A MULTI-DIMENSIONAL METHOD FOR THE ANALYSIS OF HUMAN BLOOD PLASMA METABOLOME

A MULTI-DIMENSIONAL METHOD FOR THE ANALYSIS OF HUMAN BLOOD PLASMA METABOLOME

By

CATHERINE AMOATENG, B.Sc.

A Thesis Submitted to the school of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Science

McMaster University

MASTER OF SCIENCE (2011) McMaster University

(Chemistry) Hamilton, Ontario

TITLE: A Multi-dimensional Method for the Analysis of Human Blood Plasma Metabolome

AUTHOR: Catherine Amoateng, B.Sc. (University of Cape Coast)

SUPERVISOR: Professor B.E. McCarry

NUMBER OF PAGES: xx, 129

ABSTRACT

The comprehensive analysis of human blood plasma metabolome has been completed using derivatization gas chromatography-mass spectrometry (GC-MS) analysis, liquid chromatography-mass spectrometry (LC-MS) analysis and a comprehensive LC-GC-MS analysis approach wherein LC fractions were collected, derivatized and analyzed using GC-MS. In all cases blood plasma samples were deproteinized using solvent precipitation prior to chromatography and MS analysis.

In GC-MS analyses, the progress of all derivatization reactions was monitored by adding 9-anthracenemethanol and 1,3-diphenylacetone to all reaction mixtures; their conversions to 9-anthracenemethanol trimethylsilyl ether and the oxime derivative of 1,3-diphenylacetone were used as measures of the completion of these derivatization reactions. Any reactions with completions less than 99% were repeated.

GC-MS analysis of blood plasma samples detected 100 peaks; 44 were positively identified by comparing retention indices and mass spectra with those of authentic standards. LC-MS analyses were conducted on a HILIC column (aminopropyl phase) with MS detection in both negative ion and positive ion modes and resulted in the identification of 97 peaks; 47 were observed in the positive ion mode, 58 in the negative ion mode with 8 peaks observed in both modes.

The multi-dimensional LC-GC approach was not designed as a routine analytical method; rather the purpose of this approach was to see how many compounds could be observed in the sample and to obtain better quality mass spectra and retention index values. The LC separation afforded 16 fractions which upon derivatization GC-MS

analysis gave an additional 176 peaks from a total of 276 peaks. The MS data from these additional spectra can be used to develop selected ion monitoring GC-MS or tandem mass spectrometry analytical methods.

This thesis has demonstrated the power of off-line comprehensive methods to identify compounds that neither the GC nor the LC methods detected.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Brian E. McCarry for his expert guidance and financial support for my research work. I also wish to express my appreciation and thanks to members of my committee, Dr. Britz-McKibbin and Dr. Kirk Green for their valuable criticisms.

I am grateful to the staff of MLA Special Coagulation, Hamilton Health Sciences, McMaster Hospital's Division, especially Jodi Seecharan, for been there to draw my blood anytime I run out of human blood samples.

I gratefully acknowledge the support from NSERC and the Department of Chemistry and Chemical Biology.

I also wish to acknowledge the encouragement from my past and present students in ABB 464 and ABB 474; Libia Saborido Basconcillo, Tarlika Persaud, Sujan Fernado, Uwayemi Sofowote, Ken Chalcraft, Elna Deglint, Roger Luckham, and Fan Fei.

I am exceedingly thankful for the invaluable prayers from members of West Highland Baptist Church especially, Ron and Lisa Vella, Robbie and Vicki Jaster, Ken and Janine Murray.

DEDICATION

This thesis is dedicated to my husband, Erasmus Amoateng, my adorable son, Yaw Arko-Amoateng and my parents, Professor Stephen Asante-Poku and Mrs. Grace Asante-Poku.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACN: acetonitrile

AMDIS: automatic mass spectra deconvolution and identification system

APCI: atmospheric pressure chemical ionization

BioCyc: a collection of biological databases

CID: collision-induced dissociation

Da: Daltons

DIMS: direct infusion mass spectrometry

EcoCyc: encyclopedia of Escherichia coli K-12 genes and metabolism

ESI: electrospray ionization

FAME: fatty acid methyl esters

FT-ICR: Fourier transform ion cyclotron resonance

FT-IR: Fourier transform infrared

GMD: the Golm metabolome database

HMDB: human metabolome database

HumanCyc: encyclopedia of Homo sapiens genes and metabolism

Isop: isopropanol

KEGG: Kyoto encyclopaedia of genes and genomes

KNApSAcK: a comprehensive species metabolite relationship database

MeOH: methanol

MetaCyc: encyclopedia of metabolic pathway

METLIN: a metabolite mass spectral database

MMCD: Madison metabolomics consortium database

MMD: Manchester metabolome database

mRNA: messenger ribonucleic acid

MS2T: massbank : high resolution mass spectral database

MSTFA: N-methyl-N-(trimethylsilyl)trifluoroacetamide

NIST: National Institute of Standards and Technology

PAH: polycyclic aromatic hydrocarbons

PubChem: public chemical database

RSD: relative standard deviation

TMCS: trimethylchlorosilane

TMS: trimethylsilyl

TOF: time-of-flight mass analyzer

2D-GC: two dimensional gas chromatography

1. INTRODUCTION

1.1 Metabolites and the Metabolome

Metabolites are the low molecular mass compounds (<1500Da) of enzymatic reactions (1, 2). Metabolites are classified as either primary metabolites or secondary metabolites. Primary metabolites refer to those low molecular mass compounds that are directly involved in normal cell growth and cell development and include chemicals in compound classes such as amino acids, lipids, sugars, sugar alcohols and phosphorylated sugars (3). Secondary metabolites, on the other hand, are low molecular mass compounds that are not directly involved in normal cell growth but play a role in defence mechanisms of living things (3). Secondary metabolites include chemicals from the compound classes of sterols, terpenoids and alkaloids. It is estimated that there are about 90,000 – 200,000 metabolites in the plant kingdom with about 50,000 metabolites having already been characterized (4, 5). In human plasma, the number of reported metabolites is approximately 4200.

Similar to the definition of the terms genome (all genes in an organism), transcriptome (all mRNAs) and proteome (all proteins of an organism), the metabolome is defined as the total set of metabolites of an organism (1, 6, 7). Metabolites profiles are snapshots of cellular function and also define the phenotype of an organism. The concentrations of metabolites in an organism change as the organism responds to genetic expression and environmental stresses (8, 9).

The ability to measure and monitor changes in metabolite levels presents the potential for monitoring cell function in a comprehensive way that will have application

in the diagnosis and treatment of diseases, and the assessment of the health status of patients.

1.2 Classification of Metabolite Analysis

The analysis of the metabolome involves the detection and quantification of metabolites in an organism. The different analytical approaches used for metabolite analysis are grouped into four distinct approaches: target analysis, metabolite profiling, metabolite fingerprinting and metabolomics (10). The definitions given to them are:

Metabolite target analysis: Metabolite target analysis is the quantitative and qualitative analysis of one or very few related target compounds in a biological sample. The scope of metabolite identification is therefore limited to a small fraction of all metabolites. The identity of all target metabolite(s) is known and their methods for their quantitation are usually based on the use of isotope-labelled internal standards (4, 5).

Metabolite profiling: Metabolite profiling is the identification and quantitation of a pre-defined set of known and unknown metabolites from various compound classes. The pre-defined metabolites may be within a single class of compounds such as amino acids, lipids, organic acids or sugars, may be compounds in association with a specific metabolic pathway or may be from multiple compound classes. This approach is not considered to be comprehensive approach. In metabolite profiling, quantitation is done relative to a comparator sample (4, 5, 11).

Metabolite fingerprinting: Metabolic fingerprinting is a rapid, high-throughput sample screening approach for determining differences in metabolic profiles and for

classifying samples according to their biological origin. There is no separation of individual metabolites and as such metabolite quantitation is not involved (5, 11, 12).

Metabolomics: The metabolomics approach is the comprehensive analysis of all metabolites in the metabolome. The motivation for this approach is in the fact that results of a single genetic alteration on the levels of metabolites is generally not limited to one biochemical pathway and as such an unbiased and universal method of identification and quantitation of all metabolites is needed to provide a complete picture of the responses of genetic alterations or environmental stresses in biological systems (4, 5, 11).

Metabonomics, an analytical approach similar to metabolomics, is the multi-parametric detection and quantitation analysis of metabolites in studies of metabolic responses to drugs, diseases and environmental stresses (8-10). The term metabonomics was coined in the study of metabolites in human diseases and nutrition using NMR spectroscopy (13, 14) and researchers in toxicology and nutrition studies continue to use the term 'metabonomics' as compared to 'metabolomics' which is used by most researchers in the field. The difference is mainly a choice of terminology by research scientists (15).

The four above-mentioned analytical approaches are defined by the scope of their metabolite coverage. Whereas the metabolomics approach is comprehensive in nature, metabolite profiling, metabolite fingerprinting and metabolite target analysis approaches provide restrictive, non-comprehensive views of metabolite in organisms.

1.3 Analytical Techniques in Metabolomics

The analytical techniques currently being used for the analysis of metabolites can be grouped into two main categories: hyphenated mass spectrometry-based methods (MS) and spectroscopy-based methods (Figure 1-1) (16). The hyphenated mass spectrometry-based methods are further classified into two groups based on whether the metabolites are or are not separated chromatographically before detection (chromatography-based and chromatography-free MS methods).

Different technological platforms have their own strengths and weaknesses, the different technologies complement each other and none can be classified as truly comprehensive or clearly superior to any other method. For purposes of what class of compounds each technology can detect, a comparison of the following is discussed: NMR and mass spectrometry (MS) technological platform; different separation sciences in mass spectrometry based technologies; and one dimensional analysis against two dimensional analyses.

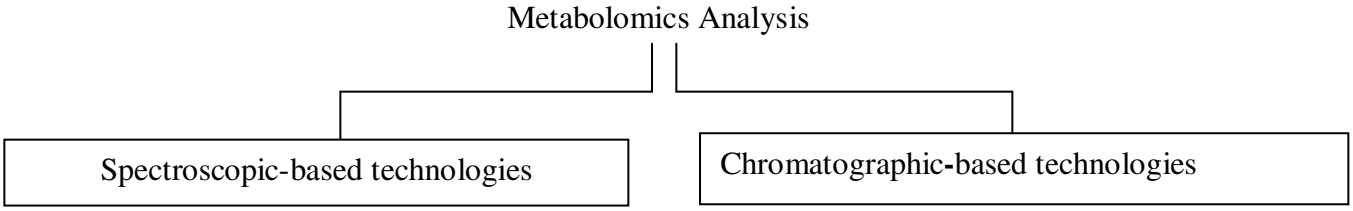


Figure 1-1: Two main categories of metabolomics analysis.

1.3.1 Spectroscopic-based techniques:

Spectroscopic techniques used in metabolite analysis include Nuclear Magnetic Resonance (NMR) and Vibrational Spectroscopy such as Fourier Transform Infrared Spectroscopy (FT-IR, (17, 18).

A) Nuclear Magnetic Resonance (NMR): The NMR technique focuses on specific chemical signatures relating to specific metabolic process. The pattern observed in the NMR spectrum is used in the classification of samples. The NMR technique requires little or no sample preparation, is high-throughput, unbiased, rapid, robust, reproducible and non-destructive but not very sensitive. The concentration range of detected compounds is μM to mM, thus allowing for the detection of only high abundant metabolite. Averages of about 30 metabolites are detected per sample (6). The technique is being applied in biochemical analysis in medical research. Investigations into biofluids such as urine, cerebrospinal fluid, saliva, cyst fluid, bile, blood plasma and tissue extracts have been reported. (19-22)

B) Vibrational Spectroscopy: The most commonly used vibrational spectroscopy technique in metabolomics is infrared (IR) and Raman spectroscopy. Similar to NMR, these vibrational techniques provide spectral fingerprints of biological samples. Both techniques provide very rapid (less than a minute per sample for both sample preparation and analysis time), reagentless, non-destructive, non-selective, unbiased and high-throughput analysis. One major disadvantage in vibrational spectroscopy is lack of sensitivity. Vibrational spectroscopy

techniques are currently seen as starting points in the hierarchy of analytical techniques used in metabolomics (23, 24).

1.3.2 Chromatography-based techniques:

Chromatographic separations methods are usually combined with a wide range of mass spectrometry-based detection methods. The most widely used methods include gas chromatography coupled to mass spectrometry (GC-MS), liquid chromatography coupled to mass spectroscopy (LC-MS) and capillary electrophoresis coupled to mass spectroscopy (CE-MS). The strengths and weakness of these techniques are discussed below.

A) **Gas chromatography-mass spectrometry (GC-EI-MS):** GC approaches afford high resolution analyses of volatile and semi-volatile compounds: in the case of metabolites analysis chemical derivatization is required prior to analysis. The type of ionization can be either electron impact or chemical ionisation. This technique is capable of a comprehensive profiling of wide dynamic range of metabolites (picomolar to nanomolar) (25) . GC-MS methods have been used extensively in biomedical, microbial and plant research. (26, 27). GC-MS methods are easy, rapid, robust, relatively inexpensive and reproducible. The EI mass spectral fragmentation patterns aid in identification by comparison to specific public and commercial mass spectral reference libraries. The GC-MS methods involve extensive sample treatment including chemical derivatization. (28). MS detectors generally used are: quadrupole MS and TOF-MS.

B) **Liquid chromatography-mass spectrometry:** Liquid chromatography followed by electrospray ionisation (ESI) is one of the commonly used techniques for metabolite separation. (8). The types of ionization used include ESI and APCI. Unlike GC/MS sample volatility is not required and therefore the LC/MS technique is suitable for the analysis of non-volatile and thermally unstable metabolites (10, 29). The separation and detection of some hydrophilic and small ionic metabolites such as amino acids and small organic acids can be poor. (16) To achieve higher resolution in LC/MS analysis, there has been improvement in chromatographic separation through the use of smaller column particles in LC systems (30). Covering a reasonable dynamic range of six fold, over 1000 metabolite peaks can be detected with capillary LC systems (31). The introduction of a multi-dimensional LC-MS, LC x LC-MS has demonstrated the potential of an increase in metabolite coverage to about 8000 metabolites (32). One other drawback in the use of LC systems is the non availability of standard spectral libraries to facilitate identification of unknown metabolites (33). MS detectors generally used are: FT-ICR, Obitrap FT, triple quadrupole MS, TOF-MS and ion-mobility TOF-MS.

C) **Capillary electrophoresis–mass spectroscopy (CE-MS):** This technique has been named the environmental friendly technique due to minimal solvent waste production (34). The ionization methods are ESI and APPI. CE-MS is used extensively in microbial metabolome analysis (35, 36) and in biomedical research (12, 37, 38). Its limit of detection is in the lower nanomolar range (39). Other advantages of CE-MS

include minimal sample preparation, high analytical speed, high resolution, robustness and high chromatography efficiency. Traditionally, this technique has an inherent bias against certain class of compounds as it is more of a feasible method for the separation of charged compounds. Time-of-flight MS is used.

In conclusion, none of the discussed technological platforms provides a truly comprehensive profile of the metabolome (Figure 1-2). Whatever the choice of analytical platform used, interest in the field of metabolomics is growing at a fast rate (Figure 1-3). The increasing interest in the field is demonstrated in a literature survey of published articles in metabolome analyses. Querying the ISI web of science database on the use of five analytical platform applications in metabolomics or metabolite profiling, a total of 695 publications were counted for GC/MS technology, 1176 for LC/MS, 156 publications for CE/MS, 1476 publications for NMR and 32 for 2D GC for the period covering year 2000 up to the end of 2009.

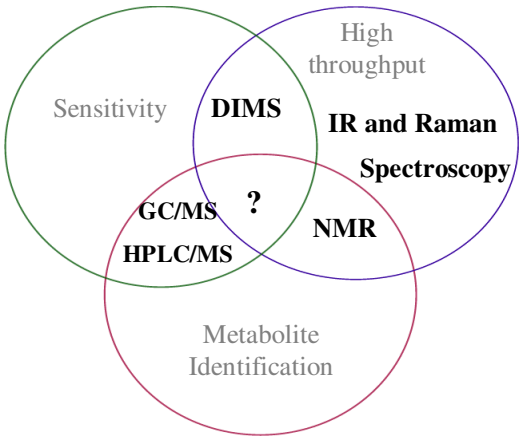


Figure 1-2: Current analytical techniques and objectives of metabolomics

Comparing the number of publications across all five analytical platforms in only year 2000 to the number of publications in only year 2009, there is close to a 100 percent increases in publication numbers. The steep increases in publication numbers can be attributed to improved multivariate statistical software packages available in the last decade.

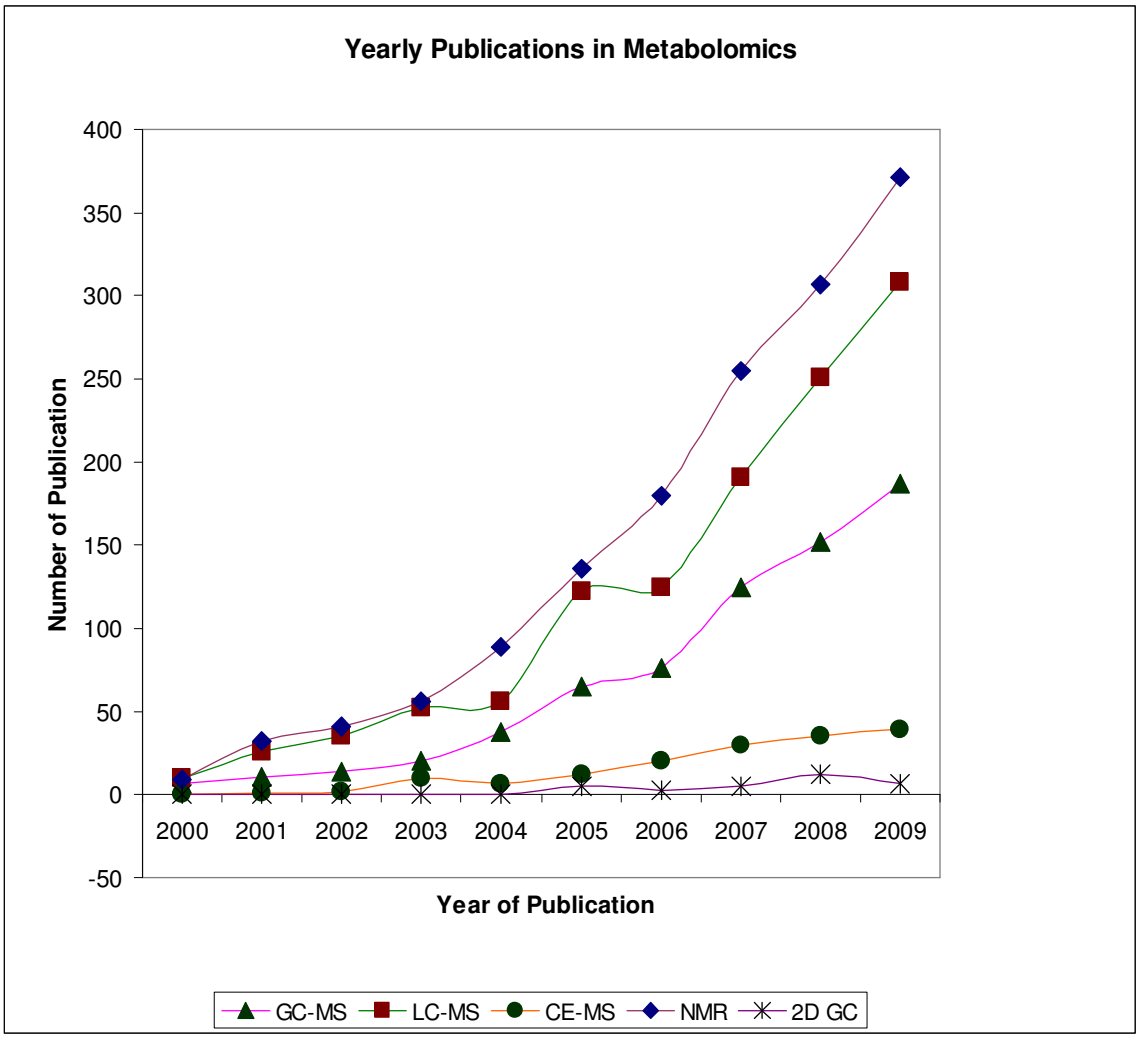


Figure 1-3: Number of published articles in metabolomics from 2000 to 2009

Table 1-1: Analysis of query results of ISI web of knowledge database in three categories: subject area, document type and country of origin.

Analysis of query results for NMR based metabolomics (metabonomics)					
Subject (Top 5)		**Document type (5 of 9)**		**Country (Top 5)**	
Analytical Chemistry	285	Articles	1128	England	453
Biochemical Research Methods	230	Review	161	USA	450
Biochemistry & Molecular Biology	203	Proceedings Paper	90	P.R. China	121
Pharmacology & Pharmacy	171	Meeting Abstract	78	Germany	99
Toxicology	121	Editorial Material	11	Netherlands	92

Analysis of query results for GC/MS based metabolomics					
Subject (Top 5)		**Document type (Top 5)**		**Country (Top 5)**	
Analytical Chemistry	187	Articles	539	USA	185
Biochemical Research Methods	142	Review	72	Germany	158
Plant Science	126	Proceedings Paper	65	England	88
Biochemistry & Molecular Biology	92	Meeting Abstract	13	P.R. China	83
Biotechnology & Applied Microbiology	82	Editorial Material	4	Japan	48

Analysis of query results for LC/MS based metabolomics					
Subject (Top 5)		**Document type (Top 5)**		**Country (Top 5)**	
Analytical Chemistry	448	Articles	922	USA	399
Biochemical Research Methods	307	Review	146	England	170
Pharmacology & Pharmacy	197	Proceedings Paper	60	Germany	141
Biochemistry & Molecular Biology	133	Meeting Abstract	41	P.R. China	115
Plant Science	112	Editorial Material	3	Japan	84

Analysis of query results for CE/MS based metabolomics					
Subject (Top 5)		**Document type (Top 5)**		**Country (Top 5)**	
Analytical Chemistry	115	Articles	130	USA	44
Biochemical Research Methods	81	Review	47	Japan	43
Biochemistry & Molecular Biology	25	Proceedings Paper	12	P.R. China	21
Pharmacology & Pharmacy	18	Meeting Abstract	5	England	17
Plant Science	13	Editorial Material	2	Germany	16

The query results can also be analysed in-terms of subject area, type of document, country of origin and/or journal type. Analysis of the first five subject areas shows dominance of biomedical or microbiological applications in all analytical platforms with toxicology research featuring largely in NMR applications (Table 1-1). In review documents, NMR

recorded 161 reviews followed by LC/MS with 146, GC/MS with 72 and CE/MS with 47. The huge number of NMR articles and review documents could be due to minimum sample preparation and historic coincidence in the use of NMR platform in toxicology research in England.

1.4 Identification of Metabolites

The identification of metabolites can be classified into four categories: [1] identification confirmed with an authentic chemical standard [2] reputedly identified by a match to a mass spectral and retention time/index database [3] belonging to a certain compound class and [4] unknown. As previously discussed, MS-based analysis are done using different types of separation (such as GC, LC, CE), ionization (EI, CI, APCI, ESI) and as well as different detection techniques. Detection of metabolite is done in any of these three data sets:

I. Parent mass only
II. Parent mass with chromatographic retention time
III. Parent mass with chromatographic retention time/index and mass spectral data

To facilitate metabolite identification, three types of MS-based metabolite databases exist and these are:

A. The raw unidentified MS (GC-MS or LC-MS) spectral data of biofluids or tissues extracts.
B. Identified MS (GC-MS or LC-MS) spectral data of biofluids or tissues extracts.

C. Reference retention time/index and MS spectral data (GC-MS or LC-MS) of pure compounds.

A list of commercial and free electronic database containing detailed information on small molecule metabolites found in the human body is listed in Table 1-2. These databases are searchable and available for use in the identification process of metabolites as well as linking compounds to molecular interactions and reaction networks for biological interpretation of systemic functions.

With the capability of MS-based techniques to detect 100s of metabolites, the following issues limit the identification of all detected metabolites in the metabolite identification process:

i. Different stand-alone software applications of propriety MS instruments
ii. Different sets of spiked-in retention time markers
iii. Unstandardized information in currently available databases
iv. Limited number of confirmed known metabolites in current databases.

Table 1-2: Some commonly used chemical, mass spectral and metabolome databases.

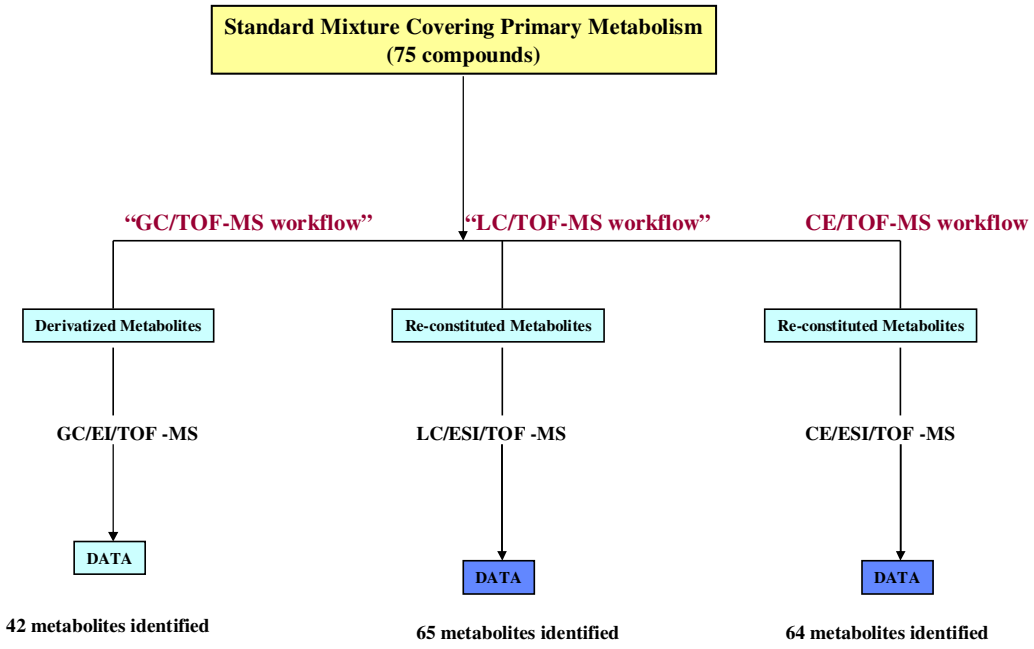
Name	Content	Reference
Compound and species specific databases		
PubChem	A collection of chemical structures and biological properties of small molecules, including three sub-databases: substance, compound and bioassay.	(40, 41)
HMDB	A comprehensive database of >6800 human body metabolites harvested from experiments and literature, including MS and NMR spectra for approximately 3000 ,metabolites and two sub-databases: DrugBank and FooDB	(1, 42)
LIPID MAPS	Structure and annotations of >9000 biological lipids in human and other mammals	(43, 44)
KNApSAcK	A cross-species metabolite database of >28,500 compounds with approximately 100,000 masses, mostly from plants and microorganisms.	(45)
MMD	A knowledgebase of 42,687 endogenous and exogenous metabolites constructed from genome-scale metabolic pathways, HMDB, KEGG, LIPID MAPS, BioCyc and DrugBank, including mass spectral libraries of 1065 authentic metabolite standards.	(46, 47)
Reference mass spectral libraries or databases		
NIST08	A commercial library of >220,000 EI mass spectra from >190,000 pure chemical compounds; containing 14,802 spectra of 3898 positive ions and 1410 negative ions and 293,247 Kovats RI values for 44,008 compounds	(48)
GMD	Mass spectra and retention indices of known plant metabolites analyzed by GC-quadruple MS and GC-TOF-MS	(49, 50)
METLIN	A database of LC-MS, LC-MS/MS and LC-FTMS mass spectra of metabolites from human and microbial species	(51, 52)
MassBank	A high-resolution MS/MS spectral database for standard chemical substances	(53)
MS2T	An MS/MS spectral tag library of phytochemical compounds	(54)
MMCD	An NMR and LC-MS spectral library of metabolite standards	(55)
Metabolic Pathway-specific databases		
KEGG	A composite database consisting of collections of pathway maps, genes, organisms, enzymes and ligands (metabolites, drugs and other small molecules)	(56)
BioCyc	A collection of 506 pathway/genome database of organisms with completely or partially sequenced genome, database of chemical compounds	(57-59)
EcoCyc	A scientific database for the bacterium Escherichia coli K-12 MG1655	(60)
MetaCyc	A database of nonredundant, experimentally elucidated metabolic pathways	(61-63)
HumanCyc	A human metabolic pathway and human genome database	(64, 65)
Reactome	A database of human metabolic pathways and biological processes involving small metabolites, biomolecules and their reactions/interactions	(66-68)

1.5 The Analytical Challenges in Metabolite Analysis

The metabolome is made up of a large number of small molecular mass compounds from various compound classes such as ionic inorganic species, hydrophilic carbohydrates, hydrophobic lipids, complex natural products, etc. These compound classes exhibit a wide diversity of physical and chemical properties. The metabolites also cover an eleven fold dynamic range with limits of detection ranging from picomolar (pM) to millimolar (mM) concentrations (8, 25). The physical and chemical diversity of the metabolome poses significant challenges for a comprehensive analysis of the wide range of metabolites.

The demands of any analytical technique in a comprehensive metabolite analysis therefore are: high sensitivity, selectivity, universality, robustness and reproducibility. The chemical complexity and wide range of compounds within the metabolome makes it unlikely that a single analytical platform could separate and identify all metabolites in a single analysis (9). To evaluate the potential of three analytical platforms, Buscher and colleagues analyzed a standardized mixture of 75 metabolites on GC-MS, LC-MS and CE-MS as described in Figure 1-4 (69). The 75 metabolites represented metabolites in central carbon and energy metabolism. In conclusion, 72 out of the 75 compounds could be detected on at least one platform and three compounds probably due to loses in sample preparations, could not be detected with any of the platforms (Figure 1-5). The group also reported that 33 out of 75 metabolites could be detected on all three platforms while very few metabolites could only be detected on any single platform. The platforms with

separation of metabolites in solution, LC and CE, had the greatest overlap with respect to metabolite coverage.



Jorg Martin Buscher et al , Anal. Chem. 81(6), 2135 – 2143, 2009

Figure 1- 4: Comparison of metabolites coverage in three different analytical platforms

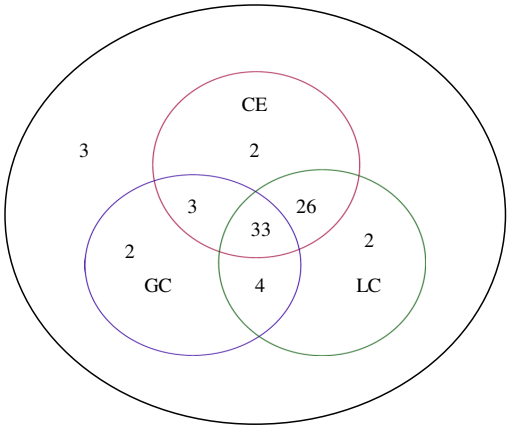


Figure 1-5: Coverage of 75 central metabolic compounds by three analytical platforms

The limited coverage in GC analytical platform is attributed to steric hindrance of silyl groups introduced in the method to increase volatility of hydroxyl group of compounds.

1.6 Multidimensional Separation Methods

With limited coverage in the use of individual technological platforms (Figures 1-4); there is a push for multidimensional analysis for the comprehensive analysis and metabolite identification.

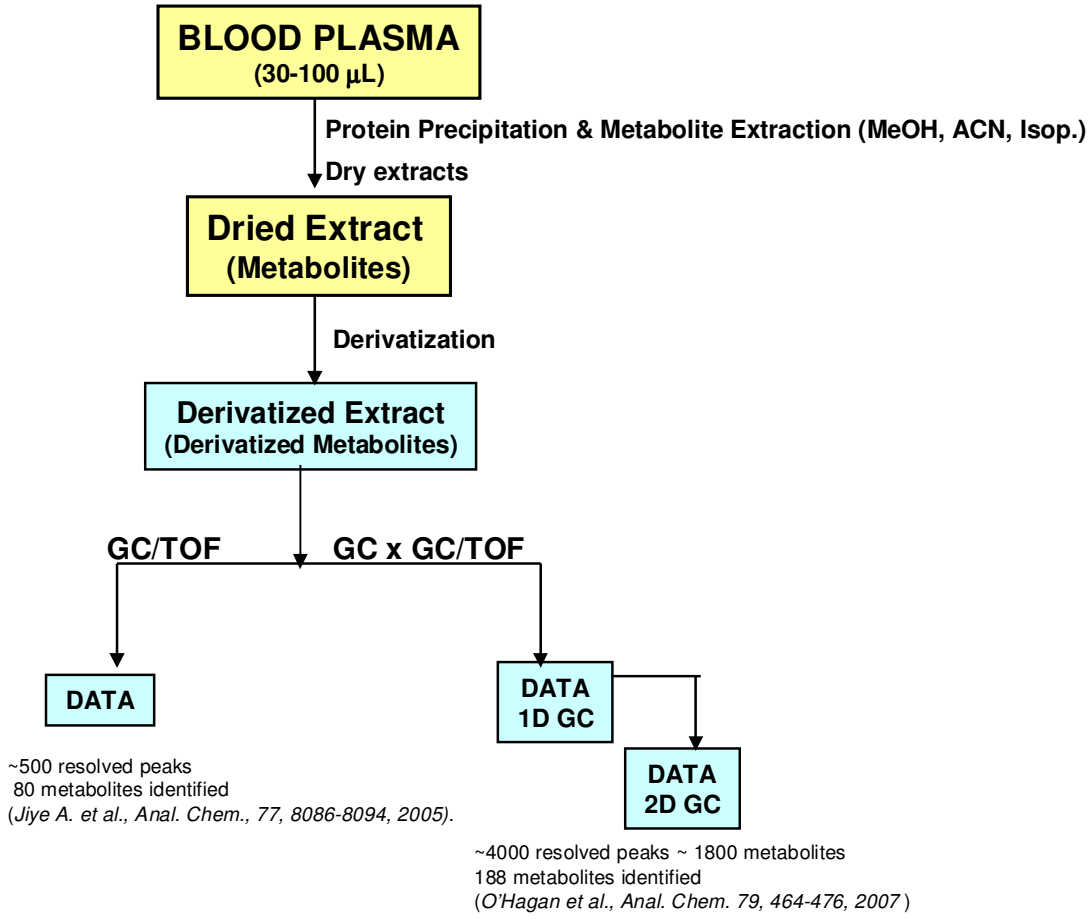


Figure 1-6: Multidimensional metabolomic analysis involving gas chromatography

The use of multidimensional methodology commonly referred to as two-dimensional chromatography or 2D chromatography (70, 71) employs either the use of two different column chemistries or two different properties of mobile phases. Multidimensional separation methods results in about 5-fold lower detection limits, a 10-fold increase in spectra purity and a 2 to 10 fold increase in the number of peaks as seen in a 2D-GC analysis of blood plasma shown in Figure 1-6 (28, 30).

1.7 Applications of Metabolomics

The present and future metabolomics applications can be put into three: personalized health care, molecular epidemiology, and uncovering biomarkers for drug discovery (Figure 1-7). In personalized health care, the metabolic profiles of an individual could be used to look for patient's susceptibilities to disease and their response to medication in order to promote healthy lifestyles and improve drug therapies (80).

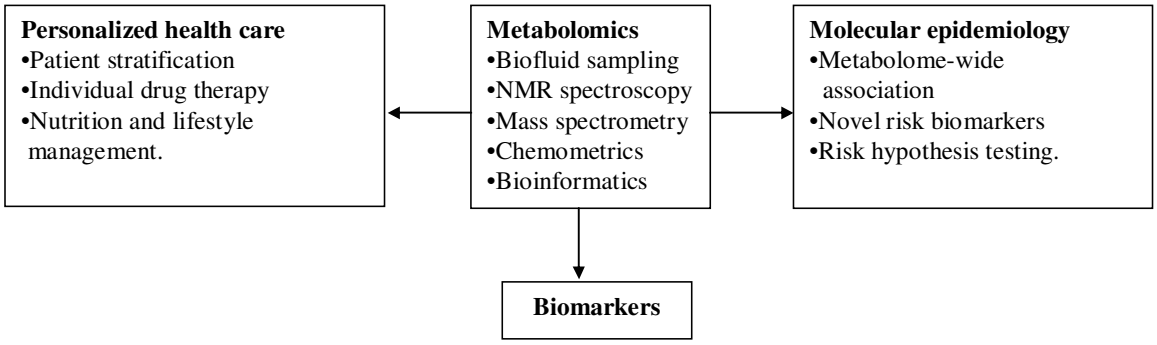


Figure 1-7: Applications of metabolomics

Metabolite screening or profiling of entire population groups in the research area of molecular epidemiology has the ability to work out susceptibilities of population groups to disease. The identification of biological targets and their association with identified biochemical pathways is and would lead to the discovery of new drugs (72). A number of metabolites have been identified as potential biomarkers for variety of diseases. For example in schizophrenia biomarkers found in cerebrospinal fluid include glucose, acetate, alanine and glutamine (73). The previously unrecognized role of formate in a series of interconnected renal ion exchanges involving chloride and sodium ions in the kidney via urine samples was discovered using NMR analysis of urine (22, 74, 75).

1.8 Blood Plasma Metabolome

The use of biological fluids such as blood plasma, cerebrospinal fluid, and saliva in physiological studies of organisms is not new. The circulation of these biological fluids helps to maintain constancy in the internal environments of organisms and as such abnormalities in an organism are reflected to a large extent in the altered composition of biological fluid. Changes in different metabolic pathways are reflected in changing levels of chemical compounds especially of low molecular mass organic compounds transported through the blood (30, 31).

The human blood is a complex fluid matrix of plasma (water, proteins, nutrients, inorganic ions and a large number of low molecular mass organic compounds covalently bounded to the proteins) and cellular components (Figure 1-8) (76).

The comprehensive analysis of blood plasma presents significant challenges. One such challenge is the precipitation of proteins and the extraction of metabolites. Of the many protocols that exist for protein precipitation, organic solvents miscible in water are the most widely used precipitation protocol.

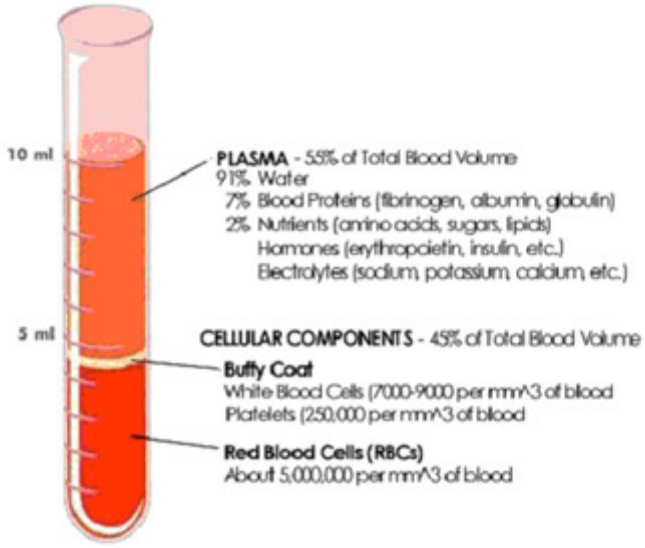


Figure 1-8: Composition of human blood (84)

The water miscible organic solvent does not only precipitate proteins but concurrently extract metabolites from plasma proteins (77). Different water miscible organic solvents extract different metabolites of different classes. No single organic solvent is known to comprehensively extract all metabolites in blood plasma. In spite of challenges, human blood plasma has been the subject of detailed chemical analyses for many years in clinical and biomedical research. Some clinical and biomedical studies in blood plasma metabolite profiling or metabolomics have involved the use of NMR spectroscopy (78),

LC-MS (79), GC-MS (71), 2D GC (70) and UPLC-MS (80) with varying degrees of success.

