Mass Spectrom.* **2011,** *22* (2), 339-347.

36. Halket, J. M.; Waterman, D.; Przyborowska, A. M.; Patel, R. K. P.; Fraser, P. D.; Bramley, P. M., *Journal of Experimental Botany* **2005,** *56* (410), 219-243.

37. Spagou, K.; Tsoukali, H.; Raikos, N.; Gika, H.; Wilson, I. D.; Theodoridis, G., *Journal of Separation Science* **2010,** *33* (6-7), 716-727.

38. Lam, M. Y.; Siu, S. O.; Lau, E.; Mao, X.; Sun, H. Z.; Chiu, P. N.; Yeung, W. B.; Cox, D.; Chu, I., *Anal Bioanal Chem* **2010,** *398* (2), 791-804.

39. Guy, P. A.; Tavazzi, I.; Bruce, S. J.; Ramadan, Z.; Kochhar, S., *Journal of Chromatography B* **2008,** *871* (2), 253-260.

40. Nordstrom, A.; Want, E.; Northen, T.; Lehtio, J.; Siuzdak, G., *Analytical Chemistry* **2007,** *80* (2), 421-429.

41. Gika, H. G.; Theodoridis, G. A.; Earll, M.; Snyder, R. W.; Sumner, S. J.; Wilson, I. D., *Analytical Chemistry* **2010,** *82* (19), 8226-8234.

42. Bijlsma, S.; Bobeldijk, I.; Verheij, E. R.; Ramaker, R.; Kochhar, S.; Macdonald, I. A.; van Ommen, B.; Smilde, A. K., *Analytical Chemistry* **2005,** *78* (2), 567-574.

43. Ahuja, S., *Journal of Pharmaceutical Sciences* **1976,** *65* (2), 163-182.

44. Huang, X.; Regnier, F. E., *Anal Chem* **2008,** *80* (1), 107-14.

45. Asan, A.; Isildak, I., *Journal of Chromatography A* **2003,** *988* (1), 145-149.

46. Santa, T.; Al-Dirbashi, O. Y.; Fukushima, T., *Drug discoveries & therapeutics* **2007,** *1* (2), 108-18.

47. Baran, R.; Bowen, B. P.; Bouskill, N. J.; Brodie, E. L.; Yannone, S. M.; Northen, T. R., *Analytical Chemistry* **2010,** *82* (21), 9034-9042.

48. Woo, H. K.; Go, E. P.; Hoang, L.; Trauger, S. A.; Bowen, B.; Siuzdak, G.; Northen, T. R., *Rapid communications in mass spectrometry : RCM* **2009,** *23* (12), 1849-55.

49. Yang, W. C.; Adamec, J.; Regnier, F. E., *Anal Chem* **2007,** *79* (14), 5150-7.

50. Soukup-Hein, R. J.; Remsburg, J. W.; Dasgupta, P. K.; Armstrong, D. W., *Anal Chem* **2007,** *79* (19), 7346-52.

51. Hegeman, A. D.; Schulte, C. F.; Cui, Q.; Lewis, I. A.; Huttlin, E. L.; Eghbalnia, H.; Harms, A. C.; Ulrich, E. L.; Markley, J. L.; Sussman, M. R., *Anal Chem* **2007,** *79* (18), 6912-21.

52. Birkemeyer, C.; Luedemann, A.; Wagner, C.; Erban, A.; Kopka, J., *Trends Biotechnol* **2005,** *23* (1), 28-33.

53. Feldberg, L.; Venger, I.; Malitsky, S.; Rogachev, I.; Aharoni, A., *Analytical Chemistry* **2009,** *81* (22), 9257-9266.

54. Fukusaki, E.; Harada, K.; Bamba, T.; Kobayashi, A., *Journal of bioscience and bioengineering* **2005,** *99* (1), 75-7.

55. Yang, W. C.; Regnier, F. E.; Sliva, D.; Adamec, J., *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2008,** *870* (2), 233-40.

56. Guo, K.; Li, L., *Analytical Chemistry* **2009,** *81* (10), 3919-3932.

57. Guo, K.; Li, L., *Analytical Chemistry* **2010,** *82* (21), 8789-8793.

58. Stanislaus, A.; Guo, K.; Li, L., *Analytica chimica acta* **2012,** *750*, 161-72.

59. Seiler, N.; Knodgen, B., *Journal of chromatography* **1978,** *145* (1), 29-39.

60. Zotou, A.; Loukou, Z.; Soufleros, E.; Stratis, I., *Chromatographia* **2003,** *57* (7-8), 429-439.

61. Minocha, R.; Long, S., *Journal of Chromatography A* **2004,** *1035* (1), 63-73.

62. Price, N. P. J.; Firmin, J. L.; Gray, D. O., *Journal of Chromatography A* **1992,** *598* (1), 51-57.

63. Cech, N. B.; Enke, C. G., *Mass spectrometry reviews* **2001,** *20* (6), 362-387.

64. Leitner, A.; Lindner, W., *PROTEOMICS* **2006,** *6* (20), 5418-5434.

65. Fekete, S.; Olah, E.; Fekete, J., *Journal of chromatography. A* **2012,** *1228*, 57-71.

66. Tanak, N.; Kobayashi, H.; Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Hosoya, K.; Ikegami, T., *Journal of chromatography. A* **2002,** *965* (1-2), 35-49.

67. Kobayashi, H.; Ikegami, T.; Kimura, H.; Hara, T.; Tokuda, D.; Tanaka, N., *Analytical Sciences* **2006,** *22* (4), 491-501.

68. Unger, K. K.; Jilge, O.; Kinkel, J. N.; Hearn, M. T. W., *Journal of Chromatography A* **1986,** *359* (0), 61-72.

69. Li, H.; Deng, Z.; Liu, R.; Loewen, S.; Tsao, R., *Food Chemistry* **2012,** *132* (1), 508-517.

70. Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K., *Nature protocols* **2010,** *5* (6), 1005-18.

71. Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K., *Nature protocols* **2013,** *8* (1), 17-32.

72. Thompson, J. J., *Philosophical Magazine* **1897,** *44*.

73. Zeleny, J., *Physical Review* **1914,** *3* (2), 69-91.

74. Yamashita, M.; Fenn, J. B., *The Journal of Physical Chemistry* **1984,** *88* (20), 4451-4459.

75. Alexandrov, M. L.; Gall, L. N.; Krasnov, N. V.; Nikolaev, V. I.; Pavlenko, V. A.; Shkurov, V. A., *Rapid Communications in Mass Spectrometry* **2008,** *22* (3), 267-270.

76. Kebarle, P.; Peschke, M., *Analytica chimica acta* **2000,** *406* (1), 11-35.

77. Bleiner, D.; Hametner, K.; Günther, D., *Fresenius J Anal Chem* **2000,** *368* (1), 37-44.

78. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B., *The Journal of Chemical Physics* **1968,** *49* (5), 2240-2249.

79. Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M., *The Journal of Chemical Physics* **1970,** *52* (10), 4977-4986.

80. Iribarne, J. V.; Thomson, B. A., *The Journal of Chemical Physics* **1976,** *64* (6), 2287-2294.

81. Constantopoulos, T. L.; Jackson, G. S.; Enke, C. G., *J Am Soc Mass Spectrom* **1999,** *10* (7), 625-34.

82. Asperger, A.; Efer, J.; Koal, T.; Engewald, W., *Journal of chromatography. A* **2001,** *937* (1-2), 65-72.

83. Stephenson, J. L., Jr.; McLuckey, S. A., *Anal Chem* **1996,** *68* (22), 4026-32.

84. Comisarow, M. B.; Marshall, A. G., *Chemical Physics Letters* **1974,** *25* (2), 282-283.

85. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S., *Mass spectrometry reviews* **1998,** *17* (1), 1-35.

86. Mathur, R.; O'Connor, P. B., *Rapid Communications in Mass Spectrometry* **2009,** *23* (4), 523-529.

87. Marshall, A. G.; Hendrickson, C. L., *International Journal of Mass Spectrometry* **2002,** *215* (1–3), 59-75.

88. Kind, T.; Fiehn, O., *BMC bioinformatics* **2007,** *8*, 105.

89. Kilgour, D. P.; Mackay, C. L.; Langridge-Smith, P. R.; O'Connor, P. B., *Anal Chem* **2012,** *84* (17), 7431-5.

90. Weber, R. J.; Southam, A. D.; Sommer, U.; Viant, M. R., *Anal Chem* **2011,** *83* (10), 3737-43.

91. Han, J.; Danell, R. M.; Patel, J. R.; Gumerov, D. R.; Scarlett, C. O.; Speir, J. P.; Parker, C. E.; Rusyn, I.; Zeisel, S.; Borchers, C. H., *Metabolomics* **2008,** *4* (2), 128-140.

Chapter 2 : UPLC-UV Quantification for the Determination of Optimal Derivatization Reaction Conditions for Isotope Labelling Metabolomics Studies

2.1 Introduction

Metabolomics is an important and growing field that endeavours to either profile or quantify all constituent small molecules (<1500 Da) within a sample of biological origin. This field of study focuses on the end of the omics cascade, after both genomics and proteomics. The data collected using this approach is important for biomarker discovery, biomarker quantification, and linking studies with other omics technologies.[1-3] Proteins and DNA have in common the fact that they have a lot of chemical similarity from molecule to molecule; they are both composed of a set of elementary units, amino acids and nucleic acids respectively. Metabolites are much more diverse, they present a challenge to analyze due to the large variation in chemical properties and the concentrations at which they exist in the diverse set of biological samples to be studied. The classes of compounds often studied using metabolomics analytical platforms include fatty acids, amino acids, carboxylic acids, carbohydrates, vitamins, and lipids.[4] In addition to the chemical diversity of the compounds, the sheer number of different metabolites presents a daunting task to any analyst. Numbers range from 600 in *S. cerivisiae* to 200000 in some plant species.[5] Currently there is no single metabolomics platform in existence that is able to detect and identify this wide diversity, nor is any able to quantify the amounts of metabolites currently found in biological samples of interest. Early quantification work of metabolites was done using gas chromatography-mass spectrometry (GC-MS) and nuclear magnetic resonance (NMR) spectroscopy.[6, 7] Mass spectrometry is the most promising of these technologies and with the invention and concurrent development of electrospray ionization (ESI) technology allowed liquid

chromatography (LC) to be used for separation in conjunction with MS for detection and identification of metabolites.[8, 9]

This chemical diversity of metabolites is more easily analysed if three chemical properties are assessed, in the example of LC-MS this would be hydrophobicity, mass analysis, and light absorption characteristics.[10, 11] Derivatizing molecules changes their chemical properties by attachment of a chemical moiety. This affects both its retention and mass characteristics. The labeling causes an increased retention for hydrophilic metabolites when using reverse phased liquid chromatography (RPLC).[12, 13] If the derivatization compound is made with two different stable isotopes it is possible to carry out differential labeling. Stable isotope labelling allows for relative quantification of metabolites with co-eluting metabolite pairs that are differentiated by the mass difference introduced through the derivatization chemistry. Our group has developed and applied this analysis strategy using dansyl chloride (DnsCl) for amine and phenol containing metabolites and *p*-dimethylaminophenacyl (DmPA) bromide for carboxylic acid containing compounds.[14-17]

One of the critical steps in isotopic labelling analysis is the derivatization reaction. Current protocols use ovens that conventionally heat reaction mixtures for fixed periods of time, often an hour or longer. Microwave-assisted heating (MAH) is the heating of a reaction through the use of microwave dielectric heating. This heating depends on the nature of the solvent and constituents of the reaction, specifically their ability to couple with the microwave field produced. For a mixture this is defined by the solvent or compounds dielectric properties.[18] Specifically, the dipoles of interacting molecule try to align with the field as it oscillates; this molecular motion then generates heat through friction and dielectric loss. The amount of heat generated here is then specifically resultant from the character of the matrix and how it interacts with the

microwave field.[19] The result of this is a rapid internal heating of the reaction media. When compared with conventional oven or oil bath heating, a MAH reaction mixture always has a much higher internal temperature profile.[20] MAH has been employed in both organic chemistry and the derivatization of compounds for LC-MS and GC-MS analysis. [21-23] The results of other studies demonstrate that using MAH can greatly reduce the amount of time required for derivatization. Reductions in time vary from 1-2 hours with no loss of accuracy or precision.[24, 25] In addition to reduced reaction times, MAH is thought to be safe for chemical reactions as it operates at a low enough frequency and therefore with low enough energy that photons produced (0.0016 eV) are too low in energy to break chemical bonds and therefore cannot cause chemical reactions.[26-28] This is important for our chemical labeling approach, as minimization of chemical byproducts is very important. Historically there has been a debate about the effect of microwave heating on chemical reactions and whether the increase in reaction rate is due to thermal effects, non-thermal effects, or a combination of the two.[29-31] Recently however the general consensus within the field of microwave organic synthesis is that increased rates and yields arise only from the improved thermal heating provided by a microwave reactor.[32-33]

The study presented in this chapter explores and compares the effects of varying reaction time, heating method, temperature, and microwave power for DnsCl labelling of metabolites for analysis by ultra-performance liquid chromatography with ultraviolet detection (UPLC-UV). Although detection by MS is the final goal of using this labelling chemistry, UV detection is useful to quickly and easily detect and quantify the products of variable reaction conditions in order to decide which are optimal.

2.2 Experimental

2.2.1 Chemicals and Reagents

All chemicals and reagents were purchased from Sigma-Aldrich Canada (Markham, ON, Canada) except those otherwise noted. LC-MS grade water and acetonitrile (ACN) were purchased from Thermo Fisher Scientific (Edmonton, AB, Canada).

2.2.2 Sample Preparation

An amino acid standard mixture was prepared by weighing solid standards and diluting in water (arginine, asparagine, aspartate, glutamine, histidine, leucine, methionine, phenylalanine, proline, tryptophan, threonine, and valine) or hydrochloric acid (1 M) (cystine and tyrosine) to 20 mM. These were then mixed to make a standard solution with 2 mM for each metabolite and 28 mM concentration in total. Following dansylation reactions 10 µL of 10 mM Dns-glycine was added as an internal standard to 90 µL of the reaction mixture that had been centrifuged at 15000 rpm for 5 minutes before injecting into the UPLC.

2.2.3 Derivatization Reactions

2.2.3.1 Conventional Labeling Reaction

The dansylation procedure follows that previously reported by Guo *et al* with some minor changes.[15] Briefly the procedure is as follows, for metabolite labelling, 50 µL of amino acid standard was mixed with 25 µL $Na_2CO_3/NaHCO_3$ (pH 9.4) buffer and reacted with 75 µL of ^{12}C-dansyl chloride (DnsCl) in ACN (12 mg/mL or 44.5 mM). The reaction mixture was then heated in a conventional oven, microwave oven or left at room temperature (~22°C). The conventional oven reactions were carried out at ambient temperature (AT) (varying between 21 and 24°C), 40°C, 60°C, or 80°C. 10 µL of sodium

hydroxide (NaOH at 250 mM) was added to each reaction in order to consume the excess DnsCl at the following time intervals of the reaction; 5, 10, 20, 30, 60, and 120 minutes. After quenching with NaOH samples were mixed and left for 10 minutes at room temperature after which 50 μL of 425 mM formic acid (FA) in H_2O: ACN (1:1) was added to acidify the samples.

2.2.3.2 Microwave Labeling Reaction

Microwave samples were prepared in the same way as described in the conventional labelling method. Microwave-assisted heating (MAH) was applied using either a CEM Discover microwave reactor (CEM, Matthews, NC, USA) or a Panasonic NN-ST661W consumer microwave (Panasonic, Mississauga, On, Canada). In order to assess the optimal MAH conditions two different types of reactions were carried out. The first is known as power control (PC), during which the samples were subjected to a constant microwave power measured in wattage (W) over specified time periods or time temperature control (TTC) during which power is applied only to maintain a temperature specified and measured by an infrared thermometer. The PC experiments were carried out at 50 , 100 , 200 , and 300 W in the CEM microwave and 120, 480 , 960 W, and 1200 W in the Panasonic microwave for 30, 60, 120, and 300 s. This resulted in 16 experiments for each of the microwave types. TTC samples were reacted at 40°C and 60°C for 5, 30 , and 60 minutes.

2.2.4 UPLC-UV Separation and Detection

Chromatography analysis was performed using a Waters Acquity UPLC with Photo Diode Array (PDA) UV detector system (Waters Canada, Mississauga, ON, Canada). The column used was a Phenomenex (Torrance, CA ,USA) Aeris 3.6 μM XB-C18 core shell particles 2.1 ID x 150 mm. Elution conditions were as follows: linear gradient from 10-43% B mobile phase over 12 minutes, 43-100% B mobile phase over 4

minutes, an isocratic hold at 100% for 1 minute, and a return to initial conditions for 3.5 minutes. Mobile phase A consists of 5% ACN in H_2O with 0.1% Formic acid (FA) and mobile phase B consists of 5% H_2O in 0.1% FA. Column temperature was maintained at 30°C and the sample manager temperature was maintained at 4°C. The flow rate was set at 650 µL/min and 5 µL of sample was injected for all analysis. Labelled metabolites were detected by UV absorbance using both 3D detection from 210-400 nm and 2D detection at 254 nm at a sampling rate of 20 points/s. The data was acquired and processed using Empower® Version 2 (Waters Canada, Mississauga, ON, Canada).

2.2.5 Method Validation

The analytical method was validated to prove acceptable values would be obtained by determining the following: selectivity and carryover, linearity, external calibration, inter-day precision, intra-day precision and carryover. Statistical analysis was performed using Excel 2010 and Origin 7.

2.2.5.1 Linearity and Reaction Reproducibility

Linearity of and reaction reproducibility were determined by comparing reactions with varying initial concentrations of metabolites (0.01, 0.1, 1, and 2 mM). The relative response ratio to the internal standard was then plotted for each of the analytes. Reproducibility was determined by comparing three replicate samples and quantifying the amounts of metabolite present.

2.2.5.2 External Calibration and Percent Yield Determination

External calibration was performed by producing calibration curves for prelabeled standards of Dns-Aspartic Acid and Dns-Threonine at 6 concentrations (0.01 mM to 2 mM). Percent yield was determined by comparing the externally determined absolute concentration with that of the theoretical 100% value of 2 mM.

2.2.5.3 Selectivity and Carryover

Selectivity and carryover were determined by doing the following experiments: instrument blank (ACN injection, 5 μL), solvent blank (50:50 H_2O: ACN w/ 0.1 % FA), reaction blank (reaction mixture with H_2O in place of amino acid mixture), and an amino acid blank (underivatized amino acid in H_2O: ACN w/ 0.1% FA). All were analyzed on the method described above with 3 replicated for each analysis. Carryover was determined by analysis of 3 replicates of a 28 mM total metabolite concentration directly followed by an instrument blank.

2.2.5.4 Intraday Precision, Interday Precision, and Method Reproducibility

Intraday precision was determined by analyzing the same quality control (QC) sample twice in one day with three replicates during each analysis. The column was washed and re-equilibrated during the analysis of these samples. Inter-day precision was evaluated by comparing the same QC sample which was analyzed before and after analysis on each separate analysis day. There were a total of 8 analysis days with 2 samples for each day for a total of 16 samples compared. Method reproducibility was determined by running 5 replicated QC samples sequentially. To determine the differences in the samples the relative peak areas were compared to determine the %RSD.

2.2.5.5 LOD and LLOQ

The limit of detection (LOD) and lower limit of quantitation (LLOQ) were determined for the three Dns-amino acids for which external calibration curves were produced. The equations $3 \times \frac{(\sigma_B)}{m}$ and $10 \times \frac{(\sigma_B)}{m}$ respectively were used to calculate LOD and LLOQ. σ_m is the standard deviation of the blank response and m is the slope of the calibration curve. Standard deviation of the blank response was determined from calibration curves produced in Excel.[34]

2.3 Results and Discussion

2.3.1 LC Gradient Method Development

The compounds assessed in this study are listed in Table 2-1. The purpose of this work was to assess a mixture of compounds chosen to represent some common metabolites, which in this case were amino acids. To determine the concentration of these compounds and therefore the amount produced under differing reaction conditions, it was important to develop a UPLC-UV quantification method. A systematic approach to method development was used to decrease the amount of time and number of method development samples required. This approach began with defining the separation goals and the nature of the sample. The separation goals of this work were to have baseline separation of each of the Dansylated compounds of interest in order to determine their amount using a UV PDA detector. All of the compounds of interest have a dansyl group attached, so their retention characteristics are very similar due to the strong hydrophobic character of the aromatic moiety. This means that these metabolites will elute in a higher % of mobile phase B. Because of past knowledge using dansylated metabolites it was decided that a gradient elution method would be used.

Table 2-1 List of compounds in standard mixture used for method development**.**

Dansyl Amino Acid	3 Letter Abbreviation	1 Letter Abbreviation	Retention Time (min)
Arginine	Dns-Arg	R	2.422
Asparagine	Dns-Asp	N	2.891
Glutamine	Dns-Gln	Q	3.315
Serine	Dns-Ser	S	3.764
Hydroxy Proline	Dns-T4H	T4H	4.218
Aspartate	Dns-Asp	D	4.270
Glutamate	Dns-Glu	E	4.271

Threonine	Dns-Thr	T	4.745
Glycine	Dns-Gly	G	5.034
Alanine	Dns-Ala	A	5.943
Proline	Dns-Pro	P	8.428
Valine	Dns-Val	V	8.858
Methionine	Dns-Met	M	9.130
Tryptophan	Dns-Trp	T	9.862
Lysine	Dns-Lys1	L	10.534
Phenylalanine	Dns-Phe	F	10.685
Isoleucine	Dns-Ile	I	10.804
Leucine	Dns-Leu	L	11.125
Cystine	Dns-Cystine	Cys	12.910
Cysteine	Dns-Cys	C	12.920
Lysine	Dns-Lys2	K	13.749
Histidine	Dns-His	H	13.962
Tyrosine	Dns-Tyr	T	14.876

[a] Lysine is labelled with both one and two dansyl molecules causing two peaks to appear in the separation.

2.3.1.2 Detection Wavelength Selection

Selection of a detection wavelength in chromatography is important for a number of reasons. First it is important to detect compounds of interest with minimal interference. Secondly we want to detect the smallest concentration of our analyte possible, and lastly we would like a linear detector response over the concentration of the analysis. UV PDA detection satisfies these conditions and therefore it was chosen due to its ease of use. MS has a lower limit of detection but is much more expensive and time consuming to perform, and in this analysis sensitivity was not the main concern. Although some amino acids (Trp, Tyr, and Phe) absorb strongly in the UV range due to the presence of aromatic

groups, other amino acids do not. When labelled with DnsCl however, the presence of the aromatic group on each of the analyte molecules causes strong absorbance at ~ 220 nm, 254 nm, and 338 nm. This is characteristic of benzene rings. These however are often red shifted due to the presence of an auxochrome near benzene rings. This happens through electron donating groups that effect the $\pi \rightarrow \pi^*$ transitions and extended conjugation which can effect the $n \rightarrow \pi^*$. Consideration of UV cut-off is also important when choosing an analysis wavelength. Water absorbs at 200 nm and acetonitrile at 190 nm so detection at lower wavelengths can cause interference and increased noise. Four example spectra of dansylated metabolites are given in Figure 2-1. It can be seen that each sample has very high absorbance at 220 nm, a second absorbance peak between 250-290 nm and a third absorbance maximum at greater than 300 nm. 254 nm was chosen as the analysis wavelength due to the distance from solvent cut-offs at short wavelengths, a relatively flat and reproducible profile, and because in most cases there is a higher absorbance at 254 nm than the 330 nm wavelength. From the three dimensional PDA spectrum in figure 2-1 B it is observed that a number of the dansyl compounds have a wide absorbance spectrum and that at 254 nm there is generally high absorbance out of the range of the solvent cut off for acetonitrile. Consideration of all of these factors led to the decision to use 254 nm as the analytical wavelength for this work.

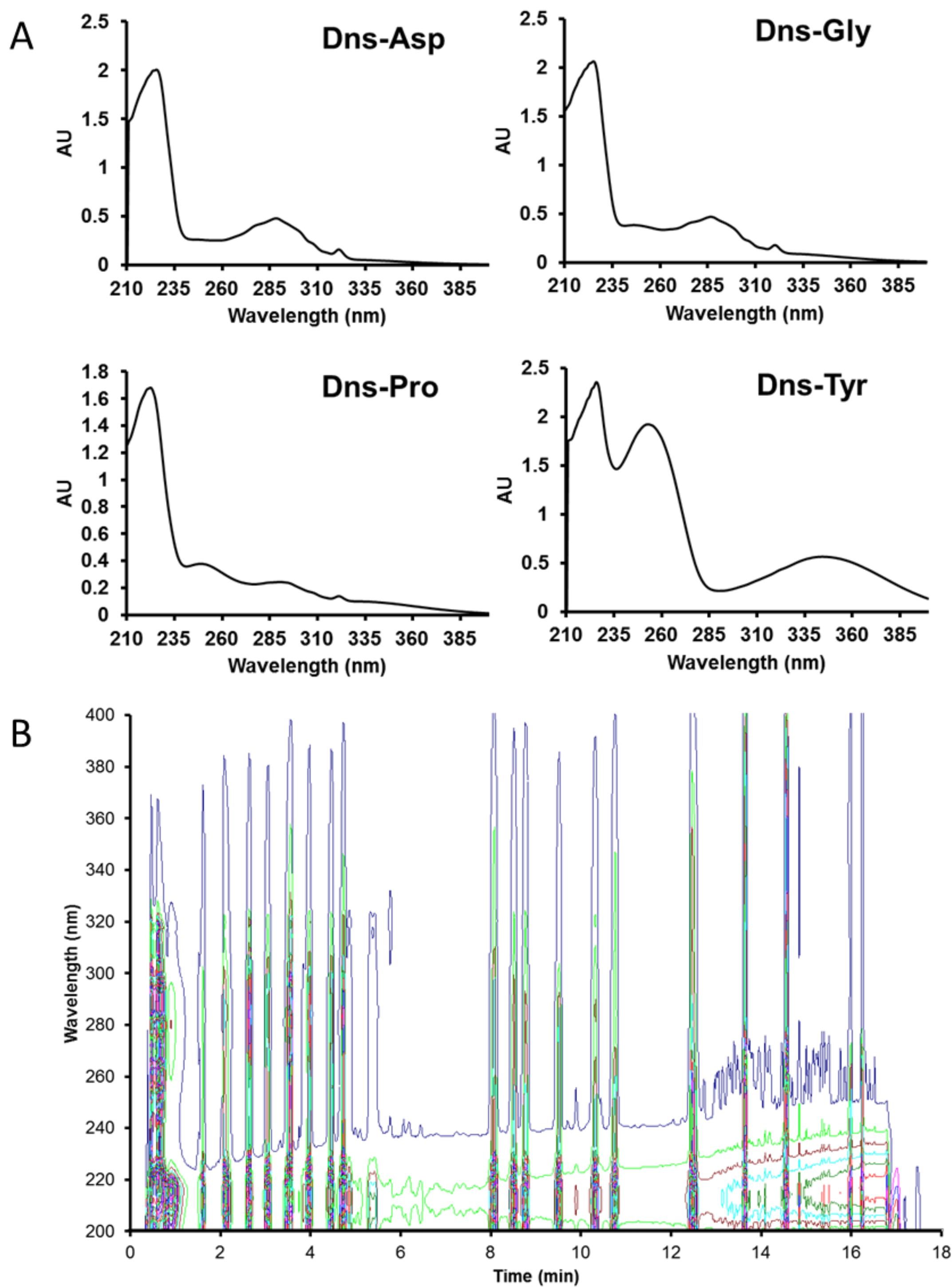


Figure 2-1 Representative UV spectra of four selected Dansylated metabolites obtained with a UV-Vis PDA detector (A). PDA spectrum of dansylated amino acid standard mixture (B).

2.3.1.3 Standard Mixture Design

To carry out a study of reaction efficiency comparing different methods it was important to design a quantitative method. A mixture of amino acids were chosen as they have been shown to react well with dansyl chloride and they are often used as models in labelling studies involving DnsCl. Due to the large number of required samples is it was important to keep the analysis time as short as possible. Initially using a purchased liquid standard mixture was explored but it was found to be unacceptable because of the presence of a number of co-eluting compounds. Figure 2-1 A shows this with the regions of a number of co-eluting compounds highlighted. Also it is interesting to note that compounds such as Dns-Asp, Dns-Gln, and Dns-Ser are not detected in this standard. Due to these issues it was decided to create a mixture for this experiment using individual amino acids. This would allow better control over concentrations and also for easier identification by matching individual retention times of standard with unknown peaks in the mixture. Figure 2-1 B shows the initial separation attempt of a mixture of all of the compounds listed in Table 2-1. All of the compounds were detected. Tyrosine was not added to the standard but it is expected to be the last eluting compound in the mixture as seen in Figure 2-1 A. This mixture of compounds had a number of components that elute too closely together and the only possibility for separation was a large increase in sample analysis time, which as mentioned earlier was important to avoid in this analysis due to the large amount of samples required.

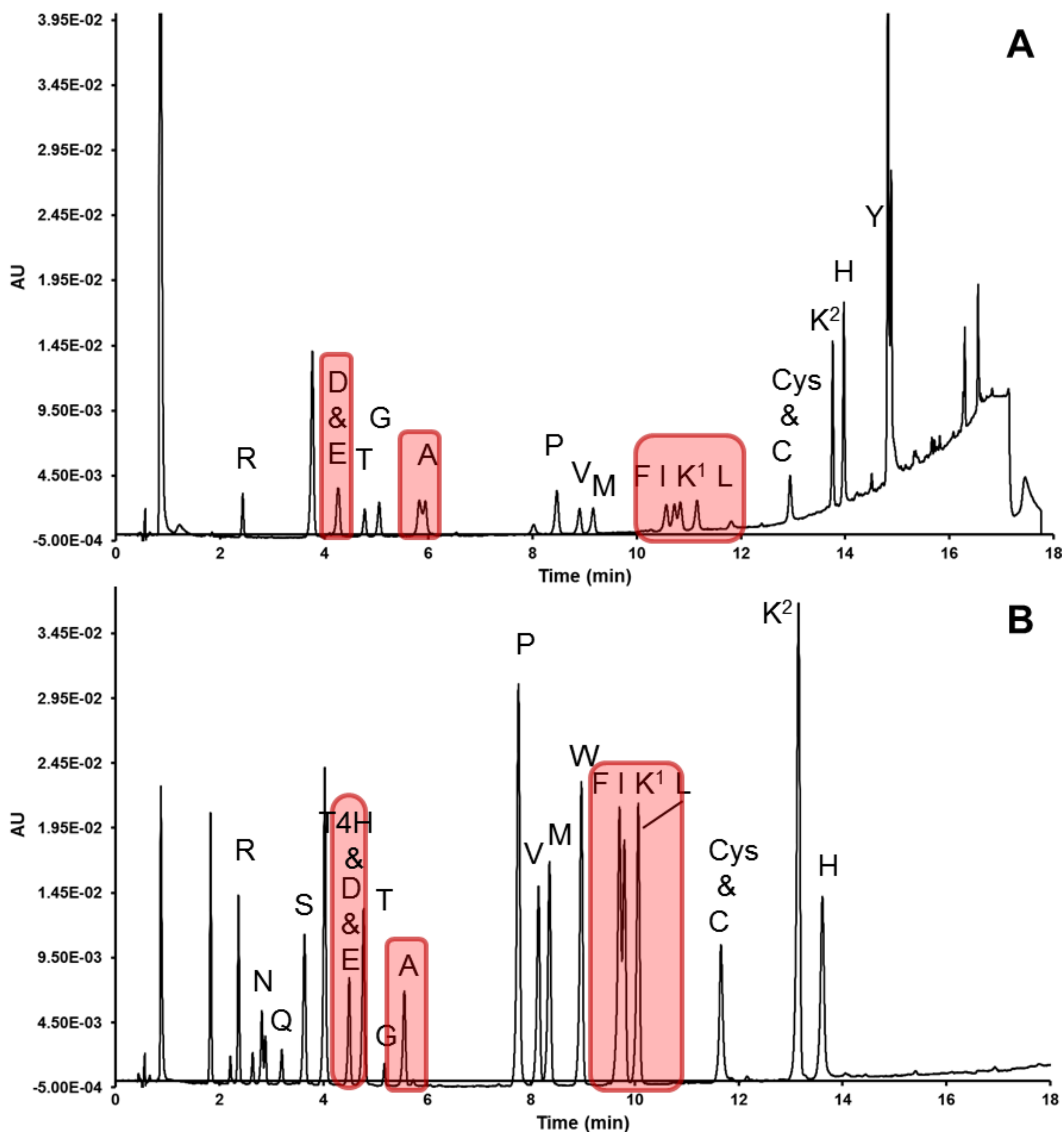


Figure 2-2 Representative UPLC-UV Chromatograms of the purchased Sigma® standard (1.25 mM each) (A) and the standard created from individual compounds (3 mM each) (B). Dansylated compounds are identified by their one letter amino acid code and a list with corresponding retention times can be found in table 2-1. Highlighted boxes in both A and B show groups of co-eluting compounds.

