# ANALYTICAL METHODS BASED ON ION MOBILITY AND MASS SPECTROMETRY FOR METABOLOMICS

Aditya Malkar

Centre for Analytical Science
Department of Chemistry
Loughborough University
Loughborough, UK
LE11 3TU

A Doctoral thesis submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy to the Loughborough University

*Dedicated to my mom..*

*“The journey itself is the reward..”*

*- Steve Jobs*

# ACKNOWLEDGEMENTS

I would like to thank dearly my supervisor Professor Colin S. Creaser for his continued support, guidance, patience and encouragement through the entirety of the PhD. You have been a source of inspiration.

I would like to thank my mom and dad for the constant emotional and financial support. Mom, words won't do justice to your place in my life. Thank you for always being there for me, from across the world, as a guide and a friend.

I am grateful to Professor C.L. Paul Thomas and Dr James C. Reynolds for their words of wisdom over the years.

I extend my gratitude to Dominic Shaw, Emma Wilson and the team at the Respiratory BRU at Nottingham City Hospital for the collection of samples and clinical opinions for the Asthma project. I would like to thank Don JL Jones and Stewart Sale at Department of Cancer Studies and Molecular Medicine, University of Leicester for the collaboration on the colorectal cancer study. I would also like to acknowledge the help and input from Phillip Watson and colleagues at SSEHS and Centre for Analytical Science at the Loughborough University for the metabolomic investigations on physiological stress.

It's a privilege to call my colleagues as my friends and family. Neil, Caitlyn, Helen, Kim, Cristina, Lauren, Matthew, Rob, Sarah, Dorota, Pareen, Vicky; you all have made these years unforgettable. Tulika, you have been there for me through it all; the distance and different time-zones. Thank you.

And I thank wholeheartedly, my grandparents. You two have been with me throughout this journey, and you will always be in my heart..

# ABSTRACT

Travelling wave ion mobility spectrometry (TWIMS) in combination with ultra-high performance liquid chromatography (UHPLC) and mass spectrometry (MS) has been applied successfully for the untargeted, global metabolic profiling of biofluids such as mouse plasma and saliva. Methods based on UHPLC-MS alone and in combination with ion mobility spectrometry (UHPLC-IM-MS) have been developed and validated for the untargeted metabolite profiling of saliva, obtained non-invasively by passive drool. Three separate metabolic profiling studies have been carried out in conjunction with bioinformatics strategies to identify potential metabolomic biomarker ions that are associated with efficacy of rice bran in colorectal cancer, physiological stress and that have the potential for the diagnosis of asthma.

The advantages offered by the utility of ion mobility in UHPLC-MS based metabolic profiling studies, including the increased analytical space, mass spectral clean-up of contaminants such as PEG post-UHPLC-IM-MS analysis, enhancement of the selectivity of targeted metabolites as well as the potential for the identification of metabolites by comparison of ion mobility drift times have been highlighted.

Ten potential metabolic biomarker ions of asthma have been identified from the moderate asthmatics from untargeted metabolite profiling of saliva by UHPLC-MS. A predictive model based on partial least squares – discriminant analysis (PLS-DA) has been constructed using these ten discriminant ions, which demonstrates good predictive capability for moderate asthmatics and controls.

Potential metabolic biomarker ions of physiological stress have been identified through untargeted metabolite profiling analysis of saliva samples collected before and after exercise by UHPLC-IM-MS. Valerolactam has been identified as a potential biomarker of physiological stress from saliva by comparison of retention time, ion mobility drift time and MS/MS spectra with a standard of δ-valerolactam.

# GLOSSARY OF TERMS

| | |
|---|---|
| %RSD | % Relative standard deviation |
| ACN | Acetonitrile |
| ANN | Artificial Neural Network |
| Da | Dalton |
| DC | Direct Current |
| DMS | Differential Mobility Spectrometry |
| EIC | Extracted Ion Chromatogram |
| ESI | Electrospray Ionisation |
| FAIMS | Field Asymmetric Waveform Ion Mobility Spectrometry |
| GC | Gas Chromatography |
| HPLC | High Performance Liquid Chromatography |
| IMS | Ion Mobility Spectrometry |
| *m/z* | Mass-to-charge ratio |
| MRM | Multiple Reaction Monitoring |
| MS | Mass Spectrometry |
| MVA | Multivariate Analysis |
| NMR | Nuclear Magnetic Resonance Spectroscopy |
| PCA | Principal Component Analysis |
| QC | Quality Control |
| Q-ToF | Quadrupole Time-of-Flight |
| RF | Radio Frequency |
| SRIG | Stacked Ring Ion Guide |
| SRM | Selected Reaction Monitoring |
| TIC | Total Ion Current |
| ToF | Time-of-Flight |
| TWIMS | Travelling Wave Ion Mobility Spectrometry |
| UHPLC | Ultra-high Performance Liquid Chromatography |

# LIST OF FIGURES

# CONTENTS

# CHAPTER ONE:

INTRODUCTION

## 1.1 MASS SPECTROMETRY AND ION MOBILITY SPECTROMETRY

The year was 1912 when J. J. Thomson conducted experiments that would mark the beginning of mass spectrometry. Since then the technique has seen numerous improvements culminating into modern day mass spectrometry. The coupling of mass spectrometry with electron and chemical ionisation enabled the analysis of volatile compounds. The discoveries of electrospray ionisation by Dole *et al.* and Fenn *et al.* and matrix-assisted laser desorption ionisation by Karas *et al*. in the 1980s paved the way for analysis of non-volatile high molecular mass compounds such as synthetic polymers, proteins, glycans, polynucleotides etc. by mass spectrometry(Dole 1968; Fenn et al. 1989; Karasek 1974). The coupling of mass spectrometers with separation techniques such as gas chromatography, liquid chromatography (LC), capillary electrophoresis (CE), ion mobility spectrometry (IMS) and field asymmetric waveform ion mobility spectrometry (FAIMS) have opened up new avenues in the world of analytical chemistry allowing mass spectrometry to become an important and a very powerful tool.

Along with developments in mass spectrometry, the past decade has seen the coming of age of ion mobility spectrometry (IMS). IMS was developed initially as a method for detecting and identifying volatile and semi-volatile compounds, principally for security and military applications. The early years of ion mobility spectrometry saw the development of stand-alone IMS devices employed for the detection of volatile explosives and chemical warfare agents. IMS is a gas phase separation technique, so the development of mass spectrometric ion sources such as electrospray (ESI) and matrix assisted laser desorption ionisation (MALDI) amongst others, has equally benefited both ion mobility and mass spectrometry. The use of these sources originally developed for MS, have enabled the analysis of not only volatile species but also condensed phase and solid state analytes by IMS. Recently, IMS has found use in the rapid analysis and characterisation of drugs, pharmaceuticals and small biomolecules etc. (Weston et al. 2005; Borsdorf & Eiceman 2006).

Ion mobility spectrometry when coupled with mass spectrometry becomes a powerful analytical tool for the analysis of complex samples and thus can explore avenues in the fields of proteomics, genomics and metabolomics (Kanu et al. 2008).

### 1.1.1 Principles of mass spectrometry

Mass spectrometry is a technique which measures the mass of a molecule based on the motion of a charged particle i.e. an ion, in an electric or magnetic field. The gaseous ions formed from the molecules or atoms of a sample are separated in space and time and are detected according to their mass-to-charge ratio i.e. *m/z*. The output from a mass spectrometer is a mass spectrum, which constitutes of the number of ions of each detected mass. Typically, a mass spectrometer consists of a sample introduction system, ionisation source, mass analyser and a detector, which are maintained under vacuum. The vacuum allows the minimisation of ion-molecule reactions, scattering, collisions and neutralisation of the ions as these would interfere with the mass spectral analysis.

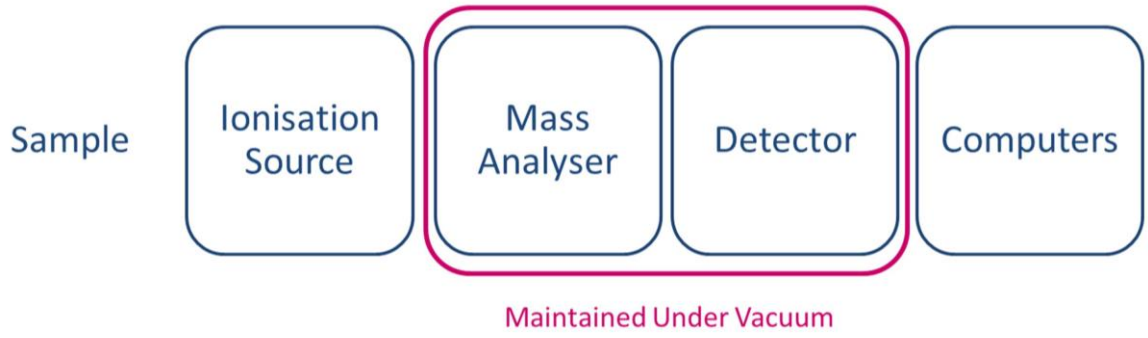



**Figure 1.1 Simplified schematic of components of a mass spectrometer.**

The sample is introduced into the ionisation region where gas phase ions are generated from the analyte. These ions are then accelerated towards the mass analyser by the use of ion optics, and electric and/or magnetic fields, which separate the ions, based on their *m/z* ratio. These separated ions strike the detector and produce an ion current, which is converted into a mass spectrum by the use of computers.

There are several different types of sources and mass analysers available, which function on various principles, but carry out the same function of producing ions and separation of ions respectively. Ionisation sources and mass analysers used in the work described in this thesis are described in the following sections.

#### *1.1.1.1 Ionisation techniques*

The samples introduced into a mass spectrometer are ionised in the ion source prior to analysis in the mass analyser. Ion source is the region where gas phase analyte ions are formed from the sample molecules.

A variety of ionisation techniques have been developed for mass spectrometry including electron ionisation, chemical ionisation, matrix assisted laser desorption ionisation (MALDI), atmospheric pressure chemical ionisation (APCI), thermospray ionisation, electrospray ionisation (ESI) etc. (de Hoffmann & Stroobant 2007). The selection of ionisation technique depends on the sample and the sample matrix; however, the function of all these ionisation techniques is the same, which is facilitating the introduction of sample into a mass analyser in a suitable form. The research presented in this thesis used electrospray ionisation, which is discussed in Section 1.1.1.2.

#### *1.1.1.2 Electrospray ionisation (ESI)*

The concept of electrospray was introduced by Dole *et al.* in 1968 (Dole 1968). However, the true potential of ESI was realised only in the late 1980s when Fenn *et al*.(Fenn et al. 1989) developed ESI as a true interface for mass spectrometry. Fenn demonstrated that multiply charged ions were obtained from large biomolecules, such as proteins, directly from solution enabling their mass spectrometric analysis at low *m/z*. ESI also enabled the hyphenation of liquid chromatography (LC) with mass spectrometry.

The development of electrospray has been significant for modern mass spectrometry as it enables the generation of singly and multiply charged ions from semi-volatile biologically significant species e.g. protein and metabolites.(Gaskell 1997)

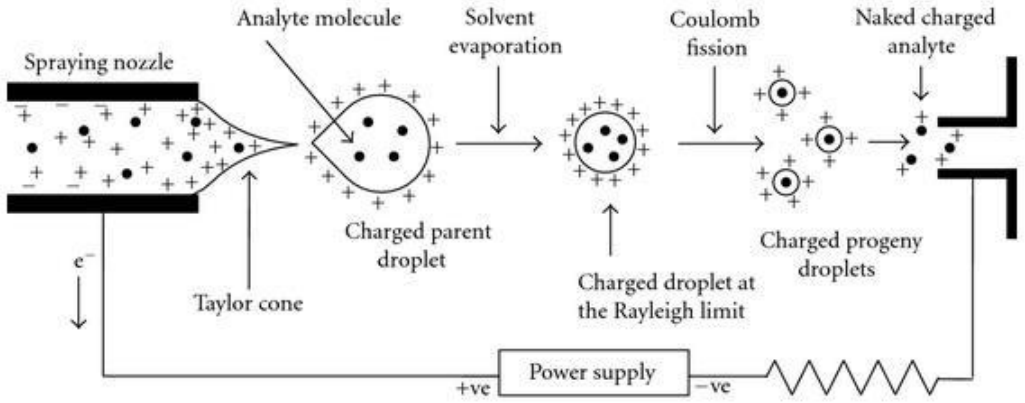



**Figure 1.2 Droplet formation in electrospray ionisation** (Banerjee & Mazumdar 2012)

The ESI interface **Figure 1.2** functions at atmospheric pressure. Ions are produced in ESI by spraying a solution of the analyte through a capillary maintained under a strong electric field. A large potential difference of 3-6 kV is applied between the capillary and a counter electrode over a small distance to produce an electric field of the order of $10^6$ $V.m^{-1}$. The high electric field induces charge separation at the tip of the capillary. In the absence of a potential applied to the liquid, the droplets fall off under the influence of gravity and no electrospray is achieved. However, with the increase in potential, at the onset voltage, which is specific to the surface tension of the solvent being used, the shape of the liquid emanating from the capillary tip starts to deform giving rise to a Taylor cone (Rohner et al. 2004) which produces the electrospray. On increasing the potential further, the accumulation of charge on the surface of the electrospray droplets coupled with evaporation of the solvent which is aided by a coaxial gas flow, causes the droplets to reduce in size until the Raleigh limit is reached (Rohner et al. 2004). At the Raleigh limit, the force of coulombic repulsion between the ions is equal to the surface tension of the solvent. The stability of these charged droplets is therefore governed by the competition between coulomb repulsion and surface tension.(Kebarle & Peschke 2000) As a result of this, the droplets break-up producing a number of smaller droplets. The coaxial gas flow also limits the dispersion of the spray in space resulting in more efficient transfer of ions to the mass spectrometer.

Two main theories have been proposed which describe the probable mechanism of production of gas phase ions from charged droplets – the charge residue model (CRM) and the ion evaporation model (IEM) (Kebarle & Peschke 2000). The charge residue

model (CRM) proposes sufficient dispersion of solvent droplets into smaller droplets until there is only one ion left in the droplet. At this stage it is believed to undergo complete solvent evaporation to produce a gas phase ion.(Dole et al. 1968) The other model, the ion evaporation model (IEM), predicts that before a droplet reached the final stages proposed by the CRM, the increased charge density in the droplet causes the coulombic forces to exceed the surface tension to expel an ion from the droplet into gaseous phase (Iribarne 1976).

The condensed phase electrospray ionisation process has advantages over early gas phase ionisation processes, such as it readily allows hyphenation with high-performance liquid chromatography (HPLC) systems. This allows separation of complex samples by HPLC prior to mass spectral analysis consequently improving the sensitivity of the platform.

#### *1.1.1.3 Mass analysers*

Following ionisation of the analyte molecules, the gaseous phase ions move to a mass analyser for separation and detection. This movement of the ions is based on a series of electrical potentials. The analyser region of the mass spectrometer separates the gas phase ions according to their mass-to-charge ratio (*m/z*). The separation on the basis of *m/z* ratio can be carried out by several different ways and, as a result there exist a variety of mass analysers each with its advantages and limitations. The types of mass analysers commonly used include magnetic sector (B), quadrupole (Q), ion trap (IT), time-of-flight (ToF), Fourier transform-ion cyclotron resonance (FT-ICR) and Orbitrap (OT) (summarised in **Table 1-1**). Two types of mass analysers were used in the work presented in this thesis, a ToF and an Orbitrap and are described in detail in the following sections.

**Table 1-1 Types of mass analysers used in mass spectrometry**

| Mass Analyser | Annotation | Principle of ion separation |
|---|---|---|
| Magnetic sector | B | Momentum |
| Quadrupole | Q | (*m/z*) Mass selective instability |
| Ion trap | IT | (*m/z*) Resonance frequency |
| Time-of-flight | ToF | Flight time of ions |
| Fourier transform ion cyclotron resonance | FTICR | (*m/z*) Resonance frequency |
| Orbitrap | OT | (*m/z*) Resonance frequency |

#### *1.1.1.3.1 Time-of-flight (ToF)*

The Cincinnati Division of the Bendix Aviation Corporation introduced the first time-of-flight (ToF) mass spectrometer in the late 1950s. Since its first introduction, there have been many revisions of the design of ToF mass analyser as described elsewhere (Guilhaus 1995; Cotter 1992; Ioanoviciu 2001). In a time-of-flight instrument, pulses of ions formed in the ion source are accelerated from the ion source by the application of a voltage into a field free region termed the flight tube. In the flight tube the ions are separated based on the differences in the time taken to reach the detector. Since, all the ions are accelerated by the same voltage; each ion possesses, at least in principle, the same kinetic energy. As a result, the velocity of each ion is a direct function of its mass. In other words, heavier ions will have a lower velocity than lighter ions and consequently will take longer (measured in terms of flight time *t*) to traverse the length of the flight tube. The *m/z* ratio of an ion is calculated from its flight time (de Hoffmann & Stroobant 2007).

The kinetic energy of an ion is equal to its potential energy following acceleration from the source, and can be described as:

$$KE = z\mathrm{eV}_{acc} = \frac{mv^2}{2}$$

Where *z* is the charge on the ion, e is the fundamental unit of charge (1.602 x $10^{-19}$ C), $V_{acc}$ is the accelerating voltage, m is the mass of the ion and *v* is the velocity of the ion. Hence, the velocity of the ion in the flight tube can be represented by:

$$v = \sqrt{\frac{2zeV_{acc}}{m}}$$

The velocity of the ion can also be represented based on its flight time (*t*) and the length (*d*) of the flight tube:

$$v = \frac{d}{t}$$

The *m/z* can then be calculated by rearrangement of the equations and substituting for *v* as follows:

$$m/z = \frac{t^2}{\left(\frac{d^2}{2V_{acc}e}\right)}$$

However, most instruments are equipped with reflectrons to effectively increase the length of the flight tube without increasing the length of the instrumentation as shown in **Figure 1.3**. As a result, the calculation of *m/z* becomes difficult. Hence, in practice, the following equation is used to determine *m/z* ratios from measured flight times:

$$m/z = at^2 + b$$

Where, 'a' and 'b' are pre-determined constant values obtained from reference compounds of known molecular weights for given instrumental conditions.

Theoretically, ToF mass analysers have an unlimited mass range due to the principles of operation. However, in reality very large molecules do not hit the detector plate with enough energy to produce a large enough current to be registered as a signal. As a result

ToF mass analysers have a limit to the mass range that can be measured. Nonetheless, ToF mass analysers are capable of functioning to the range in excess of 300kDa and the limits are constantly being pushed by instrumental developments (de Hoffmann & Stroobant 2007).

Historically, the limitation of ToF mass analysers was poor resolution which was due to the variation in the flight times of ions of the same *m/z*. This obstacle was overcome by the use of reflectrons in the ToF design. A reflectron is a series of rings or grids to which an increasing potential is applied (Alikhanov 1957). The reflectron acts as an 'ion-mirror' to reflect and focus ions of the same *m/z*. Ions with higher kinetic energies will arrive at the reflectron before those with lower kinetic energies; however they will also penetrate further into the electrostatic field of the relfectron as compared to the ions with lower kinetic energies. This arrangement allows the ions of the same *m/z*, but different kinetic energies, to arrive at the detector at the same time (Mamyrin et al. 1973). A schematic of a time-of-flight mass analyser with a reflectron is depicted in **Figure 1.3**.

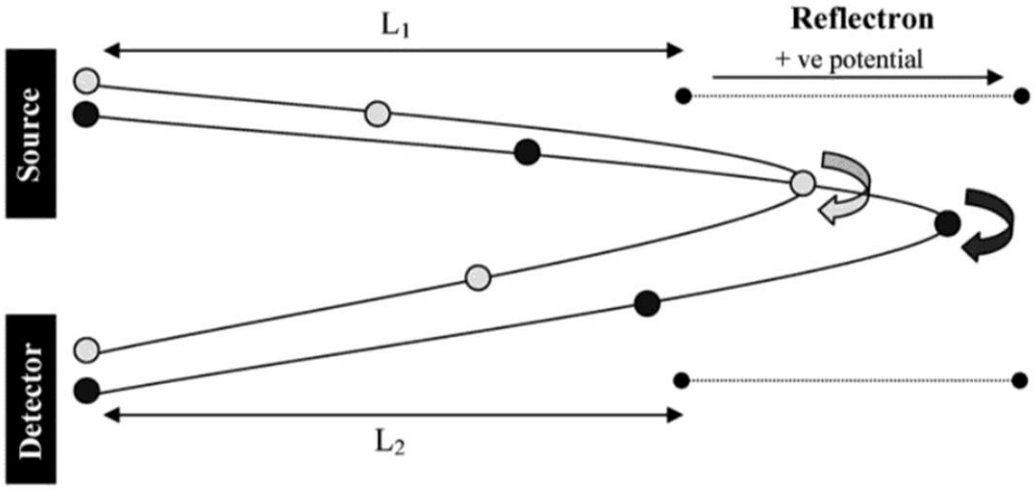



**Figure 1.3 Schematic of a time-of-flight mass analyser with a reflectron** (Rodrigues et al. 2007)**.**

The introduction of a reflectron in a ToF setup increases the resolution to the range of >10000 (FWHM) as compared to that of mere 1000 obtained using a linear ToF.

#### *1.1.1.3.2 Orbitrap*

The orbitrap is a relatively new type of mass analyser which was proposed first in 1996 in a patent by Alexander Makarov (Makarov 1996) and introduced in 2000 (Makarov 2000). The design of this mass analyzer is based on the concept of orbital trapping, which was first introduced by Kingdon (Kingdon 1923).

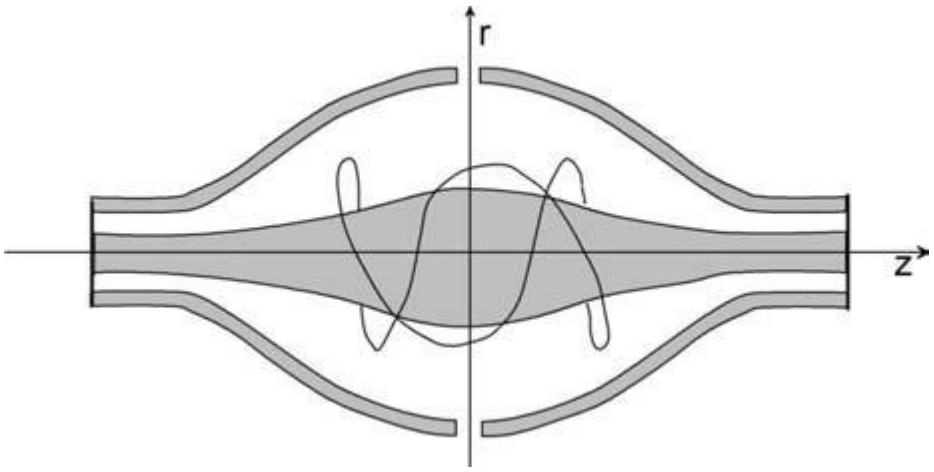



**Figure 1.4 A schematic diagram of an Orbitrap mass analyser** (de Hoffmann & Stroobant 2007)**.**

The orbitrap is essentially an electrostatic trap that uses a Fourier transform to obtain mass spectra. The mass analyser design consists of a central electrode that is spindle shaped which in turn is enclosed in a barrel shaped outer electrode consisting of two equal parts with a small gap between them (**Figure 1.4**). The ions are injected tangentially (i.e. along the axis *r* as depicted in **Figure 1.4**) through the space between the two parts of the external electrode with a kinetic energy of some keV. An opposing electrostatic voltage to the ions, of the order of several kilovolts (typically ~3.2 kV) is applied to the central spindle shaped electrode while the outer electrode is at ground potential. The electrostatic attraction of the ions towards the central electrode is compensated by the centrifugal force; as a result the ions start to oscillate in the trap in complex spiral patterns around the central electrode (in the direction of *z*-axis as shown in **Figure 1.4**) under the influence of an electrostatic field. The frequency of oscillations is inversely proportional to the square root of the *m/z*. The oscillating frequency of the ions is measured and Fourier transformed to obtain a mass spectrum (de Hoffmann & Stroobant 2007).

The main advantage offered by the orbitrap mass analyser is very high resolution rivalled only by an FTICR mass analyser. The resolving power of the orbitrap is proportional to the number of oscillations of the ion around the central electrode. An orbitrap offers dramatic increase in resolution and mass accuracy, compared to a time-of-flight (ToF) mass analyser. The orbitrap also provides high mass accuracy and increased dynamic range (Makarov 2000; Hu et al. 2005). In the studies described in this thesis an Orbitrap Exactive mass spectrometer (Thermo Scientific, UK) was used to obtain accurate mass data (as shown in **Figure 1.5**).

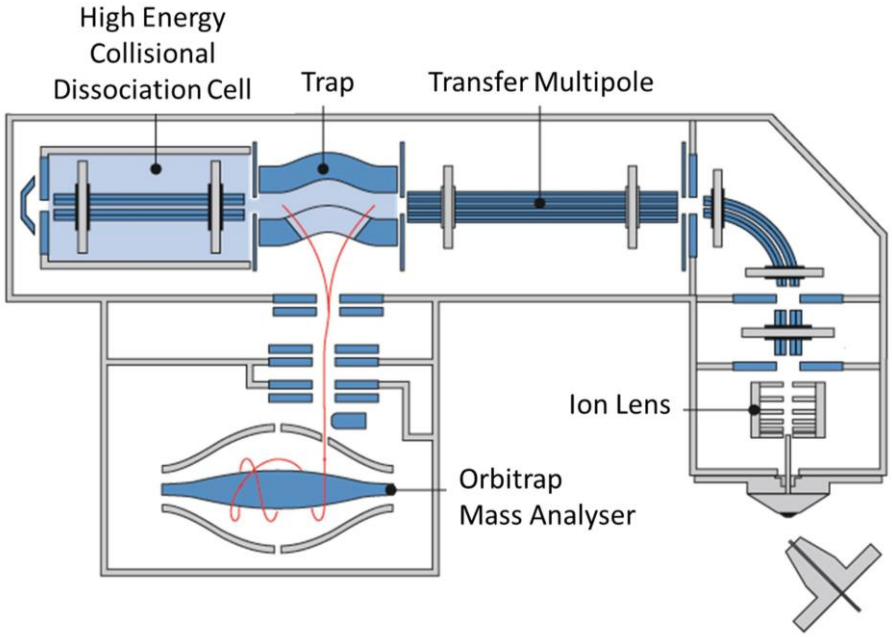



**Figure 1.5 Schematic of the Thermo Exactive Orbitrap mass spectrometer** (Thermo Scientific 2013)**.**

Charged analytes are transferred from the ion source to the orbitrap mass analyser through a series of ion optics that use a combination of voltages and a vacuum gradient. The ions can be held in the analyser for a selected duration of time and detected by the outer electrodes, which surround the orbitrap analyser.

#### *1.1.1.4 Detectors*

A detector is the final component of the mass spectrometer setup and it facilitates the transformation of the ions separated by the mass analyser into a usable signal i.e. a mass spectrum. The signal generated by the detector is proportional to the abundance of the incident ions. There are two major types of detectors used in modern day mass spectrometry electron multiplier or a multichannel plate detector. The detectors used in mass spectrometry have been reviewed elsewhere (Koppenaal et al. 2005).

### 1.1.2 Ion mobility spectrometry

#### *1.1.2.1 Overview of ion mobility spectrometry*

Modern day Ion mobility spectrometry (IMS) originated as plasma chromatography (Carr 1984; Karasek 1974). IMS is a gas phase electrophoretic technique which separates ions based on their relative mobilities in the presence of an external electric field and opposing drift gas (McDaniel & Mason 1973; Mason & McDaniel 1988). The electric field facilitates the acceleration of the ions, whereas collisions with the buffer gas decelerates them, giving rise to a constant drift velocity (Clemmer & Jarrold 1997). The mobility of an ion is a function of the collisional cross section ($\Omega$) (i.e. its size and shape), ionic charge and its reduced mass (Creaser et al. 2004). The mobility is determined by measuring the drift velocity of the ion under the influence of an electric field gradient and in the presence of a buffer gas in a drift tube. Hence, ion mobility of a particular analyte ion is characteristic of the analyte and the given buffer gas; and as a result it can be used as a detection technique for a single specific component from a mixture.

Modern day ion mobility spectrometry is subdivided into two main classes; drift tube ion mobility spectrometry (IMS) and differential mobility spectrometry (DMS) or field asymmetric waveform ion mobility spectrometry (FAIMS). Conventional IMS employs a constant, time-independent electric field applied to a drift tube and ions are separated on the basis of their interaction with opposing drift gas flow. In DMS or FAIMS, a time-dependent electric field is used to influence the ion transport properties of ions.

The development of IMS occurred over the past few decades, primarily as a method for detecting volatile and semi-volatile organic compounds in fields such as security and the military. This technique has found widespread application as a detector for chemical warfare agents in battlefields and for narcotics and explosives at airport security. IMS has been successfully employed for the detection and analysis of a wide array of chemical species such as semi-volatile compounds, carbon and silicon clusters, synthetic polymers and biomolecules (von Helden et al. 1993).

Conventional drift tube ion mobility spectrometry was not used in the work described in this thesis; however, the fundamental concepts of ion mobility are essential to understand

the mechanisms of separation by ion mobility and have been highlighted in the following sections.

### *1.1.2.2 Principles of ion mobility spectrometry*

Ion mobility spectrometry (IMS) determines the drift velocities ($v_d$) of the ionised analyte molecules in the weak electric field of the drift tube in the presence of a drift gas. The velocity of an ion in this setup can be determined from the equation:

$$v_d = KE$$

Where, $v_d$ is the ion's drift velocity (cm $s^{-1}$), K is the mobility of the ion ($cm^2$ $V^{-1}$ $s^{-1}$) and E is the electric field strength (V $cm^{-1}$). Hence, the drift velocity of the ion is directly proportional to its mobility.

The mobility (K) of an ion within any given drift cell can be defined by the Mason-Schamp equation:

$$K = \frac{3q}{16N}\sqrt{\frac{2\pi}{k\mu T_{eff}}}\frac{1+\alpha}{\Omega_{(T_{eff})}}$$

Where, *q* represents the charge state of the ion, *N* is the number density of the drift gas, *k* is the Boltzman's gas constant (1.380x10-$^{23}$ $m^2$ kg $s^{-2}$ $K^{-1}$), μ=(mM)/(m+M) is the reduced mass (kg) for the ion (m) and drift gas molecule (M), $T_{eff}$ is the effective temperature, Ω is the collisional cross section of the ion ($cm^2$) and α is the electric field correction factor. Thus, the mobility of an ion at a given temperature and gas pressure is determined by the mass, charge and the collisional cross section of the ion. However, this relationship holds true only in low electric fields i.e. where the ratio of *E/N* is ≤ 2 Townsend (Td). At higher values of *E/N* the mobility of the ion becomes dependent on the field strength, which forms the basic principle of DMS or FAIMS.

It is clear that the mobility measurements are dependent on the instrumentation and the parameters such as temperature and electric field strength used for the analysis. Any changes to the instrumentation or the parameters will affect the mobility of the ion. Changes to the temperature will alter the velocity of the ions. The number density of the drift gas will also affect the ion velocity, as increased number density of the drift gas will

cause increased collisions with the analyte ions, thus reducing the velocity of the analyte ions. As a result for comparison of ion mobility data, standardisation is required. Thus, for comparison of mobility measurements obtained under differing conditions and instrumentation, mobility is reported as reduced mobility ($K_o$), which normalises and scales the data to normal conditions of temperature (T in Kelvin) and pressure (P in Torr) of the drift gas.

$$K_o = K\left(\frac{273}{T}\right)\left(\frac{P}{760}\right)$$

Where, $K_o$ is the reduced mobility, $K$ is the mobility, $T$ is temperature (Kelvin) and $P$ pressure (Torr).

### *1.1.2.3 Travelling wave ion mobility spectrometry (TWIMS)*

The Waters Synapt HDMS mass spectrometer (Waters Corporation, Manchester, UK) was the first commercially available mass spectrometer capable of performing ion mobility measurements. A schematic of the instrument is depicted in **Figure 1.6**. The Synapt implements a new ion mobility method termed as the travelling wave ion mobility spectrometry (TWIMS).

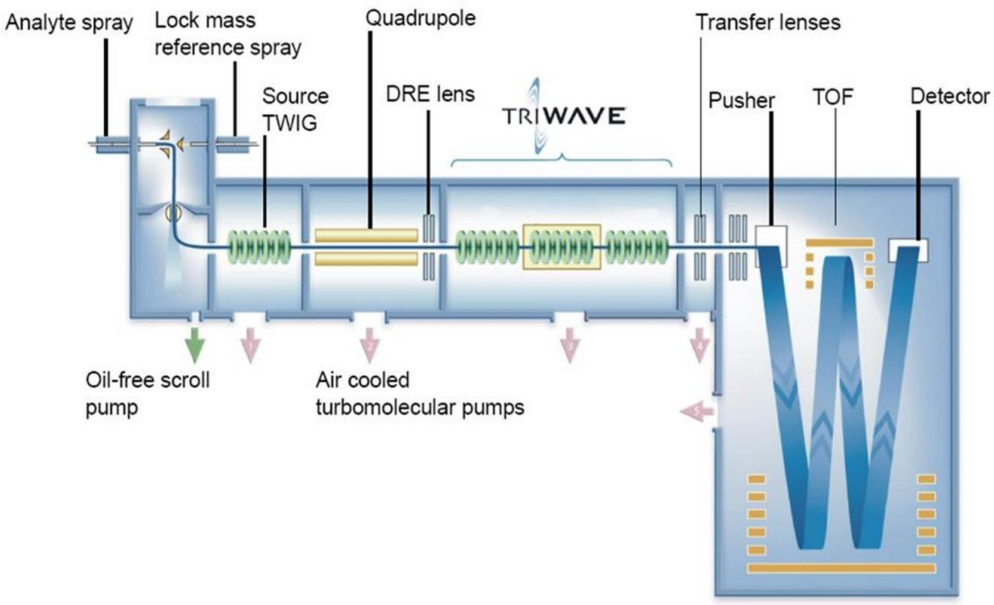



**Figure 1.6 Schematic of Waters Synapt HDMS mass spectrometer with tri-wave ion mobility drift cell** (Waters Corporation 2007)**.**

Unlike the conventional ion mobility devices, which employ a constant electric field for the propulsion of ions through a drift tube, travelling wave ion mobility spectrometers use a series of stacked ring ion guides (SRIG) in the presence of a neutral buffer gas. In a SRIG a sequence of ring-shaped electrodes is coaxially arranged as shown in **Figure 1.7**. Ions are transported through the SRIG in a manner similar to that of a quadrupole mass analyser, by the application of radio frequency (RF) voltages applied to the electrodes. Alternate ring electrodes have opposite phases of RF applied to them which radially confine the ions within the device causing them to travel along the axis.

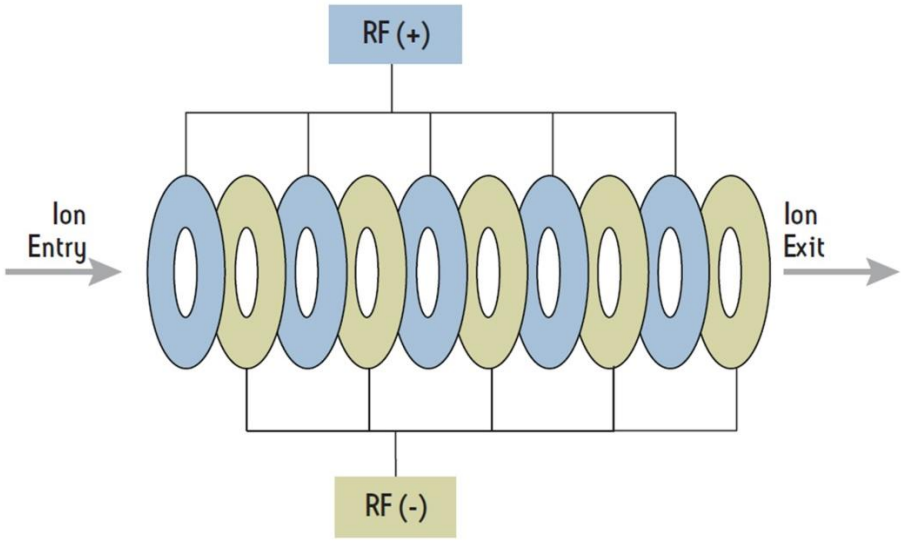



**Figure 1.7 Schematic of a stacked ring ion guide (SRIG) in a T-wave device** (Waters Corporation 2012)**.**

A direct current (DC) voltage is applied to a pair of adjacent ring electrodes. This superimposed DC voltage produces a barrier, which the ions cannot cross. If this DC voltage is stepped up to the next adjacent set of ring electrodes, it effectively moves the barrier forward a short distance and hence the superimposed moving DC potential is termed as 'travelling wave'. This causes any ions, which are in front of the barrier to move forward (Figure 1.8). Sequential stepping up the ion barrier from one end of the device to the other end produces a travelling wave which drives the ions through the device (Giles et al. 2004; Shvartsburg & Smith 2008).