A GC-MS human blood plasma metabolomics study by Jiye and colleagues identified about 80 metabolites (71). A GC-MS and LC-MS combination study by Lawton and colleagues also reported the detection of more than 300 metabolic features of which 79 were positively identified (7). For reference purposes the concentration ranges of identified metabolites of normal and diseased state blood samples in comprehensive as well as target metabolite studies have been compiled (1, 6). The important role of blood plasma in an organism, availability of referential tables and easy accessibility to samples continue to make blood plasma one of the most studied biofluid including this study.

1.9 The Problem Statement

Monitoring of metabolite levels in blood offers the promise of new opportunities for diagnosis and treatment of diseases through the identification of biomarker compounds. Nevertheless, significant challenges remain in the comprehensive analysis of metabolites using current technologies. These difficulties have been attributed to:

1. the wide dynamic range of metabolite concentrations,
2. the diversity of physical and chemical properties of metabolites,
3. the limitations of current human metabolome database,
4. the inadequacy of any single analytical technique to provide complete coverage of all metabolites and metabolite classes

5. the need to improve analytical procedures to improve detection limits and selectivity of detection.

1.10 Scope of Work

This thesis work will focus on the analysis of metabolites in human blood plasma and their comparative separation and detection using GC-MS and LC-MS methods. Owing to the problem of potential overlaps of metabolites peaks in GC-MS analysis and the poor detection of sugar compounds by LC methods, it was proposed to separate plasma extracts into fractions using HILIC prior to analysis by derivatization GC-MS. It is proposed that fractionation by HILIC analysis prior to derivatization GC-MS analysis provide a useful multi-dimensional approach that would provide more intense, higher quality mass spectra. It is proposed to examine a single sample using: [1] derivatization GC-MS analysis, [2] HILIC LC-MS analysis and [3] derivatization GC-MS analysis of HILIC LC fractions prepared from human blood plasma.

1.11 Research Goals

While a number of analytical techniques have been used for metabolomic studies, no single method provides of itself a comprehensive profile of the metabolome. The dilemma is which from the metabolomics toolbox is the most appropriate given that all the methods have different characteristics. There is no simple answer but with pragmatism a choice can be made.

A primary goal of this thesis is to compare the abilities of GC-MS and LC-MS methods to separate and detect metabolites in blood serum. The specific objectives of this study are:

(a) To evaluate the suitability of quality control standards for the derivatization steps used prior to GC-EI-MS analysis

(b) To analyse human blood plasma extracts by derivatization GC-EI-MS analysis with these quality control standards included in the protocol

(c) To analyse human blood plasma extracts by HILIC LC-ESI-MS

(d) To collect fractions of human blood plasma following HILIC LC analysis and to analyze these fractions by derivatization GC-MS

(e) To expand an in-house GC-MS electron impact mass spectral library

(f) To develop an in-house searchable GC-MS mass spectral and retention index library

2. EXPERIMENTAL

2.1 Chemicals

The following odd number fatty acids, heptanoic acid (C7), nonanoic acid (C9), undecanoic acid (C11), tridecanoic acid (C13), pentadecanoic acid (C15), nonadecanoic acid (C19), tricosanoic acid (C23), heptacosanoic acid (C27), hentiacontanoic acid (C31) were purchased from Sigma-Aldrich (Oakville, ON, Canada) and used as retention index standard precursors. DL-4-Hydroxy-3-methoxymandelic acid, melatonin, dopamine hydrochloride and L-noradrenaline hydrochloride were provided by Dr. Ram Mishra from the department of Psychiatry, Behavioural & Neurosciences (McMaster University). O-Methylhydroxylamine hydrochloride and N-trimethylsily-N-methyltrifluoroacetamide (MSTFA) were purchased from Sigma-Aldrich (Milwaukee, WI, USA)

2.2 Gases and Solvents

High purity helium carrier gas (>99.99%) and Argon (>99.99%) were purchased from VitalAire (Hamilton, ON, Canada). HPLC grade solvents were purchased from Caledon Laboratories Ltd (Georgetown, ON, Canada).Anhydrous pyridine was purchased from Sigma-Aldrich (Milwaukee, WI, USA). Pure water was produced using a Milli-Q-Reagent Water System (Millipore Corp., Billerica, MA, USA)

2.3 Collection and Storage of Human Blood Plasma

Whole human blood used was drawn using the venepuncture procedure by technicians in McMaster Children's Hospital directly into K_3EDTA lavender-top tubes. Within 15 minutes of whole blood withdrawal, cells were separated from plasma by centrifugation at a speed of 3000 g and a temperature of 4 °C for 10 min. The separated human plasma was decanted into micro-centrifuge tubes and homogenized by vortexing for 10 seconds. 200 µL of the homogenized human blood plasma was then aliquotted into micro centrifuge tubes and the aliquots stored at -80°C for four weeks. When needed, samples were thawed in an ice bath for at least 30 minutes prior to use.

2.4 Procedures for the Preparation of Standards and Human Blood Plasma

2.4.1 Preparation of Derivatization Standards

Two chemical standards, a ketone (1,3-diphenylacetone) and an alcohol (9-anthracenemethanol) were used as derivatization controls. 9-anthracenemethanol was used to monitor the silylation step where as derivatized 1,3-diphenylacetone was used to monitor the process of methoximation reaction step.

The two chemical standards were dissolved individually in pyridine to afford stock solutions of 1000 ng/µL each. Aliquots of 75µL of 9-anthracenemethanol stock solution and 125µL of 1,3-diphenylacetone stock solution were combined with 300µL of pyridine to give a combined working stock solution for which 9-anthracenemethanol had a concentration of 150 ng/µL and 1, 3-diphenylacetone had a concentration of 250 ng/µL. Working with 10µL aliquot of the combined stock solution was used in each

derivatization, 10ng of 9-anthracenemethanol and 14ng of 1,3-diphenylacetone were injected each time (Table 2-1).

Table 2-1: Mass of derivatization control standards injected

Derivatization Standards	Stock Concentration of derivatization standards in hexane (ng/µL)	Volume Taken (µL)	concentration of injected derivative(ng/µL)
1,3-diphenylacetone	250	10	14
9-anthracenemethanol	150	10	10

2.4.2 Preparation of Internal Standards

Two deuterated PAH standards were used as internal standards in the experiments: acenaphthene-d_{10} and chrysene-d_{12}. Each standard was weighed individually and dissolved in hexane to afford solutions of each standard. The concentration of each standard is summarized in Table 2-2.

Table 2-2: Concentration of PAH stock solutions

Internal Standards	Stock Concentration of Internal standards (ng/µL)	Volume Taken (µL)	Concentration in mixture (ng/µL)
Acenaphthene-d10	5000	75	1500
Chrysene-d12	2000	25	1700

2.4.3 Preparation of Retention Index Standard

To aid in the identification of compounds in human blood plasma, the retention index (RI) approach was used. Nine odd saturated fatty acids ranging from C_7 to C_{31} with the following assigned retention indices were used: heptanoic acid TMS ester (C_7) = 700, nonanoi acid TMS ester (C_9) = 900, undecanoic acid TMS ester (C_{11}) = 1100, tridecanoic acid TMS ester (C13) = 1300, pentadecanoic acid TMS ester (C15) = 1500, nonadecanoic

acid TMS (C19) = 1900, tricosanoic acid TMS ester (C23) = 2300, heptacosanoic acid TMS ester (C27) = 2700, hentriacontanoic acid TMS ester (C31) = 3100.

Individually fatty acids were weighed and prepared as tetrahydrofuran (HPLC-grade THF) solutions. From the individual stock solutions, aliquots were taken for a composite fatty acid retention index solution with concentrations shown in Table 2-3.

Table 2-3: Preparation of Fatty Acid Retention Index Standard Solution

Fatty Acid Standards	Mass (g)	Volume (µL)	Stock Concentration of Fatty Acid Standard (ng/µL)	Volume Taken (µL)	Concentration in Mixture (ng/µL) (Final Volume = 1000 µL)
Heptanoic Acid (C_7)	0.01	1000	10000	60	600
Nonanoic Acid (C_9)	0.01	1000	10000	30	300
Undecanoic Acid (C_{11})	0.01	1000	10000	25	250
Tridecanoic Acid (C_{13})	0.01	1000	10000	25	250
Pentadecanoic Acid (C_{15})	0.01	1000	10000	25	250
Nonadecanoic Acid (C_{19})	0.01	1000	10000	40	400
Tricosanoic Acid (C_{23})	0.01	1000	10000	50	500
Heptacosanoic Acid (C_{27})	0.01	1000	10000	90	900
Hentriacontanoic Acid (C_{31})	0.01	1000	10000	200	2000

2.4.4 Preparation of Retention Index and Internal Standards Composite Solution

From the combined stock solution of the retention index standards, a 100µL aliquot was taken and placed in a 500µL vial. This was blown down to dryness using Nitrogen gas. An aliquot of 25µL of the combined PAH internal standards solution was then added and the final volume was made up to 350µL by adding 35µL of MSTFA and 290µL of hexane (Table 2-4). The vial was kept in a desiccator and stored in a 4°C refrigerator until needed.

Table 2.4: Concentrations of Retention Index & Internal Standards

Fatty Acid Standards	Stock Concentration of Fatty Acid (ng/µL)	Volume (µL) Taken	Concentration in Mixture (ng/µL) (Final volume =350 µL)
Heptanoic Acid (C_7)	600	100	171
Nonanoic Acid (C_9)	300	100	86
Undecanoic Acid (C_{11})	250	100	71
Tridecanoic Acid (C_{13})	250	100	71
Pentadecanoic Acid (C_{15})	250	100	71
Nonadecanoic Acid (C_{19})	400	100	114
Tricosanoic Acid (C_{23})	500	100	143
Heptacosanoic Acid (C_{27})	900	100	257
Hentriacontanoic Acid (C_{31})	2000	100	571
Internal Standards	**Stock Concentration of Internal standards (ng/µL)**	**Volume (µL) Taken**	**Concentration in mixture (ng/µL) Final volume =350 µL**
Acenaphthene-d_{10}	1500	25	107
Chrysene-d_{12}	1700	25	121

2.4.5 Preparation of Recovery Standards

The ability of methanol and water to comprehensively extract metabolites of different compound classes from human blood plasma was tested. Six standard metabolite compounds, representing six compound classes were used as recovery compounds (Table 2-5). The six standard compounds were made up of three reference compounds which are not seen in human blood plasma and three stable isotope-labeled standards.

Table 2-5: List of recovery standards

Compound Name	Compound Class
Phenylalanine-d_8	Amino acid
m-Tyramine	Amine
Stigmasterol	Sterol
$^{13}C_6$ Glucose	Sugar
Octadecanoic acid-d_{35}	Fatty acid
Syringic acid	Organic acid

Stock solutions of the individual compounds were prepared in suitable solvents (Table 2-6). Two composite solutions, composite A and B, were made using aliquots of the individual stock solutions (Tables 2-7 and 2-8). Composite solution A was made up of aliquots from m-tyramine, stigmasterol, stearic acid and syringic acid in methanol while that of composite B was made up of phenylalanine-d_8 and $^{13}C_6$ glucose in water. The final concentrations of injected derivatized metabolites are recorded in Table 2-9.

Table 2-6: Preparation of Recovery Standards

Compound Name	Compound class	Mass (ng)	Final Volume (µL)	Solvent	Stock Concentration (ng/µL)
Phenylalanine-d_8	Amino acid	0.01	1000	H_2O	10000
m-Tyramine	Amine	0.01	1000	MeOH	10000
Stigmasterol	Sterol	0.01	1000	THF	10000
$^{13}C_6$ Glucose	Sugar	0.01	1000	H_2O	10000
Octadecanoic acid-d_{35}	Fatty acid	0.01	1000	THF	10000
Syringic acid	Organic acid	0.01	1000	MeOH	10000

Table 2-7: Concentration of composite solution A

Compound Name	Compound class	Stock Concentration (ng/µL)	Aliquot from stock (µL)	Final Volume (µL)	Composite Concentration (ng/µL)
m-Tyramine	Amine	10000	40	1000	400
Stigmasterol	Sterol	10000	50	1000	500
Octadecanoic acid-d_{35}	Fatty acid	10000	50	1000	500
Syringic acid	Organic acid	10000	50	1000	500

Table 2-8: Concentration of composite solution B

Compound Name	Compound class	Stock Concentration (ng/µL)	Aliquot from stock (µL)	Final Volume (µL)	Composite Concentration (ng/µL)
$^{13}C_6$ Glucose	Sugar	10000	30	1000	300
Phenylalanine-d_8	Amino acid	10000	50	1000	500

Table 2-9: Concentration of recovery standards injected

Compound Name	Compound class	Composite Concentration (ng/µL)	Aliquot from composite mixture(µL)	Concentration of injected derivative(ng/µL)
m-Tyramine	Amine	400	5	27
Stigmasterol	Sterol	500	5	15
Octadecanoic acid-d_{35}	Fatty acid	500	5	16
Syringic acid	Organic acid	500	5	23
$^{13}C_6$ Glucose	Sugar	300	5	24
Phenylalanine-d_8	Amino acid	500	5	24

2.5 Experimental Design for Recovery Standards

The experiment to test the extraction ability of methanol and water was done in three parts. The total solvent volume for each experiment was fixed at 190µL.

2.5.1 Experimental Design for Underivatized Recovery Standards

The experimental setup shown in Figure 2-1 is used to obtain the mass spectra of underivatized recovery standards. A 5 µL aliquot of composite solution A was added to 185 µL of hexane which was substituted for MSTFA. The solution was mixed for 10 sec. and 1 µL injected on the GC-MS.

2.5.2 Experimental Design for Derivatized Recovery Standards

The experimental setup shown in Figure 2-2 was used for recovering derivatized standards. There is minimum sample handling and the recovery here represents 100% recovery. A 5 µL aliquot of composite solution A was added to 5 µL of composite solution B. A 10 µL aliquot of O-methyl hydroxylamine hydrochloride ($MeONH_2$.HCL) in pyridine and 10 µL of derivatization control mixture (9-anthracenemethanol and 1,3-diphenylacetone) were then added. The solution was mixed by vortexing for 10 s and derivatized at 37°C for 90 min.

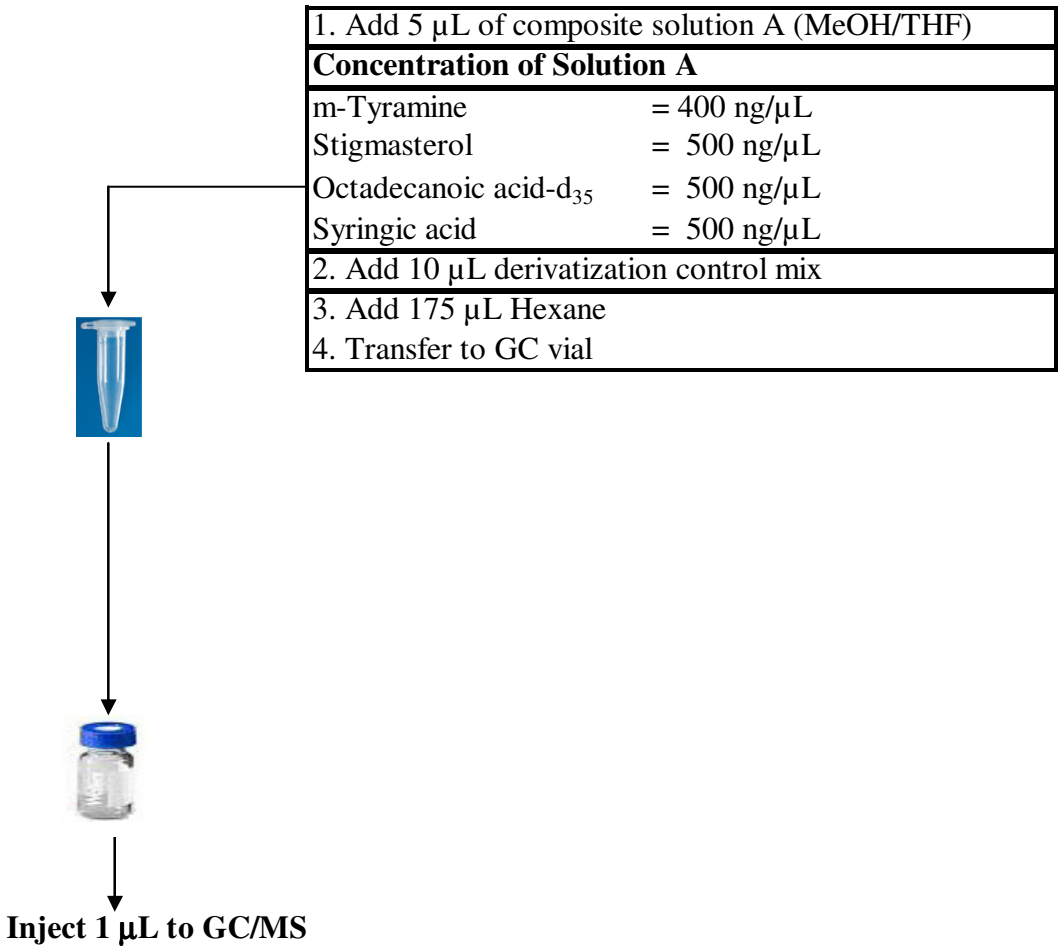


Figure 2-1: Experimental flow diagram for underivatized recovery standards

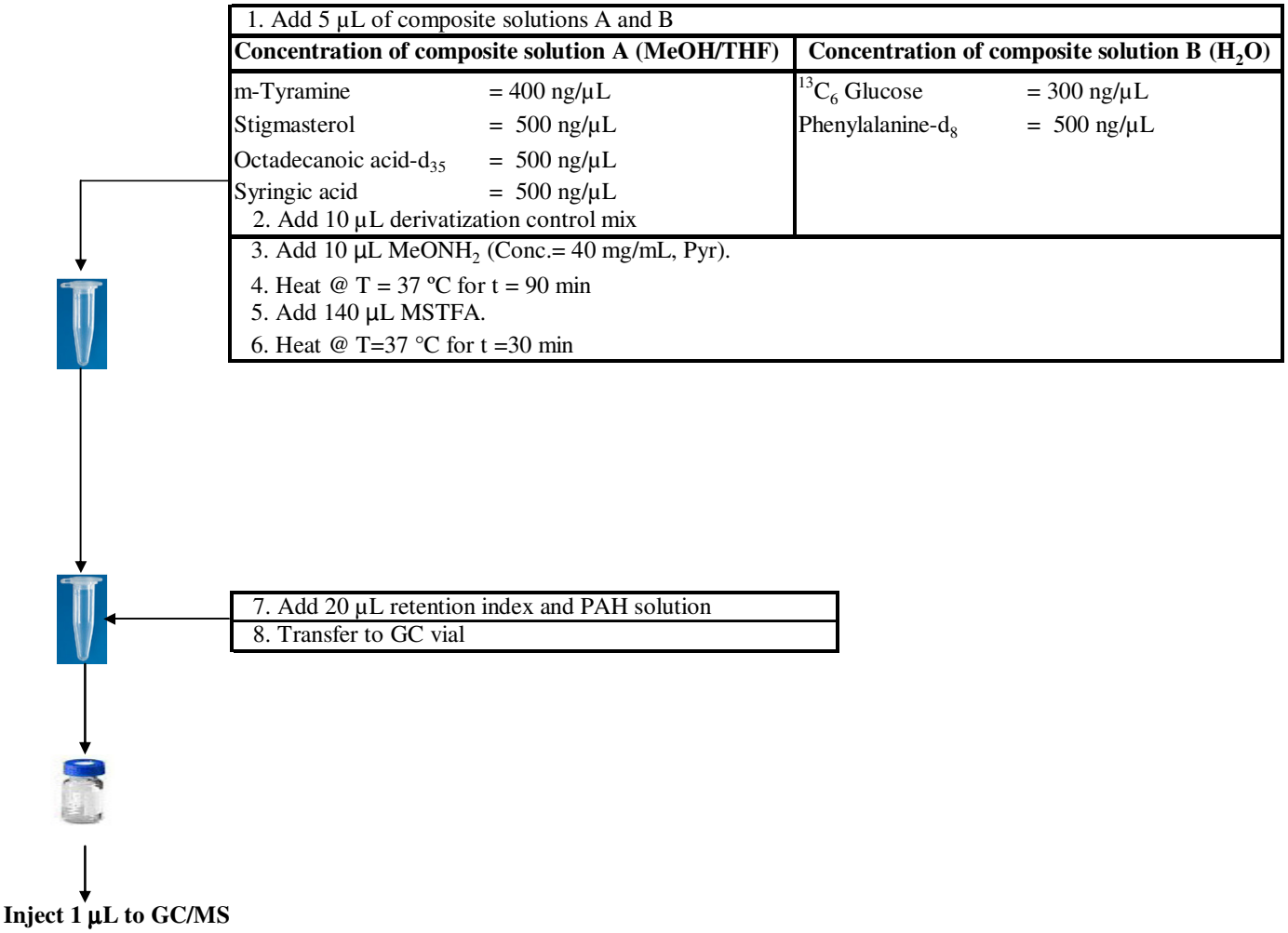


Figure 2-2: Experimental flow diagram for derivatized recovery standards

The resulting oximated solution was then silylated with 150 µL of MSTFA at 37°C for 30 min. A 20 µL aliquot solution of retention index and PAH solution was added and 1 µL injected on the GC-MS.

2.5.3 Experimental Design for Derivatized Recovery Standards in Method Blank

The experimental setup shown in Figure 2-3 represents the extraction of recovery standards in: method blank where the sample has been replaced with water and human blood plasma samples.

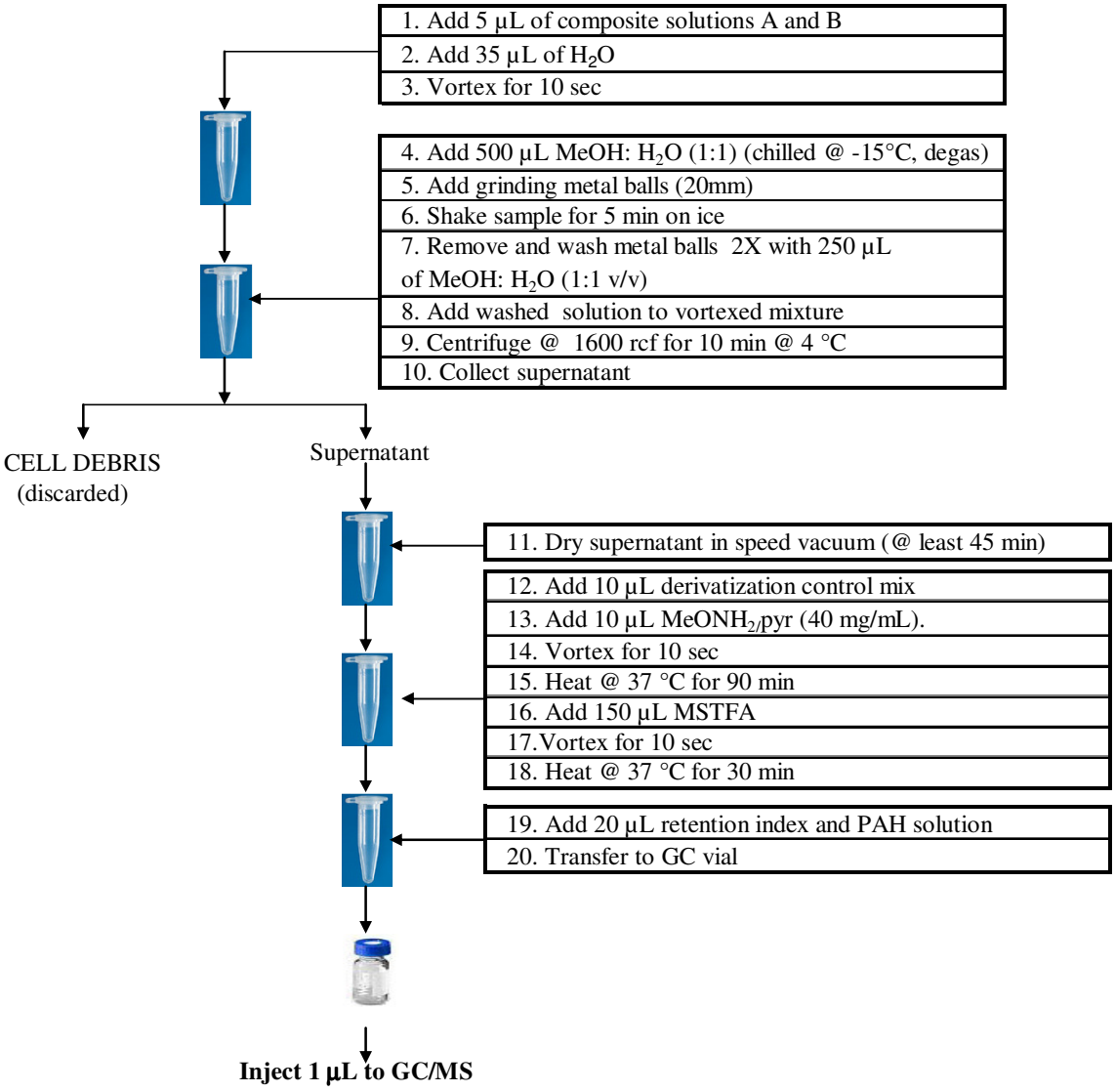


Figure 2-3: Experimental flow diagram for extraction and derivatization of recovery standards in method blank and human blood plasma sample.

For the method blank, 5 µL aliquots of both composite A and B was added to 30 µL of H_2O and vortexed mixed for 10 s. To the mixed solution was added 500 µL of degassed chilled methanol and water (1:1 v/v). Three grinding metal balls are added and solution shaken on ice for 5 min.

The three grinding metal balls are then removed and washed with two parts of 250 µL of methanol and water (1:1 v/v). The washed solution was added to the shaken solution and centrifuged at 1600 rcf for 10 min at 4°C.

The supernatant is collected and dried in a speed vacuum for at least 45 min. Methoxymation was carried out at 37° C for 90 min. The standards were subsequently trimethylsilylated at 37°C for 30min. After adding 20 µL of retention index and PAH solution, 1 µL of the derivatized standard sample was analyzed by GC-MS.

The experimental flow diagram in Figure 2-3 was also used to test for the recovery of standard compounds in composite solutions A and B in human blood plasma samples. A 200 µL aliquot of frozen human blood plasma was removed from -80° C storage and thawed on ice for about 45 min. A 30 µL aliquot of the thawed human blood plasma was added to a 10 µL mixture of recovery standard (made up of 5 µL aliquot of composite solution A and 5 µL aliquot of composite solution B) and vortexed mixed for 10 s. The recovery standards are then extracted, centrifuged, derivatized and inject on GC-MS as outlined in the experimental flow diagram in Figure 2-3.

2.6 Preparation of Metabolite Standards

Thirty eight metabolite standards that includes compounds from catecholamine, lipid, and fatty acid compound classes were weighed individually using a six-place balance (µg) and dissolved in distilled water or suitable organic solvent depending on the polarity of the standard. All solvents are deoxygenated by bubbling ultra pure argon through it. Each standard solution is sealed in 1mL vials and stored in a -20°C freezer. When needed, a vial of frozen metabolite stock solution was thawed on an ice bath. Each solution is thawed and refrozen no more than 4 times before it is discarded. The concentration of metabolite standards and the solvent used is listed in Table 2-6.

2.7 Preparation of Human Blood Plasma Metabolites

Aliquots of 200 µL frozen human blood plasma are taken out and made to thaw on ice for 45 minutes and vortex-mixed for 10 sec before use. Five 35 µL aliquots of human blood plasma were taken from the thawed 200 µL human blood plasma into micro centrifuge tubes. Using a modified Oliver Fiehn's human blood plasma metabolite extraction and derivatization method (77), 500 µL of a degassed and chilled methanol: water (1:1 v/v at -20°C) solvent mixture is added to each 35 µL human blood plasma aliquot. 500 ng/µL of Stigmasterol to be used as a recovery and quantitation standard is also added to each 35 µL human blood plasma aliquot and vortexed vigorously for 5 minutes. Subsequent steps in the extraction of metabolites from human blood plasma are shown in the method flow diagram (Figure 2-3). Remaining of the 200 µL thawed human blood plasma sample after drawing five aliquots have been taken, is discarded.

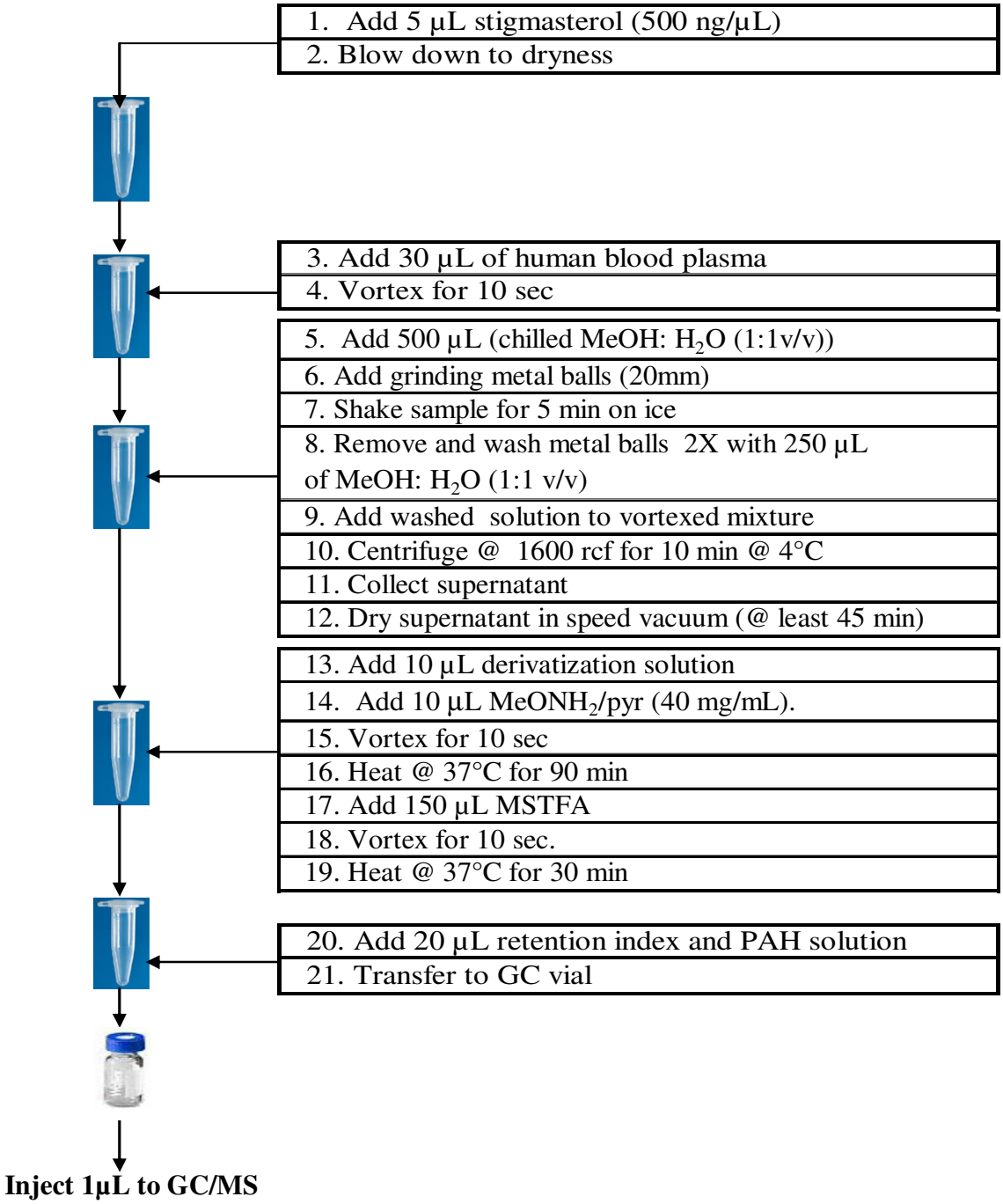


Figure 2-4: Experimental flow diagram for the extraction and derivatized of metabolites in human blood plasma.

The injection of 1 µL derivatized human blood plasma corresponded to 0.15 µL blood plasma equivalent.

2.8 Extraction of Metabolite for Multidimensional Analysis

This approach involves the preparation of human blood plasma extracts for routine LC-MS and GC-MS analysis and a GC-MS analysis of LC fractions of human blood plasma extract. The experimental flow diagram is shown in Figure 2-5.

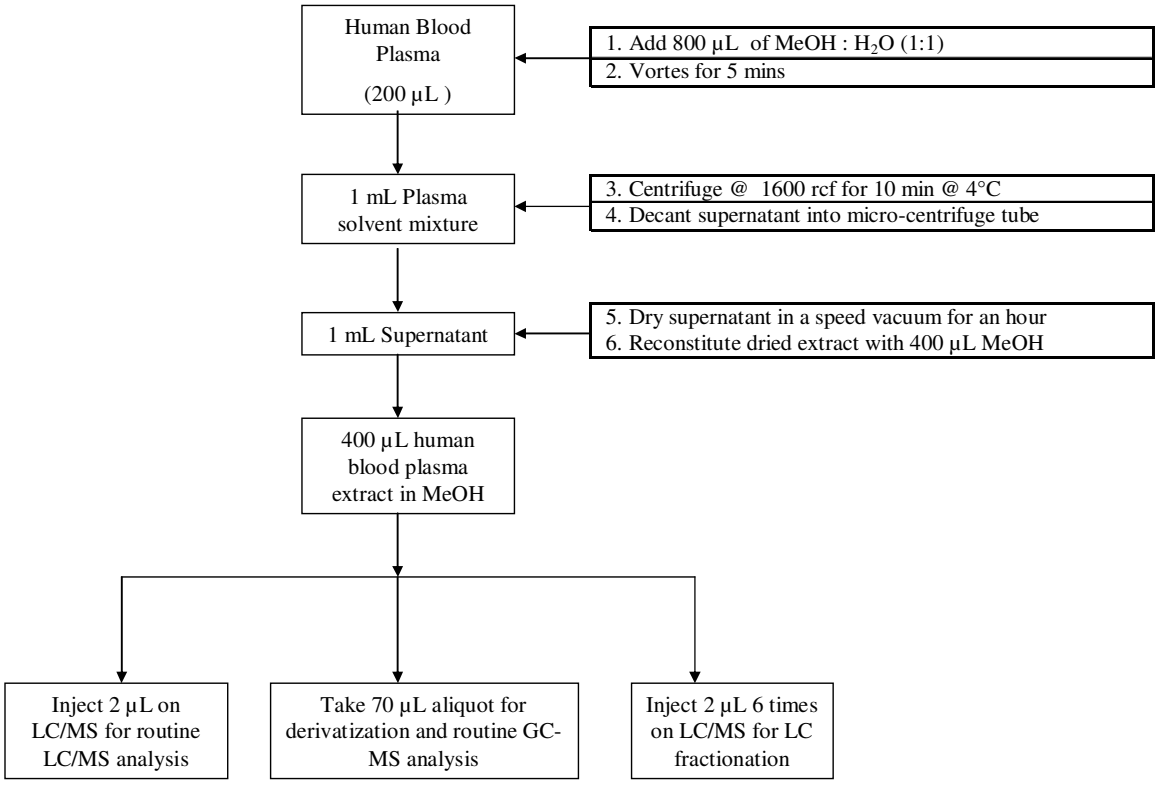


Figure 2-5: Experimental flow diagram for the multidimensional analysis of human blood plasma sample using GC/MS and LC/MS.

A 200 µL human blood plasma aliquot at -80 °C is thawed on ice for 45 min. 800 µL of a mixture of MeOH:H2O (1:1 v/v) is added and the mixture vortex mixed for 5 min. The resulting 1 mL plasma solvent mixture is centrifuged at 1600 rcf at a temperature of 4 °C for 10 min. The supernatant is decanted and dried in a speed vacuum for one hour. The dried blood plasma extract is reconstituted with 400 µL of MeOH. Aliquots are then taken for the various multidimensional analyses.

2.9 Derivatization Procedures

2.9.1 Derivatization of Metabolite Standards

A 5 µL aliquot solution of freshly prepared individual metabolite standards is put in a micro- centrifuge tube and evaporated to dryness (takes an average of 10–30 min depending on solvent of dissolution) with nitrogen gas. An aliquot solution of O-methyl hydroxylamine hydrochloride in pyridine (20 µL, 40 mg/mL) is added to the dried metabolite standard. The resulting solution is heated at 37°C for 90 min in an aluminum block heated. After 90 min, an aliquot solution of neat MSTFA (150 µL) is added and the solution heated at 37°C for 30 min. A 20 µL aliquot of retention index and PAH solution is added to the O-methyl oximated and trimethylsilylated metabolite standard. A 10 µL aliquot is then transferred to a washed GC vial and 1 µL of that is injected on the GC/MS. The micro-centrifuge tube containing the remaining derivatized metabolite solution is sealed with Teflon tape, placed in a desiccator and desiccator stored at 4°C for three days, in the event that a second analysis is needed.

2.9.2 Derivatization of Human Blood Plasma Extracts

A 70 µL aliquot of the reconstituted extract is taken for a GC/MS "whole" sample analysis (Figure 2-5). This 70 µL aliquot is equivalent to the initial 35 µL aliquot used for GC/MS analysis (Figure 2-4). The 70 µL extract is dried in a speed vacuum, followed by the addition of 10 µL of derivatization standard and 10 µL of $MeONH_2 \cdot HCL$/pyridine (40 mg/mL) and vortex mixed for 10 sec. The reaction mixture is heated at 37 °C for 90 min. Then, neat MSFTA (150 µL) was added and the reaction heated for an additional 30 minutes at 37 °C. A 10 µL aliquot is transferred to a GC vial and 1 µL injected on a GC/MS. The injection of 1 µL derivatized mixture corresponded to 0.15 µL blood plasma equivalent sample (see Figure 2-3).

2.10 GC – MS Analysis

2.10.1 Instrumentation

A Hewlett-Packard Model 5890 Series II gas chromatography (Hewlett-Packard, Palo Alto, CA, USA) coupled to a Hewlett-Packard Model 5971A Mass Selective Detector (Hewlett-Packard, Palo Alto, CA, USA) was used to analyse all samples. Analyses were performed in full scan mode using electron impact (EI) ionization with helium as the carrier gas. A DB-17ht capillary column (50% phenyl/50% methyl silicone, 30 m length X 0.25 mm i.d. X 0.15 µm film, J & W Scientific, Folsom, CA) was used for all sample analyses. A 5m retention gap (deactivated fused silica, 5 m X 0.53 mm, Chromatographic Specialties Inc., Brockville, ON) was placed on the front end of the column and connected to the DB-17ht column via a 2-way glass union (Chromatographic

Specialties). Table 2.7 summarizes the column properties and temperature program used for the analyses.

Table 2-10: Column and temperature program for GC-MS analysis.