2.3.1.4 UPLC Gradient Development

To develop a routine assay procedure for this mixture a systematic approach to method development was applied. This method is applicable to any small molecule sample mixture and would be useful in future development of separations for dansyl metabolite analysis by UPLC-UV or MS. As dansyl labeled metabolites are all hydrophobic in nature it was decided that RPLC would be used. To assess a sample compistion an initial short "fact finding" experiment was performed. The results of this can be seen in Figure 2-3 A. Each of the chromatograms in this figure are defined by a changing gradient retention factor (k*). The results of this short gradient defined how we would proceed with changing separation conditions for the following experiments. As can be seen in Figure 2-3 A there is separation of the analytes but there are many coeluting peaks, notably early in the separation from retention times of 1 to 3 minutes and at 4 and 5 minutes. A second question that was answered using this intial short gradient was: will a gradient or isocratic separation be used for this analysis? This was assessed by judging the elution time of the first and last analytes of interest. In this run the first and last peaks elute at 1.5 and 6 minutes, the difference in retention time (Δt_r) is then 4.5 minutes. Also important is the average retention time $(\Delta t_r)_{avg}$ (1.5+6)/2 = 3.75 minutes. Samples that have low values of Δt_r can be seperated isocratically. To numerically assess the samples the following approximation is used: if $\Delta t_r/t_g < 0.25$ (t_g is gradient time) then isocratic separation may be used, if $\Delta t_r/t_g > 0.45$ then gradient separation should be used. If the value falls in between, then either isocratic or gradient may be used. In this initial experiment, the value of $\Delta t_r/t_g = 0.56$ was obtained, so a gradient elution would be the most efficient method of separation. This decision is important as often an isocratic separation can be advantageous and simpler to perform with better reproducibility.

The next step is to adjust k* by altering the gradient time. The goal of these adjustments is to improve the resolution (R_s) to at least 1.5 and preferably >2. The results of two subsequent experiments can be seen in Figure 2-3 (A) and (B). The critical pair in a separation is a group of analytes that have the lowest R_s. In the case of the k*=5 separation the critical pair can be seen at 4 minutes. In this case the resolution is not acceptable. In the k*=15 separation the R_s = 0.71 and in the k* = 20 the R_s = 1.31. While this is less than the R_s = 1.5 required for baseline separation, it is the best given the time constraints already mentioned. In each of the separations in Figure 2-3 there are a number of unseparated peaks. These are either dansyl reaction by-products or perhaps impurities from the reagents used to make the standards.

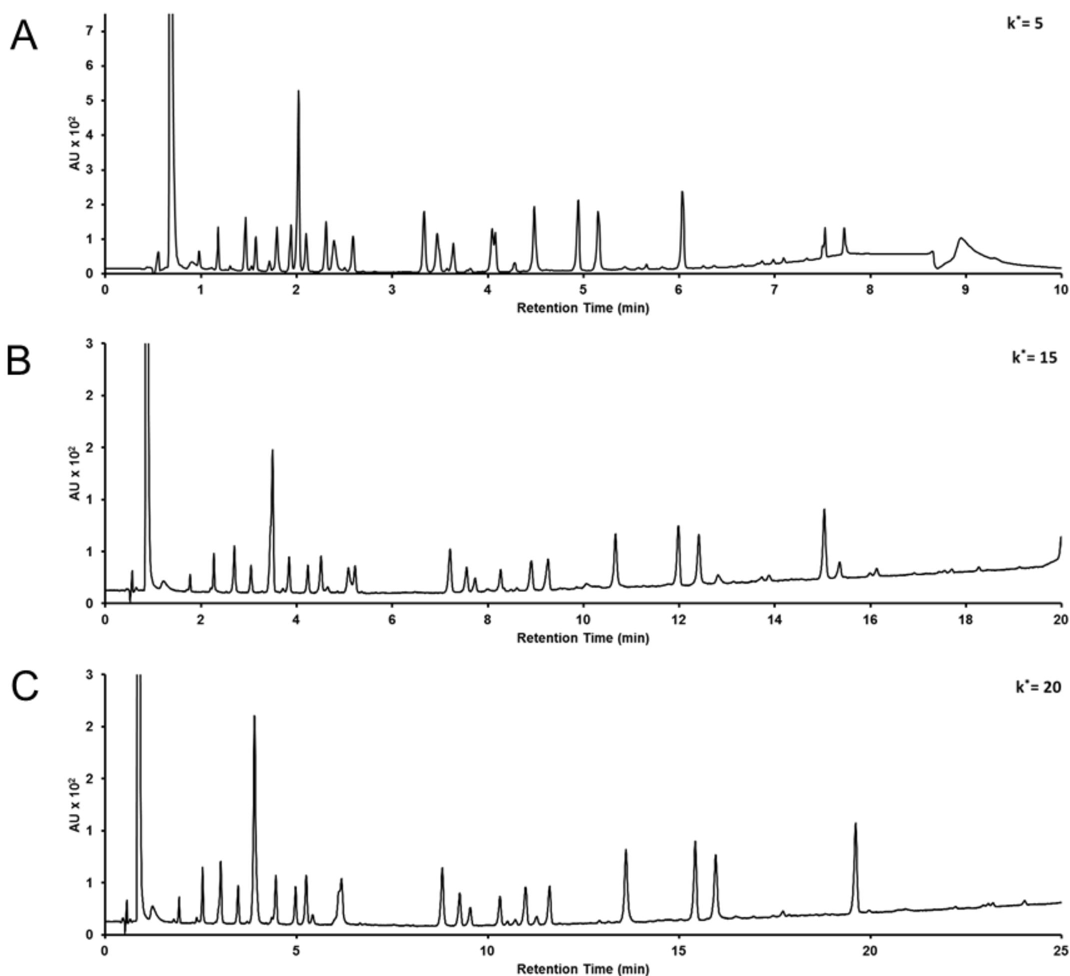


Figure 2-3 Gradient method development UPLC separation chromatograms with varying gradient times. k*=5 with an 8 minute gradient (A). k*=15 with a 19 minute gradient (B). k*=20 with a 26.5 minute gradient (C). Conditions: All methods had 10% B initial and 80% B final with a linear gradient, H_2O:ACN w/ FA; flow rate, 0.65 mL/min; column temperature = 30°C.

The next step in method development is to optimize the gradient selectivity (α*) of the separation. There are a number of chromatographic conditions that can be altered

to change the α* of a gradient separation. They are summarized in the following table along with comments regarding their effectiveness in this study:

Table 2-2 Experimental Conditions that affect α* in Reverse Phase Chromatography

Chromatographic Condition	Comment
Gradient time (t_g)	Changing up to a maximum of 10 fold. In this case 4 fold (k* = 5 - 20)
Temperature (T)	50°C attempted, decrease in resolution.
Solvent Type	Methanol attempted, decrease in resolution.
Column Type	5 cm column tried decrease in resolution.
Mobile Phase pH	Kept acidic with FA. Want to mimic MS conditions.
Mobile Phase buffer	No ionized compounds in sample.
Ion-pair reagent	No ionized compounds in sample.

Changing these conditions can result in a change of selectivity for the sample which means that some peaks will change retention times while others may not; it depends on the nature of each analyte. The first parameter changed after t_g was temperature (T). It is well characterized that changes in temperature can greatly affect the retention and selectivity of compounds in gradient elution.[35, 36] In this study, temperature was increased from 30°C to 50°C and the results can be seen in Figure 2-4. The higher temperature separation gives a small retention time reduction for all peaks as expected due to a reduction of the k* value. The most noticeable change is the reduced resolution between the peak pairs marked in Figure 2-4 (B).

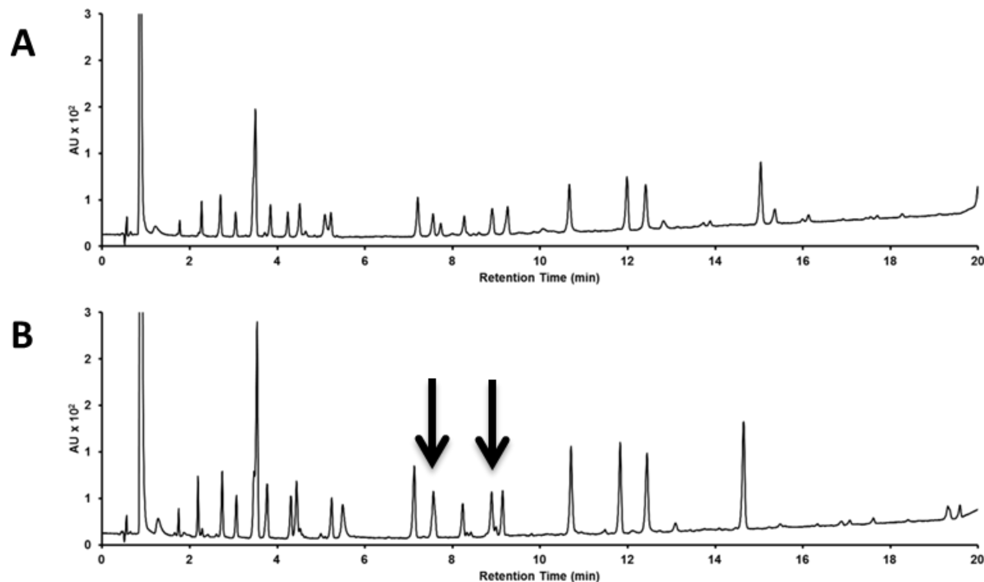


Figure 2-4 UPLC-UV chromatograms comparing separation temperatures: 30°C (A) and 50°C (B). The arrows denote where a reduction of resolution between peaks occurred at the higher temperature.

2.3.1.5 Column Selection

A shorter column was initially explored in order to determine if it could give adequate separation for this analysis. It is expected that with a shorter column there will be a reduction in the resolution between critical peaks. The plate number (N) of an analysis can be affected by the following factors: column length, flow rate, and particle size. In this study the option of changing the length and particle size of the column was limited due to a small amount of options available. It is important to note that changes in N for an analysis often do not have as large an effect on resolution as do other factors such as gradient time and temperature which have already been discussed. With comparable k* values the longer column provides improved resolution. In Figure 2-5 a comparison

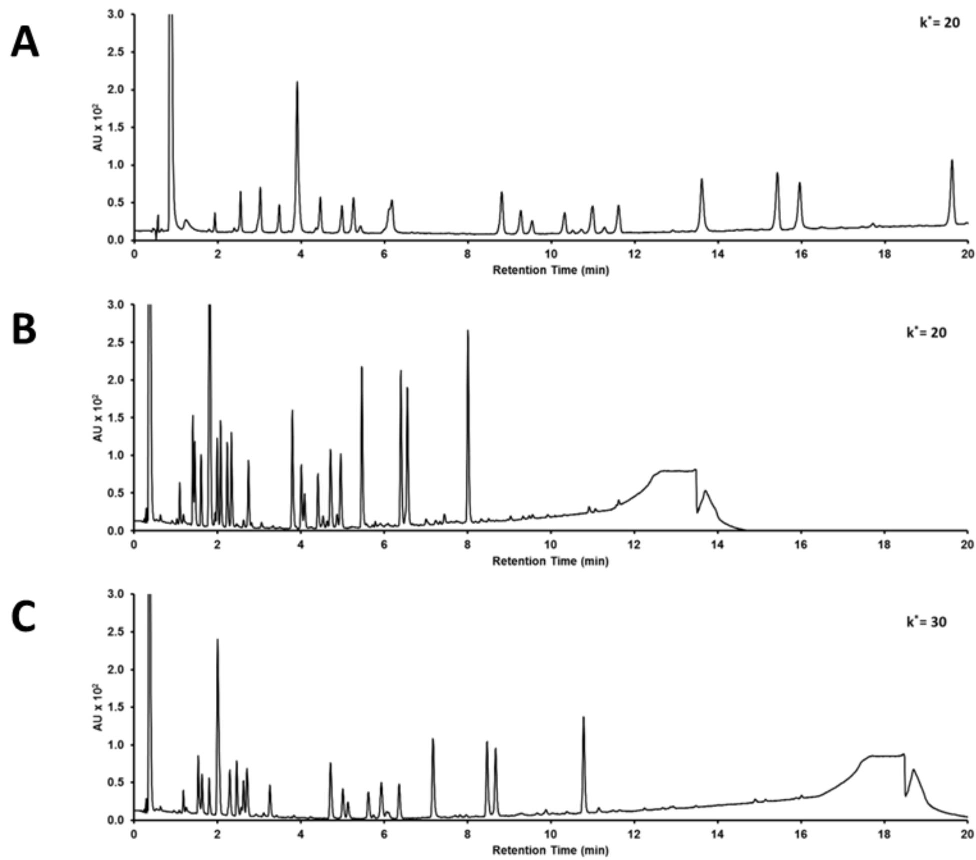


Figure 2-5 UPLC-UV chromatograms comparing separation of standards on a 15 cm column (A), and a 5 cm column (B & C). k* values are also given for comparison. All methods had a linear gradient, H_2O:ACN w/ FA; flow rate, 0.65 mL/min; column temperature = 30°C. Column = Aeris 3.6 µM XB-C18 core shell particles 2.1 ID x 150 mm

of these two different column lengths is shown. When comparing the initial portion of the separation, the 15 cm column performs much better; the 5 cm column has man co-eluting peaks (Figure 2-5 B). Even after increasing the k* value to 30 the 5 cm column cannot resolve the critical pair at 5 minutes (Figure 2-5 C). Even though the run time could still be increased to improve resolution, it was decided that the difference in the two columns was not enough to warrant further study and the 15 cm column was chosen for this analysis.

2.3.1.6 Mobile Phase Selection

If samples have peaks that are unresolvable with changes in %B a change in the type of strong solvent can offer a difference in selectivity and therefore improve the separation and resolution. The selectivity changes that arise are due to changes in interactions between the solute, B phase, and the column.[37, 38] For a change of solvent to be explored the B solvent strength (according to polarity) should remain close to that of the solvent already used. In the case of this experiment methanol was tried in place of acetonitrile. The %B range for the separation was adjusted in order to be as close to equivalent to the acetonitrile as possible using a solvent nomograph.[39] Figure 2-6 illustrates the separation of the sample mixture using acetonitrile (A) and methanol (B). The methanol separation gave reasonable separation but the width of the chromatographic peaks seemed to increase. Even after repeated equilibration and replicate experiments this was observed. Currently there is no explanation for this and the use of methanol was abandoned. Also the baseline signal at 254 nm was much larger due to methanol having a cut-off wavelength of 205 nm (compared to acetonitrile at 195). For these reasons it was decided to used acetonitrile and not methanol as the B phase solvent.

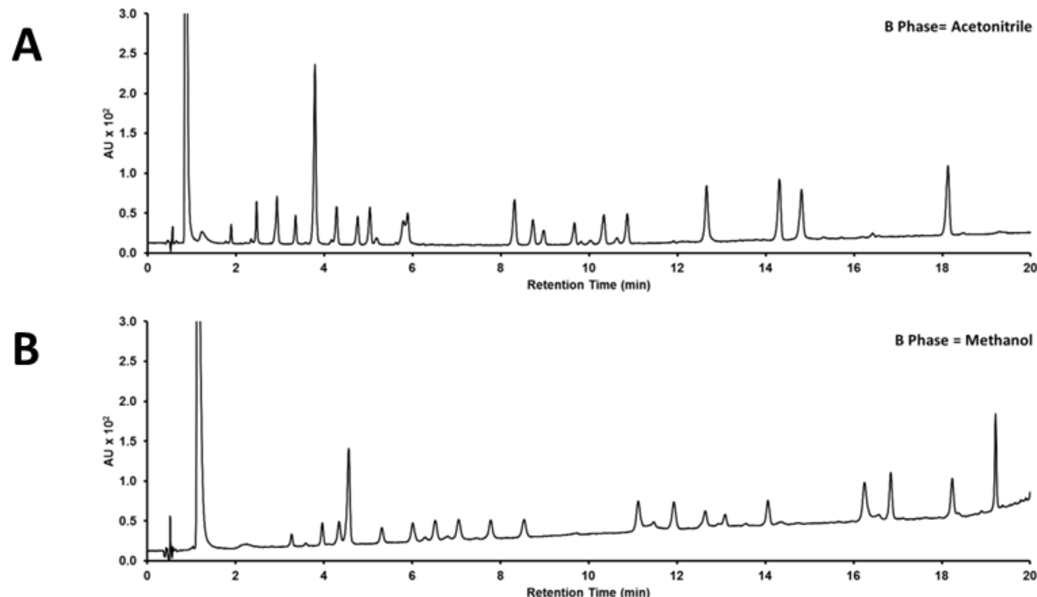


Figure 2-6 UPLC-UV chromatograms comparing acetonitrile (A) and methanol (B) used as different B phase solvent types.

2.3.1.7 Gradient Range and Shape Adjustment

To shorten the overall analysis time a segmented gradient is often used. This method takes advantage of excess resolution at the end of the sample run to decrease run time, use a high %B to remove hydrophobic impurities from the column, and/or to improve the separation. The method was developed with a linear gradient and the later eluting peaks are well resolved enough that it was possible to increase the %B change or the slope of the gradient. This decreased retention time and therefore analysis time. The final gradient method can be seen in Figure 2-7. As this analysis will be done with standard compounds the high % B isocratic hold at the end of the gradient meant to wash the column is kept short. Column equilibration time is approximately 3 minutes, at a flow rate of 0.650 mL/min. The estimated column volume is 0.234 mL so after three minutes

the column has been flushed by 1.95 mL which with a noncomplex biological mixture is considered a long enough equilibration time.

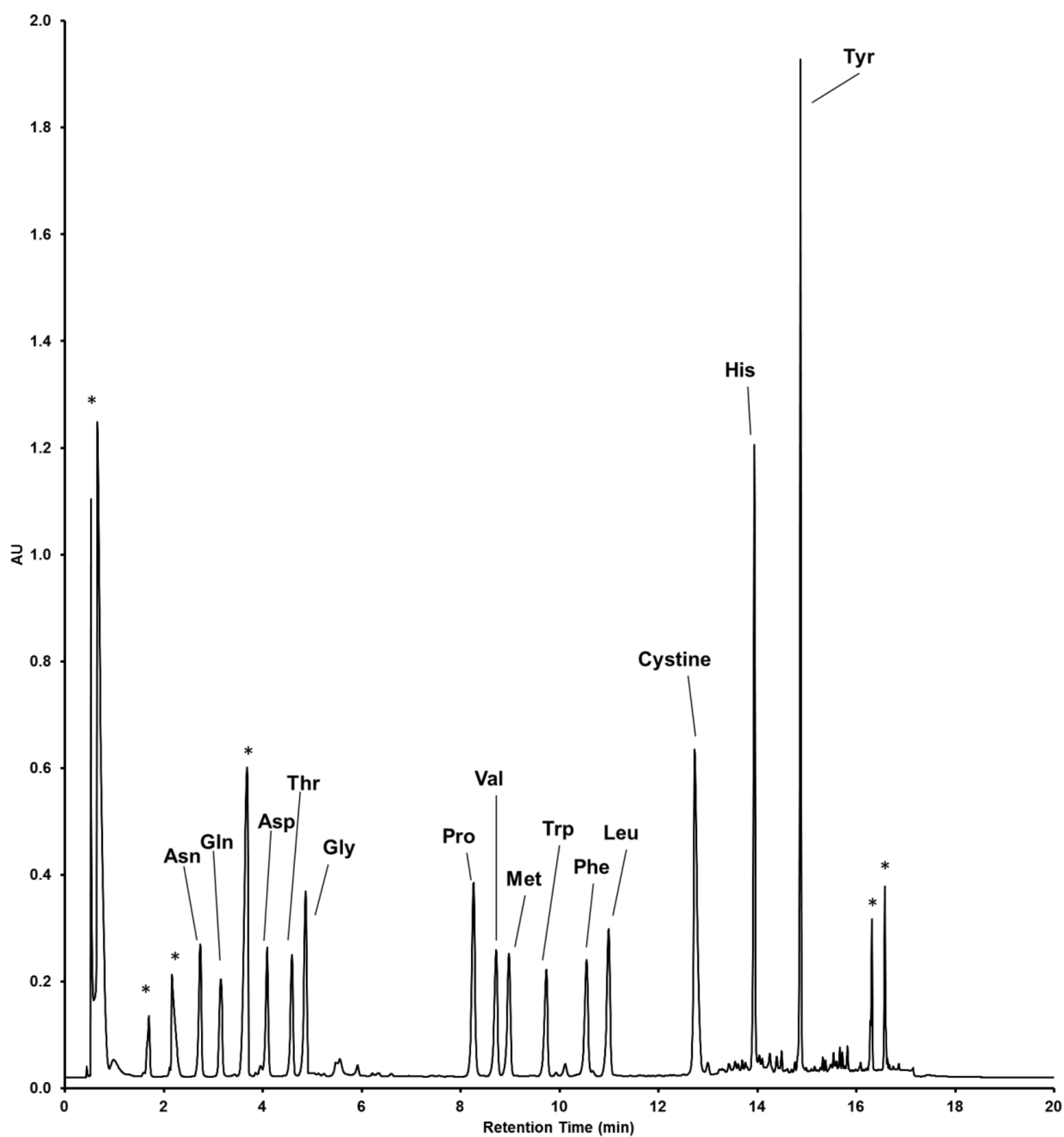


Figure 2-7 UPLC-UV chromatograms of labeled metabolites and the final UPLC method showing segmented chromatogram. Dansylation by-product peaks are denoted by *.

The method developed is useful to separate a group of compounds of interest for their quantification. This quantification will allow for a quick and easy comparison of a large number of different reaction conditions used for metabolite labelling with DnsCl.

2.3.2 Quantification of Labeled Metabolites

This developed method was applied to the analysis of a mixture of amino acids that were used to mimic a more complex mixture of metabolites that is often found in biological samples for metabolomics studies. A representative chromatogram for a separation of the final mixture is shown in Figure 2-8. In order to quantify the metabolites each of the peak areas were compared to that of the internal standard, glycine. Dns-glycine standard was added after reaction and mixed for a final concentration of 0.5 mM. This comparison will give the relative peak area described in most of the following figures. For identification of each of these peaks individual standard were analyzed using the developed method in order to assess the retention time which was then matched to that in the analysis. Often the absolute concentration of amine or phenol containing compounds in biological samples is unknown before the labeling reaction is carried out. For this reason this study was carried out using a very low ratio of DnsCl to that of the metabolite amount (amino acid standard). The molar ratio of DnsCl to amino acid standard used for all of the experiments unless otherwise mentioned is 2.4:1. It is important to note that in Figure 2-8 there are number of peaks labeled with an * that refer to suspected dansyl byproduct peaks. These are found in all dansyl reactions and artifacts and will be discussed in further detail later.

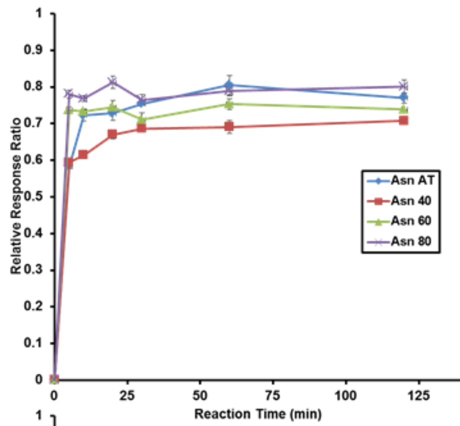
Table 2-3 Mixture of dansylated compounds for comparison of reaction conditions and retention times from the UPLC-UV analysis.

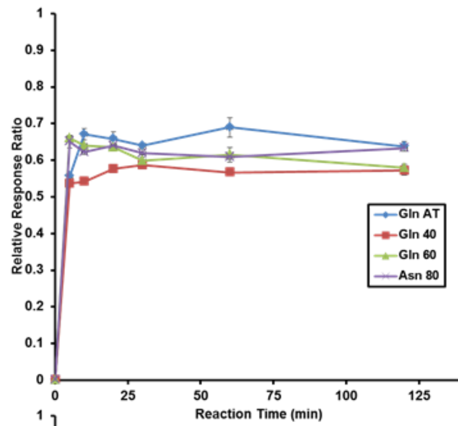
Dansyl Amino Acid	Abbreviation	Retention Time (min)
Dansyl asparagine	Dns-Asn	2.75
Dansyl aspartate	Dns-Asp	4.10
Dansyl cystine	Dns-Cystine	12.74
Dansyl glutamine	Dns-Glu	3.16
Dansyl histidine	Dns-His	13.94
Dansyl leucine	Dns-Leu	11.00
Dansyl methionine	Dns-Met	8.99
Dansyl phenylalanine	Dns-Phe	10.56
Dansyl proline	Dns-Pro	8.28
Dansyl tryptophan	Dns-Trp	9.74
Dansyl threonine	Dns-Thr	4.60
Dansyl tyrosine	Dns-Tyr	14.87
Dansyl valine	Dns-Val	8.73

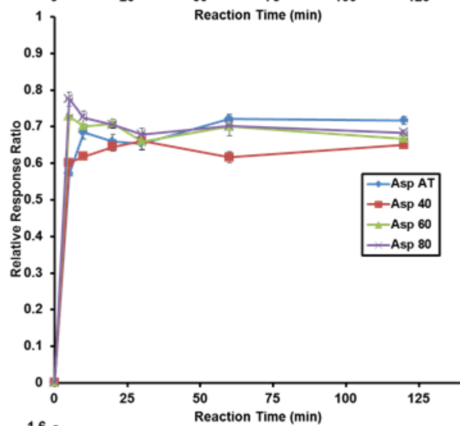
2.3.2.1 Conventional Heating Derivatization

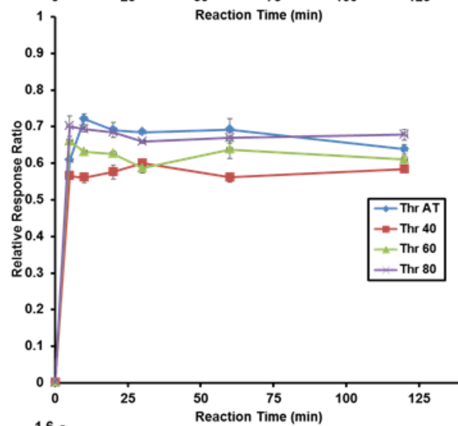
By quantifying the amount of reaction products produced during the dansylation reaction it is hoped that this would serve as a method of comparison between different reaction conditions. There are a number of variables involved in the dansylation reaction with the most important assessed in this study being reaction time, reaction temperature,

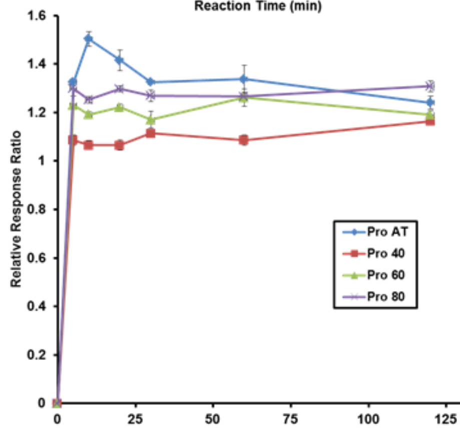
heating method, and relative component concentration. Figure 2-9 describes the results of comparing reactions heated using a conventional oven with different reaction temperatures of different amounts of time. The time difference was achieved by removing triplicate samples from the oven at the specified time points and quenching the reactions with sodium hydroxide. Conventional reactions such as those described in the literature were shown to take at least a half hour for completion when derivatizing amino acids. Reactions were often done either at room temperature or at elevated temperatures.[40-42] Following this trend and the fact that our research group is presently using a reaction time of 60 minutes at a temperature of 60°C, the experiments were chosen to analyze these conditions. Reactions were done at 5, 10, 20, 30, 60 and 120 minute intervals. It was expected that as time increased the amount of product would increase as well and most likely plateau at some point. This is the case for Dns-Asn, Dns-Asp, Dns-Gln, Dns-Pro, and Dns-Thr which can all be seen in Figure 2-9. These results suggest that the reaction time is not of major significance as many of these derivatized metabolites seem to reach a maximum concentration amount very quickly. Most of these compounds are at their maximum after only a five minute reaction time, with many that show decreases as reaction time is increased further (Dns-Met, Dns-Val, Dns-Trp, Dns-Phe, Dns-Leu, Dns-Cystine, Dns-His, and Dns-Tyr). The most notable of these is Dns-Met, which had the greatest reduction in product over time. This trend may possibly be explained by the decomposition of the ester formed between the DnsCl and methionine. Another possible explanation is that the sulfide is being oxidized during longer reactions which results in a change of retention time and making this reaction artifact unidentifiable in the chromatogram. Increases in temperature have been shown to increase the amount of by-product formed, thereby reducing the amount of derivatized product.[43]