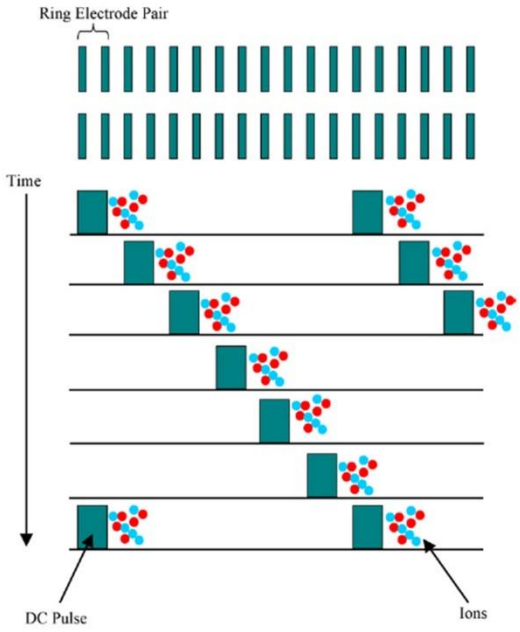



**Figure 1.8 Illustration of propulsion of ions through a T-wave device by application of a moving DC voltage barrier** (Pringle et al. 2007)**.**

The application of the DC voltage to a pair of ring electrodes is grouped and this pattern is repeated throughout the SRIG such that at any given instance, there are multiple travelling waves passing through the device (Figure 1.9).

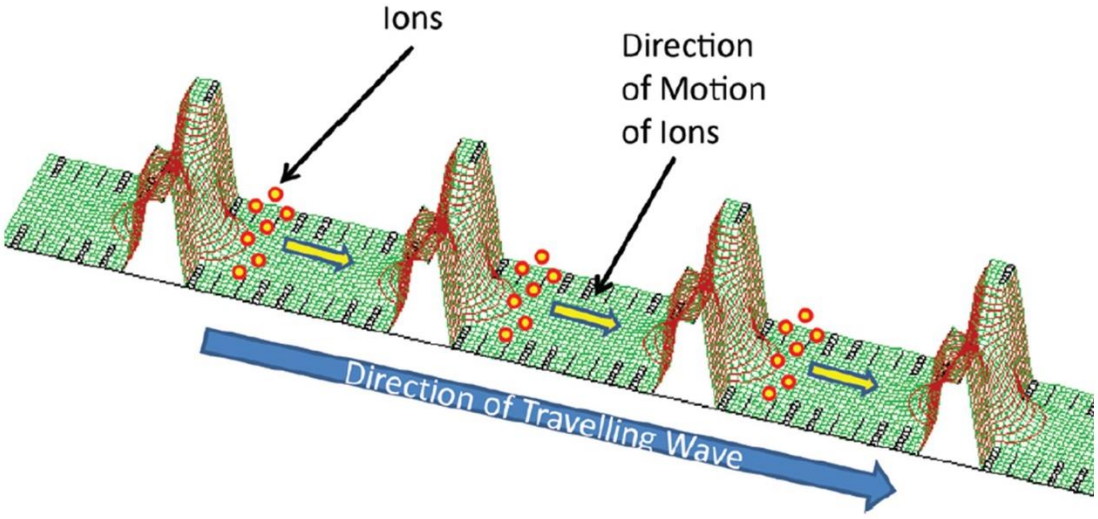



**Figure 1.9 A depiction of multiple travelling waves moving through the T-wave device and movement of ions (Giles et al. 2004).**

In the presence of a neutral gas in the TWIMS device, the transport of ions in the SRIG region is dependent on the travelling wave height (V), velocity of the travelling wave (m/s) and gas pressure (mbar). These factors influence the dwell time of the ion in the SRIG and careful manipulation of these factors can be used for separation of ions of different collisional cross section. Out of these three, the travelling wave height (V) is an important consideration. A large wave height causes ions to 'surf' the travelling wave and exit the SRIG, but with a lower wave height, the ions move along the decreasing potential gradient, whilst the interactions with the neutral buffer gas will cause 'drag' resulting in the ions rolling over the top of the wave to interact with the following travelling wave. Ions with high mobility with respect to the travelling wave will be propelled at the base of the travelling wave, whereas ions with lower mobility will roll over the top of the wave more frequently and will be transported through the SRIG at a lower velocity by subsequent travelling waves. Hence, the separation of the ions in the device is based on the interaction of the structural characteristics (i.e. size, shape) of the ions with the neutral buffer gas and the travelling wave (Pringle et al. 2007; Waters Corporation 2012). This process of interaction of the ions with the travelling wave provides mobility based separation of the ions in the SRIG forming the basis of travelling wave ion mobility spectrometry (TWIMS).

The travelling wave device in the Synapt HDMS (Waters Corporation, Manchester, UK) can be operated in a variable wave height mode i.e. the wave height can be gradually altered as the wave traverses the length of the SRIG region. This effectively increases the length of the drift tube providing better separation of the ions. Manipulation of the wave in such a manner imparts greater control over mobility based separations in TWIMS.

### 1.1.3 Ultra-high performance liquid chromatography (UHPLC)

The complex nature of biological samples demands pre-separation prior to mass spectrometric analysis. Chromatographic techniques essentially form the foundation of the separation sciences and hence are at the core of analytical chemistry. In recent years, high-performance liquid chromatography (HPLC) has become the method of choice for separations and analysis of complex mixtures in various fields such as general chemistry, pharmaceuticals, natural products, food safety, forensic analysis and '-omic' sciences (McMaster 2007). The need for separation of complex mixtures with varying physical and chemical properties along with technological advancements has led to the origins of a number of variants of liquid chromatographic techniques such as high-performance, ultra-pressure, supercritical fluid, hydrophilic interaction (HILIC) and ion chromatography. The versatility offered by liquid chromatographic techniques for the analysis of chemically diverse compounds is due to the customisation of LC through a choice of stationary phases along with additives such as buffers. Consequently, it offers greater control over separations. Selectivity for analytes is achieved in liquid chromatography by choice of columns in conjunction with modification of composition of mobile phases (Ardrey 2003). The core fundamental concepts of liquid chromatography have been extensively reviewed elsewhere (Ardrey 2003; McMaster 2007). Liquid chromatography is commonly combined with MS for high-throughput analysis of complex samples due to its high resolving power, robustness and easy hyphenation using electrospray ionisation (ESI).

Reversed phase liquid chromatography is the most commonly used form of LC separations and is based on the principle of hydrophobic interactions between analytes and a stationary phase. A non-polar stationary phase is employed which is typically a straight chain alkyl group (such as $C_{18}H_{37,}$ known as $C_{18}$ or ODS) linked covalently to a silica particle through a covalent siloxane bond. An aqueous, moderately polar mobile phase such as water combined with an organic solvent (methanol, acetonitrile etc.) is employed to separate analytes based on their hydrophobicity (Kellner et al. 2004). This results in the fast elution of polar molecules whilst less polar molecules are retained. A gradient elution can be applied to this system wherein the percentage of organic solvent in the eluent is gradually increased; the affinity of hydrophobic analytes towards the stationary phase is thus reduced resulting in faster elution times for such analytes. The

optimisation of the percentage of the organic solvent and the gradient profile is critical for ensuring maximum separation of analytes from any complex mixture prior to mass spectral analysis.

Recent years have seen evolution of traditional HPLC through technical advances giving rise to ultra-high performance liquid chromatography (UHPLC). UHPLC utilises smaller particle sizes (typically sub ~2 µm) or superficially porous particles for increased efficiency of chromatographic separation. The back-pressure generated by UHPLC columns is inversely related to the particle size of the media used to pack the column. Over the years the trend in liquid chromatography has been the development of efficient media (i.e. stationary phases) to improve chromatographic performance and reduce analysis times. This is achieved by building high performance pumps and instrumentation which can withstand high back-pressures generated by LC columns with smaller particle sizes or by using high efficiency particles such as core-shell stationary phases. These two technologies have been used in the research presented in this thesis and are outlined in the following sections:

#### *1.1.3.1 UPLC*

In HPLC columns, as the particle size decreases to less than 2.5 µm there is a significant increase in chromatographic efficiency which does not diminish even at increased mobile phase flow rates (Swartz 2005). As a result of using smaller particle size, speed and peak capacity can be increased. Ultra performance liquid chromatography employs sub 2 µm particles of materials such as bridged ethylene hybrid (BEH) $C_{18}$ as a packing material (Waters Corporation, Manchester, UK). This has been termed UHPLC or UPLC™ (Waters Corporation, Manchester, UK). UPLC takes advantage of advancements in particle chemistry combined with instrumentation which is capable of operating at higher pressures as compared to that used in HPLC systems.

#### *1.1.3.2 Poroshell*

Porous silica is widely used as a column packing material in analytical and preparative LC systems. Porous silica particles have proved useful for the analysis of macromolecules such as proteins, polypeptides and nucleic acids (Kirkland et al. 2000). The concept of pellicular particles was first proposed by Horvath *et al.* in late 1960s and

were initially intended for analysis of macromolecules (Horvath & Lipsky 1969). The evolution of porous silica particles over the years has led to the development of superficially porous silica microspheres as a packaging material for HPLC (Cohen et al. 2011). These have been termed core-shell or poroshell particles (Kirkland et al. 2000).

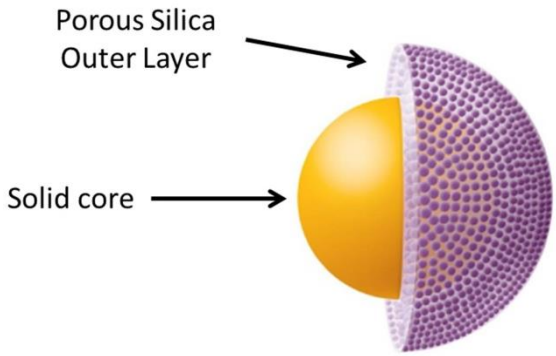



**Figure 1.10 Schematic depicting the structure of core-shell particles** (Agilent Technologies 2013)**.**

Poroshell particles consist of an impermeable solid inner core surrounded by a uniform porous silica layer and thus are significantly different to conventional porous silica particles (Kirkland et al. 2000). These particles typically have a solid core of ~1.7 µm coated by a porous layer of 0.5 ~µm thickness. These stationary phases have been claimed to offer 80-90% of the efficiency at reduced system back-pressures of the order of ~40-50% compared to that offered by sub 2µm particles (Eddinger-Anderson et al. 2010). Poroshell particles provide increased column performance comparable with UHPLC systems without generating excessive back-pressure; as a result these can be used using conventional HPLC setup without the loss of chromatographic efficiency.

However, the nature and design of these particles render them difficult for manufacturing. The preparation of thin porous shells (of the order of 0.25 to 0.5 µm) around spherical non-porous silica cores (of 1.2 to 1.7 µm diameters) is an expensive process. Along with the difficulties in manufacturing of core-shell particles, the efficient packing of the narrow bore columns with these particles presents further challenges in production of poroshell columns (Cohen et al. 2011).

The use of poroshell columns has not been limited to macromolecules and with technological advancement in the manufacturing of stationary phases, poroshell particles with smaller pore sizes which are well suited for small molecule analysis (of the order of 120 Å as opposed to 300 Å for macromolecule analysis) have been made possible, extending the analytical range of this technique (Agilent Technologies 2013).

In the research described in this thesis, both UPLC systems and Poroshell $C_{18}$ particle columns have been used in separate metabolite profiling studies providing comparable chromatographic performance.

### 1.1.4 Hyphenation of LC-MS with IM

Gas chromatography (GC) due to the nature of the technique functions with gaseous phase ions for the separation and thus can be easily interfaced with a mass spectrometer (MS). However, the interfacing of liquid chromatographic (LC) systems with a MS presents greater challenges as the analytes are dissolved in LC mobile phase and the molecular weights and chemical nature of the species separated and analysed by LC vary to a larger extent than the volatile and semi-volatile species analysed by GC. The problem of LC-MS interfacing was addressed with the introduction of the technique of electrospray ionisation (ESI), which facilitated the generation of gaseous phase ions from the condensed phase. This enables the 2-dimensional separation of components of a complex mixture based first on retention time ($t_R$) and then by the mass-to-charge ratio (*m/z*).

In the past few decades, the most significant advancement in the field of ion mobility as a valuable analytical technique has been its hyphenation with mass spectrometry (IM-MS). The hyphenation of IM with MS presents the opportunity to obtain valuable data which is not possible from IMS or MS alone. McDaniel *et al.* were the first to couple low-field ion mobility drift cells with magnetic sector MS (McDaniel et al. 1962). The hyphenation of IM with time-of-flight (ToF) and quadrupole (Q) mass spectrometers followed shortly after making MS the detector of choice for IM separations. A historical account and a review of hyphenation of various types of ion mobility such as IMS, DMS, FAIMS and TWIMS with mass spectrometry is presented by Kanu *et al.* (Kanu et al. 2008).

Both hyphenated techniques described here i.e. LC-MS and IM-MS have been immensely successful with widespread applications, the reason for this being the orthogonality of the separation techniques involved. However, there are frequent cases where compounds which co-elute by LC-MS are not resolved by the separating ability of mass spectrometry and similarly cases involving the analysis of challenging, complex mixtures which require pre-separation prior to analysis with IM-MS (Varesio et al. 2011). A solution would be the combination of all three techniques of separation. Modern day ion mobility systems, such as travelling wave devices are capable of coping with short dwell time of ions in the ion mobility drift cells and thus can easily keep up with rapid separation offered by UHPLC coupled with a fast scanning mass spectrometer. Along with these advantages, ion mobility spectrometry, being a gas phase separation technique, is orthogonal in nature to mass spectrometry and can easily be combined with UHPLC. This hyphenation of UHPLC-IM-MS has been explored for analysis of complex mixtures such as detection of trace impurities in drug products (Eckers et al. 2007) and targeted analysis of multiple compounds from a urine sample (Varesio et al. 2011).

The combination of ultra-high performance liquid chromatography, ion mobility and mass spectrometry presents itself as a powerful technique for the analysis of complex mixtures such as metabolomic samples derived from biofluids. The hyphenation of ion mobility with UHPLC-MS for metabolite profiling of biofluids is explored in the research presented in this thesis.

## 1.2 METABOLOMICS

Metabolism is the sum total of all the chemical processes occurring in the body of a living organism or within a cell that are necessary for the maintenance of life. Metabolites are the molecules which partake in the process of metabolism and include a range of chemically diverse species including molecules such as amino acids, sugars, vitamins, organic acids, as well as macromolecules such as nucleic acids and proteins. Metabolism encompasses: anabolism – the assembly of small molecules to produce larger compounds, catabolism – the breakdown of larger molecules into smaller molecules and biotransformation – the conversion and elimination of endogenous or exogenous compounds through excretion. All exogenous and endogenous compounds are subject to these processes and thus form an intricate network of metabolic pathways. The entire complement of small molecular weight metabolites (< 1500 Da) is termed the metabolome, and the term was first introduced by Oliver *et al.* (Oliver et al. 1998). It also includes interactions with metabolite products and substrates such as lipids, small peptides, vitamins, nucleic acids and other protein factors (Krastanov 2010).

The idea that changes in biological fluids or tissues are indicative of the physiological state of an organism or a diseased state, dates as far back as ancient Greece. The efforts for exploring biochemical systems enclosed within the human body have been ongoing for centuries. 'Urine charts' which linked colours, smells and tastes of urine to various conditions were widely used in ancient times as diagnostic indicators of diseased states. Modern day metabolomics explores the same basic principle of relating chemical fingerprint of biofluids or tissues to the biological state of the organisms (Nicholson & Lindon 2008).

Since the Human Genome Project in 1990s, there has been a revolution in the approaches and techniques used in molecular biology and biochemistry. The interest and advancement over the years in these related fields has led to a better understanding of the human biochemistry (National Institutes of Health 2013). The outcomes from early stages of the Human Genome Project suggested that genetic differences might be able to account for various disease processes. This led to the discipline of transcriptomics. Research in the recent years has highlighted the complex relationship between environmental interactions and genetic make-up and its impact on the phenotype of living organisms. A full understanding of these complex interactions cannot be achieved

only at genomic, transcriptomic or proteomic level alone. Even the combination of genomics and proteomics does not provide a full understanding of the cellular functions of living systems. This is due to the fact that both these techniques ignore the dynamic metabolic state of the organism (Fiehn 2002).

### 1.2.1 'Omics' sciences

The word 'omics' is derived from the Greek suffix, 'ome' which means whole or complete and refers to scientific areas in systems biology. The sciences which come under the umbrella of 'omics' e.g. genomics, proteomics, transciptomics, lipidomics and metabolomics aim to provide a global perspective on the biological functions in order to better understand life.

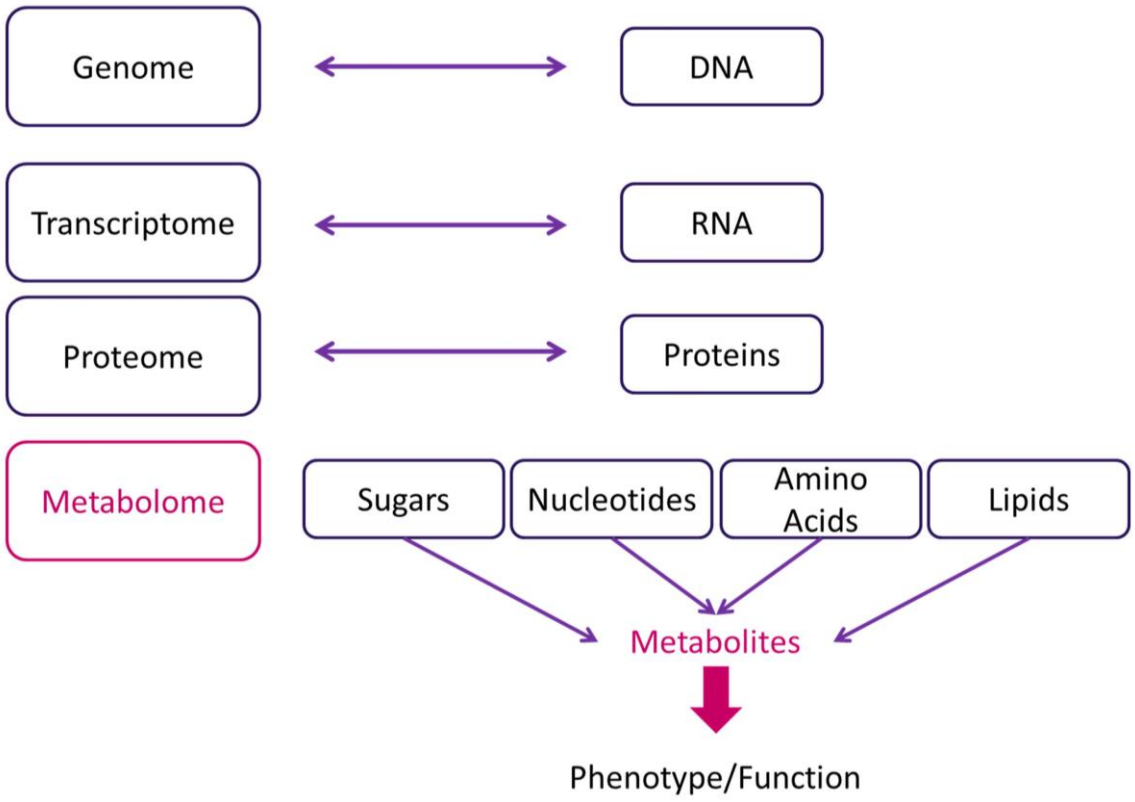



**Figure 1.11 Coverage of '-omic' sciences and place of metabolomics as the most predictive phenotype**

### 1.2.2 Overview of metabolomics

The science of investigating the metabolome started in the late 1960s, when Dalgliesh *et al.* combined GC with flame ionisation detection (FID) for investigations on metabolism (Dalgliesh et al. 1966), although the term 'metabolomics' was not used (Griffiths et al. 2010). Metabolomics arose from the work carried out in the 1980s. The field grew at a rapid pace with its coupling with advanced pattern recognition and multivariate statistical techniques (Nicholson et al. 1983). The term metabonomics was coined in 1999 by Nicholson *et al.* and was defined as "the quantitative measurement of the dynamic, multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification" (Nicholson et al. 1999). It is the systematic study of the unique chemical fingerprints of low molecular weight endogenous metabolites or metabolite profiles in biological samples (Zhang, Sun & Wang 2012).

The field of metabolomics has witnessed the parallel development of closely related fields of metabonomics and metabolic profiling.

### 1.2.3 Metabonomics, metabolomics and metabolic profiling

The terms metabonomics and metabolomics have separate origins, however, they have been used interchangeably throughout the literature. One description suggests that the term metabolomics applies to plant studies, whereas metabonomics was reserved for studies involving animals. A second approach inferred that the choice of terms was based on the analytical technique used, with metabolomics referring to studies utilising LCMS and metabonomics using NMR (Mishur & Rea 2011). Nicholson *et al.* describe the difference between metabolomics and metabonomics as 'philosophical rather than technical'. Metabonomics has been described to measure the global metabolic response of living systems to biological stimuli or genetic modification; the focus of which is to understand systematic change through time. Whereas, metabolomics aims to characterise and quantify all small molecules in given samples. However, these terms are used interchangeably in literature and the analytical methods and statistical procedures used are the same (Mishur & Rea 2011; Nicholson & Lindon 2008).

The terms metabolomics, metabonomics and metabolite profiling will be used interchangeably in the context of this thesis.

### 1.2.4 Approaches in metabolomics

Regardless of the analytical technique utilised for analysis, metabolomic studies can be divided into two distinct approaches; targeted analysis of specific metabolites and the more challenging non-targeted methods.

Targeted metabolomic studies focus on a single metabolite or a sub-set of a particular class of metabolites which are representative of a particular metabolic pathway. The strategy of analysis by this approach is to evaluate metabolic target compounds identified from the study of metabolic pathways. Hence, targeted metabolomics approach is referred to as a 'hypothesis-testing' approach. This approach proves to be inefficient at detecting previously uncharacterised compounds as only selected analytes are evaluated. Analytical methods routinely employed for targeted metabolomics include LC-MS operated in selected ion recording (SIR) or multiple reaction monitoring (MRM) modes. An account of analytical strategies utilised in targeted metabolomics is given later in this thesis and is discussed elsewhere (W. Lu et al. 2008).

Non-targeted metabolomic analysis refers to the comprehensive global metabolic profiling of all measurable analytes in a sample. In the non-targeted approach, no prior knowledge of the samples is assumed and hence is recognised as 'hypothesis-generating' approach.

A significant portion of modern metabonomics research is focussed on metabolic profiling to measure the entire metabolic content within a biological system. The reasoning behind the ambition to study the entire metabolic content was the recent advances in whole genome sequencing (genomics) and comprehensive proteome profiling (proteomics). Comprehensive profiling of the metabolome in similar fashion to that of genome and proteome will provide insight into biochemistry, which connects the genome and the proteome with a particular phenotype, while accounting for physiological development and interactions with environment, i.e. nutrients, diurnal variations etc. This will provide a more complete picture of the biochemistry for integrating the knowledge with systems biology. As unbiased monitoring of all the analytes is expected, the data generated from global metabonomic profiling by any analytical method is vast; and thus complex data interpretation strategies such as multivariate analysis and chemometrics are required. Also the number of metabolites

present in a particular system cannot be predicted. Metabolites have a diverse variety of chemical structures and properties and thus extraction, enrichment, separation and analysis steps are difficult.

### 1.2.5 Biofluids in metabolomics

Recent metabolomic studies of biomedical significance show a growing trend of the use of biofluids, cell or tissue extracts. These include plasma, whole blood, urine, cerebrospinal fluid, synovial fluid, semen, serum, tissue homogenate etc. Metabolomic studies have also investigated biofluids such as seminal fluid, amniotic fluid, lung aspirate as well as digestive fluids. However, biofluids such as urine and plasma are routinely used for various disease diagnosis and in clinical trial settings as they are easily obtainable and established procedures and protocols exist for their collection (Zhang, Sun, Wang, et al. 2012).

Various metabolomic studies have shown the potential of mapping and analysing the perturbations in biological systems through biofluid analysis for disease diagnosis and biomarker discovery (Claudino et al. 2012). The goal of modern day metabolomics is to identify novel and specific biomarkers for early disease diagnosis and to better understand disease mechanism than traditional biochemistry and histopathological methods. At the same time the emphasis is on non-invasive biomarkers and thus patient comfort. Hence, non-invasively collected biofluids such as saliva and urine are gaining precedence as compared to traditionally used diagnostic biofluids such as plasma (Zhang, Sun, Wang, et al. 2012).

### 1.2.6 Analytical techniques in metabolomic analyses

Metabolomics investigations particularly those involving global metabolite profiling consist of two keys steps; data collection of metabolite fingerprints and information mining. To facilitate this, there is a requirement for analytical platforms capable of producing global metabolite profiles from complex biological samples. An ideal analytical technique should enable the analysis to be performed directly on the sample, with minimal sample pre-treatment, high throughput, no bias towards particular classes of metabolites and should be both highly and equally sensitive to all the components in the sample. Realistically, there is no single analytical technique, which provides all the

desired properties. Hence, there is a general trend towards carrying out metabolite profiling using multiple analytical platforms to maximise the coverage of the metabolic profiles. The metabolites are present in large number, diversity, concentration, dynamic range and complexity in biological samples; and thus separation science forms an essential part of the study of all metabolites in these samples.

An overview of the key analytical techniques, which have been used for metabonomic studies, is given below. Particular emphasis is given to liquid chromatography-mass spectrometry (LC-MS) and ion mobility-mass spectrometry (IM-MS).

#### *1.2.6.1 Overview of analytical techniques in metabolomics*

Nicholson and co-workers pioneered the metabonomics approach using nuclear magnetic resonance (NMR) spectroscopy in 1999 (Nicholson et al. 1999); and since then NMR has been the predominantly used technique in metabonomic investigations along with mass spectrometry (MS). The hyphenation of MS with separation-based techniques such as gas, liquid and thin-layer chromatography (GC, LC and TLC) or capillary electrophoresis (CE) have been used as a means of detection and identification of metabolomic biomarkers (Wilson et al. 2005; An et al. 2010; Ciborowski et al. 2010; N. Li et al. 2010; Varghese et al. 2010; Jansson et al. 2009; Sreekumar et al. 2009; Issaq et al. 2008; Want et al. 2010; Kaddurah-Daouk 2006). There have been a limited number of reports of the use of techniques such as fourier transform infra-red (FT-IR) spectroscopy, fourier transform-ion cyclotron resonance-mass spectrometry (FT-ICR-MS) and other electrophoretic techniques for metabolome-related studies (Jansson et al. 2009). Numerous researchers have debated the advantages and disadvantages of chromatographic and other analytical techniques in combination with mass spectrometry (Pasikanti et al. 2008; Xiayan & Legido-Quigley 2008; Theodoridis et al. 2008; Issaq et al. 2008; Zhou et al. 2012; X. Lu et al. 2008; Dunn & Ellis 2005).

The main advantage of NMR spectroscopy is its non-destructive nature wherein the sample can be further analysed using a different analytical technique. Apart from being a rapid, cost-effective, high-throughput method; its minimal sample pre-treatment requirement is ideal for global metabolomic profiling research. It is a powerful technique for obtaining structural information on small organic molecules and thus has been employed for metabolomic studies of various biofluids such as urine, cerebrospinal fluid,

plasma, saliva etc. and has been effectively used to identify and quantify their components (Tang et al. 2009; Dunn & Ellis 2005; Nicholson et al. 1985; Nicholson et al. 1983). However, the drawback of this method is its low sensitivity, despite the use of higher magnetic field instruments for increasing the resolution, compared to other techniques such as mass spectrometry (MS). The sample volume requirements for the method are also high in comparison with other techniques. Biofluid samples contain large proportion of water in them, which causes peak broadening in NMR spectra and thus may mask detection of other metabolites of interest. In addition, some classes of analytes such as sulphates cannot be detected by NMR spectroscopy. Thus, although NMR has been the analytical workhorse for metabonomics research, the improved sensitivity and low sample volume requirements offered by hyphenated mass spectrometric (MS) methods (mainly liquid chromatographic) has invoked growing interest for metabolic profiling investigations.

Gas chromatography (GC) combined with mass spectrometry (MS) provides a very powerful method for the analysis of volatile organic compounds. With improvements in technology and fused-silica columns, it provides high resolution, high efficiency, better reproducibility and requires a small sample volume. With advancements such as two-dimensional GC (GCxGC) coupled with time-of-flight mass spectrometry (ToFMS), the analysis time is significantly reduced and detection of large number of metabolites is possible in a single analysis. A major prerequisite for GC-MS analysis is that the analytes should be volatile and thermally stable. However, most of the metabolites from biological fluids are polar and non-volatile in nature; thus these cannot be directly analysed using GC-MS and require derivatisation. This along with long chromatographic run times reduces the throughput of the method. Despite its shortcomings, gas chromatography coupled with mass spectrometry (GC-MS) provides a valuable tool for analysis of volatile components of biofluids in metabonomic studies (Pasikanti et al. 2008; Dalgliesh et al. 1966) and has been implemented for metabolite profiling studies of biofluids such as urine and plasma (Zhang, Sun, Wang, et al. 2012).

Capillary electrophoresis (CE) methods which separate analytes based on their size-to-charge ratio, have also been used with limited success for characterising the metabolome (Sugimoto et al. 2010; Ramautar et al. 2008). Recent advances in technology have enabled the hyphenation of CE with mass spectrometry using an electrospray ionisation

source, however with limited success. The main advantage of CE is its applicability to hydrophilic metabolites and its suitability for analysis of ionic and polar compounds, which may not be retained by reversed-phase liquid chromatographic columns. Although CE exhibits good sensitivity, selectivity and fast analysis times; the main shortfall of this method is difficulty in interfacing with mass spectrometry. Since the volume of sample used for this technique is in the nano-litre range, detection of low abundance metabolites is difficult to achieve. Most of the applications of this method in metabolomics involve targeted approach rather than comprehensive profiling of the metabolome. CE-MS is considered as a complementary method to liquid chromatography.

Along with hyphenating mass spectrometry and chromatographic methods, MS is also used as a high-throughput screening tool by direct infusion mass spectrometry (DIMS). In this method, biological sample extracts are directly injected or infused into an electrospray ionisation mass spectrometer. A single mass spectrum is acquired per sample and thus has high complexity. Comprehensive analysis of the metabolome is not possible using this method, and is thus used as a screening tool in metabolomics. Applications of DIMS are mainly concerned with microbial and plant metabolomics.

#### *1.2.6.2 LC-MS in metabolomics*

Liquid chromatography in tandem with mass spectrometry is widely used in many biochemical and chemical analyses especially in '-omics' sciences to analyse the composition of complex samples. The use of chromatographic technique such as LC prior to MS detection decreases the complexity of data by imparting the separation of analyte species. This has proved to be invaluable for the analysis of complex samples such as biofluids, plant and microbial extracts as well as tissue homogenates encountered in metabolomic studies (Theodoridis et al. 2008; X. Lu et al. 2008; Zhou et al. 2012).

The application of LC-MS to metabolomic studies is relatively recent, but has shown rapid and continuing increase in number of publications based on this approach. MS is already a well-established technique in drug-metabolite and small molecule analysis. MS combined with a separation technique such as LC shows potential for comprehensive metabolite profiling studies.

Most commonly used ionisation technique for interfacing LC-MS is with electrospray ionisation, which is capable of producing gas-phase ions from the liquid phase. ESI has

advantages such as excellent quantitative response, reproducibility, high sensitivity and simple sample preparation. ESI is a soft-ionisation technique and is suitable for global profiling of metabolites.

Like all the analytical techniques, LC has benefited from the advancement in technology. The use of UHPLC (Section 4.1.2) has resulted in higher chromatographic efficiencies and shorter analysis times and the advantages of UHPLC-MS for metabolic profiling studies are evident. LC-MS methods have already been successfully used for metabonomic profiling studies pertaining to biofluids such as urine and plasma for applications including nutritional studies and biomarker detection.

### 1.2.7 Application of ion mobility spectrometry (IM-MS) in metabolomics

Ion mobility spectrometry (IMS) complements traditional mass spectrometric (MS) separations as IM separates ions based on shape and collisional cross sections (CCS) of the analyte ions. The implication of this being structural isomers that share the same *m/z* and are unresolved by mass spectrometry can be separated by ion mobility (Lapthorn et al. 2012). Dwivedi *et al.* have shown the applications of combination of ion mobility with mass spectrometry (IM-MS) for detection and separation of isomers of complex carbohydrates (Dwivedi, Bendiak, et al. 2007) as well as enantiomeric separation of chiral compounds (Dwivedi et al. 2006). In addition, IMS is a rapid separation technique with typical scan times in the order of milliseconds. Qualities such as the speed of analysis and the orthogonal nature to mass spectrometry have demonstrated the potential of this technique to be used as a second-dimension of separation following GC or LC to provide enhanced coverage of complex samples without compromising the high throughput nature of chromatography-mass spectrometry based techniques (Bedair & Sumner 2008).

Despite these advantages, the applications of IM-MS in metabolomic investigations, especially for metabolite profiling studies remain limited in number, although the interest in the use of this technique is growing. Dwivedi *et al.* investigated and characterised the potential of ESI-IM-MS for separation, detection and identification of components from complex metabolic mixtures (Dwivedi, Wu, et al. 2007). Bacterial cultures of *E. coli* were analysed for metabolites in this study and over 500 metabolites were separated and detected. The potential advantages offered by ESI-IM-MS and MALDI-IM-MS were

highlighted for the analysis of complex samples such as bacterial metabolome of *E. coli* (Dwivedi, Puzon, et al. 2010). Ambient pressure ion mobility mass spectrometry has been used for global and targeted studies on neuronal metabolome of rats (Kaplan et al. 2013). This study focussed on probing the metabolome from a rat brain tissue after acute cocaine dosing to demonstrate the potential of using high throughput IM-MS separations for metabolomic studies for pharmacological investigations in drug abuse.

Harry *et al.* have explored an approach for enhancing coverage of urinary metabolome by LC-IM-MS (Harry et al. 2008). Drift tube ion mobility spectrometry combined with HPLC and MS was evaluated for profiling of urinary metabolome using rat urine and strategies for data mining were demonstrated. Other applications include metabolite profiling of human blood by high resolution ion mobility mass spectrometry (IM-MS) (Dwivedi, Schultz, et al. 2010) and analysis of breath using differential mobility spectrometry (Basanta et al. 2010) amongst a growing trend of metabolomic studies utilising IM for increased selectivity and enhanced coverage of the metabolome.

## 1.3 INTRODUCTION TO DATA MINING METHODS IN METABOLOMICS

Metabolic profiling experiments generate vast amounts of data that can be overwhelming in terms of size and complexity and the complexity of the generated dataset only increases with advancements in the sophisticated analytical technologies used to acquire the data. Hyphenated chromatography and mass spectrometry platforms can generate extended datasets, which need data mining and deconvolution before conclusions that address the biological question under scrutiny can be drawn (i.e. obtaining results which are 'fit-for-purpose'). Due to the size and the complexity of the data produced, the use of univariate statistical techniques is not suitable. Thus, data mining methods based on pattern recognition and multivariate statistics are an essential analytical tool for scrutinising metabolomic data, especially data obtained from non-targeted global metabolomic analyses (Jansen & Westerhuis 2012). The primary aim of the data mining in metabolomics is to reduce the complexity of the data whilst maintaining all the relevant information contained in the huge quantity of the data generated. In the case of metabolomic experiments focussed on non-hypothesis driven, untargeted metabolite profiling studies, the data handling and statistical data processing needs to be done with utmost care to avoid false discoveries of biomarkers (Broadhurst & Kell 2006).

Mass spectrometry has been hailed as the gold standard for the study of the wide chemical diversity and range of concentrations of the metabolome. There has been significant research into various methodologies for processing of the data obtained from mass spectrometric techniques. These data processing techniques have been reviewed extensively by Boccard *et al.* (Boccard et al. 2010) and Goodacre *et al*. (Goodacre et al. 2004), including techniques to combine data from different analytical sources, such as LC-MS and $^{1}$H-NMR, to obtain comprehensive coverage of the metabolome (Forshed et al. 2007).

### 1.3.1 Data pre-treatment

Typically, metabolomic raw data generated from chromatographic and mass spectrometric techniques consists of several hundred to several thousand chromatographic peaks and their associated mass spectra arising from multiple components in the metabolic profile. However, the dataset is also contaminated with noise and peak areas of different metabolites usually differ. Thus, to extract relevant

information from such a large and heterogeneous and contaminated dataset, several data pre-treatment steps followed by powerful regression and projection methods are required.

A crucial step prior to data mining is to assemble a primary dataset from which noise and missing data-points are removed. The datasets obtained from hyphenated chromatography-mass spectrometry platforms often require extensive pre-processing before being subjected to statistical analysis. Over a course of the analysis of an analytical sample batch, chromatographic peaks may demonstrate retention time shifts or intensity variations due to instrumental and other factors such as temperature, pH, column contamination etc. (Want et al. 2010). Hence, chromatographic peak detection, noise filtering, retention time alignment, normalisation etc. are required to achieve good quality data matrix prior to subsequent statistical analysis.