Column Properties	
Column Type	J & W DB-17ht
Stationary Phase	50% phenyl/50% methyl silicone
Column Length (m)	30
Column I.D. (mm)	0.25
Film Thickness (µm)	0.15
Carrier Gas	Helium
Flow Rate (mL/min)	0.958
Oven Temperature Program	
Initial Oven Temperature (°C)	50
Hold Time (min) @ 50°C	5
Temperature Program Rate (°C/min)	5
Solvent Delay Time on MS (min)	10
Final Oven Temperature (°C)	300
Final Hold Time (min) @ 300°C	5
Total Run Time (min)	60

2.10.2 Monitoring Performance of GC Column

To monitor the deterioration of the GC column, a 200 pg/µL PAH calibration standard was injected prior to the first injection of derivatized human blood plasma extract and randomly during sample sets to check column performance. Using the same temperature programme, the peak widths of six 252 amu PAHs, namely, benzo(b)fluoranthene, benzo(k)fluoranthene, benzo(j)fluoranthene, benzo(e)pyrene, benzo(a)pyrene and perylene, in the PAH calibration standard were monitored. Typically, an average peak width below 0.040 minutes of the six 252 amu PAHs indicated a well performing column. A 50% increase in the average peak width of the PAH's suggested a deteriorated GC column that would result in poor chromatography. To

continue analyzing derivatized human blood plasma extracts, about 50 cm of the retention gap was cut to restore good chromatography.

2.11 LC-MS Analysis

Without any further sample treatment, a 2 µL of the reconstituted extract is injected on a LC/MS. This LC /MS injection corresponded to 1.0 µL blood plasma equivalent sample.

2.11.1 Instrumentation

An Agilent 1200 RR "D" series Liquid chromatography coupled to Bruker micro TOF II ESI-TOF-MS was used for routine analysis.

Table 2-11: Column properties for routine LC analysis

Column Properties	
Column Type	Luna NH2 100A
Column Length (mm)	50
Column I.D. (mm)	2
Particle Size (µ)	3
Solvent Elution	
Solvent Program	Gradient
Solvent Type	Acetonitrile : Water (95:5) Acetonitrile : Water (55:45)
Flow Rate (µL/min)	200
Column Temperature (°C)	Room Temperature

2.11.2 LC Fractionation Collection System

A Varian Prostar 210 solvent delivery liquid chromatograph Model 330 coupled to a Varian 1200 L triple Quadrupole ESI-MS/MS (Varian Inc., Walnut Creek, CA, USA) was used to collect liquid chromatographic fractions of human blood plasma samples. Six injections of 2 µL blood plasma equivalent samples were made for the fractionation. Samples were fractionated by a 3 micron Luna 100A Aminopropyl column (150 mm length x 2.1 mm i.d., Phenomenex Inc., Torrance, CA, USA).

Table 2-12: Column properties for LC fractionation

Column Properties for Fractions	
Column Type	Luna NH2 100A
Column Length (cm)	15
Column I.D. (mm)	2.1
Particle Size (µ)	3
Solvent Elution	
Solvent Program	Isocratic
Solvent Type	Acetonitrile : Water (80:20)
Flow Rate (µL/min)	300
Column Temperature(°C)	Room Temperature

LC fractions were collected every 30 seconds (0.5 min) for the first thirteen fractions and there after every 60 seconds (1.0 min) for the 14^{th} to the 16^{th} fraction. Fractions were collected separately into test tube labelled 1 to 16 for each 2 µL injection on the LC system. The fractions were pooled together according to test tube numbers. Each pooled test tube volume was 500 µL.

2.11.3 Derivatization and GC-MS Analysis of LC Fractions

All pooled fractions were evaporated to dryness in a speed vacuum for at least 2 hours. A solution of $MeONH_2 \cdot HCL$ in pyridine (10 µL of 40 mg/mL) and 10 µL of derivatization control solution were added to each dried extract vortex mixed for 10 seconds and allowed to react at 37 °C for 90 minutes. 150 µL of neat MSTFA was added and the reaction mixtures heated for 30 minutes at 37 °C. 20 µL of retention index and PAH solution was added. 10 µL portions of all resulting derivatized fractions were transferred to GC vials and 1 µL was injected onto the GC/MS. Each GC/MS injection corresponded to 0.32 µL equivalent of human blood plasma sample.

3. ANALYTICAL METHOD DEVELOPMENT

Currently, there is no optimized protocol for the metabolic profiling of blood plasma samples using GC/MS. Experimental parameters such as solvent for deproteinazation/extraction of metabolites; temperature and the duration of derivatization before GC/MS analysis differ in publications. The two most widely used protocols are shown in Table 3-1.

Table 3-1: Two widely used GC-MS blood plasma protocols

Controlling Parameter	Fiehn et al. 2005 & 2009	Jiye et al. 2005
Protein precipitation and extraction solvent	Isopropanol:acetone	Methanol:water
Oximation Reagent	Methoxyamine·HCl	Methoxyamine·HCl
Oximation Reaction Temperature	37° C	Room Temperature
Oximation Duration (Hr.)	1.5	16
Silylation Reagent	MSTFA	MSTFA+1% TMCS
Silylation Reaction Temperature	37 °C	Room Temperature
Silylation Reaction Time (Hr.)	0.5	1.0

The differences in the two methods could be attributed to apparent losses of metabolites during extraction and also the uncertainties in the completion of derivatization reactions. There is the need therefore to the monitoring of losses due to extraction and derivatization of metabolites and to ensure reproducibility of metabolites. One way of monitoring the metabolite extraction process and the reaction steps before GC-MS is the addition of

standard compounds to assess the recovery and completeness of all adduct formation reactions before injecting samples on GC/MS. To do this, standard compounds of same chemical nature were picked from a selected group of compounds to be used as monitoring controls for extraction and reaction of metabolites.

3.1 Quality Control Standards

3.1.1 Addition of Recovery Standards

Six standard compounds (Table 3-2) representing six compound classes present in human blood plasma were introduced into the experimental method to monitor recoveries due to losses in extraction and handling of samples. The six standard compounds were studied in a series of designed metabolite recovery experiments (see experimental procedures in Figures 2-1, 2-2 and 2-3).

Table 3-2: List of compounds used as recovery standards

Compound Class	Name of Compound
Amino Acid	Phenylalanine-d_8
Amine	m-Tyramine
Sterol	Stigmasterol
Sugar	$^{13}C_6$ Glucose
Fatty Acid	Stearic Acid-d_{35}
Organic Acid	syringic Acid

Using a modified Fiehn et al. 2005 blood plasma protocol, the six standard compounds were first analyzed as underivatized compounds (see procedure in Figure 2-1) to obtain the underivatized mass spectrum of each standard. The underivatized mass spectrum was

used to monitor the degree of derivitization. Secondly, without drying the six standards in solution, the compounds were then taken through methoximation and silylation derivatization reactions (see procedure in Figure 2-2). Omitting the drying step before derivatization was to help in the reduction of losses and to maximize the recovery of standards. Thirdly, the experimental procedure outlined in Figure 2-3 was used in the recovery evaluation of added standards to water sample and blood plasma sample. Water sample (referred to as blank method procedure) and blood plasma samples were used to assess matrix effect introduced by samples. An internal standard (acenapthene-d_{10}) was used to normalize all derivatized selected trace model ions of the six standards.

The normalized peak areas of standards in the method blank and in human blood plasma were compared to the normalized peak areas of standards in the derivatized standards experimental procedure with the omitted drying step in Figure 2-2. The mean percentage recoveries of the six standard compounds are listed in Table 3-3.

The recoveries obtained for the standards in blank method (95% to 117%) and sample method (82% to 112%) protocols are high with standard deviations ± 11 and below. This is an indication that losses due to handling of standards are negligible. With a *P* value larger than 0.05, the difference between the blank method and sample method is statistically not different and the effect of sample matrix on the extraction of standards is negligible.

Table 3-3: Mean percent recoveries of standards representing six compound classes in a blank sample and blood plasma method of analysis.

Compound Class	Compound Name	Model ion (m/z)	Percentage of derivative recovered from blank sample Mean ± RSD (n=3)	Percentage of derivative recovered from blood plasma Mean ± RSD (n=3)
Amino Acid	Phenylalanine-d_8 $(TMS)_2$	274	107 ± 8.0	112 ± 7.1
Amine	m-Tyramine$(TMS)_2$	174	95 ± 6.0	94 ± 8.3
Sterol	Stigmasterol TMS	129	100 ± 5.0	100 ± 6.9
Sugar	$^{13}C_6$ Glucose MeOX1 $(TMS)_5$	323	119 ± 6.0	82 ± 7.6
Sugar	$^{13}C_6$ Glucose MeOX2 $(TMS)_5$	323	117 ± 10.0	111 ± 2.1
Fatty Acid	Myristic Acid-d_{27} TMS	312	114 ± 2.0	103 ± 2.6
Organic acid	Syringic Acid $(TMS)_2$	342	104 ± 11.3	82 ± 7.8

3.1.2 Derivatization Control: Derivatizing with MSTFA

In GC analysis, analytes need to be volatile and thermally stable. Non-volatile compounds like amino acids, sugars and organic acids are therefore converted into volatile chemical derivatives. Silylation is one of the simplest and most widely used reaction steps to increase the volatility of compounds. However, without an oximation reaction step before silylation reactions in GC/MS analysis of sugar compounds such as fructose and ribose, the silylation reaction results in the formation of open chain and cyclic compounds adding to the number and complexity of resolved peaks. In Figure 3-12, the reaction of fructose with MSTFA results in the formation of four compounds (peaks at 25.90 min., 26.20 min., 26.48 min. and 26.58min.). To prevent the sugar ring cyclization prior to silylation, oxime derivatives of sugar containing compounds are prepared before the silylation reaction step with MSTFA. Both silylation and oximation

reactions are temperature and time dependant and must be monitored to complexities in GC-MS analysis of metabolites.

3.1.2.1 Addition of 9-Anthracenemethanol

9-Anthracenemethanol was selected as a silylation monitoring standard by a former graduate student in the group (81).

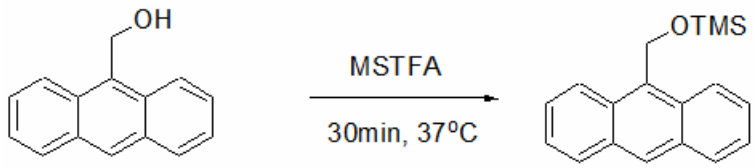
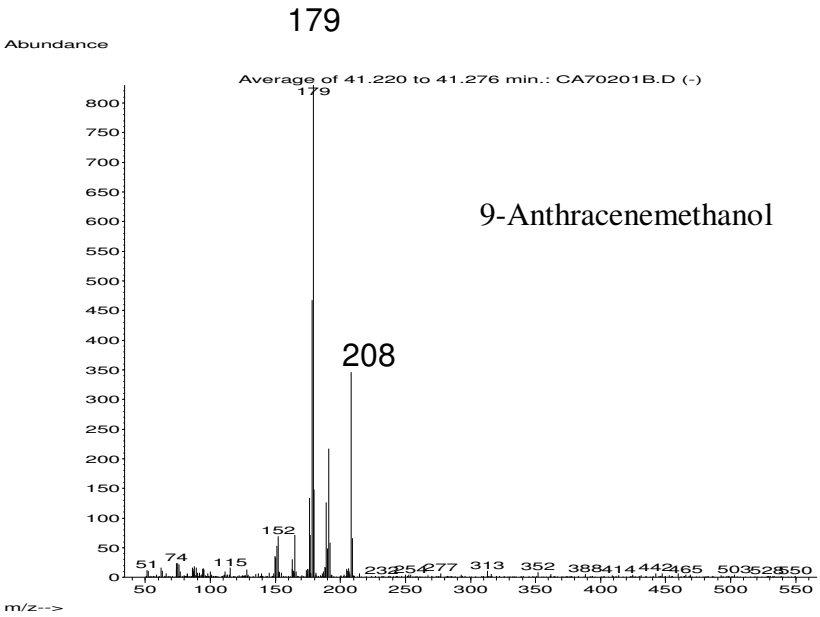


Figure 3-1: Conversion of 9-Anthracenemethanol to its TMS derivative with MSTFA

The underivatized alcohol and its trimethylsilyl derivative are separated on the GC column and their mass spectra are different (Figures 3.1 and 3.2). The detection of the underivatized form (9-anthracenemethanol) at the end of a silylation reaction shows an incomplete silylation reaction. 9-Anthracenemethanol was added to all reaction mixtures to determine the completeness of the silylation reaction step.

(A)



(B)

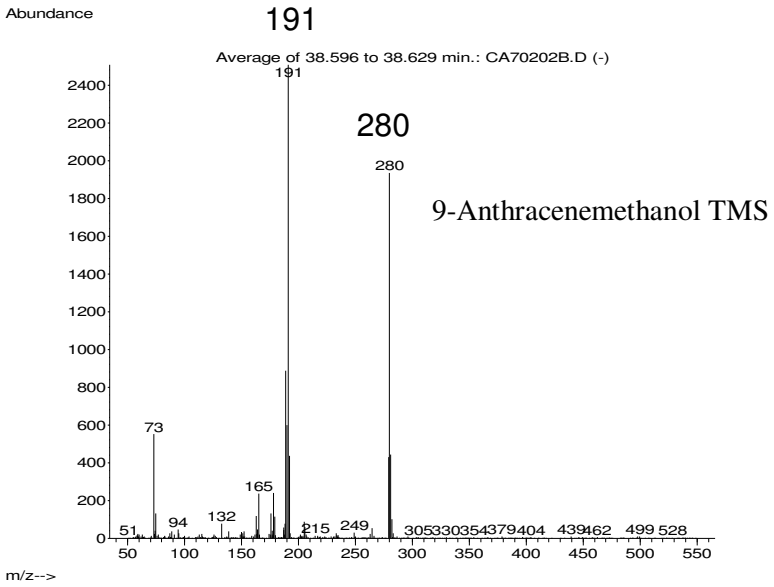


Figure 3-2: Mass spectrum of: (A) underivatized 9-anthracenemethanol and (B) trimethylsilyl ether of 9-anthracenemethanol.

3.1.3 Derivatizing with $MeONH_2 \cdot HCL$

3.1.3.1 Addition of 1,3-diphenylacetone

In selecting a standard compound to monitor the oximation reaction step, the following five ketones and one aldehyde were studied (Table 3-4).

Table 3-4: Properties of Selected Ketones and an Aldehyde

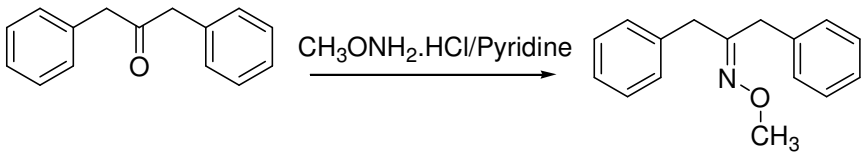
Name	Compound Class	Chemical Formula	Molecular weight	Boiling Point (°C)
6-Undecanone	Ketone	$C_{11}H_{22}O$	170	228
Isobutyrophenone	Ketone	$C_{20}H_{12}O$	148	217
2-Naphthaldehyde	Aldehyde	$C_{11}H_{10}O$	156	155
Benzophenone	Ketone	$C_{13}H_{10}O$	182	305
4,4'-dichlorobenzophenone	Ketone	$C_{13}H_8Cl_2O$	251	353
1,3-diphenylacetone	Ketone	$C_{15}H_{14}O$	210	330

The criteria for selecting an oxime control for derivatization from the list in Table 3-3 are:

A. The underivatized and derivatized compounds are both detectable by GC-MS.
B. Standard gives only one derivative and derivatization is completed within the reaction time specified in protocol.

3.1.3.2 The Detection of Underivatized and Derivatized Standard Compounds.

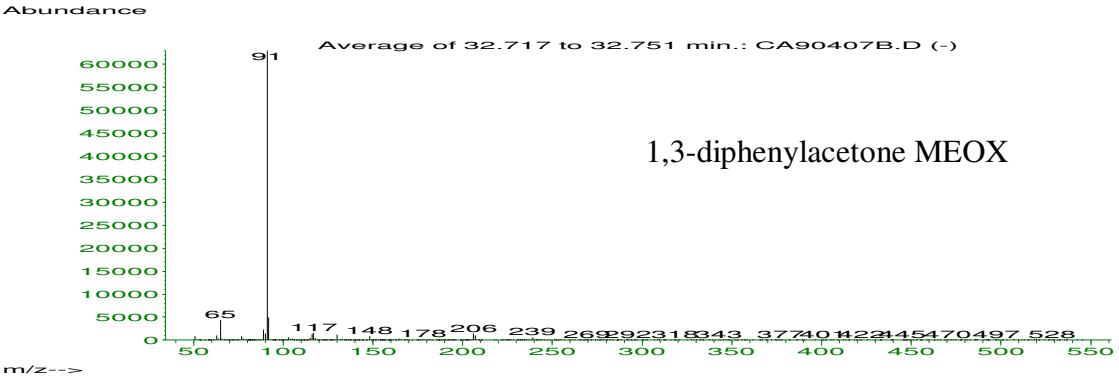
Using the derivatization protocol for standard metabolite compounds (Figure 2-4), mass spectra of the underivatized and derivatized standards of all six standards listed in Table 3-4, was detected by GC-MS. The methoximation reactions of three examples of standards listed in Table 3-4 are shown with their respective mass chromatograms in Figures 3-3 to 3-8.



(a) (b)

Figure 3-3: 1,3-Diphenylacetone O-methyl oxime reaction showing (a) 1,3-diphenylacetone and (b) derivatized 1,3-diphenylacetone.

(a)



(b)

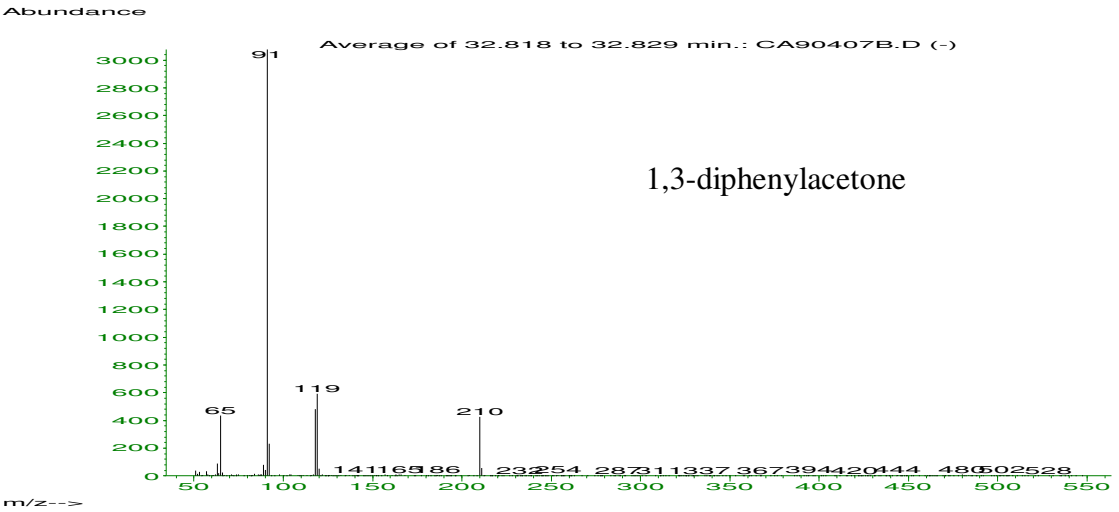
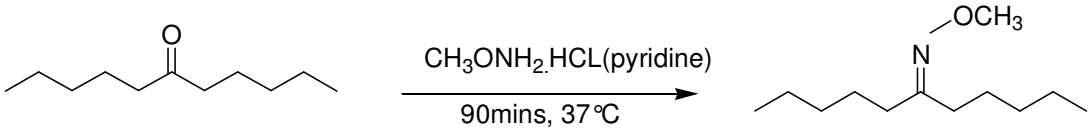


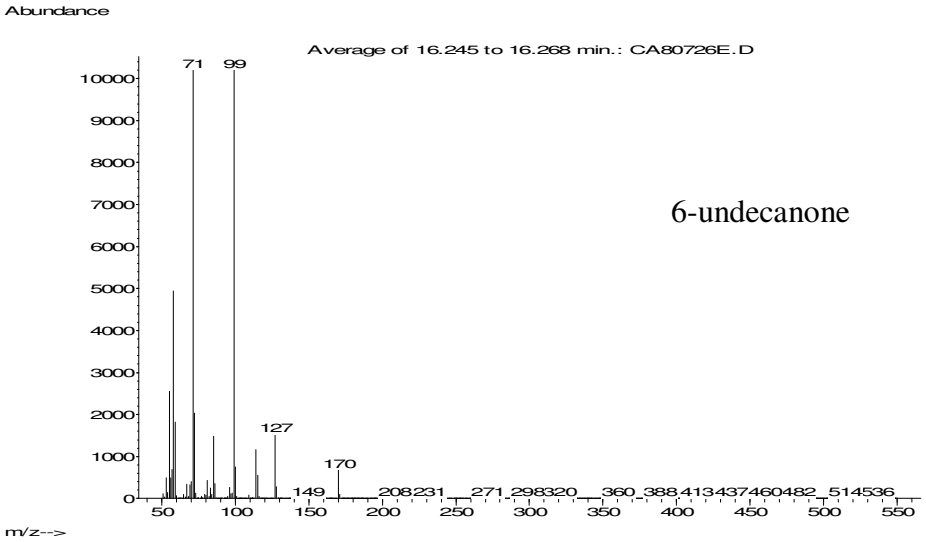
Figure 3-4: Mass spectra of (a) 1,3-diphenylacetone O-methyl oxime and (b) underivatized 1,3-diphenylacetone.



(a) (b)

Figure 3-5: 6-undecanone methoxymation reaction showing (a) 6-undecanone and (b) derivatized 6-undecanone

(a)



(b)

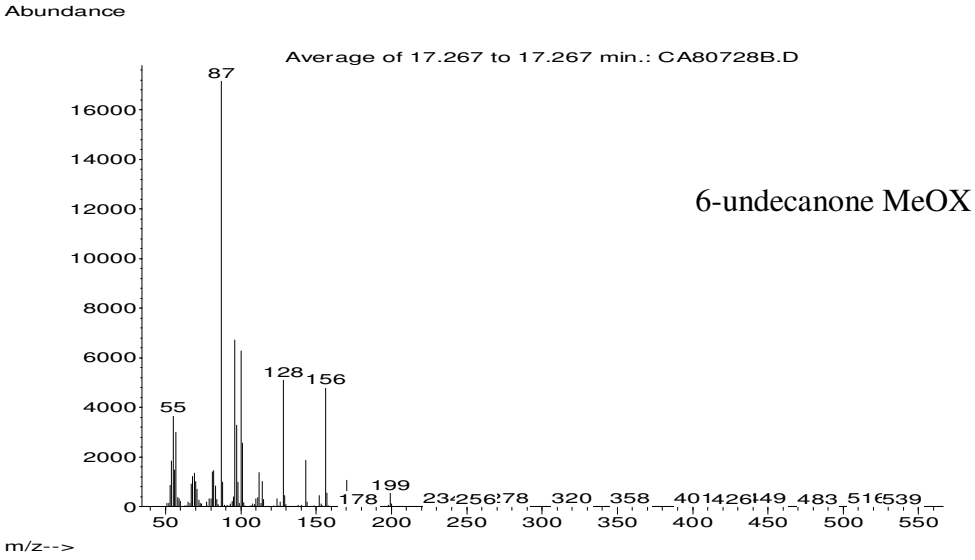
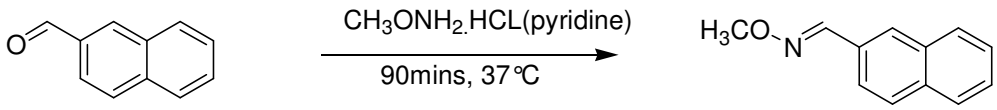


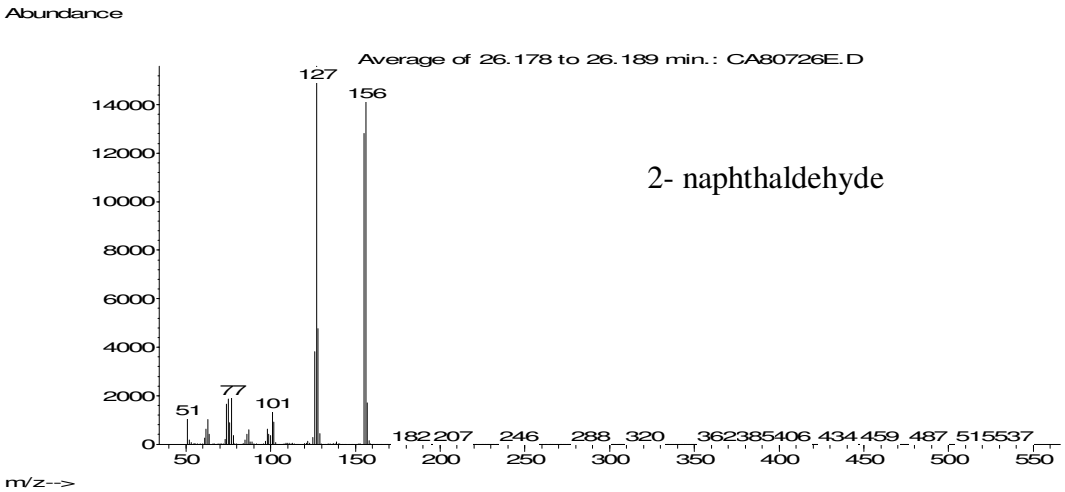
Figure 3-6: Mass spectra of (a) 6-undecanone and (b) 6-undecanone O-methyl oxime.



(a) (b)

Figure 3-7: 2-naphthaldehyde methoxymation reaction showing (a) 2-Naphthaldehyde and (b) derivatized 2-Naphthaldehyde

(a)



(b)

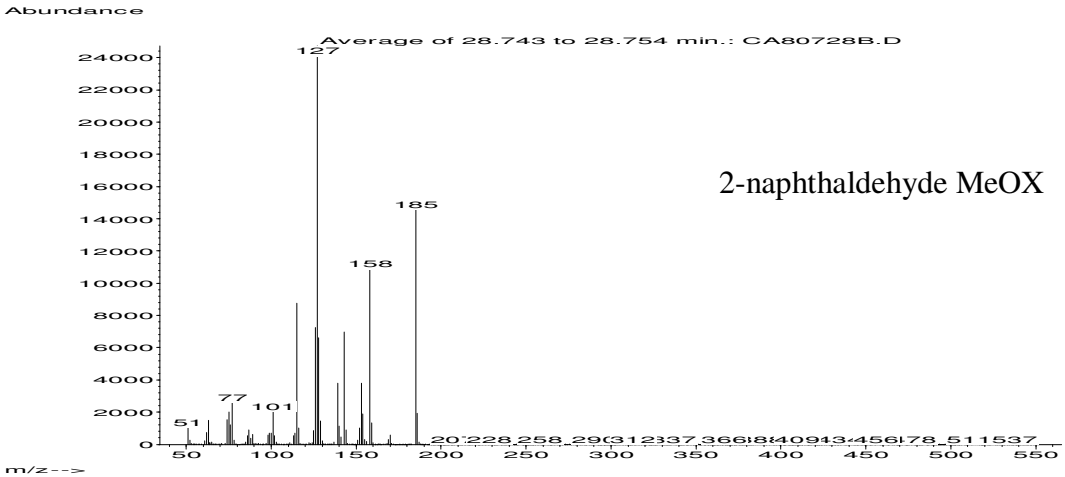


Figure 3-8: Mass spectra of (a) 2-naphthaldehyde and (b) 2-naphthaldehyde O-methyl oxime.

3.1.3.3 Number of Derivatives and Extent of Derivatization of Standards

All six standard compounds were derivatized and injected on the GC (see protocol in section 2.9.1). The derivatization duration was varied in reference to the two protocols in Table 3-1. For the Fiehn et al. 2005 and 2009 protocol derivatization duration started and varied from 0.0 min., 0.5 min., 2.5 min., 5.0 min., 10.0 min., 20 min., 45 min., to 90 min at 37°C. For the Jiye A et al 2005 protocol, derivatization duration started and varied from 60 min., 180 min., 360 min., 720 min., to 960 min at room temperature. From the TICs of the derivatized standards, the number of adducts formation of all the standards was obtained. The mass chromatograms of three selected standards at three selected derivatization duration are shown in Figures 3-9, 3-10 and 3-11.

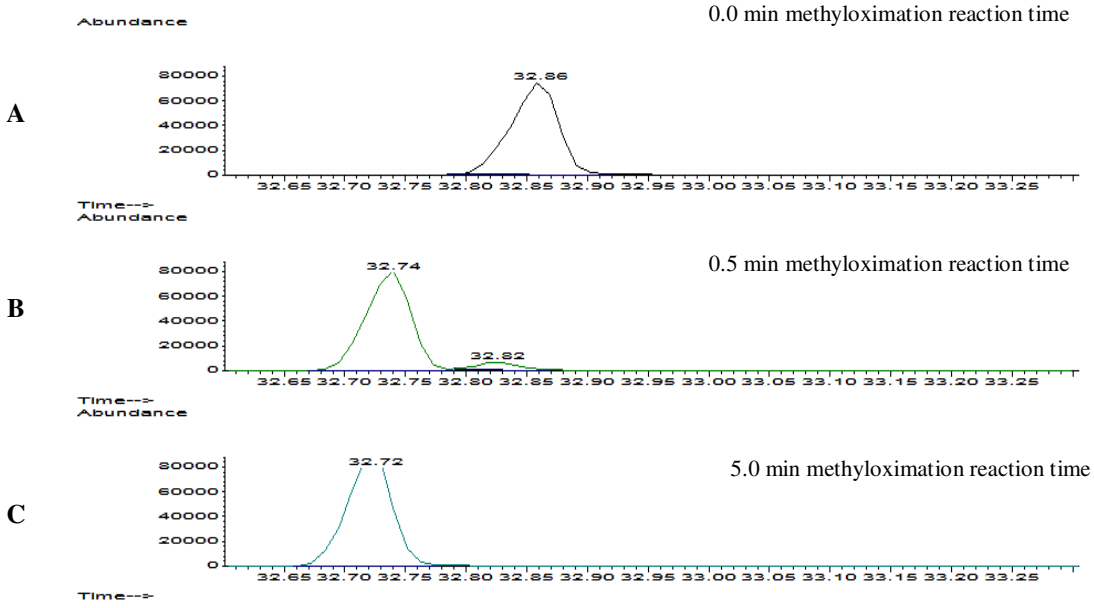


Figure 3-9: Comparing ion chromatograms of O-methyl oxime derivatives of 1,3-diphenylacetone at reaction times of (A) 0.0 min; (B) 0.5 min and (C) 5.0 min at 37°C.

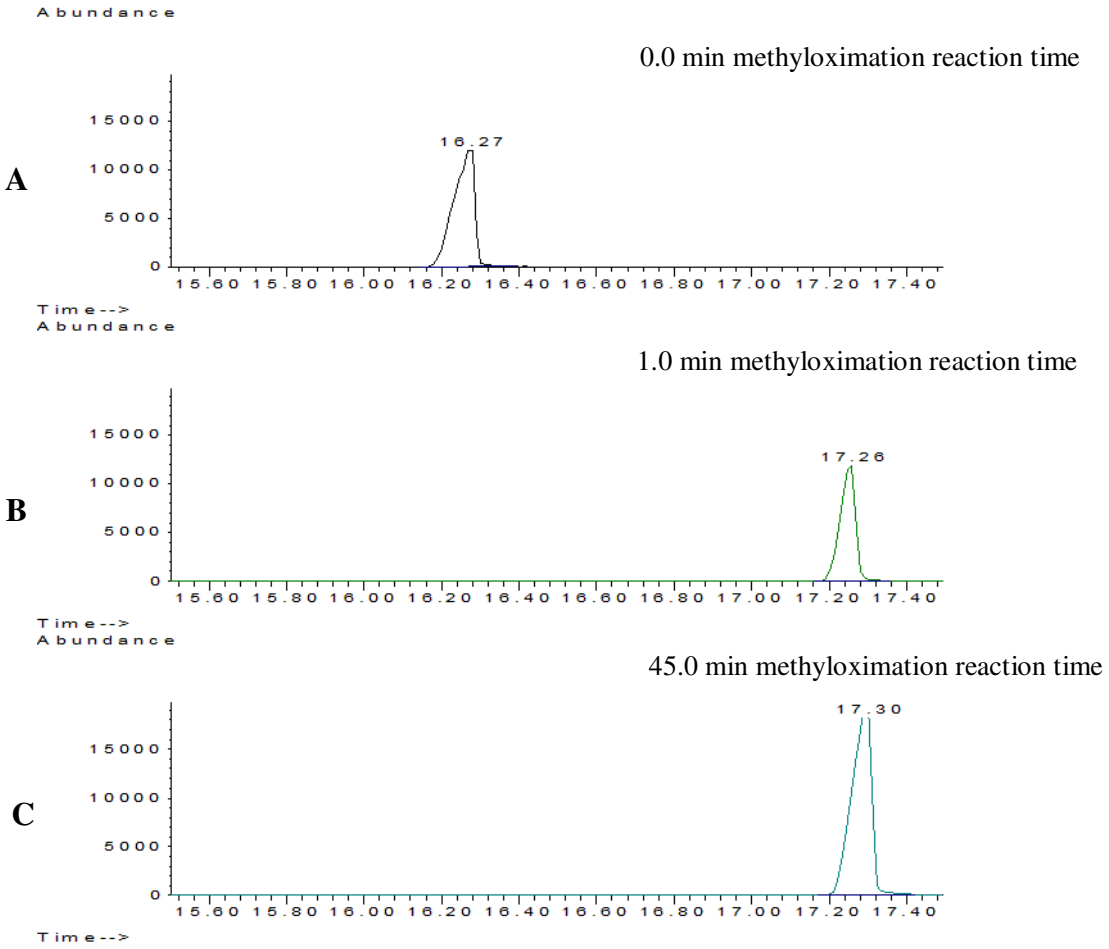


Figure 3-10: Comparing ion chromatograms of O-methyl oxime derivatives of 6-undecanone at reaction times of (A) 0.0 min; (B) 1.0 min and (C) 45.0 min at 37°C.

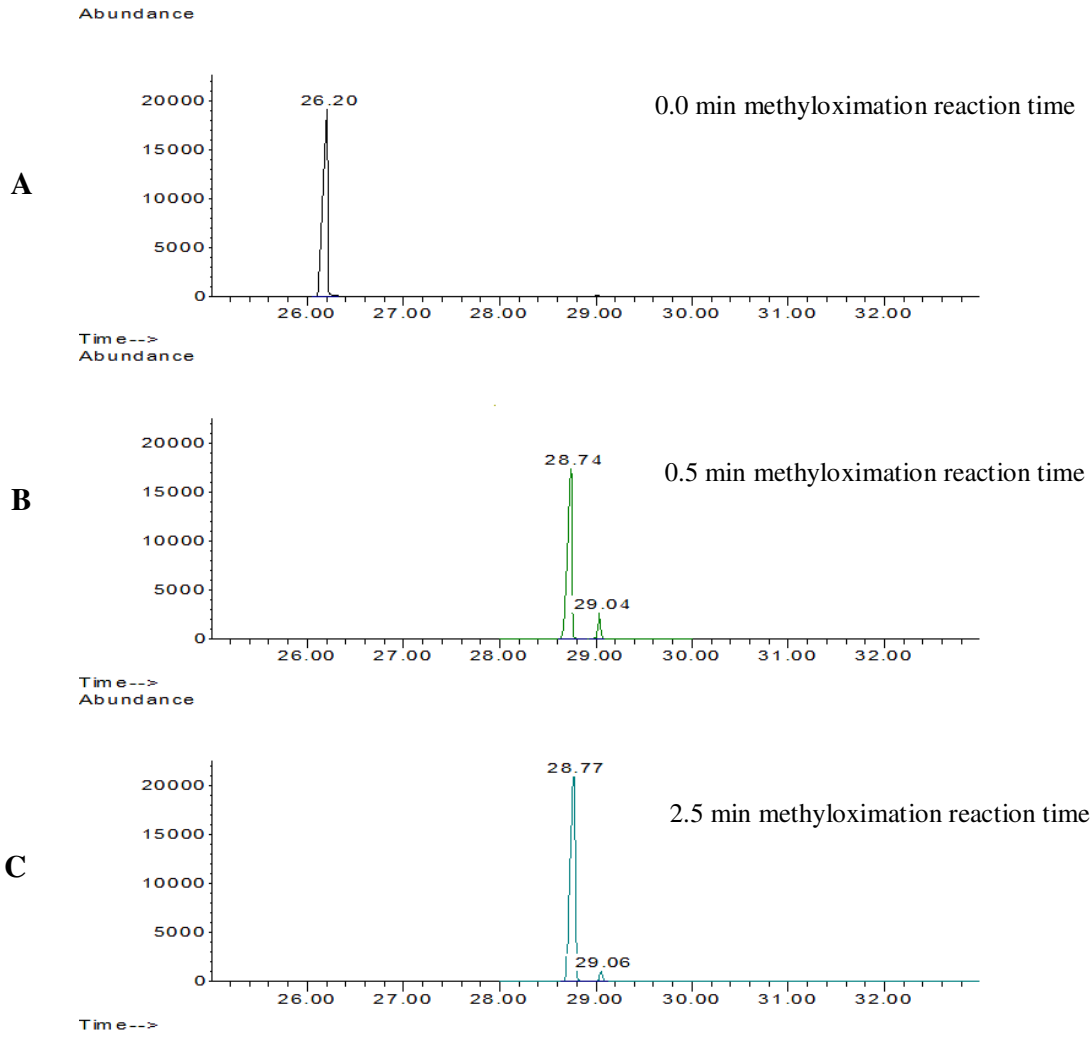


Figure 3-11: Comparing ion chromatograms of O-methyl oxime derivatives of 2-naphthaldehyde at reaction times of (A) 0.0 min; (B) 0.5 min and (C) 2.5 min. at 37°C.

A time profile analyses was done to ascertain the time it took for the complete conversion of all standards to their adduct forms in both protocols (Table 3-4 and Table 3-5).

Table 3-5: Time profile percentage conversion of aldehyde or ketones to their O-methyl oxime derivatives in pyridine solution at 37°C

Compound Name	Model Ion (m/z)	0.0 (min.)	0.5 (min.)	1.0 (min.)	2.5 (min.)	5.0 (min.)	10.0 (min.)	20.0 (min.)	45.0 (min.)	90.0 (min.)
1,3-diphenylacetone	91	100	8	5	1	0	0	0	0	0
1,3-diphenylacetone MEOX	91	0	92	95	99	100	100	100	100	100
4,4'-dichlorobenzophenone	139	100	100	100	100	99	99	99	95	92
4,4'-dichlorobenzophenone MeOX	248	0	0	0	0	1	1	1	5	8
Benzophenone	182	100	100	100	99	99	97	96	89	79
Benzophenone MeOX	211	0	0	0	1	1	3	4	11	21
2-Naphthaldehyde	127	100	12	1	0	0	0	0	0	0
2-Naphthaldehyde MeOX1	127	0	80	92	93	94	94	95	97	97
2-Naphthaldehyde MeOX2	127	0	7	7	7	6	6	5	3	3
Isobutyrophenone	105	100	100	99	99	97	95	90	74	58
Isobutyrophenone MeOX1	104	0	0	0.5	0.9	2	3	6	16	26
Isobutyrophenone MeOX2	104	0	0	0.3	0.6	1	2	4	10	16
6-Undecanone	71	100	77	41	7	5	5	4	0	0
6-Undecanone MeOX	87	0	23	59	93	95	95	96	100	100

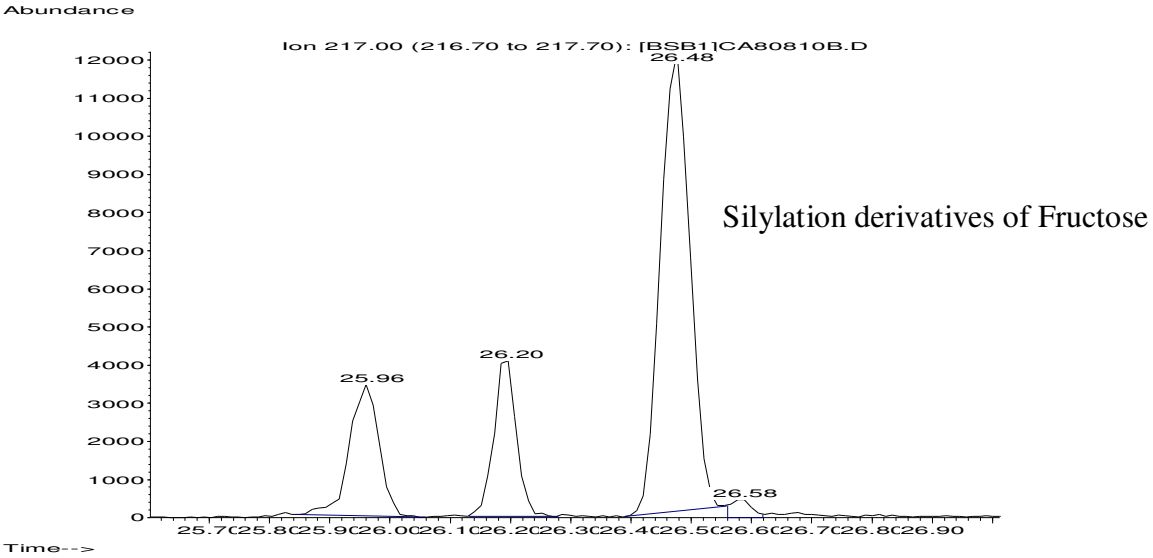
Table 3-6: Time profile percentage conversion of aldehyde or ketones to their O-methyl oxime derivatives in pyridine solution at room temperature.