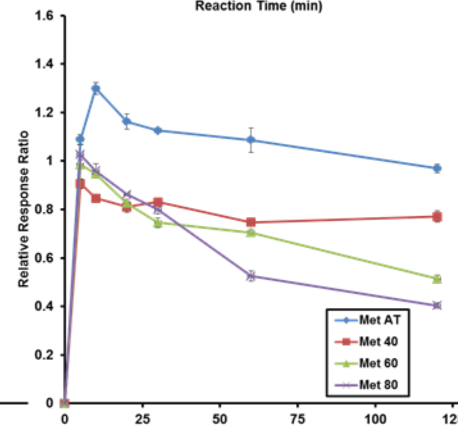

Relative Response Ratio
Reaction Time (min)
Asn AT
Asn 40
Asn 60
Asn 80

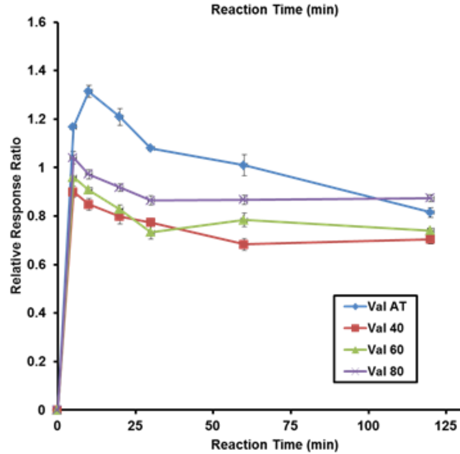

Relative Response Ratio
Reaction Time (min)
Gln AT
Gln 40
Gln 60
Asn 80

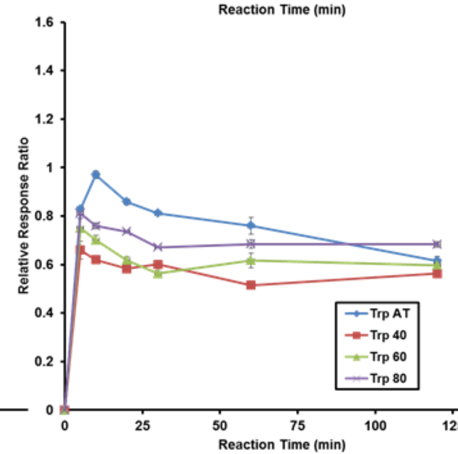

Relative Response Ratio
Reaction Time (min)
Asp AT
Asp 40
Asp 60
Asp 80


Relative Response Ratio
Reaction Time (min)
Thr AT
Thr 40
Thr 60
Thr 80


Relative Response Ratio
Reaction Time (min)
Pro AT
Pro 40
Pro 60
Pro 80


Relative Response Ratio
Reaction Time (min)
Met AT
Met 40
Met 60
Met 80


Relative Response Ratio
Reaction Time (min)
Val AT
Val 40
Val 60
Val 80


Relative Response Ratio
Reaction Time (min)
Trp AT
Trp 40
Trp 60
Trp 80

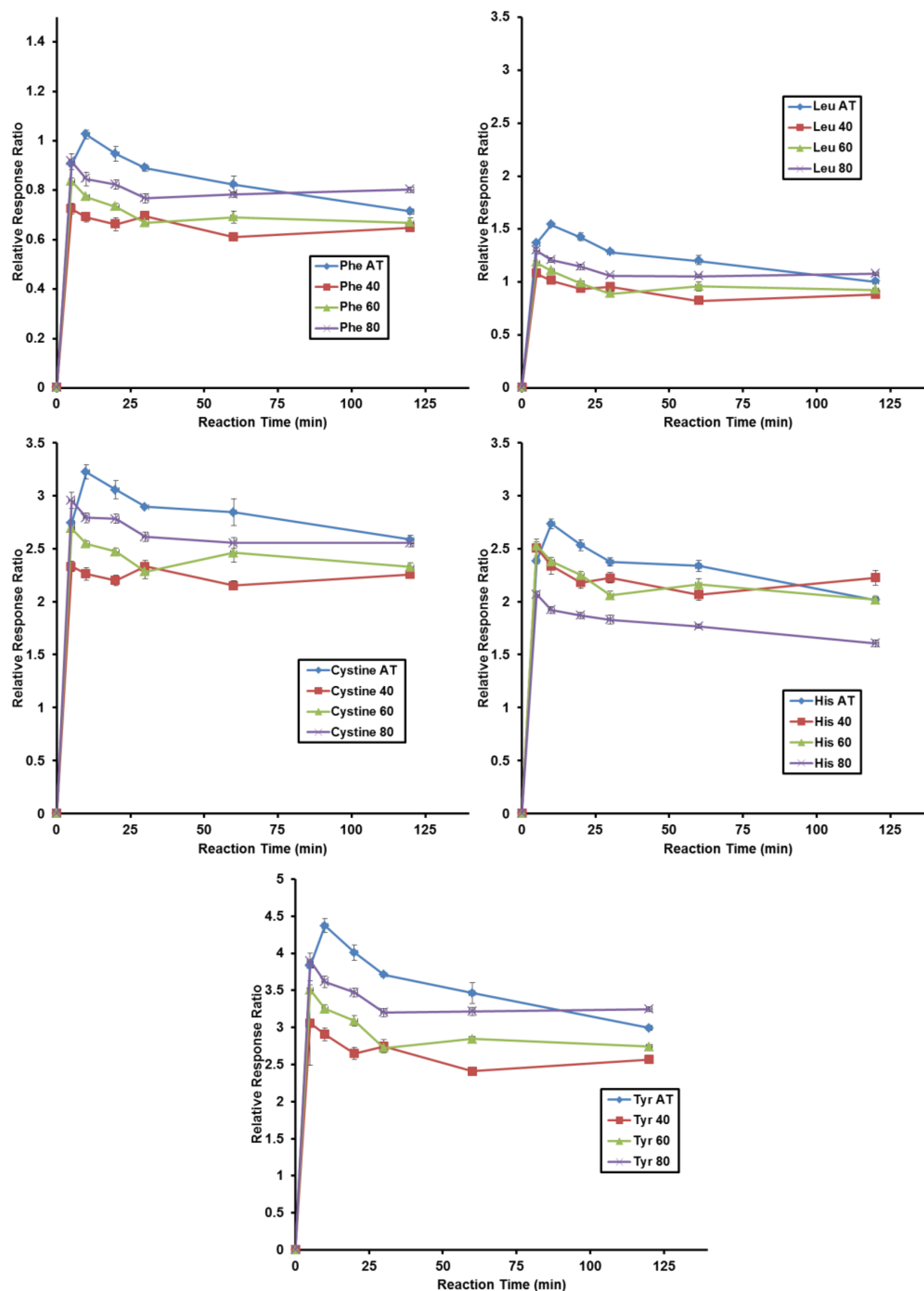


Figure 2-9 UPLC-UV quantification of labeled amino acids under different reaction temperatures and reaction times. Relative response is the peak area ratio of each labeled peak to that of the internal standard Dns-glycine added post reaction. Reaction temperatures are given in the legend with ambient temperature (AT, 21-24°C), 40°C, 60°C, or 80°C.

Degradation increases with temperature (from AT-80°C) and this would further explain why there is a reduction in the amount of Methionine at increased temperatures. Interestingly the reaction temperature seems to follow a trend of increasing product for the reactions from 40-80°C but noticeably the reactions at ambient temperature often produce a greater amount of labeled product. This was true for reaction time up 60 minutes. This can be seen in reactions at 120 minutes, which have the greatest amount of product formed at 80°C. These results show that perhaps no heating is required for the labelling of metabolites. However the metabolites used here are only a small sample size compared to the diversity that exists in real biological samples.

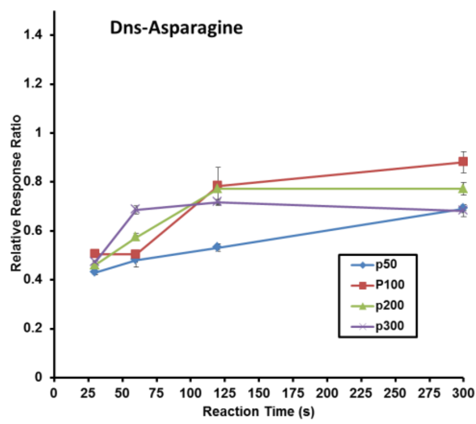
2.3.2.2 Microwave Reactor Assisted Derivatization

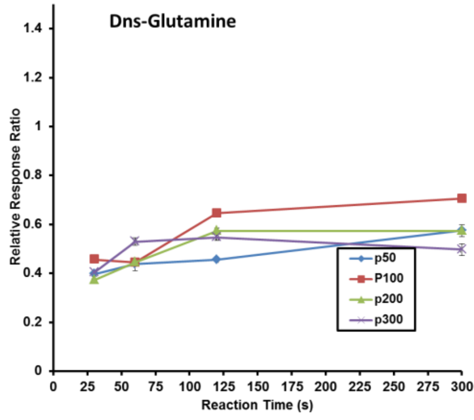
The exploration of microwave heating for derivatization was used in this study to assess if the reaction time of 60 minutes could be reduced. Reaction temperatures are also a concern and subjecting a biological sample to high temperatures over long periods of time could in theory cause degradation or side reactions that make resulting data confusing. Besides streamlining the sample preparation process it was hypothesized that using a short microwave reaction could decrease reaction artifact formation due to the reduction of time during which the reaction mixture is exposed to increased heat. When using conventional heating the vessel must be heated from the outside in, during microwave heating the solvent (in this experiment water and acetonitrile) molecules directly interact with the radiation created by the microwave. This leads to rapid heating and possibly the improved reaction yields from nonthermal microwave effects. This explanation is the reasoning behind the advent of microwave ovens being used for chemical reactions and has been demonstrated to decrease the time and improve the yield of many types of chemical reactions.[44]

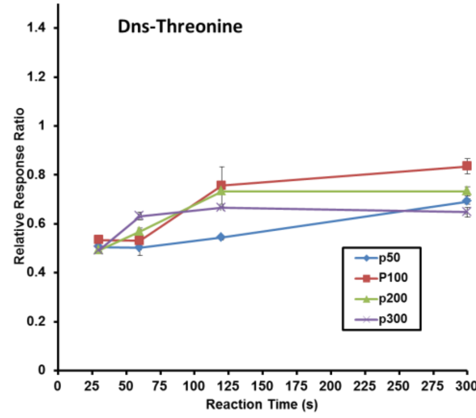
Microwave heating relies on two controllable parameters: microwave power (w) and irradiation time (s). To assess variable reaction conditions the same method for quantification of labelled metabolites was used as described above. The results of these initial experiments are shown in Figure 2-10. Each of the amino acids present in the standard mixture is shown. The amounts of product relative to the internal standard, Dns-Glycine, are plotted against the reaction time. The most notable trend is that the shorter reaction times (30 s and 60 s) have a much lower relative response than that of the longer reaction times (120 s and 300 s). A number of the compounds (Dns-Pro, Dns-Val, Dns-Trp, Dns-Phe, and Dns-Leu) have very large standard deviations at the shorter reactions times suggesting that these reactions are incomplete. As the reactions increase to 300 s there is a large reduction in the amount of derivatized product. Most notable are Dns-Met, Dns-His, and Dns-Tyr. In some cases this difference is as high as 45%. This suggests that the temperature reached during the 300 s reaction is high enough for degradation to occur. This follows the trend displayed in Figure 2-9 for which the amount of derivatized product is reduced at longer reactions times. The power levels chosen in this study were based on the maximum wattage attainable which for the CEM discover is 300 W. An increase in power level seems to cause an increase in the amount of product formed, but only to a certain point. In Figure 2-10 and Figure 2-13 it can be seen that there is much less product formed when using 50 W to irradiate the sample. Application of 100 W and 200 W produce samples with comparable amounts (Dns-Pro, Dns-Asp, Dns-Asn), some samples benefited from 100 watts (Dns-Phe, Dns-Leu, Dns-Tyr) while others (Dns-His) from a higher power setting. Figure 2-10 demonstrates that both Dns-His and Dns-Tyr, when irradiated at 300 W for 300 s show a large reduction in amount and also a large variability between samples. This likely occurs due to the same reasons discussed for Dns-Met reduction when using conventional heating. For microwave

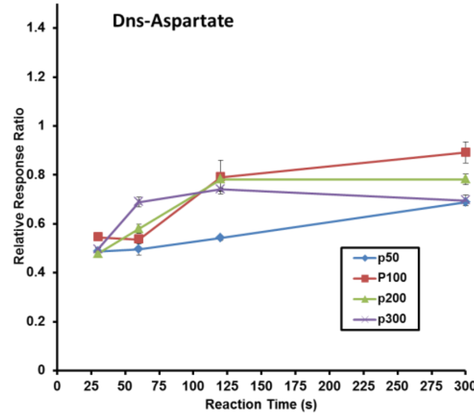
reactions it was decided that 100 W for 120 s produced a high amount of reliably labelled metabolite product.

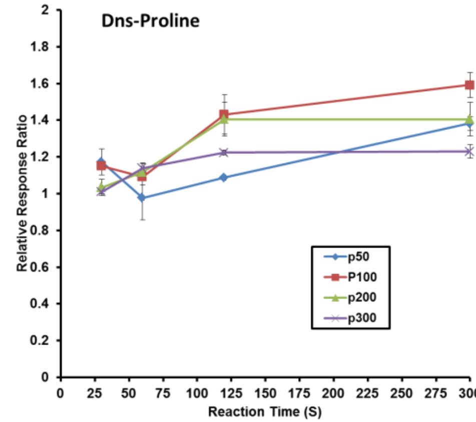
Figure 2-11 shows the results for the optimization of the CEM discover microwave reactor. In Figure 2-11 A, power levels are compared at a reaction time of two minutes. This is the shortest acceptable time with which we were satisfied that reaction of DnsCl and the metabolites would be complete. The intermediate power levels of 100 and 200 watts provide the best results. Discerning between the two is trickier with the differences between individual metabolites being small. 100 W was chosen as the more preferable power because in more complicated biological samples it is assumed that the chemical diversity may prove more challenging to label completely. In Figure 2-11 B the reaction time is compared when using a reaction power of 100 w. The best results were seen at times of 120 and 300 seconds. It was decided that 120 seconds was preferable as it would reduce the amount of time that the sample would be subjected to higher temperatures and therefore reduce thermal degradation in the sample.

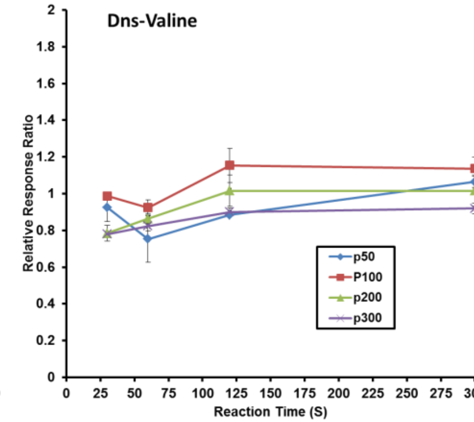

Dns-Asparagine
Relative Response Ratio
Reaction Time (s)
p50
P100
p200
p300

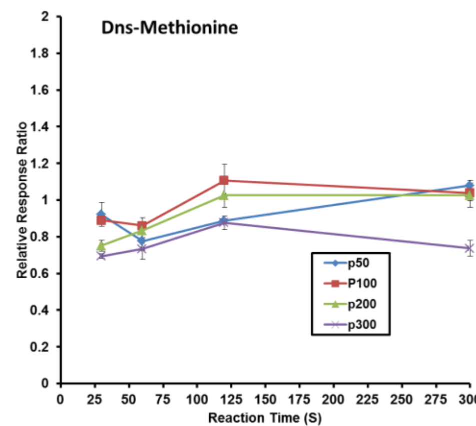

Dns-Glutamine
Relative Response Ratio
Reaction Time (s)
p50
P100
p200
p300

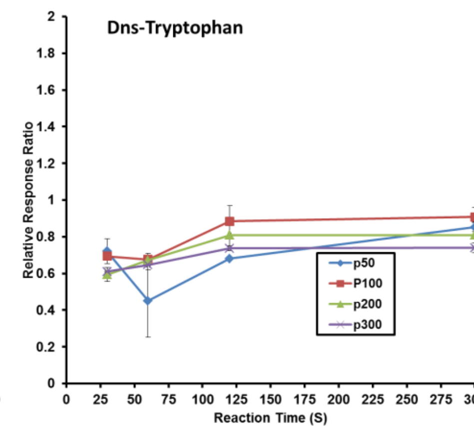

Dns-Threonine
Relative Response Ratio
Reaction Time (s)
p50
P100
p200
p300


Dns-Aspartate
Relative Response Ratio
Reaction Time (s)
p50
P100
p200
p300


Dns-Proline
Relative Response Ratio
Reaction Time (S)
p50
P100
p200
p300


Dns-Valine
Relative Response Ratio
Reaction Time (S)
p50
P100
p200
p300


Dns-Methionine
Relative Response Ratio
Reaction Time (S)
p50
P100
p200
p300


Dns-Tryptophan
Relative Response Ratio
Reaction Time (S)
p50
P100
p200
p300

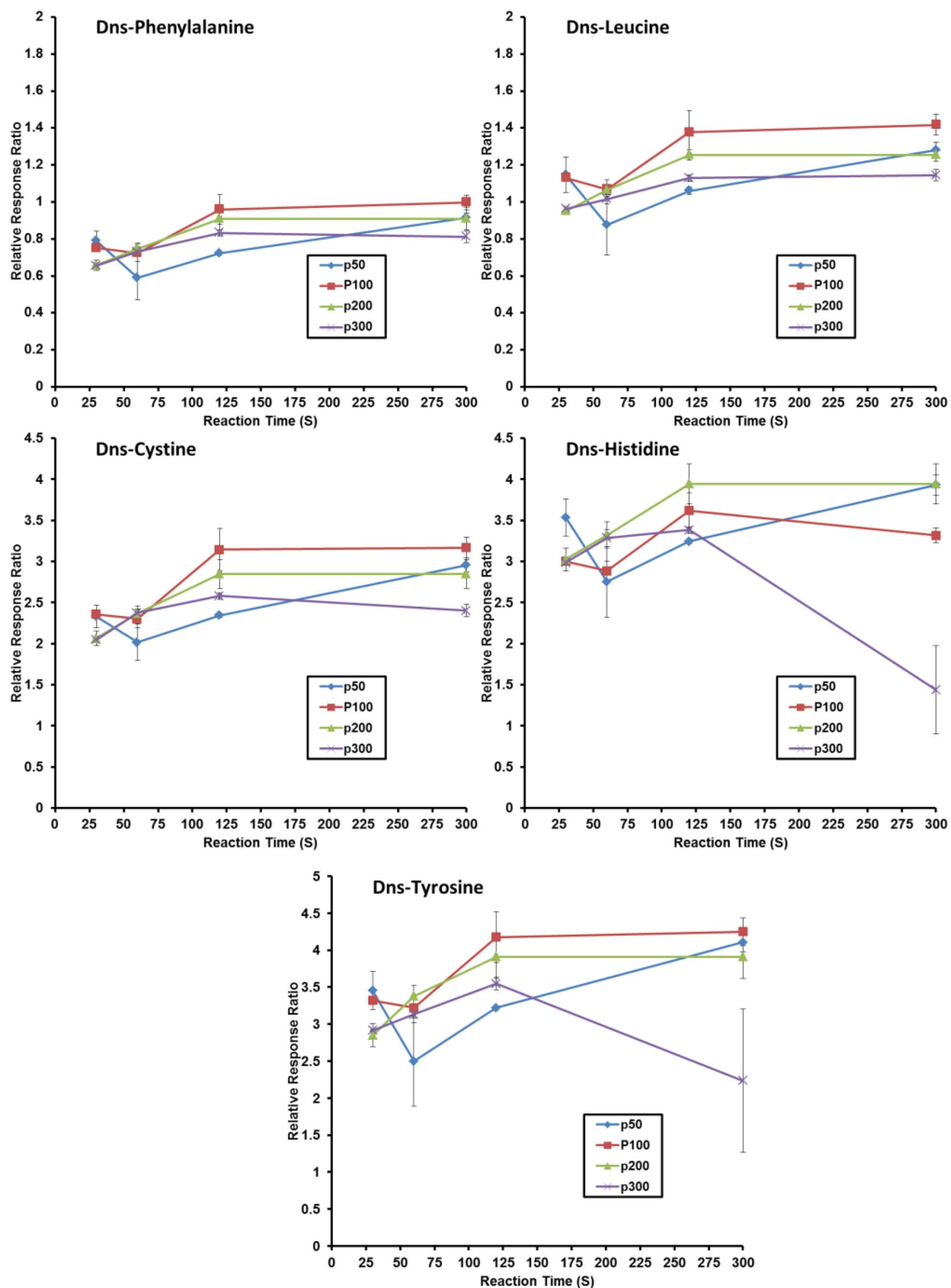


Figure 2-10 UPLC-UV quantification of labeled amino acids labeled using a CEM Discover microwave reactor. Relative response is the peak area ratio of each labeled peak to that of the internal standard Dns-Glycine. The amount of power applied (w) to each reaction is given in the legend.

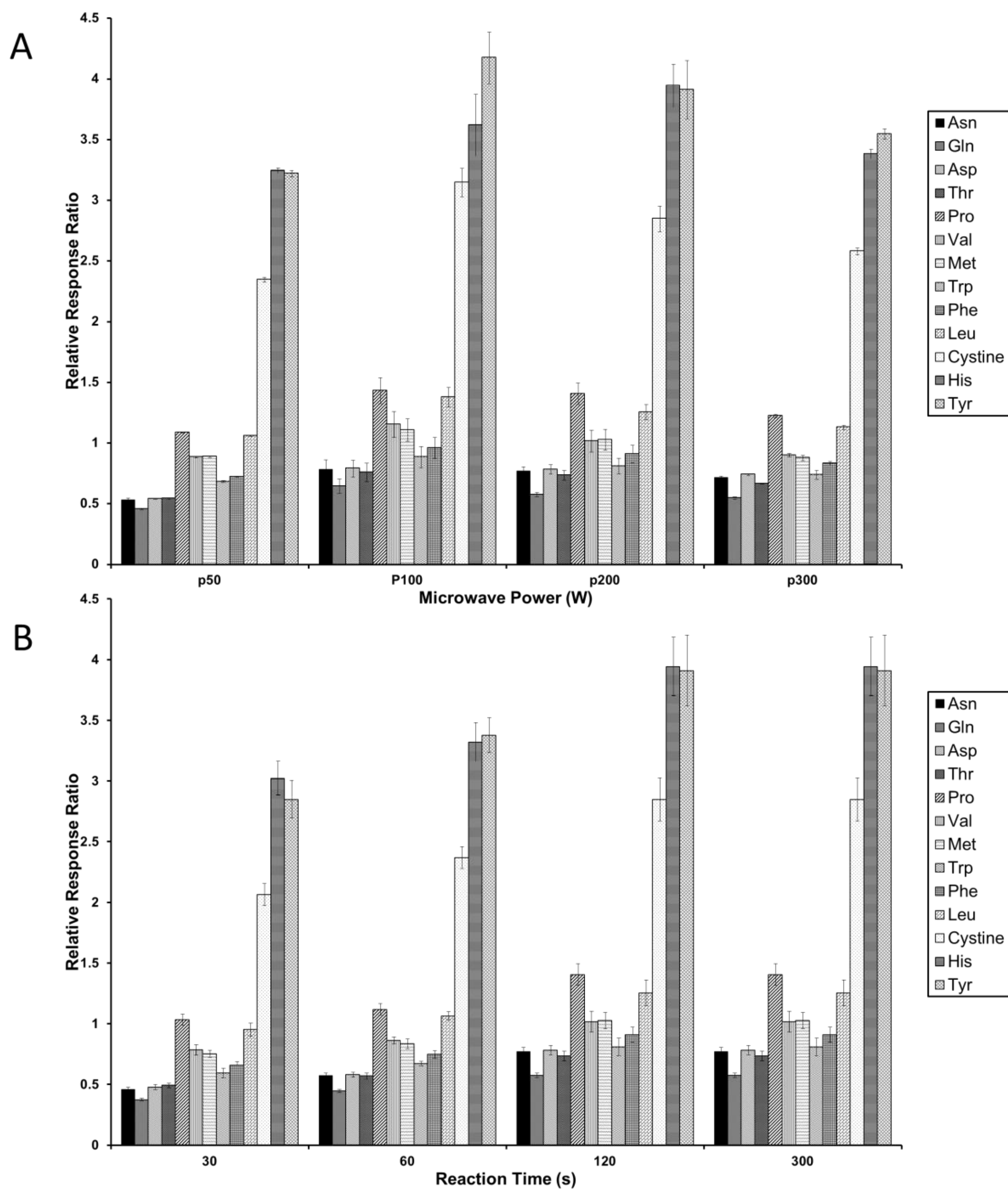


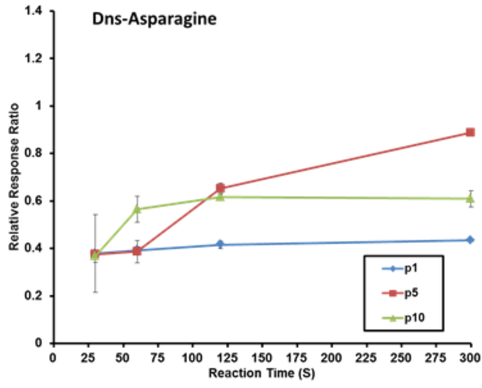
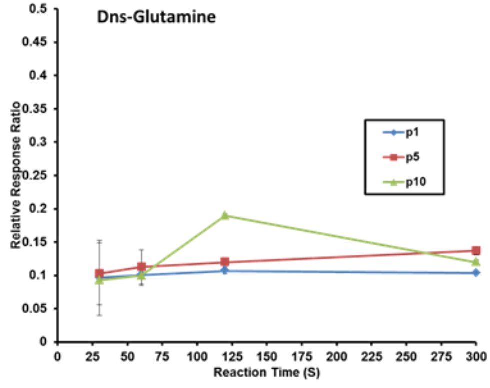
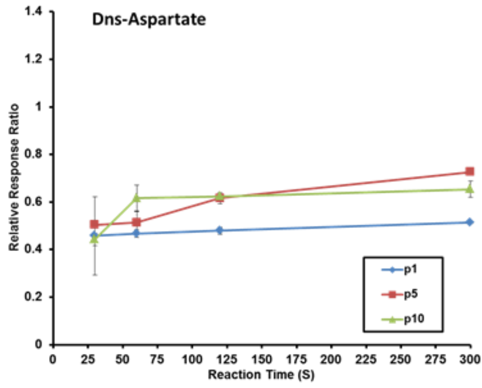
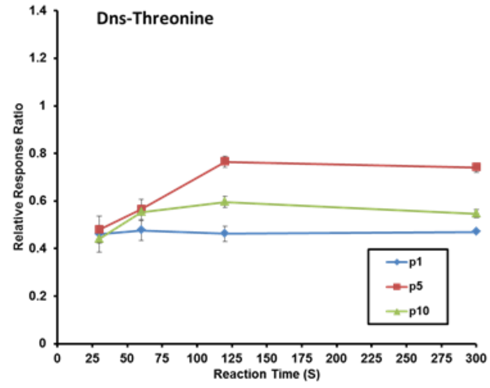
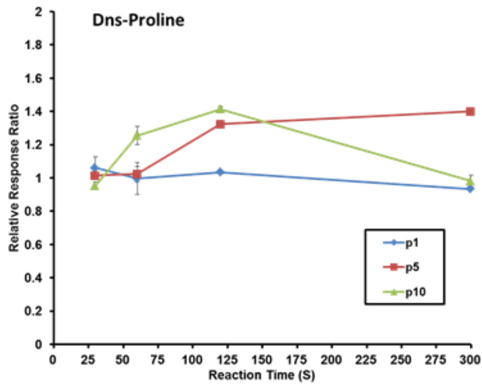
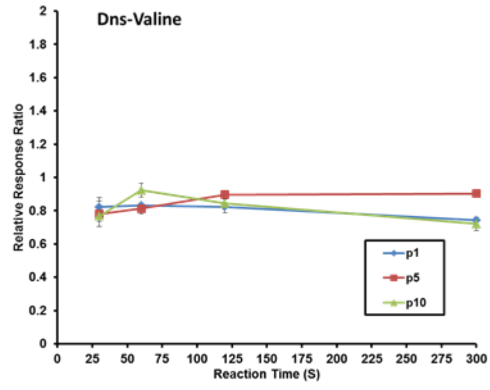
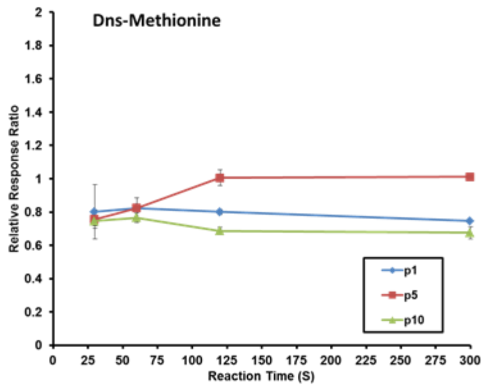
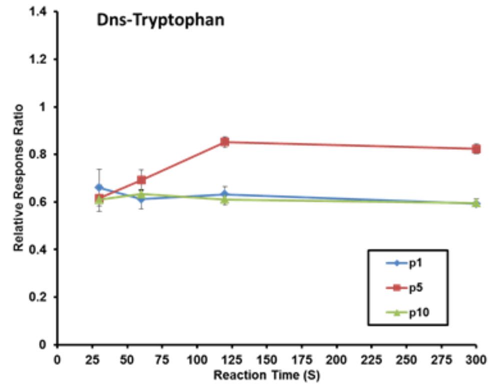
Figure 2-11 The effect of microwave power and time on the amount of dansyl labeled amino acid using a CEM Discover microwave. Power comparison was done for a reaction time of 120 s (A). Time comparison is done with 100 watts applied power (B).

2.3.2.3 Consumer Microwave Assisted Derivatization

The previous section describes the use of a dedicated microwave reactor designed for organic synthesis, digestion, and other purely chemical purposes. It is hoped that the methods explored in this paper would be widely applicable. That is easily and inexpensively reproduced. For this to be true it is important to reduce the amount of specific equipment required for sample preparation, in this case the microwave oven. For this purpose the use of a consumer microwave oven (CMO) for dansyl metabolite derivatization was explored. The peak ratio results for each individual metabolite are shown in Figure 2-12. The CMO had a much wider power range and a much higher maximum for applied power (120-1200 w). The domestic oven has no built in stirring system, which is standard in the dedicated reactor models. As well there is no temperature measurement or control. Despite these shortcomings, reactions using this type of oven show much the same trends and produce labeled samples with reasonable reproducibility as those completed using the other methods of heating. The lowest power setting (p1) produces a field of 120 watts and the results seen as the blue lines in Figure 2-12. This demonstrates that there is not much difference in the amount of derivatized product over all time points and that these amounts are generally lower than those reacted at higher power levels and therefore temperatures. Using p5 or 600 watts produces results that increase in derivatized peak area over time and also produce the greatest amount of product of the three power levels tested. The highest power level p10 or 1200 w produced reasonable amounts of product at lower time intervals but with any reaction longer than 1 minute there was a wide variation in the amount of product produced. As with the other reaction methods this is most noticeable in Dns-His and Dns-Tyr. This may suggest that these two metabolites do not react well with dansyl chloride, or that they easily degrade under extreme conditions. There is also a very noticeable trend of a major reduction in dansylated metabolite between the 120 s and 300 s reaction times when using this higher

power level. As with shorter reactions using the CEM Discover, the commercial microwave labeled samples from this time period display highly variable results with large standard deviations for triplicate experiments.

A comparison of microwave power and time for the commercial oven is given in Figure 2-13 A and B. The power comparison takes place at a reaction time of 120 s in order to be consistent with the CEM Discover results and because it produced reliable amounts. The best result appears to be 600 W, although comparable with the reaction at 1200 W, it was considered favourable due to reduced thermal heating. Keeping the temperature reached during reaction is desirable as it is thought to reduce the amount of by-product formed. The reaction time was compared using a reaction power of 600 W and the results show that much like the CEM Discover, a 120 s reaction is sufficient to produce a labelled sample. A number of metabolites in the 30 and 60 s reactions show a large variation in amount and therefore are probably not complete reactions. The 300 s reaction reached a higher temperature (see appendix for temperature profiles) so it was considered more important to reduce the maximum temperature reached in the reaction instead of increasing the reaction time for a slight increase in derivatized metabolite.