The first stage is to generate a data matrix table containing retention time & *m/z* pairs (or features) and peak areas (or peak heights in some cases). Once a data matrix has been constructed, the data needs to be further treated in order to concentrate on the features associated with biological variation in the data. LC-MS or GC-MS data generated from metabolomics profiling studies inherently possess several properties that make data analysis difficult without proper pre-treatment. For example, there are large differences in abundances, peak areas and intensities, which need to be compensated for as a detectable biological effect is usually related to the variation in concentration rather than the absolute concentration of a metabolite. The variation induced due to the biological change of interest is mixed with other types of variation such as systematic, random and arising from other technical sources. This variation of interest needs to be distilled from other sources of variation.

### 1.3.2 Normalisation

Normalisation is the first data transformation process, which aims to compensate for instrumental and systematic errors arising in the metabolic profiling workflow. These errors commonly arise during chromatography or mass spectrometry, such as variation in the intensities of the metabolites between different samples due to loss of sensitivity etc. (Redestig et al. 2009). A metabolic profiling experiment usually involves the analysis of a number of samples and even with modern high-throughput methods this requires

extended run times. As a result, during the course of the analysis of the samples, variations arise in the characteristics of the data, such as chromatographic retention times and mass spectral peak intensities. A solution to this problem is to introduce an internal standard, such that the intensity of the endogenous metabolite can be related to that of the internal standard. Stable-isotope labelled counterparts of the metabolites of interest are ideal candidates to be used as internal standards in targeted metabolite profiling, where normalisation of the data to the intensity of the internal standard is rapidly becoming the method of choice. The other option for non-targeted metabolic profiling is to normalise the data to the intensity of the total ion current (TIC) for that particular sample. This method of normalisation serves to avoid additional steps in sample pre-treatment, as well as avoiding the possibility of interaction of the added internal standard with endogenous molecules, such as ion suppression, or chemical interactions with other endogenous metabolites in the sample of interest. Normalisation is a critical step in data pre-treatment as it serves as means of circumventing systematic errors and removal of between sample variations for detection of biological differences in intensities of metabolites of interest.

### 1.3.3 Multivariate statistical approaches

Multivariate statistical approaches are well suited to the extraction of valuable significant information from the vast amount of data generated from '–omics' techniques in systems biology. Two multivariate statistical analysis methods are most commonly used in metabolomics profiling studies *viz.* principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). These are projection-based methods where latent or artificial variables are constructed through modelling from observed variables i.e. a data matrix consisting of *m/z* and retention time pairs. Techniques which have evolved from these two basic approaches, such as orthogonal projection to latent structures (O-PLS) and O2-PLS, are also gaining popularity and have been reviewed elsewhere (Fonville et al. 2010). Depending on the aim of the analysis, unsupervised or supervised approaches are employed for data processing, for example observing trends in the data, group classification and obtaining an overview of the data.

#### *1.3.3.1 Principal component analysis (PCA)*

Principal component analysis (PCA) is the first and simplest method used in metabolic profiling. It assumes no prior knowledge of the samples and hence is called an 'unsupervised' method. PCA is a multivariate statistical analysis approach that facilitates the identification of differences and similarities between groups of samples. PCA was first conceived by Pearson in the year 1901, who defined it mathematically as an orthogonal linear transformation (Pearson 1901).

PCA forms the basis of modern day multivariate data analysis and the fundamentals of the method have been described in detail by Wold (Wold 1987). It attempts to reduce the dimensionality of the data while retaining the original variability in the data. PCA transforms a number of correlated variables into smaller number of uncorrelated variables called 'principal components'. Principal components are assigned so that the first component (PC1) explains/accounts for most of the variation in the dataset, the second (PC2) accounts for the next largest variation and so on. A PCA analysis proceeds in this manner with each component accounting for progressively smaller amounts of variance. Variance is a measure of the spread of the dataset. Consequently it can be used to assess groupings, outliers, trends and homogeneity in datasets. PCA can also be used to determine if the analytical system remained within acceptible control limits by the measurement of quality control (QC) samples spread throughout the analytical sample sequence and by assessing the clustering of the samples from this group.

#### *1.3.3.2 Partial least squares discriminant analysis (PLS-DA)*

Partial least squares (PLS) regression is a multivariate method of assessing a relationship between a descriptor matrix **X** and a response matrix **Y**. However, unlike PCA, PLS regression is a 'supervised' method and is directed by the response matrix **Y** to derive components from the descriptor matrix **X** to maximise the covariance in the data. In other words, the response matrix drives out the discriminant components from the dataset to maximise separation. PLS in its simplest form is a method for relating two data matrices and building predictive models based on the same (Wold 2001).

PLS regression can be used for discrimination analysis in the form of PLS-DA. This variant of the regression method is commonly used in metabolic profiling studies for classification or class-separation purposes. PLS and other derived supervised approaches

are used to tackle subtle problems in which relationships within datasets are buried in the background of other large multiplexed effects allowing the smaller features to be uncovered. For example, in the case of biological systems, a variety of factors such as age, gender, environment, metabolic rate etc. contribute to the composition of the metabolic profile of different individuals. Not all of these factors are relevant to the understanding of a biological variation of interest and the main source of variation in the data may not address the biological question being interrogated. As a result, these should be down-weighted during the analysis and importance should be given to the relevant source of variation in the data. This can be achieved through the supervised approaches and these have proved useful in modern day metabolic profiling (Fonville et al. 2010). However, one drawback of the PLS-DA method is that it can suffer from overfitting. This usually occurs because the number of samples used in metabolic profiling experiments is usually much smaller than the number of variables extracted from the analytical data.

### 1.3.4 Scaling and transformations

The validity and accuracy of the supervised and unsupervised models can be improved by the correct use of scaling and centering. Three types of scaling are predominantly used in multivariate statistics involving metabolic profiling *viz* unit variance (UV), centering with no scaling (Ctr) and Pareto scaling (Par). The UV scaling assigns each variable a weighting of one and as a result provides equal chance to all the variables of being expressed in the model with equal importance. In terms of mass spectrometric data, this means ions with high intensity will be compressed while those with low intensity will be magnified and, as a result, metabolites from both the ends of the dynamic range in terms of their concentration will be given equal importance. In Ctr method of scaling, the importance given to a variable is related to its signal, i.e ions with low amplitude will have less influence on the model as compared to those with high signal amplitude. The drawback of this method is that metabolites which are present in lower concentrations in the samples, even if important, will not influence the model thus causing loss of potentially important sources of variation. The compromise between UV and Ctr is the Par scaling. This method of scaling allows spectral regions with low amplitude to influence the model, but only if they show variation. This is the most popularly used method for scaling in metabonomic applications.

## 1.4 SALIVA

Saliva is often regarded as the ‘mirror of the body’. It’s a perfect surrogate medium to be applied for clinical diagnostic purposes. It has an old history of being studied but its physiological importance has only been recognized recently. In the past 50 years the pace of salivary research has accelerated with the advent of new techniques and approaches such as metabolomics, genomics, proteomics etc. The research has illuminated the biochemical and physicochemical properties of saliva, comprising the multifunctional roles it has in speech, lubrication, digestion of food and maintaining oral and general health (Mandel 1993). The interest in saliva has grown exponentially in the last decade with the finding that saliva was filled with hundreds of components that may serve to detect systemic disease or evidence of exposure to various harmful substances as well as provide biomarkers of health and diseased states (Schipper et al. 2007).

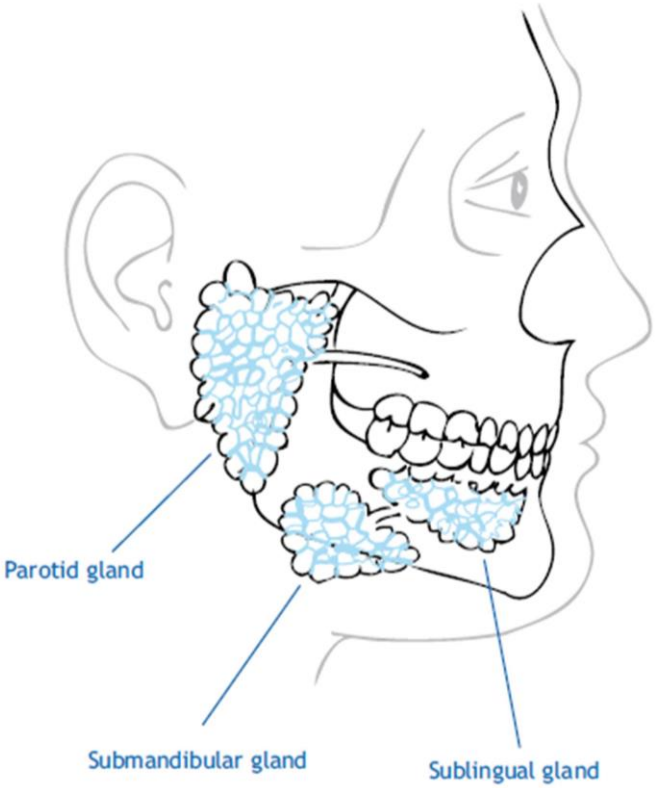



**Figure 1.12 Location of the major salivary glands in the oral cavity.**

### 1.4.1 Physiology and secretion

Human saliva is a complex bio-fluid consisting of secretions from a number of glands. These salivary glands are divided into three major salivary glands – the paired parotid glands, the submandibular glands and the sublingual glands which contribute to the largest percentage volume of saliva produced. The paired parotid glands are located opposite the maxillary first molars whereas the submandibular and the sublingual glands are situated on the floor of the mouth under the tongue **Figure 1.12**. The major salivary glands are aided by minor salivary glands located throughout the oral cavity (Carpenter 2013).

### 1.4.2 Composition of saliva

Saliva consists of secretions from major and minor salivary glands, gingival crevicular fluid, oral bacteria, their metabolites and food debris. Whole saliva consists of a variety of biologically relevant compounds including hormones, nucleic acids, peptides, proteins, metabolites, immunoglobulins, enzymes (lysozymes, lingual lipase, α-amylase etc.), mucins, electrolytes (Na, K, Ca, Mg, carbonates, phosphates etc.) and nitrogenous products such as urea and ammonia (Humphrey & Williamson 2001). Despite the complexity, it is a very dilute fluid comprising of 99% water.

Saliva is a highly variable and individualised fluid and it's chemical composition and flow is affected by various factors including age, gender, circadian rhythm, circannual cycle, psycho-emotional state, physical exercise, hydration, systemic diseases, medication, substance abuse and nutrition (De Almeida et al. 2008; Carpenter 2013).

### 1.4.3 Functions of saliva

Saliva performs a wide array of roles within the oral cavity, and its importance is not fully appreciated until it is absent or its quality or quantity is diminished (Humphrey & Williamson 2001). Saliva is responsible for a number of functions involving maintenance of oral health and homeostasis. It aids in the production of food bolus during mastication, aids in the digestive process (through enzymes such as amylase and lipase), protects the oral cavity against mechanical damage by lubrication as well as provides first line of defence against pathogens. The lubrication of oral tissue also aids in swallowing of food and speech. It forms an integral part of the sense of taste by acting as

a solvent for ions while maintaining a balanced pH in the oral cavity by acting as a buffer (Walsh 2007). Another important function of saliva is the maintenance of the health of the oral mucosa through growth factors which promote wound healing in the oral cavity and curtailing destructive enzymes. Saliva ensures the stability in the oral cavity and aids in maintaining integrity of tooth enamel (Pedersen et al. 2002).

### 1.4.4 Saliva as a diagnostic bio-fluid and its clinical applications

In the past, saliva or oral fluid has been neglected as a diagnostic medium and preference has been given to other more invasively collected biofluids such as blood, serum and plasma. Saliva, however has the advantage of allowing non-invasive sample collection and substances present in saliva are in their biologically active form (Pink et al. 2009). A good correlation between levels of compounds in saliva as compared to plasma has been reported, which makes saliva an attractive alternative to invasively obtained matrices such as whole blood, plasma and CSF.

Saliva has been used in clinical diagnosis for more than a century. Ancient traditional Chinese medicine believed saliva to be 'sibling' biofluid of blood and that these two share a common origin. It was believed that changes in saliva were indicative of the wellness of an individual. The viscosity, odour, taste as well as patient's gustatory sensation of their saliva were considered as symptoms of diseased states e.g. over-secretion of saliva was linked with heartburn. These theories were earliest marks of use of saliva as a diagnostic biofluid (Zhang et al. 2009a).

The pace of research into saliva as a diagnostic medium has accelerated in recent years with the advent of new techniques, which have provided a wealth of information regarding the biochemical and physicochemical properties of saliva. Significant advancements have been made in the development of salivary diagnostic tools (Wong 2012). The potential for early detection of systemic diseases or exposure to harmful substances has been demonstrated (Pink et al. 2009; Walsh 2007). Saliva has been used as a tool for monitoring steroid, peptide and immune markers in sport and exercise science (Papacosta & Nassis 2011). It has also been utilised as a non-invasive biomatrix in pharmacokinetic studies (Raju et al. 2013). Salivary biomarkers have found applications in occupational and environmental medicine (Soo-Quee Koh & Choon-Huat Koh 2007) as well as proving valuable in clinical applications (Zhang et al. 2009b).

Salivary α-amylase has been identified as a stress-related biomarker for autonomic nervous system activity (DeCaro 2008). Measurement of late evening salivary cortisol concentrations by tandem LC-MS has been used as a screening test for Cushing's syndrome (Raff 2011). Effects of aerobic exercise have also been investigated using a variety of analytical platforms using saliva (González et al. 2008; Papacosta & Nassis 2011). The majority of the applications listed here have utilised LC-MS or NMR as the method of choice for analysis, which highlights the trend of instrumentation utilised in salivary analysis.

The complexity, biological significance of saliva and the advancements in analytical technology combined with modern day bioinformatics strategies (Ai et al. 2012) have shown the immense potential of this matrix in metabolite profiling. The non-invasive nature of the sampling and good-correlation with levels of compounds present in blood suggests that saliva maybe a valuable alternative to blood or urine profiling for screening of biomarkers (Chiappin et al. 2007; Williamson et al. 2012).

The following table highlights some of the metabolomics/biomarker studies which have been published with saliva as the biofluid.

| **Technique** | **Application** | **Author/Publication Year** |
|---|---|---|
| **NMR** | Metabonomic analysis of periodontis | (Aimetti et al. 2011) |
| **NMR** | Diurnal analysis of salivary profiles | (Bertram et al. 2009) |
| **NMR** | Effect of dietary standardisation on metabolomic profiles | (Walsh et al. 2006) |
| **CE/MS** | Pancreatic, breast, oral cancer profiles | (Sugimoto et al. 2010) |
| **Assays** | Biomarkers in blood and saliva comparison | (Williamson et al. 2012) |
| **LC/MS** | Signatures of oral cancer and leukoplakia | (Wei et al. 2010) |
| **LC/MS** | Saliva sample preparation for metabolomics | (Álvarez-Sánchez et al. 2012) |

### 1.4.5 Metabolite profiling of saliva

The utility of saliva for metabolite profiling has been demonstrated in a variety of applications. The term 'Salivaomics' was coined in 2008 suggesting the vast amount of interest and applications related to saliva based '–omic' research including metabolomic, proteomic, transcriptomic and genomic research (Wong 2012).

A wide range of analytical platforms have been explored for the metabolite profiling studies on saliva. Bertram *et al.* demonstrated the ability of $^{1}H$ nuclear magnetic resonance (NMR) for metabolite profiling of saliva. The diurnal effect on salivary metabolite profile, as well as differences between morning and night saliva samples, were also evaluated using NMR (Bertram et al. 2009). A holistic approach was employed by Takeda *et al.* for understanding the salivary metabolome using NMR (Takeda et al. 2009). There have been examples of comparative studies between salivary, plasma and urinary metabolite profiles following an acute dietary standardisation using NMR (Walsh et al. 2006). NMR has also been used for investigating salivary metabolomic signatures of chronic periodontal disease (Aimetti et al. 2011). Techniques such as capillary electrophoresis coupled with mass spectrometry (CE-MS) have been used to identify oral, breast and pancreatic cancer specific profiles from salivary metabolite profiling (Sugimoto et al. 2010). Salivary metabolome has been successfully utilised for diagnosis of oral squamous cell carcinoma (OSCC) or oral cancer and leukoplakia by untargeted metabolite profiling using LC-MS (Wei et al. 2010). LC-MS is rapidly becoming the method of choice for untargeted metabolite profiling studies, however, the applications of this technique for profiling of saliva are limited in number.

### 1.4.6 Sampling of saliva

The complex nature of biochemical and physico-chemical properties presents challenges for the use of saliva as a research material (Schipper et al. 2007). The composition of saliva is highly variable both intra- and inter-participants and is also dependent on time of the day. Hence, standardisation of the saliva collection is an important consideration of salivary research.

### *1.4.6.1 Collection devices*

There exists a large number of commercially available sample collection devices for saliva sampling, such as synthetic swabs, micro-sponges, cotton pledgets etc. Most of these are optimised for particular assays or for specific class of compounds, such as steroids or measuring specific analytes such as DHEA-S (Whetzel & Klein 2010), cortisol (Poll et al. 2007). These devices vary in size, shape and the materials out of which they are made. Some devices are equipped with salivary stimulant materials to aid in production of saliva to obtain desirable volume of sample specimen. Most of the commercially available devices are made in such a way that they can be used to collect saliva for general-purpose analysis.

However, not all saliva sample collection devices perform the same (Smolders et al. 2009; Donzella et al. 2008). Saliva is a mixture of oral fluids secreted by different source glands e.g. submandibular gland which is situated in the lower area of the mouth between cheek and jaw. Thus, a saliva collection device, such as an adsorbent swab, which is placed in this region will have a higher concentration of analytes secreted from the submandibular gland. Hence, in such cases the saliva sample obtained will not be a representative sample of the whole saliva. Therefore, although these commercially available saliva sampling devices can be used for a variety of applications and targeted analyses, they are subject to producing interference or loss of analytes due to adsorption or reaction with the sample device material. All these factors prevent such collection devices from being a universal sample collection solution.

### *1.4.6.2 Sample collection*

Passive drool is the most widely used technique for the collection of whole saliva. It is highly recommended as it is approved to be used with almost all analytes. Since it involves no adsorbent materials during sampling, interference caused by such materials can be avoided. Sampling methodology for passive drool is simple, efficient and very cost effective as no complicated sampling devices or salivary stimulants are utilised (Donzella et al. 2008; Navazesh & Christensen 1982).

During the passive drool approach of saliva collection, the participants are instructed to sit comfortably with their heads tilted forward and keep their eyes open (as this is a factor which influences salivary flow). Participants are handed a glass collection vial for

the saliva and asked to let saliva pool in front of their mouths; the saliva is then dribbled out between slightly parted lips into the glass collection vial held near the lips at regular intervals. Stimulated saliva generated by spitting is different in composition to unstimulated saliva obtained by passive drool method and hence spitting is to be discouraged during sample collection (Navazesh & Christensen 1982).

#### *1.4.6.3 Comparison of saliva collection approaches*

The methods for collection of saliva can be classified under two major groups *viz.* stimulated and unstimulated saliva collection (Navazesh 1993; Donzella et al. 2008). The stimulated methods of saliva collection include mastication, spitting or the use of salivary stimulants such as citric acid. The alternate approach of unstimulated saliva collection includes methods such as passive drool or draining wherein the participants allow saliva formed in the oral cavity to be dribbled off the lower lip into a collection vial. Metabolite concentrations have been reported to be higher in unstimulated saliva as compared to stimulated saliva (Takeda et al. 2009). Bacterial levels have been reported to be 14 fold higher in samples collected by stimulated method of spitting as compared to unstimulated samples obtained by passive drool approach (Chiappin et al. 2007). Saliva collected by mastication e.g. from chewing on a Salivette – a cotton roll, and then recovered by centrifugation of the cotton roll produces saliva with lesser viscosity, contamination with cotton wool as well as the possibility of loss of metabolites by adsorption onto the cottonwool (Poll et al. 2007). Hence, to minimise such effects, passive drool approach of saliva collection is commonly used in metabolomic research.

#### *1.4.6.4 Storage*

The procedures used for the storage of saliva are an important consideration while using saliva for analytical and metabolomic applications. The composition of the saliva sample is prone to change due to enzymatic or bacterial activity immediately after sample collection. The bacteria contained in saliva samples continue to metabolise components from saliva post sampling procedures, producing low levels of new components. If precautions are not taken during storage, saliva can also produce degradation products from the salivary components as many of them have short biological half-lives. There is also the risk of loss of compounds from saliva due to adsorption of compounds from saliva onto surfaces of storage containers of transfer mediums used. Many components

of saliva are thermally labile i.e. susceptible to temperature changes and are affected to a greater extent.

### 1.4.7 Saliva sample preparation

Pre-treatment of saliva prior to its analysis is often required to remove interferences or for concentration of analytes for optimum detection. A popular technique for pre-treatment of saliva is the use of solid phase extraction (SPE) which facilitates the clean-up of sample as well as concentration of the sample.

The larger the number of steps in sample pre-treatment, the greater the likelihood of human error increases, which may result in loss of reproducibility of sample preparation. As a consequence the sample may not be an accurate representation of the metabolic composition of the sample. Hence, limiting the number of steps of sample pre-treatment is desirable.

Human saliva contains a host of proteins representative of various biological activities such as food digestion (α-amylase, maltase, peptidases, phosphatises etc.) and protective functions in the gastrointestinal tract (immunoglobulins - IaG, lysozymes, lactoferrins, histatins, cystatins etc.); various proteins involved with cardiovascular (kallikrein, histamine, renin, tonin) and nervous system functions (neuroleukin, nerve growth factor etc.) are also present (Grigoriev et al. 2003). Other proteins are also present which have antibacterial and/or antiviral properties. Multifunctional glycoproteins in saliva, such as mucins provide protection against mechanical damage and prevention of dehydration (Turner & Sugiya 2002). Thus, it is clear that human saliva contains significant levels of proteins and thus sample pre-treatment for metabolomic applications should focus on removal of proteins from the samples to improve the selectivity of metabolites in saliva.

With the growing interest in ‘salivaomics’, sample preparation protocols for the analysis of saliva by LC-MS are being established. Alvarez-Sanchez *et al.* compared a variety of sample preparations techniques for the metabolite profiling of saliva to obtain the maximum number of features by liquid chromatography-time-of-flight-mass spectrometry (LC-ToF-MS) (Álvarez-Sánchez et al. 2012). An investigation carried out by Polson *et al.* dealt with the optimisation of protein precipitation from biofluids. This involved the evaluation of various protein precipitation techniques, such as addition of

variety of organic solvents, acids, salts and metal ions. The effectiveness of protein precipitation by employing various ratios of organic solvents was also evaluated. Acetonitrile in the ratio of 2:1 to biofluid sample was found to be most efficient with protein removal efficiency of >96% (Polson et al. 2003). Theodoridis *et al.* suggested a minimal sample pre-treatment methodology based on the removal of particulate matter from biofluids, followed by protein precipitation using solvents such as acetonitrile for LC-MS based metabolomics (Theodoridis et al. 2008). This approach of protein precipitation combined with ultrasonication for homogenising the viscous saliva samples was utilised in the research presented in this thesis.

## 1.5 RATIONALE BEHIND THE RESEARCH

Metabolomics is a technologically driven science. In recent times, mass spectrometry (MS) has occupied a central role in metabolomics which is growing along with the development and improvements in MS technologies (Bedair & Sumner 2008). The advent of new technologies such as IMS and FAIMS, as well as the continuous advancements in instrumentation and data processing methodologies, has shown promise for metabolomics. Although technologies such as travelling wave ion mobility spectrometry (TWIMS) and FAIMS have existed for over a decade, their use in combination with mass spectrometry for metabolomic studies on biofluids remains largely under explored (Kanu & Hill 2004; Dwivedi, Wu, et al. 2007).

Modern day metabolomics has seen a trend towards the use of non-invasively obtained biofluids such as saliva, urine and breath etc. for applications such as biomarker discovery and disease diagnosis (Smolders et al. 2009). The utility of saliva as an analytical tool in metabolomics has already been demonstrated with a variety of applications (Esteban & Castaño 2009). Good correlation between saliva and the compounds present in blood have been established. These factors coupled with the non-invasive sampling of saliva makes it an attractive alternative to invasively obtained biofluids for metabolomic studies. However, there are limiting factors to the collection of saliva as mentioned in section 1.4.6, which highlight the need for establishing protocols for the collection, storage and pre-treatment of saliva prior to its effective use in metabolomic setting.

Thus, the work presented in this thesis aims to evaluate the use of ion mobility in metabolic profiling studies involving plasma and saliva, develop and validate methods using the combination of UHPLC-IM-MS as well as establish standard protocols for the collection, storage and pre-treatment of a valuable biofluid - saliva.

## 1.6 AIMS AND OBJECTIVES

The objective of the research described in this thesis was the development of novel methods based on mass spectrometry and ion mobility spectrometry for metabolite profiling and the detection of diagnostic and prognostic biomarkers of diseases. The remainder of this thesis is divided into four chapters as described below:

1. Chapter 2 describes the development and evaluation of a method based on UHPLC-IM-MS for detection of biomarkers of chemopreventive efficacy of rice bran in colorectal cancer as evaluated from a $APC^{min}$ mouse model of colon carcinogenesis. It highlights the advantages of using ion mobility for the clean-up of contaminants from mass spectrometric data post analysis.
2. Chapter 3 describes the potential of non-invasively obtained biofluid – saliva, for metabolite profiling analysis. Untargeted global metabolite profiling analysis of saliva was carried out on samples obtained from healthy controls and individuals with asthma to identify potential metabolomic biomarkers of the disease.
3. Chapter 4 describes the development and validation of a method for metabolite profiling of saliva based on UHPLC-IM-MS. The advantages of the developed method for the targeted and untargeted metabolite profiling analysis of saliva have been highlighted. The developed method was used for untargeted metabolite profiling of samples obtained after induced physiological stress to identify the metabolomic biomarkers. The advantages offered by ion mobility towards increasing the analytical space and the identification of the discriminant metabolites by comparison with standards have been presented.
4. Chapter 5 presents a conclusion to the thesis and details further work.

## 1.7 CHAPTER ONE REFERENCES


Agilent Technologies, Agilent Poroshell 120 columns for HPLC and UHPLC -Technical Note. **2013,** URL: www.agilent.com/chem/poroshell120appnotes

Ai, J.Y.; Smith, B.; Wong, D.T.W. *Int. J. Oral Sci.* **2012**, 4, 85–87.

Aimetti, M.; Cacciatore, S.; Graziano, A.; Tenori, L. *Metabolomics* **2011,** 8, 465–474.

Alikhanov, S.G. *Sov. Phys. JETP-USSR* **1957,** 4, 452–453.

De Almeida, P. V.; Gregio, A.M.T.; Machado, M.A.N.; De Lima, A.A.S.; Azevedo, L.R. *J. Contemp. Dent. Pract.* 9. **2008**, 72–80.

Álvarez-Sánchez, B.; Priego-Capote, F.; Luque de Castro, M.D. *J. Chromatogr. A.* **2012,** 1248, 178–181.

An, Z.; Chen, Y.; Zhang, R.; Song, Y.; Sun, J.; He, J.; Bai, J.; Dong, L.; Zhan, Q.; Abliz, Z. *J. Proteome Res*. **2010**, 4071–4081.

Ardrey, R.E. Liquid Chromatography - Mass Spectrometry: An Introduction. John Wiley & Sons: Chichester, U.K. **2003**.

Banerjee, S.; Mazumdar, S. *Int. J. Anal. Chem*. **2012**, Vol 2012, 1-40.

Basanta, M.; Jarvis, R.M.; Xu, Y.; Blackburn, G.; Tal-Singer, R.; Woodcock, A.; Singh, D.; Goodacre, R.; Thomas, C.L.P.; Fowler, S.J. *Analyst* **2010**, 135, 315–320.

Bedair, M.; Sumner, L. *TrAC Trends Anal. Chem*. **2008**, 27, 238–250.

Bertram, H.C.; Eggers, N.; Eller, N. *Anal. Chem*. **2009**, 81, 9188–9193.

Boccard, J.; Veuthey, J.L.; Rudaz, S. *J. Sep. Sci*. **2010**, 33, 290–304.

Borsdorf, H.; Eiceman, G. *Appl. Spectrosc. Rev*. **2006**, 41, 323–375.

Broadhurst, D.I.; Kell, D.B. *Metabolomics* **2006**, 2, 171–196.

Carpenter, G.H. *Annu. Rev. Food Sci. Technol*. **2013**, 4, 267–276.

Carr, T.W. *Plasma chromatography. Plenum Press*. **1984.**

Chiappin, S.; Antonelli, G.; Gatti, R.; De Palo, E.F. *Clin. Chim. Acta* **2007,** 383, 30–40.

Ciborowski, M.; Javier Rupérez, F.; Martínez-Alcázar, M.P.; Angulo, S.; Radziwon, P.; Olszanski, R.; Kloczko, J.; Barbas, C. *J. Proteome Res*. **2010**, 9, 4131–4137.

Claudino, W.M;, Goncalves, P.H.; di Leo, A.; Philip, P. A; Sarkar, F.H. *Crit. Rev. Oncol. Hematol.* **2012**, 84, 1–7.

Clemmer, D.E.; Jarrold, M.F. *J. Mass Spectrom*. **1997**, 32, 577–592.

Cohen, S.A.; Guiochon, G.; Gritti, F. *J. Chromatogr*. A **2011**, 1218, 1915–1938.

Cotter, R.J. *Anal. Chem.* **1992**, 64, 1027–1039.

Creaser, C.S.; Griffiths, J.R.; Bramwell, C.J.; Noreen, S.; Hill, C. A., Thomas, C.L.P. *Analyst* **2004**, 129, 984.

Dalgliesh, C.E.; Horning, E.C.; Horning, M.G.; Knox, K.L.; Yarger, K. *Biochem. J*. **1966**, 101, 792–810.

DeCaro, J. A. *Am. J. Hum. Biol*. **2008**, 20, 617–619.

Dole, M. *J. Chem. Phys*. **1968**, 49, 2240.

Dole, M.; Hines, R.L.; Mack, L.L.; Mobley, R.C.; Ferguson, L.D.; Alice, M.B. *Macromolecules* **1968**, 1, 96–97.

Donzella, B.; Talge, N.M.; Smith, T.L.; Gunnar, M.R. *Dev. Psychobiol*. **2008**, 50, 714–717.

Dunn, W.B.; Ellis, D.I. *TrAC Trends Anal. Chem*. **2005**, 24, 285–294.

Dwivedi, P.; Bendiak, B.; Clowers, B.H.; Hill, H.H. *J. Am. Soc. Mass Spectrom.* **2007**, 18, 1163–1175.

Dwivedi, P.; Puzon, G.; Tam, M.; Langlais, D.; Jackson, S.; Kaplan, K.; Siems, W.F.; Schultz, A.J.; Xun, L.; Woods, A.; Hill, H.H. *J. Mass Spectrom*. **2010**, 45, 1383–1393.

Dwivedi, P.; Schultz, A.J.; Hill H.H.Jr. *Int. J. Mass Spectrom*. **2010**, 298, 78–90.

Dwivedi, P.; Wu, C.; Matz, L.M.; Clowers, B.H.; Siems, W.F.; Hill, H.H. *Anal. Chem.* **2006**, 78, 8200–8206.

Dwivedi, P.; Wu, P.; Klopsch, S.J.; Puzon, G.J.; Xun, L.; Hill, H.H. *Metabolomics* **2007**, 4, 63–80.

Eckers, C.; Laures, A.M.; Giles, K.; Major, H.; Pringle, S. *Rapid Commun. Mass Spectrom.* **2007**, 21, 1255–1263.

Eddinger-Anderson, M.; Clary, T.; Joseph, M. Chromatography-Based Characterization of Biologics (Agilent Technologies, White Paper) **2010**.

Esteban, M.; Castaño, A. *Environ. Int.* **2009**, 35, 438–449.

Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Science* **1989**, 80. 246, 64–71.

Fiehn, O. *Plant Mol. Biol.* **2002**, 48, 155–171.

Fonville, J.M.; Richards, S.E.; Barton, R.H.; Boulange, C.L.; Ebbels, T.M.D.; Nicholson, J.K.; Holmes, E.; Dumas, M.E. *J. Chemom*. **2010**, 24, 636–649.

Forshed, J.; Idborg, H.; Jacobsson, S.P. *Chemom. Intell. Lab. Syst*. **2007**, 85, 102–109.

Gaskell, S.J. *J. Mass Spectrom* **1997**, 32, 677–688.

Giles, K.; Pringle, S.D.; Worthington, K.R.; Little, D.; Wildgoose, J.L.; Bateman, R.H. *Rapid Commun. Mass Spectrom.* **2004**, 18, 2401–2414.

González, D.; Marquina, R.; Norelis, R. *Res. Sport. Med*. **2008**, 16, 128–137.

Goodacre, R.; Vaidyanathan, S.; Dunn, W.B.; Harrigan, G.G.; Kell, D.B. *Trends Biotechnol.* **2004**, 22, 245–252.

Griffiths, W.J.; Koal, T.; Wang, Y.; Kohl, M.; Enot; D.P.; Deigner, H.P. *Angew. Chem. Int. Ed. Engl*. **2010**, 49, 5426 – 5445.

Grigoriev, I. V; Nikolaeva, L. V; Artamonov, I.D. *Biochem. Biokhimiia* **2003**, 68, 405–406.

Guilhaus, M. *J. Mass Spectrom.,* **1995**, 30, 1519–1532.

Harry, E.L.; Weston, D.J.; Bristow, A.W.T.; Wilson, I.D.; Creaser, C.S. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci*. **2008**, 871, 357–361.

Von Helden, G.; Gotts, N.G.; Bowers, M.T. *J. Am. Chem. Soc*. **1993**, 115, 4363–4364.

De Hoffmann, E.; Stroobant, V. Mass Spectrometry Principles and Applications, Third Edition. ed. John Wiley & Sons, Ltd. **2007**.

Horvath, C.; Lipsky, S.R. *J. Chromatogr. S*ci. **1969**, 7, 109.

Hu, Q.; Noll, R.J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. *J. Mass Spectrom*. **2005**, 40, 430–443.

Humphrey, S.P.; Williamson, R.T. *J. Prosthet. Dent*. **2001**, 85, 162–169.

Ioanoviciu, D. *Int. J. Mass Spectrom*. **2001**, 206, 211–229.

Iribarne, J. V. *J. Chem. Phys*. **1976**, 64, 2287.

Issaq, H.J.; Abbott, E.; Veenstra, T.D. *J. Sep. Sci*. **2008**, 31, 1936–1947.

Jansen, J.J.; Westerhuis, J.A. *Metabolomics* **2012**, 8, 1–2.

Jansson, J.; Willing, B.; Lucio, M.; Fekete, A.; Dicksved, J.; Halfvarson, J.; Tysk, C.; Schmitt-Kopplin, P. *PLoS One* **2009**, 4, 6386.

Kaddurah-Daouk, R. *PLoS Med*. **2006**, 3, 363.

Kanu, A.B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill H.H.Jr. *J. Mass Spectrom*. **2008,** 43, 1–22.

Kanu, A.B.; Hill, H.H.; **2004,** *LPI April/May*.

Kaplan, K.A.; Chiu, V.M.; Lukus, P.A.; Zhang, X.; Siems, W.F.; Schenk, J.O.; Hill H.H.Jr. *Anal. Bioanal. Chem*. **2013**, 405, 1959–1968.

Karasek, F.W. *Anal. Chem.* **1974**, 46, 710–720.

Kebarle, P.; Peschke, M. *Anal. Chim. Acta* **2000**, 406, 11–35.

Kellner, R.; Mermet, J.; Otto, M.; Widmer, H.M. Analytical Chemistry, Second. ed. Wiley **2004**

Kingdon, K.H. *Phys. Rev.* **1923**, 21, 408–418.

Kirkland, J.J; Truszkowski, F. A.; Dilks, C.H.; Engel, G.S. J. *Chromatogr. A* **2000**, 890, 3–13.

Koppenaal, D.W.; Barinaga, C.J.; Denton, M.B.; Sperline, R.P.; Hieftje, G.M.; Schilling, G.D.; Andrade, F.J.; Barnes, J.H. *Anal. Chem*. **2005**, 77, 418–427.

Krastanov, A. *Biotechnol. Biotechnol. Equip*. **2010**, 24, 1537–1543.

Lapthorn, C.; Pullen, F.; Chowdhry, B.Z. *Mass Spectrom. Rev*. **2012.**

Li, N.; Liu, W.; Li, W.; Li, S.; Chen, X.; Bi, K.; He, P. *Clin. Biochem*. **2010**, 43, 992–997.

Lu, W.; Bennett, B.D.; Rabinowitz, J.D. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci*. **2008**, 871, 236–242.

Lu, X.; Zhao, X.; Bai, C.; Zhao, C.; Lu, G.; Xu, G. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci*. **2008**, 866, 64–76.