Compound Name	Model Ion (m/z)	60.0 (min.)	180.0 (min.)	360.0 (min.)	720.0 (min.)	960.0 (min.)
1,3-diphenylacetone	91	0	0	0	0	0
1,3-diphenylacetone MEOX	91	100	100	100	100	100
4,4'-dichlorobenzophenone	139	100	82	41	23	15
4,4'-dichlorobenzophenone MeOX	248	0	18	59	77	85
Benzophenone	182	21	63	29	13	8
Benzophenone MeOX	211	79	37	71	87	92
2-Naphthaldehyde	127	100	0	0	0	0
2-Naphthaldehyde MeOX1	127	0	97	98	98	98
2-Naphthaldehyde MeOX2	127	0	3	2	2	2
Isobutyrophenone	105	69	23	12	6	3
Isobutyrophenone MeOX1	104	13	48	51	56	59
Isobutyrophenone MeOX2	104	18	29	33	37	38
6-Undecanone	71	82	0	0	0	0
6-Undecanone MeOX	87	18	100	100	100	100

Using the criteria outlined in section 3.1.3.1 to select a monitoring compound for the oximation reaction step it was observed that the underivatized and derivatized forms of the ketones and aldehyde under study (Table 3-3) were all detected in the method. While 2-naphthaldehyde and isobutyrophenone derivatized into two derivatives, 1,3-diphemylacetone, 4,4'-dichlorobenzophenone, benzophenone and 6-undecanone gave only one derivative each (Table 3-4).

The ketones and aldehyde under study were taken through eleven (11) methoxymation reaction times and the effect of varying methoxymation reaction times recorded in Table 3-4. For the only aldehyde in the mix, although none of the aldehyde, 2-naphthaldehyde was detected at 2.5 min, the methoximation reaction resulted in two (2) derivatives (Figure 3-13) which would add to the complexity of the chromatogram and as a result, 2- naphthaldehyde would not be a perfect standard for monitoring the methoxymation reaction step. With the ketones, 4,4'-dichlorobenzophenone, benzophenone and isobutyrophenone would not be selected because by the 90 min, 92% of 4,4'-dichlorobenzophenone, 79% of benzophenone and 58% of isobutyrophenone still remains. 1,3-diphenylacetone was 100% derivatized by the 5th minute, gave only one derivative, elutes in the middle of the chromatogram and in a relatively non-congested region of the chromatogram. 6-undecanone is completely derivatized by the 45th minute. 1,3-diphenylacetone was selected and added as a standard to monitor the methoxymation reaction step in the method.

(A)



(B)

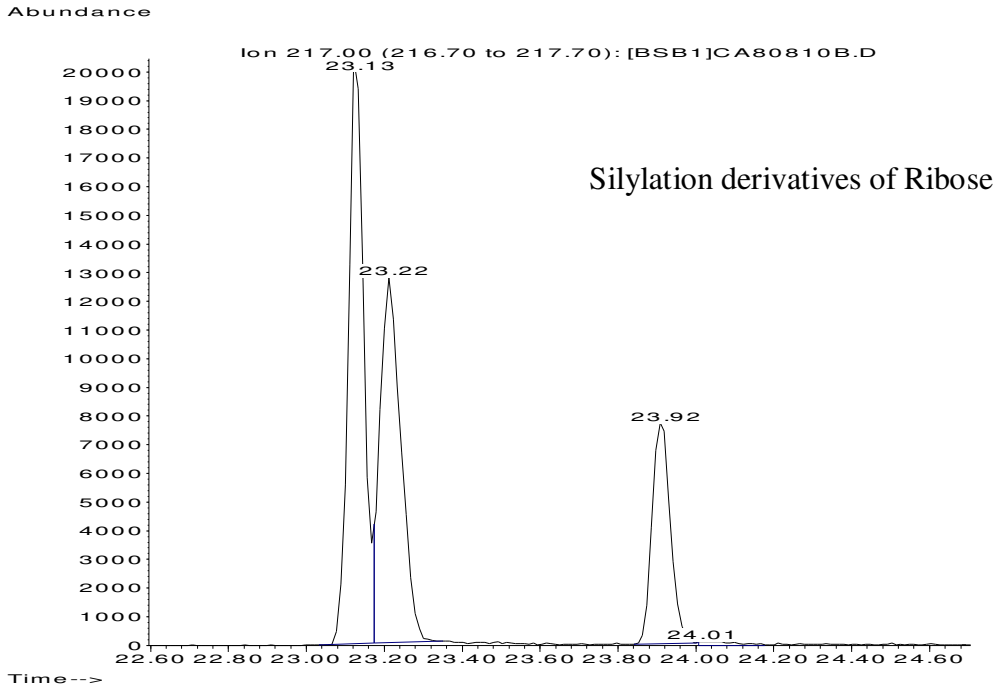


Figure 3-12: Mass chromatograms of: (A) fructose $(TMS)_5$ and (B) ribose $(TMS)_4$

Table 3-7: Time profile percentage conversion of Ribose $(TMS)_4$ and Fructose $(TMS)_5$ to their O-methyl oxime derivatives in pyridine solution at 37°C.

Compound Name	Model Ion (m/z)	0.0 (min.)	0.5 (min.)	1.0 (min.)	2.5 (min.)	5.0 (min.)	10.0 (min.)	20.0 (min.)	45.0 (min.)	90.0 (min.)
Ribose $(TMS)_4$	217	100	100	30	0	0	0	0	0	0
Ribose MEOX1 $(TMS)_4$	217	0	0	65	87	86	86	85	81	81
Ribose MEOX2 $(TMS)_4$	217	0	0	5	13	14	14	15	19	19
Fructose $(TMS)_5$	217	100	100	100	20	0	0	0	0	0
Fructose MeOX1 $(TMS)_5$	217	0	0	0	50	59	59	59	59	58
Fructose MeOX2 $(TMS)_5$	217	0	0	0	30	41	41	41	41	42

Table 3-8: Time profile percentage conversion of Ribose $(TMS)_4$ and Fructose $(TMS)_5$ to their O-methyl oxime derivatives in pyridine solution at room temperature.

Compound Name	Model Ion (m/z)	60.0 (min.)	180.0 (min.)	360.0 (min.)	720.0 (min.)	960.0 (min.)
Ribose $(TMS)_4$	217	0	0	0	0	0
Ribose MEOX1 $(TMS)_4$	217	98	96	93	92	90
Ribose MEOX2 $(TMS)_4$	217	2	4	7	8	10
Fructose $(TMS)_5$	217	0	0	0	0	0
Fructose MeOX1 $(TMS)_5$	217	59	58	57	53	51
Fructose MeOX2 $(TMS)_5$	217	41	42	43	47	49

Figure 3-12 shows the formation of four anomeric derivatives of fructose and ribose in a silylation reaction. A simple time and temperature experiment of fructose and ribose using the two protocols (Table 3-1) show the various degrees of completion of the oxime reactions (Tables 3-7 and 3-8). By the 2.5 minute and 37°C, there is conversion of all silylated ribose to oxime silylated ribose derivative and at 5.0 minute there is the complete conversion of silylated fructose to the oxime silylated fructose (Table 3-7).

From Table 3-8, silylated ribose and fructose are converted to their oxime silyl derivative at 60 minutes at room temperature.

3.2 Inclusion of Mass Spectra and Retention Indices of New Standards in our in-house Mass Spectral and Retention Index Library

The mass spectra of 39 derivatives of 38 chemical standards were added to our in-house mass spectra and retention index library of 113 derivatives of 83 standards. All standards (new and old) were derivatized and analyzed on a DB-17ht column using nine odd fatty acid standards as retention index markers (C_7 – C_{31} with assigned retention index values from 700 to 3100). Using the TIC of each standard, background mass spectra were subtracted and mass spectra of S/N ratios of above 100 to 1 were added to the in-house mass spectra and retention index library. The lists of standards are listed in Table 3-8.

3.3 Developing a Searchable Metabolite Database

An in-house searchable metabolite database has been developed for 160 metabolites. It has built-in tools for querying, extracting, viewing mass spectrum of metabolites and adding new metabolites (Figure 3-6). The entire database includes text, sequence, structure and mass chromatogram data. Data fields include compound name, chemical formula, retention index, retention index tolerance, peak list and mass spectra plots. Search parameters are independent of each other. Retention index parameter has an associated tolerance that allows a range of the retention indices to be searched.

Table 3-9: List of derivatized authentic standards that were added to in-house mass spectral database and retention index library.

Standard Compound Before Derivation	Compound Class	Solvent	Standard Compound After Derivation	Molecular [a] Formula of Deriv.	Molecular [b] Mass of Deriv.	Major Fragment Ions and their Relative Intensity
Methyl Nonanoate	Fatty acid methyl Ester	THF	Methyl Nonanoate	$C_{10}H_{20}O_2$	172	74(100), 87(80), 172(30)
Methyl Undecanoate	Fatty acid methyl Ester	THF	Methyl Undecanoate	$C_{12}H_{24}O_2$	200	74(100), 87(80), 200(30)
Urea	Amine	MeOH	Urea(TMS)$_2$	$C_7H_{20}N_2OSi_2$	204	147(100), 189(65), 204(2)
Methyl Laurate	Fatty acid methyl Ester	THF	Methyl Laurate	$C_{13}H_{26}O_2$	214	74(100), 87(80), 214(30)
Methyl Tridecanoate	Fatty acid methyl Ester	THF	Methyl Tridecanoate	$C_{14}H_{28}O_2$	228	74(100), 87(80), 228(30)
Methyl Myristate	Fatty acid methyl Ester	THF	Methyl Myristate	$C_{15}H_{30}O_2$	242	74(100), 87(80), 242(30)
Methyl Pentadecanoate	Fatty acid methyl Ester	THF	Methyl Pentadecanoate	$C_{16}H_{32}O_2$	256	74(100), 87(80), 256(30)
Succinic Acid D_4	Organic acid(deuterated)	H_2O	Succinic Acid D_4(TMS)$_2$	$C_{10}H_{22}D_4O_4Si_2$	266	147(100), 251(25), 266(1)
Methyl Palmitoleate	Fatty acid methyl Ester	THF	Methyl Palmitoleate	$C_{17}H_{34}O_2$	270	74(100), 87(80), 270(30)
Methyl Heptadecanoate	Fatty acid methyl Ester	THF	Methyl Heptadecanoate	$C_{18}H_{36}O_2$	284	74(100), 87(80), 284(30)
Methyl Stearate	Fatty acid methyl Ester	THF	Methyl Stearate	$C_{19}H_{38}O_2$	298	74(100), 87(85), 298(35)
Myristic Acid	Fatty acid	THF	Myristic Acid TMS	$C_{17}H_{36}O_2Si$	300	285(100), 300(10), 117(70), 132(30)
Melatonin	Amine	MeOH	MelatoninTMS	$C_{16}H_{24}N_2O_2Si$	304	232(100), 245(85), 289(0.5), 30420)
Methyl Nonadecanoate	Fatty acid methyl Ester	THF	Methyl Nonadecanoate	$C_{20}H_{40}O_2$	312	74(100), 87(80), 312(30)
Phenylalanine-D_8	Amino acid(deuterated)	H_2O	Phenylalanine-D_8 (TMS)$_2$	$C_{15}H_{27}D_8NO_2Si_2$	317	219(100), 200(80), 302(5)
Homovanillic Acid	Organic acid	H_2O	Homovanillic Acid (TMS)$_2$	$C_{15}H_{26}O4Si_2$	326	209(100), 311(70), 326(85), 117(5), 132(4)
Methyl Arachidate	Fatty acid methyl Ester	THF	Methyl Arachidate	$C_{21}H_{42}O_2$	326	74(100), 87(80), 326(30)
Myristic Acid D_{27}	Fatty acid(deuterated)	THF	Myristic Acid D_{27} TMS	$C_{17}H_{36}D_{27}O_2Si$	327	312(100), 327(10), 132(35)
Methyl Heneicosanoate	Fatty acid methyl Ester	THF	Methyl Heneicosanoate	$C_{22}H_{44}O_2$	340	74(100), 87(80), 340(30)
Syringe Acid	Organic acid	MeOH	Syringe acid (TMS)$_2$	$C_{15}H_{26}O_5Si_2$	342	327(100), 312(80), 342(85)
Aspartic Acid D_3	Amino acid(deuterated)	H_2O	Aspartic Acid D_3 (TMS)$_3$	$C_{13}H_{31}D_3NO_4Si_3$	352	235(100), 337(2), 352(0.8)
Linoleic Acid	Fatty acid	THF	Linoleic Acid TMS	$C_{21}H_{40}O_2Si$	352	74(100), 87(80), 352(30)
Meta-Tyramine	Amine	MeOH	m-Tyramina (TMS)$_3$	$C_{17}H_{35}NOSi_3$	353	174(100), 338(10), 86(20)
Oleic Acid	Fatty acid	THF	Oleic Acid TMS	$C_{21}H_{42}O_2Si$	354	74(100), 87(80), 354(30)
Elaidic	Fatty acid	THF	Elaidic TMS	$C_{21}H_{42}O_2Si$	354	74(100), 87(80), 354(30)
Methyl Arachidonate	Fatty acid methyl Ester	THF	Methyl Arachidonate	$C_{23}H_{46}O_2$	354	74(100), 87(80), 354(30)
Stearic Acid	Fatty acid	THF	Stearic Acid TMS	$C_{21}H_{44}O_2Si$	356	341(100), 356(15), 117(60), 132(30)
Methyl Tricosanoate	Fatty acid methyl Ester	THF	Methyl Tricosanoate	$C_{24}H_{48}O_2$	368	74(100), 87(80), 368(30)
Methyl Lignocerate	Fatty acid methyl Ester	THF	Methyl Lignocerate	$C_{25}H_{50}O_2$	382	74(100), 87(80), 382(30)
Stearic Acid D_{35}	Fatty acid(deuterated)	THF	Stearic Acid D_{35} TMS	$C_{21}H_{44}D_{35}O_2Si$	391	376(100), 391(20), 132(50)
Serotonin	Amine	H_2O	Serotonin (TMS)$_3$	$C_{19}H_{36}N_2OSi_3$	392	174(100), 377(15), 218(8)
Methyl Pentacosanoate	Fatty acid methyl Ester	THF	Methyl Pentacosanoate	$C_{26}H_{52}O_2$	396	74(100), 87(80), 396(30)
4-4Hydroxy-3- methoxy Mandelic	Organic acid	THF	Mandelic Acid (TMS)$_3$	$C_{18}H_{34}O_5Si_3$	414	297(100), 399(15), 117(1), 132(1)
Dopamine	Amine	H_2O	Dopamine (TMS)$_4$	$C_{20}H_{43}NO_2Si_4$	441	174(100), 426(25)
Cholesterol	Sterol	THF	Cholesterol TMS	$C_{30}H_{54}OSi$	458	129(100), 329(85), 353(45), 368(50), 458(25)
Serotonin	Amine	H_2O	Serotonin (TMS)$_4$	$C_{22}H_{44}N_2OSi_4$	464	174(100), 449(18), 290(20), 200(10), 218(10)
Stigmasterol	Sterol	THF	Stigmasterol TMS	$C_{32}H_{56}OSi$	484	129(95), 394(30), 484(35), 469(2)
Norepinephrine	Amine	H_2O	Norepinephrine (TMS)$_5$	$C_{23}H_{51}NO_3Si_5$	529	174(100), 355(25), 514(2)
$^{13}C_6$ Glucose	Sugar	H_2O	$^{13}C_6$ Glucose MeOX (TMS)$_5$	$C_{22}H_{53}D_6NO_2Si_5$	575	323(100), 207(90), 220(45)

a: Molecular formula of Deriv.: Molecular formular derivative of the chemical standard
b: Molecular mass Deriv: Molecular mass of derivatised standard compound

Figure 3-13: A screenshot image of our in-house metabolite database showing a search and data display tools describing methyl stearate.

Compound classes of metabolites included in the database are, amino acids, lipids (sterols, terpenoids, fatty acids and fatty acid methyl esters), carboxylic acids, carbohydrates and catecholamines. New metabolites can be added to the database application program without any risk to the security and integrity of the old data.

4. RESULTS AND DISCUSSION

The comprehensive metabolite profiles of human blood plasma generated by GC-MS and LC-MS can be quite complex, showing many overlapping peaks (71, 77, 82)-(83).

In this chapter, data discussion is laid out as follows:

a) Analysis of GC-MS data of derivatized human blood plasma extracts
    i. Derivatization of extracts
    ii. Deconvolution of co-eluting peaks
    iii. Criteria for the identification of metabolites
    iv. Identified compounds and compound classes
b) Analysis of LC-MS data of human blood plasma extracted including identified compounds and compound classes
c) Analysis of GC-MS data from derivatized LC fractions prepared from human blood plasma extract including identified compounds and compound classes
d) Comparison of GC-MS blood plasma data to LC-MS blood plasma data
e) Comparison of GC-MS 'whole' blood plasma data to GC-MS data from LC fractions.

4.1 GC-MS Data Analysis of Human Blood Plasma Samples

The comprehensive metabolite identification of metabolite in human blood plasma was performed using methods reported by Palazoglu et al. (83) and Jiye et al. (83, 84). In the Jiye et al. methods, MS files are exported into MATLAB software for data pre-treatment such as baseline correction and chromatogram alignment. Using ChromaTOF software, peaks are automatically detected and deconvoluted with peak widths set to 2 sec and signal-to-noise ratios of 5 and above. Using this approach Jiye et al. detected about 500 peaks and identified 80 peaks using an in-house library. In the Palazoglu et al. study agilent ChemStation software was used for data pre-treatment while a combination of AMDIS and the AgilentFiehn metabolite library was used to deconvolute and identify metabolites; a total of 102 metabolites were identified in human blood plasma samples.

From these two studies, at least two independent parameters, retention indices or locked retention times and mass spectral data were used in combination for compound identification. Following these metabolite identification approaches, a combination of the following was used in this work for finding and positively identifying metabolites in human blood plasma sample extracts:

- ChemStation and AMDIS mass spectral deconvolution software programs were used for data pre-treatment and mass spectral deconvolution of peaks.
- Retention index and mass spectral data were used for the identification of metabolites.

4.2 Monitoring Derivatization Reactions

Metabolites were extracted from deproteinated human blood plasma samples using MeOH and H_2O in a volume ratio of 1:1 and with the addition of stigmasterol as a recovery and internal standard. This method is a modification of the protein precipitation and metabolite extraction protocols of Fiehn *et al* (77) and Jiye *et al.* (71). The human blood plasma sample extracts were derivatized using O-methyl hydroxylamine hydrochloride in pyridine followed by silylation using MSTFA. Two chemical compounds for monitoring the derivatization reactions were added prior to derivatization. After derivatization nine odd numbered fatty acids retention index standards and two PAH internal standards were added before analyzing the samples by GC-MS in full scan mode on a DB-17ht column. A typical GC-MS total ion chromatogram of a derivatized sample of human blood plasma is shown in Figure 4-1.

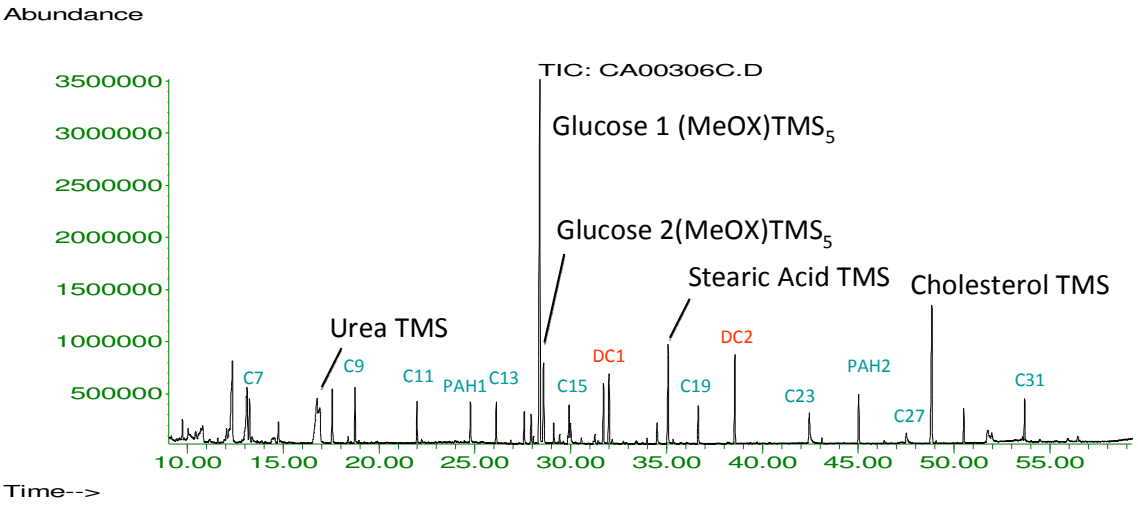


Figure 4-1: GC-MS chromatogram showing major peaks, derivatization controls (DC1 and DC2), internal standards (PAH 1 and PAH 2) and retention index standards (C_7 – C_{31}).

To ascertain whether there had been complete derivatization of all metabolites, the extent of derivatization of 1, 3-diphenylacetone and 9-anthracenemethanol was monitored in every sample.

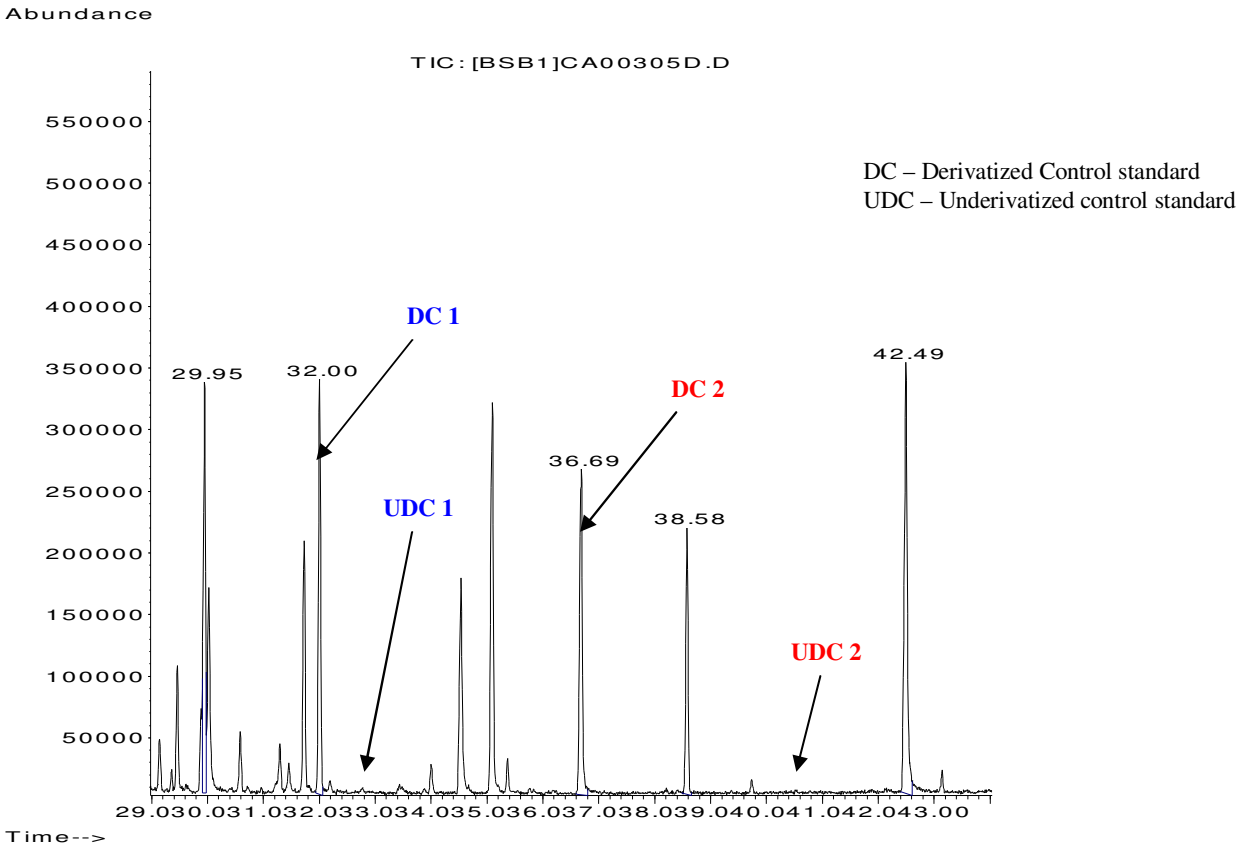


Figure 4-2: A fourteen-minute portion of the chromatogram in Figure 4-1 between 29 and 43 minutes.

If peaks corresponding to the underivatized and derivatized forms of these two compounds were detected, the analysis was considered void.

4.3 Deconvolution of Co-eluting Peaks

Using Agilent's ChemStation software, background subtraction and matching of the resulting mass spectra with spectra in our in-house retention index and mass spectral library allowed us to identify many compounds. Among the major peaks were glucose, cholesterol, urea and stearic acid (Figure 4-1). The mass spectral match factors for these abundant metabolites exceeded 90 using our in-house mass spectral library.

For the analysis of less abundant metabolites and co-eluting peaks, automatic mass spectral deconvolution (AMDIS) was used. AMDIS is an automated GC-MS identification program developed by NIST. AMDIS automatically finds peaks and deconvolutes the mass spectra of co-eluting compounds using model ion traces that best describe the presence of unique peaks. Using a combination of medium and high AMDIS settings for resolution, sensitivity and peak shape, together with two adjacent peak subtractions, the mass spectra, retention times, retention indices, ion traces and signal to noise ratios of all peaks, discrete and co-eluting, in a given GC-MS data file were obtained. The AMDIS report includes information on retention time, peak width, signal-to-noise ratio of model ion traces, mass-to-charge ratio of model ion traces and peak intensities. Peaks with signal-to-noise ratios below 5 and peak widths less than 2 scans were rejected from the data set. Using this protocol a total number of 100 peaks were identified in the full scan GC-MS data file of the derivatized human blood plasma extract.

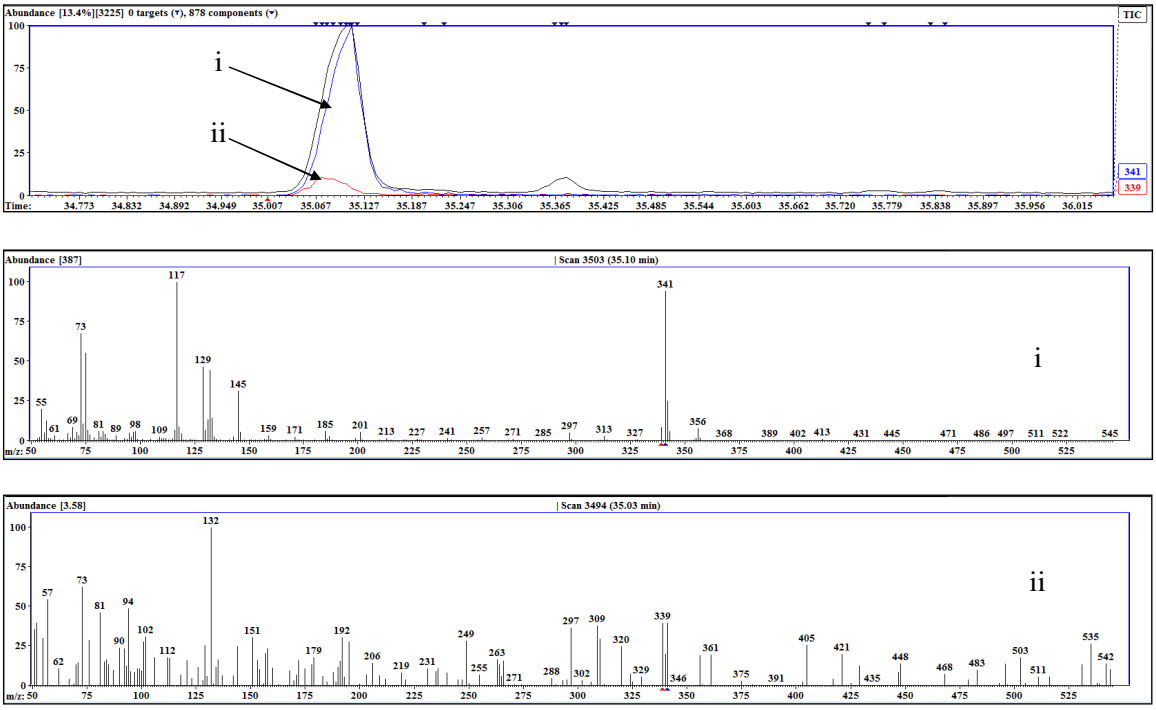


Figure 4-3: Deconvolution of two co-eluting peaks in AMDIS showing unique mass spectral components with ion traces m/z 339 in (i) and 341 in (ii)

Of the 100 deconvoluted peaks, 44 peaks were identified positively by matching their retention indices to retention indices and mass spectra of standards in our in-house library; 8 peaks were tentatively identified by matching to retention indices and mass spectra in the Human Metabolome Database library (1). There were another 48 peaks which remain unidentified. An example of an identification using AMDIS is shown in Figure 4.4. A listing of 100 peaks identified is provided in Table 4-1.

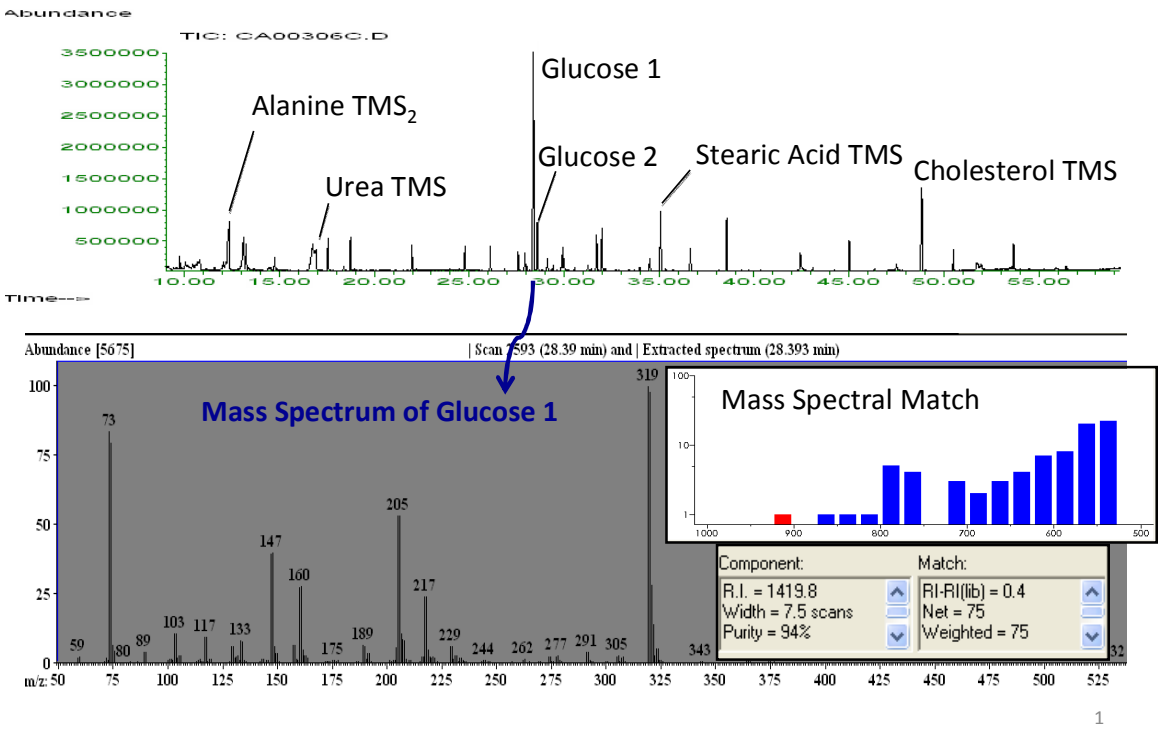


Figure 4-4: Identification of glucose 1 using AMDIS and our in-house MS library and retention index database.

Table 4-1: Retention indices and identifications of 100 deconvoluted peaks from GC-MS data analysis of human blood plasma extract.

No.	R.I.(DB-17ht, Fatty Acids)[a]	R.I.(DB-17ht, n-Alkanes)[b]	Identified derivatized compound[c]	m/z[d]	S/N(m/z)[e]
CA00305D-P1-U1	551.8		unknown	130	32
CA00305D-P2-T1	578.0		*lactic acid (TMS)* $_2$	129	8
CA00305D-P3-T2	612.5		*hydroxylamine TMS*	152	9
CA00305D-P4-U2	620.2		unknown	274	7
CA00305D-P5-U3	631.7		unknown	119	31
CA00305D-P6-N1	640.3		**alanine (TMS)$_2$**	116	12
CA00305D-P7-N2	645.6		**glycine (TMS)$_2$**	204	5
CA00305D-P8-N3	646.7		**pyruvic MEOX TMS**	174	15
CA00305D-P9-U4	676.8		unknown	132	7
CA00305D-P10-N4	691.2		**norvaline TMS**	129	7
CA00305D-P11-U5	704.4		unknown	196	7
CA00305D-P12-N5	720.6		**valine (TMS)$_2$**	144	6
CA00305D-P13-N6	734.1		**ethanolamine (TMS)$_3$**	174	10
CA00305D-P14-N7	766.3		**glyerol (TMS)$_3$**	205	20
CA00305D-P15-U6	773.5		unknown	204	58
CA00305D-P16-N8	776.4		**norleucine TMS**	158	6
CA00305D-P17-N9	778.5		**leucine (TMS)$_2$**	158	13
CA00305D-P18-T3	802.0	1217.3	*itaconic acid TMS*	215	10
CA00305D-P19-T4	802.1	1217.3	*citraconic acid TMS*	215	9
CA00305D-P20-N10	802.9	1218.3	**isoleucine (TMS)$_2$**	158	9
CA00305D-P21-N11	824.6	1243.5	**serine (TMS)$_2$**	132	5
CA00305D-P22-N12	838.7	1259.7	**proline (TMS)$_2$**	142	5
CA00305D-P23-T5	850.1	1272.9	*threonine (TMS)* $_2$	219	10
CA00305D-P24-N13	859.9	1284.3	**urea TMS**	189	23
CA00305D-P25-N14	861.9	1286.9	**benzoic acid TMS**	179	14
CA00305D-P26-N15	887.7	1316.8	**serine (TMS)$_3$**	218	7
CA00305D-P27-U7	887.8	1316.7	unknown	73	10
CA00305D-P28-N16	905.0	1336.6	**succinic acid (TMS)$_2$**	247	8
CA00305D-P29-N17	906.1	1337.7	**threonine (TMS)$_3$**	218	15
CA00305D-P30-U8	907.0	1338.8	unknown	117	5
CA00305D-P31-U9	942.9	1380.4	unknown	327	9
CA00305D-P32-U10	954.4	1393.7	unknown	267	20
CA00305D-P33-U11	962.8	1403.4	unknown	133	14
CA00305D-P34-U12	973.6	1415.9	unknown	172	5
CA00305D-P35-N18	1001.0	1447.6	**capric acid TMS**	129	5
CA00305D-P36-U13	1006.7	1454.0	unknown	217	5
CA00305D-P37-U14	1056.0	1511.0	unknown	207	5
CA00305D-P38-U15	1102.8	1564.7	unknown	75	5
CA00305D-P39-U16	1117.9	1581.0	unknown	204	9
CA00305D-P40-U17	1118.9	1582.1	unknown	355	8
CA00305D-P41-T6	1120.4	1583.8	*adenine (TMS)* $_2$	263	5
CA00305D-P42-U18	1137.7	1602.5	unknown	117	5
CA00305D-P43-U19	1152.8	1618.8	unknown	103	6
CA00305D-P44-N19	1175.9	1644.0	**lyxose MEOX2-(TMS)$_4$**	233	6
CA00305D-P45-U20	1182.7	1651.4	unknown	147	7
CA00305D-P46-U21	1193.8	1663.4	unknown	191	5
CA00305D-P47-N20	1197.0	1666.9	**arabinose MEOX1 (TMS)$_4$**	103	9
CA00305D-P48-N21	1198.3	1668.2	**glutamine (TMS)$_4$**	156	20
CA00305D-P49-N22	1200.1	1671.2	**ribose MeOX2 (TMS)$_4$**	147	14
CA00305D-P50-N23	1202.3	1672.6	**lauric acid TMS**	117	7

No.	R.I.(DB-17ht, Fatty Acids)[a]	R.I.(DB-17ht, n-Alkanes)[b]	Identified derivatized compound[c]	m/z[d]	S/N(m/z)[e]
CA00305D-P51-N24	1205.4	1676.4	**xylitol (TMS)$_5$**	103	9
CA00305D-P52-N25	1206.8	1677.4	**lyxose MeOX1-(TMS)$_4$**	103	6
CA00305D-P53-N26	1208.3	1681.7	**ribose MeOX1-(TMS)$_4$**	103	8
CA00305D-P54-N27	1212.1	1683.2	**ribose MEOX2-(TMS)$_4$**	217	6
CA00305D-P55-U22	1215.6	1687.0	unknown	75	5
CA00305D-P56-U23	1222.1	1694.1	unknown	102	9
CA00305D-P57-N28	1231.9	1704.7	**homocysteine (TMS)$_3$**	219	16
CA00305D-P58-U24	1232.2	1705.0	unknown	219	29
CA00305D-P59-N29	1237.2	1710.5	**phenylalanine (TMS)$_2$**	218	9
CA00305D-P60-U25	1238.3	1711.6	unknown	147	7
CA00305D-P61-U26	1245.1	1719.0	unknown	117	5
CA00305D-P62-N30	1245.4	1719.3	**fucose MEOX1-(TMS)$_4$**	117	5
CA00305D-P63-U27	1250.2	1724.6	unknown	221	6
CA00305D-P64-U28	1299.3	1777.7	unknown	117	22
CA00305D-P65-U29	1338.7	1817.1	unknown	74	5
CA00305D-P66-U30	1375.4	1853.7	unknown	312	15
CA00305D-P67-N31	1395.6	1873.9	**fructose MEOX1-(TMS)$_5$**	217	16
CA00305D-P68-U31	1396.0	1874.3	unknown	103	52
CA00305D-P69-N32	1401.1	1879.4	**myristic acid TMS**	117	14
CA00305D-P70-N33	1418.0	1896.2	**glucose MEOX1-(TMS)$_5$**	205	81
CA00305D-P71-N34	1419.8	1898.0	**galactose MEOX2-(TMS)$_5$**	319	36
CA00305D-P72-N35	1421.0	1899.2	**glucose MEOX2-(TMS)$_5$**	205	172
CA00305D-P73-U32	1431.7	1909.8	unknown	205	24
CA00305D-P74-U33	1441.5	1919.6	unknown	217	6
CA00305D-P75-N36	1457.5	1935.6	**tyramine (TMS)$_3$**	338	8
CA00305D-P76-U34	1474.2	1952.3	unknown	73	31
CA00305D-P77-U35	1484.0	1962.0	unknown	204	6
CA00305D-P78-U36	1496.4	1974.5	unknown	147	20
CA00305D-P79-U37	1504.6	1981.9	unknown	205	44
CA00305D-P80-N37	1536.1	2009.7	**sedoheptulose MeOX1 (TMS)$_6$**	217	5
CA00305D-P81-U38	1537.4	2010.8	unknown	103	15
CA00305D-P82-U39	1538.1	2011.5	unknown	204	17
CA00305D-P83-U40	1539.0	2012.2	unknown	273	19
CA00305D-P84-U41	1576.3	2045.1	unknown	217	8
CA00305D-P85-U42	1589.9	2057.2	unknown	205	11
CA00305D-P86-N38	1605.4	2070.9	**palmitic acid TMS**	117	54
CA00305D-P87-N39	1611.2	2076.0	**palmitelaidic acid TMS**	129	5
CA00305D-P88-U43	1666.0	2124.3	unknown	327	5
CA00305D-P89-N40	1706.9	2160.5	**heptadecanoic acid TMS**	117	7
CA00305D-P90-U44	1740.4	2190.1	unknown	102	17
CA00305D-P91-U45	1771.9	2223.3	unknown	376	51
CA00305D-P92-N41	1801.4	2224.3	**oleic acid TMS**	132	7
CA00305D-P93-N42	1805.5	2261.8	**stearic acid TMS**	341	69
CA00305D-P94-N43	1821.2	2279.8	**linoleic acid TMS**	132	12
CA00305D-P95-U46	2110.2	2576.9	unknown	244	11
CA00305D-P96-U47	2702.6	3110.6	unknown	117	7
CA00305D-P97-N44	2820.4	3199.9	**Cholesterol TMS**	129	110
CA00305D-P98-T7	2835.0	3213.6	*alpha-tocopherol TMS*	502	50
CA00305D-P99-T8	3080.4	3397.0	*octacosanoic acid TMS*	74	37
CA00305D-P100-U48	3237.1	3528.4	unknown	207	32

a: Retention index values using fatty acids on a DB-17ht column
b: Retention index values using n-alkanes on a DB-17ht column
c: Known compounds are in bold; tentative compounds are in italics and the rest are unknowns
d: Selected model ion that best describes peak
e: signal to noise ratio of selected fragment ion

4.4 Comparison of Retention Indices with those in our in-house Retention Index Library.

The retention indices of peaks determined in this work were compared to retention indices in our in-house library. The differences in retention indices of 35 positively identified compounds in the sample compared to retention indices in our in-house library ranged between -0.3 to +0.7 retention index units (Figure 4-5). These differences are very small; most amino acid showed the lowest differences while sugars and sugar alcohols showed greater differences (Table 4-2). These data confirm that our in-house database of retention index values is extremely useful in compound identification.