Dns-Asparagine
Relative Response Ratio
Reaction Time (S)
p1
p5
p10

Dns-Glutamine
Relative Response Ratio
Reaction Time (S)
p1
p5
p10

Dns-Aspartate
Relative Response Ratio
Reaction Time (S)
p1
p5
p10

Dns-Threonine
Relative Response Ratio
Reaction Time (S)
p1
p5
p10

Dns-Proline
Relative Response Ratio
Reaction Time (S)
p1
p5
p10

Dns-Valine
Relative Response Ratio
Reaction Time (S)
p1
p5
p10

Dns-Methionine
Relative Response Ratio
Reaction Time (S)
p1
p5
p10

Dns-Tryptophan
Relative Response Ratio
Reaction Time (S)
p1
p5
p10

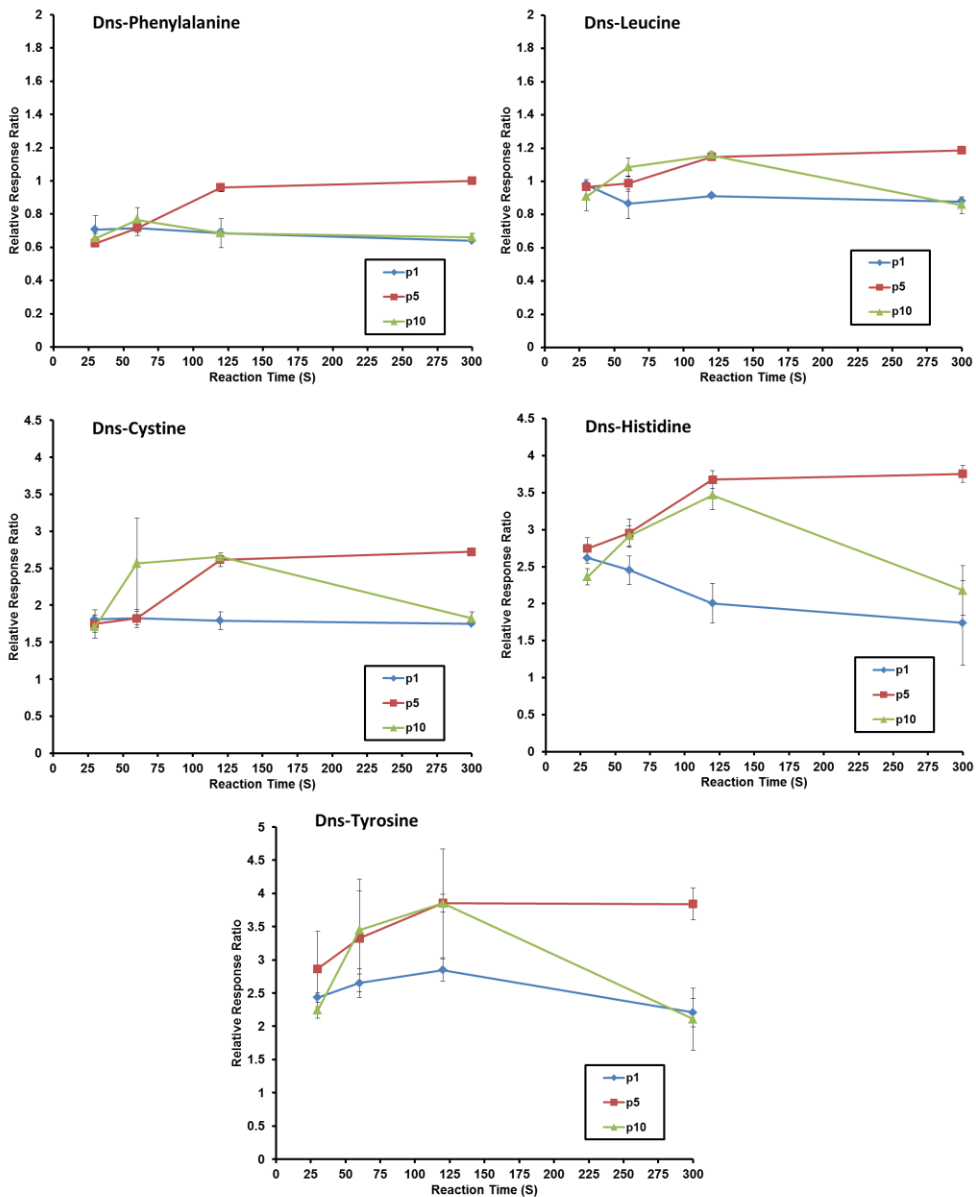


Figure 2-12 UPLC-UV quantification of labeled amino acids labeled using a Panasonic consumer microwave. Relative response is the peak area ratio of each labeled peak to that of the internal standard Dns-Glycine. The amount of power applied (power level) is given in the legend.

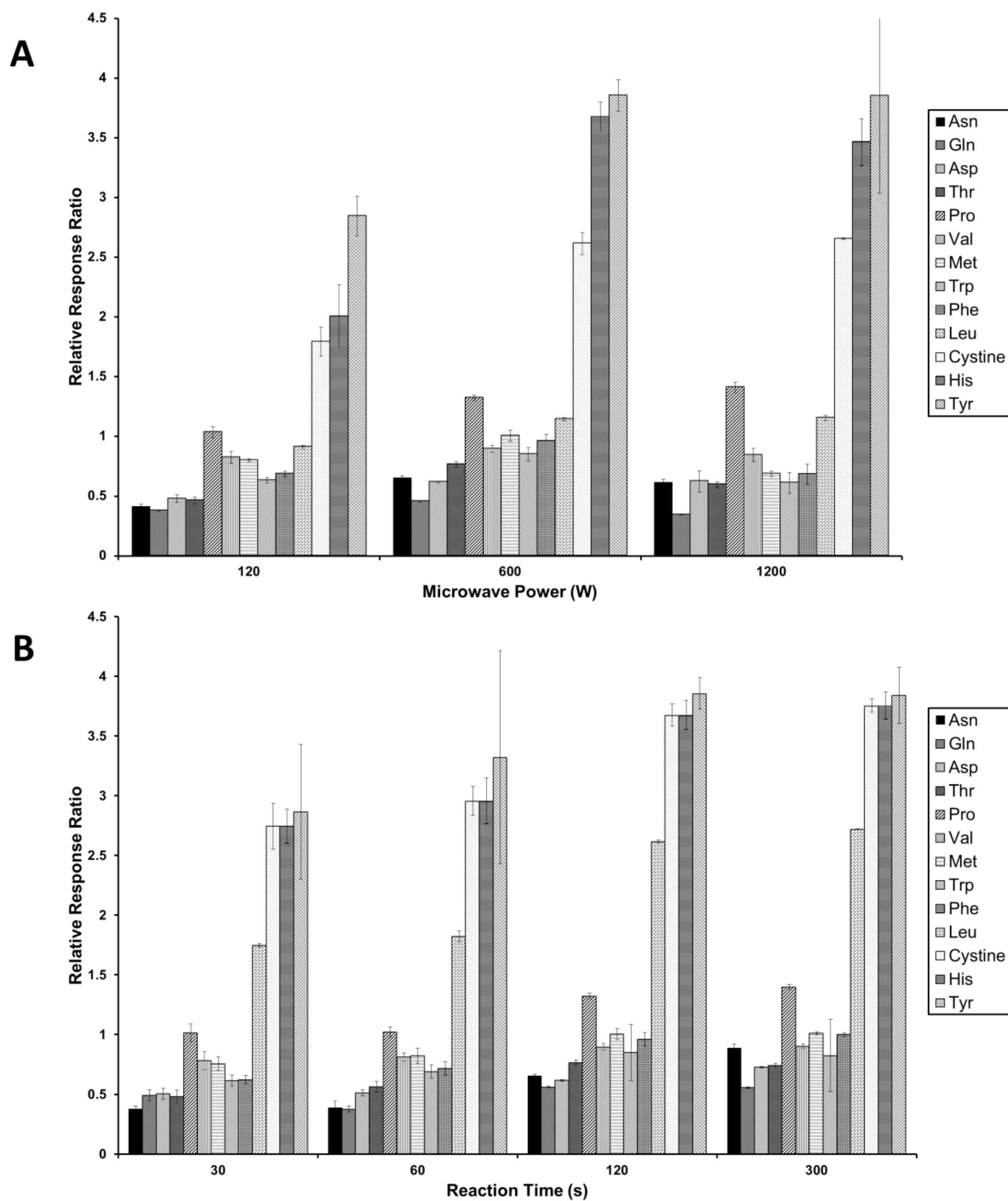


Figure 2-13 The effect of microwave power and time on the amount of dansyl labeled amino acid using a Panasonic consumer microwave. Power comparison was done for a reaction time of 120 s (A). Time comparison is done with 200 W applied microwave power (B).

2.3.2.4 Comparison of Reaction Methods

In order to evaluate and compare each of these methods, the peak area ratio results of the CEM Discover, Panasonic microwave, traditional heating method, and ambient temperature or no heating are displayed in Figure 2-14. Each of the compounds assessed in this study is represented including one of the dansyl byproduct peaks, termed in this figure DnsBP and appears at 3.8 minutes in Figure 2-8. As with the other experiments all reactions were performed in triplicate and standard deviations are shown. The microwave reactions were both 120 s, while the conventional heating and no heating reactions were reacted for 60 minutes. The general trend is that the microwave reactions have higher or equal relative response ratio when compared with the other methods. When comparing the two microwave methods the peak area ratios using the CEM Discover are generally higher. It is possible that the more controlled and consistent microwave environment produced these results. The DnsBP peak values show an increase with the conventional heating method with is consistent with other results. The MAH methods both show reduced amounts of DnsBP but greater than that of ambient temperature. This demonstrates that heating is a factor in by-product formation. It is currently unknown whether or not this would affect stable isotope labeling metabolomics studies. These results demonstrate that microwave heating is a viable option that produces derivatized metabolite samples of equal or greater quality to that of conventional heating or no heating methods currently being employed for this type of derivatization.

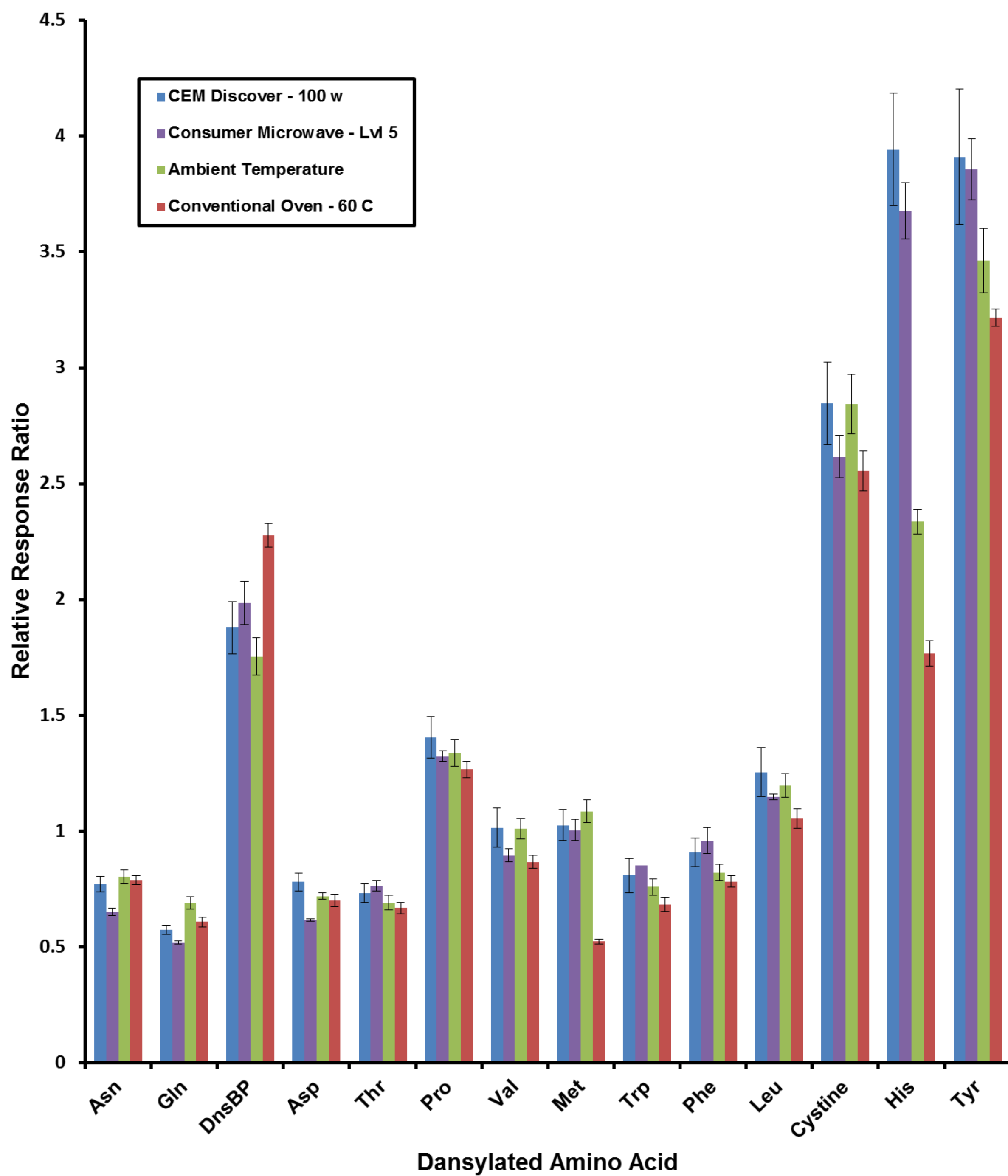


Figure 2-14 Comparison of quantification results for various derivatization methods. The conventional heating methods were both reacted for 60 minutes while the microwave reactions took place over 2 minutes.

2.3.3 Method Validation

2.3.3.1 Linearity and Reaction Reproducibility

To assess the differences in reaction conditions it was important to develop a rigorous quantitative method that would allow for both relative and absolute determination of derivatized compounds produced with differing conditions. Relative quantification was carried out for all compounds in the mixture by using the relative peak area response of each derivatized metabolite peak divided by the peak are of the internal standard Dns-Gly. All of the metabolites assessed in this study were reproducibly labeled at the three higher concentrations (28, 10, and 1 mM total metabolite concentration) with RSD values <10%. This experiment was carried out for both a conventionally heated and a microwave assisted reaction and the results are given in Tables 2-4 and 2-5. At an initial total metabolite concentration of 0.1 mM the reproducibility is poor with many RSD values >10%. This is most likely due to the fact that this concentration of metabolite is close to or below the LLOQ for the method. Calibration curves for both reaction methods are given in Figures 2-15 and 2-16. As with previous results the microwave assisted heating method produced samples with higher relative response ratios and reproducibility comparable to that of the longer conventional heating method. Linear response is observed across the concentration range tested with almost all R^2 values greater than 0.998 for both heating methods. It is expected that the lowest concentration tested was too low for UPLC-UV detection and that using a slightly higher concentration would produce better results.

2.3.3.2 External Calibration and % Yield Determination

To determine the % yield of the dansylation reaction an external calibration method was used for the absolute quantification of a number of the labeled metabolites found in the mixture. The two standards used were Dns-aspartic acid and Dns-threonine.

Table 2-4 Repeatability (RSD), R^2 and relative response vs. internal standard (Dns-Gly) at four different concentrations of amino acid standard (n=3). Total metabolite concentrations are 28, 10, 1, and 0.1 mM. Reaction conditions are: Conventional heat; 60 minutes, 60°C.

Compound	Equation and R^2	Relative Response Ratio to IS	RSD (%) (*n*=3)
Dns-Asn	y = 0.0205x + 0.0066	0.6	3.1
	R^2=0.998	0.2	7.5
		0.02	6.1
		0.003	16.6
Dns-Gln	y = 0.0203x + 0.0007	0.569	0.8
	R^2=1.000	0.2	2.6
		0.02	8.3
		0.003	10.9
Dns-Asp	y = 0.0212x + 0.0048	0.6	1.5
	R^2=0.9999	0.2	9.6
		0.03	6.5
		0.004	8.5
Dns-Thr	y = 0.0187x + 0.0049	0.5	4.2
	R^2=0.9986	0.2	7.4
		0.02	4.9
		0.002	15.6
Dns-Pro	y = 0.0359x + 0.0128	1.0	3.9
	R^2=0.9989	0.4	8.3
		0.05	7.8
		0.005	8.9
Dns-Val	y = 0.0238x + 0.0032	0.7	7.1
	R^2=0.9995	0.3	7.7
		0.02	8.6
		0.002	18.6
Dns-Met	y = 0.0277x - 0.0095	0.8	11.8
	R^2=0.9983	0.2	7.0
		0.02	7.9
		0.002	19.2
Dns-Trp	y = 0.0187x - 0.0014	0.5	6.3
	R^2=1.000	0.1	8.1
		0.01	6.3

		0.001	11.9
Dns-Phe	y = 0.022x + 0.0050	0.6	8.3
	R^2=0.9985	0.2	7.3
		0.02	5.0
		0.002	24.4
Dns-Leu	y = 0.0285x + 0.0047	0.7	9.2
	R^2=0.9993	0.3	7.5
		0.02	6.6
		0.003	15.0
Dns-Cystine	y = 0.0065x - 0.0018	1.8	5.7
	R^2=0.9999	0.6	8.2
		0.05	8.2
		0.005	16.6
Dns-His	y = 0.0776x + 0.0033	2.1	4.4
	R^2=0.9998	0.8	13.9
		0.07	6.4
		0.007	16.3
Dns-Tyr	y = 0.0773x - 0.0134	2.1	6.8
	R^2=0.9995	0.7	8.3
		0.06	6.5
		0.01	9.5

Table 2-5 Repeatability (RSD), R^2 and relative response vs. internal standard (Dns-Gly) at four different concentrations of amino acid standard (n=3). Total metabolite concentrations are 28, 10, 1, and 0.1 mM. Reaction conditions are: CEM Discover, 200 s, and 100 watts.

Compound	Equation and R^2	Relative Response Ratio to IS	RSD (%)
Dns-Asn	y = 0.0286x + 0.0014	0.6	3.1
	R^2=0.9998	0.2	7.5
		0.03	6.1
		0.003	16.6
Dns-Gln	y = 0.0203x + 0.0007	0.6	0.8
	R^2=1.000	0.2	2.6
		0.02	8.3
		0.003	10.9
Dns-Asp	y = 0.0295x + 0.0101	0.6	1.5
	R^2=0.9993	0.2	9.6
		0.03	6.5
		0.004	8.5
Dns-Thr	y = 0.0257x + 0.0102	0.5	4.2
	R^2=0.9967	0.2	7.4
		0.02	4.9
		0.002	15.6
Dns-Pro	y = 0.0501x + 0.0093	0.8	3.2
	R^2=0.9999	0.3	4.1
		0.03	6.1
		0.003	16.6
Dns-Val	y = 0.0333x + 0.0074	0.6	2.8
	R^2=0.9987	0.2	7.2
		0.02	11.8
		0.003	15.4
Dns-Met	y = 0.0307x - 0.0098	0.8	1.9
	R^2=0.9972	0.3	4.6
		0.04	6.5
		0.005	8.5

Dns-Trp	y = 0.0261x + 0.0002	0.7	4.5
	R^2=0.9997	0.3	3.1
		0.03	4.9
		0.003	15.6
Dns-Phe	y = 0.0307x + 0.0098	1.4	4.0
	R^2=0.9972	0.5	4.8
		0.06	7.8
		0.007	8.9
Dns-Leu	y = 0.0398x + 0.0102	0.9	7.3
	R^2=0.9984	0.4	4.5
		0.03	8.6
		0.003	18.6
Dns-Cystine	y = 0.1128x - 0.0205	1.1	11.8
	R^2=0.9999	0.4	3.4
		0.03	7.9
		0.004	19.2
Dns-His	y = 0.1392x - 0.0021	0.8	6.5
	R^2=0.9998	0.3	5.9
		0.02	6.3
		0.9	
Dns-Tyr	y = 0.1401x - 0.009	0.9	5.1
	R^2=0.9998	0.4	5.0
		0.03	5.0
		0.003	24.4

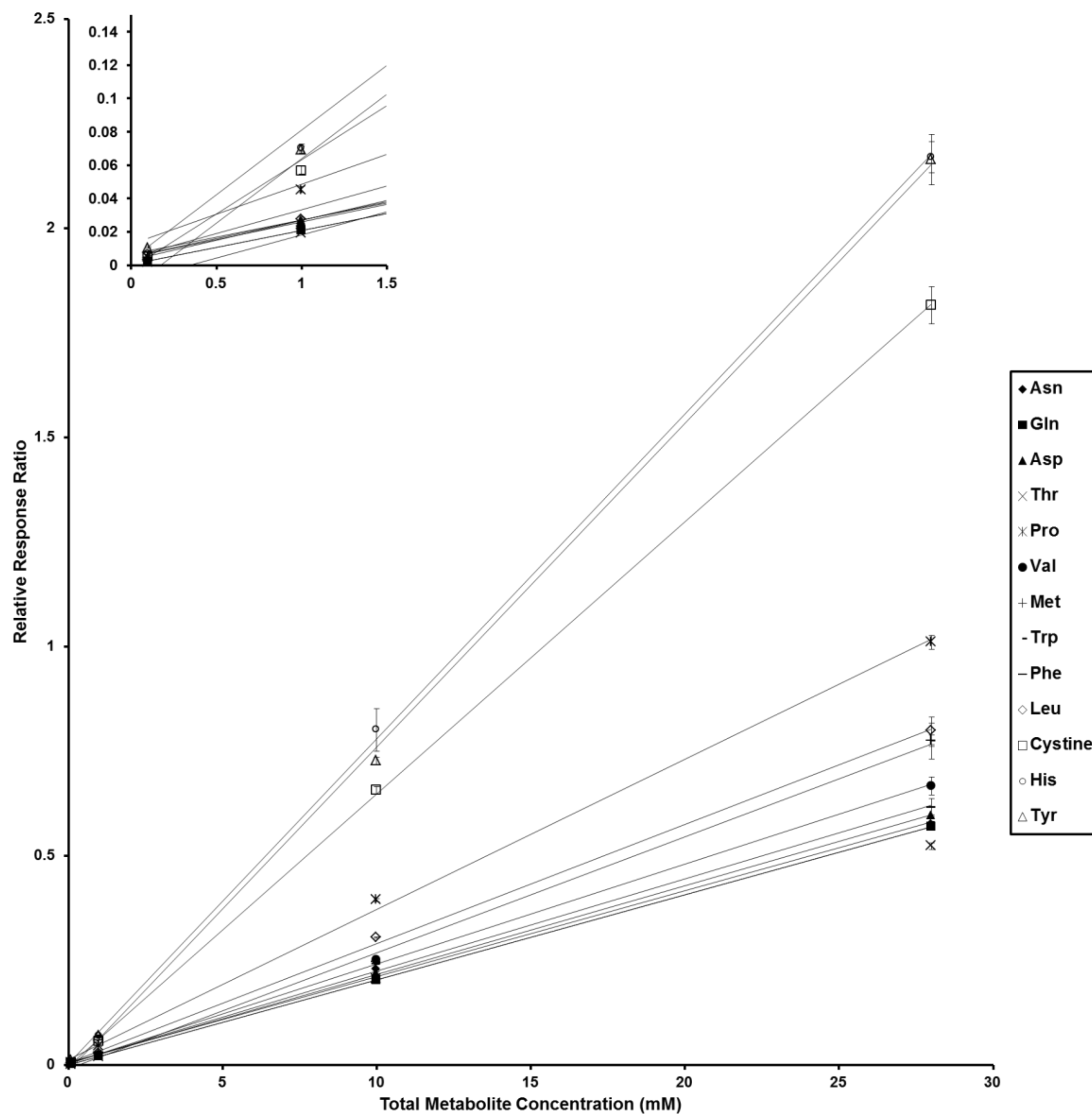


Figure 2-15 Quantification of amino acid standards using UPLC-UV over the range of 0.1 mM-28 mM total metabolite concentration. Samples derivatized using conventional heating at 60°C for 60 minutes.

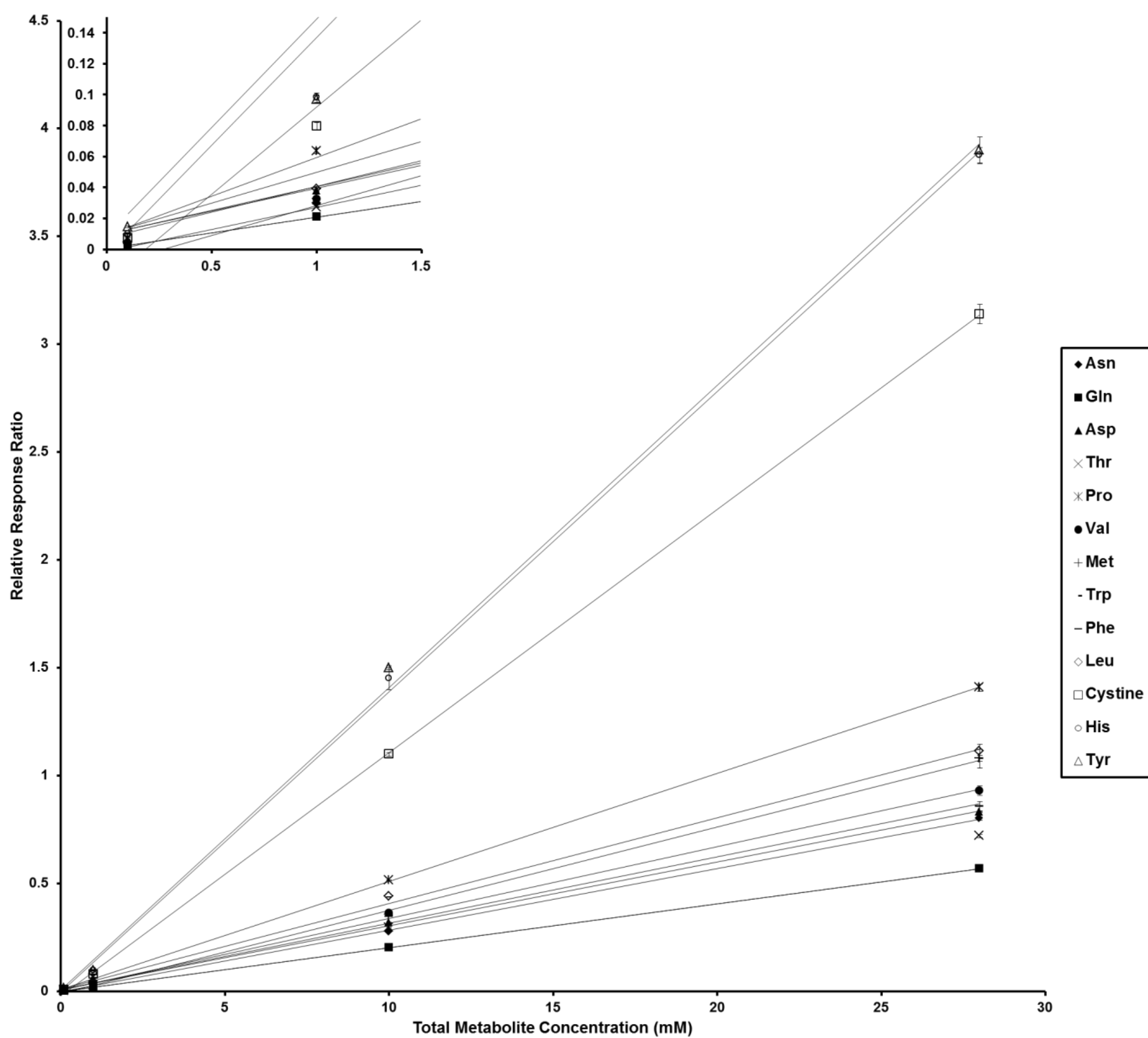
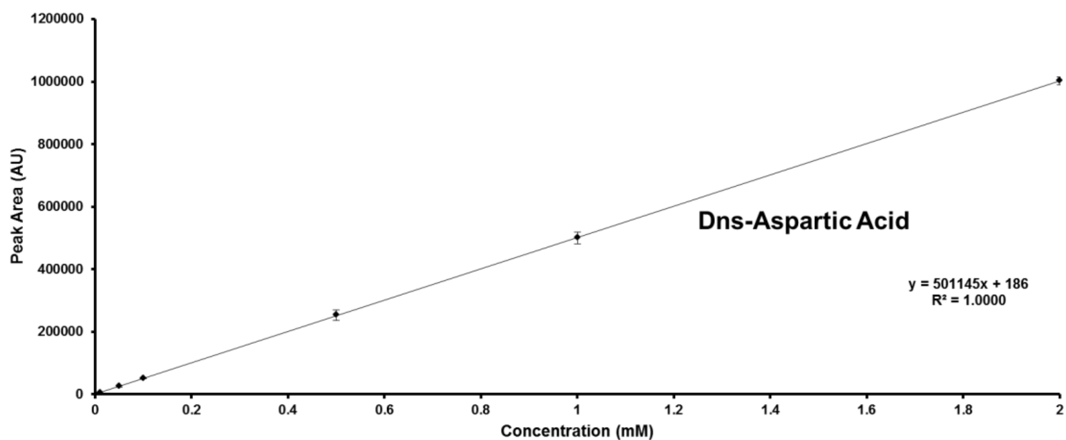
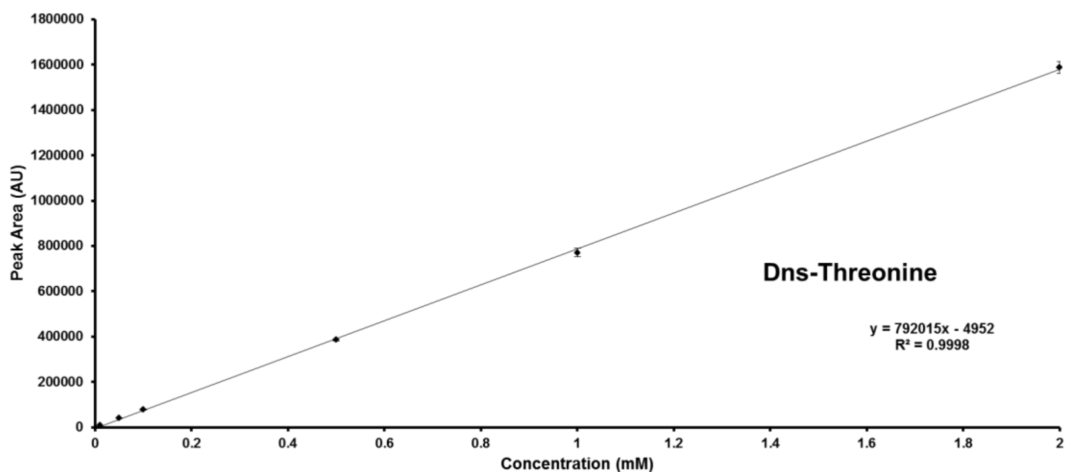


Figure 2-16 Quantification of amino acid standards using UPLC-UV over the range of 0.1 mM- 28 mM total metabolite concentration. Samples derivatized using CEM discover microwave heating for 2 minutes at 100 watts of power.





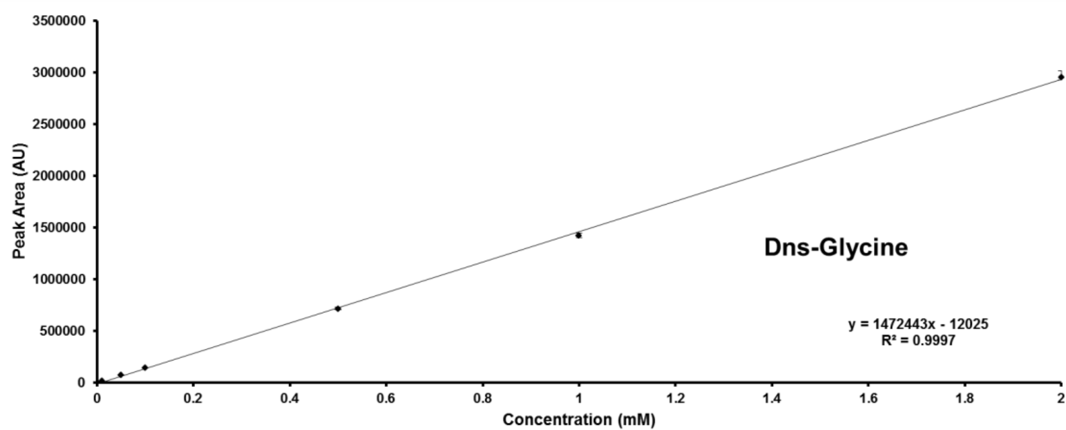


Figure 2-17 External calibration curves for Dns-aspartic acid, Dns-threonine, and Dns-Glycine with concentrations ranging from 0.01 mM to 2 mM using a 5 µL injection volume.

Table 2-6 Results of external calibration absolute quantification using purchased predansylated standards. Initial concentration is the measured amount of metabolite added initially to the reaction; final concentration is the amount calculated using the calibration curve. Microwave reaction is 200 s at 100 W, conventional reaction is 60 minutes at 60°C.