Makarov, A. **1996**. Mass Spectrometer, US Patent: 5886364.

Makarov, A. *Anal. Chem*. **2000**, 72, 1156–1162.

Mamyrin, B.A.; Karataev, V.I.; Shmikk, D.V.; Zagulin, V.A. *J. Exp. Theor. Phys.* **1973**, 64, 82–89.

Mandel, I.D. *Crit. Rev. oral Biol. Med. an Off. Publ. Am. Assoc. Oral Biol*. **1993**, 4, 599–604.

Mason, E.A.; McDaniel, E.W. *NASA STI/Recon Tech. Rep. A* **1988**, 89, 15174.

McDaniel, E.W.; Martin, D.W.; Barnes, W.S. *Rev. Sci. Instrum*. **1962**, 33, 2.

McDaniel, E.W.; Mason, E.A. Mobility and diffusion of ions in gases. John Wiley and Sons, Inc., New York **1973**

McMaster, M.C. HPLC : A practical user's guide, Second Ed., Analytical Chemistry. Hoboken, N.J. : John Wiley. **2007.**

Mishur, R.J.; Rea, S.L. *Mass Spectrom. Rev*. **2011**, 1–26.

National Institutes of Health, **2013**. National Human Genome Research Institute.

Navazesh, M. *Ann. N. Y. Acad. Sci*. **1993**, 694, 72–77.

Navazesh, M.; Christensen, C.M. *J. Dent. Res*. **1982**, 61, 1158–1162.

Nicholson, J.; Timbrell, J.; Sadler, P. *Mol. Pharmacol*. **1985**, 27, 644–651.

Nicholson, J.K.; Buckinghamt, M.J.; Sadler, P.J. *Biochem. J*. **1983,** 211, 605–615.

Nicholson, J.K.; Lindon, J.C. *Nature* **2008**, 455, 1054–1056.

Nicholson, J.K.; Lindon, J.C.; Holmes, E. *Xenobiotica* **1999**, 29, 1181–1189.

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

Papacosta, E.; Nassis, G.P. *J. Sci. Med. Sport* **2011,** 14, 424–434.

Pasikanti, K.K.; Ho, P.C.; Chan, E.C.Y. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci*. **2008**, 871, 202–211.

Pearson, K. *Philos. Mag*. **1901**, 2, 559–572.

Pedersen, A.M.; Bardow, A.; Jensen, S.B.; Nauntofte, B. *Oral Dis*. **2002**, 8, 117–129.

Pink, R.; Simek, J.; Vondrakova, J.; Faber, E.; Michl, P.; Pazdera, J.; Indrak, K. *Biomed. Pap. Med. Fac. Univ. Palacky. Olomouc. Czech. Repub*. **2009**, 153, 103–110.

Poll, E.-M.; Kreitschmann-Andermahr, I.; Langejuergen, Y.; Stanzel, S.; Gilsbach, J.M.; Gressner, A.; Yagmur, E. *Clin. Chim. Acta*. **2007**, 382, 15–19.

Polson, C.; Sarkar, P.; Incledon, B.; Raguvaran, V.; Grant, R. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **2003**, 785, 263–275.

Pringle, S.D.; Giles, K.; Wildgoose, J.L.; Williams, J.P.; Slade, S.E.; Thalassinos, K.; Bateman, R.H.; Bowers, M.T.; Scrivens, J.H. *Int. J. Mass Spectrom.* **2007**, 261, 1–12.

Raff, H. *Pituitary* **2011**, 15, 64–70.

Raju, K.S.R.; Taneja, I.; Singh, S.P.; Wahajuddin. *Biomed Chromatogr*. **2013**, 27: 1354–1366.

Ramautar, R.; Mayboroda, O. A.; Deelder, A.M.; Somsen, G.W.; de Jong, G.J. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci*. **2008**, 871, 370–374.

Redestig, H.; Fukushima, A.; Stenlund, H.; Moritz, T.; Arita, M.; Saito, K.; Kusano, M. *Anal. Chem*. **2009**, 81, 7974–7980.

Rodrigues, J.A.; Taylor, A.M.; Sumpton, D.P.; Reynolds, J.C.; Pickford, R.; Thomas-Oates, J. *Adv. Carbohydr. Chem. Biochem*. **2007**, 61, 59–141.

Rohner, T.C.; Lion, N.; Girault, H.H. *Phys. Chem. Chem. Phys.* **2004**, 6, 3056.

Schipper, R.G.; Silletti, E.; Vingerhoeds, M.H. *Arch. Oral Biol*. **2007**, 52, 1114–1135.

Shvartsburg, A. A; Smith, R.D. *Anal. Chem*. **2008**, 80, 9689–9699.

Smolders, R.; Schramm, K.W.; Nickmilder, M.; Schoeters, G. *Environ. Health* **2009,** 8, 8.

Soo-Quee Koh, D.; Choon-Huat Koh, G. *Occup. Environ. Med*. **2007,** 64, 202–210.

Sreekumar, A.; Poisson, L.M.; Rajendiran, T.M.; Khan, A.P.; Cao, Q.; Yu, J.; Laxman, B.; Mehra, R.; Lonigro, R.J.; Li, Y.; Nyati, M.K.; Ahsan, A.; Kalyana-Sundaram, S.; Han, B.; Cao, X.; Byun, J.; Omenn, G.S.; Ghosh, D.; Pennathur, S.; Alexander, D.C.; Berger, A.; Shuster, J.R.; Wei, J.T.; Varambally, S.; Beecher, C.; Chinnaiyan, A.M. *Nature* **2009**, 457, 910–914.

Sugimoto, M.; Wong, D.T.; Hirayama, A.; Soga, T.; Tomita, M. *Metabolomics* **2010**, 6, 78–95.

Swartz, M.E. *J. Liq. Chromatogr. Relat. Technol*. **2005**, 28, 1253–1263.

Takeda, I.; Stretch, C.; Barnaby, P.; Bhatnager, K.; Rankin, K.; Fu, H.; Weljie, A.; Jha, N.; Slupsky, C. *NMR Biomed.* **2009,** 22, 577–584.

Tang, H.; Xiao, C.; Wang, Y. *Magn. Reson. Chem.* **2009**, 47 Suppl 1, S157–162.

Theodoridis, G.; Gika, H.; Wilson, I. *TrAC Trends Anal. Chem*. **2008**, 27, 251–260.

Thermo Scientific, Exactive Plus Orbitrap Mass Spectrometer [WWW Document]. Planetorbitrap. **2013,** URL http://planetorbitrap.com/exactive-plus#tab:schematic (accessed 7.14.13).

Turner, R.J.; Sugiya, H. *Oral Dis*. **2002**, 8, 3–11.

Varesio, E.; Le Blanc, J.C.Y.; Hopfgartner, G. *Anal. Bioanal. Chem*.**2011**, 402, 8, 2555-2564

Varghese, R.S.; Cheema, A.; Cheema, P.; Bourbeau, M.; Tuli, L.; Zhou, B.; Jung, M.; Dritschilo, A.; Ressom, H.W. J. *Proteome Res.* **2010**, 9, 2786–2793.

Walsh, L.J. *Int. Dent. South Africa (Australasian Ed*). **2007**, 2, 16–30.

Walsh, M.C.; Brennan, L.; Malthouse, J.P.G.; Roche, H.M.; Gibney, M.J. *Am. J. Clin. Nutr*. **2006**, 84, 531–539.

Want, E.J.; Wilson, I.D.; Gik, H.; Theodoridis, G;, Plumb, R.S.; Shockcor, J.; Holmes, E.; Nicholson, J.K., Gika, H. *Nat. Protoc.* **2010**, 5, 1005–1018.

Waters Corporation, Waters Synapt High Definition Mass Spectrometry System - Operator's Guide. **2007,** 7150012990th ed.

Waters Corporation, **2012**. T-WAVE DEVICES.

Wei, J.; Xie, G.; Zhou, Z.; Shi, P.; Qiu, Y.; Zheng, X.; Chen, T.; Su, M.; Zhao, A.; Jia, W. *Int. J. cancer*. **2010**, 129, 2207–2217.

Weston, D.J.; Bateman, R.; Wilson, I.D.; Wood, T.R.; Creaser, C.S. *Anal. Chem.* **2005**, 77, 7572–7580.

Whetzel, C. A.; Klein, L.C. *BMC Res. Notes* **2010**, 3, 204.

Williamson, S.; Munro, C.; Pickler, R.; Grap, M.J.; Elswick, R.K. *Nurs. Res. Pract.* **2012**, 1-4.

Wilson, I.D.; Plumb, R.; Granger, J.; Major, H.; Williams, R.; Lenz, E.M. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci*. **2005**, 817, 67–76.

Wold, S. *Chemom. Intell. Lab. Syst*. **1987**, 37–52.

Wold, S. *Chemom. Intell. Lab. Syst.* **2001**, 58, 109–130.

Wong, D.T.W. *J. Am. Dent. Assoc*. **2012**, 143.

Xiayan, L.; Legido-Quigley, C. *Electrophoresis* **2008**, 29, 3724–3736.

Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. *J. Proteomics* **2012**, 75, 1079–1088.

Zhang, A.; Sun, H.; Wang, X. *Appl. Biochem. Biotechnol*. **2012**, 168, 1718–1727.

Zhang, L.; Xiao, H.; Wong, D.T. *Mol. diagnosis Ther*. **2009,** 13, 245–259.

Zhou, B.; Xiao, J.F.; Tuli, L.; Ressom, H.W. *Mol. Biosyst*. **2012,** 8, 470–481.

# CHAPTER TWO:

# METABOLOMIC ANALYSIS OF THE CHEMOPREVENTIVE EFFICACY OF RICE BRAN IN COLORECTAL CANCER BY UHPLC-IM-MS.

## 2.1 INTRODUCTION

Colorectal cancer is the third most common cancer in the world accounting for over 9% of all cancer incidences (Haggar & Boushey 2009). It is the fourth most common cause of cancer related deaths worldwide. There is a wide variation in frequency of occurrences of colorectal cancer across the world due to factors such as differences in diet, consumption of red and processed meat, fibre intake, alcohol, body weight and physical activity. (Cancer Research UK 2011). Countries with the highest incidence rates include Australia, New Zealand, Canada, the United States, and parts of Europe. The countries with the lowest risk include China, India, and parts of Africa and South America (Haggar & Boushey 2009). Food and nutrition have a highly important role in the causation and prevention of colorectal cancers. There has been a significant amount of research into the chemo-preventive role of dietary fibre in protecting against benign and malignant colorectal cancers according to the WCRF, AICR report of 2007 (WCRF/AICR 2007).

Several studies linking colorectal cancer with dietary factors have been published. These include the impact of intake of polycylic aromatic hydrocarbons (PAH's) (Obayashi et al. 2009), dietary flavonoids (Cooke et al. 2006) etc.

In 1971 Burkitt highlighted low dietary fibre intake as a risk factor for colorectal cancer for the first time (Burkitt 1993). According to the WCRF/AICR report, there have been at least sixteen cohort studies and ninety-one case-control studies investigating the effect of dietary fibre on colon carcinogenesis. Most of these studies concluded that increase in dietary fibre reduced the risk of colorectal cancer (WCRF/AICR 2007). Conflicting point of views on the relationship between dietary fibre intake and colorectal cancer have been presented in reviews elsewhere (Obrador 2007; Peters et al. 2003; Sengupta et al. 2001). Despite extensive research, there has not been a definitive conclusion on the role of dietary fibre in protecting against colorectal cancer.

A study carried out by Verschoyle *et al.* showed that rice bran supplementation in diet at a concentration of 30% w/w interfered with adenoma development in a preclinical model of colon carcinogenesis based on the $APC^{min}$ mice (**Figure 2.1**) (Verschoyle et al. 2007).

### *2.1.1* APC$^{min}$ mice

**Figure 2.1 APC$^{min}$ mouse**

In 1989, Moser *et al.* discovered a mutation in the mouse genome that led to multiple intestinal neoplasia (Moser et al. 1989). The mutation was observed in progeny, which were derived from mice treated with ethylnitrosourea (**Figure 2.2**), which is a mutagen. These mice developed numerous adenomas throughout the intestinal tract. This phenomenon has been referred to as multiple intestinal neoplasia (MIN) and the gene responsible for the mutation was determined to be adenomatous polyposis coli (APC). Hence this model of colon carcinogenesis derived the name APC$^{min}$. This model is well characterised and has shown phenotypic and histopathological similarities with man. Thus it has been widely used as a pre-clinical model of colorectal cancer and consequently to test the therapeutic efficacy of new drugs and cancer preventive strategies (Leclerc et al. 2004).

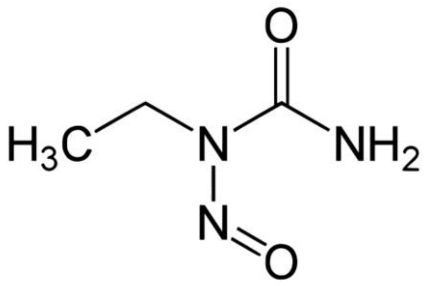



**Figure 2.2 Ethylnitrosourea - a highly potent mutagen**

### 2.1.2 Fibre and its role

Fibre exerts several effects in the gastrointestinal tract. In principle, there are three main mechanisms by which dietary fibre is believed to interfere with colorectal carcinogenesis. Firstly, fibre is thought to exert 'physical' effects such as increasing the faecal bulk and thus hastening faecal transit and binding potentially co-carcinogenic bile salts. Secondly, fibre can cause changes in the intestinal microflora. Lastly, carbohydrates can undergo fermentation and produce short chain fatty acids such as butyric acid, which are believed to exert anti-carcinogenic effects. The anti-cancer properties of fibre can be due to any of these mechanisms or a combination of two or more; but precise mechanisms for its probable protective role towards colorectal cancer are still not clearly understood.

The analytical methods used for the metabolite profiling of colorectal cancer have been critically reviewed elsewhere (Williams et al. 2013). In contrast to already published works, this study presents the first instance of the use of ion mobility spectrometry in combination with UHPLC-MS for metabolite profiling of mouse plasma.

The aim of the current study was to develop statistical models based on plasma metabolite profiling and to identify potential metabolomic biomarker ions of the efficacy of rice bran in colorectal cancer using ultra-high performance liquid chromatography-ion mobility-mass spectrometry (UHPLC-IM-MS).

## 2.2 CHAPTER TWO AIMS AND OBJECTIVES

The purpose of the work presented in this chapter was to explore novel methods for the metabolic profiling of plasma obtained from $APC^{min}$ mice using travelling wave ion mobility spectrometry (TWIMS) as an added dimension of separation in conjunction with UHPLC-MS. The particular objectives were:

- To develop a method based on UHPLC-IM-MS for metabolic profiling of plasma obtained from $APC^{min}$ mice.
- To explore the potential of ion mobility spectrometry for the improvement of mass spectral quality by selective removal of contamination such as PEG from the sample.
- To employ the developed method for the metabolomic analysis of the chemopreventive efficacy of rice bran in colorectal cancer.
- To identify potential metabolomic biomarker ions of the efficacy of rice bran in preventing colorectal cancer.

## 2.3 EXPERIMENTAL

### 2.3.1 Chemicals and reagents

Acetonitrile (analytical grade), water (analytical grade) and formic acid (>99.9%) were obtained from ThermoFisher Scientific (Loughborough, UK). L-carnitine (>99%) and hydrocortisone (>99%) were obtained from Sigma-Aldrich (Gillingham, UK). A 400 ng/mL stock solution of leucine enkephalin was obtained from Waters Corporation (Manchester, UK) and used to prepare a 2 ng/ml working solution of leucine enkephalin in water:acetonitrile (50:50) (v/v) to be used as a LockMass (*m/z* 556.2771) solution. A standard mixture of 0.01 mg/mL L-carnitine and hydrocortisone was prepared in water:acetonitrile (95:05) (v/v) with 0.1% formic acid.

### 2.3.2 $APC^{min}$ mice bran supplementation and plasma collection

The animal studies were carried out by the academic staff at the Department of Cancer Studies and Molecular Medicine (CSMM) at the University of Leicester. The animal studies included rice bran supplementation of mice, plasma collection, and assessment of intestinal neoplasia along with proteomic studies on mice plasma.

A mixed population of male and female $APC^{min}$ mice (n=35) were administered with rodent feed supplemented with rice bran at concentrations of 0%, 5%, 15% and 30% w/w. Based on the percentage of bran supplemented the mice were classified into groups starting with Group A for 0% rice bran supplemented (control group) to group D which had 30% rice bran supplemented in their diet respectively **Table 2-1**. These mice were administered with the supplemented rodent feed from the age of 4 weeks till the termination of the study at the end of 16 weeks. At the end of this period the mice were culled by cardiac exsanguination and whole blood was collected. Blood was processed into plasma and immediately stored at $-80^{o}C$. The gastrointestinal tracts were extracted and washed with phosphate buffered saline (PBS) to remove food content and then were opened longitudinally to reveal adenomas. These adenomas were counted and their sizes were measured using digital callipers to determine the total tumour burden.

**Table 2-1 Summary of demographic of the study**

| Group | Supplemented Bran (%) | Total (n) | Male | Female |
|---|---|---|---|---|
| **Group A** | 0 | 7 | 4 | 3 |
| **Group B** | 5 | 4 | 4 | 0 |
| **Group C** | 15 | 16 | 9 | 7 |
| **Group D** | 30 | 8 | 4 | 4 |
| ***Total*** | | ***35*** | *21* | *14* |

### 2.3.3 Plasma sample pre-treatment

Mouse plasma sample pre-treatment, method development and metabolomic profiling analysis by UHPLC-IM-MS and subsequent data processing was carried out at Centre for Analytical Science, Loughborough University by the author (Aditya Malkar). Mice plasma samples stored at -80°C were thawed at room temperature prior to sample clean-up by protein precipitation. Each thawed plasma sample was vortexed for 30 seconds followed by ultra-sonication for 1 minute to improve homogeneity. Protein precipitation was achieved by the addition of 200 µL acetonitrile to 100 µL of plasma i.e. 2:1 ratio of precipitant to plasma (Polson et al. 2003). The mixture was vortexed for 30 seconds followed by ultra-sonication for 1 minute and then subjected to centrifugation at 10,000 g for 10 minutes at ambient temperature. Precipitated proteins from the sample were discarded as a pellet at the bottom of the microcentrifuge tube. The supernatant was transferred to a fresh microcentrifuge tube and was subjected to concentration by evaporation using Turbovap LV concentration workstation (Caliper Life Sciences, Hopkinton, MA, USA). The supernatant was evaporated to near dryness prior to being reconstituted in 50 µL water/acetonitrile (95:05) (v/v) with 0.1% formic acid. The reconstituted plasma extract was placed in a 200 µL polypropylene micro-insert (Supelco, UK) in an autosampler vial (2 mL) for UPLC-MS analysis. An overview of the sample pre-treatment procedure is outlined in **Figure 2.3**.

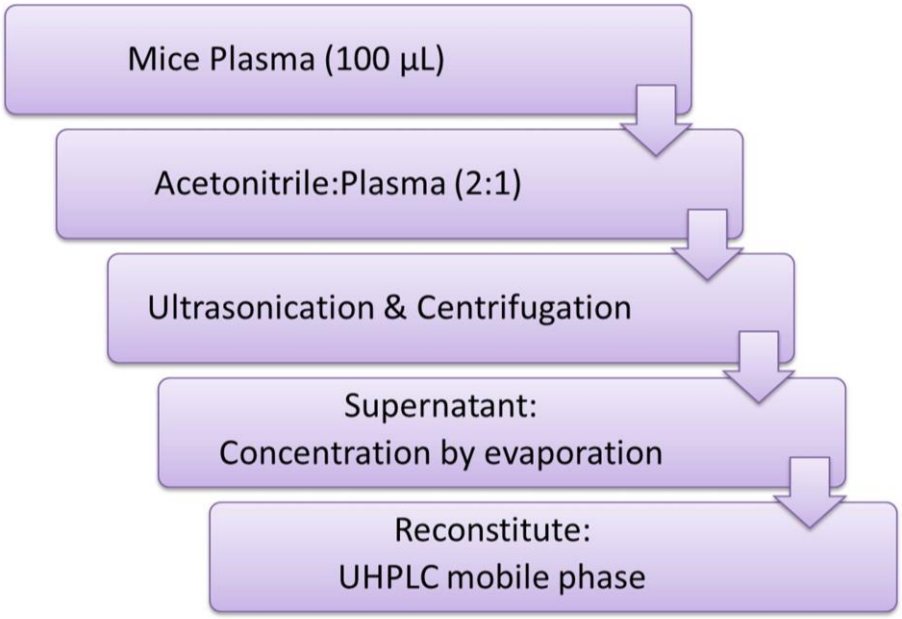



**Figure 2.3 Overview of sample pre-treatment for mice plasma**

### 2.3.4 Instrumentation

The liquid chromatographic separations for the metabolomic analysis of mice plasma were performed on a Waters AQUITY UPLC chromatograph with an in-built auto-sampler (Waters Corporation, Manchester, UK) fitted with a Waters AQUITY BEH (bridged ethylene hybrid) $C_{18}$ column (2.1 mm x 100 mm, 1.7 µm; Waters Corporation, Manchester, UK) maintained at 40°C. A VanGuard BEH $C_{18}$ pre-column (Waters Corporation, Manchester, UK) was attached prior to the analytical column. The UPLC system was coupled with a Waters Synapt HDMS ion-mobility mass spectrometer (Waters Corporation, Manchester, UK) equipped with an electrospray interface. MassLynx v4.1 (Waters Corporation, Manchester, UK) was used for controlling the setup and for data acquisition.

#### *2.3.4.1 UPLC conditions*

Mobile phase A consisted of 0.1% aqueous formic acid (v/v) and mobile phase B was 0.1% formic acid in acetonitrile (v/v) with the flow rate set to 0.5 ml/min. The mouse plasma extract (10 µL injected) was analysed by the chromatographic gradient: 5% B (0-1 min), increased to 35% B (1-10 min). This was followed by a column clean-up phase built into the method to reduce carry over and condition the column for analysis of

subsequent samples, in which the mobile phase was increased to 95% B (10-11 min) and maintained at 95% B (11-12 min) before returning to initial conditions (13-15 min).

#### *2.3.4.2 MS conditions*

Electrospray ionisation conditions for the MS, with the ion source operated in positive ion mode were: capillary voltage 3.0 kV; cone voltage 30 V; source temperature 120°C; desolvation temperature 300°C; desolvation gas, $N_2$ gas flow 600 L/hr; cone gas flow 30 L/hr. Data was acquired in centroid mode from 50 – 1200 Da with an inter-scan delay of 0.02 sec.

#### *2.3.4.3 Ion mobility conditions*

The tri-wave (TWIMS) drift cell conditions were set at 30 mL/min drift gas ($N_2$) with a variable travelling wave height ramp of 6.0 – 13.0 V and wave velocity of 300 m/s. The mass spectrum acquisition rate was 0.065 ms/scan.

### 2.3.5 Method development

The chromatographic gradient programme was optimised for the analysis of mice plasma samples using 0.01 mg/mL mixture solution of L-carnitine and hydrocortisone in acetonitrile/water (05:95%) with 0.1% formic acid. The TWIMS conditions were optimised online by infusing a 2 ng/mL solution of leucine enkephalin. The optimisation of ion mobility parameters is discussed in depth in chapter 4. The starting volume of mouse plasma used for sample pre-treatment was optimised by evaluating varying volumes in the range of 50 µL – 400 µL. Sample pre-treatment was carried out as outlined in Section 2.3.3 with an UHPLC injection volume of 10 µL. Based on the signal intensity and to avoid the overloading of the detector, 100 µL volume was chosen for analysis.

### 2.3.6 Quality control of the methodology

A quality control (QC) sample was prepared from equal-volume sub-aliquots of all the samples from all the sample groups of the study (Group A to D). A series of 5 consecutive extracts derived from the QC samples were analysed prior to the sample run for conditioning of the chromatographic column. Randomisation of the sample list was

carried out prior to analysis to remove bias. A pooled mice plasma extract was analysed after every 5 samples as a quality control sample (QC). Blanks were analysed after every sample to ensure no-carryover between samples runs.

### 2.3.7 Data processing and statistical analysis

Data mining was carried out using MarkerLynx XS (Waters Corporation, Manchester, UK). The parameters used for peak alignment were: retention time range 0-10 min, retention time tolerance was set to 0.1 min; mass range was limited to 50-500 Da and isotopic peaks were excluded from the analysis. The data matrix generated by the MarkerLynx software was normalised to the summed total ion intensity per chromatogram. The data were pareto scaled prior to subjecting the data matrix to multivariate statistical analysis. Multivariate statistical analysis including PCA and PLS-DA was carried out using SIMCA-P + (Umetrics, Sweden). Ion mobility data was visualised and processed using Driftscope v2.1 (Waters Corporation, Manchester, UK).

## 2.4 RESULTS AND DISCUSSION

### 2.4.1 Physiological data from the mice

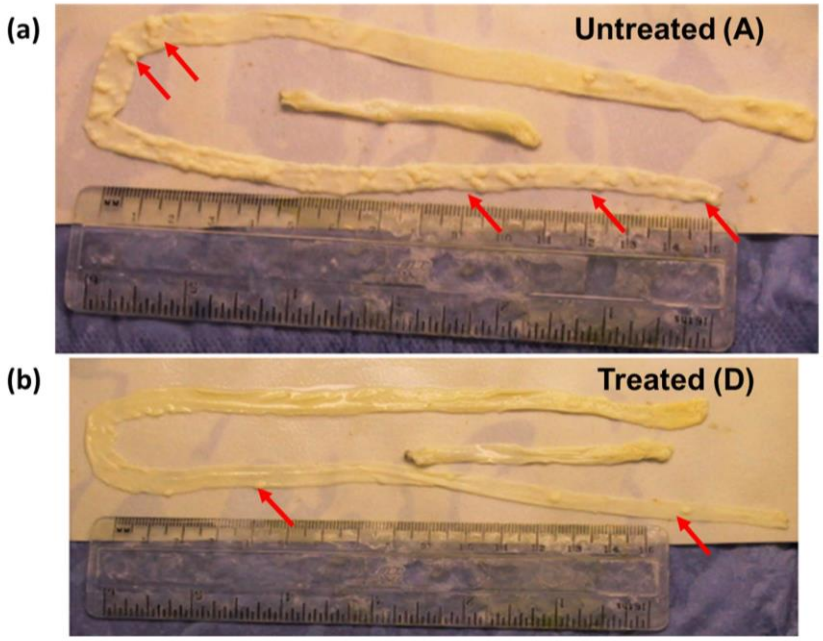



**Figure 2.4 Small intestine and colon of APCmin mice from Group A (0% bran supplemented control population) and Group D (30% bran supplemented population)**

An example of a small intestine and colon of an $APC^{min}$ mouse from Group A (control population) and another from Group D (30% bran supplemented) population is shown in **Figure 2.4**. A significant reduction in the number and size of the adenomas can be observed in Group D (**Figure 2.4 b**) when compared with Group A (**Figure 2.4 a**). The number and burden of adenomas was evaluated from the intestines and colons of mice from all the groups.

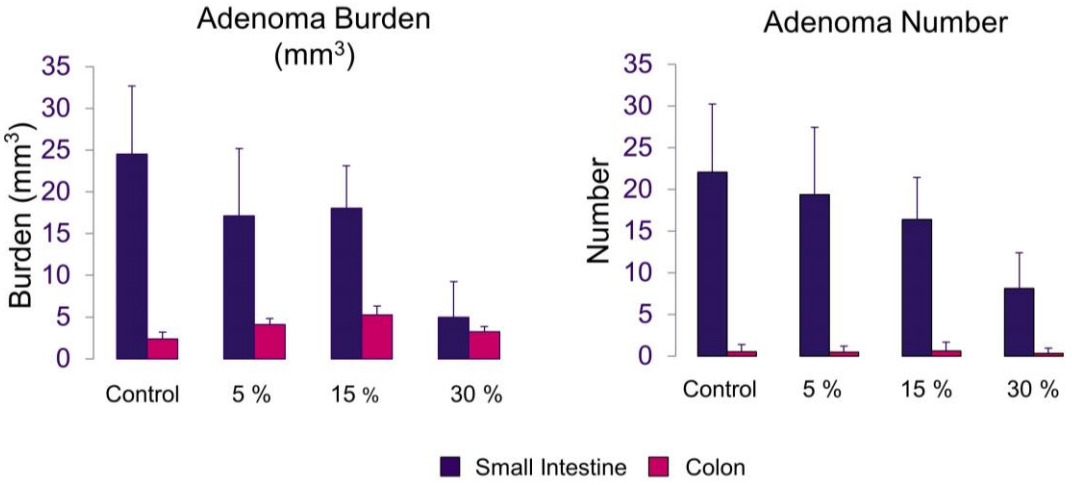



**Figure 2.5 Effect of lifetime dietary fibre administration on the burden and number of adenomas for APCmin mice**

**Figure 2.5** shows that there is a dose related gradual decrease in the number and burden of adenomas from 0% to 30% bran supplementation. The group D population shows statistically significant decrease both in number and burden of the adenomas compared with the group A population, supporting the evidence that fibre at 30% concentration in the diet interferes with adenoma development in the $APC^{min}$ mouse model. Hence, for subsequent multivariate analysis on the metabolomic data obtained, Group A and Group D were focussed upon.

### 2.4.2 Metabolite profiling of mice plasma by UHPLC-IM-MS

Protein precipitation was the only step carried out in preparation of plasma samples for metabolomic analysis by UHPLC-IM-MS. This was achieved by the addition of acetonitrile in the ratio 2:1 with plasma (Polson et al. 2003). The main aim of the sample pre-treatment performed was to extract small-molecular-mass compounds while removing large amount of proteins from plasma, which would otherwise interfere with the metabolomic analysis. Mass spectrometric data acquisition was carried out in positive ionisation mode (ESI+). A typical total ion current (TIC) obtained from the analysis of mice plasma extract is shown in **Figure 2.6** (a).

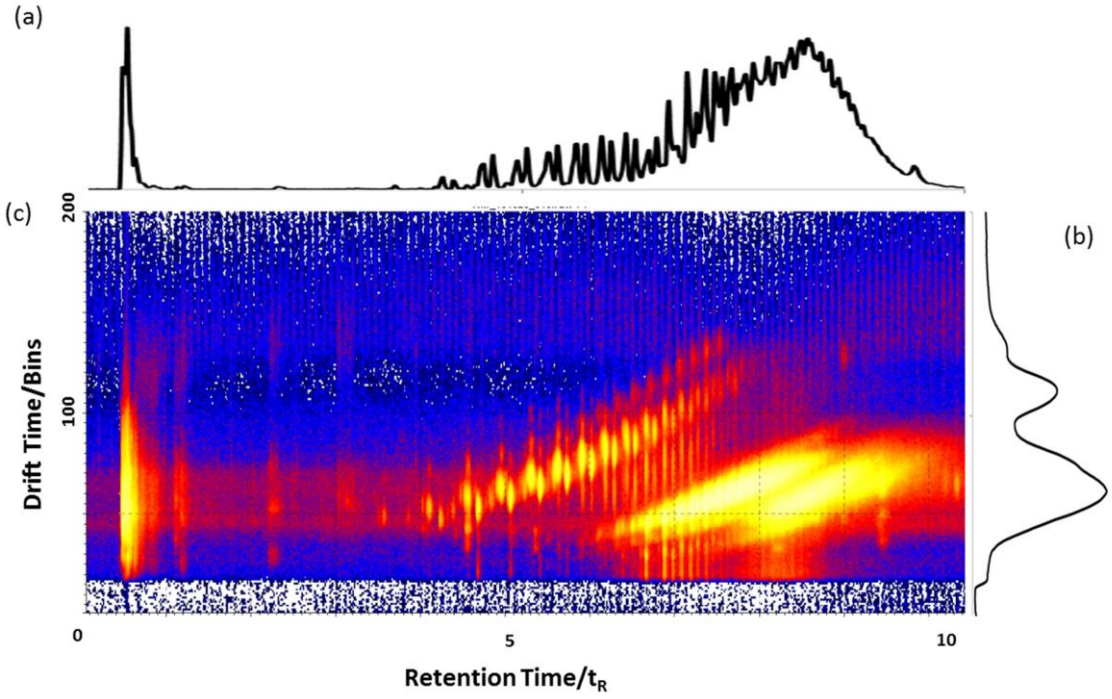



**Figure 2.6 Fig 1. UPLC-IM-MS metabonomic analysis of mice plasma extract. (a) UHPLC total ion chromatogram, (b) total ion mobility drift time plot summed over the entire chromatographic run, (c) retention time vs ion mobility drift time heatmap.** (Data produced on Synapt HDMS)

The UHPLC-IM-MS analysis of mice plasma generated a complex dataset consisting of mass-to-charge ratio, ion mobility drift time and chromatographic retention time. **Figure 2.6** shows an example profile of UHPLC-IM-MS analysis of mice plasma extract. The complexity of the three dimensional dataset obtained from the analysis within a short time frame of 10 minutes is illustrated (excluding the column clean-up and re-

equilibration phase). It is clear that the manual interrogation of the data and comparison of profiles obtained from different groups of samples would be tedious, time consuming and unfruitful. Hence, the need for using multivariate statistics to find discriminating metabolites between the sample groups is evident.

A three dimensional illustration of the structure of the data is depicted in **Figure 2.7**, generated by Waters DriftScope v2.1 (Waters Corporation, Manchester, UK). The orthogonal nature of the ion mobility domain of separation can be seen with respect to the chromatographic retention time ($t_R$). The two dimensional retention time vs ion mobility drift time plot (**Figure 2.6** c) shows examples of multiple species contributing to single chromatographic peak, which are separated by ion mobility. This demonstrates the enhanced separation of the complex plasma samples achieved by the use of ion mobility spectrometry within the same timeframe as a UHPLC-MS run.

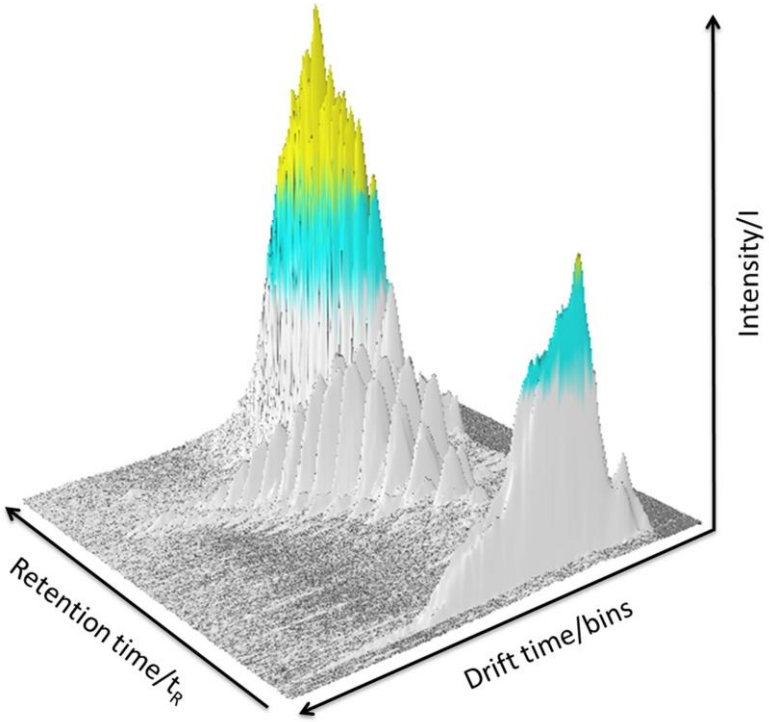



**Figure 2.7 UHPLC-IM-MS analysis of mice plasma extract, a 3-dimensional representation illustrating the orthogonal structure of the ion mobility domain (drift time) with respect to chromatographic retention time ($t_R$).** (Data produced on Synapt HDMS)

### 2.4.3 Improvement of mass spectral quality using ion mobility

Polyethylene glycols (PEG) are widely used in pharmaceutical and biological analyses for various purposes such as stabilisation of proteins for separation, storage and purification (Zhao & O'Connor 2007). However, they are also a common source of contamination in mass spectrometric systems, which cause significant challenges in the analysis of analytes of interest. Common sources of PEG are volumetric and storage equipment contaminated with PEG, laboratory glassware cleaning products such as Neutracon, Triton-X, Tween etc. and plastics utilised in analytical laboratory environments.

PEG, if present in samples, generates significant interference with mass spectrometric experiments causing loss of signal and by obscuring the signal from analytes of interest. PEG can be spotted in a mass spectrum as it shows a common repeat of 44 Da with a characteristic Gaussian profile (Fenn et al. 1989). Removal of PEG from samples containing small molecules is a difficult task by traditional methods of sample cleanup such as solid-phase extraction, dialysis, ion exchange catridges and titanium dioxide addition. Such methods of sample clean-up would also require additional steps of sample pre-treatment for metabolomic analysis and would affect the holistic extraction of metabolites from samples (Eckers et al. 2007; Castro-Perez et al. 2007).