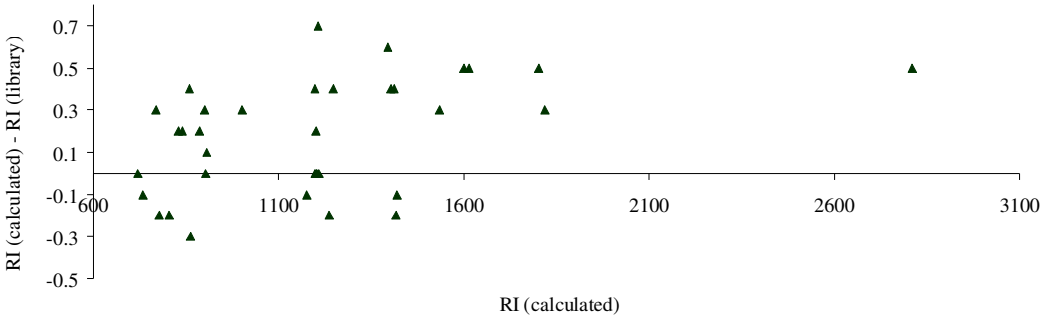


Figure 4-5: Deviation of retention indices of 35 identified compounds from retention index values in our in-house retention index library.

4.5 Reproducibility of Retention Index Values

The reproducibility of retention index values of compounds was determined using five separate aliquots of human blood plasma extracts which were derivatized and analyzed separately same day.

Table 4-2: Mean retention index values and percent relative standard deviations of 35 metabolite derivatives in human blood plasma samples.

Identified Compound	RI(Calculated)[a] (Mean)	S.D.[b] (n = 5)	%RSD[c]	RI(Library (Mean)	S.D.[d] (n = 7~12)	%RSD[e]
Valine $(TMS)_2$	719.5	0.40	0.06	719.7	0.8	0.11
Ethanolamine $(TMS)_3$	733.1	0.21	0.03	733.2	1.3	0.18
Glycerol $(TMS)_3$	768.5	0.92	0.12	768.6	1.6	0.21
Leucine $(TMS)_2$	778.2	0.21	0.03	782.4	0.7	0.08
Isoleucine $(TMS)_2$	804.8	0.40	0.05	805.9	0.5	0.06
Serine $(TMS)_2$	829.5	0.25	0.03	829.0	1.3	0.15
Proline $(TMS)_2$	841.7	0.78	0.09	841.0	0.5	0.06
Urea TMS	860.1	0.96	0.11	860.3	0.4	0.06
Benzoic Acid TMS	862.7	0.62	0.07	864.3	1.2	0.14
Serine $(TMS)_3$	887.5	0.12	0.01	887.1	1.2	0.14
Methyl Caprate	900.5	0.15	0.02	900.0	0.8	0.60
Succinic Acid $(TMS)_2$	904.5	0.64	0.07	904.1	0.4	0.05
Threonine $(TMS)_3$	905.9	0.67	0.07	906.9	0.8	0.09
Capric acid TMS	1000.6	0.31	0.03	1000.0	0.4	0.12
Lyxose MeOX2 - $(TMS)_4$	1176.1	0.21	0.02	1176.7	5.6	0.48
Arabinose MeOX1 - $(TMS)_4$	1197.1	0.44	0.04	1197.0	1.5	0.12
Glutamine $(TMS)_4$	1197.8	0.40	0.03	1198.5	1.4	0.12
Lauric acid TMS	1200.8	0.95	0.08	1200.0	0.5	0.30
Ribose MeOX2 - TMS	1201.1	0.95	0.08	1199.6	3.7	0.31
Xylitol $(TMS)_5$	1205.0	0.35	0.03	1204.5	4.7	0.39
Lyxose MeOX1 - $(TMS)_4$	1206.3	0.46	0.04	1205.5	5.6	0.47
Ribose MeOX1 - TMS	1208.2	0.32	0.03	1207.6	3.3	0.28
Phenylalanine $(TMS)_2$	1238.2	0.32	0.03	1238.1	1.1	0.09
Fucose MeOX1	1248.2	0.26	0.02	1243.7	3.3	0.27
Fructose MeOX1 - $(TMS)_5$	1395.3	0.70	0.05	1394.0	3.4	0.24
Myristic Acid TMS	1402.6	0.74	0.05	1403.0	1.0	0.09
Glucose MeOX1 - TMS5	1411.5	0.12	0.01	1409.7	2.4	0.17
Galactose MeOX2 - $(TMS)_5$	1415.9	0.36	0.03	1415.7	2.1	0.15
Glucose MeOX2 - $(TMS)_5$	1420.8	0.12	0.01	1415.7	2.1	0.15
Sedoheptulose MeOX	1533.7	0.32	0.02	1532.9	2.7	0.18
Palmitic acid TMS	1601.2	0.75	0.05	1600.0	2.0	0.12
Palmitelaidic Acid TMS	1612.2	0.74	0.05	1612.0	1.3	0.80
Stearic Acid TMS	1804.4	0.81	0.05	1803.0	0.8	0.50
Oleic Acid TMS	1819.9	0.51	0.03	1819.0	1.3	0.16
Cholesterol TMS	2811.2	0.61	0.02	2810.0	0.9	0.67

a: Mean retention index value calculated
b: Standard deviation of retention index value calculated
c: Percentage relative standard deviation of retention index value calculated
d: Standard deviation of retention index value in Library
e:Percentage relative stanadrad deviation of retention index value in Library

The experimental protocol follows the description in the flow diagram in Figure 2-4. The retention indices of 35 known compounds were selected for this retention index reproducibility study. The retention index values were determined using the retention times of model ions selected by the AMDIS program. Mean retention index values and percent relative standard deviation for the 35 peaks are listed in Table 4-2. Percent relative standard deviations in these retention index values ranged between 0.01 % and 0.12% with a mean value of 0.04%. This study demonstrated that the retention index values of the identified compounds in the sample were not statistically different from values in the retention index. Thus, matching retention indices to values in an in-house library is a very valuable tool in confirming peak identities between samples.

4.6 The Dependence of Mass Spectral Matches on Signal-to-Noise Ratios

The quality of mass spectral matches between experimental mass spectra and spectra in a database can be grouped into three categories: matches < 70 %, matches between 70% to 85% and matches that exceed 85%. An investigation of the data from 35 known compounds revealed that peaks with signal/noise ratios of less than 15 gave mass spectral matches that were below 70%. Peaks with signal/noise ratios of 15 and 25 gave mass spectral matches between 70% and 85% while peaks with signal-to-noise ratios greater than 25 gave mass spectral matches of 85% or greater (Figure 4-6). This demonstrates clearly that the signal-to-noise ratios of a peak are a determining factor in the quality of mass spectral matches. Glucose and cholesterol showed the highest signal-to-noise ratios and had mass spectral matches well above 90 %.

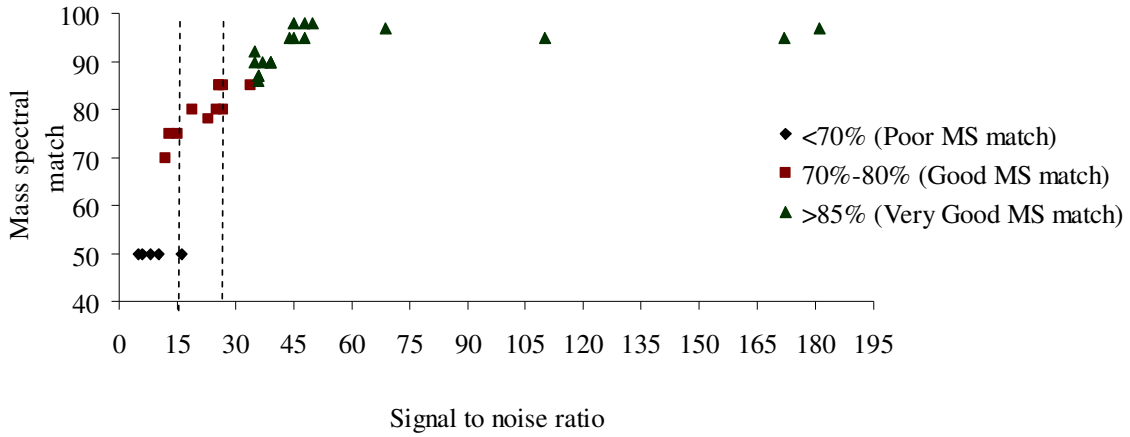


Figure 4-6: Comparing mass spectra quality to signal-to-noise ratios

4.7 Tentative Identification of Metabolites

For the identification of compounds not listed in our in-house library, mass spectra and retention index values were matched against the mass spectra and retention indices obtained from libraries accessible on the web. These libraries used include the Golm metabolite Database (49) and the Human Metabolite Database (1). These two public libraries were built using GC-MS and an n-alkane retention index data DB-5 (5% phenyl silicone) column. To be able to compare our data to data in these public libraries, we needed to convert our RI data into their RI data. Calculated retention indices of twenty selected metabolites identified in human blood plasma extracts (Table 4-3) using the list of n-alkanes analyzed on a DB-17ht and the retention indices obtained from the human metabolome database were compared (Figure 4-7).

A retention index and retention time graph of n-alkane series on DB-17ht and DB-5ms columns of the twenty selected compounds shows a close relationship of retention indices (Figure 4-11). The delta retention index is -87 to + 73. Using a tolerance of ± 100, eight compounds were tentatively identified and listed in Table 4-1 (85).

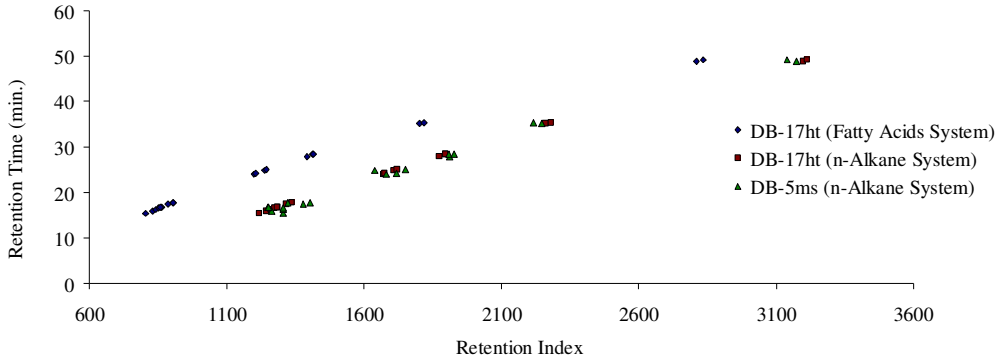


Figure 4-7: Comparing retention indices of two series of standards (fatty acids and n-alkane series) on two columns; DB-17ht and BD-5ms.

Table 4-3: Comparison of retention indices of 20 standards on DB-17ht and DB-5ms columns.

RT(min)	Name of Derivatized Compound	Fatty Acid system RI (DB-17)	n-Alkane System RI (DB-17)	n-Alkane System RI(HMDB: DB-5)	Δ RI (DB-17 and DB-5 n-Alkane)
15.5	Isoleucine TMS2	805.0	1218.3	1305.0	-86.7
15.9	Serine TMS2	828.8	1243.5	1264.0	-20.5
16.2	Proline TMS2	841.2	1259.7	1305.0	-45.3
16.5	Threonine TMS2	851.0	1274.3	1305.0	-30.7
16.8	Benzoic Acid TMS	862.8	1286.9	1250.0	36.9
16.7	Urea TMS	859.0	1284.3	1254.0	30.3
17.3	Serine TMS3	887.0	1316.8	1378.0	-61.2
17.7	Succinic Acid TMS2	904.0	1336.6	1323.0	13.6
17.7	Threonine TMS3	905.8	1337.7	1405.0	-67.3
24.1	Ribose MeOX2 - TMS	1200.0	1670.2	1681.0	-10.8
24.2	Xylitol TMS5	1205.2	1675.9	1717.0	-41.1
24.9	Phenylalanine TMS2	1237.6	1710.5	1639.0	71.5
25.0	Fucose MeOX1	1243.0	1719.4	1751.0	-31.6
28.0	Fructose MeOX1 - TMS5	1394.0	1873.8	1912.0	-38.2
28.4	Glucose MeOX1 - TMS5	1411.0	1896.3	1929.0	-32.7
28.4	Galactose MeOX2 - TMS5	1416.0	1897.9	1908.0	-10.1
35.1	Stearic Acid TMS	1803.0	2261.9	2248.0	13.9
35.4	Oleic Acid TMS	1818.0	2279.6	2217.0	62.6
48.9	Cholesterol TMS	2810.0	3200.0	3175.0	25.0
49.1	alpha-tocopherol TMS	2835.0	3213.6	3141.0	72.6

4.8 Identification of Unknown Metabolites

The retention indices and mass spectra of forty-seven peaks could not be matched to retention index values or mass spectra in our library or to either of the public libraries. These compounds remain as "unknowns". However, the mass spectra of some of these compounds showed significant similarities to spectra of sugars, sugar alcohols and fatty acid methyl esters (see appendix 1). Some of these unknowns were given tentative compound class identifications.

4.9 Elution Region of Compound Classes in a GC-MS Chromatogram

A typical GC-MS chromatogram of a derivatized human blood plasma extract sample is shown in Figure 4:8. Four compound classes are detected across the chromatogram. The elution of fatty acids as their trimethylsilyl derivatives and also as methyl esters (FAMEs) are seen from C_7 to C_{31} retention index standards. Amino acid

TMS derivatives eluted between C_7 and C_{31} retention index standards. The sugars and sugar alcohols TMS derivatives eluted between C_{12} and C_{25} retention index standards while sterols and terpenoids TMS derivatives elute between C_{25} and C_{31} retention index standards.

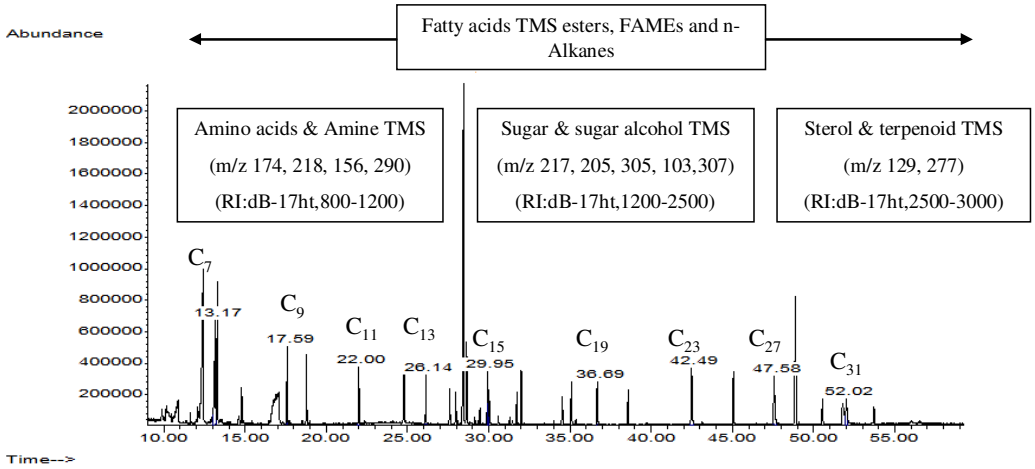


Figure 4:8 Total ion current chromatogram of derivatized human blood plasma extract showing the elution regions of some common compound classes.

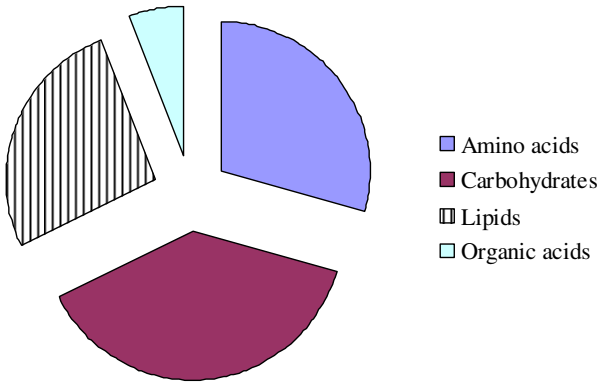


Figure 4-9: Contribution of compound classes in human blood plasma extract using GC-MS.

4.10 Peak Area Dynamic Range of Derivatives (GC-MS of "Whole" Sample)

The peak areas of 35 deconvoluted peaks in a typical GC-MS chromatogram (e.g. Figure 4-8) were normalised to acenaphthene d_{10}. The resulting data showed that these 35 peaks had peak areas which ranged over 2 orders of magnitude (Figure 4-10).

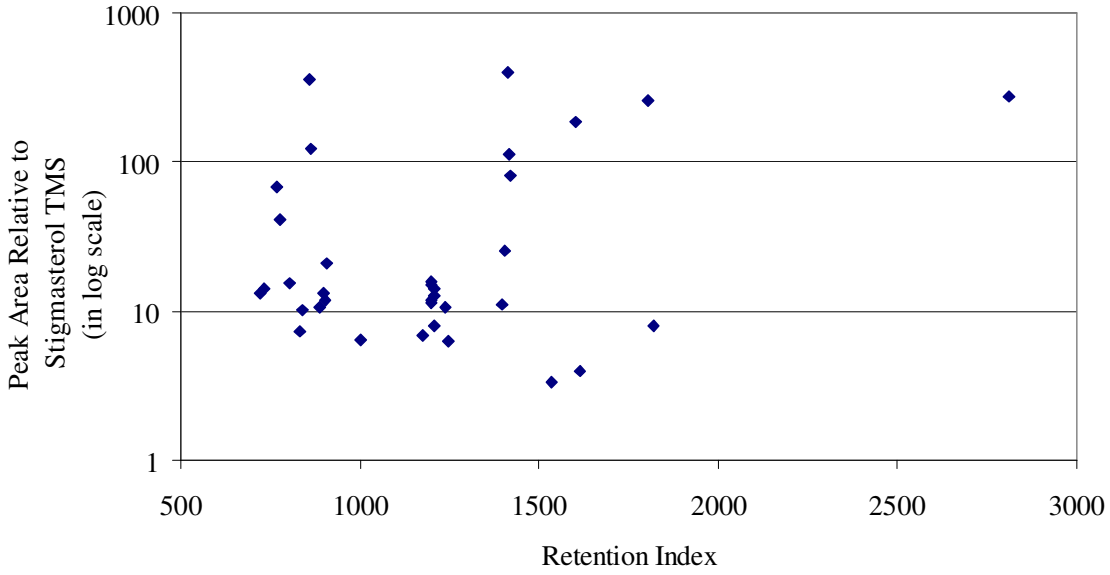


Figure 4-10: Range of peak areas of 35 peaks identified in human blood plasma extracts using GC-MS.

4.11 LC-MS Analysis of Human Blood Plasma Extracts

The human blood plasma extracted was then analyzed using ESI-LC-MS on a HILIC column in both positive ion and negative ion modes.

The total number of peaks observed in plasma (Figure 4-11) was 97 (Tables 4-4 and 4-5) with peak widths ranging between 0.2 min. and 0.8 min (see selected mass chromatogram in Figures 4-12 and 4-13). Forty-seven peaks were observed in the positive ion mode while 58 were observed in the negative ion mode; 8 peaks were observed in both positive and negative ion modes. The total number of peaks with confirmed identifications is 44. These identities were confirmed by accurate masses (within a tolerance range of 2 ppm), retention times and characteristic CID spectra of authentic standards.

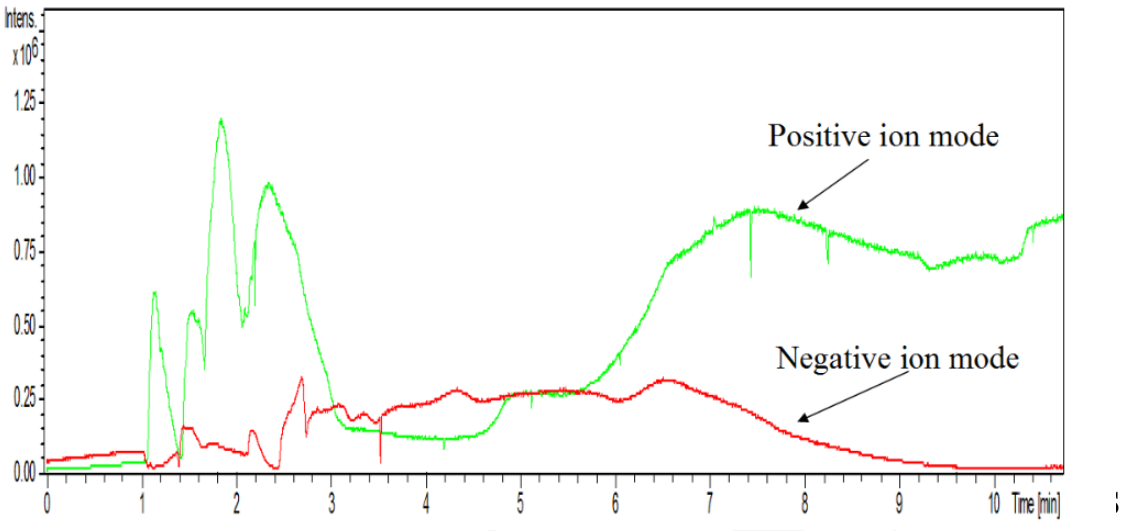


Figure 4-11: Total ion current chromatograms of human blood plasma extract obtained from gradient HILIC LC-MS analysis.

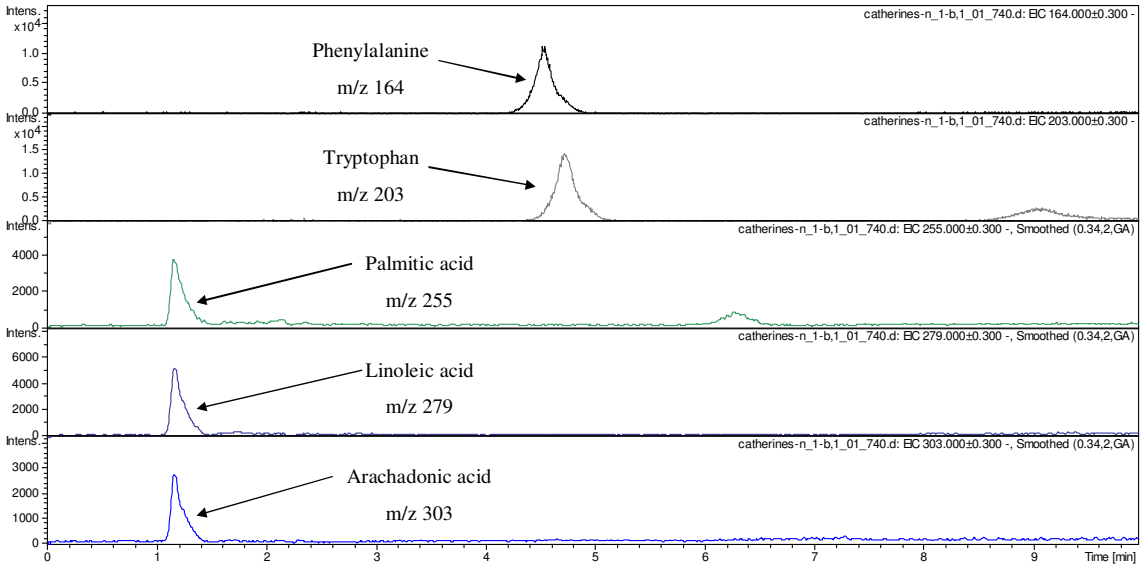


Figure 4-12: Selected extracted ion chromatograms in LC-MS negative ion mode

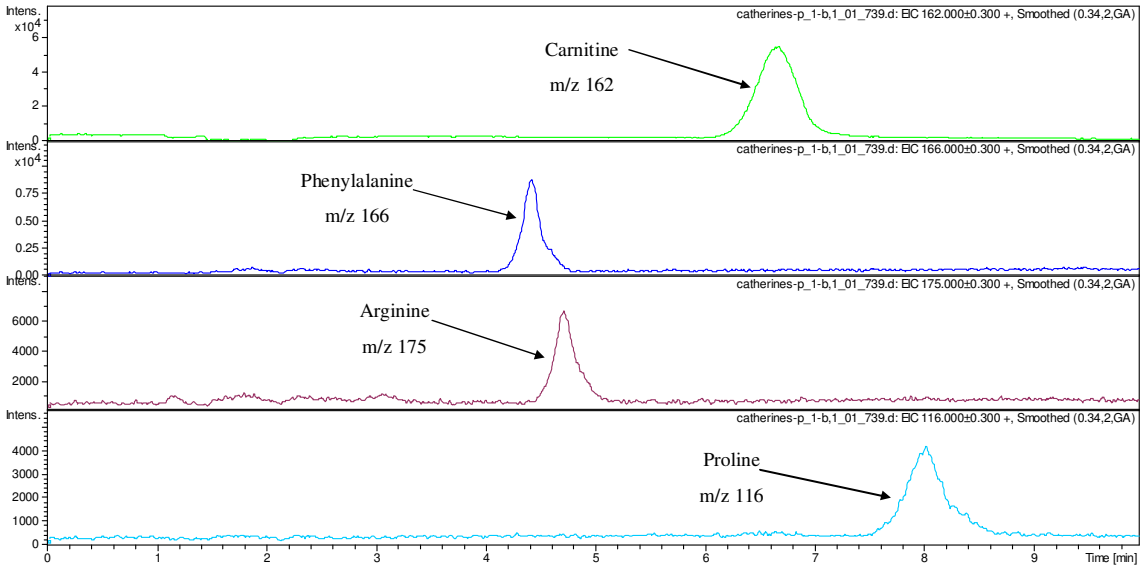


Figure 4-13: Selected extracted ion chromatograms in LC-MS positive ion mode.

Table 4-4: Retention times of peaks identified in LC-MS positive ion mode

Retention Time (min.)	Molecular mass of $(M+H)^+$	Chemical formula	Compound identification [a]
1.14	357.31	$C_{20}H_{42}N_2OP$	
1.14	391.29		
1.16	263.24	$C_{18}H_{31}O$	
1.16	312.29		
1.17	181.08	$C_7H_9N_4O_2$	**theophyline/theobromine**
1.39	188.07	$C_{11}H_{10}NO_2$	**indoleacrylic acid**
1.46	355.30	$C_{20}H_{39}N_2O_3$	
1.46	371.32	$C_{22}H_{43}O_4$	
1.46	374.33		
1.46	384.26	$C_{21}H_{38}NO_5$	*3-hydroxy-5,8-tetradecadiencarnitine*
1.46	394.23	$C_{10}H_{29}N_{13}O_2P$	
1.47	381.30		
1.55	258.11		
1.59	244.19	$C_{13}H_{26}NO_3$	
1.60	243.23	$C_{15}H_{31}O_2$	*pentadecanoic acid*
1.85	156.08	$C_6H_{10}N_3O_2$	**histidine**
2.08	230.04		
2.13	158.96	CN_2NaO_6	**sodium formate**
2.13	227.08		
2.14	211.10	$C_8H_{16}N_2NaO_3$	**glycyl-leucine**
2.17	256.27	$C_{16}H_{34}NO$	**palmitic amide**
2.67	136.05	$C_4H_7N_3NaO$	**creatinine (Na^+)**
2.79	226.18	$C_{13}H_{24}NO_2$	
2.93	185.13	$C_{11}H_{21}S$	
3.26	160.13	$C_8H_{18}NO_2$	**aminooctanoic acid**
4.40	166.09	$C_9H_{12}NO_2$	**phenylalanine**
4.63	205.10	$C_{11}H_{13}N_2O_2$	**tryptophan**
4.74	175.12	$C_6H_{15}N_4O_2$	**arginine**
4.80	132.08	$C_4H_{10}N_3O_2$	**leucine**
4.84	133.10	$C_6H_{14}NO_3$	
4.85	229.16	$C_{11}H_{21}N_2O_3$	**leucyl-proline**
5.27	132.10	$C_6H_{14}NO_2$	**isoleucine**
5.65	150.06	$C_5H_{12}NO_2S$	**methionine**
5.77	144.10	$C_7H_{14}NO_2$	**proline betaine**
5.89	118.09	$C_5H_{12}NO_2$	**valine**
6.66	162.11	$C_7H_{16}NO_3$	**carnitine**
7.21	182.08	$C_9H_{12}NO_3$	**tyrosine**
7.99	116.07	$C_5H_{10}NO_2$	**proline**
9.32	239.15	$C_{11}H_{24}N_2NaS$	
8.00	140.07	$C_5H_{11}NNaO_2$	**aminopentanoic acid**
8.13	199.17	$C_{12}H_{23}O_2$	*dodecenoic acid*
8.19	147.11	$C_6H_{15}N_2O_2$	**lysine**
8.42	280.09		
9.08	282.28		
9.17	165.05	$C_9H_9O_3$	
9.18	176.07	$C_5H_{10}N_3O_4$	**guanidinosuccinic acid**
9.19	233.23		

a: Known compounds are in bold; tentative compounds are in italics and the rest are unknowns

Table 4-5: Retention times of peaks identified in LC-MS negative ion mode

Retention Time (min.)	Molecular mass of $(M-H)^-$	Chemical formula	Compound identification [a]
1.08	313.08		
1.09	165.02		
1.11	239.02	$C_6H_{11}N_2O_4S_2$	*cystine*
1.14	255.23	$C_{16}H_{31}O_2$	**palmitic acid**
1.15	281.25	$C_{18}H_{33}O_2$	**oleic acid (or isomer)**
1.15	250.14	$C_9H_{20}N_3O_5$	
1.16	277.07		
1.16	279.23	$C_{18}H_{31}O_2$	**linoleic acid**
1.16	303.23	$C_{20}H_{31}O_2$	**arachadonic acid**
1.16	327.23	$C_{22}H_{31}O_2$	*docosahexaenoic acid*
1.42	265.15		
1.65	133.06	$C_4H_9N_2O_3$	*canaline*
1.82	157.12	$C_7H_{14}N_3O$	
2.24	239.09		
2.20	261.04	$C_6H_{14}O_9P$	**mannitol phosphate**
2.45	88.03	$C_3H_6NO_2$	**alanine/sarcosine**
2.46	121.03	$C_2H_5N_2O_4$	
2.52	112.05	$C_4H_6N_3O$	**creatinine**
2.71	116.07	$C_5H_{10}NO_2$	**valine**
4.28	209.07	$C_7H_{13}O_7$	**sedoheptulose**
4.53	164.07	$C_9H_{10}NO_2$	**phenylalanine**
4.53	165.07		
4.54	209.08	$C_{10}H_{14}N_2OP$	
4.57	173.10	$C_6H_{13}N_4O_2$	**arginine**
4.71	203.08	$C_{11}H_{11}N_2O_2$	**tryptophan**
5.13	130.09	$C_6H_{12}NO_2$	**leucine**
5.59	196.07	$C_6H_{15}NO_4P$	
5.58	130.05	$C_5H_8NO_3$	**isoleucine**
5.60	146.08	$C_6H_{12}NO_3$	
5.63	206.10	$C_8H_{16}NO_5$	
5.78	367.16	$C_{19}H_{27}O_5S$	*testosterone sulfate*
6.04	145.10	$C_6H_{13}N_2O_2$	**lysine**
6.57	161.03	$C_6H_9O_3S$	
6.57	179.06		
6.59	89.02	$C_3H_5O_3$	
6.59	119.03	$C_4H_7O_5$	
6.60	101.06	$C_5H_9O_2$	**valeric acid**
6.60	215.03	$C_6H_{12}ClO_6$	**hexose (Cl adduct)**
6.62	225.06	$C_7H_{13}O_8$	**hexose (formate)**
7.17	162.02	$C_5H_8NO_3S$	**acetylcystiene**
7.20	89.02	$C_3H_5O_3$	
7.20	119.03	$C_4H_7O_6$	
7.22	179.06	$C_6H_{11}O_6$	**hexose**
7.23	149.03	$C_2H_6N_4O_2P$	
7.39	180.07	$C_9H_{10}NO_3$	**tyrosine**
7.40	187.01	$C_7H_7O_6$	**2-methylaconitate**
7.40	153.02	$C_7H_5O_4$	**2-pyrocatechuic acid**
7.45	127.05	$C_5H_7N_2O_2$	**dihydrothymine**
7.58	118.05	$C_4H_8NO_3$	**threonine**
7.62	167.02	$C_5H_3N_4O_3$	**uric acid**
8.13	145.06	$C_5H_9N_2O_3$	**glutamine**
8.35	127.05	$C_5H_7N_2O_2$	**dihydrothymine**
8.54	103.04	$C_4H_7O_3$	**hydroxybutyric acid**
8.72	293.18		
9.07	302.10		
9.15	242.08	$C_9H_{12}N_3O_5$	**cytidine**
9.20	201.04	$C_3H_9N_2O_8$	
9.25	357.03		

a: Known compounds are in bold; tentative compounds are in italics and the rest are unknowns

The 44 identified metabolites were classified into four groups: amino acids and conjugates, organic acids, lipids and sugar derivatives. Each compound class was represented by 5 or more compounds. Amino acids and conjugates were highly represented while sugar derivatives (carbohydrate-bearing compounds) were poorly represented.

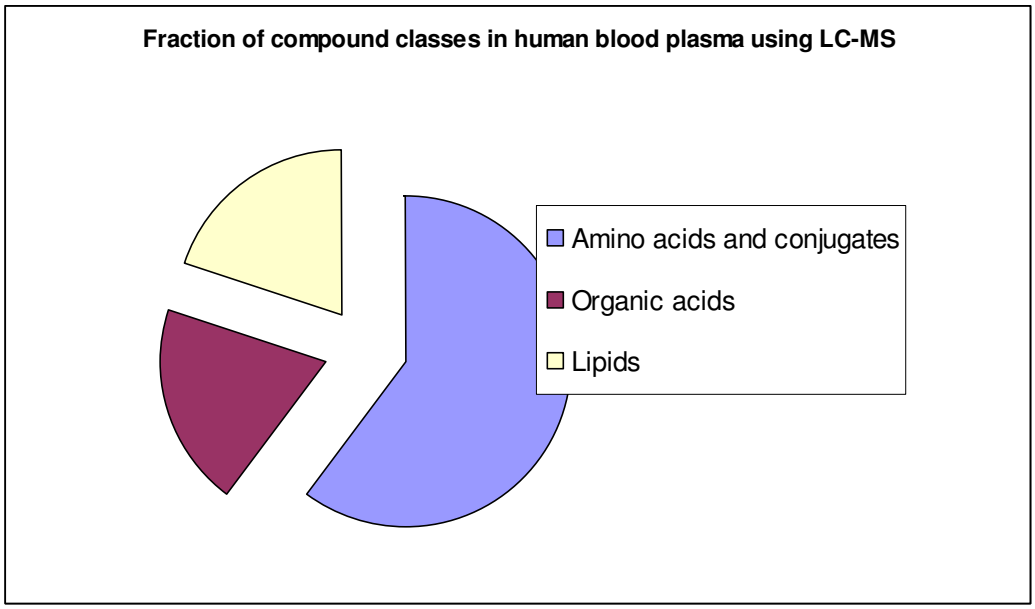


Figure 4-14: Contribution of compound classes in human blood plasma extracts using LC-MS

The range of peak areas in the LC-MS analyses was found to be about four orders of magnitude as shown in Figure 4-15.