Dansyl Amino Acid	[Initial] (mM)	[Final] mM	%Yield	%RSD	Reaction Method
Dns-Aspartate	2	1.60	80.1	3.1	Microwave
	1	0.76	76.8	0.6	
	0.1	0.077	77.5	5.9	
	0.01	NQ	NQ	NQ	
Dns-Threonine	2	1.55	77.7	1.9	
	1	0.76	76.0	1.0	
	0.1	0.076	76.0	8.7	
	0.01	NQ	NQ	NQ	
Dns-Glycine	2	1.44	90.1	2.7	
	1	0.68	91.1	1.6	
	0.1	0.069	89.4	1.4	
	0.01	NQ	NQ	NQ	
Dns-Aspartate	2	1.50	75.3	3.1	Conventional
	1	0.72	72.2	0.6	
	0.1	0.072	72.8	5.9	
	0.01	NQ	NQ	NQ	
Dns-Threonine	2	1.38	69.1	1.9	
	1	0.67	67.6	1.0	
	0.1	0.06	67.0	8.8	
	0.01	NQ	NQ	NQ	
Dns-Glycine	2	1.37	91.0	2.7	
	1	0.66	92.2	1.6	
	0.1	0.066	91.0	2.8	
	0.01	NQ	NQ	NQ	

* NQ refers to Non Quantifiable Values

The calibration curves for each of these compounds were produced with 6 varying concentrations from 0.01 mM to 2 mM. The calibration curves can be seen in Figure 2-17. The results for the absolute quantification of Dns-Asp and Dns-Thr, Dns- Glycine are shown in Table 2-6. This was done as a comparison between microwave and conventional reactions with amino acid standard initially in reaction over a range of four

concentrations (from 0.01-2 mM). In each case the microwave reaction (100 w for 200s) performed well with yields ranging from 76-95%. It should be noted that the high yields given for the low concentrations are close to the LLOQ for the method. The conventional heating method (60°C for 60 minutes) produces yields much lower than that of the microwave reactions with yields ranging between 69-96%. Dns-Asp had a much higher % yield than that of Dns-Thr and shows stable yields at 3 of the 4 tested concentrations. Dns-Gly had the best performance when labelled with both MAH and conventional heating. Dns-Thr in comparison shows reduced yields when concentration is reduced. The reasons for this are unknown but this demonstrates that there is variability among the concentration of derivatized compounds and therefore the efficiency of their derivatization reactions. It is hypothesized that the presence of water may reduce the reaction efficiency. Water in the reaction mixture cannot be avoided as the end goal is the analysis of biofluids. In a similar study of biogenic amines there were similar derivatized product recovery amounts at MAH reaction times shorter than 5 minutes. This study had no comparison with conventional heating however.[45] It is important that the reaction time of our microwave reactions be kept shorter due to high temperatures reached for reactions longer than 5 minutes. This is of importance as one of the key factors in our analysis is detection of true metabolites not reaction byproducts or degraded metabolites.

2.3.3.3 Selectivity and Carryover

Evaluation of selectivity of the analytes was determined by the analysis of an instrument blank, a solvent blank, a reaction blank, and an underivatized amino acid standard mixture blank. Carryover was analyzed by running replicate instrument blank samples following the analysis of higher concentration derivatized amino acid standard samples. Chromatograms of these can be seen in Figure 2-18. Figure 2-18 (A & B) demonstrate only a small peak at approximately 15 minutes that does not interfere with

any of the peaks of interest. The internal standard Dns-Glycine was added to both the reaction blank and unreacted amino acid standard in Figure 2-18 (C & D). Large early peaks that elute before two minutes in these blanks did not interfere with the analysis of compounds of interest. There are some small peaks that elute later in the chromatograms that could come from either reaction by-product or perhaps some unlabeled aromatic amino acids (e.g. Tryptophan) are absorbing at the chosen detection wavelength.

2.3.3.4 Inter-day Precision, Intra-day Precision, and Method Reproducibility

Precision of the method was determined by the analysis of replicate QC samples and the results are presented in Table 2-7. Results are all expressed as %RSD values determined using the number of replicates given on the table. All the given results are below 15% and most fall below 10% which indicates that the method is satisfactory.

Table 2-7 Intra-day, inter-day and method reproducibility of the analytical method.

Compound	Intra-day %RSD (*n*=6)	Inter-day %RSD (*n*=5)	Method %RSD (*n*=16)
Dansyl asparagine	2.7	4.3	7.6
Dansyl aspartate	3.6	7.5	7.7
Dansyl cystine	2.2	5.2	8.1
Dansyl glutamine	5.2	7.7	3.5
Dansyl histidine	2.9	3.2	7.5
Dansyl leucine	4.4	9.6	9.2
Dansyl methionine	4.5	5.3	3.5
Dansyl phenylalanine	2.9	9.3	7.8
Dansyl proline	7.6	3.8	3.8
Dansyl tryptophan	8	12.2	8.2
Dansyl threonine	8.2	6.3	8.9
Dansyl tyrosine	2.5	3.4	4.4
Dansyl valine	1.6	3	1.3

2.3.3.5 LOD and LLOQ

The limit of detection (LOD) and lower limit of quantitation (LLOQ) were determined by calculating the standard deviation of the blank from the y error and the slope. The LOD values were between 0.06-0.09 mM and the LLOQ values between 0.21 and 0.30 mM. These values correspond well with similar analysis in which the determine LOD values were for Dns-Asp, Dns-Thr, and Dns-Gly were 0.009, 0.002, and 0.001 mM respectively.[133] It should be noted that the authors used a 20 μL injection volume compared with the 5 μL injection volume used in this analysis. This is most likely the reason for the difference between the values of the two methods. Many other methods employ fluorometric detection if dansyl labeled compounds, in these cases LOD of 1.5 pmol/injection are reported.[46] In our analysis the LOD is 300-400 pmol/injection which is 2 orders of magnitude larger but this acceptable due to the inherent sensitivity difference in UV and fluorescence detection. The sensitivity of the analysis with MS used as a detector would be much better than using either UV or fluorescence detection. Another study of dansylated biogenic amines uses a method much like the one developed here. There results show detection limits that are an order of magnitude lower than those found using our method.[45] There is no mention of injection volume however and it would be possible that our method would be more sensitive at increased injection volumes. This information along with the relative error of measured replicates for the LLOQ, %RSD and linear dynamic range of the method for the measured metabolites can be found in Table 2-7. The values of %RE range from -6.52 – 13.23 and demonstrate that the method is valid at the LLOQ. Values measured during the evaluation of reaction conditions were completed using a much higher concentration than the LLOQ and all values fell within the linear dynamic range determined with these experiments.

Table 2-8 LLOD and LLOQ of internal standard and standards used for external calibration.

Compound	LLOQ (mM)	Measured Concentration (mM)	%RE	%RSD	LOD (mM)	Linear Dynamic Range (mM)
Dns-Aspartic Acid	0.21	0.22	6.12	4.11	0.06	0.21-2.0
Dns- Threonine	0.28	0.26	-6.52	1.89	0.08	0.28-2.0
Dns-Glycine	0.30	0.34	13.23	3.45	0.09	0.30-2.0

2.4 Conclusions

This chapter describes the development and utilization of a UPLC-UV method for the relative and absolute quantification of dansyl labelled metabolite standard compounds. The UPLC method was demonstrated to be sensitive and provide a linear response. As well the method proved reliable in terms of accuracy, reproducibility, and selectivity. This derivatization chemistry is key to our group's research in metabolomics profiling, relative, and absolute quantification metabolomic studies. Due to the importance of this step in the overall analysis, the evaluation of reaction conditions involved is very important. Evaluation of conventionally heated reactions demonstrates that ambient temperatures are preferable to higher temperatures. Microwave assisted heating was explored as an alternative to that of conventional heating. Results from this study demonstrate that using short microwave reaction times for metabolite labelling produced samples with a higher comparative derivatized metabolite concentration. In addition to this, both an organic reactor style microwave and a consumer microwave were tested and compared with conventional heating and both performed better than conventional heating. With the hope that dansyl isotopic labelling will be used at some point in clinical methods, the use of microwave assisted heating could provide a short and reliable results for samples preparation.

2.5 Literature Cited

1. Goodacre, R.; Vaidyanathan, S.; Dunn, W. B.; Harrigan, G. G.; Kell, D. B., *Trends in Biotechnology* **2004,** *22* (5), 245-252.

2. Fiehn, O., *Plant Mol Biol* **2002,** *48* (1-2), 155-171.

3. Bino, R. J.; Hall, R. D.; Fiehn, O.; Kopka, J.; Saito, K.; Draper, J.; Nikolau, B. J.; Mendes, P.; Roessner-Tunali, U.; Beale, M. H.; Trethewey, R. N.; Lange, B. M.; Wurtele, E. S.; Sumner, L. W., *Trends in Plant Science* **2004,** *9* (9), 418-425.

4. Dunn, W. B.; Ellis, D. I., *TrAC Trends in Analytical Chemistry* **2005,** *24* (4), 285-294.

5. Mungur, R.; Glass, A. D.; Goodenow, D. B.; Lightfoot, D. A., *Journal of biomedicine & biotechnology* **2005,** *2005* (2), 198-214.

6. Brindle, J. T.; Antti, H.; Holmes, E.; Tranter, G.; Nicholson, J. K.; Bethell, H. W.; Clarke, S.; Schofield, P. M.; McKilligin, E.; Mosedale, D. E.; Grainger, D. J., *Nature medicine* **2002,** *8* (12), 1439-44.

7. Chamberlin, B. A.; Sweeley, C. C., *Clinical chemistry* **1987,** *33* (4), 572-6.

8. Dettmer, K.; Aronov, P. A.; Hammock, B. D., *Mass spectrometry reviews* **2007,** *26* (1), 51-78.

9. Villas-Boas, S. G.; Mas, S.; Akesson, M.; Smedsgaard, J.; Nielsen, J., *Mass spectrometry reviews* **2005,** *24* (5), 613-46.

10. Price, N. P. J.; Firmin, J. L.; Gray, D. O., *Journal of Chromatography A* **1992,** *598* (1), 51-57.

11. Hayman, A. R.; Gray, D. O., *Phytochemistry* **1989,** *28* (2), 673-675.

12. Santa, T., *Biomedical chromatography : BMC* **2011,** *25* (1-2), 1-10.

13. Shortreed, M. R.; Lamos, S. M.; Frey, B. L.; Phillips, M. F.; Patel, M.; Belshaw, P. J.; Smith, L. M., *Analytical Chemistry* **2006,** *78* (18), 6398-6403.

14. Stanislaus, A.; Guo, K.; Li, L., *Analytica chimica acta* **2012,** *750*, 161-72.

15. Guo, K.; Li, L., *Analytical Chemistry* **2009,** *81* (10), 3919-3932.

16. Guo, K.; Li, L., *Analytical Chemistry* **2010,** *82* (21), 8789-8793.

17. Guo, K.; Peng, J.; Zhou, R.; Li, L., *Journal of chromatography. A* **2011,** *1218* (23), 3689-94.

18. Kappe, C. O., *Angewandte Chemie (International ed. in English)* **2004,** *43* (46), 6250-84.

19. Stadler, A.; Pichler, S.; Horeis, G.; Kappe, C. O., *Tetrahedron* **2002,** *58* (16), 3177-3183.

20. Schanche, J.-S., *Mol Divers* **2003,** *7* (2-4), 291-298.

21. Cardenes, L.; Ayala, J. H.; Gonzalez, V.; Afonso, A. M., *Journal of chromatography. A* **2002,** *946* (1-2), 133-40.

22. Damm, M.; Rechberger, G.; Kollroser, M.; Kappe, C. O., *Journal of chromatography. A* **2009,** *1216* (31), 5875-81.

23. Deng, C.; Yin, X.; Zhang, L.; Zhang, X., *Rapid communications in mass spectrometry : RCM* **2005,** *19* (16), 2227-34.

24. Meyer, H.; Liebeke, M.; Lalk, M., *Analytical Biochemistry* **2010,** *401* (2), 250-259.

25. Zuo, Y.; Zhang, K.; Lin, Y., *Journal of Chromatography A* **2007,** *1148* (2), 211-218.

26. Mingos, D. M. P.; Baghurst, D. R., *Chemical Society Reviews* **1991,** *20* (1), 1-47.

27. Stuerga, D.; Delmotte, M., Wave–Material Interactions, Microwave Technology and Equipment. In *Microwaves in Organic Synthesis*, Wiley-VCH Verlag GmbH & Co. KGaA: 2004; pp 1-33.

28. Gabriel, C.; Gabriel, S.; H. Grant, E.; H. Grant, E.; S. J. Halstead, B.; Michael P. Mingos, D., *Chemical Society Reviews* **1998,** *27* (3), 213-224.

29. Gedye, R. N.; Smith, F. E.; Westaway, K. C., *Canadian Journal of Chemistry* **1988,** *66* (1), 17-26.

30. de la Hoz, A.; Diaz-Ortiz, A.; Moreno, A., *Chem Soc Rev* **2005,** *34* (2), 164-78.

31. Perreux, L.; Loupy, A., *Tetrahedron* **2001,** *57* (45), 9199-9223.

32. Baghbanzadeh, M.; Škapin, S. D.; Orel, Z. C.; Kappe, C. O., *Chemistry – A European Journal* **2012,** *18* (18), 5724-5731.

33. Kappe, C. O.; Pieber, B.; Dallinger, D., *Angewandte Chemie International Edition* **2013,** *52* (4), 1088-1094.

34. Vial, J.; Jardy, A., *Analytical Chemistry* **1999,** *71* (14), 2672-2677.

35. Neue, U. D.; Mazzeo, J. R., *Journal of Separation Science* **2001,** *24* (12), 921-929.

36. Dolan, J. W., *Journal of Chromatography A* **2002,** *965* (1–2), 195-205.

37. Snyder, L. R.; Carr, P. W.; Rutan, S. C., *Journal of Chromatography A* **1993,** *656* (1–2), 537-547.

38. Vitha, M.; Carr, P. W., *Journal of Chromatography A* **2006,** *1126* (1–2), 143-194.

39. Lloyd R. Snyder, J. J. K., John W. Dolan, *Introduction to Modern Liquid Chromatography*. Wiley: 2009; p 960.

40. Levina, N. B.; Nazimov, I. V., *Journal of Chromatography A* **1984,** *286* (0), 207-216.

41. Tapuhi, Y.; Schmidt, D. E.; Lindner, W.; Karger, B. L., *Analytical Biochemistry* **1981,** *115* (1), 123-129.

42. Wiedmeier, V. T.; Porterfield, S. P.; Hendrich, C. E., *Journal of Chromatography B: Biomedical Sciences and Applications* **1982,** *231* (2), 410-417.

43. Parris, N.; Gallelli, D., *Journal of Liquid Chromatography* **1984,** *7* (5), 917-924.

44. Caddick, S., *Tetrahedron* **1995,** *51* (38), 10403-10432.

45. Linares, R. M.; Ayala, J. H.; Afonso, A. M.; González Díaz, V., *Journal of Chromatography A* **1998,** *808* (1–2), 87-93.

46. Kang, X.; Xiao, J.; Huang, X.; Gu, Z., *Clinica Chimica Acta* **2006,** *366* (1–2), 352-356.

47. Fürst, P.; Pollack, L.; Graser, T. A.; Godel, H.; Stehle, P., *Journal of Chromatography A* **1990,** *499* (0), 557-569.

Chapter 3 : Evaluation of Reaction Conditions for Differential Isotope Labeling for Metabolomic Profiling of Biological Samples by Liquid Chromatography Mass Spectrometry

3.1 Introduction

Metabolomics is a field that endeavors to characterize small molecules and their interactions in biological systems for both understanding and clinical application. This class of molecules encompasses primary and secondary endogenous metabolites, small peptides, dietary, pharmaceutical, or molecules from environmental sources. For this type of study there are two techniques, nuclear magnetic resonance (NMR) and mass spectrometry (MS), which are used to accurately and precisely identify and quantify metabolites in a wide range of biological sample types.[1-2] While these two techniques are powerful, the complexity of biological samples makes their study challenging. In order to reduce the complexity of samples these detection/identification techniques are coupled with a method of separation. Gas chromatography, liquid chromatography, and capillary electrophoresis are the best and traditionally most used techniques in the field.[3-6] Although other methods are used, liquid chromatography mass spectrometry (LC-MS) has become the tool of choice for both targeted and profiling metabolomics. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) has very high mass resolution and good sensitivity; therefore it is suitable for metabolomics studies.[7-8] By using this instrument accurate mass for each detected metabolite is determined, this can lead to putative and verified identification. [9-10]

Many metabolites exist in very low concentrations and are obtained in limited amounts. In order to meet this challenge many have applied an isotopic labeling approach, first in the field of proteomics and recently in metabolomics.[11-14] Isotopic

labeling is useful in LC-MS analysis for three reasons, the first is that tags can be designed to label specific functional groups thereby reducing detection and identification complexity. The second is that the separation of metabolites is improved, especially those that are hydrophilic. A third reason is that the chemical tag is more active during electrospray ionization (ESI) due to enhanced chargeability, increased hydrophobicity, and the fact that most analytes will elute in higher percentage of organic solvent.[15] To detect these compounds our group has developed and applied a number of differential isotope labeling methods such as danyslation with which these compounds can be more effectively separated, sensitively detected, quickly identified, and accurately quantified.[16-19]

The number of metabolites detected by MS will depend on the numbers that are labeled. There have been a number of studies on the optimization and effectiveness of dansyl labeling for various applications.[20-22] As profiling metabolomics looks to ascertain the greatest amount of information on the metabolite composition of a sample reaction conditions gleaned from these studies may not prove ideal. The goal of this work was to explore the effect of changing reaction conditions, such as temperature, time and heating method on the results of metabolic profiling experiments of biological samples.

3.2 Experimental

3.2.1 Chemicals and Reagents

All chemicals and reagents were purchased from Sigma-Aldrich Canada (Markham, ON. Canada) except those otherwise noted. LC-MS grade water and acetonitrile (ACN) were purchased from Thermo Fisher Scientific (Edmonton, AB, Canada).

3.2.2 Sample Collection and Processing

3.2.2.1 Urine

Human urine was collected from 3 healthy volunteers who were not correlated based on diet or any other classification. Informed consent was given by the volunteers and ethics approval was obtained from University of Alberta in compliance with the University Health Information Act. Volunteers had ages ranging from 24-35 years in age. Urine was collected as a second morning void sample with midstream collection. It was ensured that volunteers had fasted for at least 8 hours before sample collection. Immediately after collection samples were centrifuged, filtered with a 0.22 μM filter, aliquoted and stored at -80°C until required for analysis.

3.2.2.2 Plasma

Human plasma was collected from 3 healthy volunteers who were not correlated based on diet or any other classification. Informed consent was given by the volunteers and ethics approval was obtained from University of Alberta in compliance with the University Health Information Act. Volunteers were between the ages of 24-35 years in age. Blood was collected by Dynalife DX diagnostic laboratory services (Edmonton, AB) in the morning after volunteers had fasted for at least 8 hours. Blood was immediately centrifuged in order to separate plasma from the cellular portion of the blood. Plasma was aliquoted and stored at -80°C until further processing. Before derivatization 90 μL of methanol was added to 30 μL of plasma in order to precipitate proteins. After mixing, sample was centrifuged at 15000 rpm for 15 minutes. Supernatant equal to approximately 90 μL was vacuum dried until no solvent remained. The sample was then reconstituted in 50 μL of 1:1 H_2O: ACN and stored on ice before reaction with DnsCl.

3.2.3 Derivatization Reactions

3.2.3.1 Conventional Heating

The dansylation procedure follows that previously reported by Guo *et al.*[15] Briefly the procedure is as follows, for metabolite labelling, 50 µL of processed urine, plasma, or standard was mixed with 25 µL $Na_2CO_3/NaHCO_3$ (pH 9.4) buffer and reacted with 75 µL of ^{12}C-dansyl chloride (DnsCl) in ACN (12 mg/mL or 44.5 mM). The reaction mixture was then heated in a conventional oven or left at room temperature (~22°C). The conventional oven reactions were carried out at ambient temperature (AT) (varied between 21 and 24°C), 40°C, 60°C, or 80°C. 10 µL of sodium hydroxide (NaOH at 250 mM) was added to each reaction in order to consume the excess DnsCl after 60 minutes. After quenching with NaOH, samples were mixed and left for 10 minutes at room temperature after which 50 µL of 425 mM formic acid (FA) in H_2O: ACN (1:1) was added to acidify the samples.

3.2.3.2 Microwave Assisted Heating

Microwave samples were prepared in the same way as described in the conventional labelling method. Microwave-assisted heating (MAH) was applied using either a CEM Discover microwave reactor (CEM, Matthews, NC). In order to assess the optimal MAH conditions two different types of reactions were carried out. The first is known as power control (PC), during which the samples were subjected to a constant microwave power measured in wattage (W) over specified time periods or time temperature control (TTC) during which power is applied only to maintain a temperature specified and measured by an infrared thermometer. The PC experiments were carried out at 100 w and 200 in the CEM Discover microwave and 120 W, 480 W, 960 W, and 1200 W in the Panasonic microwave for 30 s, 60 s, 120 s, and 300 s. TTC samples were

reacted at 40°C and 60°C for 60 minutes. After heating the protocol followed that of the conventionally heated samples.

3.2.4 UPLC-UV

Chromatography analysis for total labeled metabolite determination was performed according to a protocol previously developed in our group.[17, 23] Briefly the method is described here: using a Waters Acquity UPLC with Photo Diode Array (PDA) UV detector system (Waters Canada, Mississauga, ON). The column used was a Phenomenex (Torrance, CA) Kinetix 1.7 μM C18 core shell particles 2.1 x 50 mm with a 100 Å pore size. Elution conditions were as follows: isocratic elution from for 1 minute, 0 -95 %B mobile phase over 0.1 minutes, an isocratic hold at 95 %B for 1.49 minutes, 95-0 %B over 0.5 minutes, followed by an isocratic hold for 3 minutes. Mobile phase A consists of 5% ACN in H_2O with 0.1% FA and mobile phase B consists of 5% H_2O in 0.1% FA. Column temperature was maintained at 35°C and the sample manager was maintained at 4°C. The flow rate was set at 450 μL/min and 5 μL of sample was injected. Before injection samples both plasma and urine samples were diluted by half with 1:1 H_2O: ACN. Labelled metabolites were detected by UV absorbance using both 3D detection from 210-400 nm and 2D detection at 338 nm at a sampling rate of 20 points/s. The data was acquired and processed using Empower® Version 2 (Waters Canada, Mississauga, ON).

3.2.5 LC-FTICR

An Agilent 1100 series HPLC system (Agilent, Palo Alto, CA) and an Agilent Zorbax Eclipse C18 column (2.1 mm × 100 mm, 1.8 μm particle size, 100 Å pore size) were used for online LC-MS. Mobile phase A consists of 5% ACN in H_2O with 0.1% FA and mobile phase B consists of 5% H_2O in 0.1% FA. Elution conditions were as follows: 20-35% B for 3.5 minutes, 35-65% B for 12.5 minutes, 65-95% B for 3 minutes, 95-99%

B for 3 minute, and an isocratic hold at 99% B for 4 minutes. All gradients were linear. The flow rate was 180 µL/min, and the sample injection volume was 2 µL. The flow from HPLC was split 1:3 and a 60 µL/min flow was directed to the ESI source of a Bruker 9.4 Tesla Apex-Qe Fourier-transform ion-cyclotron resonance (FT-ICR) mass spectrometer (Bruker, Billerica, MA), while the rest of the flow was delivered to waste. All MS spectra were obtained in the positive ion mode.

3.2.6 Method Validation

3.2.6.1 Selectivity and Carryover

Selectivity and carryover were determined by doing the following experiments: instrument blank (ACN injection, 5 µL), solvent blank (50:50 H_2O: ACN w/ 0.1 % FA), reaction blank (reaction mixture with H_2O in place of amino acid mixture or a biological sample), and an amino acid standard (derivatized amino acid in H_2O: ACN w/ 0.1% FA). All were analyzed using the method described above. Carryover was determined by analysis of blank injection following three subsequent injections of a labeled biological sample.

3.2.6.2 Linearity, Interday/Intraday Precision, and Method Reproducibility

Linearity and precision for the UPLC-UV quantification method were determined by producing calibration plots from diluted labeled biological samples. Precision and reproducibility were determined by examining replicate samples during the same day, 2 weeks apart, and by analysis of three replicate injections of the same sample.

Linearity of and reaction reproducibility of the LC-FTICR method were determined by comparing reactions by varying the dilution of the biological samples. The relative response ratio to the internal standard was then plotted for each of the analytes. Reproducibility was determined by comparing three replicate samples and comparing

peak pair counts and selected peak intensity. Intra-day precision was determined by analyzing the same QC sample three times in one day. Inter-day precision was evaluated by comparing the same QC sample which was analyzed before analysis on each separate analysis day. There was a total of 3 analysis days included in this determination. Method reproducibility was determined by running 3 replicated QC samples sequentially. To determine the differences in total peak pair number and selected peak intensities were compared to determine the %RSD.

3.2.7 Data Processing and Statistical Analysis

Raw LC-MS data files were converted to comma separated value (.csv) files using a script designed for use in DataAnalysis (Bruker Daltonics). IsoMS, a program created within our research group, was used to analyze the .csv files in order to filter the mass values that result in $^{12}C/^{13}C$ dansyl labeled ion pairs or putative labeled metabolite molecules. In addition to determining the mass and retention the program produces an intensity difference between the ion pair. The processed IsoMS files were then aligned by retention time using another in house program. The output of this program gives putative metabolite hits, their possible Human Metabolome Database (www.hmdb.ca) matches, and relative intensity ratios of the $^{12}C/^{13}C$ dansyl labeled ion pairs. For statistical analysis the web tool MetaboAnalyst 2.0 was used.[24-25]

3.3 Results and Discussion

3.3.1 Assessment of Derivatization Conditions for Dansyl Labelling of Biofluids

There derivatization of metabolites using DnsCl is an important step in our groups overall approach to the qualitative and quantitative approach to the analysis of the metabolome. This reaction has historically been carried out at increased temperatures or for long periods of time with temperature varying from ambient to as high as 100°C.[20, 26-

[27] There have been a number of studies that assess the effect that dansyl reaction temperature has an effect on the labeling efficiency of different molecules.[28] Considering the derivatization effectiveness is important in our analysis due to the diversity characteristic of biological samples. An important consideration for our strategy is the identification of as many metabolites as possible within a biological sample, for this we must consider labeling a wide diversity of molecules found in nature. In addition to this it is important for the reaction to be specific to primary amines, secondary amines, and phenolic compounds. This allows for the reduction of complexity when deducing the identity of putative metabolites from m/z values alone. It should be noted that a putative identification is often used to define a match with one or more measured characteristics (often high resolution m/z and isotopic pattern) but does not compare these with a separate reference standard. The other type is definitive identification in which reference standard are used to compare measured characteristics (often retention time and m/z).[29] As this study involves large data sets with a lot of peak pairs determined it would not be feasible to conduct any large scale definitive identification.

3.3.1.1 UPLC UV Analysis for Total Metabolite Determination

Biofluid samples such as urine and plasma often contain many metabolites that are very polar, this leads to decreased retention times in reverse phase liquid chromatography (RPLC) that often results in a large peak of polar metabolites eluting at the beginning of a gradient analysis that is often complicated to detect and analyze by MS. By derivatizing these compounds with DnsCl is causes their retention time to change to much later time because of the hydrophobicity of the dansyl moiety. In addition the ESI response is improved greatly due increased surface activity and chargeability. This labeling is also beneficial as the dansyl moiety absorbs well in the UV region.

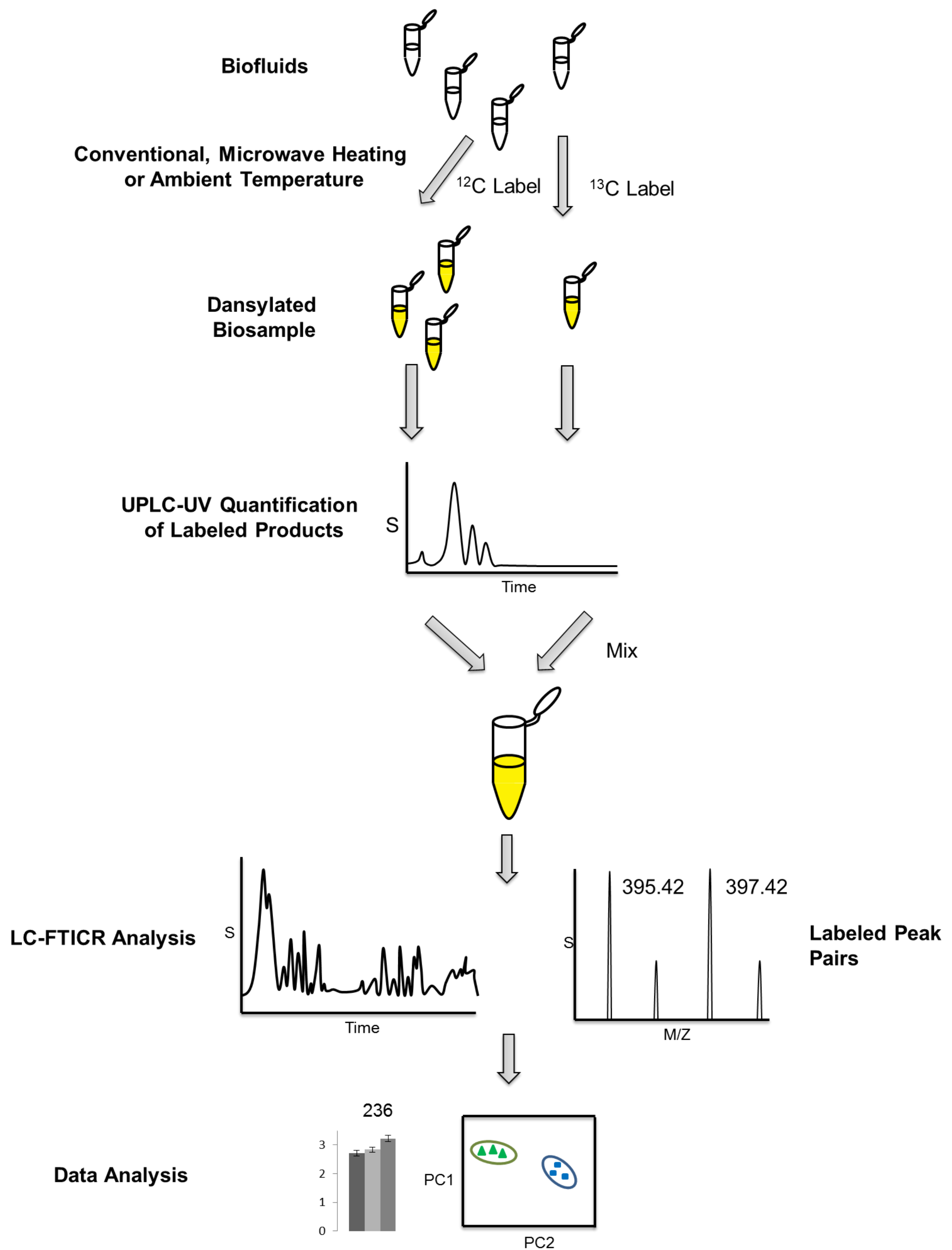


Figure 3-1 Workflow for LC-MS metabolomic analysis of biofluids using dansyl chloride for differential labeling of biological samples.