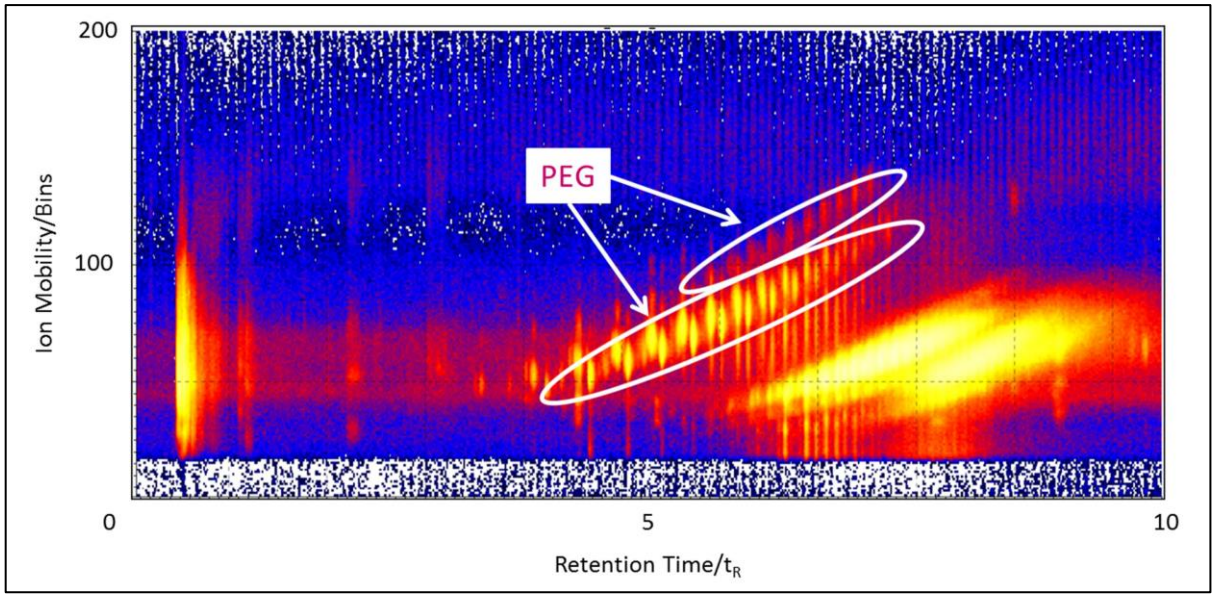



**Figure 2.8 Visualisation and selection of PEG series by ion mobility drift time demonstrating charge state separation of singly and doubly charged PEGs obscuring the metabolomic information by suppression of signal .**

The contamination of samples with PEGs was suspected in the mice plasma samples used in this study. A possible source of contamination was thought to be plastic micro-centrifuge tubes used for the storage and transport of mice plasma samples. The ion mobility heat map obtained from the UHPLC-IM-MS analysis of mice plasma extract facilitates easy detection and visualisation of the PEG contamination within the mice plasma extract samples as highlighted in **Figure 2.8**. The charged state separation of the PEG present in the mice plasma samples can be seen in the circled regions of **Figure 2.8**.

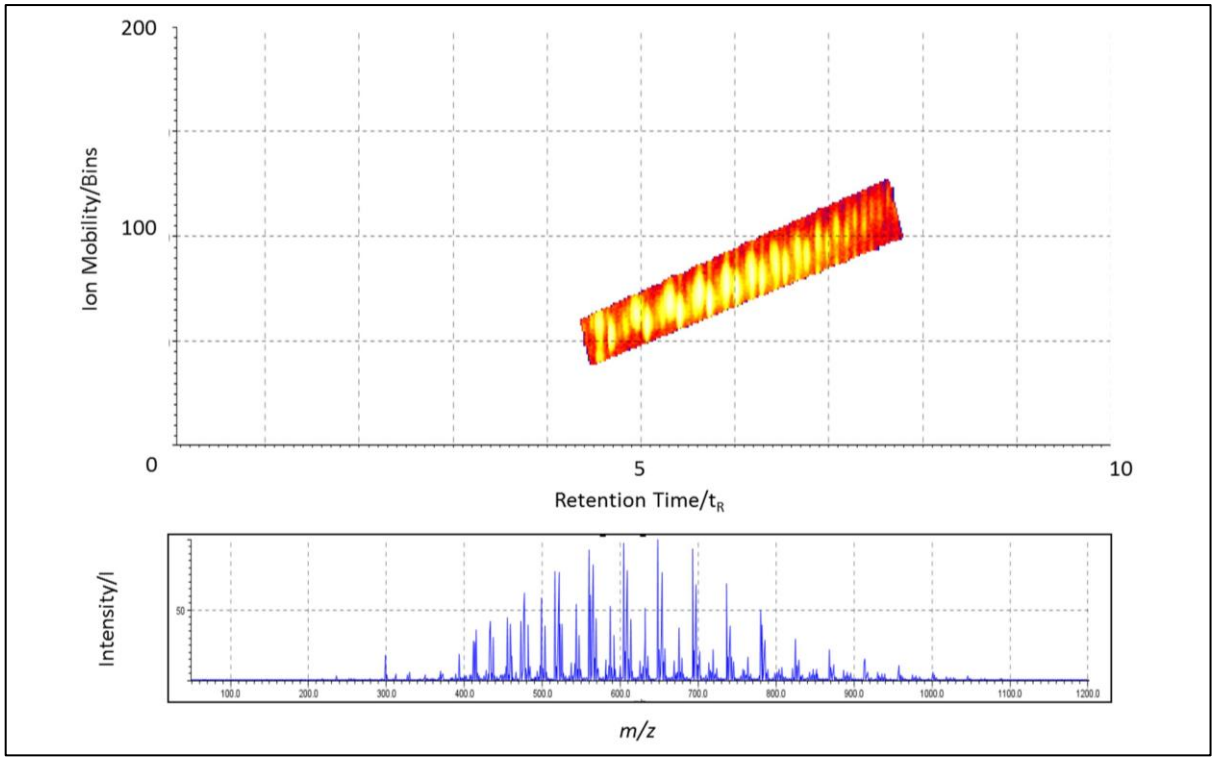



**Figure 2.9 UHPLC-IM-MS analysis of mice plasma extract showing post-analysis selective extraction of PEG using ion mobility drift time.** (Data produced on Synapt HDMS)

A specific region of the chromatographic retention time ($t_R$) vs ion mobility drift time (bins) can easily be selected from the generated heatmap and the mass spectra for that region extracted. **Figure 2.9** shows the selection of the region of the heatmap containing a series of PEG ions displaying extracted mass spectrum for the region. The PEG ion distribution extends from approximately *m/z* 300 – 1000 with high peak intensity in the region around *m/*z 600. Low level metabolite ions in this region would therefore be obscured by the PEG ions. The selected regions shown in **Figure 2.8** and **Figure 2.9** can be removed from the dataset using DriftScope software (Waters Corporation, Manchester, UK) and the remnant dataset can be exported back to MassLynx (Waters Corporation, Manchester, UK) for further data interrogation. The removal of PEG signal by using the ion mobility drift time information enables access to lower-level components of the sample by the removal of PEG signals which would otherwise obscure the components in the given mass range.

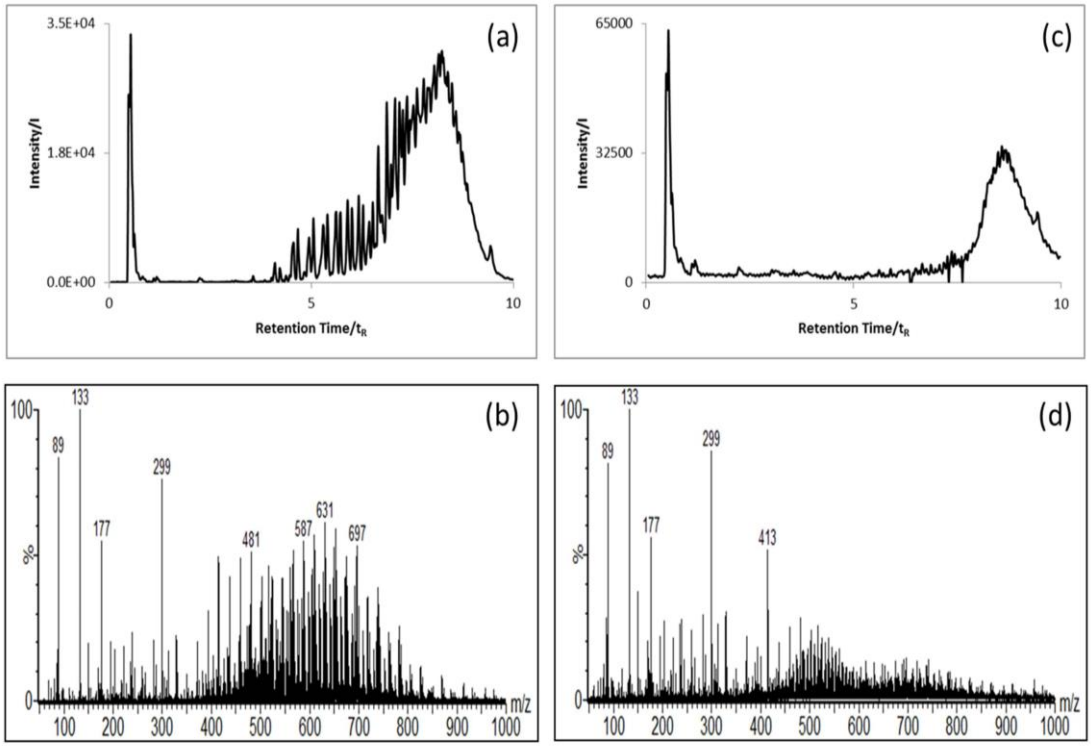



**Figure 2.10 Improvement of spectral quality by UHPLC-IM-MS: (a) UHPLC total ion current (TIC) for mice plasma extract without post analysis clean-up of PEG using drift time information, (b) mass spectrum associated with 0-10 min of the UHPLC chromatogram showing distribution of PEG series, (c) UHPLC total ion current (TIC) for mice plasma extract selectively cleaned up for PEG using drift time and (d) cleaned-up mass spectrum associated with 0-10 min of chromatogram using drift time to selectively remove PEG showing mass spectral improvement.** (Data produced on Synapt HDMS)

The improvement of the mass spectral quality by the use of ion mobility spectrometry can be seen from **Figure 2.10**. The total ion current (TIC) obtained by the UHPLC-IM-MS analysis of mice plasma extract without subjecting to post-analysis ion mobility clean-up of the chromatogram is shown in **Figure 2.10 (a)**, with its associated mass spectrum shown in **Figure 2.10 (b)**. The series of PEGs present in the sample were selectively removed by post-analysis processing of the UHPLC-IM-MS data as described above. **Figure 2.10 (c)** shows the PEG cleaned-up TIC of the UHPLC-IM-MS analysis of mice plasma extract with its associated mass spectrum shown in **Figure 2.10 (d)**. The mass spectral quality improvement can be seen with the reduction in the intensity of PEG series ions increasing the signal-to-noise for previously obscured components of the

mass spectrum in the region. Mass spectral clean-up with ion mobility by this method can prove to be a valuable tool for analysis of targeted metabolites from complex samples as well as for untargeted metabolite profiling experiments by improving the detection of smaller features from the UHPLC-IM-MS analysis.

### 2.4.4 Software limitations

Currently used software for mining the UHPLC-IM-MS metabolomic dataset e.g. MarkerLynx XS (Waters Corporation; used in the present study) were not equipped to deal with the ion mobility dimension of the dataset during multivariate statistical analysis step. As a result, the workaround used for this problem has been to collapse the ion mobility dimension during the first stage of data processing resulting in significant amount of potentially valuable unused data.

Attempts were made to generate an in-house spreadsheet to convert the ion mobility data into numerical values to be included in the data matrix. However, this yielded insignificant results in the preliminary test and it was deemed that extensive specialist programming will be required for accurately and efficiently converting and incorporating the ion mobility data into the data matrix.

A significant shortfall of the present data mining tools commercially available for metabolomic studies has been highlighted here.

### 2.4.5 Potential metabolomic biomarker ions of the chemopreventive efficacy of rice bran in colorectal cancer

The multidimensional dataset generated by the UHPLC-IM-MS analysis (**Figure 2.7**) presented a challenge for the data analysis software (MarkerLynx XS, Waters Corporation, Manchester, UK in this case) traditionally used for metabolite profiling studies, as the software are equipped to deal only with two dimensional datasets e.g. chromatographic retention time and mass-to-charge ratio (*m/z*). The commercially available software are not capable of integrating the third dimension of separation - ion mobility, while generating the data matrix required for multivariate statistical analysis. As a result, ion mobility dimension was collapsed to reduce the complexity of the data with the intention to 'open up' the ion mobility dimension post multivariate statistics.

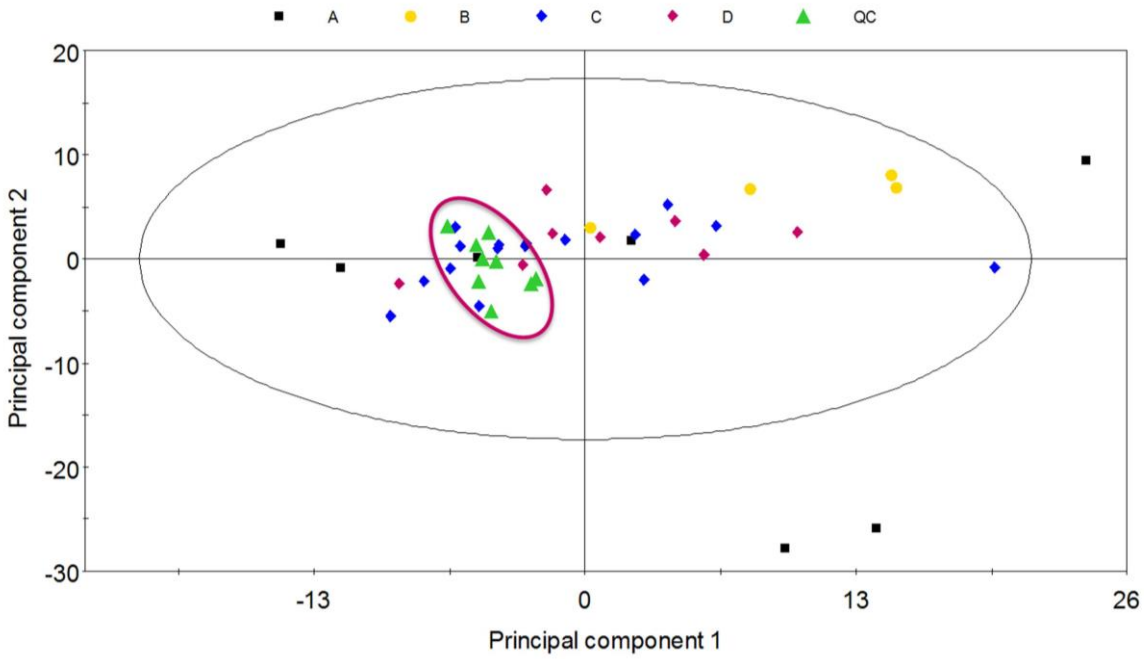



**Figure 2.11 Principal component analysis (PCA) of all mice plasma extracts (Group A, B, C, D and QC) samples. Component 1 (20.51%) and component 2 (13.46%). Model based on K=339 variables and N=44 observations.**

The quality of the data generated from chromatography-mass spectrometric techniques is of paramount importance in untargeted global metabolite profiling studies. It is important to establish that the data generated during metabolomic studies is of high quality if the studies are to yield valuable insights from a biological interpretation point of view. The data generated has to be acceptable in terms of chromatographic retention time

reproducibility and mass spectrometric sensitivity. As these studies aim to unravel small differences in the analysed samples, it is important to establish that the source of variation in the data is due to the biological question being asked and not due to instrumental variations or variable system performance.

Principal component analysis (PCA), the most commonly employed unsupervised pattern recognition technique, enables the reduction of the dimensionality while maintaining the maximum variability of the data. **Figure 2.11** shows the PCA plot of all the samples from the study (Group A – control to Group D – 30% bran fed mice population) and quality control (QC) samples. The QC samples were generated as described in Section 2.3.6, by pooling together aliquots from all the samples used in the analysis so as to provide a representative mean sample containing all the analytes which will be encountered during the analysis. The tight clustering of the QC samples towards the middle of the PCA plot (circled in **Figure 2.11**) provides evidence of reliable data quality, which is fit-for-purpose for metabolite profiling studies. The numbered QC samples showed no particular trend and were distributed in a random order within the cluster of samples indicating there was no significant drift in system performance which would cause a bias in the analysis. In case of a system performance drift, trailing of QC samples would have been observed in the PCA plot.

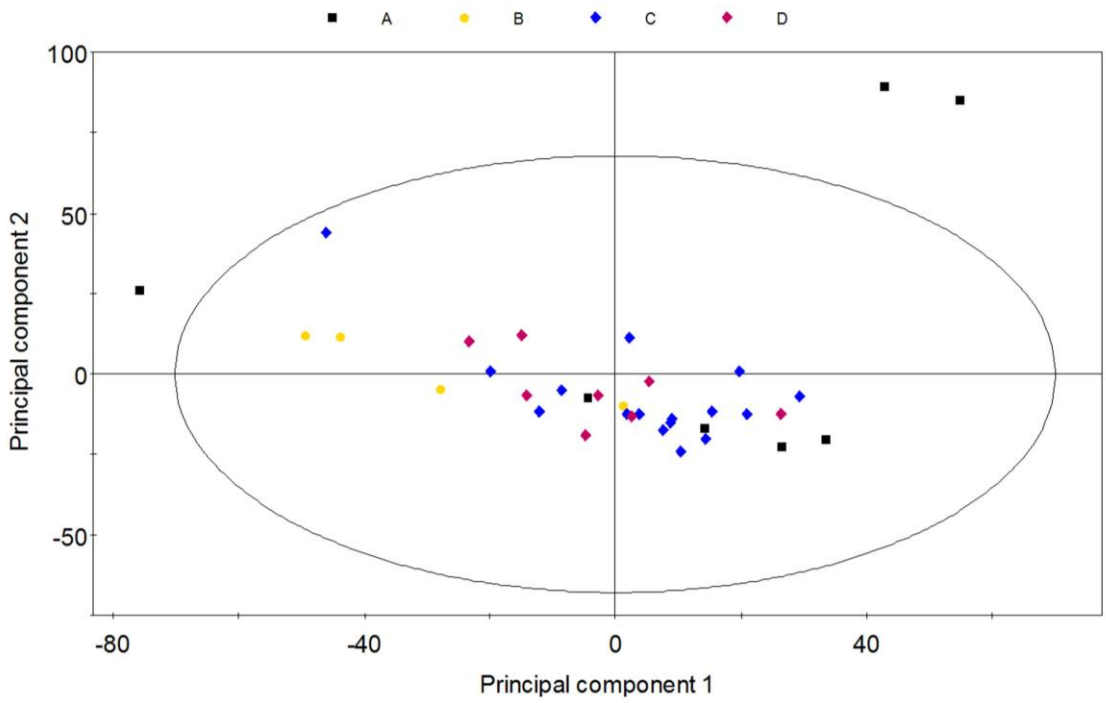



**Figure 2.12 Principal component analysis (PCA) of all mice plasma extracts (Group A, B, C and D) samples. Component 1 (17.56%) and component 2 (16.44%). Model based on K=332 variables and N=35 observations.**

The QC samples were removed from the model after the initial inspection to establish reliable data quality and the rest of the samples (Groups A to D) were analysed by the unsupervised PCA approach (**Figure 2.12**). However, these showed no significant separation of the groups of samples in a model based on K=332 variables. Thus, from the data described in Section 2.4.1, it was decided to focus on the two extreme groups from the study i.e. Group A – the 0% bran supplement fed population (control group) and Group D – the 30% bran fed population which had shown statistical evidence of reduction in number of adenomas.

The comparison of Group A and Group D by PCA showed partial separation with clustering of Group D samples towards the middle of the PCA plot (**Figure 2.13**).

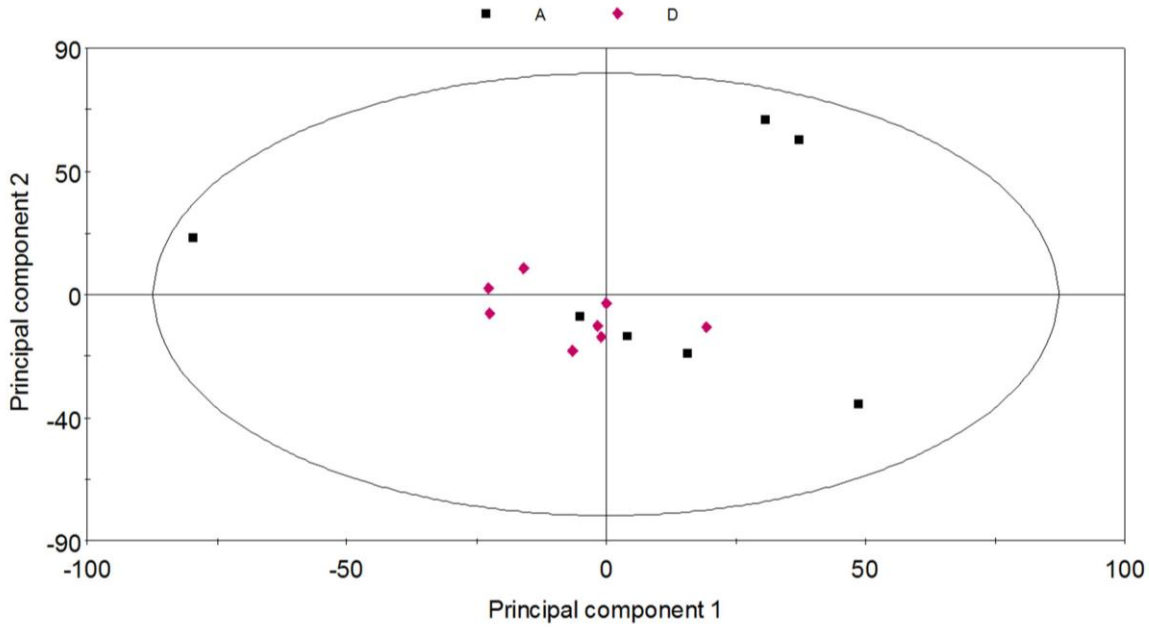



**Figure 2.13 Principal component analysis (PCA) of Group A and Group D samples. Component 1 (22.41%) and component 2 (19.18%). Model based on K=313 variables and N=15 observations.**

Supervised pattern recognition approaches such as partial least squares – discriminant analysis (PLS-DA) and orthogonal partial least squares – discriminant analysis (OPLS-DA) are often applied in metabolomic studies to tackle problems of subtle nature such as to uncover complex relationships buried in the background of larger multi-perplexed effects (Fonville et al. 2010).

The current study involved plasma taken from a well-established $APC^{Min}$ mice model of colon carcinogenesis. These mice were kept in a very controlled environment of a laboratory with the same age and maintained on the same diet, which would limit the factors responsible for metabolic variation. However, the population still had minor sources of variation such as gender, as the population was a mix of male and female mice, as well as individuality contributing to different rates of metabolism. As a result, a supervised OPLS-DA model was required to draw out the biological sources of variation in the data present in the cohort.

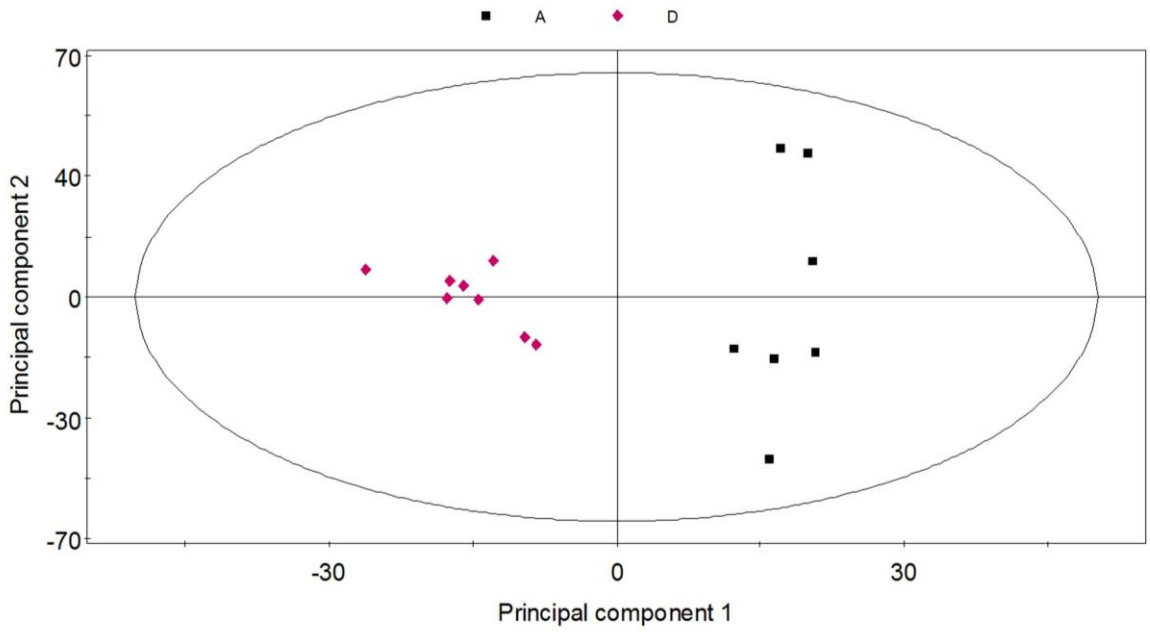



**Figure 2.14 OPLS-DA analysis of Group A and Group D samples. Model based on K=315 variables and N=15 observations. Component 1 (18.79%) and component 2 (7.76%).**

An OPLS-DA model based on K=315 variables produced a marked separation in the sample groups indicating that there were differences in the data between Group A and Group D populations. Group A i.e. the control population showed a wider within-sample-group variation as can be seen from the spread of Group A samples across the principal component 2 in comparison to Group D samples which are restricted to a smaller range. Group D (30% bran supplemented) population shows tighter clustering as can be seen from **Figure 2.14**.

Group A (N=7) consisted of 4 male and 3 female mice whereas Group D (N=8) contained 4 male and 4 female mice. The OPLS-DA plot does not show sub-clustering of the samples from either Group A or Group D, indicating that the effect due to bran supplementation is non-gender biased and affects male and female mice in the same respect.

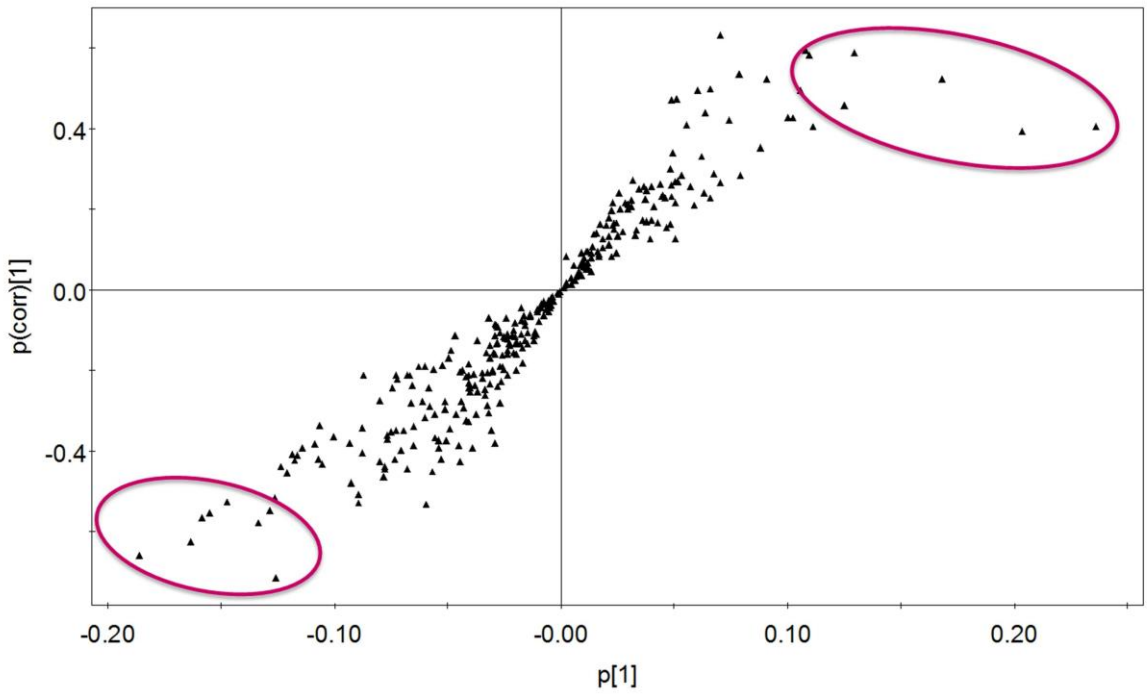



**Figure 2.15 S-plot generated from PLS-DA model showing distribution of the potential biomarker ions.**

A loadings S-plot was constructed using the OPLS-DA model shown in **Figure 2.14**, to segregate potential biomarker ions of efficacy. The S-plot is a statistical tool for visualisation of the distribution of the data from an OPLS-DA model of two classes. It is a popular method to filter out putative biomarker ions from the dataset as it is in essence a plot of the contribution of the variable towards class separation on the x-axis (modelled covariation i.e. p1) and the confidence or reliability of the variables on the y-axis (modelled correlation i.e. p(corr)1).

The S-plot obtained from the OPLS-DA analysis of Group A and Group D samples is shown in **Figure 2.15**. The variables from Group D are shown in the positive first quadrant whereas those from Group A are shown in the third (negative) quadrant of the plot. A number of discriminant metabolite ions from either ends of the S-plot were selected and raw data was interrogated by analysing the extracted ion chromatograms (EIC) for the obtained individual ions. The chromatographic peak shapes of the EICs were analysed to ensure that only the ions, which produced valid chromatographic peaks with good peak symmetry and intensity, were retained for further processing and to narrow down the list of potential biomarker ions. Emphasis was given to scrutinising the raw data to verify the authenticity of the potential biomarker ions. The variables obtained

from either ends of the s-plot (highlighted in **Figure 2.15**) were systematically tested for their trend-plots and chromatographic peak shapes to distil and reduce the number of potential biomarker ions.

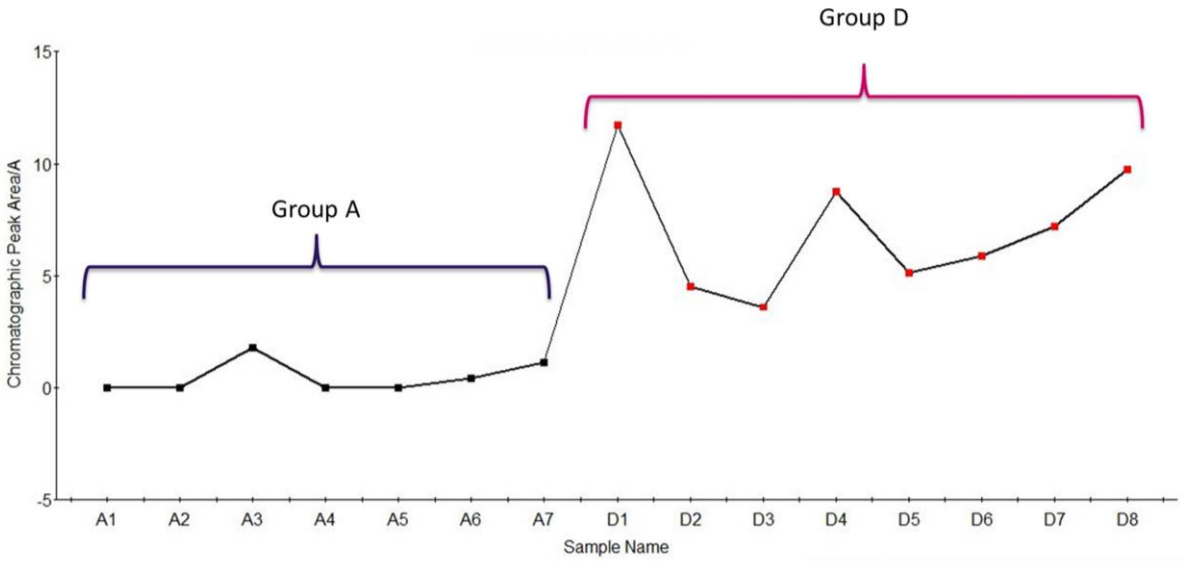



**Figure 2.16 Trend plot showing relative intensities of the potential biomarker ion 212.95±0.02 across samples from Group A and Group D.**

**Figure 2.16** shows an example trend plot of one such discriminant metabolite ion obtained after careful scrutinising of the variables from the s-plot. A trend plot facilitates the easy visualisation and comparison of the normalised chromatographic peak areas of the selected variable's distribution across the samples from the study. The **Figure 2.16** shows a plot of all the samples from Group A and Group D vs the normalised chromatographic peak areas for the variable (potential biomarker ion) with *m/z* 212.95±0.02. The figure shows a clear distinction between the individual intensities for the ion between the two sample groups. The normalised chromatographic peak areas for the variable are lower in Group A as compared to Group D and hence it can be inferred that the ion is up-regulated in Group D samples due to bran supplementation.

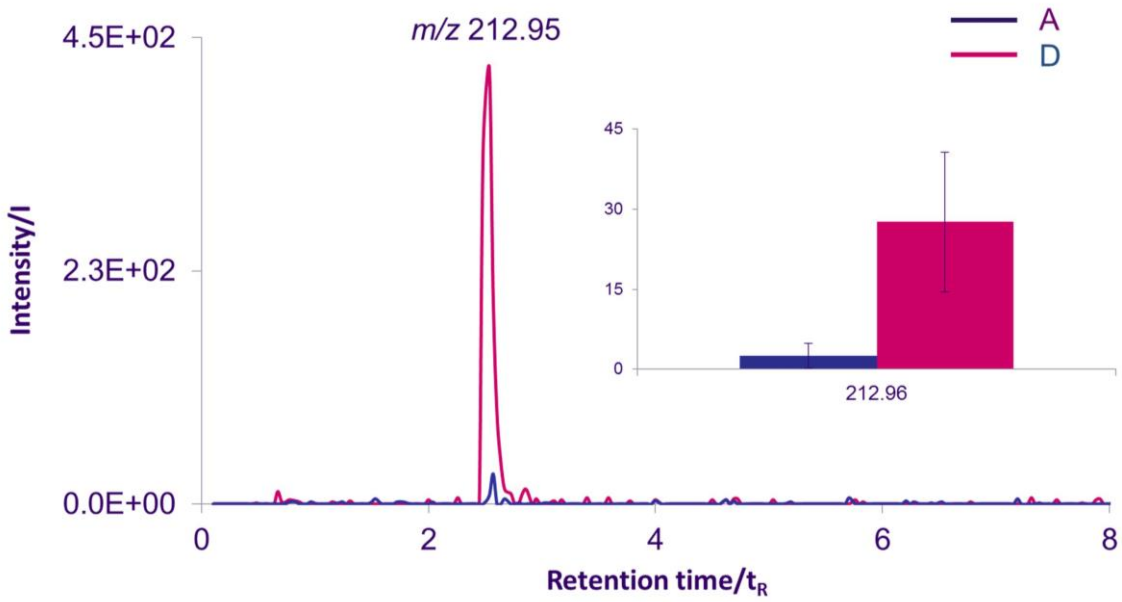



**Figure 2.17 Example extracted ion chromatogram (EIC) of potential biomarker ion 212.95±0.02 showing up-regulation.** (Data produced on Synapt HDMS)

The raw data for the potential biomarker ion was inspected by using the extracted ion chromatogram for *m/z* 212.95±0.02. The Figure 2.17 shows comparison of example EIC from an example of a Group A sample and a Group D sample, providing evidence of up-regulation of the ion in the latter Group of samples. The change in the chromatographic peak areas was statistically evaluated for all the samples from the study pertaining to Group A and Group D. A clear statistical difference was observed from the histogram with error bars (inserted image in **Figure 2.17**) thus confirming the variable as a potential biomarker ion of efficacy.

Other ions were similarly interrogated from the s-plot using their EICs and are summarised in **Table 2-2**.

**Table 2-2 Potential biomarker ions of efficacy obtained from OPLS-DA comparison of Group A and Group D mice plasma extract samples**

| No. | Retention Time (min) | *m/z* ±0.02 | Regulation |
|---|---|---|---|
| **1** | 2.52 | 148.99 | Up-regulated |
| **2** | 7.71 | 189.02 | Up-regulated |
| **3** | 2.54 | 212.95 | Up-regulated |
| **4** | 3.19 | 218.01 | Up-regulated |
| **5** | 2.54 | 218.97 | Up-regulated |
| **6** | 0.92 | 160.07 | Down-regulated |
| **7** | 4.48 | 226.97 | Down-regulated |

The identification of the potential metabolomic biomarker ions listed in the table was beyond the scope of this research. Further work is required to identify the discriminant metabolite ions obtained from this study. A logical approach is to carry out MS/MS analysis on the individual ions obtained as an outcome of this work to understand the structure of individual molecules and to obtain a list of possible matches. Authentic standards of the shortlist of these candidate molecules could then be compared with the discriminant metabolites by LC-MS and LC-MS/MS to ascertain their identities.