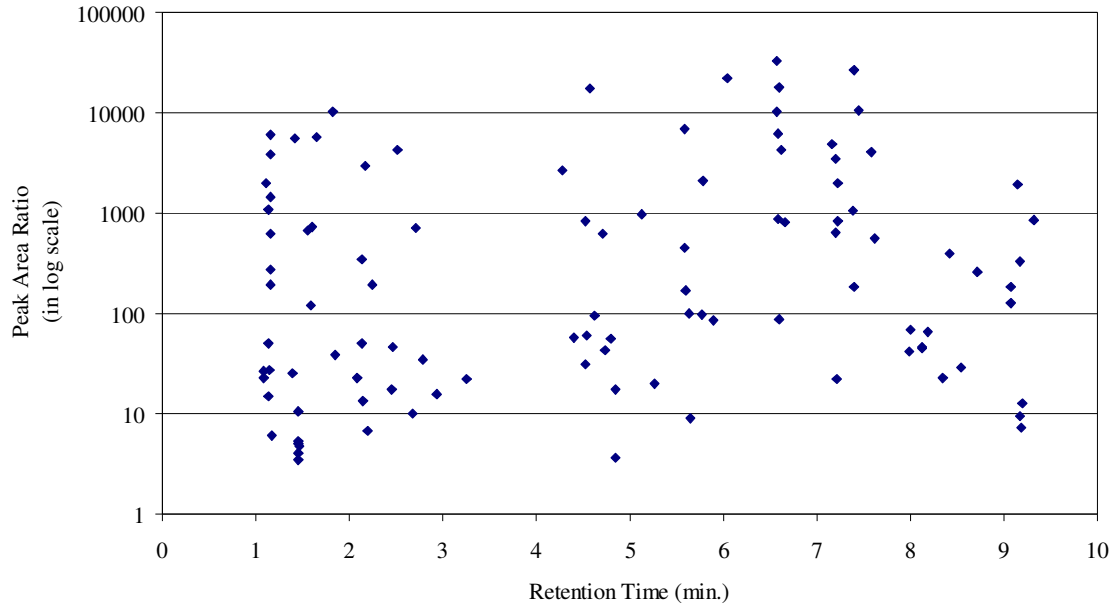


Figure 4-15: Graph showing the range of peak areas observed in both positive and negative ion in LC-MS analyses of human blood plasma extracts.

4.12 GC-MS Analysis of LC Fractions of Human Blood Plasma Extracts

In order to explore the plasma metabolome in greater detail it was decided to collect the HILIC LC eluent into fractions, concerntrate these fractions and analyse the fractions using derivatization GC-MS. It was hoped that the fractionation analysis protocol would afford a degree of separation not possible in either the LC or GC modes alone. Moreover, the quality of GC-MS mass spectra of weak and unknown components should be significantly better. The improved quality of these mass spectra would be useful in future work.

To this end the extract of human blood plasma was fractionated several times using the HILIC LC column according to Figure 4-16. Each of the 16 fractions was

derivatized and analyzed by GC-MS on a DB-17ht column. All 16 chromatograms were analyzed using AMDIS; particular case was used in identifying compounds which appeared in more than one fraction.

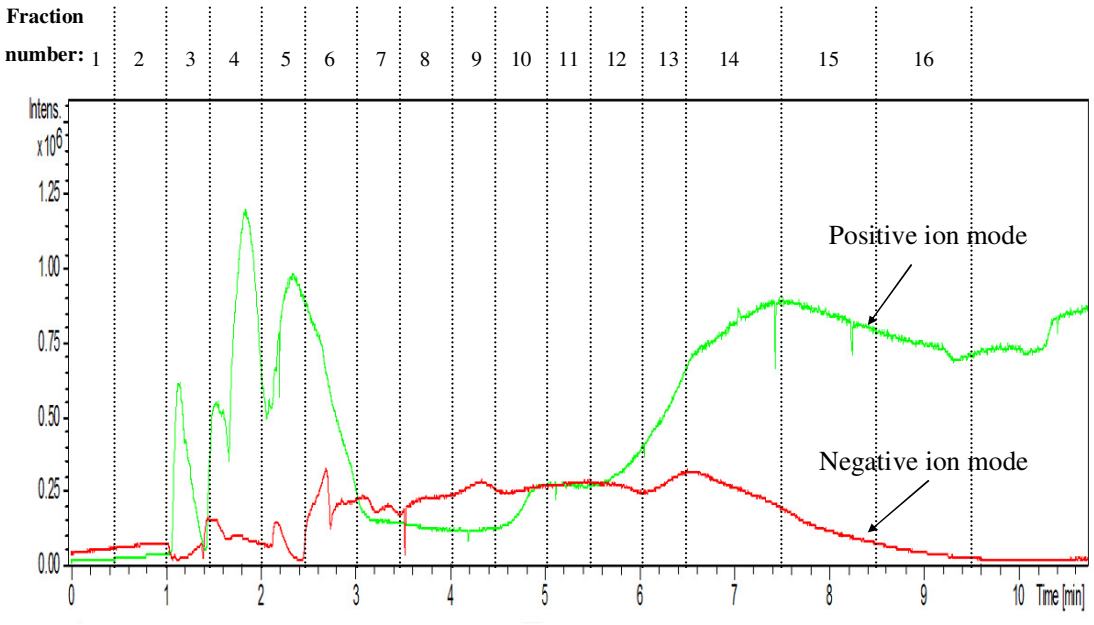


Figure 4-16: Positive and negative total ion current chromatograms showing fractionation regions in LC.

The AMDIS deconvolution of the 16 LC fractions resulted in the detection of a grand total of 393 peaks. Thirteen compounds were found in more than one LC fraction (Figure 4-6) and five compounds were smeared all through the 16 LC fractions analyzed. These are: benzoic acid, myristic acid, palmitic acid, stearic acid and an unknown. A total of 276 unique compounds were found after the elimination of repeating peaks. Of these unique peaks, about 50% is found concentrated in fractions 3 to 6. At least one peak from four compound classes – amino acids, lipids, carbohydrates and organic acid was found in

the unique LC fractions from 3 to 6. Deconvoluted peaks in each fraction has been re-constructed in excel graphs in groups of three and four fractions (Appendix 3).

Table 4-6: Table showing the number of unique GC-MS peaks per LC fraction and their compound class (es)

Total number of GC-MS peaks in LC fractions	Number of GC-MS peaks unique to a given fraction	Compound class(es) of peaks in fraction
13	8	lipids[a]
15	9	amino acids; lipids[a]
35	26	amine; lipids[a]
40	30	lipids[a]; carbohydrates[c]; sugar alcohol
56	44	amine; carbohydrates[c]; organic acids[b]; lipids[a]
40	30	amino acids; carbohydrates[c]; lipids[a]
23	16	lipids[a]; carbohydrates[c]
20	14	amino acids; lipids[a]
29	22	amino acids; lipids[a]
22	14	amino acids; lipids[a]
17	9	lipids[a]; amino acids
21	15	amino acids; carbohydrates[c]
23	17	amino acids; carbohydrates[c]; lipids[a]
12	6	amino acids; lipids[a]
14	8	amino acids; lipids[a]
13	8	lipids[a]

a: Compound class includes fatty acids, fatty acid methyl esters, sterols and terpenoids
b: Compound class is mainly carboxylic acids
c: Compound class includes monosaccharides and disaccharides

Table 4-7: Compounds found in more than one LC fraction analyzed by GC-MS.

RI(DB-17ht, fatty acid series)	Compound identity	Fraction Number	% of total Peak Area
561.6	Unknown	5	25
562.3	Unknown	6	24
561.4	Unknown	7	25
561.4	Unknown	8	26
768.6	Glycerol $(TMS)_3$	3	0.7
768.9	Glycerol $(TMS)_3$	4	90.8
768.5	Glycerol $(TMS)_3$	5	7.5
768.5	Glycerol $(TMS)_3$	6	1.0
781.5	Leucine $(TMS)_2$	9	14
781.9	Leucine $(TMS)_2$	10	59
781.7	Leucine $(TMS)_2$	11	27
858.7	Urea TMS	2	39
859.0	Urea TMS	3	58
859.0	Urea TMS	4	3
904.0	Succinic Acid $(TMS)_2$	5	42
904.0	Succinic Acid $(TMS)_2$	6	54
904.0	Succinic Acid $(TMS)_2$	7	4
1197.5	Glutamine $(TMS)_4$	9	3
1198.0	Glutamine $(TMS)_4$	10	19
1197.8	Glutamine $(TMS)_4$	11	61
1198.0	Glutamine $(TMS)_4$	12	17
1394.0	Fructose MeOX1 - $(TMS)_5$	4	97
1393.6	Fructose MeOX1 - $(TMS)_5$	5	3
1410.9	Glucose MeOX1 - $(TMS)_5$	5	98
1410.5	Glucose MeOX1 - $(TMS)_5$	6	2
1421.0	Glucose MeOX2 - $(TMS)_5$	5	99
1420.5	Glucose MeOX2 - $(TMS)_5$	6	1
2606.1	Unknown	10	70
2605.4	Unknown	11	30
2807.4	Cholesterol TMS	3	7
2810.0	Cholesterol TMS	4	91
2809.8	Cholesterol TMS	5	2
811.6	Glycine TMS3	13	7
810.3	Glycine TMS3	14	84
810.7	Glycine TMS3	15	9
1743.70	Methyl stearate	3	77
1743.90	Methyl stearate	4	23

4.13 Improvements in Mass Spectra Matches for Peak Identification.

The mass spectra matches of 35 selected identified peaks are grouped into three: < 70 %, 70% - 80% and > 85%. An investigation of the data revealed that two compounds had MS matches below 70%, 10 compounds had MS matches between 70% to 80% and 23 compounds had MS matches greater than 85%. Glucose and cholesterol recorded the highest signal/noise ratios (280 and 127 respectively) with mass spectra matches close to 100%.

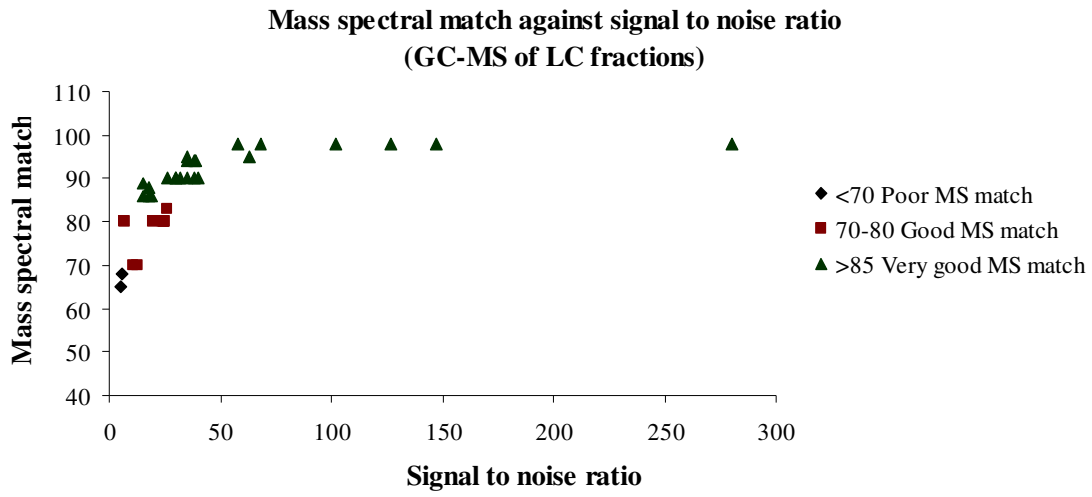


Figure 4-17: Mass spectra matches and signal-to-noise ratios of 35 selected derivatized metabolites in human blood plasma sample.

4.14 Peak Area Dynamic Range of Derivatives (GC-MS of LC Fractions)

Normalizing all identified derivatized peak areas of the 16 LC fractions to the peak area of added internal standard, acenapthene-d_{10}, the general pattern of major, minor and trace peaks was repeated. Grouping the LC fractions into groups of three to four, peak area dynamic range of compounds were found to be between five to six orders of magnitude. Compounds in fractions 4 to 6 recorded the highest peak area dynamic range (six orders of magnitude). This is the group that includes glucose, cholesterol and glycerol derivatives belonging to the compound class of the most abundant compounds in human blood plasma.

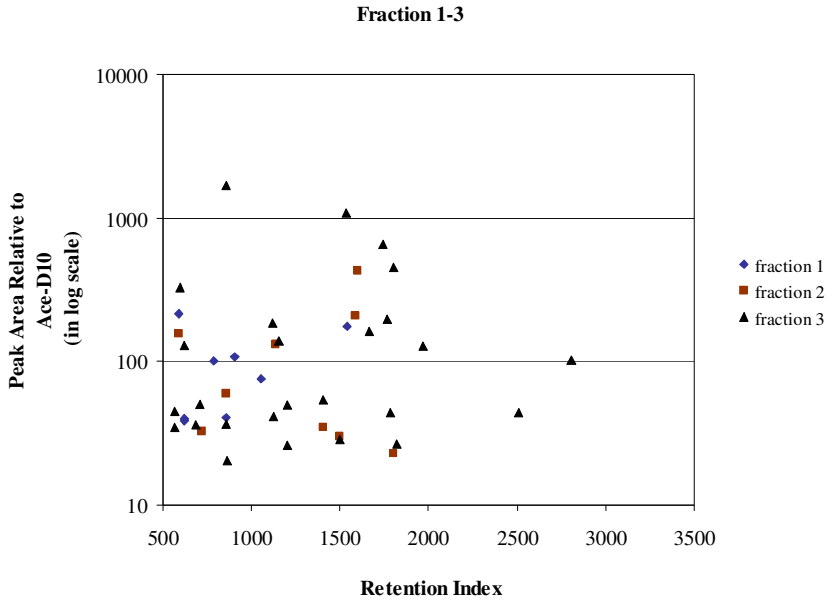


Figure 4-18: Peak area dynamic range of compounds of LC fractions 1 to 3.

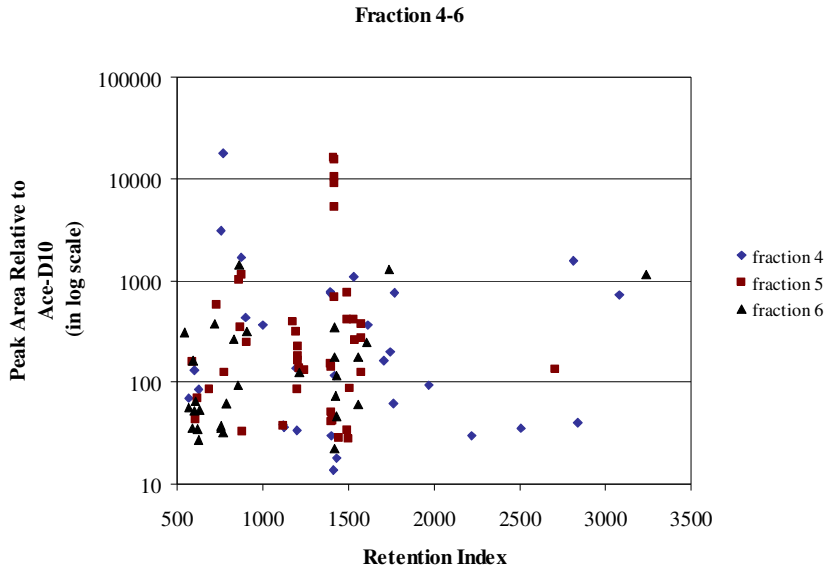


Figure 4-19: Peak area dynamic range of compounds of LC fractions 4 to 6.

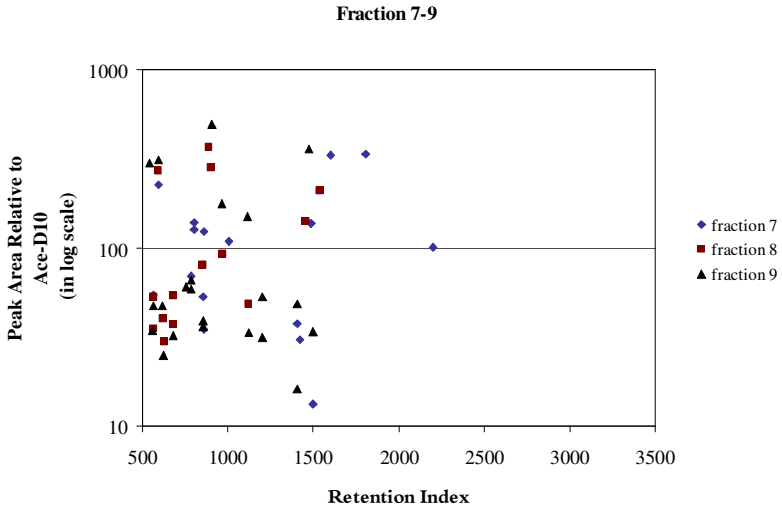


Figure 4-20: Peak area dynamic range of compounds of LC fractions 7 to 9.

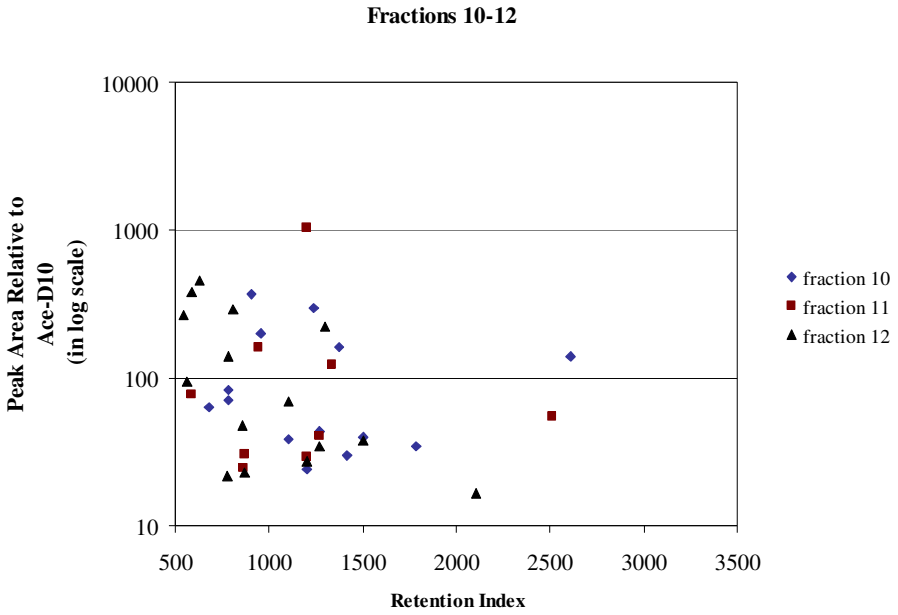


Figure 4-21: Peak area dynamic range of compounds of LC fractions 10 to 12.

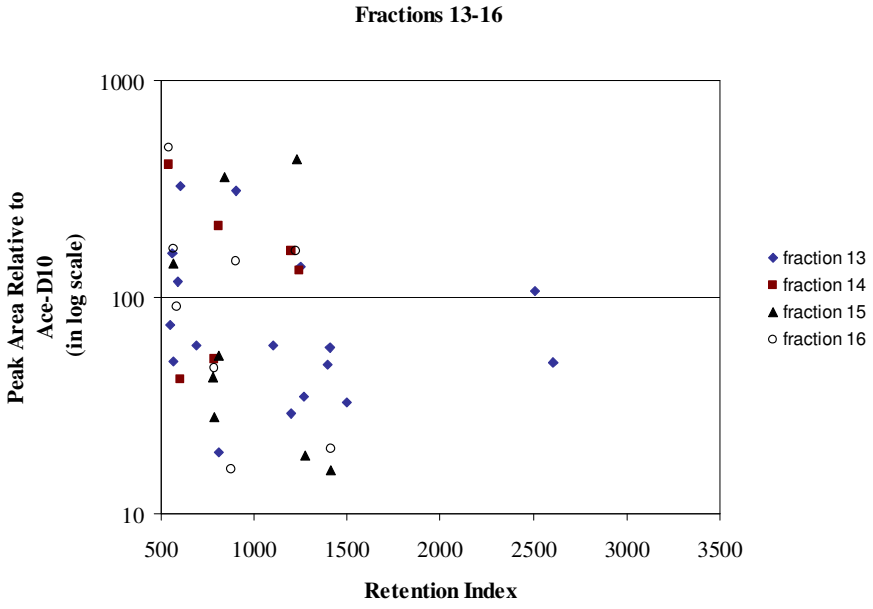


Figure 4-22: Peak area dynamic range of compounds of LC fractions 13 to 16.

4.15 Comparison of GC-MS "Whole" Sample to LC-MS Blood Plasma data

Classifying the identified compounds in GC-MS by major chemical groups resulted in four compound classes (Table 4-7).

Table 4-8: Percentage compound class representation in GC-MS and LC-MS of positively identified compounds.

Compound Class	% in GC-MS	% in LC-MS
Amino acids	29	60
Carbohydrates	38	0
Organic acids	27	20
Lipids[a]	6	20
a: compound class includes fatty acids, sterols and fatty acid methyl esters		

The major chemical groups in LC-MS resulted in three compound classifications (Table 4-7). The difference in compound class identification suggests the extent of metabolite coverage in the use of GC-MS and LC-MS. The absence of carbohydrate compound class (which represent the most abundant compounds in the blood plasma) in LC-MS and the high percentage of amino acid compounds reveals the strengths and weakness of the two analytical platforms used in the analysis of metabolites.

4.16 Comparison of Data : GC-MS "whole" Sample to GC-MS of LC Fractions of Blood Plasma.

The observation of the TIC of the LC fractions of "whole sample" shows a reduction of number of peaks per sample due to a reduction of peak overlaps as compared to the TIC of GC-MS of "whole sample" (Figure 4-23). The reduction in peak overlaps

and background noise resulted in doubling the number of peaks of human blood plasma analyzed by GC-MS (Table 4-9).

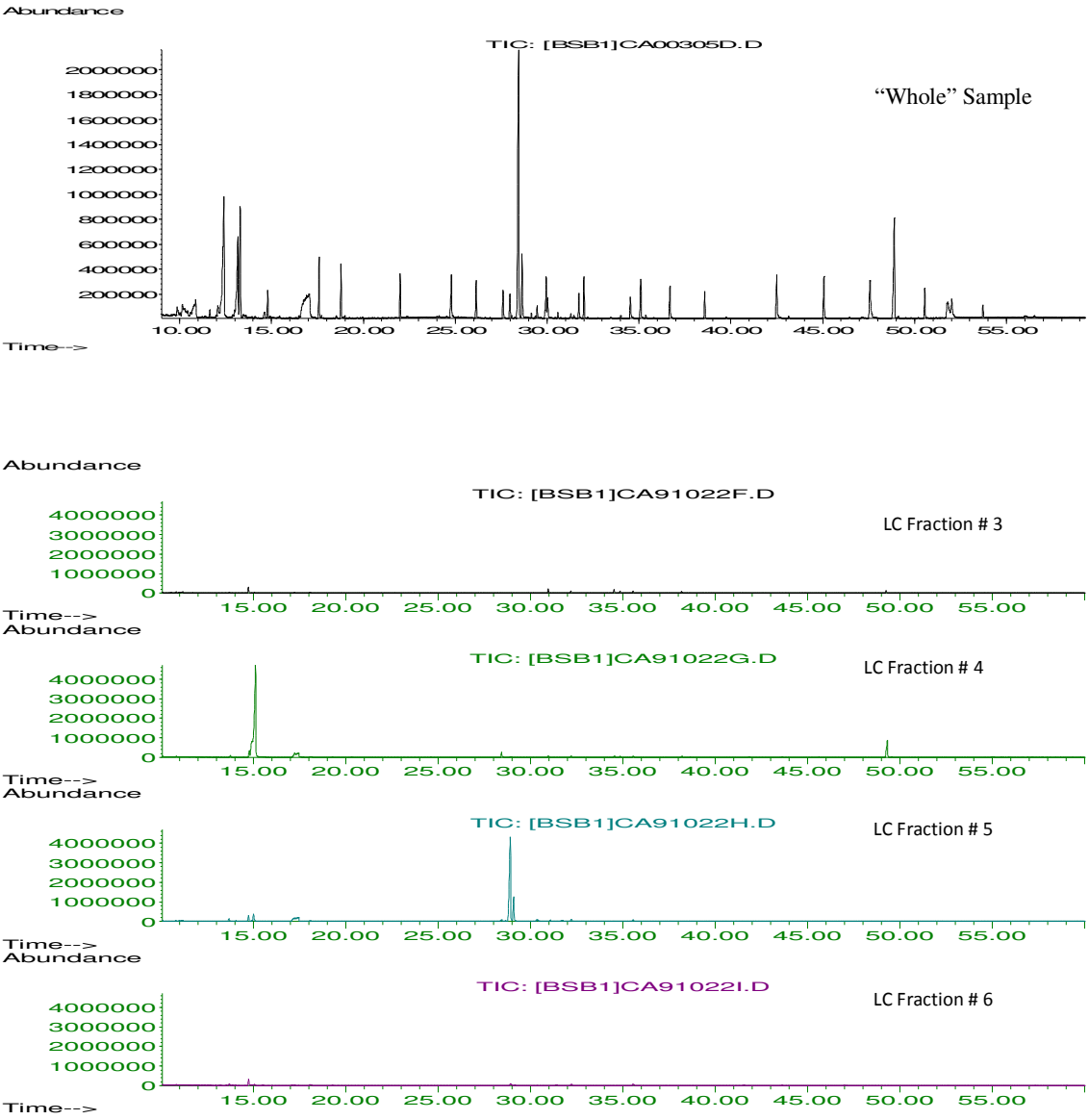


Figure 4-23: TIC of "whole" sample and four LC fractions showing the reduction in peak numbers and the elution of peaks in different portions of the chromatogram.

Comparing the total number of peaks in GC-MS of "whole sample" to the number of peaks generated in the GC-MS of LC fractions, an additional 176 new peaks were detected as result of LC fractionation of blood plasma samples (Table 4-9). Of the 176 peaks, 18 were additional positively identified peaks while 166 peaks were identified as unknowns.

Table 4-9: A summary table comparing the number of peaks generated by GC-MS analysis of "whole" and LC-fractionated blood plasma samples.

Analytical Technique	Total Number of Peaks	Positively Identified Peaks	Tentatively Identified Peaks	Unknown Peaks
GC-MS of "Whole Sample"	100	44	8	48
GC-MS of LC Fractions	276	62	8	206[a]

a: Number includes unknown peaks in GC-MS of "Whole Sample"

Comparing the number of positively identified compounds in the "whole sample" to those in the LC fractionated samples, additional compounds identified in the amino acid group showed a 90% increase in number of peaks, while the additional identified compounds in lipids compound class had a 67% increase followed by a 15% increase in the carbohydrate class of compounds and a 50% increase in the organic acid compound class (Table 4-10)

Table 4-10: A summary table of the number of compounds positively identified in four compound classes of peaks generated by GC-MS of "whole sample" and LC fractions

Compound Class	Number of Compounds ("Whole Sample")	Number of Compounds (LC Fractions)
Amino Acids	10	19
Carbohydrates	13	15
Lipids[a]	9	15
Organic Acids	2	3

a: compound class includes fatty acids, sterols and terpenoids, and fatty acid methyl esters

The signal-to-noise ratios improved for many compounds when the mass chromatograms from the analysis of "whole sample" and LC fractions were compared. An example of such a comparison is shown by comparing Figure 4-24 to Figure 4-25. In Figure 4-24, the ethanolamine TMS_3 and valine TMS_2 peaks had signal-to-noise ratios of 10 and 6 respectively. In the LC fractions, ethanolamine TMS_3 elutes in fraction number 5 with a three times improved signal-to-noise ratio of 30 while valine TMS_2 elutes in fraction number 6 with a five times improved signal-to-noise ratio of 30.

4.17 Comparison of Data: GC-MS of LC fractions to LC-MS of Human Blood Plasma.

Different compound classes elute in different regions in the GC-MS and LC-MS analysis. The elution regions of the different compound classes in GC-MS of "whole sample" were shown in Figure 4-8. For example, carbohydrates elutes early in the HILIC LC analysis (fractions 4-7, Figure 4-26).

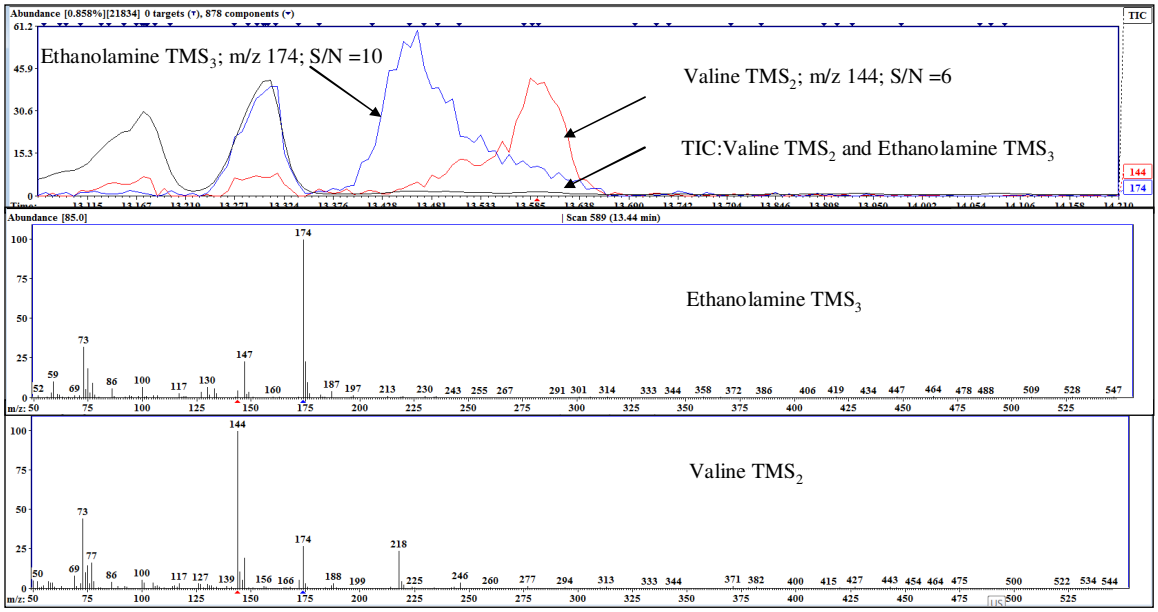


Figure 4-24: Mass chromatogram showing the deconvolution of two peaks in the GC-MS analysis of "whole" sample, ethanolamine TMS_3 and valine TMS_2 and corresponding S/N ratios.

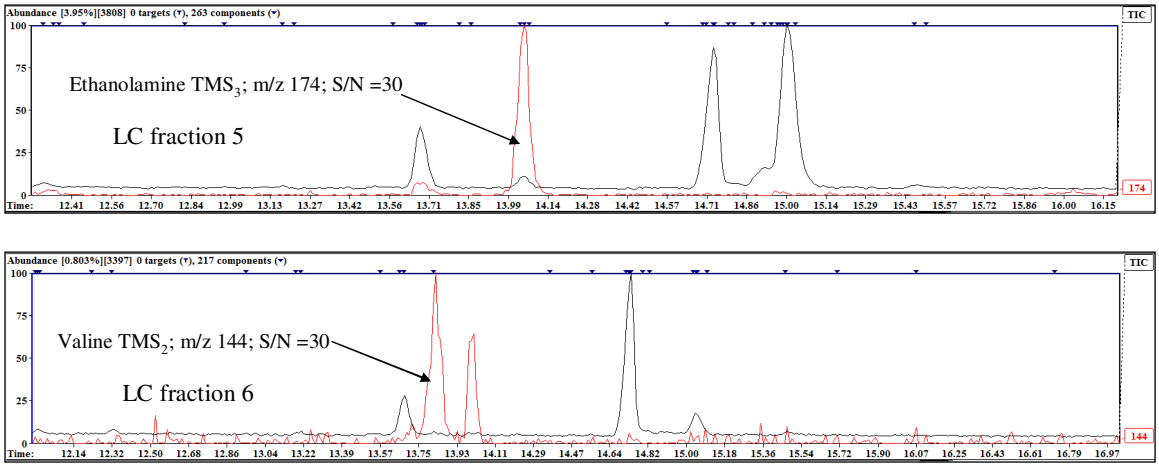


Figure 4-25: Mass chromatogram showing the elution of ethanolamine TMS_3 and valine TMS_2 in different LC fractions and with improved S/N ratios.

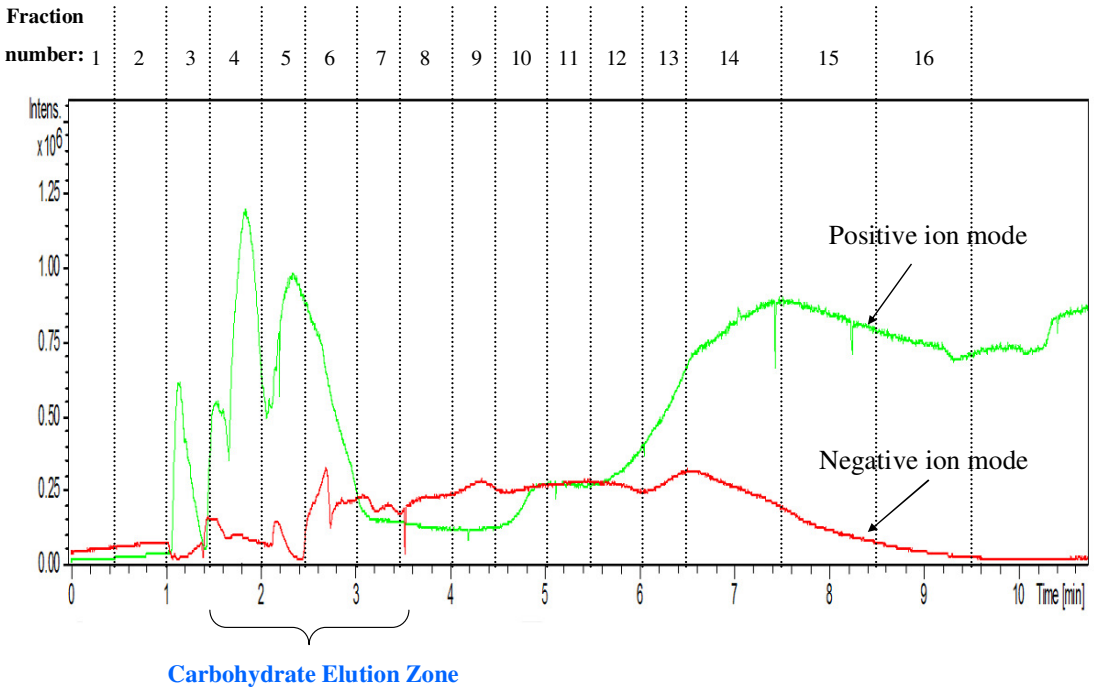


Figure 4-26: Total ion current chromatograms of LC fractions showing the elution region (1.5 min. to 3.5min.) of carbohydrates in fractions 4-7.

The analysis of the 16 LC fractions of the sample by GC-MS did show the elution of four major compound classes in the LC-MS analysis (see Table 4-6). Carbohydrates are not detected in LC-MS. However the analysis of the LC fractions show that had it been detected it would have eluded from between 1.5 minutes to 3.5 minutes as shown in Figure 4-26.

5.0 Conclusions and Future Work

A protocol for the analysis of human blood plasma by derivatization GC-EI-MS has been developed. This protocol includes sample handling, metabolite extraction and derivatization methods. The protocol included the introduction of recovery and derivatization standards that were useful for the determination of losses due to sample handling, extraction and matrix effects in addition to allowing an easy way to monitor completion of derivatization reactions.

The use of agilent chemStation software for data pre-treatment was complemented by NIST's AMDIS retention indices and mass spectral deconvolution software package which allowed matching mass spectra and retention indices of deconvoluted chromatographic peaks to mass spectra in our in-house library.

There are serious shortcoming when using public mass spectra and retention index libraries to make positive identification of compounds. The development of an in-house library remains the best option for positive identification of compounds. However, while RI values of identified compounds in this study highly correlate to those in our in-house library the same could not always be said of mass spectral matches. Low abundance compounds gave very poor signal-to-noise ratios and poor mass spectra matches. We showed that good mass spectral matches were obtained for peaks with signal-to-noise ratios greater than 15.

Through the identification of different compound classes in the data generated by GC-MS and LC-MS, the weakness and the strengths of both analytical techniques have been demonstrated. The use of GC-MS has a much wider coverage than LC-MS in the

detection of compounds from various compound classes. LC-MS was very poor for the detection of carbohydrate compounds and showed variable detection of amino acids due to the wide range of relative response factors in ESI.

The number of peaks detected in the GC-MS analysis of "whole sample" was 100 while the GC-MS analysis of HILIC LC fractions was 276. Of the 100 peaks detected in the GC-MS of the "whole sample", 44 were positively identified. In the multi-dimensional method, an additional 18 peaks were identified. This result is in line with published data in articles where 20-40% of detected peaks are positively identified.

A searchable retention index and mass spectral database outside of agilent's chemStation and AMDIS software has been developed. This database is linked to the names, retention index values, chemical formula and mass spectra information of the 160 metabolite compounds in our in-house library. The compounds in the library are mainly amino acids, carbohydrates, lipids, amines and organic acids. The low percentage number of positively identified compounds is due to the limited number of compounds in the in-house library. It is recommended that the number of compounds in the in-house library be expanded by adding the retention indices and mass spectra of authentic standards and also by correlating these retention indices and mass spectra to those in public libraries.

Future work will focus on; [1] increasing coverage of metabolite peaks. This would be done in using two-dimensional analytical techniques like the 2D-GC. The superiority of the 2D-GC to 1D-GC in reducing chromatographic co-elution, enhancing spectral purity and improving structural elucidation of peaks would result in an increase coverage of metabolite peaks, [2] expanding GC-EI-MS retention time index and mass

spectral database in our in-house library, to enable us identify more unknown chemical derivatives. This would be done by the addition of retention indices and mass spectral data of authentic chemical standards using GC-EI-MS.

6. References

1. 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. HMDB: the human metabolome database. *Nucleic Acids Research* **2007,** *35,* D521-D526.

2. Hollywood, K.; Brison, D.R.; Goodacre, R. Metabolomics: Current technologies and future trends. *Proteomics* **2006,** *6,* 4716-4723.

3. Pearson, H. Meet the human metabolome. *Nature* **2007,** *446,* 8-8.

4. Fiehn, O. Metabolomics - the link between genotypes and phenotypes. *Plant Molecular Biology* **2002,** *48,* 155-171.

5. Fiehn, O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics* **2001,** *2,* 155-168.

6. Psychogios, N.; Hau, D.D.; Peng, J.; Guo, A.C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.; Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Xia, J.G.; Knox, C.; Dong, E.; Huang, P.; Hollander, Z.; Pedersen, T.L.; Smith, S.R.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J.W.; Goodfriend, T.; Wishart, D.S. The Human Serum Metabolome. *Plos One* **2011,** *6,*

7. Lawton, K.,A.; Berger, A.; Mitchell, M.; Milgram, K.E.; Evans, A.,M.; Guo, L.; Hanson, R.,W.; Kalhan, S.,C.; Ryals, J.,A.; Milburn, M.,V. Analysis of the adult human plasma metabolome.(Clinical report). *Pharmacogenomics* **2008,** *9,* 383(15).

8. Dunn, W.B. and Ellis, D.I. Metabolomics: Current analytical platforms and methodologies. *Trac-Trends in Analytical Chemistry* **2005,** *24,* 285-294.

9. Villas-Boas, S.G.; Mas, S.; Akesson, M.; Smedsgaard, J.; Nielsen, J. Mass spectrometry in metabolome analysis. *Mass Spectrometry Reviews* **2005,** *24,* 613-646.

10. Dunn, W.B.; Bailey, N.J.C.; Johnson, H.E. Measuring the metabolome: current analytical technologies. *Analyst* **2005,** *130,* 606-625.

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

Daouk, R.; Mendes, P.; van Ommen, B.; Lindon, J.C.; Sansone, S. The metabolomics standards initiative (MSI). *Metabolomics* **2007,** *3,* 175-178.

12. García-Pérez, I.; Vallejo, M.; García, A.; Legido-Quigley, C.; Barbas, C. Metabolic fingerprinting with capillary electrophoresis. *Journal of Chromatography A* **2008,** *1204,* 130-139.