Previously developed protocols were used for this study in which the amount of dansyl labeled metabolite present in a sample is quantified using UPLC-UV.[16]

The UPLC-UV method relies on the fact that the sample should contain a large amount of labeled metabolites in comparison with other compounds. To analyze these compounds the UV absorption at a wavelength specific to the Dansyl group is used. The characteristic absorption spectra of a number of dansyl compounds are shown in Figure 2-1 A. It can be seen that there are peak maximum at ~220 nm, ~254 nm, and ~338 nm all of which correspond to the $\pi \rightarrow \pi^*$ transitions characteristic of aromatic groups.[30] The longest wavelength was used because it is specific to the dansyl group, it has lower absorptivity than the other two wavelengths. This was chosen to avoid saturation as relatively few other functional groups absorb at this wavelength. Although the maximum absorbance is only approximate for dansyl labeled compounds this wavelength is still used because it averages the higher and lower values.

To assess the amounts of labeled compounds in a given sample with UV detection it is important to be able to integrate the peak area of the chromatographic peak belonging to the metabolites. To achieve a single peak a step gradient was used. By stepping the gradient from 0-95%B the majority of labeled analytes are quickly eluted and detected in one large peak. This peak can be seen in Figure 3-1 A for urine and B for the plasma sample at a retention time between 1.5 and 2 minutes. The earlier eluting peak in these chromatograms belongs to dansyl hydroxide (Dns-OH), which due to increased polar character elutes much quicker than other labeled components of the mixture. In Figure 3-1 there is a large intensity difference (~ 6 fold) between the urine and plasma samples, this is assumed to be due a decreased metabolite total concentration in plasma compared with urine. This is in correlation with Figure 3-10 in which the number of

putative metabolite matches in plasma is much less than in urine. In Figure 3-1 C and D are calibration curves prepared by diluting a pooled and labeled urine or plasma sample.

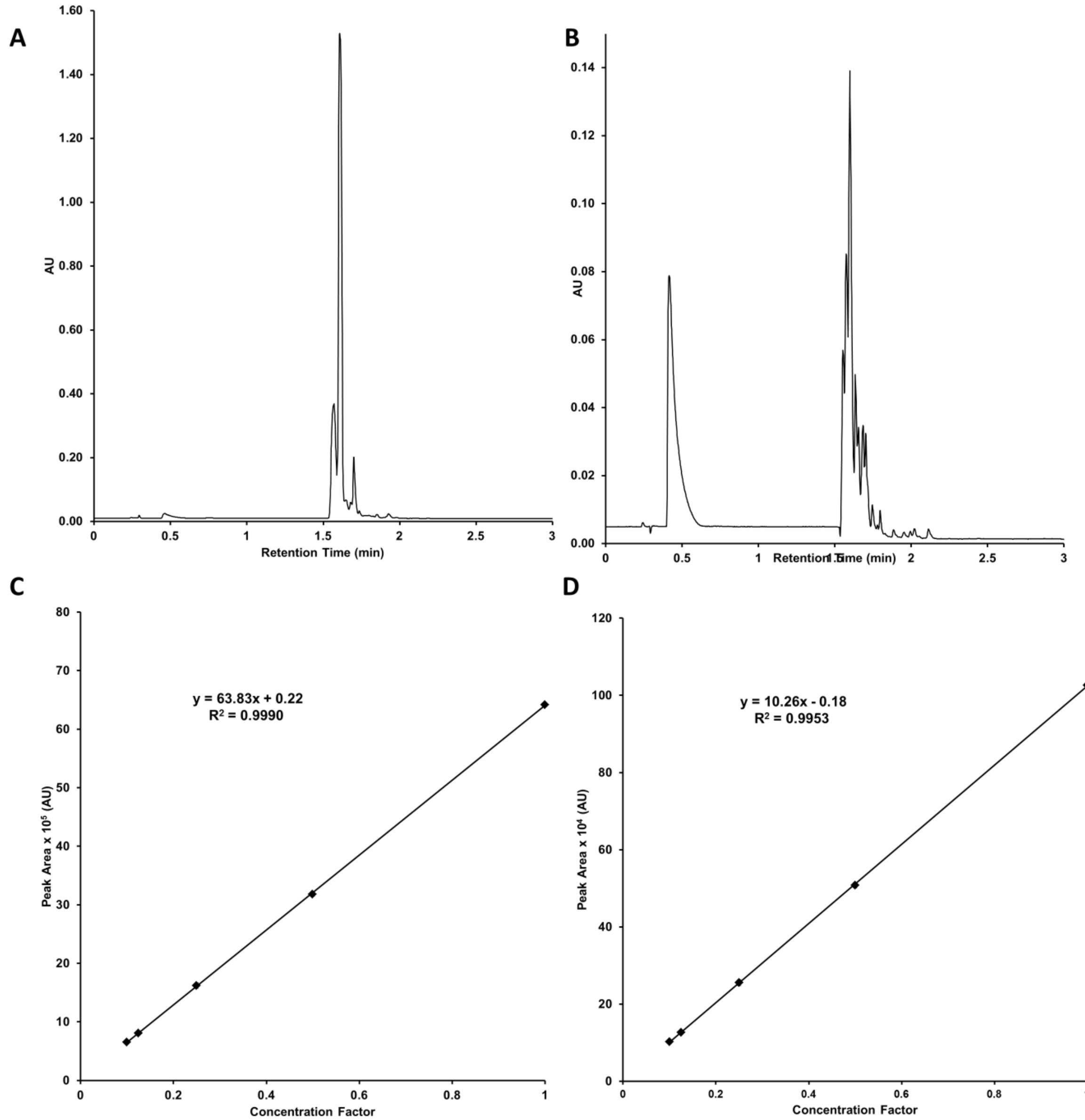


Figure 3-2 Representative UPLC-UV spectrum for a ^{12}C dansyl labeled urine sample (A) and a labeled plasma sample (B). Calibration curves for urine (C) and plasma (D).

Dilution simulates a reduced concentration from the native absorbance detected and demonstrates the linearity of the quantification method. A plot of labeled amino acid standard was prepared with a total metabolite concentration of 28 mM but the slope is different from both the urine and plasma calibration curves and therefore cannot be used as a quantitative measurement for this study. That being stated, the urine and plasma calibration curves have much different slopes indicating that these samples have very different metabolite profiles. This is expected as the concentrations of metabolites and also the variation within these two biofluids is very different.[31]

Following derivatization each individual ^{12}C or ^{13}C biofluid sample is assessed using this method. The peak area obtained is considered indicative of the total metabolite concentration within each sample. The method used in this thesis involves mixing isotopically separate samples. Subsequent relative quantification of the isotopic pair is carried out. Therefore it is important when mixing samples that the overall concentration of labeled samples is as close as possible. In order to achieve this we compare the peak area of each individual sample. The peak area obtained using the step gradient analysis is also useful for comparison of varying reaction conditions. In an initial study of a variety or dansyl derivatization conditions the effect of temperature and reaction method were assessed. Conventional heating was used for reactions at 40, 60, and 85°C, another replicate was reacted at ambient temperature (23°C measured) and the final assessed condition was using a CEM Discover microwave reactor to heat a sample to 40°C. All of the reactions were carried out for a period of 60 minutes. The microwave reaction differs from those carried out in Chapter 2. Instead of applying a steady microwave radiation to the sample, the power was varied to maintain a temperature that was measured using a reaction vessel with an infrared sensor inserted. Short microwave reactions were also carried out and will be discussed in further detail in Section 3.3. Figure 3-2 A shows the

quantification results of each individual reaction tube. Each sample was injected in triplicate for a total of 12 injections for each reaction type and 60 overall injections for the experiment. The orange bars are the samples that had ^{13}C DnsCl added; the black bars represent the ^{12}C replicates. This quantification step is then used to determine the volumes required for mixing a sample with a 1:1 ratio of ^{12}C/^{13}C labeled metabolites. It was decided that any sample with a greater than 5% difference between the ^{12}C/^{13}C samples would have its mixing volume altered. The sample list with UV peak area, %difference, and subsequent mixing volume is given in Table 3-1. As can be seen from the after mixing section there is a good correlation between the samples after mixing with all %RSD values below 6%. The amount of sample after mixing can be an indicator of the effectiveness of the labeling reaction. From Figure 3-2 B it can be seen that the largest peak areas results and therefore labeled metabolite amounts result from the 85°C reaction and ambient temperature reaction. This is the trend that was observed with conventional reactions of simpler metabolite mixtures in Chapter 2. The reaction using MAH is comparable to the reaction at 60°C. This indicates that it may be a more effective heating method than conventional heating. Although not as noticeable, the same trend can be seen in Figure 3-2 A for samples before mixing.

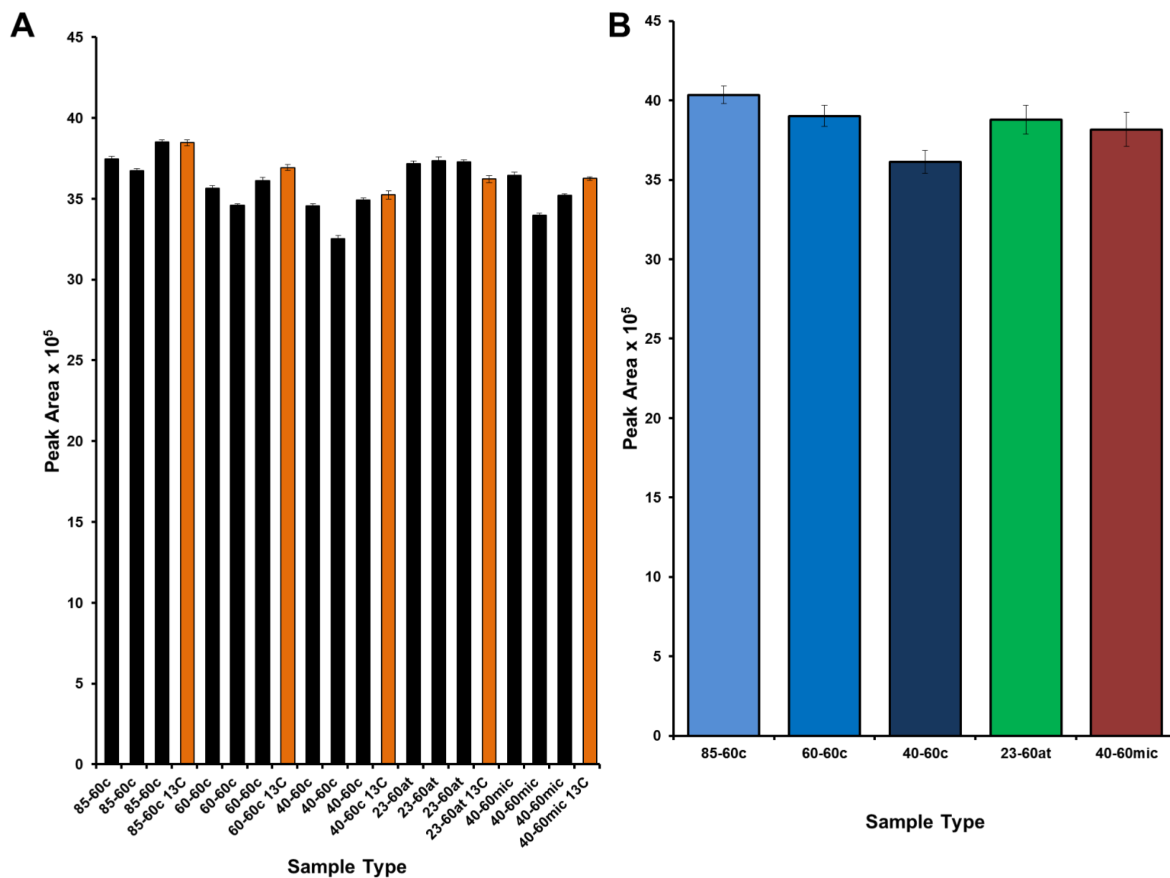


Figure 3-3 UPLC-UV total peak area of dansyl metabolites for urine samples before $^{12}C/^{13}C$ mixing (A) and after mixing (B). Reaction conditions used for each sample are given as time-temperature-reaction method. (c = conventional oven, at = ambient temperature, mic = microwave

Table 3-1 Peak area results before and after mixing of labeled urine samples with different reaction conditions. Volumes given are prior to the mixing of aliquots for LC-MS analysis.

		Before Mixing	
Reaction Type	**Average Peak Area**	**^{12}C %Difference from ^{13}C**	**Mixing Volume (µL of $^{12}C/^{13}C$)**
85-60c	3746485	-2.60	25/25
85-60c	3675654	-4.44	25/25
85-60c	3852636	0.16	25/25
85-60c ^{13}C	3846592		
60-60c	3564891	-3.46	25/25
60-60c	3458951	-6.33	23.5/26.5
60-60c	3613216	-2.15	25/25
60-60c ^{13}C	3692613		
40-60c	3456656	-1.90	25/25
40-60c	3251648	-7.72	23/27
40-60c	3491352	-0.92	25/25
40-60c ^{13}C	3523698		
23-60at	3715897	2.60	25/25
23-60at	3735432	3.14	25/25
23-60at	3726931	2.90	25/25
23-60at ^{13}C	3621866		
40-60mic	3646666	0.55	25/25
40-60mic	3399820	-6.26	23.5/26.5
40-60mic	3521987	-2.89	25/25
40-60mic ^{13}C	3626699		
		After Mixing	
Reaction Type	**Average Peak Area**	**%RSD**	
85-60c	4034481	2.76	
60-60c	3901568	3.41	
40-60c	3611843	4.02	
23-60at	3878185	4.65	
40-60mic	3816987	5.56	

3.3.1.2 LC-FTICR-MS Analysis of Variable Reaction Conditions

Determination of the total metabolite amount in a biological sample is interesting but it is not the end goal of our group's metabolite profiling and quantification analysis. To determine the identity and amounts of metabolites in a sample our group uses high-performance liquid chromatography (HPLC) for separation and Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) for detection. Because metabolites are labeled with a hydrophobic dansyl moiety, their retention is often increased when compared with the native molecule. This allows for the analysis of otherwise highly polar molecules that would often have early elution time, as can be seen in the spacing of eluting peaks in of a chromatogram (Figure 3-4). Each metabolite labeled with $^{12}C/^{13}C$ DnsCl will have coeluting 'peak pair' that is detectable using high resolution mass spectrometry. Other approaches use only separation and m/z detection for the analysis of biological samples, our group's approach has an extra level of depth due to the coeluting peak pair.[32-34] An in-house designed software program known as IsoMS is used to identify and compile a peak pair list from raw retention and m/z data exported from DataAnalysis (Bruker) software. In addition to aiding in identification of compounds by mass, the relative intensity of the $^{12}C/^{13}C$ labeled peaks is used for a semiquantitative approach to metabolomics analysis and classification of samples. In this initial analysis the comparison of the number of peak pairs determined was used to quantify the effectiveness or variability imposed by different Dns derivatization conditions. The list of peaks pairs corresponding to unique metabolites was aligned using a Java script software designed in our group. This alignment allows for the comparison of putative metabolites between the separate reaction conditions and the replicates of each analysis. Also assessed were chromatography profiles (one and two dimensional) for any differences found between the labelling methods. This can be challenging as the chromatograms are

complicated and the data is not specific. Comparisons were made using total ion counts, which reveals little about the differentially labeled metabolites contained in the sample.

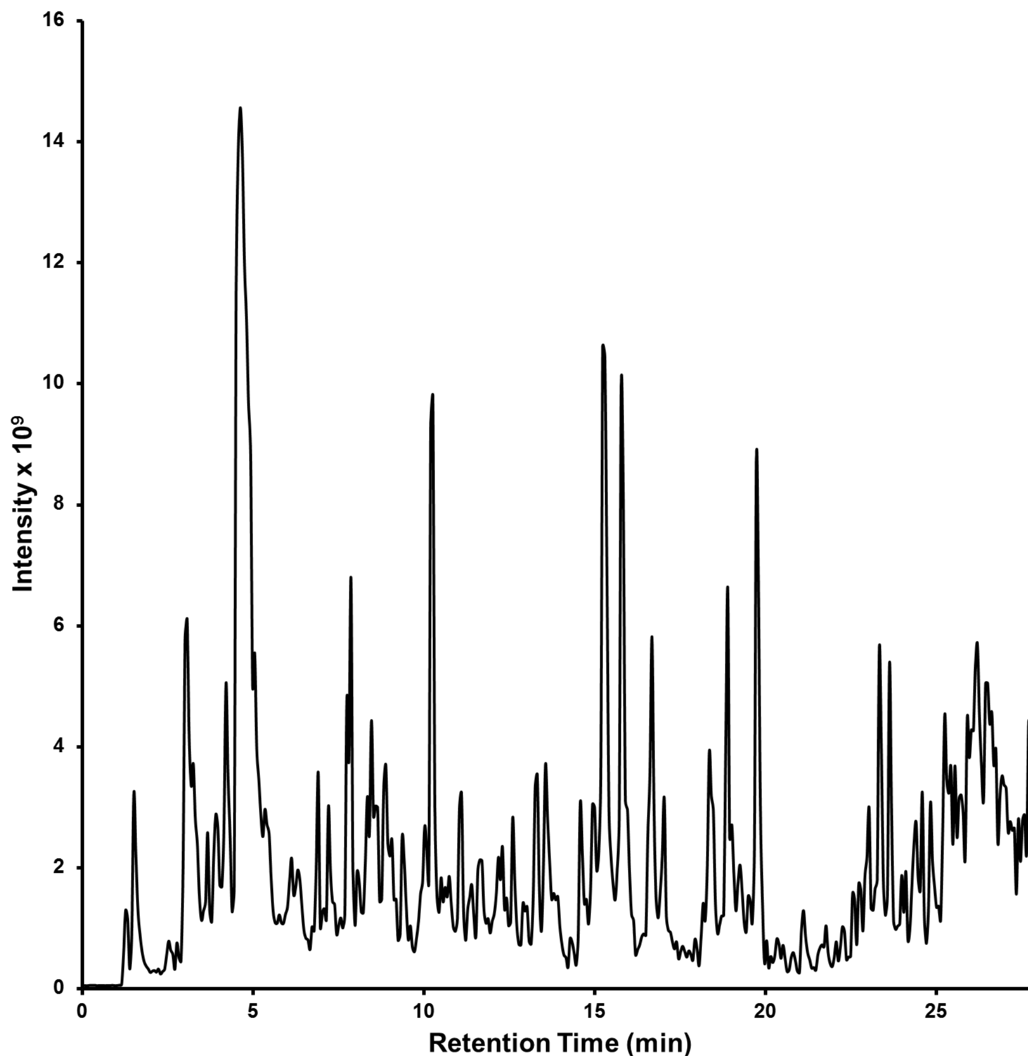


Figure 3-4 Total ion chromatogram from LC-FTICR analysis of dansyl labeled urine. Agilent Zorbax Eclipse C18 column (2.1 mm × 100 mm, 1.8 μm particle size. Mobile phase A consists of 5% ACN in H2O with 0.1% FA mobile phase B consists of 5% H2O in 0.1% FA. Elution conditions were: 20-35% B for 3.5 minutes, 35-65% B for 12.5 minutes, 65-95% B for 3 minutes, 95-99% B for 3 minute, and an isocratic hold at 99% B for 4 minutes. All gradients were linear. Flow rate was 180 μL/min.

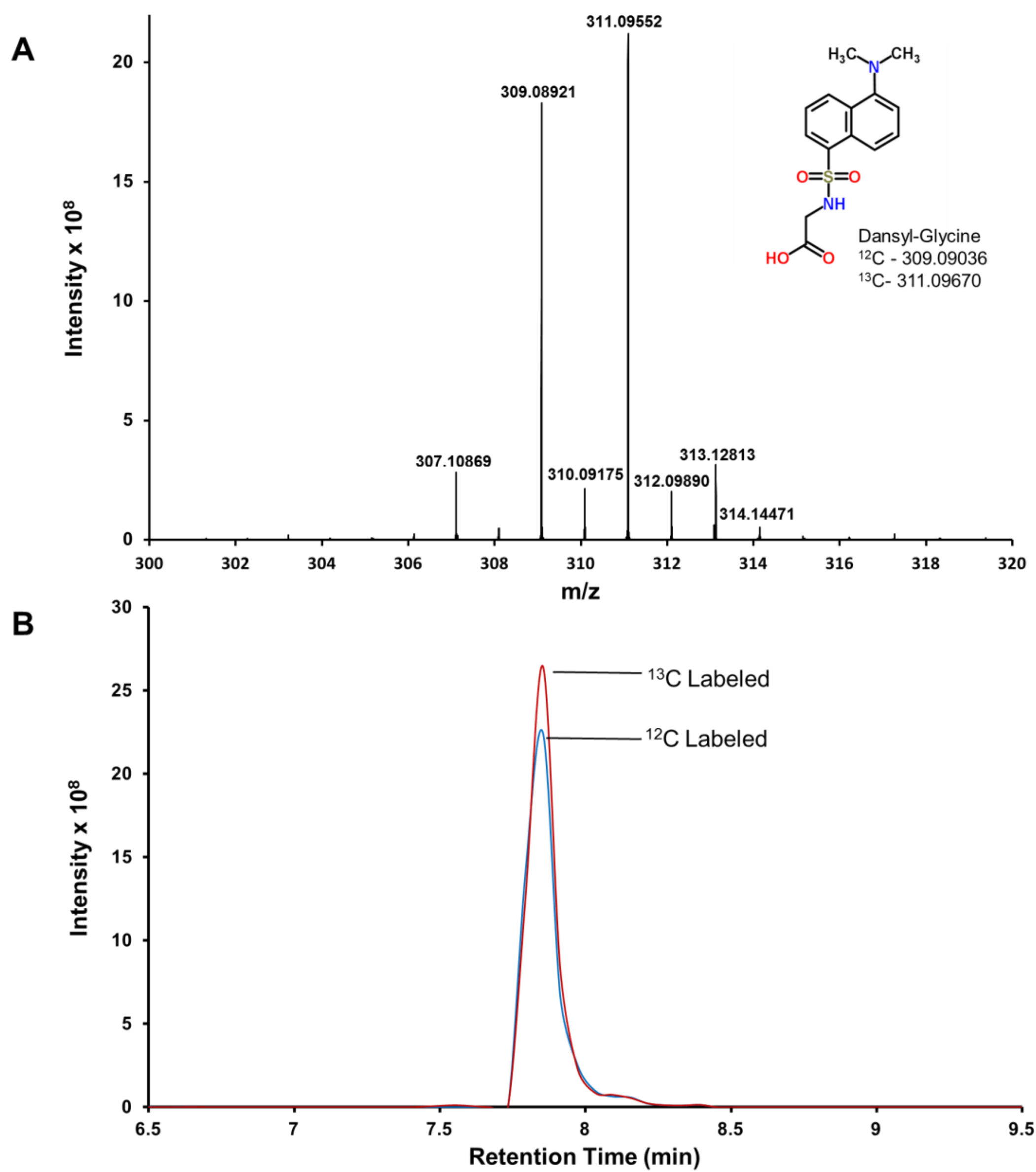


Figure 3-5 Mass spectrum of dansyl labeled glycine from analysis of labeled urine sample (A). Extracted ion chromatogram of both ^{12}C and ^{13}C labeled glycine (B).

Figure 3-3 shows a representative chromatogram of a urine sample labeled using reaction conditions that are currently the standard for all biological samples within our research group. The early baseline signal is due to the dead time of the chromatographic system. Peaks eluting between 0 and 5 minutes correspond to more polar compounds interact with the strongly polar mobile phase A. Distributing metabolites by HPLC separation in a biological sample is important when using MS for detection because there is a need for the reduction of sample complexity and ion suppression in the ESI source. Most of the metabolites of interest are contained between 4-24 minutes in Figure 3-3.Compounds eluting later than 25 minutes often hydrophobic byproducts or impurities that can carryover if sample preparation is done incorrectly. The large peak at 5 minutes belongs to DNS-NH_2, this is consistent with urine having a high concentration of ammonia.

To detect each metabolite that has been differentially labeled our approach takes advantage of the mass difference between sample aliquots labeled with ^{12}C DnsCl and those labeled with^{13}C DnsCl. In Figure 3-4 A, a mass spectrum taken from an LC-MS analysis of urine, the two largest peaks belong to Dns-glycine with both a light and a heavy tag attached. This metabolite ‘coding’ allows for the detection of metabolites that demonstrate this specific spectral pattern. The high resolution of the FT-ICR MS allows for the detection of compounds along with their multiply charged and adduct ions. The mass filter built into IsoMS is able to remove these ions and give a final peak pair number that is representative of the dansyl labeled metabolomic profile of a sample. By knowing that only metabolites displaying this spectral pattern will be detected, our approach adds an extra level of confidence for metabolomic profiling analysis.

The ratio of ^{12}C to ^{13}C labeled samples can be used as a method of relative quantification of metabolites. For a given sample set with many different subjects or sample types, all of the samples would be pooled to produce one standard which would be labeled with ^{13}C. Each individual sample would then be separately labeled with ^{12}C DnsCl and then mixed after UPLC-UV quantification. The ratio of light to heavy mass spectral peak intensities (Figure 3-4 A) would then allow for the determination of the amount of each specific metabolite for each specific sample in relation to the standard sample that had been previously heavy labeled. These values can then further be used as a relative concentration in multivariate statistical analysis.

Stable isotope labeled (SIL) compounds are very common in LC-MS analysis, especially as internal standard for absolute and relative quantification analysis. A problem that occurs when deuterated samples are used for analysis using LC is that they will not coelute with each other and also may not ionize at the same time.[35-36] This causes further problems in the fact that these two molecules will experience different ionization environments due to the presence of other analytes or due to the sample matrix differences.[37] By using ^{13}C labeling chemistry the isotopic effect is minimized. It is important to note that the limited availability of commerical standards and their prohibitive cost increases the attractiveness of using our groups labeling strategy.

One of the main goals of this differential labelling metabolomics approach is to detect as many labeled metabolites as possible in a given analysis. To assess this we use the mass spectrum to scan for peak pairs. As shown in Figure 3-5 A a peak pair is indicative of a putative metabolite detected during an LC-MS run. An easy comparison to make between different samples is to look at the raw number of peak pairs found using IsoMS. The peak pair results for different reaction condition are given in Figure 3-6 A.

Reactions heated conventionally in an oven are shown in blue. The trend appears to be that the

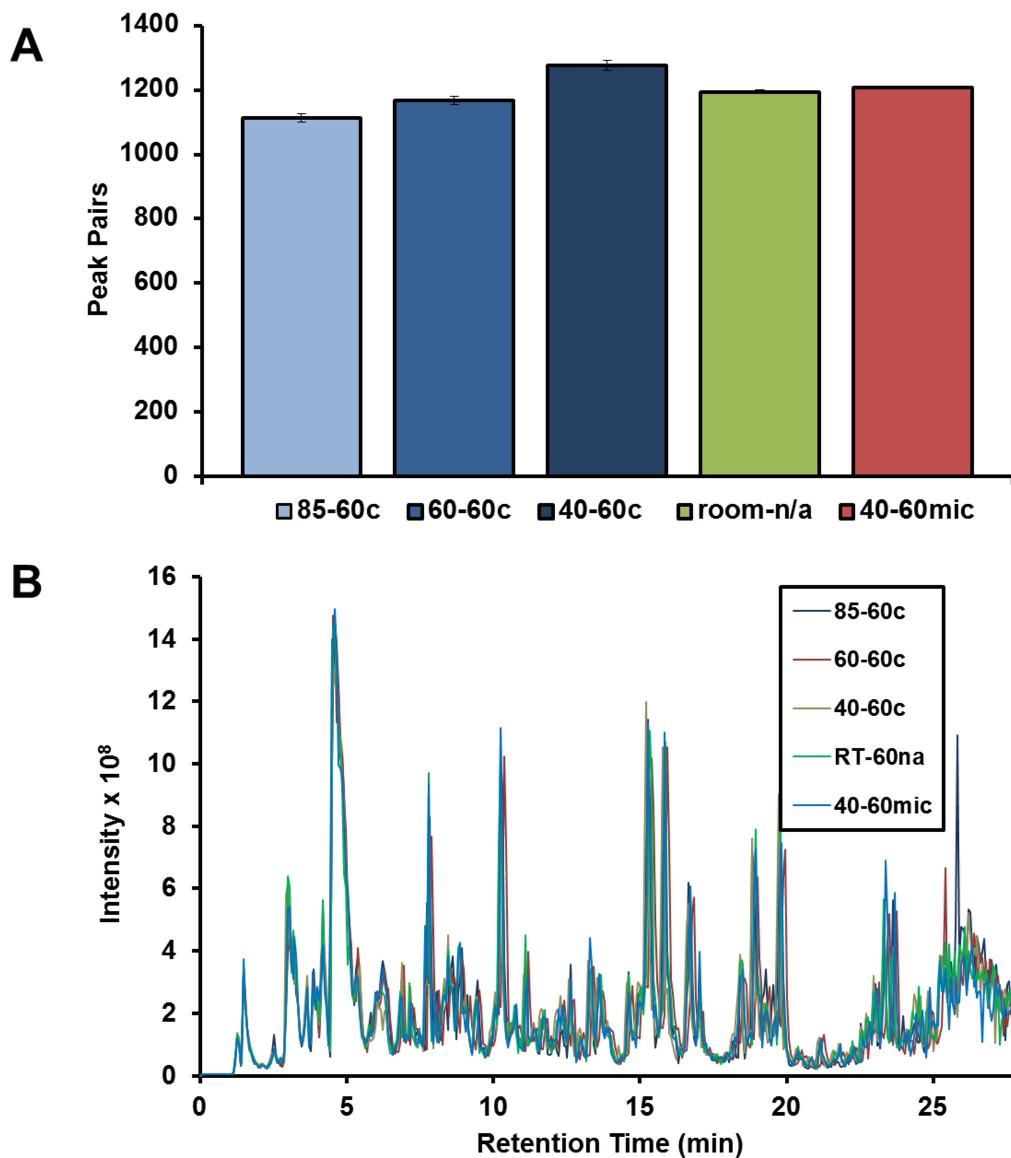


Figure 3-6 Number of differentially labeled peak pairs obtained using (A). Extracted ion chromatograms for each different reaction type (B).

number of peak pairs decreases with an increase in derivatization reaction temperature. The ambient temperature reaction produces a number of peak pairs close to that of the sample derivatized with MAH and that of the 60°C reaction. Interestingly the trend is the reverse of the results from quantification of total labeled metabolite using UPLC-UV. This is not unreasonable as there are a number of studies that show an increase in temperature can improve the derivatization reaction yield.[22] Reaction yield is not the most important metric for determination of the optimal reaction conditions. The results in Figure 3-6 highlight this fact, the number of peak pairs is much greater with an average of 1275 peak pairs found with a 40°C reaction which is more than with any of the other methods. Some possible explanations for this are that perhaps there is reduced sample degradation at lower temperatures. It is possible that some metabolites are not completely labeled and will be separately detected two or more times by the IsoMS peak picking software. A final explanation is that this temperature lies at the optimal balance point for labelling molecules that have variable derivatization yields dependant on temperature. A number of compounds demonstrated this behaviour when independently assessed in Chapter 2. Figure 3-6 B plots the total ion chromatogram from one of the three replicates of each reaction type given in the legend. Analysis of these chromatograms yielded no major differences and retention and peak shape is consistent. Differences at later retention times (i.e. those later than 25 minutes) are often due to the high percent of B mobile phase during the end of the chromatographic run. This causes hydrophobic compounds that are essentially stuck on the column to elute and the makeup of these eluents will be highly variable from injection to injection. Due to the slow scan rate of FT-ICR-MS chromatographic peaks are not always ideal and detection of peak pairs is harder.