The biological significance of the discriminant metabolites obtained cannot be evaluated at this stage without the knowledge of the identities of the ions.

## 2.5 CONCLUSIONS

The work presented here is the first application of combination of ion mobility (IM) with UHPLC-MS for metabolic profiling analysis of mouse plasma. This chapter also presents the first application of the developed methodology for the metabolomic analysis of chemopreventive efficacy of rice bran in colorectal cancer. The use of ion mobility spectrometry for post-analysis clean-up of contaminants such as PEG is also demonstrated.

The combination of high resolution UHPLC, ion mobility, mass spectrometry and powerful bioinformatics tools allows metabolomic biomarkers of efficacy to be identified from complex bio-fluids such as plasma derived from pre-clinical animal studies. This approach may prove useful for other pre-clinical studies to identify new biomarkers of efficacy. The utility of untargeted global metabolite profiling analysis by rapid UHPLC-IM-MS analysis and bioinformatics modelling for such studies has been demonstrated. The presence of PEG or other contamination in mass spectrometric samples is a difficult problem to deal with especially with respect to untargeted metabolic profiling studies. This study has preliminarily identified a panel of potential biomarker ions of efficacy following administration of rice bran extract with the view to assessing, validating and ultimately extending to a human patient intervention study. This work demonstrates that inclusion of ion mobility separation in the UHPLC-MS analysis allows clean-up of samples post-UHPLC-IM-MS analysis using ion mobility drift time information.

The existing established model of colon carcinogenesis based on $APC^{Min/+}$ mice has provided valuable insight and preliminary results over the years on human colon cancer (Leclerc et al. 2004). The next step would be to carry out the study with a larger cohort of samples utilising the potential biomarker ions obtained from bioinformatics models generated in the current study and using a targeted instead of an untargeted profiling approach. However, for more advanced and significant research, sampling needs to be carried out on human volunteers with colorectal cancers.

Plasma as a bio-fluid has been used in clinical and pre-clinical research. In recent times, the number of studies utilising plasma as a bio-fluid for the detection of proteomic and metabolomic biomarkers of diseases has seen a significant increase (Z.-Y. Li et al. 2010;

Pereira et al. 2009). However, the collection of plasma from human volunteers requires trained individuals to carry out sampling and sample handling. Plasma, being an invasively collected bio-fluid, also induces participant/patient discomfort and thus makes sampling difficult especially in participants with trypanophobia i.e. needle phobia. As a result, there is a growing trend in utilisation of non-invasively obtained bio-fluids such as saliva and urine (Raju et al. 2013).

The complete elimination of contamination from samples would be desirable. As a result, stringent protocols for collection and storage of bio-fluid samples such as plasma and saliva need to be established. The contamination of bio-fluid samples with plasticisers, PEG and other contaminants can be controlled by careful selection of materials and containers used for handling of metabolomic samples during storage and pre-treatment. This has been explored in detail in Chapter 3 and Chapter 4.

## 2.6 CHAPTER TWO REFERENCES


Burkitt, D.P. *Dis. Colon Rectum* **1993**, 36, 1071–1082.

Cancer Research UK, Colorectal cancer. **2011**, URL: http://www.cancerresearchuk.org/cancer-info/cancerstats/world/colorectal-cancer-world/ (accessed 23.5.13).

Castro-Perez, J.; Yu, K.; Shockcor, J. *Nat. Methods Appl. Notes*. **2007**.

Cooke, D.; Schwarz, M.; Boocock, D.; Winterhalter, P.; Steward, W.P.; Gescher, A.J.; Marczylo, T.H. *Int. J. Cancer* **2006**, 119, 2213–2220.

Eckers, C.; Laures, A.M.; Giles, K.; Major, H.; Pringle, S. *Rapid Commun. Mass Spectrom.* **2007**, 21, 1255–1263.

Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Science* **1989**, 80. 246, 64–71.

Fonville, J.M.; Richards, S.E.; Barton, R.H.; Boulange, C.L.; Ebbels, T.M.D.; Nicholson, J.K.; Holmes, E.; Dumas, M.E. *J. Chemom*. **2010**, 24, 636–649.

Haggar, F.A.; Boushey, R.P. *Clin. Colon Rectal Surg*. **2009,** 22, 191–197.

Leclerc, D.; Deng, L.; Trasler, J.; Rozen, R. *J. Cell. Biochem*. **2004**, 93, 1242–1254.

Li, Z.Y.; Zheng, X.Y.; Gao, X.X.; Zhou, Y.Z.; Sun, H.F.; Zhang, L.Z.; Guo, X.Q.; Du, G.H.; Qin, X.M. *Rapid Commun. Mass Spectrom*. **2010**, 24, 3539–3546.

Moser, A.M.Y.R.; Pitot, H.C.; Dove, W.F.; Clo, D. *Science* **1989**, 80. 247, 322–324.

Obayashi, Y.; Isobe, T.; Subramanian, A.; Suzuki, S.; Tanabe, S. *Environ. Res.* **2009**, 227–238.

Obrador, A. *Br. J. Nutr.* **2007**, 96, S46–S48.

Pereira, H.; Martin, J.F.; Joly, C.; Sébédio, J.L.; Pujos-Guillot, E. *Metabolomics* **2009**, 6, 207–218.

Peters, U.; Sinha, R.; Chatterjee, N.; Subar, A.F.; Ziegler, R.G.; Kulldorff, M.; Bresalier, R.; Weissfeld, J.L.; Flood, A.; Schatzkin, A.; Hayes, R.B. *Lancet* **2003**, 361, 1491–1495.

Polson, C.; Sarkar, P.; Incledon, B.; Raguvaran, V.; Grant, R. *J. Chromatogr. B.* Analyt. Technol. Biomed. Life Sci. **2003**, 785, 263–275.

Raju, K.S.R.; Taneja, I.; Singh, S.P.; Wahajuddin. *Biomed. Chromatogr*. **2013**, 27: 1354–1366

Sengupta, S.; Tjandra, J.J.; Gibson, P.R. *Dis. Colon Rectum* **2001**, 44, 1016–1033.

Verschoyle, R.D.; Greaves, P.; Cai, H.; Edwards, R.E.; Steward, W.P.; Gescher, A J. *Br. J. Cancer* **2007**, 96, 248–254.

WCRF/AICR, Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective. Washington DC, **2007.**

Williams, M.D.; Reeves, R.; Resar, L.S.; Hill, H.H. *Anal. Bioanal. Chem*. **2013,** 405, 5013–5030.

Zhao, C.; O'Connor, P.B. *Anal. Biochem.* **2007,** 365, 283–285.

# CHAPTER THREE:

# UNTARGETED METABOLITE PROFILING OF SALIVA BY ULTRA HIGH PERFORMANCE LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY (UHPLC-MS) FOR DISCOVERY OF POTENTIAL BIOMARKERS OF ASTHMA.

## 3.1 INTRODUCTION

Asthma is one of the most common chronic inflammatory diseases of the respiratory tract. About 300 million people suffer from asthma worldwide. According to 2004 data an estimated 32 million people in Europe are affected by asthma out of which 6 million suffer from severe symptoms. In the UK, around 5.4 million people are currently receiving treatment for asthma which includes around 1.1 million children (1 in 11) and 4.3 million adults (1 in 12). There were 1143 deaths in the UK due to asthma in the year 2010. An estimated 75% of hospital admissions for asthma are avoidable and as many as 90% of the deaths from asthma are preventable (Asthma UK 2013). Thus, there is the need to improve the understanding of asthma as well as to develop solutions for early detection of the disease.

Asthma is a chronic respiratory condition in which a person's airways become narrow making it difficult to breathe. Currently there is no cure for this disease. Asthma is typically characterised by coughing, wheezing or shortness of breath. When allergens or environmental factors cause a spike in these symptoms, it is referred to as an Asthma attack. During an attack, large airways called bronchi react to a trigger like an allergen with contracting spasms. The bronchi inflame and produce mucous further narrowing the airways and lead to the symptoms of asthma. Asthma attacks may last a few minutes or linger for a few days. Symptoms can usually be relieved using asthma medication. It is believed that there is a strong genetic component to predisposition to asthma (Mackay & Rosen 2001; Holgate 2008).

Asthma is a complex syndrome, which is characterised by variable degree of airflow obstruction, bronchial hyper-responsiveness and airway inflammation (Anderson 2008). It is chronic in nature and both genetic (atopy) and environmental factors (Holgate 2008) contribute towards its onset and manifestation. Asthma is a heterogeneous disease with common inflammatory features, but with several different phenotypes. The heterogeneity of asthma is not only limited to the symptoms, but it also relates to different response to therapies (Holgate 2008). Theories on mechanism of asthma and its immunohistopathology have been reviewed (Mackay & Rosen 2001).

The diagnosis of asthma in primary care is based on the presence or absence of airflow limitation, and its reversibility to inhaled bronchodilators, as measured by peak expiratory flow or spirometry. According to the guidelines, using spirometric values or peak flow as the prime outcome of interest is of limited value in patients with normal or near normal pre-treatment lung function since there is little room for measurable improvement (British Thoracic Society & Scottish Intercollegiate Guidelines Network 2012); furthermore it is estimated that 30% of patients with a diagnosis of asthma have no evidence of the disease (Shaw et al. 2012; Boulet 2003; Marklund et al. 1999). Biomarkers of airway inflammation and airway hyperresponsiveness obtained from various invasive and non-invasive physiological sources such as sputum, blood, serum, urine, and exhaled breath have been used to diagnose of asthma (Wadsworth et al. 2011). These tests differ in their sensitivity, specificity and patient acceptance. More importantly they cannot be used below certain lung function thresholds or ages, limiting their application and widespread adoption. Consequently there is a need for simple, non-invasive procedure and biomarker for the accurate identification of asthma (Szefler et al. 2012).

Historically, the diagnosis of asthma was largely based on pulmonary function tests, spirometry and tissue biopsies. In recent times, biomarkers obtained from various invasive and non-invasive physiological sources such as sputum, blood, serum, urine, exhaled breath condensate etc. have been preferred for accurate diagnosis of various subtypes of asthma (Wadsworth et al. 2011). Diagnosis of asthma based on tissue biopsies, induced sputum tests, serum protein tests, methacholine challenge tests etc. are either too invasive and involve patient discomfort or require high level of expertise for sample collection and analysis. Consequently there is a need for simple, non-invasive procedures that yield biomarkers for the clinical diagnosis of asthma (Szefler et al. 2012).

Metabolic profiling is the systematic study of the chemical fingerprints of low molecular mass endogenous metabolites in biological samples. Changes in metabolite concentrations reflect the biochemical effects produced in an organism as a function of disease state or therapeutic intervention (Nicholson & Lindon 2008). As a result metabolic biomarkers of diseases can aid early detection of diseases and provide better understanding of both disease occurrence and progression. The application of metabolic

profiling to the discovery of biomarkers of asthma has been explored in various biofluids and is reviewed elsewhere (Adamko et al. 2012; Szefler et al. 2012).

Previously carried out studies include the metabolomic analysis of asthma in urine by NMR (Saude et al. 2011) and childhood asthma by LC-MS (Mattarucchi et al. 2012). 2-D gel electrophoresis has been applied for the identification of novel biomarkers of asthma and chronic obstructive pulmonary disorder (COPD) from plasma (Verrills et al. 2011). Exhaled breath condensate has also been analysed by NMR to prospect for biomarkers of childhood asthma (Carraro et al. 2007). However, the use of saliva for the metabolomic analysis of asthma remains largely underexplored with rare applications such as the analysis of salivary cytokines and pathogens by PCR and ELISA (Wexler et al. 2009).

The work described here is the first instance of the use of saliva and metabolite profiling by UHPLC-MS as a potential tool for asthma diagnosis. In the present study, non-targeted metabolomic analysis of saliva samples from asthmatic and healthy control individuals was carried out by ultra-high performance liquid chromatography-mass spectrometry (UHPLC-MS). Data processing was performed using a commercially available software package for the discovery of potential metabolomic biomarker ions of asthma by comparing moderate asthmatic and healthy control saliva samples. Predictive models based on partial least squares-discriminant analysis (PLS-DA) were obtained from the analysis and employed for classification of saliva samples from mild/early stage asthmatic individuals.

## 3.2 CHAPTER THREE AIMS AND OBJECTIVES

The purpose of the work presented in this chapter was to explore the potential of human saliva as a non-invasive medium for metabolomic study of asthma. The specific objectives were:

- To investigate the potential of saliva as a biofluid for metabolomic study by UHPLC-MS.
- To carry out a metabolomic investigation to distinguish moderate asthmatic population from healthy control population via the analysis of metabolic profiles obtained by UHPLC-MS.
- To construct predictive models from moderate asthmatic and healthy control population and to test these models on mild asthmatic population.

## 3.3 EXPERIMENTAL

### 3.3.1 Chemicals

Acetonitrile (analytical grade), water (analytical grade) and formic acid (>99%) were purchased from Fisher Scientific (Loughborough, UK). L-carnitine (>99%) and hydrocortisone (>99%) were obtained from Sigma-Aldrich (Gillingham, UK). A standard mixture of 0.005 mg/mL L-carnitine and hydrocortisone was prepared in water:acetonitrile (95:05) (v/v) with 0.1% formic acid. A ToF Reference Mass Solution kit (Agilent Technologies, Santa Clara, CA, USA) consisting of purine (*m/z* 121.0509) and Hexakis(1H, 1H, 3H-tetrafluoropropoxy)phosphazine (*m/z* 922.0098) was used to prepare a reference mass solution for acquiring accurate mass data.

### 3.3.2 Demographic for the metabolomic study

Fifty-seven volunteers (male and female), aged between 21 and 76 participated in the study. The participants included 9 moderate asthmatic individuals (3 male and 6 female), 27 mild asthmatic individuals (14 male and 13 female) and 21 healthy individuals (7 male and 14 female) (summarised in **Figure 3.1**). The classification of moderate and mild asthmatics was based on established diagnostic tests for asthma viz. forced expiratory volume ($FEV_1$%PRED), sputum eosinophil count (EOS) and methacholine challenge (PC20). These tests are explained in detail in Section 3.3.3. The participant information, including diagnostic data and medication regimes are outlined in **Figure 3.1** and **Table 3-1**.

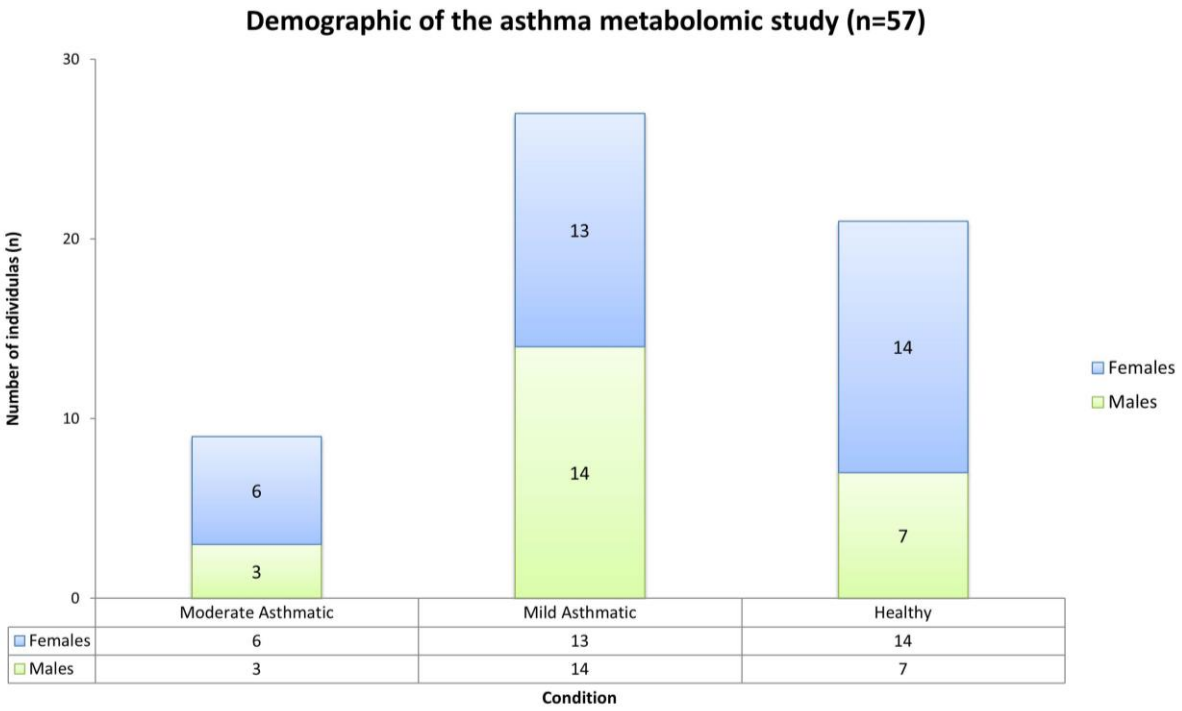



**Figure 3.1 Overview of demographic of the asthma metabolomic study.**

### 3.3.3 Overview of the tests for diagnosis of asthma

The population used in the study was well characterised for asthma based on a range of diagnostic tests (Appendix 7.6) and clinical diagnosis carried out at the Respiratory Biomedical Research Unit (BRU) at the Nottingham City Hospital (Nottingham, UK). The sub-classification of asthmatics into mild asthmatics and moderate asthmatics was carried out based primarily on three tests as described below.

#### *3.3.3.1 $FEV_1$%PRED*

Forced expiratory volume in 1 second ($FEV_1$) is a measure of lung function. It is the volume exhaled during the first second of a forced expiratory manoeuvre started from a level of total lung capacity. Total lung capacity (TLC) being the maximum volume to which the lungs can be expanded with the greatest inspiratory effort. $FEV_1$ is the most commonly used index for assessing airway obstruction, bronchoconstriction or bronchodilation. Vital capacity (VC) is the maximum amount of air that a person can expel from the lungs after first filling the lungs to the maximum extent. $FEV_1$ expressed as a percentage of vital capacity i.e. $FEV_1$%VC is the standard index for assessing and quantifying airflow limitation. Forced vital capacity (FVC) is the volume of air that can

forcibly be blown out after full inspiration, measured in litres. FVC is the most basic manoeuvre in spirometry tests.

$FEV_1$/FVC ($FEV_1$%) is the ratio of $FEV_1$ to FVC. In healthy adults this should be approximately 75–80%. In obstructive diseases (asthma, COPD, chronic bronchitis, emphysema) $FEV_1$ is diminished because of increased airway resistance to expiratory flow; the FVC may be decreased as well, due to the premature closure of airway in expiration, just not in the same proportion as $FEV_1$ (for instance, both $FEV_1$ and FVC are reduced, but the former is more affected because of the increased airway resistance). This generates a reduced value (<80%, often ~45%). In restrictive diseases (such as pulmonary fibrosis) the $FEV_1$ and FVC are both reduced proportionally and the value may be normal or even increased as a result of decreased lung compliance.

A derived value of $FEV_1$% is $FEV_1$% predicted, which is defined as $FEV_1$% of the patient divided by the average $FEV_1$% in the population for any person of similar age, sex and body composition. If the $FEV_1$%Pred value is >80%, the person in less likely to have asthma. If the value is less than 80% (<80%) then the person is more likely to have asthma. This is one of the three criteria used for segregation of asthmatic population into moderate and mild classes for this study.

#### *3.3.3.2 Blood Eosinophil Level (EOS)*

The blood eosinophil level (EOS) is a blood test which measures eosinophil or white blood cells count. Eosinophils become active in presence of allergic diseases, infections and other such medical conditions. Most common causes of increase in eosinophil count are allergic reaction of parasitosis (an infection caused or transmitted by a parasite). If the blood eosinophil level is <3% the person is less likely to have asthma and if the level is >3% the likelihood/severity of asthma is more.

#### *3.3.3.3 Methacholine challenge (PC20)*

Methacholine challenge is a test for hyper-reactivity of the smooth muscles of the lungs, often used to determine the absence or presence and severity of asthma. Lung function drops as a result of inhaling the chemical methacholine, which constricts the muscles around the airways in the lungs. The participants are made to inhale nebulised methacholine to determine the responsiveness of the airways of the participant to determine lung function. In the outcomes of this test, false-positive results are common and often caused by cigarette smoking, viral infections and presence of allergies. False-negative results are rare if the test is performed correctly, so a negative test can be used to reliably rule out the presence of asthma. If a participant can tolerate >8mg/ml of nebulised methacholine concentration, he/she is deemed less likely to have asthma; on the other hand a person responding to <8mg/ml of methacholine is more likely to have asthma.

**Table 3-1 Patient diagnostic and medication data for moderate asthmatics and healthy control population**

| | Condition | Gender | Age | Height (cm) | Weight (kg) | FEV1%PRED | EOS | PC20 (mg/ml) | BDP Equiv (µg) | V1 BDP Daily | Daily Medication Routine (1) | Daily Medication Routine (2) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | Moderate Asthmatic | Female | 60 | 162 | 58 | 59 | 10.75 | 0.03 | 1000 | 1000 | Seretide 125 1BD | Serevent 25 1BD |
| **2** | Moderate Asthmatic | Female | 66 | 158 | 62.3 | 67 | N/A | 2.3 | 400 | 400 | Symbicort 100/6 1BD | |
| **3** | Moderate Asthmatic | Male | 60 | 174.5 | 90.5 | 67 | 6 | 2.52 | 200 | 200 | Seretide 50 1OD | Serevent 25 1OD |
| **4** | Moderate Asthmatic | Female | 28 | 171 | 81 | 66 | N/A | 0.5 | 400 | 400 | Clenil Modulite 50 2BD | Serevent 25 2BD |
| **5** | Moderate Asthmatic | Male | 52 | 163 | 63 | 65 | 4.75 | 0.59 | 400 | 400 | Seretide 50 1BD | Seretide 25 1BD |
| **6** | Moderate Asthmatic | Female | 42 | 149 | 69.2 | 88 | 1.5 | 2 | 400 | 400 | Clenil Modulite 50 2BD | Serevent 25 2BD |
| **7** | Moderate Asthmatic | Female | 62 | 173 | 80.6 | 63 | 0.75 | 1 | 400 | 400 | QVAR 50 1BD | Serevent 25 1BD |
| **8** | Moderate Asthmatic | Female | 41 | 171 | 146.4 | 68 | 0.8 | 6.35 | 800 | 800 | Symbicort 200 1BD | Oxis 6 1BD |
| **9** | Moderate Asthmatic | Male | 66 | 173 | 101 | 58 | 19.5 | 2 | 2000 | 2000 | Seretide 125 2BD | |
| **10** | Control | Female | 47 | 167 | 47.6 | | | | | | | |
| **11** | Control | Male | 27 | 181 | 75.4 | | | | | | | |
| **12** | Control | Female | 65 | NR | NR | | | | | | | |
| **13** | Control | Male | 21 | 184 | 69 | | | | | | | |
| **14** | Control | Female | 21 | 166 | 87.3 | | | | | | | |
| **15** | Control | Female | 24 | 158 | 67.9 | | | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **16** | Control | Female | 37 | NR | NR |
| **17** | Control | Female | 54 | 158 | 48 |
| **18** | Control | Female | NR | NR | NR |
| **19** | Control | Female | 23 | 170 | 78.2 |
| **20** | Control | Female | 21 | 167 | 62.4 |
| **21** | Control | Female | 23 | 159 | 55.7 |
| **22** | Control | Female | 21 | 172 | 53.3 |
| **23** | Control | Female | NR | NR | NR |
| **24** | Control | Male | 28 | 163 | 60 |
| **25** | Control | Male | NR | NR | NR |
| **26** | Control | Female | 26 | 178 | 63 |
| **27** | Control | Male | 61 | 185 | 95 |
| **28** | Control | Male | 25 | 173 | 95 |
| **29** | Control | Female | 61 | 165 | 90 |
| **30** | Control | Male | 21 | NR | NR |

### 3.3.4 Saliva sample collection

This research followed the good clinical practices and protocols as approved by the Nottingham Research Ethics committee (REC reference number 10/H0408/34). Samples were collected by Emma Wilson, Wendy Gerrard-Tarpey and Glenn Hearson at the Respiratory BRU at the Nottingham City Hospital (Nottingham, UK) using the protocol described in (Appendix 7.1, 7.2 and 7.3).

Participants were asked not to consume alcohol for 12 hours prior to the sampling of saliva; only unflavoured water was permitted to be consumed. Participants were also asked to refrain from brushing their teeth, chewing gum, having a large meal, drinking tea or coffee and smoking for an hour prior to the sample collection.

Sampling of whole saliva was carried out by the passive drool approach (Chiappin et al. 2007). Each participant was seated with his or her head tilted forward causing saliva to pool in front of his or her mouth. Saliva was dribbled out of the mouth and collected into a glass collection vial (30mL Screw Top Universal Storage Vials, Chromacol, UK) prior to transferring as 600 µL aliquots into micro-centrifuge tubes (2.0 mL, LoBind Eppendorf, UK) for storage. Equal volume sub-aliquots (600 µL) of saliva from all the participants were pooled together, mixed and re-aliquoted as 600 µL aliquots to be used for method development and quality control (QC). Samples were stored at -80ºC immediately after aliquoting to arrest biological activity. The sample collection and storage was carried out according to the sampling protocol (Appendix 7.2 and 7.3).

Eppendorf LoBind microcentrifuge tubes were specifically chosen to be utilised for storage and processing of saliva as they can be stored at -80$^{o}$C as well as offer minimum interference with the analysis in terms of leechables from the surface of the tubes or loss of sample/metabolites due to adsorption on the surface.

### 3.3.5 Sample pre-treatment

The sample pre-treatment, method development and metabolic profiling analysis by UHPLC-MS and the metabolomic data processing was carried out at Centre for Analytical Science, Loughborough University, by the author (Aditya Malkar).

Samples stored at -80ºC were thawed to room temperature before subjecting to sample clean-up by protein precipitation. Each thawed saliva sample was vortexed for 30 seconds followed by ultra-sonication for 1 minute to breakdown mucous substances in saliva and to improve homogeneity. Protein precipitation was achieved by the addition of 1.2 mL acetonitrile to 600 µL of saliva i.e. 2:1 ratio of precipitant to saliva (Polson et al. 2003). The mixture was vortexed for 30 seconds followed by ultra-sonication for 1 minute and then subjected to centrifugation at 10,000 g for 10 minutes at ambient temperature. Precipitated proteins from the sample were removed as a pellet at the bottom of the microcentrifuge tube. The supernatant was transferred to a fresh microcentrifuge tube and evaporated to near dryness (~5 µL) using Turbovap LV concentration workstation (Caliper Life Sciences, Hopkinton, MA, USA) prior to being reconstituted in 300 µL water/acetonitrile (95:05) (v/v) with 0.1% formic acid. The reconstituted saliva extract was placed in a polypropylene micro-insert (Supelco, UK) in an autosampler vial (2 mL) for UHPLC-MS analysis. The sample pre-treatment protocol was followed throughout the study (Appendix 7.4)

### 3.3.6 Instrumentation

UHPLC-MS analyses were carried out using an Agilent 1200 series UHPLC interfaced with an Agilent 6230 time-of-flight mass spectrometer fitted with a JetStream ESI source operated in positive ion mode (Agilent Technologies, Santa Clara, CA, USA).

UHPLC separation was carried out using a Poroshell 120 EC-C18 column (2.1 x 75 mm, 2.7 µm; Agilent Technologies, Santa Clara, CA, USA). Poroshell particles have a solid silica core (1.7 µm) and a porous silica outer layer (0.5 µm thickness), with an end capped C18 bonded phase coated onto the porous outer layer of the particle (Kirkland et al. 2000).

Mobile phase A consisted of 0.1% aqueous formic acid (v/v) and mobile phase B was 0.1% formic acid in acetonitrile (v/v). The chromatographic gradient programme was

optimised using a solution of L-carnitine and hydrocortisone (0.005 mg/mL). The mobile phase flow rate was set to 0.5 mL/min. Saliva extracts (10 μL injected) were analysed using an optimised chromatographic gradient: 5% B (0-1 min), increased to 30% B (1-10 min). This was followed by a column clean-up phase, built into the method to reduce carry over and condition the column for analysis of subsequent samples, in which the mobile phase was increased to 95% B (10-11 min) and maintained at 95% B (11-12 min) before returning to initial conditions (13-15 min).

MS data was acquired with the instrument mode set to extended dynamic range (2 GHz). Mass spectra were acquired in the range *m/z* 50-1200. The scan rate of the MS was 10 scans/s. Source conditions for the UHPLC were: sheath gas temperature and flow, 350 ºC and 12 L/min; drying gas temperature and flow, 150 ºC and 10 L/min; nebuliser pressure, 35 psig; transfer capillary, 4000 V; skimmer voltage, 65 V; fragmentor voltage, 150 V. Continuous internal calibration was performed to achieve desired mass accuracy using signals from infused reference ion solution at *m/z* 121.0509 (protonated purine) and *m/z* 922.0098 [protonated hexakis(1H, 1H, 3H-tetrafluoropropoxy)phosphazine].

Mass Hunter Qualitative Analysis software version B.05.00 (Agilent Technologies, Santa Clara, CA, USA) was used for processing raw data. Metabonomic data analysis including multivariate statistical analysis was carried out using Mass Profiler Professional software (Agilent Technologies, Santa Clara, CA, USA).

### 3.3.7 Method development and quality control

Method blanks were prepared by subjecting 600 uL of water to the entire sample pre-treatment process and were analysed to assess contamination arising from polypropylene microcentrifuge tubes (LoBind Ependorf, UK) and polypropylene micro-inserts used during storage and sample pre-treatment. The effect of injection volumes in the range of 1-20 μL on chromatographic peak areas was analysed for the methodology using pooled saliva extract samples, which were analysed in triplicate. The chromatographic peak areas of the extracted endogenous metabolite ions were plotted against the injection volumes to determine the optimum injection volume for the analysis. The reproducibility of the retention times ($t_R$) and chromatographic peak areas of extracted metabolite ions were assessed from replicate analyses of pooled saliva samples (n=6), using the optimised injection volume.

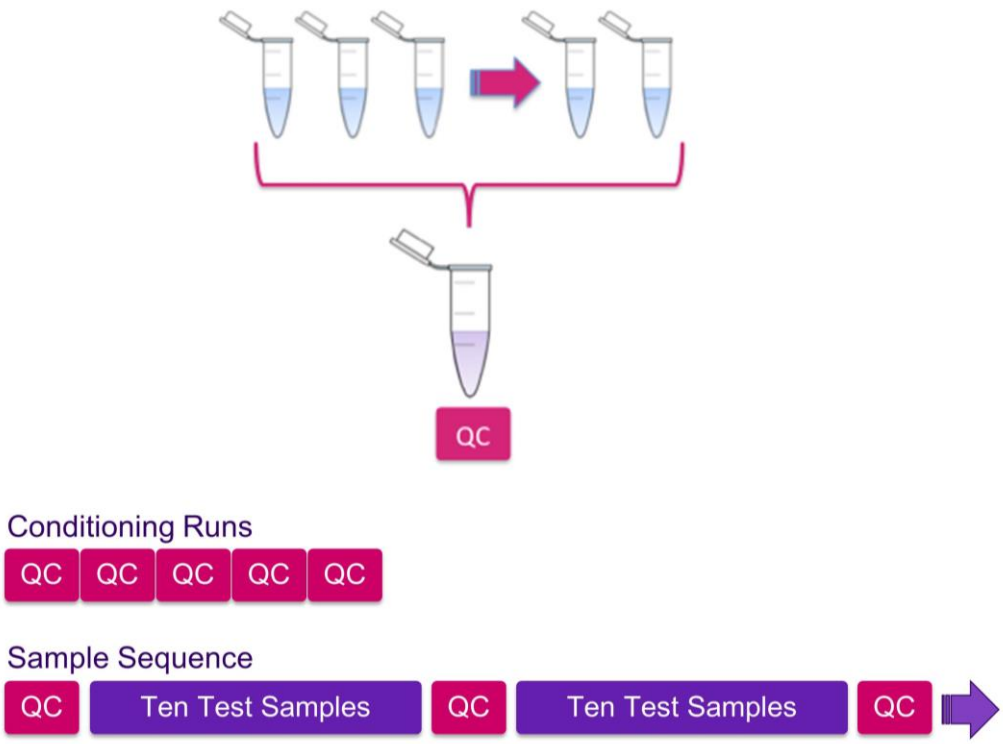



**Figure 3.2 Schematic representation of the analytical protocol for analysis of saliva by UHPLC-MS.**

A series of five consecutive pooled saliva extracts were analysed prior to the analysis of control and asthmatic saliva samples for chromatographic column conditioning purpose. The metabolomic sample list was randomised prior to analysis. A quality control (QC) sample was batch processed and analysed after every 5 samples to monitor the stability of the system and data quality.

### 3.3.8 Data analysis

Mass Hunter Qualitative Analysis software version B.05.00 (Agilent Technologies, Santa Clara, CA, USA) was used for processing raw data. Molecular feature extraction (MFE) algorithm was used to extract unknown molecular features (MF) and compile a data matrix for each sample. Molecular features were extracted using the following thresholds: *m/z* range 50-600 Da; retention time range 0.01-10.00 min and mass spectral peak height of ions (centroid data) 100 cps. The abundance value for each molecular feature was calculated by MassHunter software as the sum of intensities of corresponding isotopic and adducts peaks.

The generated data matrix consisting of the molecular features from all the samples were exported into Mass Profiler Professional software Version 12.05 (Agilent Technologies, Santa Clara, CA, USA) for metabolomic data analysis including chromatographic peak alignment across multiple UHPLC-MS data files and multivariate statistical analysis. Data filtering and alignment was carried out using the following parameters: minimum absolute abundance of molecular features 7000 cps, mass alignment window 10 ppm and retention time alignment window 0.15 min.

Stepwise reduction of the number of molecular features was performed based on their abundance values and frequency of occurrence in sample classes. Aligned molecular features were filtered to reduce the dimensionality of the dataset prior to principal component analysis (PCA). In the first instance, only the molecular features present in at least 85% of samples in only one condition (i.e. asthmatic or healthy control) were retained for further processing. In the next step, molecular features were filtered based on p-value calculated for each using a student's T-test. Lower p-value signifies higher significant difference between the conditions and $P \leq 0.1$ was chosen as the filtering criterion to ensure that only the molecular features, which differ in their respective conditions with a statistical significance of 90%, were retained. The final criterion for filtration of molecular features was the fold change (FC). Molecular features, which satisfied the threshold of $FC \geq 2.0$, were retained for further processing by PCA and to build predictive models.

## 3.4 RESULTS AND DISCUSSION

The non-targeted analysis of salivary metabolome by UHPLC-MS was performed in this study on samples collected from moderate and mild asthmatics (n=16) and healthy controls (n=21). The resulting dataset were pre-processed and then compared using unsupervised multivariate statistical analysis using principal component analysis to identify discriminant molecular features. Initial investigations were directed towards the development of a reliable and reproducible method for the metabolic profiling of saliva.

### 3.4.1 UHPLC-MS Analysis

Saliva extracts were analysed by UHPLC-MS following a protein precipitation step. Protein precipitation was the only sample clean-up process employed in order to minimise the effects of discrimination and unpredictable behaviour of unknown metabolites during sample preparation (Álvarez-Sánchez et al. 2012). The chromatographic gradient was adjusted using a solution of L-carnitine and hydrocortisone (0.005 mg/mL) such that highly polar L-carnitine was slightly retained on the column whereas less polar hydrocortisone eluted towards the end of the 10 minute chromatographic gradient **Figure 3.3**. This enabled all the metabolites from saliva extract to be separated within 10 minutes; hence a cycle time of 15 minutes was achieved with column clean-up followed by equilibration, which is suitable for high throughput metabolomic studies.