13. Nicholson, J.K.; Lindon, J.C.; Holmes, E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **1999,** *29,* 1181-1189.

14. Nicholson, J.K. and Lindon, J.C. Systems biology - Metabonomics. *Nature* **2008,** *455,* 1054-1056.

15. Robertson, D.G. Metabonomics in toxicology: A review. *Toxicological Sciences* **2005,** *85,* 809-822.

16. Jun Han, Raju Datla, Sammy Chan, Christoph H Borchers Mass spectrometry-based technologies for high-throughput metabolomics. *Bioanalysis* **2009,** *1,* 1665.

17. Robertson D.,Lindon J.,Nicholson J.K.,Holmes E., Metabonomics in Toxicity Assessment. *CRC Press:Boca Raton,FL,* **2005,**

18. Vaidyanathan S.,Harrigan G.G.,Goodacre R., Metabolome Analyses:Strategies for System Biology. *Springer:New York* **2005,**

19. Robertson, D.G.; Reily, M.D.; Sigler, R.E.; Wells, D.F.; Paterson, D.A.; Braden, T.K. Metabonomics: Evaluation of nuclear magnetic resonance (NMR) and pattern recognition technology for rapid in vivo screening of liver and kidney toxicants. *Toxicological Sciences* **2000,** *57,* 326-337.

20. Coen, M.; Lenz, E.M.; Nicholson, J.K.; Wilson, I.D.; Pognan, F.; Lindon, J.C. An integrated metabonomic investigation of acetaminophen toxicity in the mouse using NMR spectroscopy. *Chemical Research in Toxicology* **2003,** *16,* 295-303.

21. Nicholson, J.K.; Foxall, P.J.D.; Spraul, M.; Farrant, R.D.; Lindon, J.C. 750-MHZ H-1 AND H-1-C-13 NMR-SPECTROSCOPY OF HUMAN BLOOD-PLASMA. *Analytical Chemistry* **1995,** *67,* 793-811.

22. Holmes, E.; Wilson, I.D.; Nicholson, J.K. Metabolic phenotyping in health and disease. *Cell* **2008,** *134,* 714-717.

23. Johnson, H.E.; Broadhurst, D.; Goodacre, R.; Smith, A.R. Metabolic fingerprinting of salt-stressed tomatoes. *Phytochemistry* **2003,** *62,* 919-928.

24. Mordechai, S.; Sahu, R.K.; Hammody, Z.; Mark, S.; Kantarovich, K.; Guterman, H.; Podshyvalov, A.; Goldstein, J.; Argov, S. Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma. *Journal of Microscopy-Oxford* **2004,** *215,* 86-91.

25. Psychogios, N.; Hau, D.D.; Peng, J.; Guo, A.C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.; Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Xia, J.; Knox, C.; Dong, E.; Huang, P.; Hollander, Z.; Pedersen, T.L.; Smith, S.R.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J.W.; Goodfriend, T.; Wishart, D.S. The Human Serum Metabolome. *Plos One* **2011,** *6,* e16957.

26. Yuan, K.; Kong, H.; Guan, Y.; Yang, J.; Xu, G. A GC-based metabonomics investigation of type 2 diabetes by organic acids metabolic profile. *Journal of Chromatography B* **2007,** *850,* 236-240.

27. Sanchez, D.H.; Siahpoosh, M.R.; Roessner, U.; Udvardi, M.; Kopka, J. Plant metabolomics reveals conserved and divergent metabolic responses to salinity. *Physiol. Plantarum* **2008,** *132,* 209-219.

28. Koek, M.M.; Muilwijk, B.; van Stee, L.L.P.; Hankemeier, T. Higher mass loadability in comprehensive two-dimensional gas chromatography–mass spectrometry for improved analytical performance in metabolomics analysis. *Journal of Chromatography A* **2008,** *1186,* 420-429.

29. Wu, Z.; Huang, Z.; Lehmann, R.; Zhao, C.; Xu, G. The Application of Chromatography-Mass Spectrometry: Methods to Metabonomics. *Chroma* **2009,** *69,* 23-32.

30. Issaq, H.J.; Chan, K.C.; Janini, G.M.; Conrads, T.P.; Veenstra, T.D. Multidimensional separation of peptides for effective proteomic analysis. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2005,** *817,* 35-47.

31. Jia, L.; Liu, B.; Terabe, S.; Nishioka, T. Two-Dimensional Separation Method for Analysis of Bacillus subtilis Metabolites via Hyphenation of Micro-Liquid Chromatography and Capillary Electrophoresis. *Anal. Chem. 76,* 1419-1428.

32. Wang, Y.; Wang, J.; Yao, M.; Zhao, X.; Fritsche, J.; Schmitt-Kopplin, P.; Cai, Z.; Wan, D.; Lu, X.; Yang, S.; Gu, J.; Häring, H.U.; Schleicher, E.D.; Lehmann, R.; Xu, G. Metabonomics Study on the Effects of the Ginsenoside Rg3 in a β-Cyclodextrin-Based Formulation on Tumor-Bearing Rats by a Fully Automatic Hydrophilic

Interaction/Reversed-Phase Column-Switching HPLC–ESI-MS Approach. *Anal. Chem.* **2008,** *80,* 4680-4688.

33. Eva Maria Lenz and Ian D. Wilson Analytical Strategies in Metabolomics. *Proteome Research* **2007,** *6,* 443-458.

34. Issaq, H.; Van, Q.; Waybright, T.; Muschik, G.; Veenstra, T. Analytical and statistical approaches to metabolomics research. *Journal of separation science* **2009,** *32,* 2183.

35. Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T. Quantitative Metabolome Analysis Using Capillary Electrophoresis Mass Spectrometry. *Journal of Proteome Research* **2003,** *2,* 488-494.

36. Saito, N.; Robert, M.; Kitamura, S.; Baran, R.; Soga, T.; Mori, H.; Nishioka, T.; Tomita, M. Metabolomics approach for enzyme discovery. *Journal of Proteome Research* **2006,** *5,* 1979-1987.

37. Ramautar, R.; Somsen, G.W.; de Jong, G.J. CE-MS in metabolomics. *Electrophoresis* **2009,** *30,* 276-291.

38. Monton, M.R.N. and Soga, T. Metabolome analysis by capillary electrophoresis-mass spectrometry. *Journal of Chromatography A* **2007,** *1168,* 237-246.

39. Lapainis, T.; Rubakhin, S.S.; Sweedler, J.V. Capillary Electrophoresis with Electrospray Ionization Mass Spectrometric Detection for Single-Cell Metabolomics. *Anal. Chem.* **2009,** *81,* 5858-5864.

40. Wang, Y.; Xiao, J.; Suzek, T.O.; Zhang, J.; Wang, J.; Bryant, S.H. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* **2009,** *37,* W623-W633.

41. PubChem-compound and species-specific database. *http://pubchem.ncbi.nih.gov.libaccess.lib.mcmaster.ca/*

42. 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. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res.* **2009,** *37,* D603-D610.

43. Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E.A.; Glass, C.K.; Merrill, A.H., Jr.; Murphy, R.C.; Raetz, C.R.H.; Russell, D.W.; Subramaniam, S. LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* **2007,** *35,* D527-D532.

44. Lipid Maps-compound ana species-specific database. *http://www.lipidmaps.org/data/databases.html*

45. KNApSAcK-compound and species-specific database. *http://kanaya.naist.jp/KNApSAcK/*

46. Brown, M.; Dunn, W.B.; Dobson, P.; Patel, Y.; Winder, C.L.; Francis-McIntyre, S.; Begley, P.; Carroll, K.; Broadhurst, D.; Tseng, A.; Swainston, N.; Spasic, I.; Goodacre, R.; Kell, D.B. Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *Analyst* **2009,** *134,* 1322-1332.

47. MMD-compound and species-specific database. *http://dbkgroup.org/MMD/*

48. NIST08-reference spectral lidrary/database. *www.nist.gov/srd/nist1.htm*

49. Kopka, J.; Schauer, N.; Krueger, S.; Birkemeyer, C.; Usadel, B.; Bergmuller, E.; Dormann, P.; Weckwerth, W.; Gibon, Y.; Stitt, M.; Willmitzer, L.; Fernie, A.R.; Steinhauser, D. GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics* **2005,** *21,* 1635-1638.

50. Kopka, J. Current challenges and developments in GC-MS based metabolite profiling technology. *J. Biotechnol.* **2006,** *124,* 312-322.

51. Smith, C.A.; O'Maille, G.; Want, E.J.; Qin, C.; Trauger, S.A.; Brandon, T.R.; Custodio, D.E.; Abagyan, R.; Siuzdak, G. METLIN - A metabolite mass spectral database. *Therapeutic Drug Monitoring* **2005,** *27,* 747-751.

52. METLIN-reference spectral library/database. *http://metlin.scripps.edu/*

53. MassBank-reference spectral library/database. *www.massbank.jp/*

54. Matsuda, F.; Yonekura-Sakakibara, K.; Niida, R.; Kuromori, T.; Shinozaki, K.; Saito, K. MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. *Plant Journal* **2009,** *57,* 555-577.

55. Cui, Q.; Lewis, I.A.; Hegeman, A.D.; Anderson, M.E.; Li, J.; Schulte, C.F.; Westler, W.M.; Eghbalnia, H.R.; Sussman, M.R.; Markley, J.L. Metabolite identification via the Madison Metabolomics Consortium Database. *Nat. Biotechnol.* **2008,** *26,* 162-164.

56. Kanehisa M,Goto S,Hattori M, From genomics to chemical genomics:new development in KEGG. *Nucleic Acids Research* **2006,** *34,* D354-D357.

57. Karp, P.D.; Ouzounis, C.A.; Moore-Kochlacs, C.; Goldovsky, L.; Kaipa, P.; Ahren, D.; Tsoka, S.; Darzentas, N.; Kunin, V.; Lopez-Bigas, N. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic Acids Res.* **2005,** *33,* 6083-6089.

58. Krummenacker, M.; Paley, S.; Mueller, L.; Yan, T.; Karp, P.D. Querying and computing with BioCyc databases. *Bioinformatics* **2005,** *21,* 3454-3455.

59. BioCyc-pathway-specific database. *http://biocyc.org*

60. Karp P.D.,Collado-Vidas J.,Paulsen I.,Gunsalus R. EcoCyc-pathway-specific database http://ecocyc.org/. **2011,**

61. Caspi, R.; Foerster, H.; Fulcher, C.A.; Hopkinson, R.; Ingraham, J.; Kaipa, P.; Krummenacker, M.; Paley, S.; Pick, J.; Rhee, S.Y.; Tissier, C.; Zhang, P.; Karp, P.D. MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res.* **2006,** *34,* D511-D516.

62. Krieger, C.J.; Zhang, P.F.; Mueller, L.A.; Wang, A.; Paley, S.; Arnaud, M.; Pick, J.; Rhee, S.Y.; Karp, P.D. MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res.* **2004,** *32,* D438-D442.

63. Karp, P.D.; Riley, M.; Saier, M.; Paulsen, I.T.; Collado-Vides, J.; Paley, S.M.; Pellegrini-Toole, A.; Bonavides, C.; Gama-Castro, S. The EcoCyc database. *Nucleic Acids Res.* **2002,** *30,* 56-58.

64. Romero, P.; Wagg, J.; Green, M.L.; Kaiser, D.; Krummenacker, M.; Karp, P.D. Computational prediction of human metabolic pathways from the complete human genome. *Genome Biology* **2005,** *6,*

65. HumanCyc-pathway-specific database. *http://humancyc.org/*

66. Matthews, L.; Gopinath, G.; Gillespie, M.; Caudy, M.; Croft, D.; de Bono, B.; Garapati, P.; Hemish, J.; Hermjakob, H.; Jassal, B.; Kanapin, A.; Lewis, S.; Mahajan, S.; May, B.; Schmidt, E.; Vastrik, I.; Wu, G.; Birney, E.; Stein, L.; D'Eustachio, P. Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Res.* **2009,** *37,* D619-D622.

67. Vastrik, I.; D'Eustachio, P.; Schmidt, E.; Gopinath, G.; Croft, D.; de Bono, B.; Gillespie, M.; Jassal, B.; Lewis, S.; Matthews, L.; Wu, G.; Birney, E.; Stein, L.

Reactome: a knowledge base of biologic pathways and processes (vol 8, pg 39, 2007). *Genome Biol.* **2009,** *10,* 402.

68. Vastrik, I.; D'Eustachio, P.; Schmidt, E.; Joshi-Tope, G.; Gopinath, G.; Croft, D.; de Bono, B.; Gillespie, M.; Jassal, B.; Lewis, S.; Matthews, L.; Wu, G.; Birney, E.; Stein, L. Reactome: a knowledge base of biologic pathways and processes. *Genome Biol.* **2007,** *8,* R39.

69. Büscher, J.M.; Czernik, D.; Ewald, J.C.; Sauer, U.; Zamboni, N. Cross-Platform Comparison of Methods for Quantitative Metabolomics of Primary Metabolism. *Anal. Chem.* **2009,** *81,* 2135-2143.

70. Hagan, S.O.; Dunn, W.B.; Knowles, J.D.; Broadhurst, D.; Williams, R.; Ashworth, J.J.; Cameron, M.; Kell, D.B. Closed-Loop, Multiobjective Optimization of Two-Dimensional Gas Chromatography/Mass Spectrometry for Serum Metabolomics. *Anal. Chem.* **2007,** *79,* 464-476.

71. A, J.; Trygg, J.; Gullberg, J.; Johansson, A.I.; Jonsson, P.; Antti, H.; Marklund, S.L.; Moritz, T. Extraction and GC/MS Analysis of the Human Blood Plasma Metabolome. *Anal. Chem.* **2005,** *77,* 8086-8094.

72. Nicholson, J.K. Omics dreams of personalized healthcare. *J. Proteome Res.* **2006,** *5,* 2067-+.

73. Holmes, E.; Tsang, T.M.; Huang, J.T.-.; Leweke, F.M.; Koethe, D.; Gerth, C.W.; Nolden, B.M.; Gross, S.; Schreiber, D.; Nicholson, J.K.; Bahn, S. Metabolic profiling of CSF: Evidence that early intervention may impact on disease progression and outcome in schizophrenia. *Plos Medicine* **2006,** *3,* 1420-+.

74. Holmes, E.; Loo, R.L.; Stamler, J.; Bictash, M.; Yap, I.K.S.; Chan, Q.; Ebbels, T.; De Iorio, M.; Brown, I.J.; Veselkov, K.A.; Daviglus, M.L.; Kesteloot, H.; Ueshima, H.; Zhao, L.C.; Nicholson, J.K.; Elliott, P. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* **2008,** *453,* 396-U50.

75. Sindic, A.; Chang, M.H.; Mount, D.B.; Romeroa, M.F. Renal physiology of SLC26 anion exchangers. *Current Opinion in Nephrology and Hypertension* **2007,** *16,* 484-490.

76. Martini F, O.W. *Fundamentals of anatomy & physiology.* Pearson Benjamin cummings: San Francisco CA, 2006; pp. 640-643.

77. Fiehn O., K.T., Metabolite profiling in blood plasma. *Methods in molecular biology* **2007,** *358,* 3-7.

78. Lindon, J.C.; Holmes, E.; Nicholson, J.K. Metabonomics in pharmaceutical R & D. *Febs Journal* **2007,** *274,* 1140-1151.

79. Perwaiz, S.; Tuchweber, B.; Mignault, D.; Gilat, T.; Yousef, I.M. Determination of bile acids in biological fluids by liquid chromatography-electrospray tandem mass spectrometry. *J. Lipid Res.* **2001,** *42,* 114-119.

80. 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.; Kellt, D.B.; HUSERMET Consortium Development of a Robust and Repeatable UPLC-MS Method for the Long-Term Metabolomic Study of Human Serum. *Anal. Chem.* **2009,** *81,* 1357-1364.

81. Dam, D. Characterization of organic constituents of environmental films by GC/MS. *M.SC. Thesis* **2005,**

82. Ding, J.; Sorensen, C.M.; Zhang, Q.; Jiang, H.; Jaitly, N.; Livesay, E.A.; Shen, Y.; Smith, R.D.; Metz, T.O. Capillary LC Coupled with High-Mass Measurement Accuracy Mass Spectrometry for Metabolic Profiling. *Anal. Chem.* **2007,** *79,* 6081-6093.

83. Palazoglu, M. and Fiehn, O. Metabolite identification in blood plasma using GC-MS and the Agilent Fiehn GC-MS metabolomics RTL library. **April 1, 2009,**

84. A, J.; Trygg, J.; Gullberg, J.; Johansson, A.I.; Jonsson, P.; Antti, H.; Marklund, S.L.; Moritz, T. Extraction and GC/MS Analysis of the Human Blood Plasma Metabolome. *Anal. Chem.* **2005,** *77,* 8086-8094.

85. Kopka, J.; Schauer, N.; Krueger, S.; Birkemeyer, C.; Usadel, B.; Bergmüller, E.; Dörmann, P.; Weckwerth, W.; Gibon, Y.; Stitt, M.; Willmitzer, L.; Fernie, A.R.; Steinhauser, D. GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics* **2005,** *21,* 1635-1638.

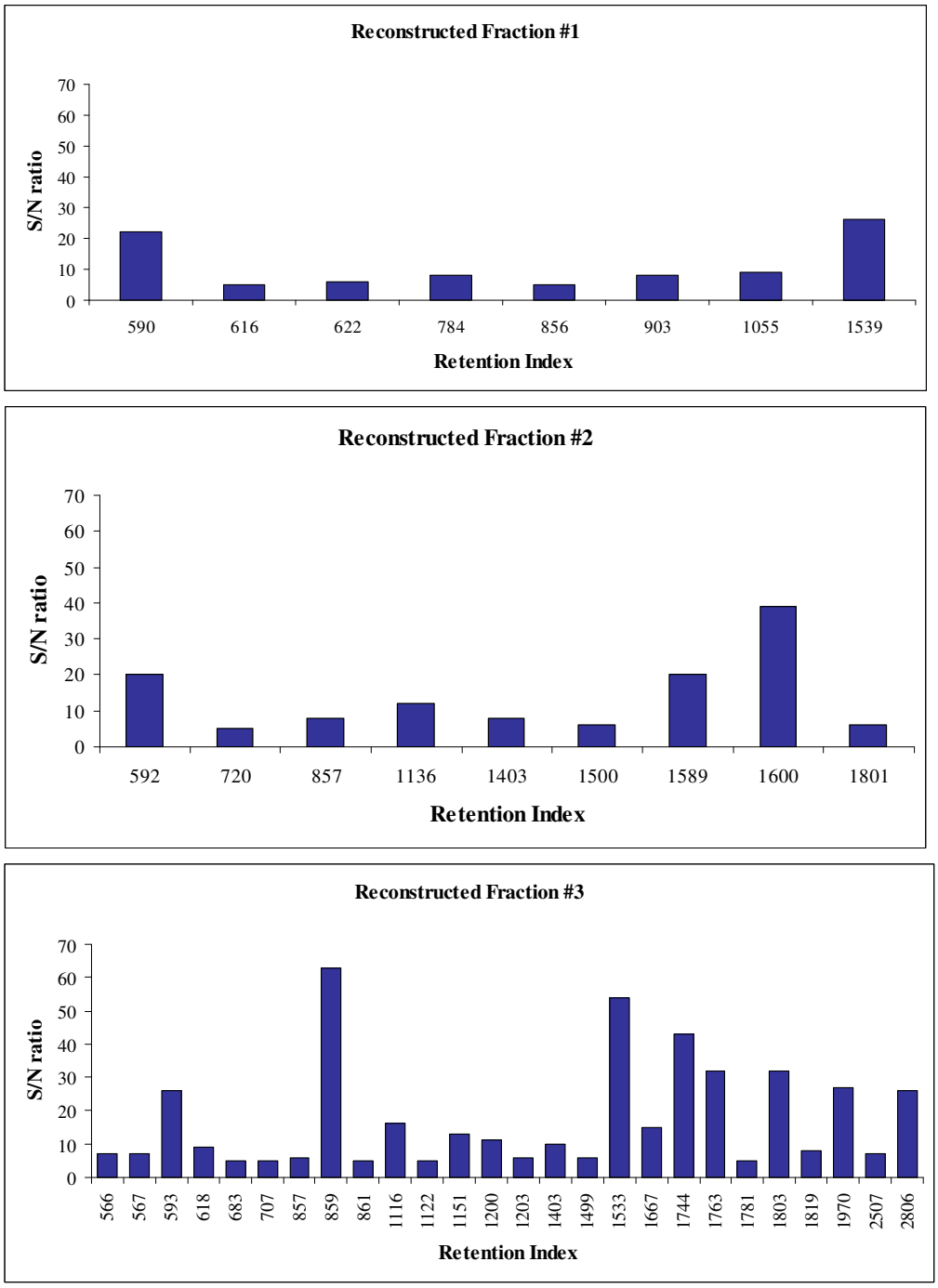
7. Appendices

Appendix 1: List of unknown peaks in GC-MS "Whole" Sample

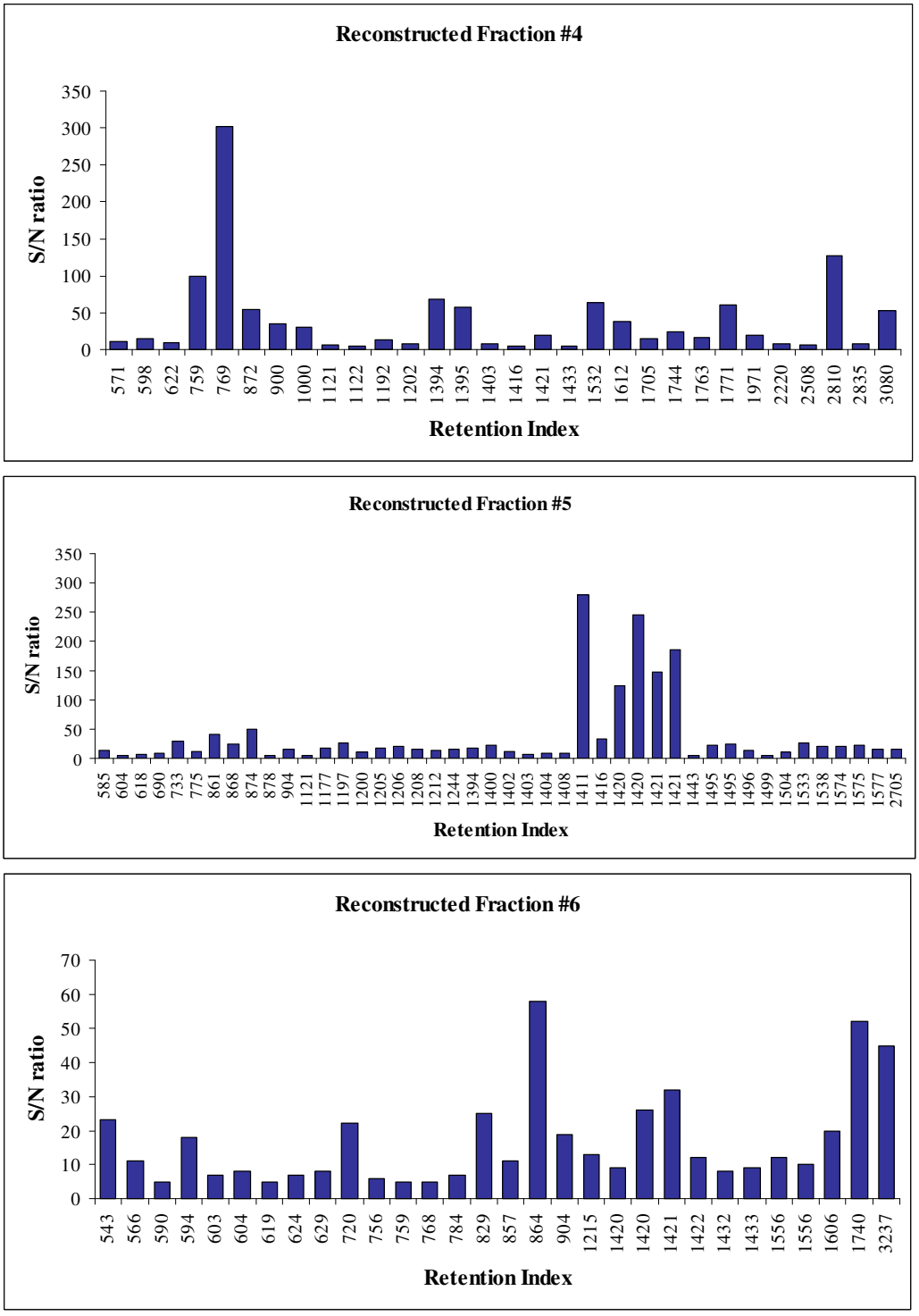
Number of unknown peaks in GC-MS of "Whole Sample": 48					
No.	**R.I.(DB-17ht, Fatty Acids)[a]**	**R.I.(DB-17ht, n-Alkanes)[b]**	**Identified derivatized compound[c]**	**m/z[d]**	**S/N(m/z)[e]**
CA00305D-P1-U1	551.8		unknown	130	32
CA00305D-P4-U2	620.2		unknown	274	7
CA00305D-P5-U3	631.7		unknown	119	31
CA00305D-P9-U4	676.8		**Fatty acid**	132	7
CA00305D-P11-U5	704.4		unknown	196	7
CA00305D-P15-U6	773.5		**Sugar alcohol**	204	58
CA00305D-P27-U7	887.8	1316.7	unknown	73	10
CA00305D-P30-U8	907.0	1338.8	**Fatty acid**	117	5
CA00305D-P31-U9	942.9	1380.4	unknown	327	9
CA00305D-P32-U10	954.4	1393.7	unknown	267	20
CA00305D-P33-U11	962.8	1403.4	unknown	133	14
CA00305D-P34-U12	973.6	1415.9	unknown	172	5
CA00305D-P36-U13	1006.7	1454.0	unknown	217	5
CA00305D-P37-U14	1056.0	1511.0	unknown	207	5
CA00305D-P38-U15	1102.8	1564.7	unknown	75	5
CA00305D-P39-U16	1117.9	1581.0	**Sugar alcohol**	204	9
CA00305D-P40-U17	1118.9	1582.1	unknown	355	8
CA00305D-P42-U18	1137.7	1602.5	**Fatty acid**	117	5
CA00305D-P43-U19	1152.8	1618.8	unknown	103	6
CA00305D-P45-U20	1182.7	1651.4	unknown	147	7
CA00305D-P46-U21	1193.8	1663.4	unknown	191	5
CA00305D-P55-U22	1215.6	1687.0	unknown	75	5
CA00305D-P56-U23	1222.1	1694.1	unknown	102	9
CA00305D-P58-U24	1232.2	1705.0	unknown	219	29
CA00305D-P60-U25	1238.3	1711.6	unknown	147	7
CA00305D-P61-U26	1245.1	1719.0	**Fatty acid**	117	5
CA00305D-P63-U27	1250.2	1724.6	unknown	221	6
CA00305D-P64-U28	1299.3	1777.7	**Fatty acid**	117	22
CA00305D-P65-U29	1338.7	1817.1	**FAME**	74	5
CA00305D-P66-U30	1375.4	1853.7	**Fatty acid**	312	15
CA00305D-P68-U31	1396.0	1874.3	unknown	103	52
CA00305D-P73-U32	1431.7	1909.8	**Sugar**	205	24
CA00305D-P74-U33	1441.5	1919.6	**Sugar/ sugar alcohol**	217	6
CA00305D-P76-U34	1474.2	1952.3	unknown	73	31
CA00305D-P77-U35	1484.0	1962.0	**Sugar alcohol**	204	6
CA00305D-P78-U36	1496.4	1974.5	unknown	147	20
CA00305D-P79-U37	1504.6	1981.9	**Sugar**	205	44
CA00305D-P81-U38	1537.4	2010.8	unknown	103	15
CA00305D-P82-U39	1538.1	2011.5	**Sugar alcohol**	204	17
CA00305D-P83-U40	1539.0	2012.2	unknown	273	19
CA00305D-P84-U41	1576.3	2045.1	**Sugar/ sugar alcohol**	217	8
CA00305D-P85-U42	1589.9	2057.2	**Sugar**	205	11
CA00305D-P88-U43	1666.0	2124.3	unknown	327	5
CA00305D-P90-U44	1740.4	2190.1	unknown	102	17
CA00305D-P91-U45	1771.9	2223.3	unknown	376	51
CA00305D-P95-U46	2110.2	2576.9	unknown	244	11
CA00305D-P96-U47	2702.6	3110.6	**Fatty acid**	117	7
CA00305D-P100-U48	3237.1	3528.4	unknown	207	32

a: Retention index values using fatty acids on a DB-17ht column
b: Retention index values using n-alkanes on a DB-17ht column
c: Compound classes are in bold; all others are unknown
d: Selected model ion that best describes peak
e: signal-to-noise ratio of selected fragment ion

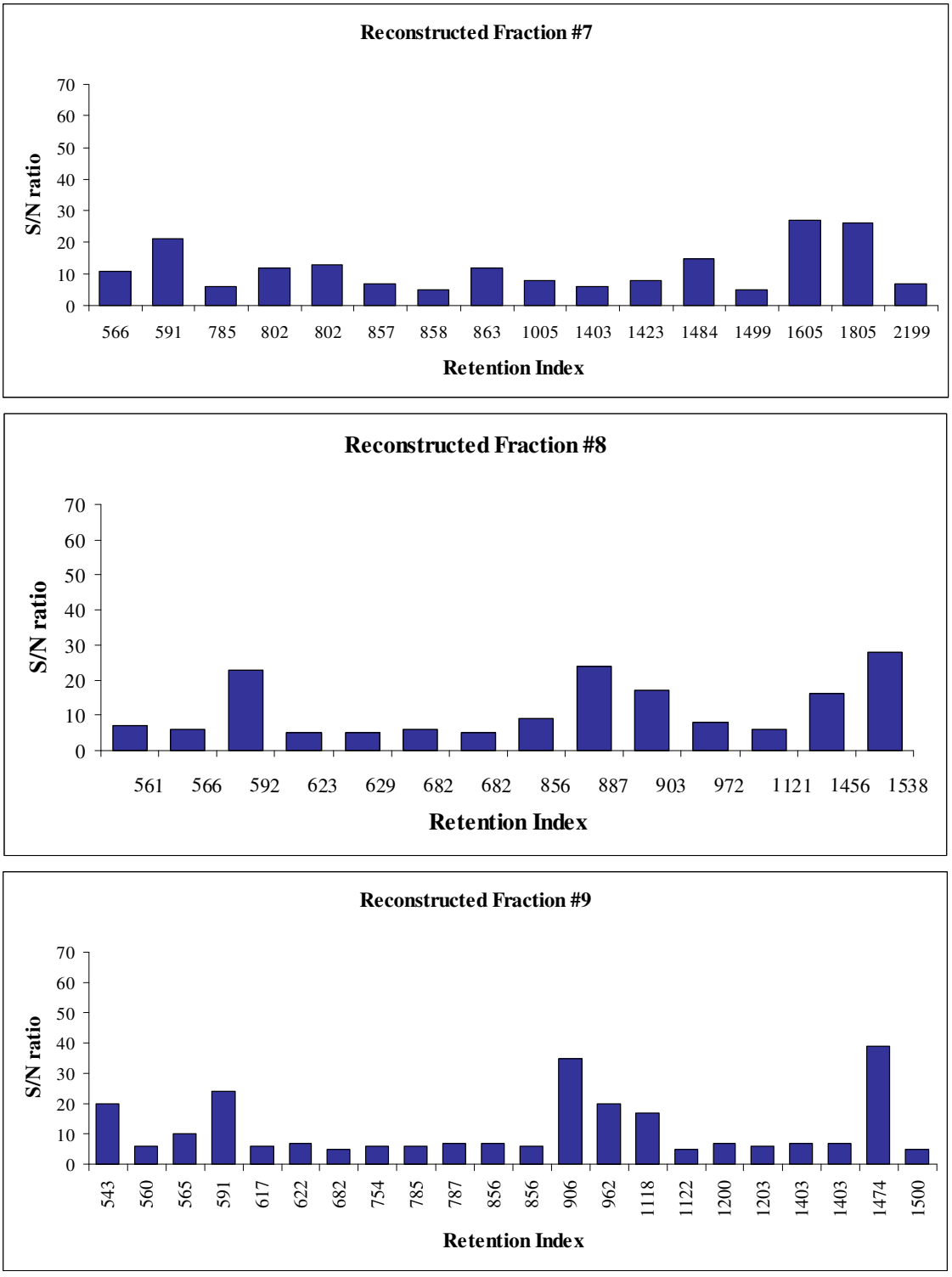
Appendix 2: Signal-to-noise ratios of peaks in LC fractions 1 – 16



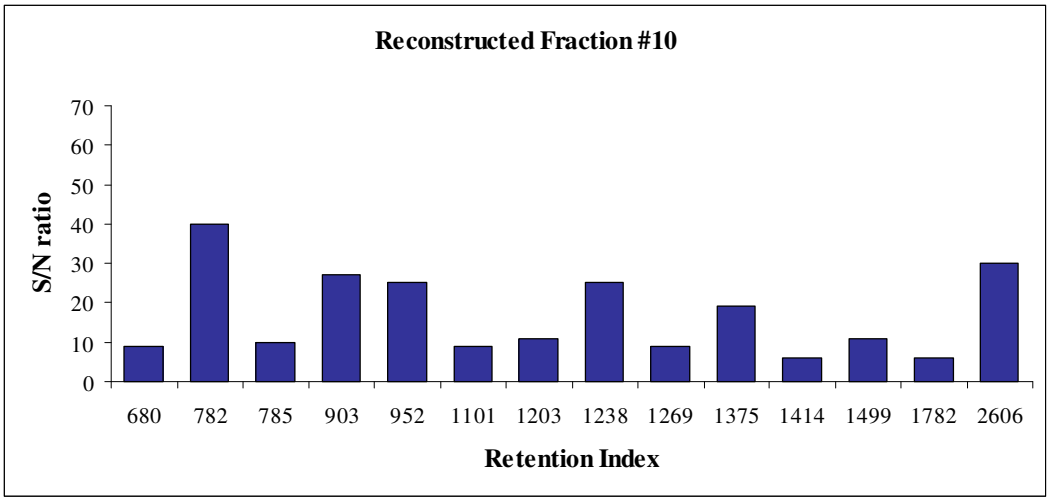
Appendix 2-1: Re-constructed GC-MS chromatogram of LC fraction numbers 1–3 showing number of peaks and their signal-to-noise ratios.

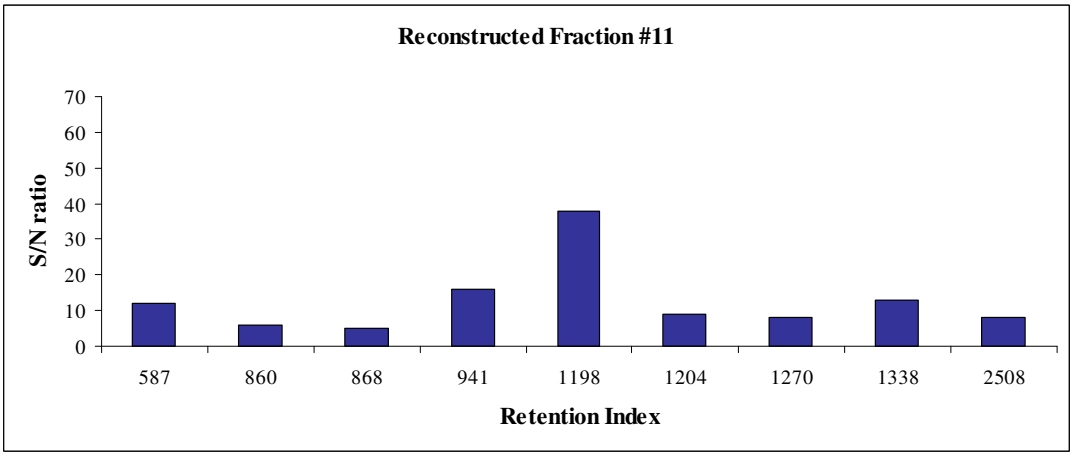


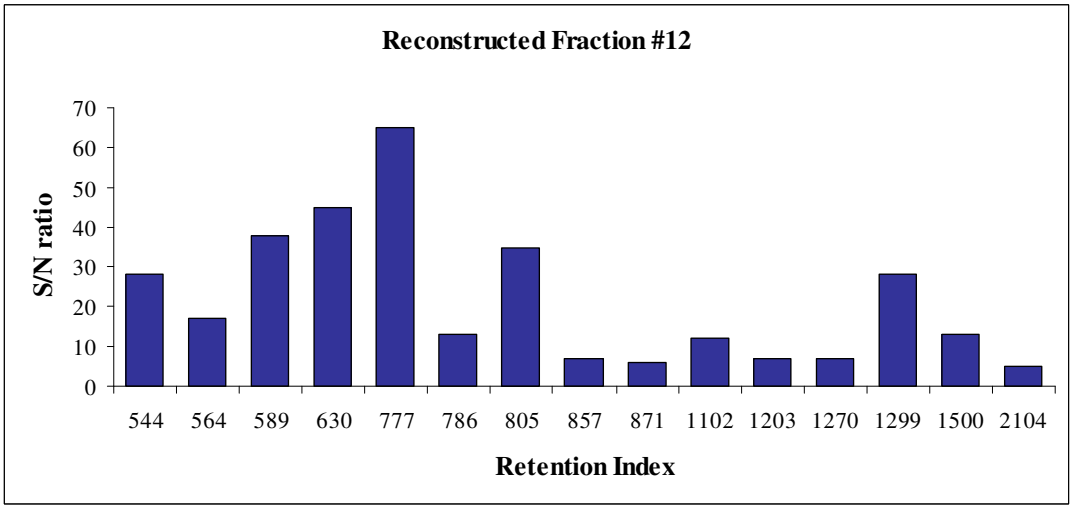
Appendix 2-2: Re-constructed GC-MS chromatogram of LC fraction numbers 4-6 showing number of peaks and their signal-to-noise ratios.



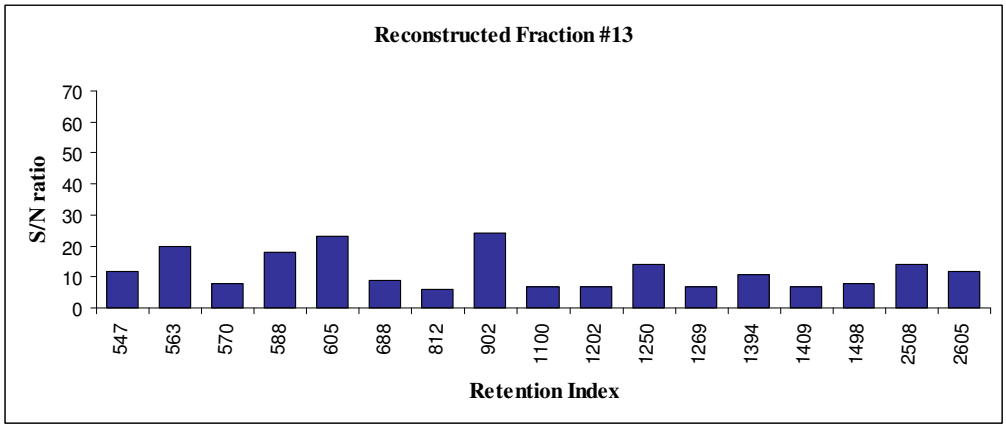
Appendix 2-3: Re-constructed GC-MS chromatogram of LC fraction numbers 7– 9 showing number of peaks and their signal-to-noise ratios.