The LC-MS analysis of biofluids for metabolites generates data sets that include the retention time and mass of each specific labeled metabolite and often

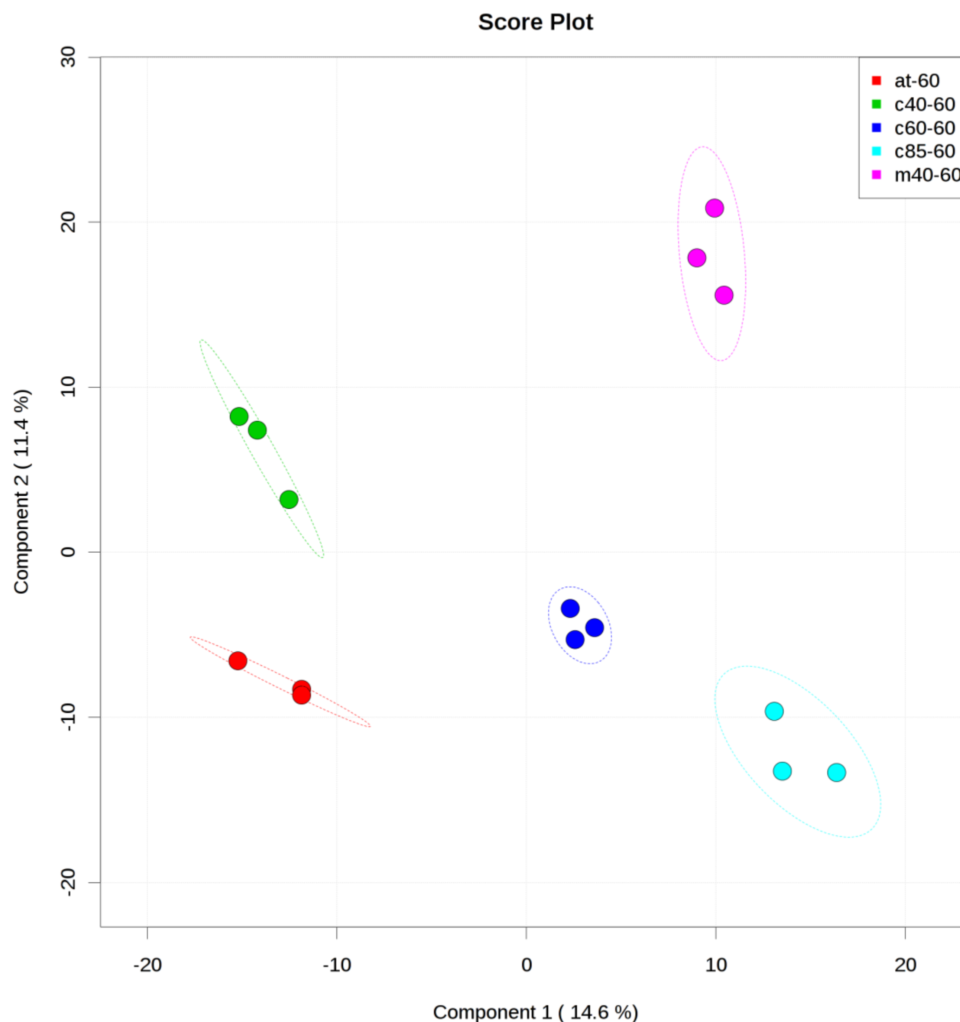


Figure 3-7 Partial least squares discriminant analysis (PLSDA) analysis of urine samples using variable reaction conditions.

with replicate samples. This type of data is multivariate. The analysis of multivariate data can be categorized as supervised and unsupervised. Principal component analysis (PCA) is an unsupervised method while partial least squares discriminant analysis (PLSDA) a supervised learning technique. This means that PLSDA uses class labels assigned to the samples to construct a model for interpretation of the IsoMS peak pair data. This method

uses the data to produce a projection that has the largest covariance between the data and the class labels (in this case the reaction methods used). It also can find the discriminant metabolites which are the main cause for the separation.[38] These analyses are for discrimination and classification. In the case of our data we wanted to look at classification of each of the separate reaction conditions. We performed PLSDA on the data set to determine if differences did exist between samples of urine labeled using different reaction conditions. Figure 3-7 shows the score plot results of this analysis. Each of the different reaction types separate well from one another under multivariate analysis.

Thus by changing the reaction conditions of DnsCl derivatization we can see not only an increase in the number of determined peak pairs corresponding to metabolites but also the production of variability in which samples are detected using our approach. Lower temperature reactions produced a larger amount of metabolite matches and therefore would be optimal for profiling metabolomics.

3.3.2 Urine and Plasma Samples with Varying Reaction Temperatures

The previous section led us to believe that lower derivatization reaction temperatures could provide complementary if not better analysis results for metabolomic profiling. In order to further explore this hypothesis we used urine and blood plasma from 3 different volunteers. The analysis was carried out in the same way outlined previously but with a more in-depth analysis of the final LC-MS results using statistical methods and different metabolite database searching. Our group currently uses a reaction temperature of 60°C, so this was used to compare with the lower temperature labeled samples.

3.3.2.1 UPLC-UV Quantification

Quantification of metabolites in the samples was important for both the proper mixing of ^{12}C and ^{13}C labeled samples. UPLC-UV is also used for comparison between samples from different volunteers and the comparison between the two different labeling temperatures employed in this study. The results for the samples before mixing are not

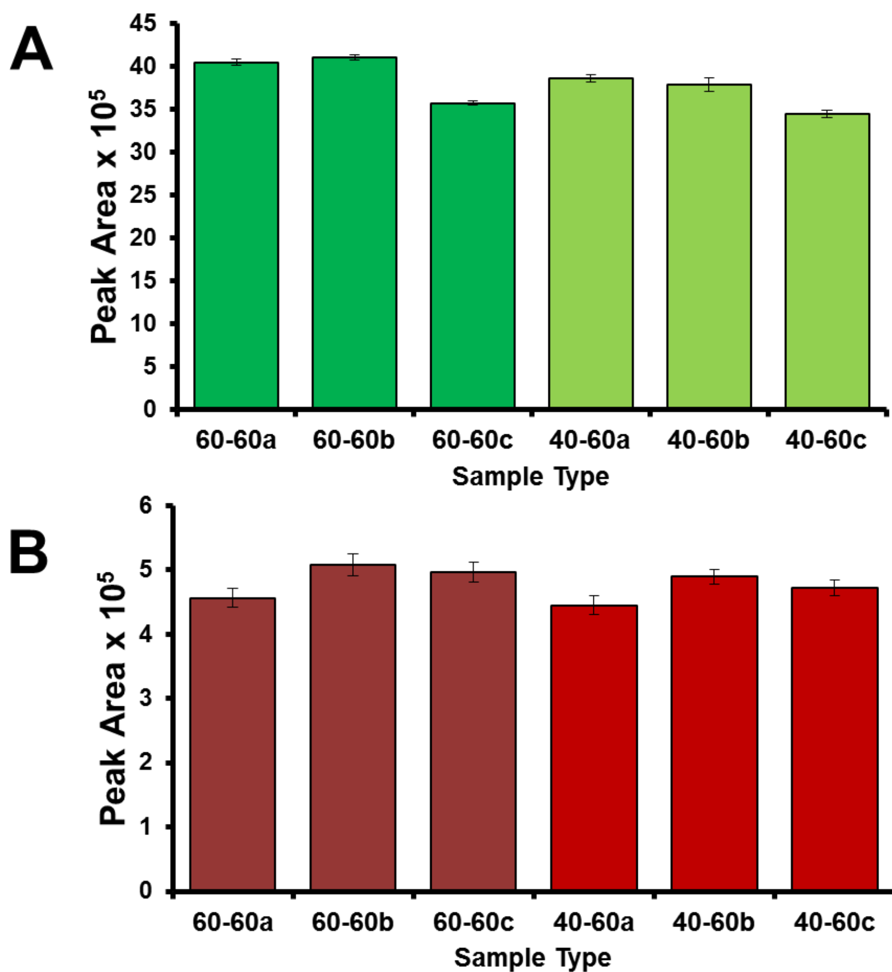


Figure 3-8 UPLC-UV quantitation of urine (A) and plasma (B) samples after derivatization with $^{12}C/^{13}C$ DnsCl and mixing of the separate reactions.

shown but they resemble the results in Figure 3-3 A. Only a small number of samples had variable mixing volumes in order to reduce variation in the amount loaded and analysed by LC-MS. Figure 3-7 A shows the results of UV quantification after mixing of ^{12}C and ^{13}C urine samples from three different volunteers using two different reaction temperatures. As was true in Figure 3-3 B the higher temperature reaction has a greater peak area corresponding to possibly more complete labeling. Each separate sample had some variation from one another as expected with fluctuation in the concentration of urine metabolites from person to person. Figure 3-7 B shows the UV quantification of plasma samples. The same trend exists with the higher temperature reaction producing a greater overall peak area for all samples. There is an order of magnitude difference in UV response between urine and plasma samples. This is interesting but not necessarily important as the sample preparation for each of these biofluids is very different. The peak area size is indicative of a larger amount of labeled metabolite but is not indicative of better labeling as there is more than one measure of effectiveness. By carrying out this analysis before and after mixing we can see that the replicate samples all agree within 5% of each other and therefore are thought of as a good measure of the variability of the biofluid samples.

3.3.2.1 LC-FTICR-MS Analysis of Urine and Plasma Samples

Once the samples had been mixed and analyzed by UPLC-UV, the next step in the analysis is to separate and detect the dansyl labeled metabolites using LC-MS. The samples are analyzed using the same protocol as the initial reaction condition assessment experiment. Figure 3-8 A shows overlaid total ion chromatograms of urine samples. Only one replicate from each of the individuals at each reaction temperature is shown. As shown previously there is very little change in the chromatograms of different samples demonstrating that there are only small changes in the metabolites makeup of each of the

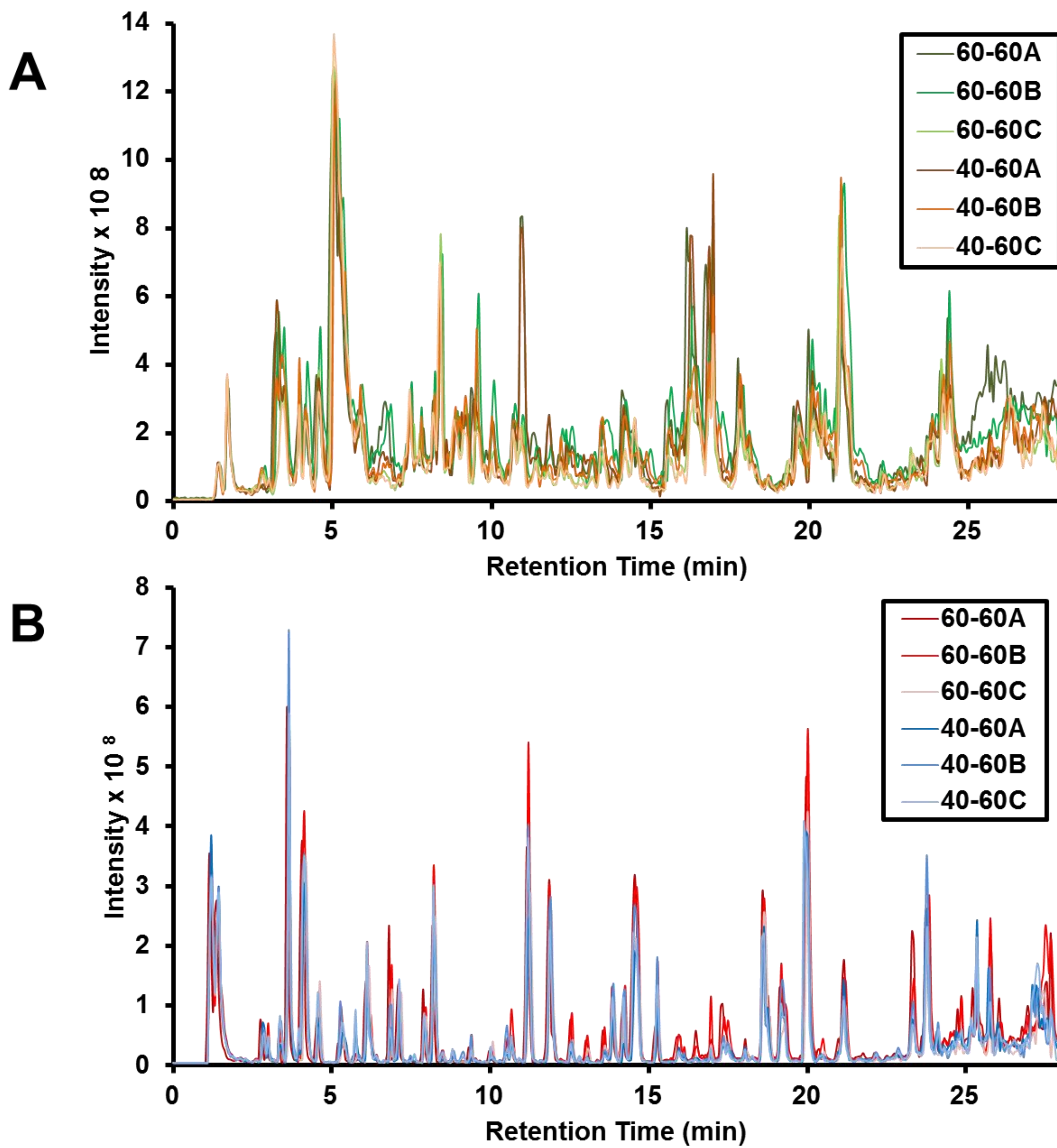


Figure 3-9 LC-MS total ion chromatograms comparing replicate samples of urine (A) and plasma (B).

samples. The total ion chromatograms for the plasma samples are shown in Figure 3-8 B. As with the urine samples, there is little variation from sample to sample and between reaction conditions. It is easy to see the difference between the urine and plasma samples. The urine samples have overall a much higher baseline than the chromatogram of plasma. This correlates with the UPLC-UV results discussed in the previous section. There is a large first peak in each chromatogram (1.2 minutes) that corresponds to the elution of Dns-OH , a byproduct of the reaction. In the plasma sample Dns-OH has a much higher relative amount to some of the other large peaks when compared with the urine sample. Another peak at 5 minutes in the urine sample corresponds to the by-product/ammonia peak of Dns-NH_2; this peak is the largest in any urine analysis chromatogram. It is interestingly of much lower abundance in the plasma sample. This is due to the high concentration of ammonia in the urine sample. The amount of information that can be gleaned from the chromatograms alone is minimal due to the complexity of the samples; in order to further assess the reaction conditions MS data must be considered.

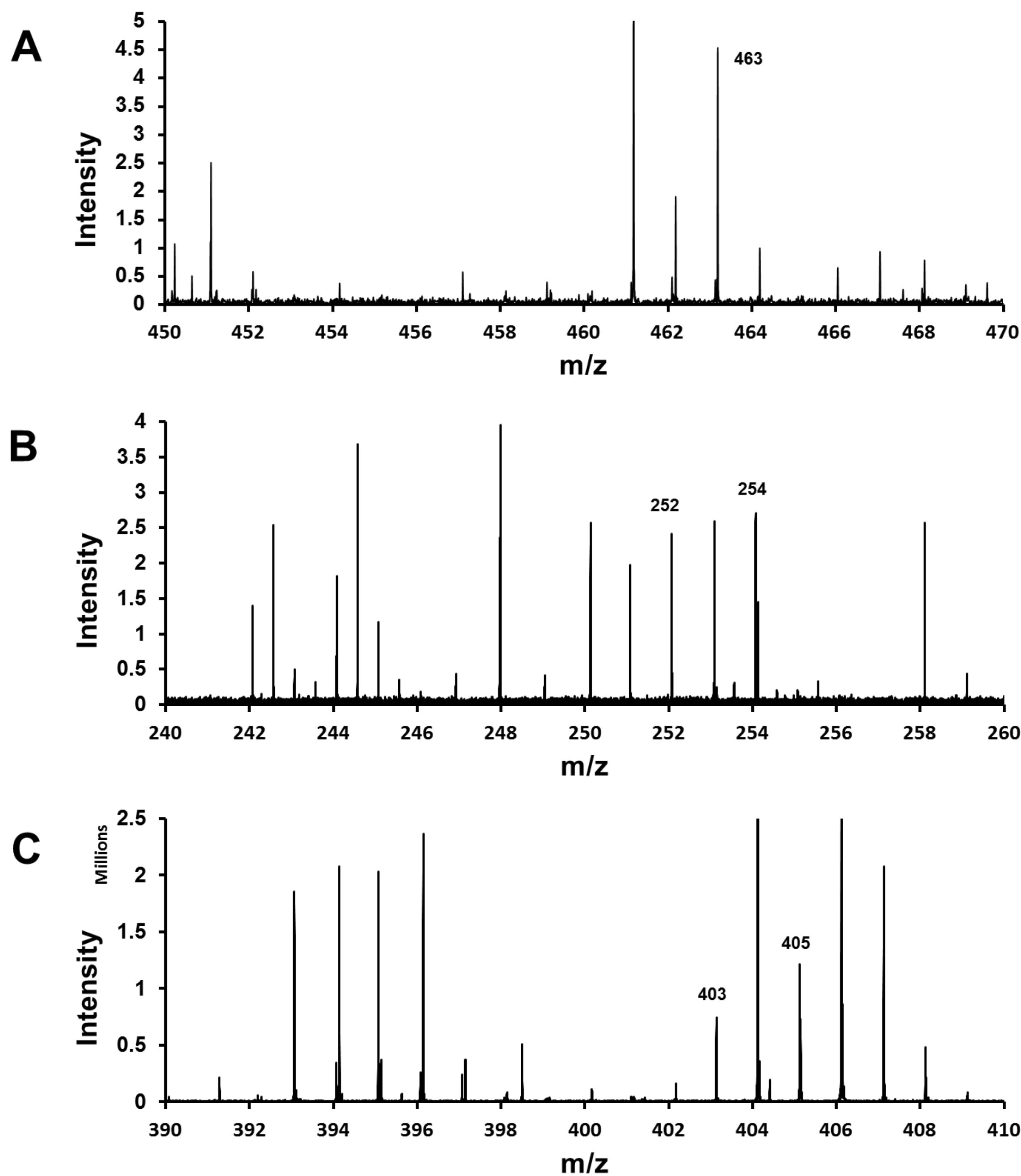


Figure 3-10 Spectra of peak pair classification level 1 (A), level 2 (B), and level 3 (C). Mass labels define which peaks are being discussed.

As previously discussed the number of total peak pairs determined is indicative

of the effectiveness of metabolite labeling for a given sample. In this experiment we used it as a comparison metric for the different reaction conditions and samples analyzed. The peak pairs are determined from the raw LC-MS data by using IsoMS. This program discriminates peak pairs using three different levels of classification. Level one is the most confident identification. This is done as a filtering system because it may be very easy to incorrectly pick two peaks with a 2 or 4 m/z difference that does not actually correspond to a peak pair. This is especially true of the complex biological samples analyzed in this work. IsoMS will analyze the natural isotopic pattern of each metabolite in order to determine its charge state. The first level of peak is categorized as having a known mass difference and both the heavy and light peaks containing natural isotopic peaks (Figure 3-9 A). A second level peak is asssigned when the heavy labeled peak has no discernible natural isotopic pattern and therefore the confidence of the peak pair match is much less than level 1 (Figure 3-9 B). The third level is for peak pairs with no natural isotopic peaks present. There is little confidence in these being true peak pairs and are often left out of statistical analysis.

The number of peak pairs was used to determine the effectiveness of the labelling reaction. The results for both urine and plasma samples are summarized in Figure 3-10 A and B respectively. The results are given for total peak pairs , level 1, level 2, level 3, and level 1 and 2 combined. The level 3 peak pairs are considered of very low confidence and are often not considered. The results for urine samples from three different individuals show the same trend as in the initial experiments presented in the previous section. For each separate sample when the reaction was carried out at 40°C the amount of peak pairs detected is greater than the reaction at 60°C. This is true for both level 1 and 2 peak pairs. This is important because if there were more level 2 or 3 peak pairs detected in the 40°C

labeled sample this could be indicative of decreased reaction yield or incomplete reactions. The same trend follows for the dansyl labeled plasma samples. The number of peak pairs found in plasma is much less than that of urine, indicating that there is less diversity in the phenol/amine metabolome of this biofluid. In both sample types there are only a small proportion of the peak pairs detected that are level 3. These generally would not be used for relative or absolute quantitative metabolomics but could be useful in profiling when the goal is the detection of many metabolites. The number of peak pairs found is consistently greater for the samples labeled at a lower temperature. This finding is important as this could lead to increased metabolite detection when samples are labeled with lower temperature reactions.

To further compare the difference in samples labeled at low and high temperatures, the number of unique peak pair matches arising from each specific reaction

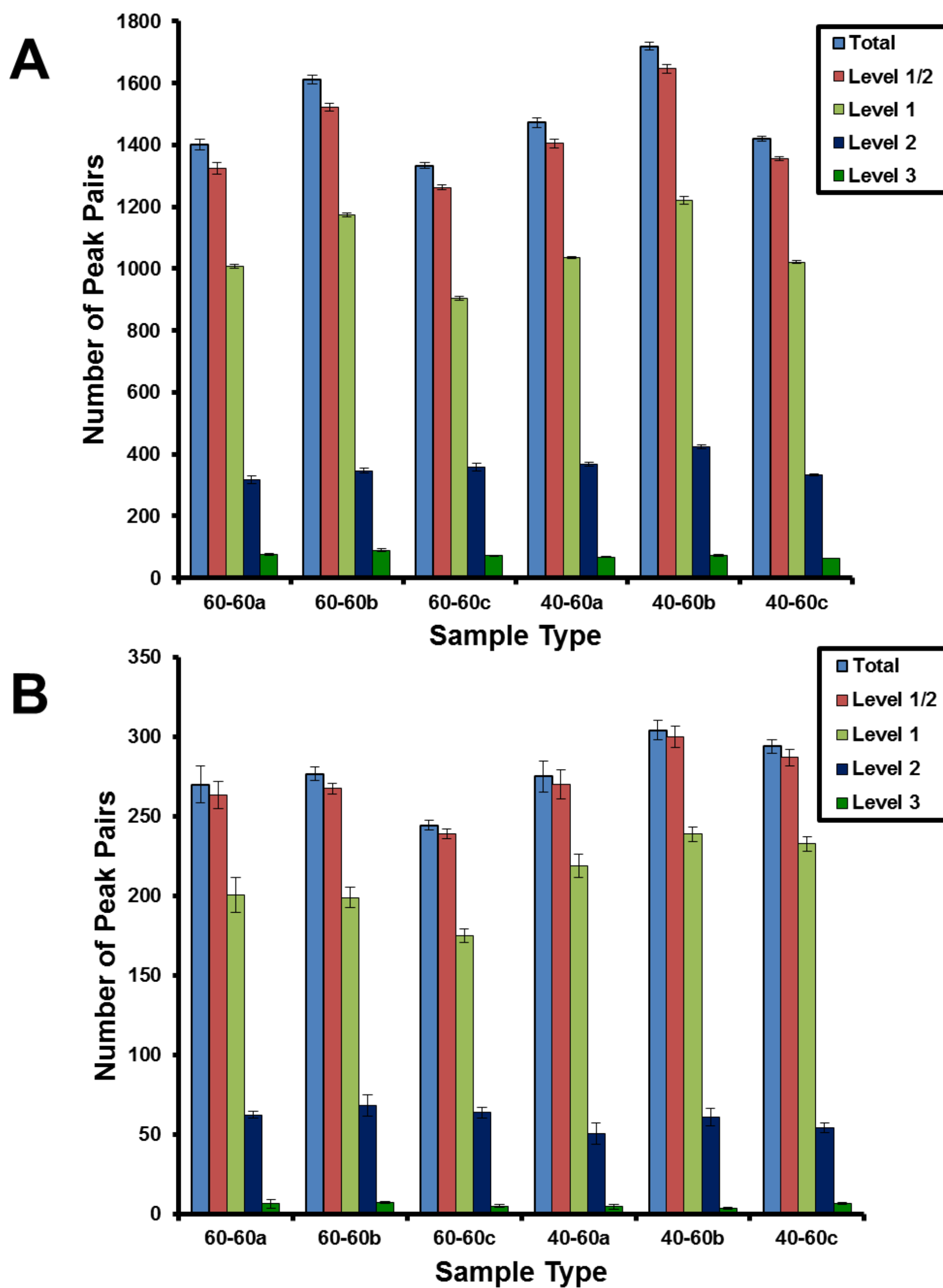


Figure 3-11 Peak pair counts from samples of urine (A) and plasma (B). The data presented is separated into peak pair level and the color coding is explained in the legend.

temperature was assessed and this data is presented in Figure 3-11 A and B for urine and plasma respectively. The number of matches refers to the times that each labeled metabolite appears in a triplicate analysis of 3 individuals volunteer's samples. Nine matches

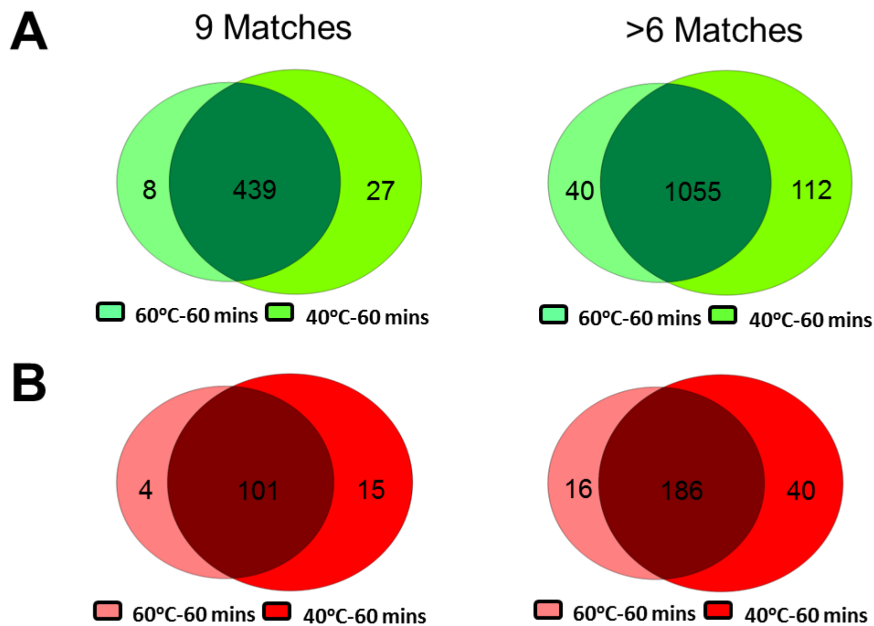


Figure 3-12 Number of metabolites found in each of the replicates on the left and those found in at least 6 on the right. Urine (A) and plasma (B) samples.

is considered to be quite confident as these metabolites show up in each of the replicate samples and from three different individuals. Also shown are the number of peak pairs found in 6 or more samples, which we are less confident in but nonetheless are thought to be true metabolites. For both plasma and urine samples there are more unique metabolites found in the 40°C labeled samples than those found in the samples labeled at 60°C samples. The total numbers are given in Figure 3-11 A and B for urine and plasma

respectively. The results for 9 matches are presented in Table 3-2 for urine and Table 3-3 for plasma. Many of the detected metabolites shown in these tables had no matches when using the HMDB search engine. This is not uncommon as this database is not all inclusive. The group number is assigned by the alignment software and without a metabolite match acts as an identifier for each peak pair.

Table 3-2 Labeled metabolites that appear in 9 of 9 urine samples analyzed and are unique to the reaction temperature used. Those with more than one name have multiple HMDB matches.

60°C Reaction				
Group	**Avg. rt(seconds)**	**Avg. Intensity**	**Molecular Weight**	**Name**
893	1055	1995333	187.0893	
2234	1222	237111	268.0730	Homolanthionine
				DHAP(6:0)
				Formononetin
2710	613	292556	146.0109	
3213	1339	313556	161.0429	
3250	1146	260444	162.0305	3 Hydroxycoumarin
3689	1290	323889	173.1003	
4471	594	266333	200.0744	
5209	1609	1102222	235.1875	

40°C Reaction				
Group	**Avg. rt(seconds)**	**Avg. Intensity**	**Molecular Weight**	**Name**
862	1548	415000	185.0184	
1026	1138	472000	100.0127	
1438	177	3482744	113.0553	
1648	1407	2528889	120.0536	
1723	1222	2013333	244.0622	
2062	918	317667	130.0227	
2120	524	129889	263.0748	
2121	840	92156	263.0754	
2392	1172	264444	275.0695	
3025	842	1286133	154.0696	
3135	1610	338444	316.0859	
3253	1380	466889	162.0384	
3288	1412	552556	164.0429	
3418	533	6911111	168.0236	
3420	425	1504444	168.0237	
3863	480	238733	180.0585	
3960	370	1437889	182.0740	2-Hydroxyfluorene
3963	441	592333	182.1121	
4451	524	264667	398.1258	
4465	1345	317222	200.0185	
5072	1324	373778	226.1257	

5531	280	400400	251.0928	
5870	1532	617000	272.1845	
5887	621	204222	275.0688	
6074	1471	338667	289.1749	
6335	1530	231222	315.1888	
6549	835	149667	358.0784	

Table 3-3 Labeled metabolites that appear in 9 of 9 plasma samples analyzed and are unique to the reaction temperature used. Those with more than one name have multiple HMDB matches.