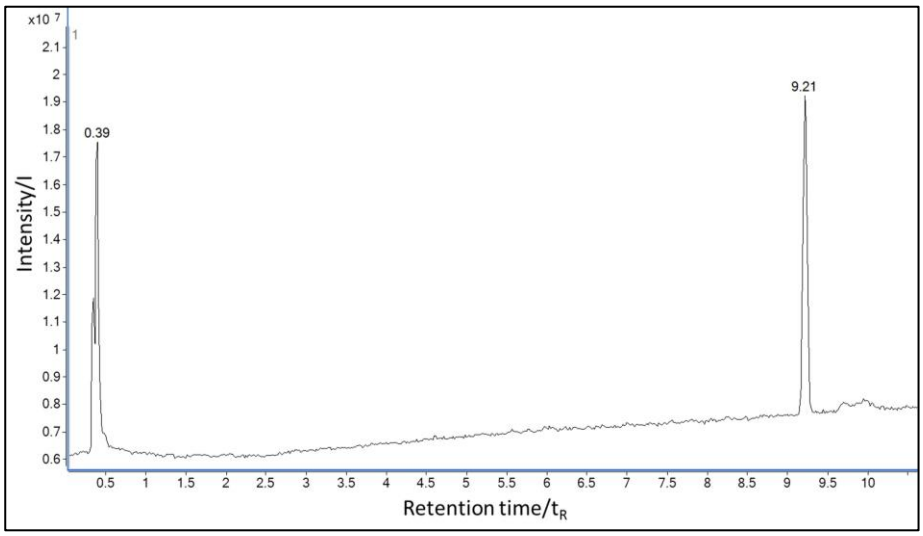



**Figure 3.3 UHPLC-MS analysis of a mixture solution of L-carnitine and hydrocortisone (0.05 mg/ml).** (Data produced on Agilent 6230)

The effect of injection volume was assessed for injection volumes in the range of 1-20 µL. Chromatographic peak areas of five extracted metabolite ions, spread across the entire 10 min gradient were randomly chosen to assess the dynamic range of the analytical system. Chromatographic peak areas showed a good linear response for injection volumes in the entire range of 1-15 µL as shown in **Figure 3.4**.

| | *m/z* (±0.02) | $t_R$/min |
|---|---|---|
| a | 267.13 | 1.9 |
| b | 195.09 | 3.1 |
| c | 716.39 | 5.8 |
| d | 416.36 | 7.2 |
| e | 652.42 | 9.2 |

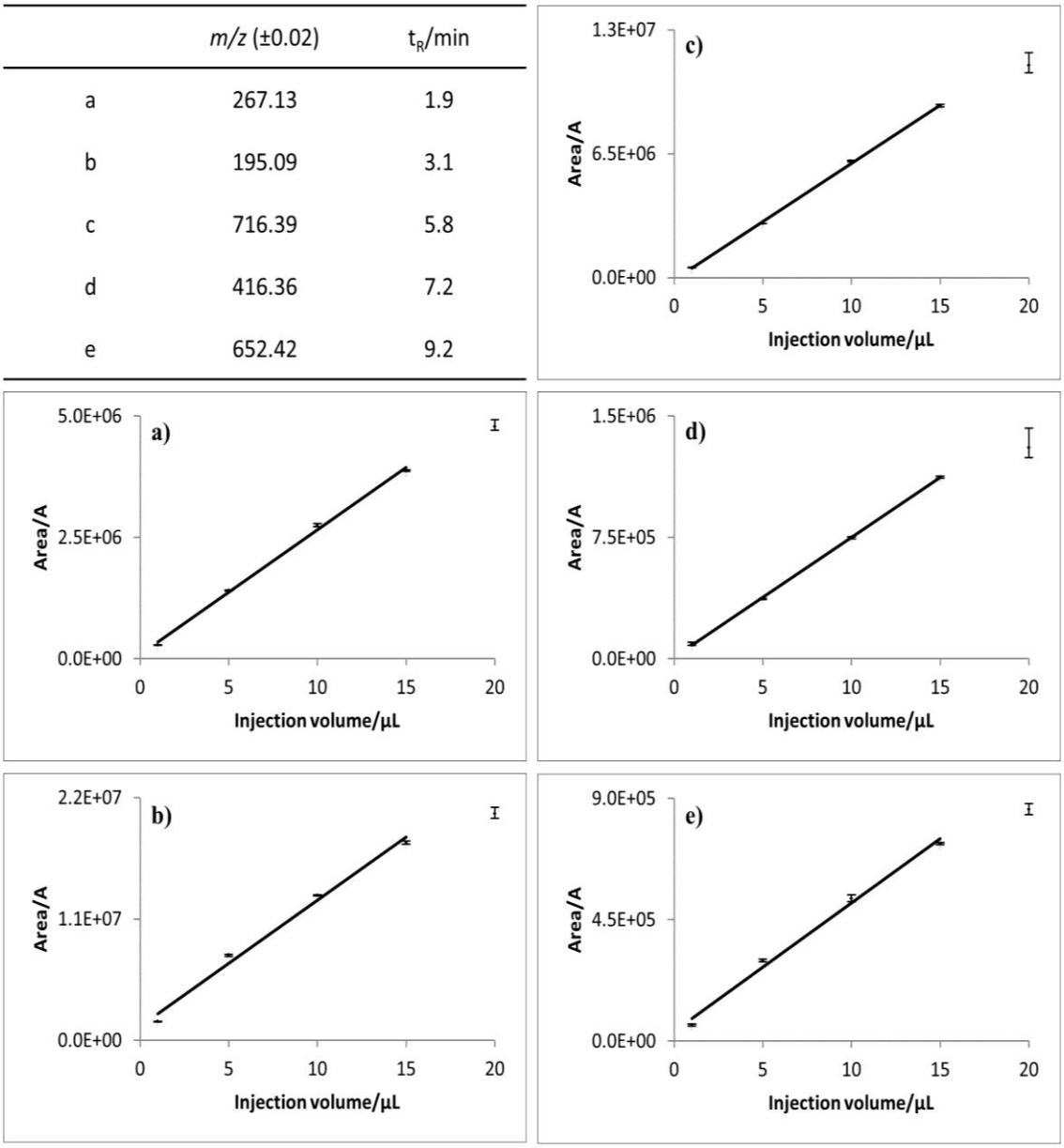



**Figure 3.4 Effect of injection volume range analysis of the methodology. Chromatographic peak areas of a random selection of ions spread across the entire chromatographic range were plotted against the injection volumes. Samples were analysed in triplicates (n=3) denoted by the error.** (Data produced on Agilent 6230)

An injection volume of 10µL was selected from these experiments to be used for the analysis of saliva extract, which ensured that the analytical method was capable of detecting small positive or negative changes in the metabolite concentrations. Endogenous metabolite ions were selected for the validation of the method, which ensured that the matrix effects were accounted for during data processing.

Reproducibility of the retention times of the metabolite ions are important parameters. They influence the outcome of the data alignment performed across all the samples. Based on replicate measurements (n=6) of the same five endogenous metabolites, the reproducibility of the retention time was determined to be less than 0.4%, demonstrating excellent UHPLC reproducibility (**Table 3-2**). The reproducibility of the chromatographic peak areas is also important for the identification of up-regulated and down-regulated metabolite ions. **Table 3-2** also shows an inter-day %RSD of less than 4% for the peak areas of metabolite ions over the range *m/z* 267 ($t_R$=1.96 min) to *m/z* 416 ($t_R$=9.21 min). Hence, it is demonstrated that relatively small variations in metabolite areas between control and asthmatic samples can be clearly distinguished.

**Table 3-2 Reproducibility analysis of retention time and chromatographic peak areas (n=6)**

| Assignment | *m/z* | Retention Time | | Area | |
|---|---|---|---|---|---|
| | (±0.02) | Mean $t_R$/min | RSD% | Mean Area/A | RSD% |
| M1 | 267.13 | 1.96 | 0.32 | 2005909 | 3.73 |
| M2 | 195.09 | 3.15 | 0.18 | 12909485 | 2.09 |
| M3 | 716.39 | 5.80 | 0.10 | 5891274 | 3.58 |
| M4 | 652.42 | 7.02 | 0.08 | 590003 | 3.81 |
| M5 | 416.36 | 9.21 | 0.09 | 561820 | 2.76 |

A typical UHPLC-MS profile for a saliva extract overlaid with a method blank is shown in **Figure 3.5**, depicting the complexity of the dataset. Method blanks demonstrated no contamination due to the polypropylene micro-centrifuge tubes and UHPLC micro-inserts used for storage, pre-treatment and analysis. Employing the optimised gradient programme for the analysis enabled the elution of all detectable metabolites from the saliva sample in 10 minutes with a typical chromatographic peak base width of 9 seconds.

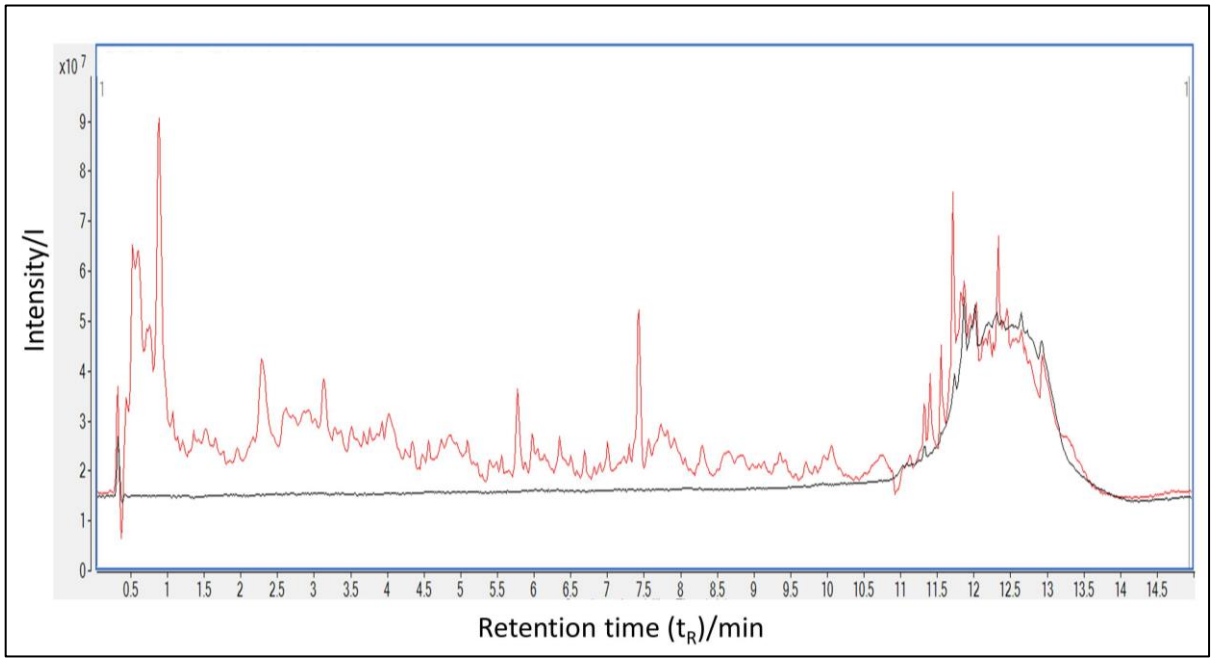



**Figure 3.5 Example of a total ion chromatogram (TIC) of saliva extract (red trace) and a blank (black trace).** (Data produced on Agilent 6230)

### 3.4.2 Data reduction and multivariate analysis

Visual inspection or manual processing of the data obtained from UHPLC-MS measurements would be time consuming and inefficient given the complexity of the data **Figure 3.5**. Hence, a molecular feature extractor (MFE) algorithm was employed, which enabled the automated extraction of ions corresponding to compounds present in saliva samples. This algorithm combines ion species such as multiply charged species, isotopes and adducts generated from a single compound into a single molecular feature (MF) to be utilised as a single variable in subsequent analysis.

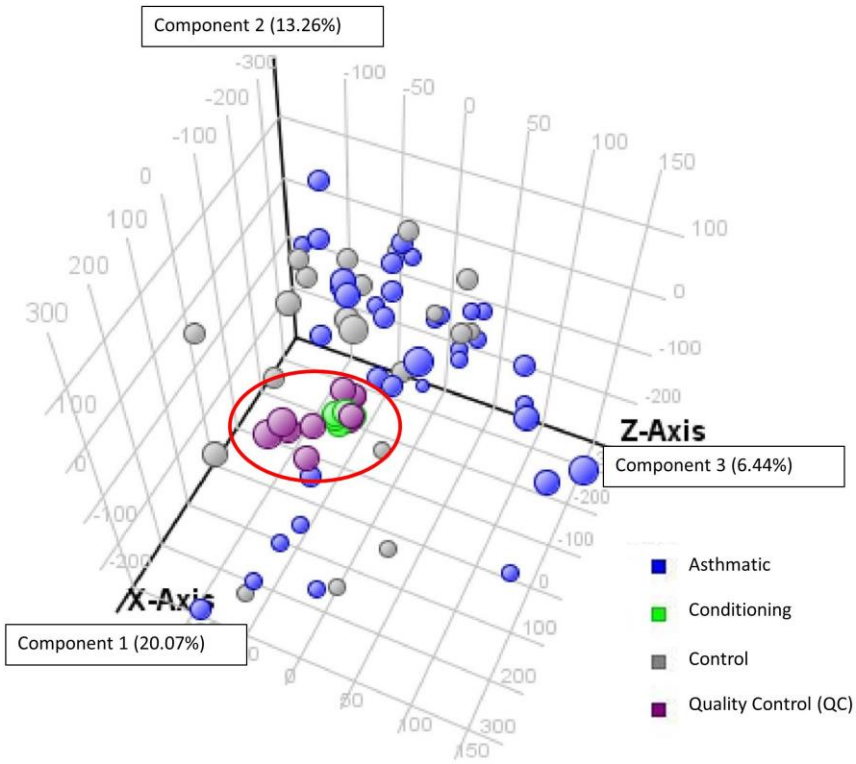



**Figure 3.6 A 3-D Principal component analysis (PCA) of saliva samples from the study coloured according to sample groups. X-Axis: Component 1 (20.07%), Y-Axis: Component 2 (13.26%) and Z-Axis: Component 3 (6.44%).**

A total of 35011 molecular features were extracted from all the samples. Out of these, 14267 were aligned across all the samples from the study. Principal component analysis (PCA) was carried out using these 14267 molecular features on all the sample groups from the study including conditioning runs and quality control (QC) samples. Clustering of QC samples (circled in **Figure 3.6**) provides evidence of consistent and reliable data quality across all the samples run, which is required for metabolic profiling studies. However, **Figure 3.6** shows no clear separation of asthmatic and healthy population; hence, further data reduction was required to identify discriminant metabolites between control and asthmatic samples.

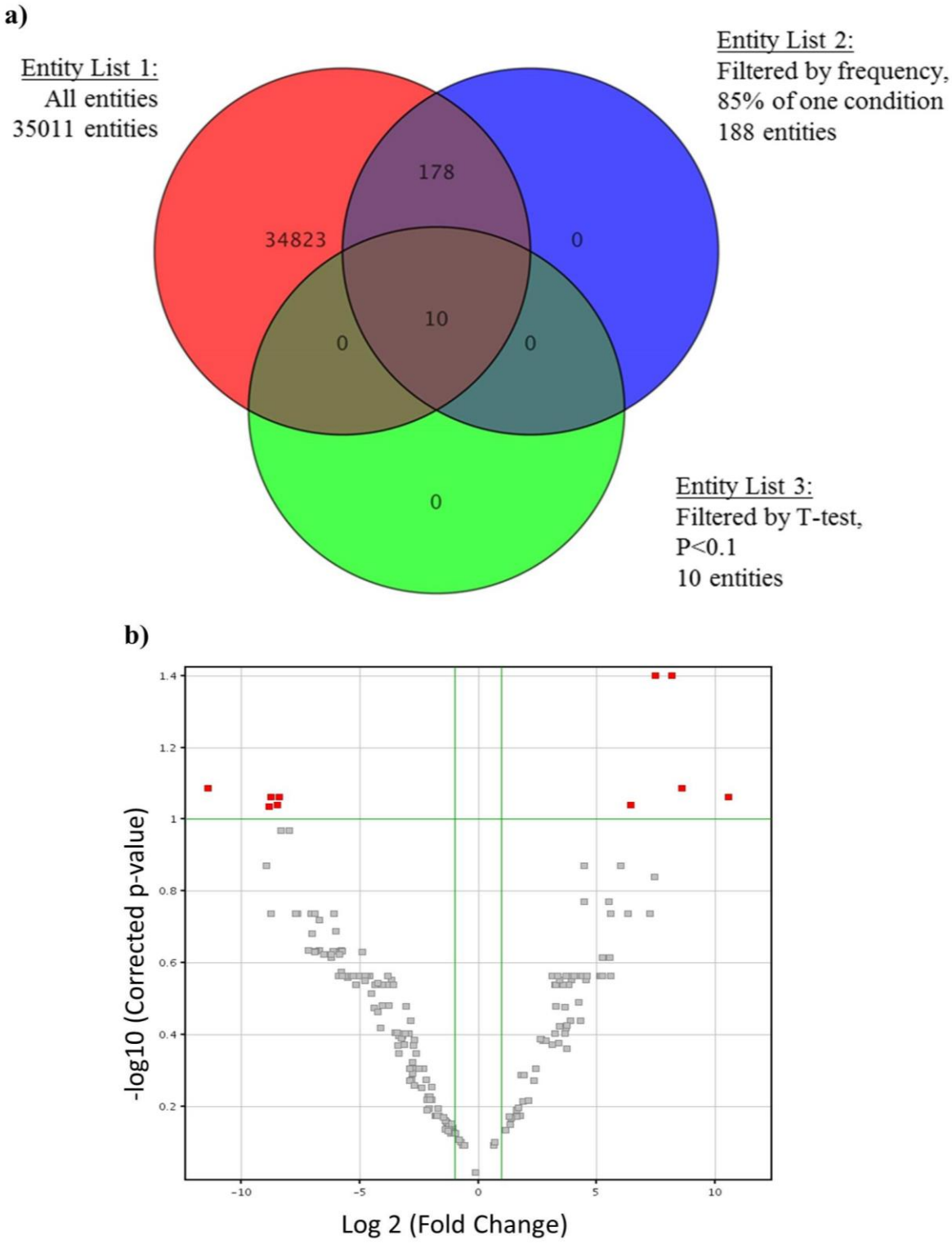



**Figure 3.7 a) Summary of data reduction based on frequency of molecular features in samples, b) Volcano plot illustrating filtering of molecular features based on the results of T-test and fold change analysis (P value cut off 0.1 and fold change cut off 2.0)**

Moderate asthmatics (n=9) were compared against healthy individuals (n=21) in the first instance. The dimensionality of the dataset was reduced prior to PCA by filtering the data based on frequency of occurrence of molecular features and outcomes of t-test (summarised in a Venn diagram in **Figure 3.7** a). In the first step, all molecular features which were present in at least 85% of samples of at least one group (i.e. moderate asthmatic or healthy) were retained. A high cut-off of 85% was chosen as a rigorous measure to account for the distribution of the individual groups' population with respect to factors such as age, gender etc. A high cut-off ensures that the potential biomarker ions obtained at the end of the study would be independent of factors such as age or gender, and would be universally applicable to all the population. This reduced the number of MFs from 14627 to 188. In the next phase, the MFs were filtered based on their p-values obtained from a t-test and fold change analysis. MFs, which satisfied the p-value cut-off value of 0.1 and fold change value of FC≥2.0 were retained as illustrated from the volcano plot in **Figure 3.7** b. This reduced the number of MFs from 188 to 10. A PCA plot was constructed based on the filtered 10 molecular features for moderate asthmatic and healthy control individuals (**Figure 3.8**). This unsupervised model shows clear separation of moderate asthmatics from healthy control samples, with the exception of a single asthma patient sample.

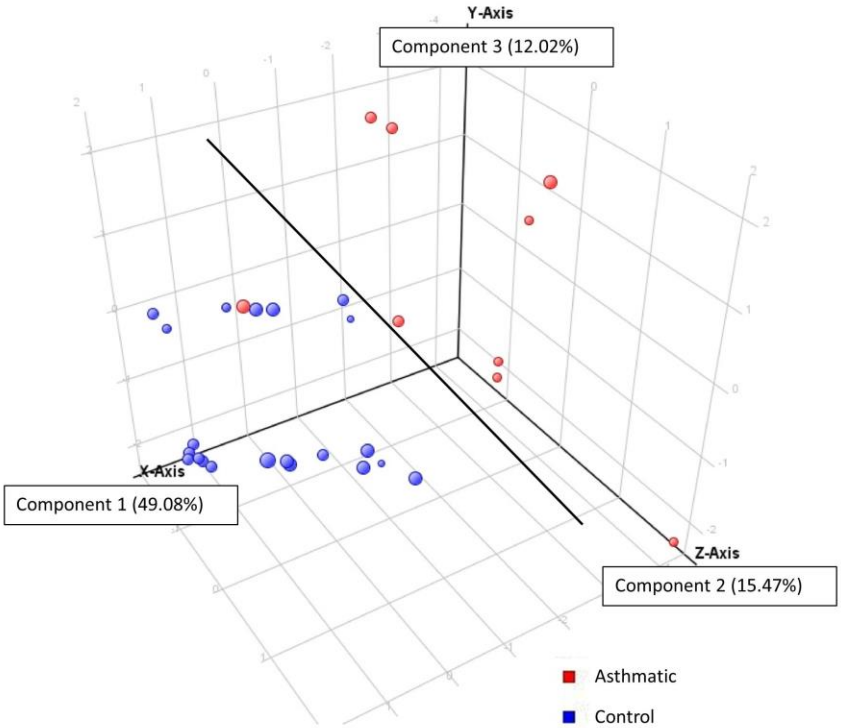



**Figure 3.8 A 3-D Principal component analysis (PCA) of saliva samples from moderate asthmatics (red) and healthy control samples (blue). X-Axis: Component 1 (49.08%); Y-Axis: Component 2 (15.47%) and Z-Axis: Component 3 (12.02%).**

The moderate asthmatic patients were on different treatment regimens (Appendix 7.6) and therefore the ability to distinguish between moderate asthma and healthy controls is not treatment specific. Hence, it can be concluded that the molecular features responsible for the class separation are representative of the endogenous factors pertaining to asthma and not due to the effect of treatment.

The next step was to assess the contribution of the 10-discriminant features towards class separation. The correlation and covariance of the features was assessed from an s-plot (**Figure 3.9**). MFs characteristic of asthmatics (i.e. up-regulated) are represented in the first quadrant of the plot and the MFs characteristic of healthy control individuals (i.e. down-regulated) are shown in the third quadrant of the plot.

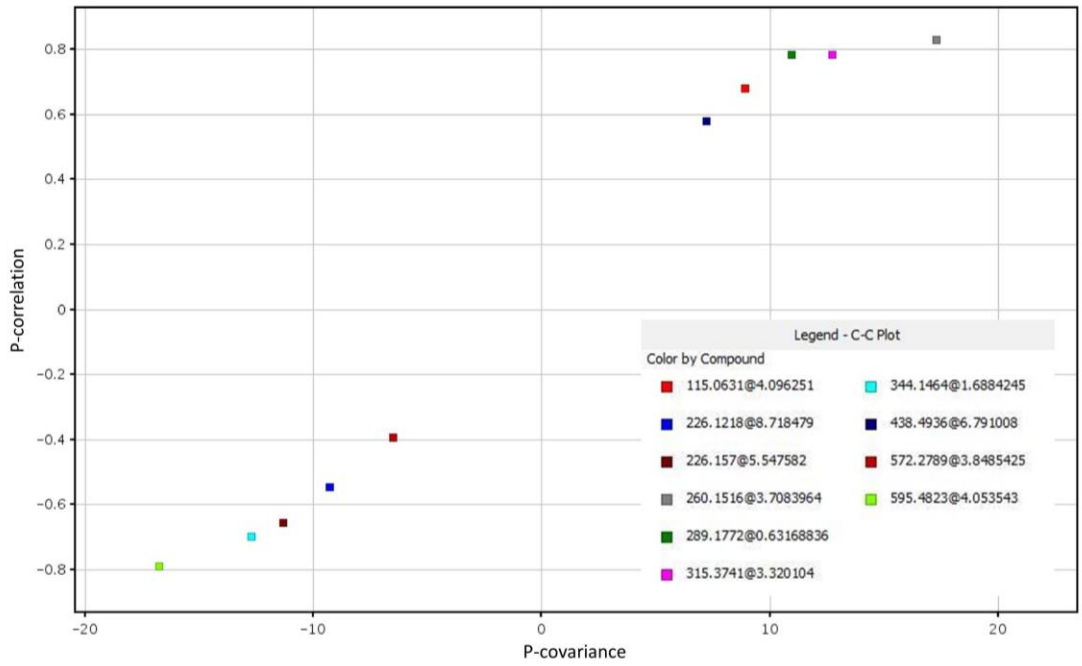



**Figure 3.9 A S-plot representing the correlation and covariance of the molecular features responsible for class separation.**

**Table 3-3** summarises the potential metabolic biomarker ions identified from this study. Accurate mass measurement data and isotopic ratio data on the discriminant metabolite ions were obtained by analysing a pooled saliva extract sample using high resolution mass spectrometry and utilised to obtain tentative elemental compositions on the discriminant metabolite ions with a mass error window of 3 ppm.

**Table 3-3 Discriminant molecular features (MFs) obtained from moderate asthmatics versus control samples**

| Marker No. | Retention Time (min) | *m/z* (±0.02) | Regulation | Tentative Elemental Composition |
|---|---|---|---|---|
| 1 | 0.4 | 116.07 | Up | $C_5H_9NO_2$ |
| 2 | 2.2 | 261.14 | Up | $C_{12}H_{23}N_3O_5$ |
| 3 | 1.1 | 290.17 | Up | $C_{11}H_{20}N_2O_5$ |
| 4 | 3.6 | 316.22 | Up | $C_{15}H_{29}N_3O_4$ |
| 5* | 4.4 | 439.46 | Up | N/A |
| 6 | 3.4 | 227.13 | Down | $C_{12}H_{18}O_4$ |
| 7 | 4.7 | 573.26 | Down | $C_{26}H_{44}N_4O_4S_3$ |
| 8 | 0.5 | 596.33 | Down | $C_{28}H_{41}N_{11}O_2S$ |
| 9 | 1.7 | 345.12 | Down | $C_{22}H_{17}ClN_2$ |
| 10 | 5.5 | 227.14 | Down | $C_{11}H_{18}N_2O_3$ |

*No elemental composition could be postulated within the set 3 ppm tolerance window

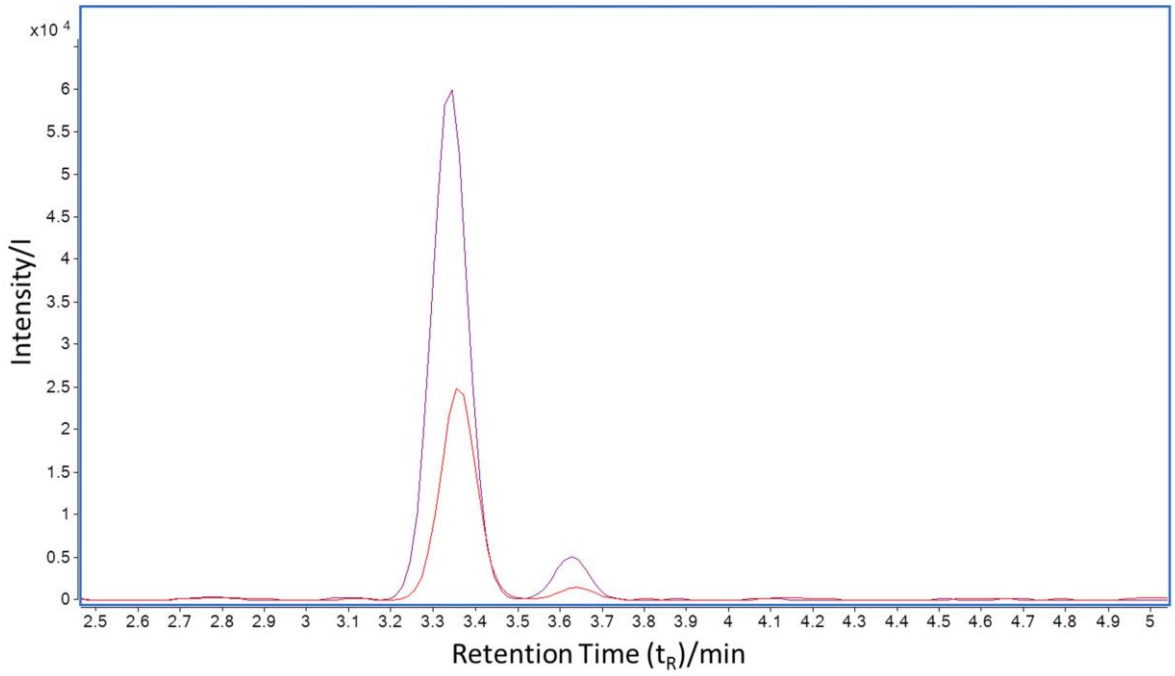



**Figure 3.10 Extracted ion chromatogram (EIC) for selected potential biomarker ion *m/z* 316.22 ± 0.02 showing change in intensity between moderate asthmatic (purple trace) and healthy control individuals (red trace).** (Data produced on Agilent 6230)

An example of a potential biomarker ion (*m/z* 316.22 ± 0.02) is shown in **Figure 3.10**, which shows up-regulation of the ion. The change in the chromatographic peak areas can be observed from the two overlaid example EICs. The EIC shows the presence of two chromatographic peaks for a single species at $t_R$=3.35 and $t_R$=3.65; with both peaks depicting up-regulation. Box and whisker plots were constructed for all the potential biomarker ions to visualise the trend of the biomarker ions in the entire demographic of the study. Box and whisker plots from up-regulated ions **Figure 3.11** and down-regulated ions **Figure 3.12** show marked differences in their normalised chromatographic peak areas.

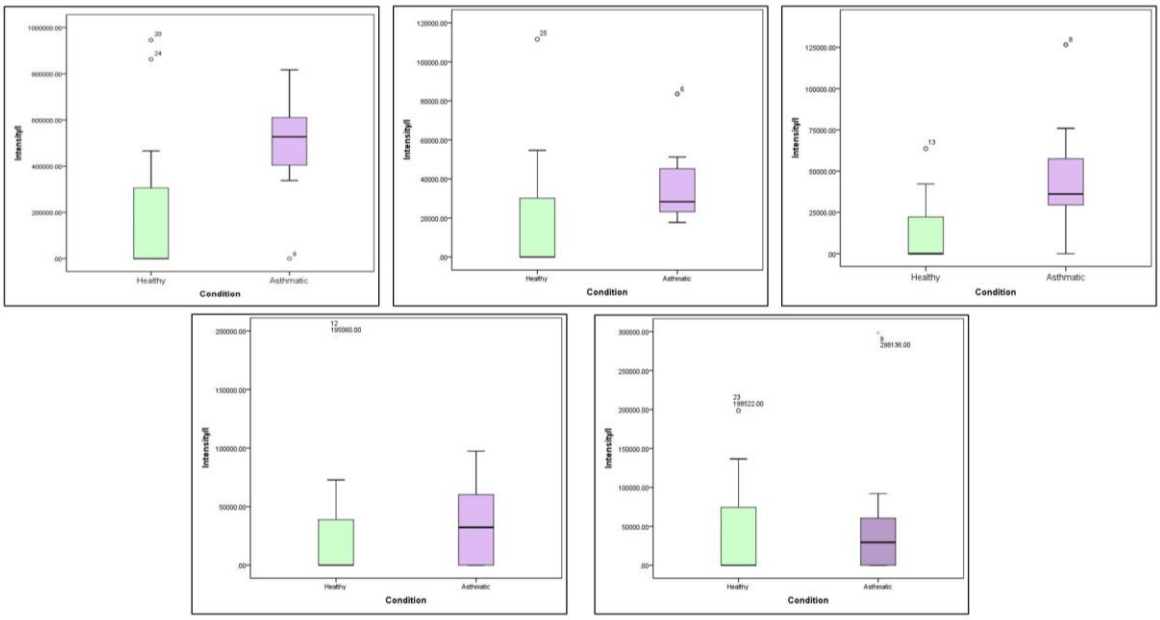



**Figure 3.11 Box and Whisker plots for Up-regulated biomarker ions (clockwise from top left: *m/z* 116.07, 261.14, 290.17, 316.22 and 439.46).**

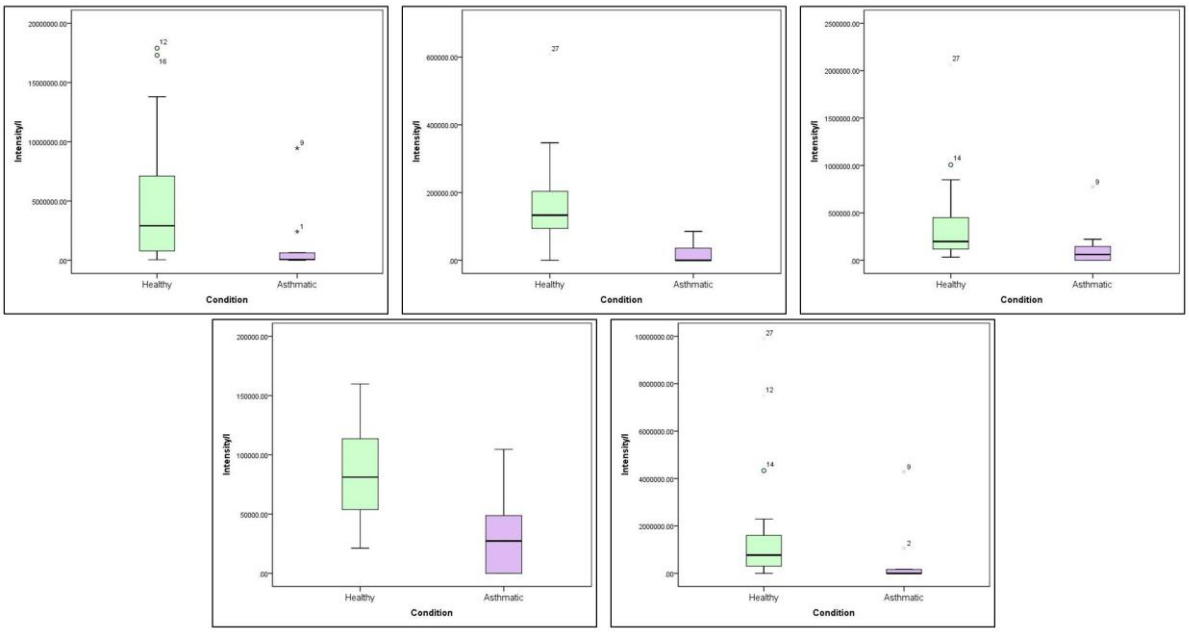



**Figure 3.12 Box and Whisker plots for Down-regulated potential biomarker ions (clockwise from top left: *m/z* 227.13, 573.26, 596.33, 345.12 and 227.14).**

### 3.4.3 Predictive models

The discrimination potential offered by the data was evaluated using a widely utilised supervised pattern recognition method of partial least squares discriminant analysis (PLS-DA). A statistical model based on PLS-DA was constructed and validated for classification of asthmatic samples. The results of sample classification are presented in **Table 3-4** in terms of recognition and prediction abilities, representing the percentage of the samples correctly classified during model training and cross-validation. The overall recognition ability during training of the model was 80% and the prediction ability for model cross-validation was 96.7%. The model therefore shows high predictive capabilities for distinguishing between moderate asthmatics and healthy controls by non-invasive sampling of saliva combined with UHPLC-MS metabolite profiling.

The model created using moderate asthmatic samples was used to classify mild (early stage) asthmatic samples. The success rate of detection of mild asthma was low (55.6 %). However, this is not a true representation of the predictive capabilities as a model built on moderate asthmatics and healthy individuals was tested on mild asthmatics due to restricted patient numbers. Further refinement of the model is required to distinguish between healthy controls, mild and moderate asthmatics.

**Table 3-4 Overview of classification results obtained by PLS-DA model.**

| | **Predicted Asthmatic** | **Predicted Controls** | **Accuracy (%)** |
|---|---|---|---|
| **Model Training** | | | |
| **True Moderate Asthmatic** | 8 | 1 | 88.9 |
| **True Controls** | 5 | 16 | 76.2 |
| **Recognition ability** | | | 80.0 |
| **Model cross-validation** | | | |
| **True Moderate Asthmatic** | 9 | 0 | 100.0 |
| **True Controls** | 1 | 20 | 95.2 |
| **Recognition ability** | | | 96.7 |
| **Model testing using mild asthmatics** | | | |
| **True Mild Asthmatic** | 27 | 15 | 55.6 |
| **True Controls** | 0 | 12 | |
| **Recognition ability** | | | 55.6 |

## 3.5 CONCLUSIONS

In this pilot study, the robust methodology developed, based on protein precipitation and UHPLC-MS has been used in combination with multivariate statistical analysis for the identification of potential metabolomic biomarkers of asthma from saliva. A predictive PLS-DA model was created based on the discriminant metabolites obtained from the comparison of moderate asthmatics versus healthy control samples which clearly distinguished between the two groups. This model was evaluated for the recognition of asthma in samples obtained from mild asthmatic population.

The potential benefits of a simple, non-invasive sampling technique for asthma screening have been highlighted. The technique can be used in conjunction with or as a preliminary alternative to the existing diagnostic tests such as forced expiratory volume (FEV1%), sputum eosinophil count and methacholine challenge. The work presented here demonstrates the advantages and the potential of non-invasively obtained saliva as a biofluid to diagnose and monitor the response to treatment of asthma.