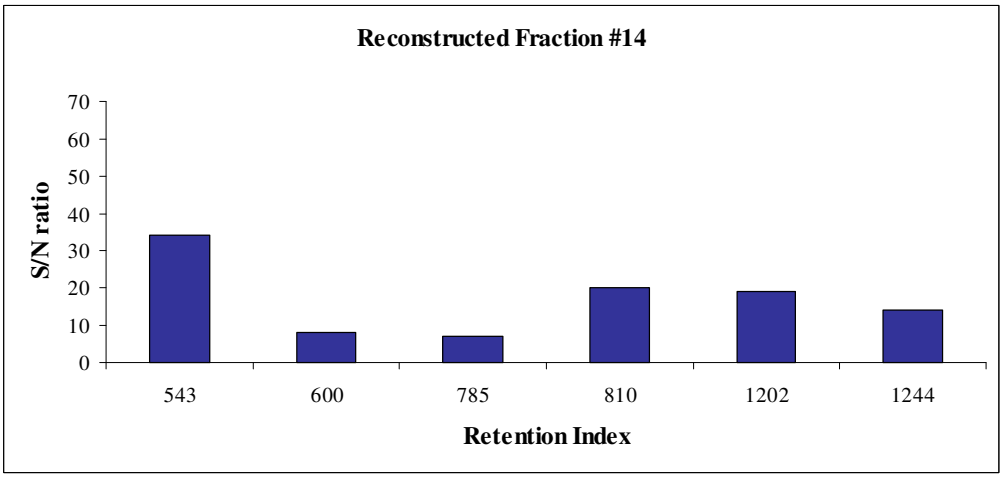


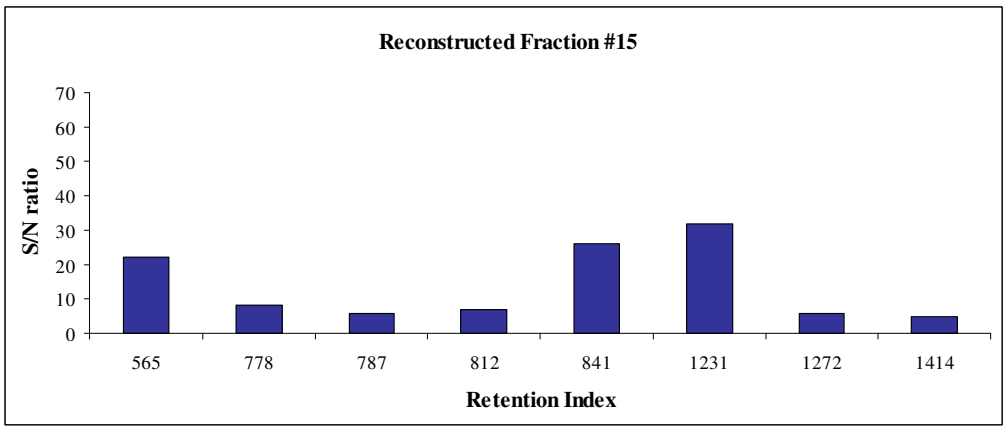




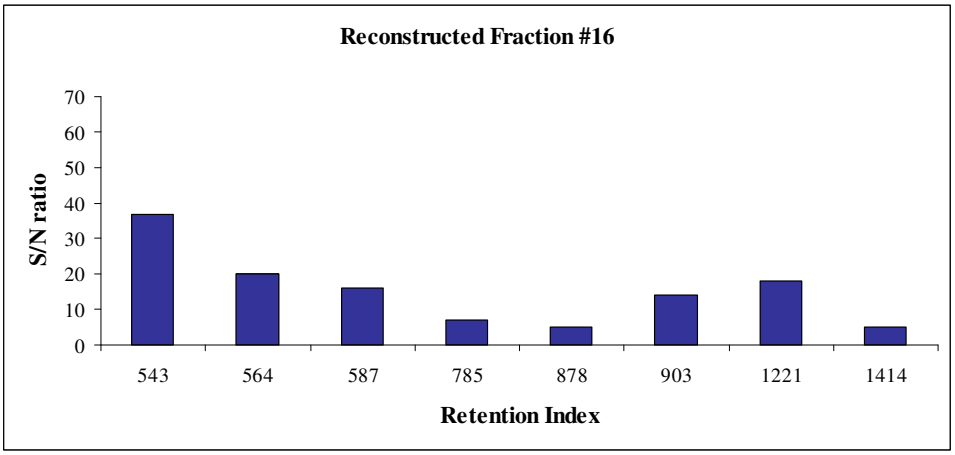
Appendix 2-4: Re-constructed GC-MS chromatogram of LC fraction numbers 10 –12 showing number of peaks and their signal-to-noise ratios.







Appendix 2-5: Re-constructed GC-MS chromatogram of LC fraction numbers 13 – 15 showing number of unique peaks and their signal-to-noise ratios.



Appendix 2-6: Re-constructed GC-MS chromatogram of LC fraction number 16 showing number of peaks and their signal-to-noise ratios.

Appendix 3: List of peaks in GC-MS of LC fractions of human blood plasma extract.

Total number of peaks: 276					
No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound [a]	m/z	S/N (m/z)
CA91022D-P1-U1	1	590.1	Unknown	117	22
CA91022D-P2-T1	1	616.0	*Hydroxylamine TMS*	152	5
CA91022D-P3-K1	1	644.7	**Glycine TMS_2**	147	6
CA91022D-P4-U2	1	784.5	Unknown	357	8
CA91022D-P5-U3	1	856.5	Unknown	232	5
CA91022D-P6-U4	1	903.2	Unknown	129	8
CA91022D-P7-U5	1	1055.3	Unknown	207	9
CA91022D-P8-U6	1	1539.4	Unknown	319	26
CA91022E-P1-U1	2	591.6	Unknown	117	20
CA91022E-P2-U2	2	720.3	Unknown	174	5
CA91022E-P3-U3	2	856.5	Unknown	147	8
CA91022E-P4-U4	2	1136.4	Unknown	117	12
CA91022E-P5-U5	2	1403.3	Unknown	133	8
CA91022E-P6-U6	2	1499.8	Unknown	299	6
CA91022E-P7-U7	2	1589.3	Unknown	205	20
CA91022E-P8-K1	2	1600.0	**Palmitic acid TMS**	117	39
CA91022E-P9-K2	2	1801.0	**Oleic Acid TMS**	132	6
CA91022F-P1-U1	3	565.6	Unknown	110	7
CA91022F-92-U2	3	566.6	Unknown	258	7
CA91022F-P3-U3	3	592.9	Unknown	117	26
CA91022F-P4-U4	3	618.3	Unknown	314	9
CA91022F-P5-U5	3	682.6	Unknown	257	5
CA91022F-P6-U6	3	705.9	Unknown	196	5
CA91022F-P7-U7	3	857.1	Unknown	189	6
CA91022F-P8-K1	3	859.0	**Urea TMS**	189	63
CA91022F-P9-U8	3	860.6	Unknown	179	5
CA91022F-P10-U9	3	1116.2	Unknown	147	16
CA91022F-P11-U10	3	1121.8	Unknown	218	5
CA91022F-P12-U11	3	1150.8	Unknown	103	13
CA91022F-P13-K2	3	1200.4	**Lauric acid TMS**	117	11
CA910-P14-U12	3	1202.9	Unknown	249	6
CA91022F-P15-K3	3	1403.0	Unknown	285	10
CA91022F-P16-U13	3	1499.0	Unknown	117	6
CA91022F- P17-K4	3	1533.3	**Methyl palmitate**	74	54
CA91022F-P18-U14	3	1666.5	Unknown	327	15
CA91022F-P19-K5	3	1743.7	**Methyl stearate**	74	43
CA91022F-P20-U15	3	1763.3	Unknown	124	32
CA91022F-P21-U16	3	1781.1	Unknown	149	5
CA91022F-P22-K6	3	1803.0	**Stearic Acid TMS**	341	32
CA91022F-P23-K7	3	1818.8	**Linoleic Acid TMS**	132	8
CA91022F-P24-U17	3	1970.3	Unknown	79	27
CA91022F-P25-U18	3	2507.0	Unknown	69	7
CA91022F-P26-U19	3	2805.7	Unknown	353	26

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound [a]	m/z	S/N (m/z)
CA91022G-P1-T1	4	577.1	*Lactic acid TMS 2*	129	11
CA91022G-P2-U1	4	598.1	Unknown	117	15
CA91022G-P3-U2	4	622.1	Unknown	193	9
CA91022G-P4-U3	4	759.0	Unknown	205	99
CA91022G-P5-K1	4	768.9	**Glycerol TMS_3**	205	302
CA91022G-P6-U4	4	872.0	Unknown	148	54
CA91022G-P7-K2	4	900.3	**Methyl Caprate**	74	35
CA91022G-P8-K3	4	1000.2	**Capric acid TMS**	129	30
CA91022G-P9-T2	4	1120.9	*Adenine TMS 2*	263	6
CA91022G-P10-U5	4	1121.9	Unknown	127	5
CA91022G-P11-U6	4	1191.6	Unknown	191	13
CA91022G-P12-U7	4	1202.2	Unknown	247	7
CA91022G-P13-K4	4	1394.0	**Fructose MeOX1 - TMS_5**	217	68
CA91022G-P14-K5	4	1395.2	**Shikimic acid TMS_4**	133	57
CA91022G-P15-U8	4	1403.0	Unknown	318	8
CA91022G-P16-K6	4	1415.6	**Galactose MeOX2 - TMS_5**	319	5
CA91022G-P17-U9	4	1421.1	Unknown	205	19
CA91022G-P18-U10	4	1432.8	Unknown	319	5
CA91022G-P19-U11	4	1532.2	Unknown	74	64
CA91022G-P20-K7	4	1611.5	**Palmitelaidic Acid TMS**	311	38
CA91022G-P21-K8	4	1705.3	**Heptadecanoic acid TMS**	117	15
CA91022G-P22-U12	4	1744.0	Unknown	74	24
CA91022G-P23-U13	4	1763.5	Unknown	67	16
CA91022G-P24-U14	4	1771.4	Unknown	376	60
CA91022G-P25-U15	4	1971.2	Unknown	79	20
CA91022G-P26-U16	4	2219.9	Unknown	79	7
CA91022G-P27-U17	4	2508.5	Unknown	123	6
CA91022G-P28-K9	4	2810.0	**Cholesterol TMS**	368	127
CA91022G-P29-T3	4	2835.2	*Alpha-tocopherol TMS*	502	7
CA91022G-P30-T4	4	3080.2	*Octacosanoic acid TMS*	74	53
CA91022H-P1-U1	5	585.1	Unknown	191	13
CA91022H-P2-U2	5	603.9	Unknown	281	6
CA91022H-P3-U3	5	618.1	Unknown	177	7
CA91022H-P4-K1	5	690.1	**Norvaline TMS**	218	9
CA91022H-P5-K2	5	733.2	**Ethanolamine TMS_3**	174	30
CA91022H-P6-K3	5	774.5	**Norleucine TMS**	158	12
CA91022H-P7-U4	5	860.9	Unknown	147	41
CA91022H-P8-U5	5	868.5	Unknown	189	24
CA91022H-P9-U6	5	873.6	Unknown	204	49
CA91022H-P10-U7	5	878.2	Unknown	117	6
CA91022H-P11-U8	5	904.1	Unknown	117	16
CA91022H-P12-U9	5	1121.5	Unknown	263	6
CA91022H-P13-K4	5	1176.5	**Lyxose MeOX2 - TMS_4**	233	18

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound [a]	m/z	S/N (m/z)
CA91022H-P14-K5	5	1196.8	**Arabinose MeOX1 - TMS_4**	103	25
CA91022H-P15-K6	5	1200.0	**Ribose MeOX2 - TMS**	103	10
CA91022H-P16-K7	5	1205.2	**Xylitol TMS_5**	103	18
CA91022H-P17-K8	5	1206.8	**Lyxose MeOX1 - TMS_4**	103	20
CA91022H-P18-K9	5	1208.6	**Ribose MeOX1 - TMS**	103	15
CA91022H-P19-K10	5	1211.5	**Ribose MEOX2-TMS_4**	217	13
CA91022H-P20-K11	5	1245.9	**Fucose MeOX1**	117	15
CA91022H-P21-U10	5	1393.8	Unknown	319	18
CA91022H-P22-K12	5	1400.3	**Galactose MeOX1 - TMS_5**	160	23
CA91022H-P23-K13	5	1402.0	**Sorbitol TMS_6**	319	12
CA91022H-P24-K14	5	1402.5	**Myristic acid TMS**	117	7
CA91022H-P25-K15	5	1404.0	**Fructose MeOX2 - TMS_5**	307	9
CA91022H-P26-U11	5	1407.6	Unknown	244	8
CA91022H-P27-K16	5	1412.9	**Glucose MeOX1 - TMS_5**	205	280
CA91022H-P28-K17	5	1415.9	**Lysine TMS_4**	220	33
CA91022H-P29-U12	5	1419.5	Unknown	320	123
CA91022H-P30-U13	5	1420.5	Unknown	217	246
CA91022H-P31-K18	5	1421.0	**Glucose MeOX2 - TMS_5**	205	147
CA91022H-P32-U14	5	1421.5	Unknown	103	185
CA91022H-P33-U15	5	1442.6	Unknown	217	5
CA91022H-P34-U16	5	1494.6	Unknown	204	22
CA91022H-P35-U17	5	1495.0	Unknown	217	24
CA91022H-P36-U18	5	1496.3	Unknown	74	13
CA91022H-P37-U19	5	1499.0	Unknown	75	6
CA91022H-P38-U20	5	1504.3	Unknown	363	11
CA91022H-P39-K19	5	1532.6	**Sedoheptulose MeOX-TMS_6**	217	26
CA91022H-P40-U21	5	1537.8	Unknown	204	21
CA91022H-P41-U22	5	1574.3	Unknown	204	21
CA91022H-P42-U23	5	1575.5	Unknown	217	23
CA91022H-P43-U24	5	1576.8	Unknown	240	15
CA91022H-P44-U25	5	2705.4	Unknown	117	15
CA91022I-P1-U1	6	543.0	Unknown	216	23
CA91022I-P2-U2	6	565.6	Unknown	119	11
CA91022I-P3-U3	6	590.5	Unknown	58	5
CA91022I-P4-U4	6	593.9	Unknown	190	18
CA91022I-P5-U5	6	603.0	Unknown	370	7
CA91022I-P6-U6	6	603.9	Unknown	281	8
CA91022I-P7-U7	6	619.2	Unknown	274	5
CA91022I-P8-U8	6	624.3	Unknown	248	7
CA91022I-P9-K1	6	646.1	**Pyruvic acid MeOX-TMS**	174	8
CA91022I-P10-K2	6	719.6	**Valine TMS_2**	144	22
CA91022I-P11-U9	6	756.1	Unknown	217	6
CA91022I-P12-U10	6	759.3	Unknown	228	5

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound [a]	m/z	S/N (m/z)
CA91022I-P13-U11	6	768.1	Unknown	218	5
CA91022I-P14-U12	6	784.4	Unknown	127	7
CA91022I-P15-K3	6	824.8	**Serine TMS_2**	132	25
CA91022I-P16-U13	6	856.6	Unknown	299	11
CA91022I-P17-K4	6	861.9	**Benzoic Acid TMS**	179	58
CA91022I-P18-K5	6	904.0	**Succinic Acid TMS_2**	247	19
CA91022I-P19-U14	6	1214.6	Unknown	75	13
CA91022I-P20-U15	6	1419.6	Unknown	103	9
CA91022I-P21-U16	6	1420.3	Unknown	319	26
CA91022I-P22-U17	6	1420.9	Unknown	160	32
CA91022I-P23-U18	6	1421.9	Unknown	129	12
CA91022I-P24-U19	6	1431.9	Unknown	205	8
CA91022I-P25-U20	6	1433.0	Unknown	205	9
CA91022I-P26-U21	6	1555.6	Unknown	205	12
CA91022I-P27-U22	6	1556.2	Unknown	217	10
CA91022I-P28-U23	6	1606.0	Unknown	313	20
CA91022I-P29-U24	6	1740.1	Unknown	147	52
CA91022I-P30-U35	6	3236.9	Unknown	147	45
CA91022J-P1-U1	7	566.1	Unknown	305	11
CA91022J-P2-U2	7	591.1	Unknown	249	21
CA91022J-P3-U3	7	784.6	Unknown	190	6
CA91022J-P4-T1	7	802.1	*Citraconic acid TMS*	215	12
CA91022J-P5-T2	7	802.3	*Itaconic acid TMS*	215	13
CA91022J-P6-U4	7	856.6	Unknown	189	7
CA91022J-P7-U5	7	857.6	Unknown	187	5
CA91022J-P8-U6	7	861.6	Unknown	205	12
CA91022J-P9-U7	7	1004.8	Unknown	217	8
CA91022J-P10-U8	7	1403.1	Unknown	247	6
CA91022J-P11-U9	7	1423.2	Unknown	160	8
CA91022J-P12-U10	7	1483.6	Unknown	204	15
CA91022J-P13-U11	7	1499.2	Unknown	319	5
CA91022J-P14-U12	7	1605.4	Unknown	74	27
CA91022J-P15-U13	7	1804.6	Unknown	341	26
CA91022J-P16-U14	7	2199.2	Unknown	262	7
CA91022K-P1-U1	8	561.4	Unknown	102	7
CA91022K-P2-U2	8	566.2	Unknown	373	6
CA91022K-P3-U3	8	592.3	Unknown	190	23
CA91022K-P4-U4	8	623.3	Unknown	199	5
CA91022K-P5-U5	8	628.6	Unknown	174	5
CA91022K-P6-U6	8	682.0	Unknown	218	6
CA91022K-P7-U7	8	682.3	Unknown	273	5
CA91022K-P8-U8	8	856.3	Unknown	133	9
CA91022K-P9-K1	8	887.0	**Serine TMS_3**	218	24

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound [a]	m/z	S/N (m/z)
CA91022K-P10-U9	8	902.9	Unknown	173	17
CA91022K-P11-U10	8	972.4	Unknown	172	8
CA91022K-P12-U11	8	1121.5	Unknown	196	6
CA91022K-P13-K2	8	1456.4	**Tyramine TMS_3**	338	16
CA91022K-P14-U12	8	1537.6	Unknown	147	28
CA91022L-P1-U1	9	542.9	Unknown	297	20
CA91022L-P2-U2	9	550.8	Unknown	130	6
CA91022L-P3-U3	9	564.8	Unknown	156	10
CA91022L-P4-U4	9	591.3	Unknown	133	24
CA91022L-P5-U5	9	617.0	Unknown	205	6
CA91022L-P6-U6	9	622.0	Unknown	134	7
CA91022L-P7-U7	9	681.8	Unknown	205	5
CA91022L-P8-U8	9	754.1	Unknown	103	6
CA91022L-P9-U9	9	785.0	Unknown	147	6
CA91022L-P10-U10	9	786.6	Unknown	350	7
CA91022L-P11-T1	9	850.1	*Threonine TMS_2*	219	7
CA91022L-P12-U11	9	856.3	Unknown	229	6
CA91022L-P13-K1	9	905.8	**Threonine TMS_3**	218	35
CA91022L-P14-U12	9	961.5	Unknown	133	20
CA91022L-P15-U13	9	1117.9	Unknown	204	17
CA91022L-P16-U14	9	1121.8	Unknown	274	5
CA91022L-P17-U15	9	1200.1	Unknown	156	7
CA91022L-P18-U16	9	1203.0	Unknown	247	6
CA91022L-P19-U17	9	1402.6	Unknown	205	7
CA91022L-P20-U18	9	1403.3	Unknown	295	7
CA91022L-P21-U19	9	1473.6	Unknown	103	39
CA91022L-P22-K2	9	1500.3	**Pentadecanoic acid TMS**	323	5
CA91023D-P1-U1	10	677.8	Unknown	132	9
CA91023D-P2-K1	10	779.9	**Leucine TMS_2**	158	40
CA91023D-P3-U2	10	784.7	Unknown	207	10
CA91023D-P4-U3	10	903.1	Unknown	218	27
CA91023D-P5-U4	10	952.3	Unknown	217	25
CA91023D-P6-U5	10	1101.5	Unknown	75	9
CA91023D-P7-U6	10	1202.8	Unknown	173	11
CA91023D-P8-K2	10	1237.6	**Phenylalanine TMS_2**	218	25
CA91023D-P9-U7	10	1269.4	Unknown	83	9
CA91023D-P10-U8	10	1374.8	Unknown	312	19
CA91023D-P11-U9	10	1413.8	Unknown	117	6
CA91023D-P12-U10	10	1499.0	Unknown	117	11
CA91023D-P13-U11	10	1781.9	Unknown	149	6
CA91023D-P14-U12	10	2606.1	Unknown	83	30
CA91023E-P1-U1	11	587.5	Unknown	133	12
CA91023E-P2-U2	11	859.9	Unknown	73	6

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound [a]	m/z	S/N (m/z)
CA91023E-P3-U3	11	867.9	Unknown	73	5
CA91023E-P4-U4	11	942.5	Unknown	327	16
CA91023E-P5-K1	11	1197.8	**Glutamine TMS_4**	156	38
CA91023E-P6-U5	11	1203.6	Unknown	249	9
CA91023E-P7-U6	11	1270.1	Unknown	163	8
CA91023E-P8-U7	11	1337.6	Unknown	74	13
CA91023E-P9-U8	11	2508.0	Unknown	69	8
CA91023F-P1U1	12	544.1	Unknown	218	28
CA91023F-P2-U2	12	563.8	Unknown	103	17
CA91023F-P3-U3	12	588.7	Unknown	219	38
CA91023F-P4-U4	12	630.4	Unknown	89	45
CA91023F-P5-U5	12	777.2	Unknown	159	65
CA91023F-P6U6	12	785.6	Unknown	116	13
CA91023F-P7-K1	12	805.0	**Isoleucine TMS_2**	158	35
CA91023F-P8-U7	12	856.7	Unknown	147	7
CA91023F-P9-U8	12	871.1	Unknown	191	6
CA91023F-P10-U9	12	1101.7	Unknown	103	12
CA91023F-P11-U10	12	1203.2	Unknown	175	7
CA91023F-P12-U11	12	1270.5	Unknown	257	7
CA91023F-P13-U12	12	1298.6	Unknown	117	28
CA91023F-P14-U13	12	1499.5	Unknown	75	13
CA91023F-P15-U14	12	2104.2	Unknown	244	5
CA91023G-P1-U1	13	546.6	Unknown	184	12
CA91023G-P2-U2	13	563.4	Unknown	184	20
CA91023G-P3-U3	13	569.5	Unknown	73	8
CA91023G-P4-U4	13	588.3	Unknown	190	18
CA91023G-P5-K1	13	639.4	**Alanine TMS_2**	116	23
CA91023G-P6-U5	13	688.1	Unknown	219	9
CA91023G-P7-U6	13	811.6	Unknown	127	6
CA91023G-P8-U7	13	902.5	Unknown	172	24
CA91023G-P9-U8	13	1099.6	Unknown	73	7
CA91023G-P10-U9	13	1201.6	Unknown	205	7
CA91023G-P11-U10	13	1249.5	Unknown	221	14
CA91023G-P12-U11	13	1268.9	Unknown	114	7
CA91023G-P13-U12	13	1394.4	Unknown	103	11
CA91023G-P14-U13	13	1408.9	Unknown	73	7
CA91023G-P15-U14	13	1497.6	Unknown	75	8
CA91023G-P16-U15	13	2508.3	Unknown	81	14
CA91023G-P17-U16	13	2605.3	Unknown	72	12
CA91023H-P1-U1	14	542.6	Unknown	230	34
CA91023H-P2-U2	14	600.5	Unknown	281	8
CA91023H-P3-U3	14	785.2	Unknown	281	7
CA91023H-P4-K1	14	810.3	**Glycine TMS_3**	174	20

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound [a]	m/z	S/N (m/z)
CA91023H-P5-U4	14	1202.3	Unknown	204	19
CA91023H-P6-U5	14	1243.6	Unknown	117	14
CA91023I-P1-U1	15	564.8	Unknown	102	22
CA91023I-P2-U2	15	778.1	Unknown	75	8
CA91023I-P3-U3	15	787.3	Unknown	219	6
CA91023I-P4-U4	15	811.7	Unknown	248	7
CA91023I-P5-K1	15	838.2	**Proline TMS_2**	142	26
CA91023I-P6-K2	15	1231.4	**Homocysteine TMS_3**	219	32
CA91023I-P7-U5	15	1272.2	Unknown	267	6
CA91023I-P8-U6	15	1414.1	Unknown	161	5
CA91023J-P1-U1	16	542.6	Unknown	205	37
CA91023J-P2-U2	16	564.4	Unknown	119	20
CA91023J-P3-U3	16	586.5	Unknown	191	16
CA91023J-P4-U4	16	785.5	Unknown	244	7
CA91023J-P5-U5	16	877.5	Unknown	97	5
CA91023J-P6-U6	16	903.2	Unknown	210	14
CA91023J-P7-U7	16	1221.3	Unknown	102	18
CA91023J-P8-U8	16	1413.9	Unknown	73	5
a: Known compounds are in bold; tentative compounds are in italics and the rest are unknowns					

Appendix 4: List of unknown peaks in GC-MS of LC fractions

Total number of unknown peaks: 214 (inclusive of 48 uknown peaks from GC-MS only)					
No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound	m/z	S/N (m/z)
CA91022D-P1-U1	1	590.1	Unknown	117	22
CA91022D-P4-U2	1	784.5	Unknown	357	8
CA91022D-P5-U3	1	856.5	Unknown	232	5
CA91022D-P6-U4	1	903.2	Unknown	129	8
CA91022D-P7-U5	1	1055.3	Unknown	207	9
CA91022D-P8-U6	1	1539.4	Unknown	319	26
CA91022E-P1-U1	2	591.6	Unknown	117	20
CA91022E-P2-U2	2	720.3	Unknown	174	5
CA91022E-P3-U3	2	856.5	Unknown	147	8
CA91022E-P4-U4	2	1136.4	Unknown	117	12
CA91022E-P5-U5	2	1403.3	Unknown	133	8
CA91022E-P6-U6	2	1499.8	Unknown	299	6
CA91022E-P7-U7	2	1589.3	Unknown	205	20
CA91022F-P1-U1	3	565.6	Unknown	110	7
CA91022F-92-U2	3	566.6	Unknown	258	7
CA91022F-P3-U3	3	592.9	Unknown	117	26
CA91022F-P4-U4	3	618.3	Unknown	314	9
CA91022F-P5-U5	3	682.6	Unknown	257	5
CA91022F-P6-U6	3	705.9	Unknown	196	5
CA91022F-P7-U7	3	857.1	Unknown	189	6
CA91022F-P9-U8	3	860.6	Unknown	179	5
CA91022F-P10-U9	3	1116.2	Unknown	147	16
CA91022F-P11-U10	3	1121.8	Unknown	218	5
CA91022F-P12-U11	3	1150.8	Unknown	103	13
CA910-P14-U12	3	1202.9	Unknown	249	6
CA91022F-P15-K3	3	1403.0	Unknown	285	10
CA91022F-P16-U13	3	1499.0	Unknown	117	6
CA91022F-P18-U14	3	1666.5	Unknown	327	15
CA91022F-P20-U15	3	1763.3	Unknown	124	32
CA91022F-P21-U16	3	1781.1	Unknown	149	5
CA91022F-P24-U17	3	1970.3	Unknown	79	27
CA91022F-P25-U18	3	2507.0	Unknown	69	7
CA91022F-P26-U19	3	2805.7	Unknown	353	26
CA91022G-P2-U1	4	598.1	Unknown	117	15
CA91022G-P3-U2	4	622.1	Unknown	193	9
CA91022G-P4-U3	4	759.0	Unknown	205	99
CA91022G-P6-U4	4	872.0	Unknown	148	54
CA91022G-P10-U5	4	1121.9	Unknown	127	5
CA91022G-P11-U6	4	1191.6	Unknown	191	13
CA91022G-P12-U7	4	1202.2	Unknown	247	7
CA91022G-P15-U8	4	1403.0	Unknown	318	8
CA91022G-P17-U9	4	1421.1	Unknown	205	19
CA91022G-P18-U10	4	1432.8	Unknown	319	5
CA91022G-P19-U11	4	1532.2	Unknown	74	64
CA91022G-P22-U12	4	1744.0	Unknown	74	24
CA91022G-P23-U13	4	1763.5	Unknown	67	16
CA91022G-P24-U14	4	1771.4	Unknown	376	60
CA91022G-P25-U15	4	1971.2	Unknown	79	20
CA91022G-P26-U16	4	2219.9	Unknown	79	7
CA91022G-P27-U17	4	2508.5	Unknown	123	6

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound	m/z	S/N (m/z)
CA91022H-P1-U1	5	585.1	Unknown	191	13
CA91022H-P2-U2	5	603.9	Unknown	281	6
CA91022H-P3-U3	5	618.1	Unknown	177	7
CA91022H-P7-U4	5	860.9	Unknown	147	41
CA91022H-P8-U5	5	868.5	Unknown	189	24
CA91022H-P9-U6	5	873.6	Unknown	204	49
CA91022H-P10-U7	5	878.2	Unknown	117	6
CA91022H-P11-U8	5	904.1	Unknown	117	16
CA91022H-P12-U9	5	1121.5	Unknown	263	6
CA91022H-P21-U10	5	1393.8	Unknown	319	18
CA91022H-P26-U11	5	1407.6	Unknown	244	8
CA91022H-P29-U12	5	1419.5	Unknown	320	123
CA91022H-P30-U13	5	1420.5	Unknown	217	246
CA91022H-P32-U14	5	1421.5	Unknown	103	185
CA91022H-P33-U15	5	1442.6	Unknown	217	5
CA91022H-P34-U16	5	1494.6	Unknown	204	22
CA91022H-P35-U17	5	1495.0	Unknown	217	24
CA91022H-P36-U18	5	1496.3	Unknown	74	13
CA91022H-P37-U19	5	1499.0	Unknown	75	6
CA91022H-P38-U20	5	1504.3	Unknown	363	11
CA91022H-P40-U21	5	1537.8	Unknown	204	21
CA91022H-P41-U22	5	1574.3	Unknown	204	21
CA91022H-P42-U23	5	1575.5	Unknown	217	23
CA91022H-P43-U24	5	1576.8	Unknown	240	15
CA91022H-P44-U25	5	2705.4	Unknown	117	15
CA91022I-P1-U1	6	543.0	Unknown	216	23
CA91022I-P2-U2	6	565.6	Unknown	119	11
CA91022I-P3-U3	6	590.5	Unknown	58	5
CA91022I-P4-U4	6	593.9	Unknown	190	18
CA91022I-P5-U5	6	603.0	Unknown	370	7
CA91022I-P6-U6	6	603.9	Unknown	281	8
CA91022I-P7-U7	6	619.2	Unknown	274	5
CA91022I-P8-U8	6	624.3	Unknown	248	7
CA91022I-P11-U9	6	756.1	Unknown	217	6
CA91022I-P12-U10	6	759.3	Unknown	228	5
CA91022I-P13-U11	6	768.1	Unknown	218	5
CA91022I-P14-U12	6	784.4	Unknown	127	7
CA91022I-P16-U13	6	856.6	Unknown	299	11
CA91022I-P19-U14	6	1214.6	Unknown	75	13
CA91022I-P20-U15	6	1419.6	Unknown	103	9
CA91022I-P21-U16	6	1420.3	Unknown	319	26
CA91022I-P22-U17	6	1420.9	Unknown	160	32
CA91022I-P23-U18	6	1421.9	Unknown	129	12
CA91022I-P24-U19	6	1431.9	Unknown	205	8
CA91022I-P25-U20	6	1433.0	Unknown	205	9
CA91022I-P26-U21	6	1555.6	Unknown	205	12
CA91022I-P27-U22	6	1556.2	Unknown	217	10
CA91022I-P28-U23	6	1606.0	Unknown	313	20
CA91022I-P29-U24	6	1740.1	Unknown	147	52
CA91022I-P30-U35	6	3236.9	Unknown	147	45

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound	m/z	S/N (m/z)
CA91022J-P1-U1	7	566.1	Unknown	305	11
CA91022J-P2-U2	7	591.1	Unknown	249	21
CA91022J-P3-U3	7	784.6	Unknown	190	6
CA91022J-P6-U4	7	856.6	Unknown	189	7
CA91022J-P7-U5	7	857.6	Unknown	187	5
CA91022J-P8-U6	7	861.6	Unknown	205	12
CA91022J-P9-U7	7	1004.8	Unknown	217	8
CA91022J-P10-U8	7	1403.1	Unknown	247	6
CA91022J-P11-U9	7	1423.2	Unknown	160	8
CA91022J-P12-U10	7	1483.6	Unknown	204	15
CA91022J-P13-U11	7	1499.2	Unknown	319	5
CA91022J-P14-U12	7	1605.4	Unknown	74	27
CA91022J-P15-U13	7	1804.6	Unknown	341	26
CA91022J-P16-U14	7	2199.2	Unknown	262	7
CA91022K-P1-U1	8	561.4	Unknown	102	7
CA91022K-P2-U2	8	566.2	Unknown	373	6
CA91022K-P3-U3	8	592.3	Unknown	190	23
CA91022K-P4-U4	8	623.3	Unknown	199	5
CA91022K-P5-U5	8	628.6	Unknown	174	5
CA91022K-P6-U6	8	682.0	Unknown	218	6
CA91022K-P7-U7	8	682.3	Unknown	273	5
CA91022K-P8-U8	8	856.3	Unknown	133	9
CA91022K-P10-U9	8	902.9	Unknown	173	17
CA91022K-P11-U10	8	972.4	Unknown	172	8
CA91022K-P12-U11	8	1121.5	Unknown	196	6
CA91022K-P14-U12	8	1537.6	Unknown	147	28
CA91022L-P1-U1	9	542.9	Unknown	297	20
CA91022L-P2-U2	9	550.8	Unknown	130	6
CA91022L-P3-U3	9	564.8	Unknown	156	10
CA91022L-P4-U4	9	591.3	Unknown	133	24
CA91022L-P5-U5	9	617.0	Unknown	205	6
CA91022L-P6-U6	9	622.0	Unknown	134	7
CA91022L-P7-U7	9	681.8	Unknown	205	5
CA91022L-P8-U8	9	754.1	Unknown	103	6
CA91022L-P9-U9	9	785.0	Unknown	147	6
CA91022L-P10-U10	9	786.6	Unknown	350	7
CA91022L-P12-U11	9	856.3	Unknown	229	6
CA91022L-P14-U12	9	961.5	Unknown	133	20
CA91022L-P15-U13	9	1117.9	Unknown	204	17
CA91022L-P16-U14	9	1121.8	Unknown	274	5
CA91022L-P17-U15	9	1200.1	Unknown	156	7
CA91022L-P18-U16	9	1203.0	Unknown	247	6
CA91022L-P19-U17	9	1402.6	Unknown	205	7
CA91022L-P20-U18	9	1403.3	Unknown	295	7
CA91022L-P21-U19	9	1473.6	Unknown	103	39
CA91023D-P1-U1	10	677.8	Unknown	132	9
CA91023D-P3-U2	10	784.7	Unknown	207	10
CA91023D-P4-U3	10	903.1	Unknown	218	27
CA91023D-P5-U4	10	952.3	Unknown	217	25
CA91023D-P6-U5	10	1101.5	Unknown	75	9

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound	m/z	S/N (m/z)
CA91023D-P7-U6	10	1202.8	Unknown	173	11
CA91023D-P9-U7	10	1269.4	Unknown	83	9
CA91023D-P10-U8	10	1374.8	Unknown	312	19
CA91023D-P11-U9	10	1413.8	Unknown	117	6
CA91023D-P12-U10	10	1499.0	Unknown	117	11
CA91023D-P13-U11	10	1781.9	Unknown	149	6
CA91023D-P14-U12	10	2606.1	Unknown	83	30
CA91023E-P1-U1	11	587.5	Unknown	133	12
CA91023E-P2-U2	11	859.9	Unknown	73	6
CA91023E-P3-U3	11	867.9	Unknown	73	5
CA91023E-P4-U4	11	942.5	Unknown	327	16
CA91023E-P6-U5	11	1203.6	Unknown	249	9
CA91023E-P7-U6	11	1270.1	Unknown	163	8
CA91023E-P8-U7	11	1337.6	Unknown	74	13
CA91023E-P9-U8	11	2508.0	Unknown	69	8
CA91023F-P1U1	12	544.1	Unknown	218	28
CA91023F-P2-U2	12	563.8	Unknown	103	17
CA91023F-P3-U3	12	588.7	Unknown	219	38
CA91023F-P4-U4	12	630.4	Unknown	89	45
CA91023F-P5-U5	12	777.2	Unknown	159	65
CA91023F-P6U6	12	785.6	Unknown	116	13
CA91023F-P8-U7	12	856.7	Unknown	147	7
CA91023F-P9-U8	12	871.1	Unknown	191	6
CA91023F-P10-U9	12	1101.7	Unknown	103	12
CA91023F-P11-U10	12	1203.2	Unknown	175	7
CA91023F-P12-U11	12	1270.5	Unknown	257	7
CA91023F-P13-U12	12	1298.6	Unknown	117	28
CA91023F-P14-U13	12	1499.5	Unknown	75	13
CA91023F-P15-U14	12	2104.2	Unknown	244	5
CA91023G-P1-U1	13	546.6	Unknown	184	12
CA91023G-P2-U2	13	563.4	Unknown	184	20
CA91023G-P3-U3	13	569.5	Unknown	73	8
CA91023G-P4-U4	13	588.3	Unknown	190	18
CA91023G-P6-U5	13	688.1	Unknown	219	9
CA91023G-P7-U6	13	811.6	Unknown	127	6
CA91023G-P8-U7	13	902.5	Unknown	172	24
CA91023G-P9-U8	13	1099.6	Unknown	73	7
CA91023G-P10-U9	13	1201.6	Unknown	205	7
CA91023G-P11-U10	13	1249.5	Unknown	221	14
CA91023G-P12-U11	13	1268.9	Unknown	114	7
CA91023G-P13-U12	13	1394.4	Unknown	103	11
CA91023G-P14-U13	13	1408.9	Unknown	73	7
CA91023G-P15-U14	13	1497.6	Unknown	75	8
CA91023G-P16-U15	13	2508.3	Unknown	81	14
CA91023G-P17-U16	13	2605.3	Unknown	72	12
CA91023H-P1-U1	14	542.6	Unknown	230	34
CA91023H-P2-U2	14	600.5	Unknown	281	8
CA91023H-P3-U3	14	785.2	Unknown	281	7
CA91023H-P5-U4	14	1202.3	Unknown	204	19
CA91023H-P6-U5	14	1243.6	Unknown	117	14

No.	Fraction #	RI(DB-17ht fatty acid system)	Identification of derivatised compound	m/z	S/N (m/z)
CA91023I-P1-U1	15	564.8	Unknown	102	22
CA91023I-P2-U2	15	778.1	Unknown	75	8
CA91023I-P3-U3	15	787.3	Unknown	219	6
CA91023I-P4-U4	15	811.7	Unknown	248	7
CA91023I-P7-U5	15	1272.2	Unknown	267	6
CA91023I-P8-U6	15	1414.1	Unknown	161	5
CA91023J-P1-U1	16	542.6	Unknown	205	37
CA91023J-P2-U2	16	564.4	Unknown	119	20
CA91023J-P3-U3	16	586.5	Unknown	191	16
CA91023J-P4-U4	16	785.5	Unknown	244	7
CA91023J-P5-U5	16	877.5	Unknown	97	5
CA91023J-P6-U6	16	903.2	Unknown	210	14
CA91023J-P7-U7	16	1221.3	Unknown	102	18
CA91023J-P8-U8	16	1413.9	Unknown	73	5