60°C Reaction				
Group	**Avg. rt(seconds)**	**Avg. Intensity**	**Molecular Weight**	**Name**
284	1454	320000	240.0078	
349	355	1610444	129.0894	
776	1648	434667	222.1618	
810	1607	307089	234.1983	
40°C Reaction				
Group	**Avg. rt(seconds)**	**Avg. Intensity**	**Molecular Weight**	**Name**
70	1299	568111	160.0836	D-Alanyl-D-alanine
150	1113	181667	180.0621	3-Deoxyarabinohexonic acid
				Myoinositol
				D-Mannose
				D-Galactose
				D-Glucose
				Beta-D-Glucose
				D-Fructose
				Allose
				L-Sorbose
				Alpha-D-Glucose

				D-Tagatose
				Beta-D-Galactose
				Scyllitol
				L-Gulose
				D-chiro-inositol
				Paraxanthine
				Theophylline
				Theobromine
285	1465	173444	240.0080	
363	982	615778	131.0574	5-Amino-2-oxopentanoic acid
				4-Hydroxy-L-proline
				3-Hydroxy-L-proline
				L-Glutamic-gamma-semialdehyde
				5-Aminolevulinic acid
				Propionylglycine
				N-Acetyl-L-alanine
				Hydroxyproline
433	1197	238667	141.0789	
462	1079	682667	145.0733	2-Keto-6-aminocaproate
				(S)-5-Amino-3-oxohexanoate
				4-Acetamidobutanoic acid
				Allysine

				N-Butyrylglycine
				Isobutyrylglycine
566	531	4540000	162.0520	3-Hydroxymethylglutaric acid
				3-Hydroxyadipic acid
				2-Hydroxyadipic acid
				2(R)-Hydroxyadipic acid
				Glucosan
569	345	193333	162.0525	3-Hydroxymethylglutaric acid
				3-Hydroxyadipic acid
				2-Hydroxyadipic acid
				2(R)-Hydroxyadipic acid
				Glucosan
573	1031	308333	325.0386	
578	313	465556	164.0678	Rhamnose
				L-Fucose
				3-deoxyfructose
				1,5-Anhydrosorbitol
				Beta-D-Fucose
				L-Rhamnulose
				2-Deoxygalactopyranose

648	1065	475444	177.0455	N-Formyl-L-methionine
716	101	488667	201.9826	
718	399	698222	202.1058	
730	508	545922	203.0898	
739	676	270444	206.0421	Lipoic acid
				Methylisocitric acid
				Homocitric acid
				2-Methylcitric acid
806	1041	1023667	232.0845	

3.3.2.2 Multivariate Analysis

The data from these experiments was used for multivariate analysis by principal component analysis (PCA). This is a multivariate method used for the reduction of data complexity. PCA takes, in our case retention and concentration, and converts them to PCA scores that show most variability between samples. This produces score plots like those shown in Figure 3-12 A and B which can be useful to look for patterns or trends within a set of data. The score plots

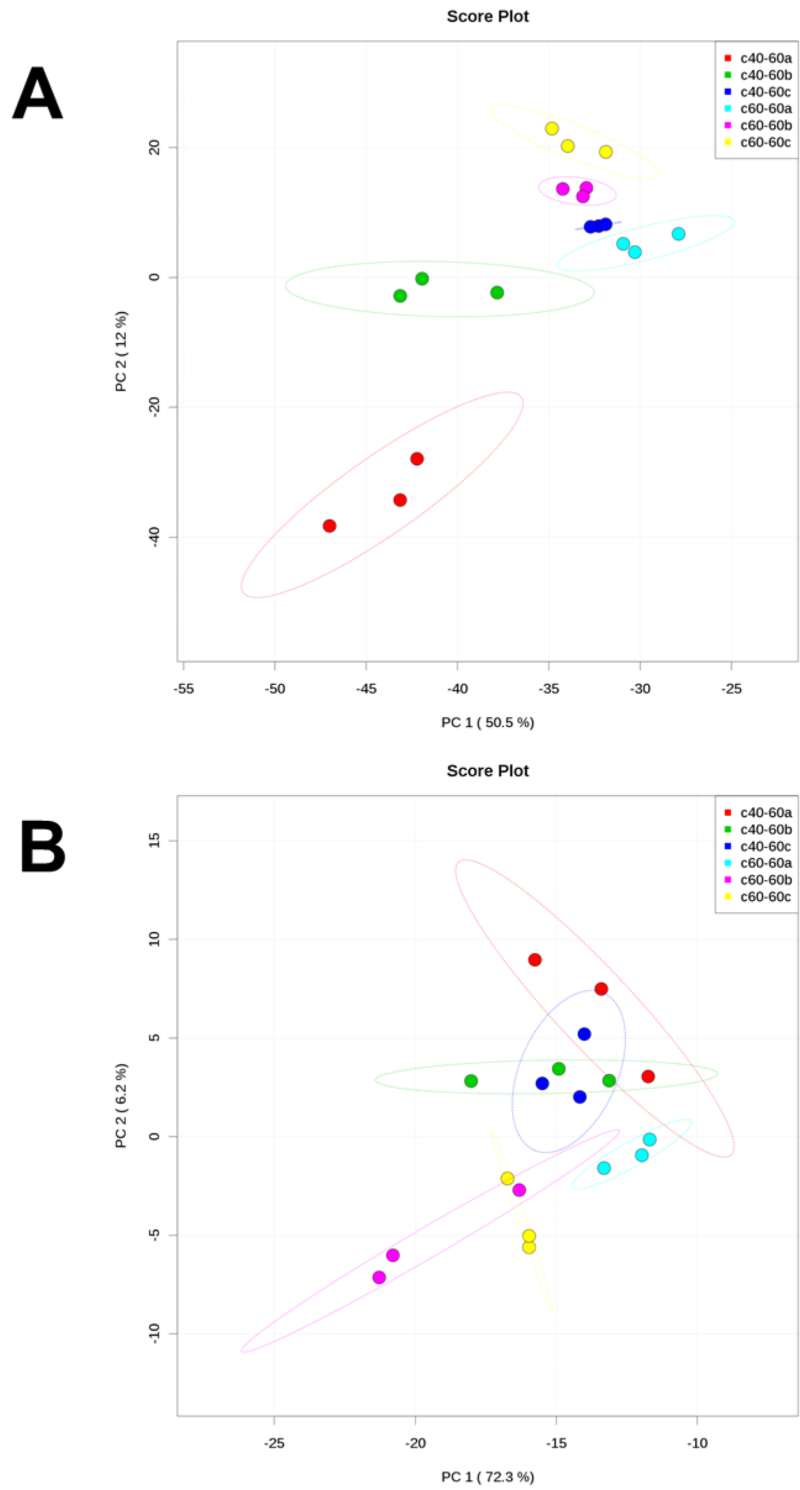


Figure 3-13 PCA score plots for urine samples (A) and plasma samples (B).

for both urine and plasma are shown in Figure 3-12 A and B for, urine and plasma respectively. The score plot for the urine sample shows good separation of samples labeled at different temperatures. The samples labeled at 40°C are shown in green, red, and blue while those labeled at 60°C are yellow, pink, and cyan. There is some clustering to separate the samples labeled at either temperature but each sample does separate from its replicate at a different temperature. The score plot for the plasma samples displays the same separation between the two reaction temperatures. There is however less defined separation between the urine samples of the same temperature. The results shown here demonstrate a difference when using different reaction temperatures for metabolite analysis.

3.3.3 Comparison of Conventional and Microwave Assisted Heating

Microwave labeling was discussed in the previous chapter and briefly during the initial conditions assessment. To study the results of using this method of derivatization we performed experiments that were intended to explore the use of MAH for rapid derivatization reactions. Urine from one volunteer was used to compare the various reaction conditions used.

3.3.3.1 UPLV-UV Quantification

As with the other experiments UPLC-UV quantification of total metabolites was carried out before and after derivatization. The results were much like those of previous experiments with the peak areas determined being very close to one another. The reaction at 60°C conventional heating followed by the reaction at 40°C conventional heat then by the p200 microwave and p100 microwave reactions. This is consistent with results from section 3.3.1.

3.3.3.2 Total Peak Pair Comparison

An assessment of the total number of peak pairs and the peak pair levels was carried out for the peak pair data extracted using IsoMS. As with the previous experiments the number of peak pairs found in the samples labeled with 40°C sample was greater than the other reactions. The short microwave reactions at powers of 200 and 100 w yielded peak pair values that were less than that of either of the conventional reactions. Also the amount of variation within each sample was higher in the microwave samples than in the conventionally labeled samples. These microwave reaction times were chosen from the optimized conditions determined in Chapter 2. The reaction time for those

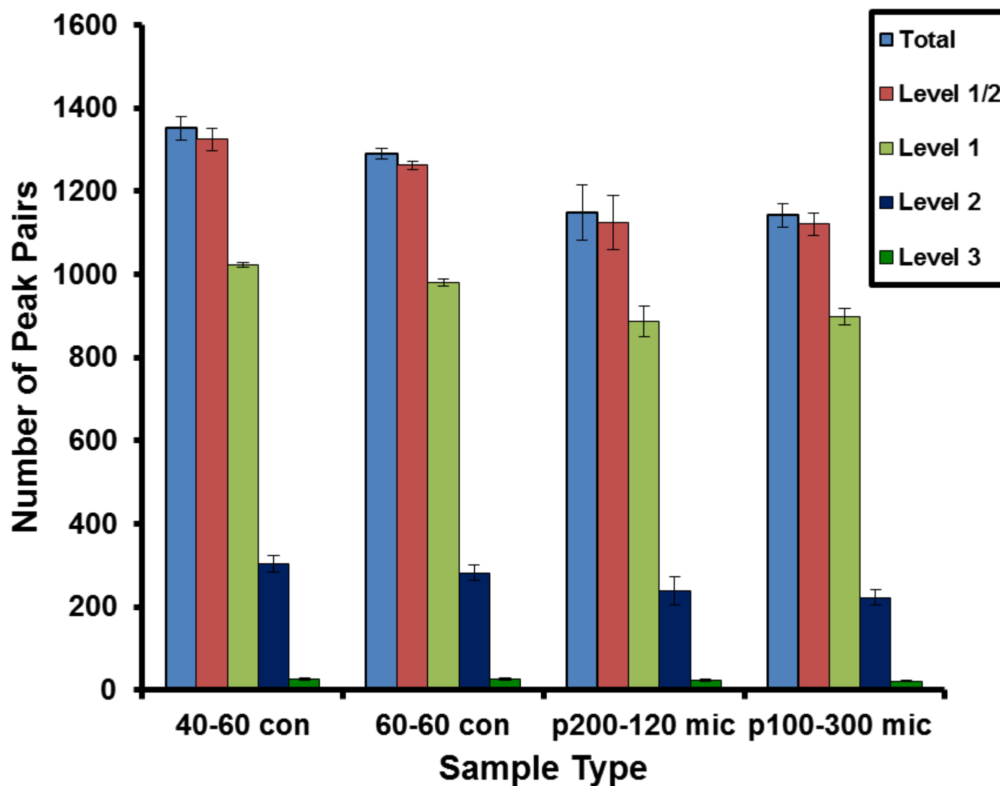


Figure 3-14 Peak pair amount and level comparison for a samples derivatized using both conventional and MAH heating. Reaction times for microwave reactions are given in seconds not minutes (i.e. 120 and 300 s)

samples reacted using MAH was kept short in order to determine if this type of reaction would be a viable option for profiling metabolomics using this approach. When using a microwave reactor if applying constant power, as was done for these experiments, the reaction temperature will increase linearly. At reaction times of five minutes the measured temperature is greater than 80°C. As discussed in Chapter 2 increased temperatures are suspected of causing metabolite degradation. In order to avoid these excess temperatures the reaction times were kept short. Microwave heating can also be applied with a measured temperature kept constant over specific time. It can be seen in Figure 3-6 A this type of reaction yields peak pair numbers comparable to the conventionally heated reactions but with no noticeable improvement. When comparing the two microwave reaction conditions used the only noticeable difference is that the longer lower power reaction seemed to yield a much less variable sample based on the standard deviation measures from triplicate samples.

3.3.3.3 Multivariate Analysis

To assess the data further multivariate analysis was carried out. It is hoped by separating the different reaction methods it would demonstrate that by applying different reaction methods there is a noticeable difference in the peak pairs measured using LC-MS. A PCA analysis (Figure 3-14 A) was carried out to compare the results of the microwave and conventional heating method. Samples labeled at two temperatures are separated from those labeled using microwave heating under two different conditions. This indicates there is differences in the metabolites identified using both of these methods. A further assessment of the reaction conditions was carried out using PLS-DA. Each of the different reaction types separates from one another indicating further that using different reaction methods causes an difference in both metabolites detected and their amount.

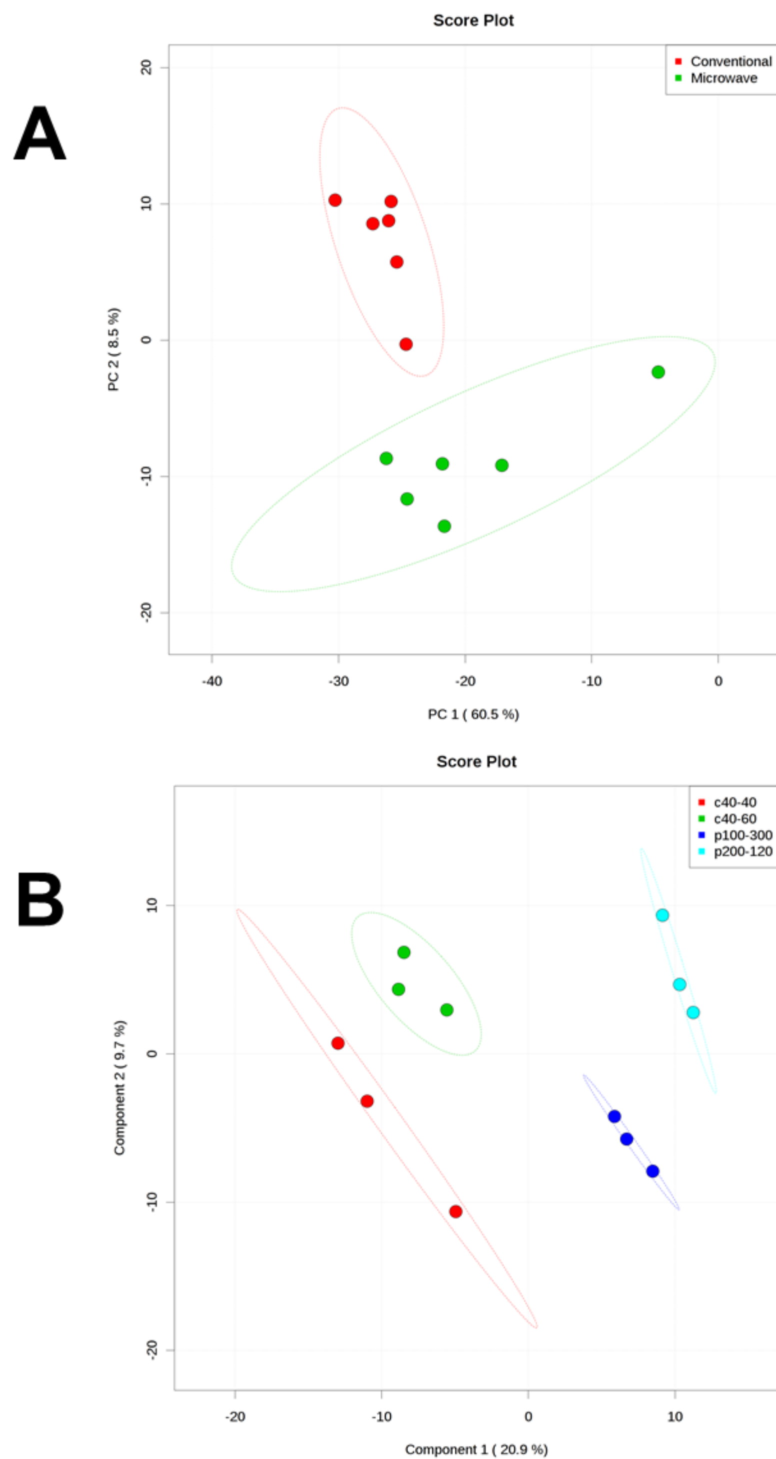


Figure 3-15 PCA plot of sample labeled using conventional heating in red and MAH in green (A). PLS-DA plot of each separate reaction (B).

3.3.3.4 Unique Metabolite Matches

The makeup of metabolites detected by peak pairs for different reaction conditions was demonstrated using multivariate analysis in the previous section. A further analysis is the assessment of the unique peak pairs that are found using each reaction type. These numbers were determined by comparing the number of peak pair matches that occurred in 2 or more of 3 sample replicates. The conventional reactions produced samples seem to have more unique peak pair matches. It is interesting however that there are peak pairs found exclusively in the MAH reactions.

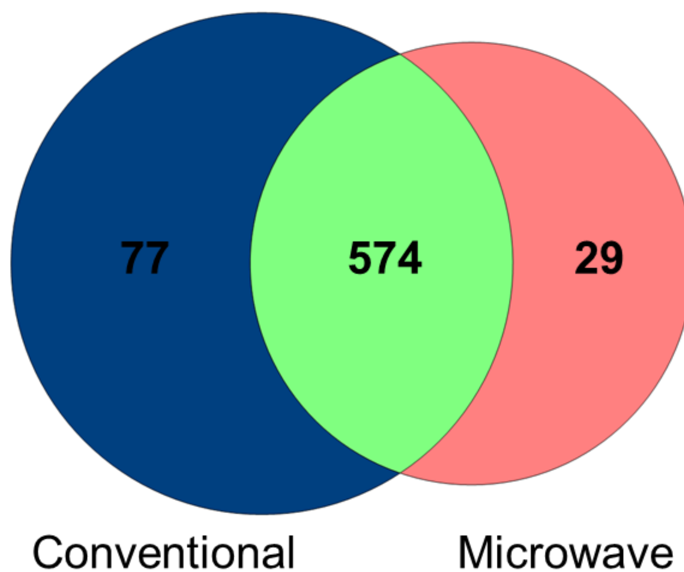


Figure 3-16 Comparison of unique peak pairs found in MAH reactions and conventionally heated reactions. The number in the middle is common peak pairs found in the two samples. These peak pairs were found in at least 2 of the 3 replicates assessed for each reaction type.

3.3.4 Method Validation

3.3.4.1 Selectivity and Carryover

Method selectivity was assessed by analyzing and acetonitrile injection blank, a reaction mixture blank with added metabolites, a 20 amino acid mixture dansyl labeled, and processed but unlabeled urine. Carryover was assessed by injecting an acetonitrile blank after three subsequent labeled urine sample runs. The chromatograms for each of these can be seen in Figure 3-15. The blank chromatogram only shows signal at the end of the run indicating there is some hydrophobic contamination on the column. This has been further assessed and there are only a few peak pairs detected in this portion of the chromatogram so this was of little concern. The analysis of the reaction mixture with no added sample or amino acid standard displays low background except for some expected peaks from dansyl reaction byproducts.[39] During every sample analysis a mixture of 20 labeled amino acids was analyzed. This serves as an indicator of retention shift from week to week and also of any changes in intensity of some select peaks. The results here show the total ion chromatogram with separation of most of the amino acids barring a few closely eluting compounds (Dns-Ile and Dns-Leu for example). A raw urine sample was run and is shown in Figure 3-15 D. There are about 40 detected valid peak pairs in this sample. These are what we usually deem as background and can be added to filters in the IsoMS software in order for removal from future samples. It is also important to note that in a comparison between the overall intensity of the dansyl labeled amino acid and the urine samples, it is almost double for the labeled sample than that of non labeled urine. Finally a blank sample was analyzed after three replicate labeled urine injections in order to assess carryover (Figure 3-17 E).

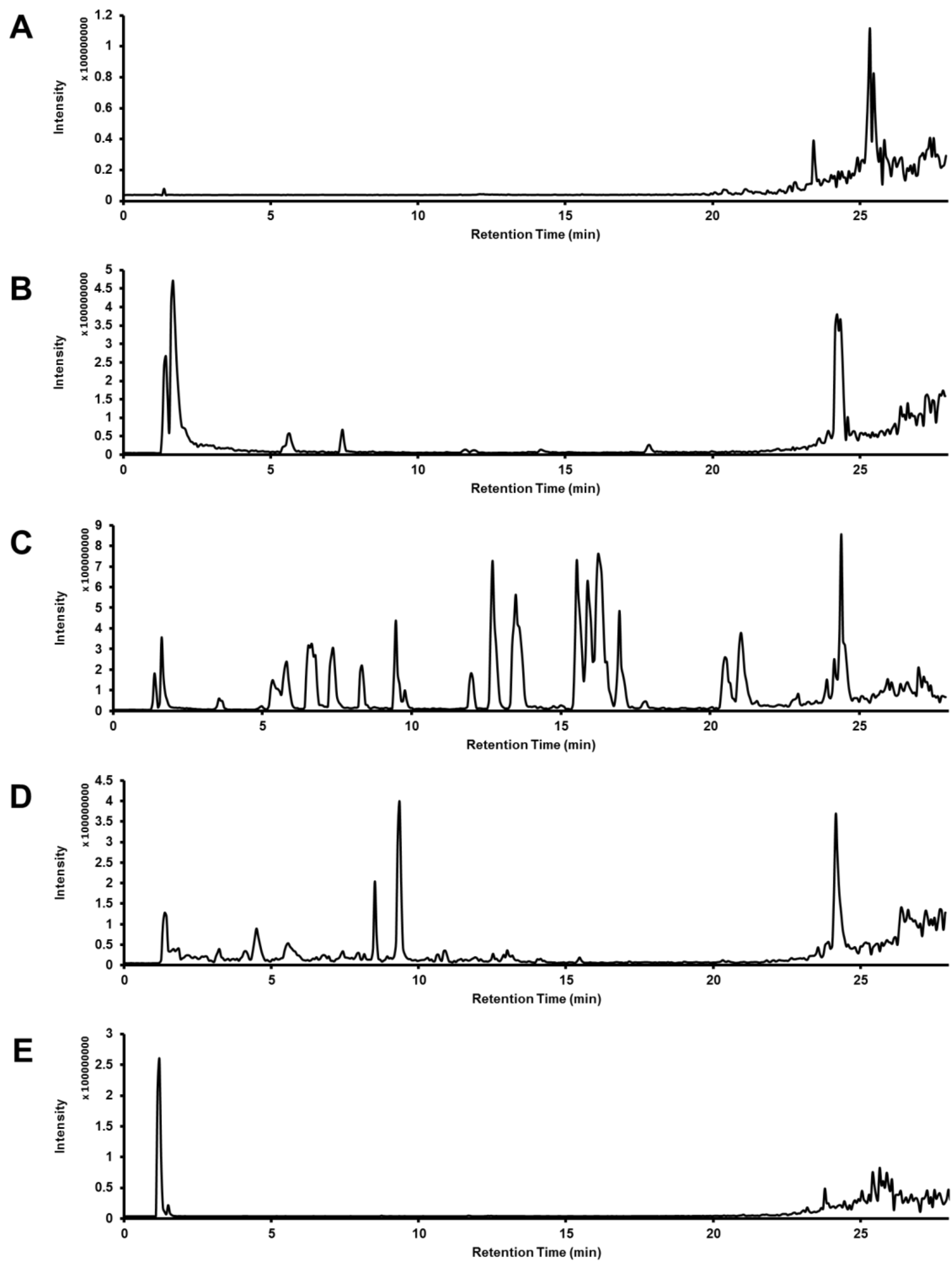


Figure 3-17 Representative total ion chromatograms for acetonitrile blank (A), reaction mixture blank (B), labeled amino acid standard (C), raw urine (D), and ACN blank after three successive labeled urine blank samples to assess carryover (E).

3.4 Conclusion

A number of derivatization conditions for differential isotope labeling using dansyl chloride were assessed. There is a variation of these different samples upon analysis using LC-MS. The reaction that produced the most amount of peak pairs was a 40°C reaction applied with conventional heat for 60 minutes. Although determination of a greater number of peak pairs is important it is not considered the only measure of reaction effectiveness. Both the lower and higher temperature reactions each label unique metabolites. It could be that there is no one optimal condition for metabolite labeling and that if instrument time was plentiful employing multiple reaction conditions for each sample with replicate analysis would yield the greatest and truest amount of information.

Microwave assisted heating was applied in two different ways, constant power and constant temperature. When assessed based on raw peak pair number, the microwave reactions underperform the conventionally heated reactions. However the short microwave reactions with constant power application produced reasonable results in a much shorter time than that of the conventional reactions.

3.5 Literature Cited

1. Patti, G. J.; Yanes, O.; Siuzdak, G., *Nature Reviews Molecular Cell Biology* 2012, 13 (4), 263-269.

2. Johnson, C. H.; Gonzalez, F. J., *Journal of Cellular Physiology* 2012, *227* (8), 2975-2981.

3. Cloarec, O.; Campbell, A.; Tseng, L. H.; Braumann, U.; Spraul, M.; Scarfe, G.; Weaver, R.; Nicholson, J. K., *Analytical Chemistry* 2007, *79* (9), 3304-3311.

4. Kuhara, T.; Shinka, T.; Inoue, Y.; Ohse, M.; Zhen-Wei, X.; Yoshida, I.; Inokuchi, T.; Yamaguchi, S.; Takayanagi, M.; Matsumoto, I., *Journal of Chromatography B: Biomedical Sciences and Applications* 1999, *731* (1), 141-147.

5. Shushan, B., *Mass spectrometry reviews* 2010, *29* (6), 930-944.

6. Maxwell, E. J.; Chen, D. D. Y., *Analytica chimica acta* 2008, *627* (1), 25-33.

7. Marshall, A. G., *International Journal of Mass Spectrometry* 2000, *200* (1–3), 331-356.

8. Hughey, C. A.; Rodgers, R. P.; Marshall, A. G., *Anal Chem* 2002, *74* (16), 4145-

9. Aharoni, A.; Ric de Vos, C. H.; Verhoeven, H. A.; Maliepaard, C. A.; Kruppa, G.; Bino, R.; Goodenowe, D. B., *Omics : a journal of integrative biology* 2002, *6* (3), 217-34.

10. Tohge, T.; Nishiyama, Y.; Hirai, M. Y.; Yano, M.; Nakajima, J.; Awazuhara, M.; Inoue, E.; Takahashi, H.; Goodenowe, D. B.; Kitayama, M.; Noji, M.; Yamazaki, M.; Saito, K., *The Plant journal : for cell and molecular biology* 2005, *42* (2), 218-35.

11. Gygi, S. P.; Rist, B.; Griffin, T. J.; Eng, J.; Aebersold, R., *Journal of proteome research* 2002, *1* (1), 47-54.

12. Tsukamoto, Y.; Santa, T.; Yoshida, H.; Miyano, H.; Fukushima, T.; Hirayama, K.; Imai, K.; Funatsu, T., *Biomedical chromatography : BMC* 2006, *20* (10), 1049-55.

13. Yang, W. C.; Adamec, J.; Regnier, F. E., *Anal Chem* 2007, *79* (14), 5150-7.

14. Armenta, J. M.; Cortes, D. F.; Pisciotta, J. M.; Shuman, J. L.; Blakeslee, K.; Rasoloson, D.; Ogunbiyi, O.; Sullivan Jr, D. J.; Shulaev, V., *Analytical Chemistry* 2010, *82* (2), 548-558.

15. Guo, K.; Li, L., *Analytical Chemistry* 2009, *81* (10), 3919-3932.

16. Wu, Y.; Li, L., *Analytical Chemistry* 2013, *85* (12), 5755-5763.

17. Zheng, J.; Dixon, R. A.; Li, L., *Anal Chem* 2012, *84* (24), 10802-11.

18. Guo, K.; Bamforth, F.; Li, L., *J. Am. Soc. Mass Spectrom.* 2011, *22* (2), 339-347.

19. Peng, J.; Li, L., *Analytica chimica acta* (0).

20. Gros, C.; Labouesse, B., *European Journal of Biochemistry* 1969, *7* (4), 463-470.

21. Cardenes, L.; Ayala, J. H.; Gonzalez, V.; Afonso, A. M., *Journal of chromatography. A* 2002, *946* (1-2), 133-40.

22. Jia, S.; Kang, Y. P.; Park, J. H.; Lee, J.; Kwon, S. W., *Journal of Chromatography A* 2011, *1218* (51), 9174-9182.

23. Wu, Y.; Li, L., *Analytical Chemistry* 2012, *84* (24), 10723-10731.

24. Xia, J.; Mandal, R.; Sinelnikov, I. V.; Broadhurst, D.; Wishart, D. S., *Nucleic Acids Res.* 2012, *40* (W1), W127-133.

25. Xia, J.; Psychogios, N.; Young, N.; Wishart, D. S., *Nucleic Acids Res.* 2009, *37* (suppl_2), W652-660.

26. Wang, Z.; Xu, H.; Fu, C., *Journal of Chromatography A* 1992, *589* (1–2), 349-352.

27. Smith, M. A.; Davies, P. J., *Plant Physiology* 1985, *78* (1), 89-91.

28. Maraschiello, C.; Miranda, E.; Millan, E.; Floriano, P.; Vilageliu, J., *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* 2003, *791* (1-2), 1-11.

29. Brown, M.; Wedge, D. C.; Goodacre, R.; Kell, D. B.; Baker, P. N.; Kenny, L. C.; Mamas, M. A.; Neyses, L.; Dunn, W. B., *Bioinformatics* 2011, *27* (8), 1108-1112.

30. Allen, J. P., *Biophysical Chemistry*. Blackwell Publishing: 2008.

31. Lenz, E. M.; Bright, J.; Wilson, I. D.; Morgan, S. R.; Nash, A. F., *Journal of pharmaceutical and biomedical analysis* 2003, *33* (5), 1103-15.

32. Tolstikov, V. V.; Fiehn, O., *Analytical biochemistry* 2002, *301* (2), 298-307.

33. Plumb, R. S.; Stumpf, C. L.; Gorenstein, M. V.; Castro-Perez, J. M.; Dear, G. J.; Anthony, M.; Sweatman, B. C.; Connor, S. C.; Haselden, J. N., *Rapid Communications in Mass Spectrometry* 2002, *16* (20), 1991-1996.

34. Zelena, E.; Dunn, W. B.; Broadhurst, D.; Francis-McIntyre, S.; Carroll, K. M.; Begley, P.; O'Hagan, S.; Knowles, J. D.; Halsall, A.; Wilson, I. D.; Kell, D. B., *Analytical Chemistry* 2009, *81* (4), 1357-1364.

35. Wade, D., *Chemico-biological interactions* 1999, *117* (3), 191-217.

36. Turowski, M.; Yamakawa, N.; Meller, J.; Kimata, K.; Ikegami, T.; Hosoya, K.; Tanaka, N.; Thornton, E. R., *Journal of the American Chemical Society* 2003, *125* (45), 13836-13849.

37. Zhang, R.; Sioma, C. S.; Thompson, R. A.; Xiong, L.; Regnier, F. E., *Anal Chem* 2002, *74* (15), 3662-9.

38. Barker, M.; Rayens, W., *Journal of Chemometrics* 2003, *17* (3), 166-173.

39. Parris, N.; Gallelli, D., *Journal of Liquid Chromatography* 1984, *7* (5), 917-924.

Chapter 4 : Conclusions and Future Work

This work focused on the qualitative and quantitative analysis of the results of metabolite labeling using isotopic derivatization reagent targeted at amine and phenol containing compounds. Both biological and simple standard samples were used for labeling and analyzed using both LC-MS and UPLC-UV. By using dansyl derivatization both the UV and MS detection of metabolites is improved. In addition this approach allows for the relative quantification, peak pair identification, and putative metabolite identification of metabolites in urine and blood plasma samples.

The analysis of simple standard mixtures was the assessed in Chapter 2. A quantification method for dansylated compounds using UPLC with UV detection was developed and validated using a number of methods. This method was then used to quantify the amount of derivatized metabolite after reaction under variable conditions. Using conventional heating it was found that the amount of product generally increases with temperature. Increases in reaction time demonstrate no remarkable changes in the amount of derivatized metabolite. This seems reasonable as the reaction is expected to happen quickly. Ambient temperature was found to be the most favourable method when compared with any conventional heating. Also evaluated were the effects of using microwave assisted heating for derivatization. Metabolites could be effectively labeled using rapid microwave heating. The quantified amount of labeled sample was greater than that of conventional reactions that require a much longer reaction time.

In Chapter 3 a differential isotope labelling using $^{12}C/^{13}C$ dansyl chloride with LC-MS analysis was used to assess the variation of derivatization reaction conditions. Again temperature, time, and heating method were varied, and the results assessed and compared. This differs from the previous chapter in that the results were evaluated less on a pure quantitative basis and more on their applicability to metabolomic profiling studies.

This was achieved through analysis of peak pair matches, putative metabolite assignments, and multivariate statistical analysis. When using peak pair number as an indicator the optimal reaction temperature was 40°C at a time of 60 minutes. Multivariate analysis demonstrates that there is a true difference in the number of labeled metabolites detected using this method and other conventional heating methods. Both urine and blood plasma from multiple sources were used to make this conclusion. Also assessed was the number of unique peak pairs found using these different heating temperatures. Both 40°C and 60°C reactions have unique peak pairs and putative metabolite matches. Microwave heating was also assessed using two different application methods. When comparing peak pair numbers obtained the rapid short method underperformed and the longer method designed to emulate conventional heating had similar results to that of the conventional heating reactions. Microwave heating did provide unique putative metabolites that were not detected when conventional heating was used for metabolite labeling.

From this work it is recommended that a temperature of 40°C be used for labeling of metabolites in both urine and plasma. Using microwave assisted heating for metabolite labeling in urine samples was not as effective as conventional heating but did produce unique results. If the analysis is of the targeted nature, with only a small number of compounds being assessed then higher temperatures would possibly produce better yields depending on the metabolites in question. Microwave assisted heating was found to label dansyl reactions with simple standard mixtures more effectively than that of conventional heating. Again this would be applicable for a targeted analysis.

Future work stemming from this thesis would be focused on further assessment of the difference in metabolite makeup between samples labeled at different temperatures. Perhaps these differences arise from lack of labeling efficiency (incomplete labeling) or there are true metabolites only labeled at specific temperatures. To evaluate this

hypothesis, standards for the putative metabolite matches would need to be obtained and their derivatization under different labeling conditions would be assessed. Our group has a number of different labeling chemistries currently in use or development. Using the same approach as in this study, their reaction conditions could be assessed and hopefully optimized. Finally it would be interesting to develop the application of dansyl labelling chemistry with microwave derivatization for targeted metabolomics analysis. There are a number of metabolites and metabolic pathways of interest that could perhaps be studied more effectively using a dansyl approach.