To our knowledge, this is the first study to employ a passive drool saliva test combined with mass spectrometry as a potential diagnostic tool for asthma. We found this approach could discriminate between asthma and health controls; a model derived from these 10 discriminant metabolite ions classified asthma with an accuracy of 80% and 96.7% for model training and cross-validation respectively. These figures compare favourably to current tests used for the diagnosis of asthma. Tests of airflow obstruction are widespread however suffer from limited sensitivity and specificity. One study found the sensitivities and specificities for FEV1/FVC ratio to be 61%, 60% respectively; bronchodilator response 49%, 70% and peak flow amplitude percent mean to be 43%, 75%(Hunter et al. 2002). Due in part to the limited tests available, the diagnosis and assessment of treatment response is still based on symptoms with high rates of misdiagnosis(Luks et al. 2010). More invasive tests such as bronchial challenge or induced sputum are available but are limited by their invasive nature, cost and need for expertise. They also only reflect one aspect of the asthmatic process(Grainge et al. 2011). Despite these tests a proper diagnosis remains a challenge even in specialist asthma clinics(Robinson et al. 2003).

The benefits of a simple, non-invasive sampling technique for asthma diagnosis are obvious and a great deal of work has focussed on the development of breath sampling, including exhaled nitric oxide (Shaw et al. 2007) , measurement of volatile organic compounds(Montuschi et al. 2007; Ibrahim et al. 2011) and the e-nose (Dragonieri et al. 2009) . The discriminatory ability of these tests varies. Importantly the passive drool technique can be used in conjunction with or as a preliminary alternative to the existing diagnostic tests such as FEV1%, has good discriminatory ability, and provides metabolic information on the status of the airway. The saliva sampling kit is inexpensive and the passive drool technique can be performed in the clinic by a wide range of patients, including children. Samples can also be stored and transported with no need for immediate analysis. The use of UHPLC-MS is increasing in clinical laboratories because of the high throughput and specificity of the technique and the development of mass spectrometry-based assays for the simultaneous determination of panels of biomarkers has been shown to offer a rapid and routine approach to clinical diagnosis.

Our pilot study has similar limitations to others in the field being cross sectional with relatively few participants. We cannot fully exclude a treatment effect, however the patients with asthma were on different treatment regimens and drugs; hence the molecular features responsible for the class separation are most likely representative of the endogenous factors pertaining to asthma and not due to a drug signal. The identification, validation and the biological significance of the discriminant molecular features obtained from the samples were beyond the scope of this pilot but will form part of future work. The discriminant metabolic features identified in the study also need to be validated and quantified using a larger cohort (n>100) of moderate and/or severe asthmatics versus healthy control populations using a targeted metabolite analysis approach.

In conclusion the passive drool technique combined with UHPLC-MS is a novel non-invasive method of accurately differentiating patients with asthma from healthy controls. Further longitudinal work is required to fully assess the discriminatory ability of this test.

## 3.6 CHAPTER THREE REFERENCES


Adamko, D.J.; Sykes, B.D.; Rowe, B.H. *Chest* **2012,** 141, 1295–1302.

Álvarez-Sánchez, B.; Priego-Capote, F.; Luque de Castro, M.D. *J. Chromatogr. A* **2012**, 1248, 178–181.

Anderson, G.P. *Lancet* **2008**, 372, 1107–1119.

Asthma UK, Asthma Facts and FAQs, **2013**, URL: http://www.asthma.org.uk/asthma-facts-and-statistics (accessed 7.7.13).

Boulet, L.P. *Curr. Allergy Asthma Rep*. **2003**, 3, 166–171.

British Thoracic Society, Scottish Intercollegiate Guidelines Network, **2012**. British Guideline on the Management of Asthma: A national clinical guideline.

Carraro, S.; Rezzi, S.; Reniero, F.; Héberger, K.; Giordano, G.; Zanconato, S.; Guillou, C.; Baraldi, E. *Am. J. Respir. Crit. Care Med*. **2007**, 175, 986–990.

Chiappin, S.; Antonelli, G.; Gatti, R.; De Palo, E.F. *Clin. Chim. Acta* **2007**, 383, 30–40.

Dragonieri, S.; Tongoussouva, O.; Zanini, A.; Imperatori, A.; Spanevello, A. *Monaldi Arch. Chest Dis.* **2009**, 71, 119–126.

Grainge, C.L.; Lau, L.C.K.; Ward, J.A.; Dulay, V.; Lahiff, G.; Wilson, S.; Holgate, S.; Davies, D.E.; Howarth, P.H. *N. Engl. J. Med*. **2011**, 364, 2006–2015.

Holgate, S.T. *Clin. Exp. Allergy* **2008**, 38, 872–897.

Hunter, C.J.; Brightling, C.E.; Woltmann, G.; Wardlaw, A.J.; Pavord, I.D. *Chest* **2002**, 121, 1051–1057.

Ibrahim, B.; Basanta, M.; Cadden, P.; Singh, D.; Douce, D.; Woodcock, A.; Fowler, S.J. *Thorax* **2011**, 66, 804–809.

Kirkland, J.J.; Truszkowski, F. A,; Dilks, C.H.; Engel, G.S. *J. Chromatogr. A* **2000**, 890, 3–13.

Luks, V.P.; Vandemheen, K.L.; Aaron, S.D. *Eur. Respir. J. Off. J. Eur. Soc. Clin. Respir. Physiol.* **2010**, 36, 255–260.

Mackay, I.R.; Rosen, F.S. *N. Engl. J. Med.* **2001**, 344, 350–362.

Marklund, B.; Tunsäter, A.; Bengtsson, C. *Fam. Pract*. **1999**, 16, 112–116.

Mattarucchi, E.; Baraldi, E.; Guillou, C. *Biomed. Chromatogr.* **2012,** 26, 89–94.

Montuschi, P.; Santonico, M.; Mondino, C.; Pennazza, G.; Mantini, G.; Martinelli, E.; Capuano, R.; Ciabattoni, G.; Paolesse, R.; Di Natale, C.; Barnes, P.J.; D'Amico, A. *Chest* **2007**, 137, 790–796.

Nicholson, J.K.; Lindon, J.C. *Nature* **2008**, 455, 1054–1056.

Polson, C.; Sarkar, P.; Incledon, B.; Raguvaran, V.; Grant, R. *J. Chromatogr. B.* Analyt. Technol. Biomed. Life Sci. **2003**, 785, 263–275.

Robinson, D.S.; Campbell, D.A.; Durham, S.R.; Pfeffer, J.; Barnes, P.J.; Chung, K.F. *Eur. Respir. J. Off. J. Eur. Soc. Clin. Respir. Physiol*. **2003**, 22, 478–483.

Saude, E.J.; Skappak, C.D.; Regush, S.; Cook, K.; Ben-Zvi, A.; Becker, A.; Moqbel, R.; Sykes, B.D.; Rowe, B.H.; Adamko, D.J. *J. Allergy Clin. Immunol.* **2011**, 127, 757–764.

Shaw, D.; Green, R.; Berry, M.; Mellor, S.; Hargadon, B.; Shelley, M.; Mckenna, S.; Thomas, M.; Pavord, I.; Shaw, D. *Prim. Care Respir. J.* **2012**, 21, 283–287.

Shaw, D.E.; Berry, M. A.; Thomas, M.; Green, R.H.; Brightling, C.E.; Wardlaw, A.J.; Pavord, I.D. *Am. J. Respir. Crit. Care Med*. **2007**, 176, 231–237.

Szefler, S.J.; Wenzel, S.; Brown, R.; Erzurum, S.C.; Fahy, J. V.; Hamilton, R.G.; Hunt, J.F.; Kita, H.; Liu, A.H.; Panettieri, R. A.; Schleimer, R.P.; Minnicozzi, M. *J. Allergy Clin. Immunol.* **2012**, 129, S9–23.

Verrills, N.M.; Irwin, J.A.;, He, X.Y.; Wood, L.G.; Powell, H.; Simpson, J.L.; McDonald, V.M.; Sim, A.; Gibson, P.G. *Am. J. Respir. Crit. Care Med*. **2011**, 4–6.

Wadsworth, S.; Sin, D.; Dorscheid, D. *J. Asthma Allergy* **2011**, 4, 77–86.

Wexler, P.J.; Sequeira W.L.; Helmerhorst, E.J.; Blicharz, T.M.; Hayman, R.B.; Chow C.; Oppenheim, F.G.; Walt, D.R.; Little, F.F. *Proc. Am. Thorac. Soc*. **2009**, 6, 334.

# CHAPTER FOUR:

# THE APPLICATION OF ULTRA PERFORMANCE LIQUID CHROMATOGRAPHY-ION MOBILITY-MASS SPECTROMETRY (UPLC-IM-MS) FOR GLOBAL METABOLITE PROFILING OF SALIVA.

## 4.1 INTRODUCTION

The health benefits of aerobic exercise for humans are well known. Exercise produces a vast amount of salutary effects but the understanding of how these effects occur is limited. It is a well-established fact that exercise triggers changes in composition and concentration of numerous metabolites (Lewis et al. 2010). A number of important investigations on the effect of saliva on human biochemistry have been carried out using plasma and serum (González et al. 2008). As an alternative to such invasively collected bio-matrices, saliva presents itself as an attractive and effective medium to evaluate the effects of exercise or in other words, physiological stress. Exercise has been shown to affect the salivary levels of steroid hormones, immunoglobulins (such as SIgA, SIgM, SIgG etc.), antimicrobial proteins and enzymes such as α-amylase (DeCaro 2008), lysozyme and lactoferrin etc. Targeted analysis of cortisol as a biomarker of stress from saliva is carried out routinely using a variety of assays and analytical methods. Catabolic and anabolic effects of exercise on human body's stress response have also been assessed from saliva by monitoring cortisol, testosterone and dehydroepiandrosterone (Papacosta & Nassis 2011). It is clear that the measurement of physiological biomarkers in whole saliva can provide a significant tool for assessing the immunological and endocrinological status associated with exercise (Pollard 1995; Perogamvros et al. 2009). A holistic approach to this by untargeted metabolite profiling can provide valuable information towards improving the understanding of biochemistry associated with exercise.

The fundamentals of UHPLC-IM-MS and reviews of metabolite profiling of saliva have been covered in Chapter one. Ion mobility spectrometry is a technique which separates ionised molecules on the basis of their collision cross sections (i.e. size and shape) in the gas phase and has been extensively reviewed (Creaser et al. 2004; Kanu et al. 2008) as well as an account of the same is presented in Chapter One. The application of ion mobility spectrometry hyphenated with mass spectrometry (IM-MS) to metabolite profiling has been reported (Dwivedi, Wu, et al. 2007; Harry et al. 2008), but the use of ion mobility-mass spectrometry combined with ultra-high performance liquid chromatography (Swartz 2005) has not been explored.

The routinely used combination of mass spectrometry (MS) with a chromatographic separation technique such as ultra-high performance liquid chromatography (UHPLC)

presents a powerful tool for the analysis of metabonome of complex biological systems. In modern metabolomic studies, UHPLC-MS is a favoured technique as three-dimensional data i.e. retention time ($t_R$), mass-to-charge ratio (*m/z*) and intensities (I) is obtained. As a result hyphenated techniques provide with better resolution of analyte metabolites. However, there is a potential for further improvement of the analytical space by the use of ion mobility (IM) hyphenated with chromatography-mass spectrometry systems. The combination of ion mobility with UHPLC-MS presents an opportunity for complex metabolomic samples to be separated on the basis of retention time, ion mobility drift time and mass-to-charge ratio in a single chromatographic run. The use of ion mobility dimension of separation does not demand any additional sample preparative steps nor does it affect the high throughput offered by UHPLC. Thus this combination provides a valuable opportunity for exploring complex biofluids such as saliva.

There have been no previously reported studies for the metabolic profiling of saliva using hyphenated UHPLC-IM-MS. This study describes the first instance of the development and validation of a method for the targeted and untargeted metabolic profiling of saliva by ultra-high performance liquid chromatography-ion mobility-mass spectrometry (UHPLC-IM-MS). The method has been applied to the profiling of metabolites and the discovery of potential discriminating biomarker ions of exercise-induced physiological stress.

## 4.2 CHAPTER FOUR AIMS AND OBJECTIVES

The purpose of the work presented in this chapter was to explore novel methods for the metabolic profiling of human saliva employing the use of travelling wave ion mobility spectrometry (TWIMS) as an additional dimension of separation in conjugation with UHPLC-MS. The specific objectives were:

- To explore the potential of travelling wave ion mobility spectrometry combined with ultra-high performance liquid chromatography and mass spectrometry for metabolite profiling analysis of saliva.
- To evaluate the advantages offered by ion mobility for targeted and untargeted metabolite profiling analysis.
- To carry out untargeted global metabolite profiling investigation for the discovery of potential discriminating biomarker ions of physiological stress using the developed method.

## 4.3 EXPERIMENTAL

### 4.3.1 Materials and samples

#### *4.3.1.1 Chemicals*

Acetonitrile (analytical grade), water (analytical grade) and formic acid (>99.9%) were obtained from ThermoFisher Scientific (Loughborough, UK). L-carnitine (>99%) and hydrocortisone (>99%) were obtained from Sigma-Aldrich (Gillingham, UK). A 400 ng/mL stock solution of leucine enkephalin was obtained from Waters Corporation (Manchester, UK) and used to prepare a 2 ng/mL working solution of leucine enkephalin in water:acetonitrile (50:50) (v/v) as a LockMass (*m/z* 556.2771) solution. A standard mixture of 0.01 mg/mL L-carnitine and hydrocortisone was prepared in water:acetonitrile (95:05) (v/v) with 0.1% formic acid. A standard solution of 0.01 mg/mL L-phenylalanine was prepared in water:acetonitrile (95:05) (v/v) with 0.1% formic acid.

#### *4.3.1.2 Saliva sample collection and storage*

Sampling of saliva for this study was carried out under the ethical approval obtained from the local ethical advisory committee (Ethical Advisory Committee, Loughborough University, Loughborough, LE11 3TU) with reference number G10-P24. Participants were healthy, non-smoking males, aged 18-35 and were recruited from Loughborough University staff and students (n=10). Samples were collected in the exercise laboratories of the School of Sport, Health and Exercise Sciences (SSEHS) by PhD students and staff of the Department of Chemistry – Helen J. Martin, Matthew A. Turner and Pareen Patel and SSEHS – Phillip Watson.

Saliva was collected by passive drool from participants after an overnight fast. Participants were seated with their head tilted forward, allowing saliva to pool in front of the mouth. Saliva was allowed to be dribbled out of the mouth into a glass collection vial (30 mL Chromacol, UK) at regular intervals. This process was repeated until a sufficient colume of saliva (i.e. 1 mL) was obtained (Chiappin et al. 2007). Saliva collected from all the participants was re-aliquoted as 500 μL aliquots into microcentrifuge tubes (2.0 mL, LoBind Eppendorf, UK). 500 μL saliva from all the participants was pooled together and re-aliquoted as 500 μL aliquots to be used for method development and quality

control (QC). Samples were aliquoted prior to storage to minimise freeze-thaw cycles. The saliva samples were stored at -80°C immediately after aliquoting to arrest biological activity.

#### *4.3.1.3 Sample pre-treatment*

The sample pre-treatment, method development and metabolic profiling analysis by UHPLC-IM-MS and the data analysis was carried out by the author (Aditya Malkar).

Saliva samples stored at -80°C were thawed at room temperature for 30 minutes prior to sample clean-up by protein precipitation. Each thawed saliva sample was vortexed for 30 seconds followed by ultra-sonication for 1 minute to breakdown mucous substances in saliva and to improve homogeneity. Protein precipitation was achieved by the addition of 1 mL acetonitrile to 500 µL of saliva i.e. 2:1 ratio of precipitant to saliva (Polson et al. 2003). The mixture was vortexed for 30 seconds followed by ultra-sonication for 1 minute and then subjected to centrifugation at 10,000 g for 10 minutes at ambient temperature. Precipitated proteins from the sample were removed as a pellet at the bottom of the microcentrifuge tube. The supernatant was transferred to a fresh microcentrifuge tube and evaporated to near dryness (~5 µL) using a Turbovap LV concentration workstation (Caliper Life Sciences, Hopkinton, MA, USA) prior to being reconstituted in 50 µL water/acetonitrile (95:05) (v/v) with 0.1% formic acid. The reconstituted saliva extract was placed in a 200 µL polypropylene micro-insert (Supelco, UK) in an autosampler vial (2 mL) for UHPLC-IM-MS analysis.

### 4.3.2 Instrumentation

The liquid chromatographic separations were performed on a Waters ACQUITY UPLC chromatograph with an in-built auto-sampler (Waters Corporation, Manchester, UK) fitted with a Waters ACQUITY BEH (bridged ethylene hybrid) C18 column (2.1 mm x 100 mm, 1.7 µm; Waters Corporation, Manchester, UK) maintained at 40°C. A VanGuard BEH C18 pre-column (Waters Corporation, Manchester, UK) was attached in series with the analytical column. The UPLC system was coupled with a Waters Synapt HDMS ion-mobility mass spectrometer (Waters Corporation, Manchester, UK). MassLynx 4.1 (Waters Corporation, Manchester, UK) was used for controlling the setup and for data acquisition. A Waters ACQUITY UPLC chromatograph (Waters

Corporation, Manchester, UK) coupled with a LTQ Orbitrap Mass Spectrometer (ThermoFisher Scientific, UK) was used to acquire accurate mass measurements on discriminant metabolites.

### 4.3.3 Method development

#### *4.3.3.1 UHPLC conditions*

Mobile phase A consisted of 0.1% aqueous formic acid (v/v) and mobile phase B was 0.1% formic acid in acetonitrile (v/v). The chromatographic gradient programme was optimised using a solution of L-carnitine and hydrocortisone (0.01 mg/mL). The mobile phase flow rate was set to 0.5 mL/min. The saliva extract (10 μL injected) was analysed by the optimised chromatographic gradient: 5% B (0-1 min), increased to 35% B (1-10 min). This was followed by a column clean-up phase built into the method to reduce carry over and condition the column for analysis of subsequent samples, in which the mobile phase was increased to 95% B (10-11 min) and maintained at 95% B (11-12 min) before returning to initial conditions (13-15 min).

#### *4.3.3.2 MS conditions*

Electrospray ionisation conditions for the Synapt HDMS, with the ion source operated in positive ion mode were: capillary voltage 3.0 kV; cone voltage 30 V; source temperature 120°C; desolvation temperature 300°C; desolvation gas, $N_2$ gas flow 600 L/hr; cone gas flow 30 L/hr. Mass spectra were acquired with a scan rate of 1 sec/scan with a 0.02 sec inter-scan delay. Conditions for LTQ-Orbitrap MS were: heated electrospray ionisation (HESI) source operated under positive polarity mode; source voltage 4.5 kV; capillary voltage 35V; capillary temperature 300ºC and sheath gas flow 45 L/hr.

#### *4.3.3.3 Optimisation of the ion-mobility separation*

The tri-wave (T-wave) drift cell conditions were set at 30 mL/min drift gas ($N_2$) with a variable travelling wave height ramp of 7.0-15.0 V and wave velocity of 300 m/s. The mass spectrum acquisition rate was 0.065 ms/scan. Optimisation of tri-wave parameters was carried out on-line by directly infusing 2 ng/mL solution of Leucine Enkephalin. The optimised variable travelling wave height ramp was set to 6.0-13.0 V. Saliva extract was analysed prior to and after optimisation of ion-mobility separation.

#### *4.3.3.4 Dynamic range determination and reproducibility*

Injection volumes (1-20 µL) of pooled saliva extract were analysed by UHPLC-IM-MS. Chromatographic peak areas of selected endogenous metabolite ions eluting at various retention times throughout the chromatogram were extracted and plotted against the injection volume of saliva.

The reproducibility of the analytical method was assessed using the optimised injection volume (10 µL) determined from the dynamic range experiment. Six pooled saliva extract samples were analysed consecutively by UHPLC-IM-MS. Chromatographic retention times and chromatographic peak areas of extracted endogenous metabolite ions were used to establish the reproducibility of the analytical method.

## 4.4 ANALYSIS OF BEFORE-EXERCISE AND AFTER-EXERCISE SALIVA SAMPLES

Participants exercised for 30 minutes on a cycle ergometer (Lode Corival, Groningen, The Netherlands) at a power output of 2 Watts/Kg of body mass of the participant. The exercise intervention and the sampling were carried out in an environmentally controlled chamber set at a temperature of 25$^{o}$C with a relative humidity set to 50%. Two saliva samples were collected per participant, one before exercise and the other 30 minutes after the end of the exercise.

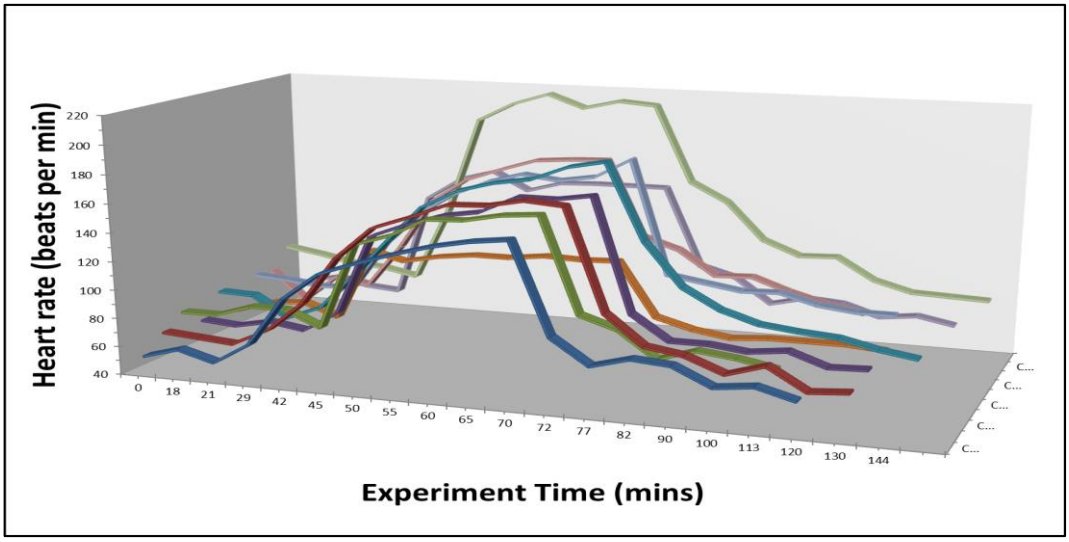



**Figure 4.1 Heart rate data depicting evidence of induced physiological stress** (Reproduced with permission) (Malkar et al. 2012)

Heart rate data of the participants was used as means to monitor the extent of physiological stress of the participants. Heart rate data was processed by Pareen Patel. **Figure 4.1** provides evidence of physiological stress being induced in all the participants due to exercise on the cycle ergometer.

### 4.4.1 Analytical protocol

The analytical protocol used for the analysis of samples is shown schematically in Section 3.3.7 - **Figure 3.2**. A series of five consecutive injections of a pooled saliva extract were run prior to the sample analysis to condition the chromatographic column. The analysis of replicate injections showed little retention time drift following this conditioning procedure. The sample list was randomised using Microsoft Excel prior to analysis and a pooled saliva extract was analysed after every 5 saliva samples as a quality control sample (QC). The sensitivity of the instrument was monitored on-line using the LockMass signal. A separate method blank study was carried out as a part of initial UHPLC-IM-MS method development for this work. Method blanks demonstrated no contamination due to polypropylene micro-centrifuge tubes or polypropylene micro-inserts used for this study. Method blanks also demonstrated no carry-over between saliva sample runs.

### 4.4.2 Data mining and multivariate statistical analysis

Data mining was carried out using MarkerLynx XS (Waters Corporation, Manchester, UK). The parameters used were retention time range 0-10 minutes; retention time tolerance was set to 0.1 min; mass range 50-500 Da; mass tolerance 0.05 Da; isotopic peaks were excluded for analysis; minimum spectral peak intensity was set to 50 counts. The resulting three dimensional data matrix containing arbitrarily assigned peak index (consisting of *m/z* and retention time pairs), sample names (observations) and their normalised chromatographic peak areas (variables) was exported to SIMCA-P+ (Umetrics, Sweden) for further data processing. Data was normalised to the total ion intensity within the particular sample for each feature (i.e. *m/z*_retention-time pair) using MarkerLynx XS (Waters Corporation, Macnhester, UK) using chromatographic peak area. This method of normalisation adjusts all the ion intensities so that the sum of intensities within each sample is the same. Multivariate statistical analysis, principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA), were carried out on normalised data to identify discriminant metabolites in saliva between the two groups of samples i.e. before and after exercise.

## 4.5 RESULTS AND DISCUSSION

### 4.5.1 Method development

The UHPLC chromatographic gradient programme used for the analysis of saliva extract was optimised using 0.01 mg/mL solution of L-carnitine and hydrocortisone, such that the highly polar metabolite, L-carnitine, was just retained on the column ($t_R$=0.52 min), whereas the low polarity hydrocortisone eluted towards the end of the 10 min chromatographic gradient ($t_R$=8.48 min) as shown in **Figure 4.2**. This was followed by a column clean-up and re-equilibration phase (10-15 minutes). The total cycle time was 15 minutes, which is suitable for high throughput global metabolic profiling workflows.

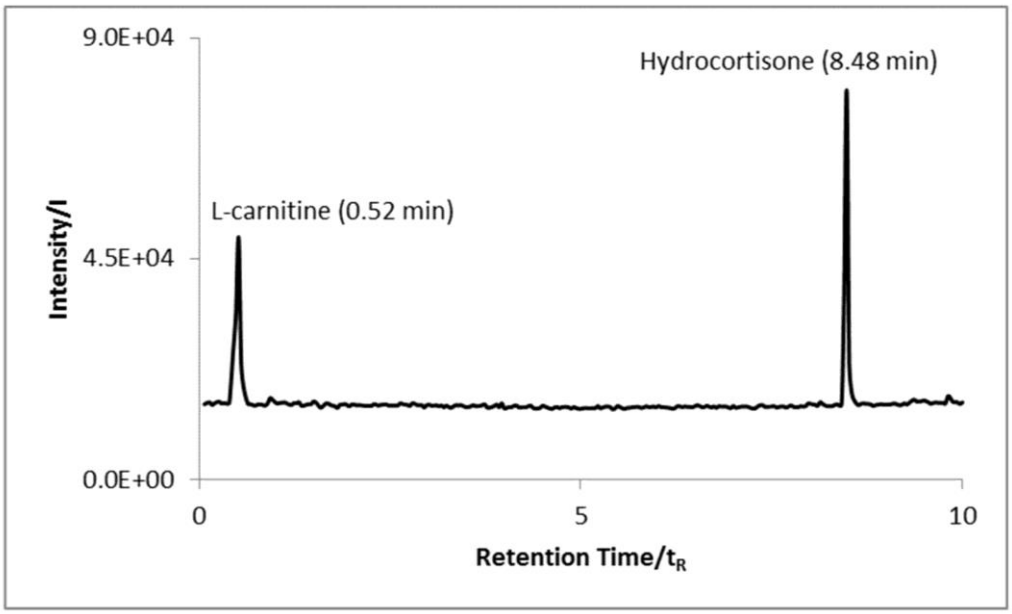



**Figure 4.2 Optimisation of UHPLC gradient based on L-carnitine and hydrocortisone solution (0.01 mg/mL).** (Data produced on Synapt HDMS)

Protein precipitation was the only sample clean-up procedure applied to the saliva prior to UHPLC-IM-MS analysis, in order to minimise the effect of discrimination and the unpredictable behaviour of unknown metabolites during sample preparation (Álvarez-Sánchez et al. 2012). The selection of endogenous metabolites for validating the method ensured that the matrix effects were accounted for during data processing.

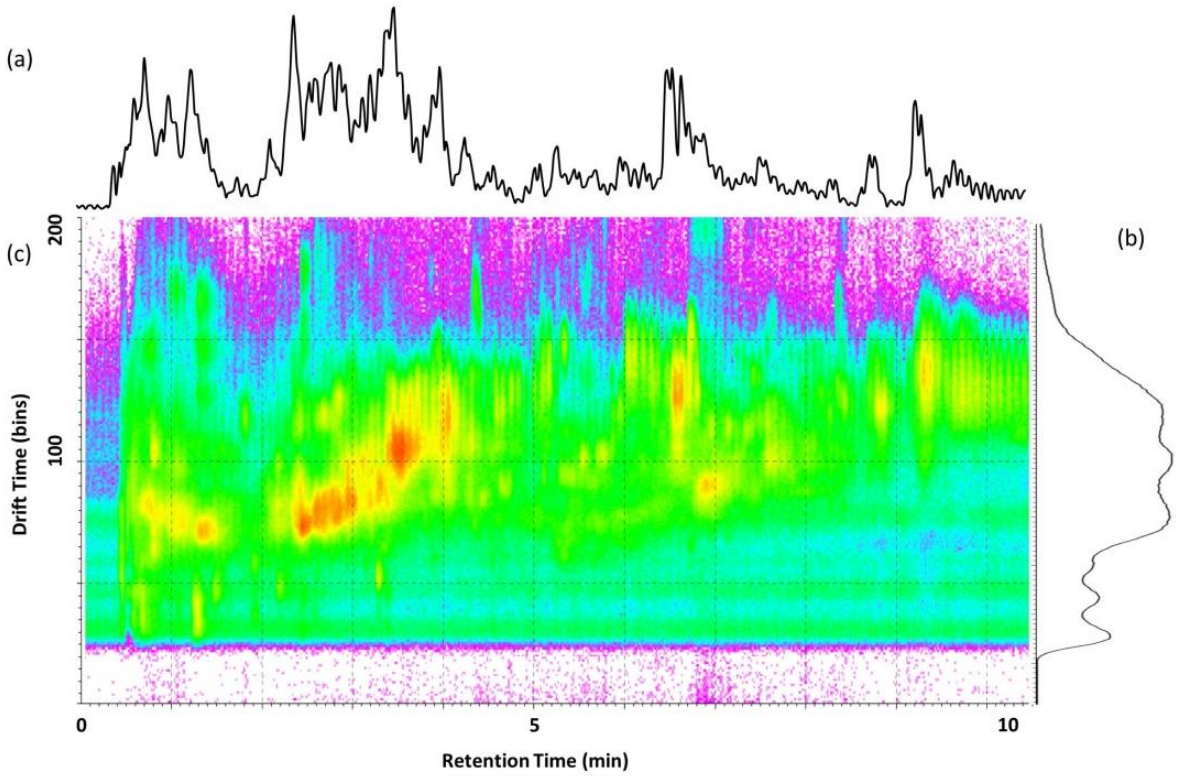



**Figure 4.3 UPLC-IM-MS metabonomic analysis of saliva extract. (a) UHPLC total ion chromatogram, (b) total ion mobility drift time plot summed over the entire chromatographic run, (c) retention time vs ion mobility drift time heatmap.** (Data produced on Synapt HDMS)

A typical UHPLC-IM-MS profile for a saliva extract is shown in **Figure 4.3** illustrating the complexity of the three-dimensional nested dataset obtained from the analysis. The two dimensional retention time vs ion mobility drift time plot shows multiple species contributing to single chromatographic peaks (**Figure 4.3** c), demonstrating that the combination of ion mobility with UHPLC allows further separation of co-eluting species. Ion mobility spectrum is plotted as bins, where each bin represents an acquired mass spectrum (65 μs/scan), with 200 bins covering the drift time range of 0-13 ms. Spectra were accumulated to yield 200 mass spectra and one ion mobility spectrum per second giving 6-7 data points across a typical UHPLC peak.

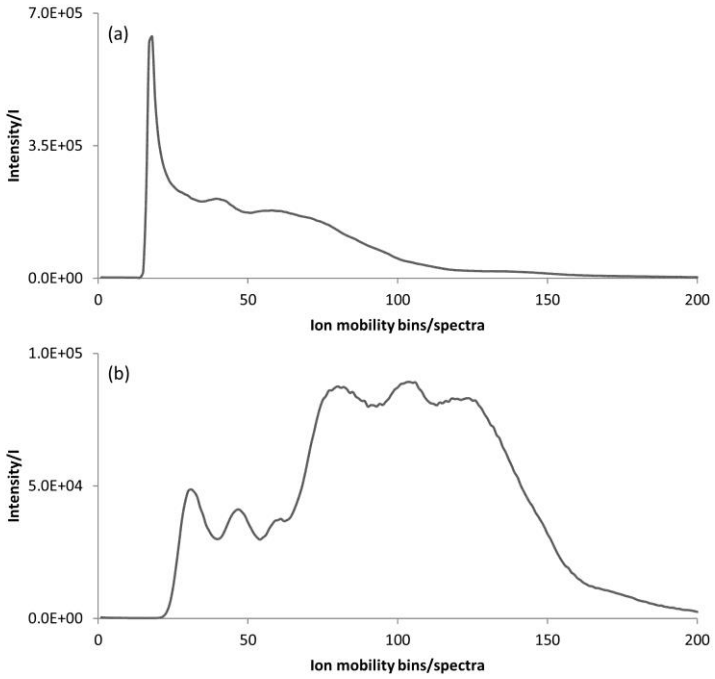



**Figure 4.4 UPLC-IM-MS analysis of saliva extract before and after optimisation of the T-wave IM parameters. (a) total ion mobility drift time plot (bins) prior to optimisation of IM parameters, (b) total ion mobility drift time plot (bins) after optimisation of IM parameters.** (Data produced on Synapt HDMS)

The major parameters that affect the separation of ions in the T-wave drift cell are IMS gas pressure, IMS wave velocity and IMS wave height. The ion mobility separation shown in **Figure 4.3** was improved by adjusting the variable wave height ramp as a wave gradient effectively increases the length of the IMS drift cell (Shvartsburg & Smith 2008).

**Figure 4.4 (a)** shows IM separation prior to optimisation, with a wave height ramp of 7.0-15.0 V. The majority of the components of the saliva extract have a drift time in the first 100 bins/spectra, associated with IM drift times of <2.6 ms. Ideally, only the first 20 bins/spectra and the last 20 bins/spectra should be empty with analytes distributed across the remaining bins/spectrum (0.13-11.7 ms) of an overall drift time plot for optimum separation. This can be seen in **Figure 4.4 (b)**, with an optimised wave height ramp of 6.0-13.0 V. Under the optimised UHPLC and ion mobility conditions, analytes that are unresolved and co-eluting in the retention time dimension may be resolved in the ion

mobility dimension as a result of the increased peak capacity of UHPLC-IM-MS as can be seen in **Figure 4.5**.

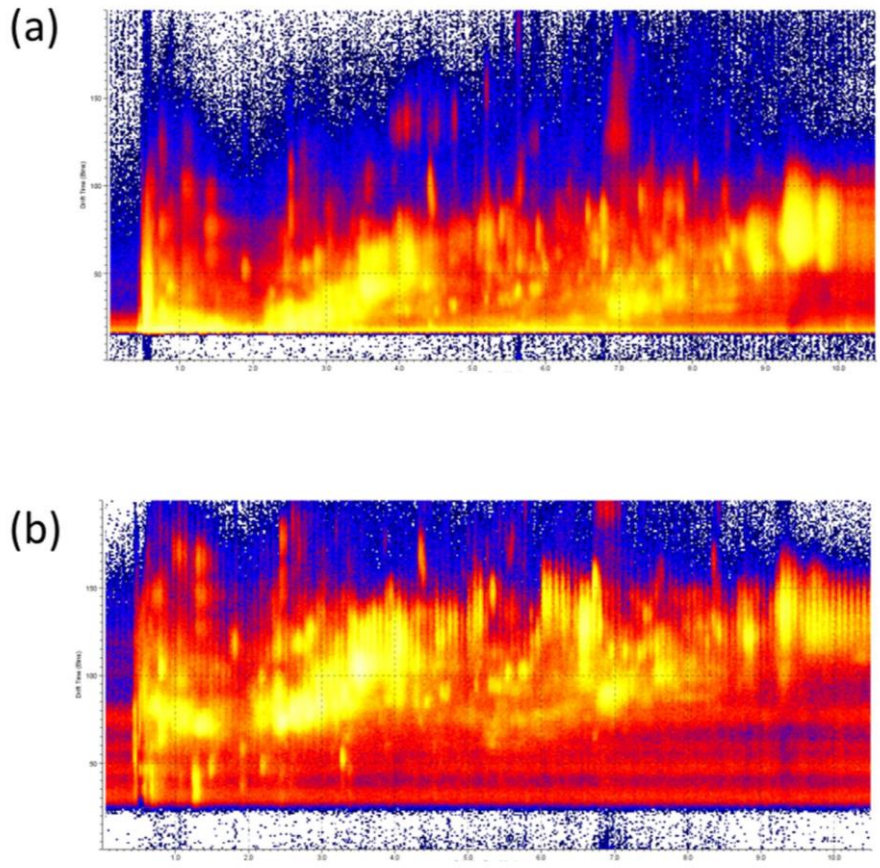



**Figure 4.5 Analysis of saliva extract before (a) and after optimisation (b) of ion mobility parameters shows improved separation of components in ion mobility dimension.** (Data produced on Synapt HDMS)

The effect of UHPLC injection volume was assessed for injection volumes in the range of 1-20 µL. Six randomly extracted metabolite ions, spread across the chromatographic range were selected to assess the dynamic range of the analytical system. Chromatographic peak areas showed a linear response for injection volumes in the range 1µL to 10µL, with the response levelling out above 15 µL for some metabolite ions (**Figure 4.6**). Hence, a 10 µL injection volume of saliva extract was used for subsequent analysis and global metabolic profiling studies as it lies within the dynamic range of the analytical method.

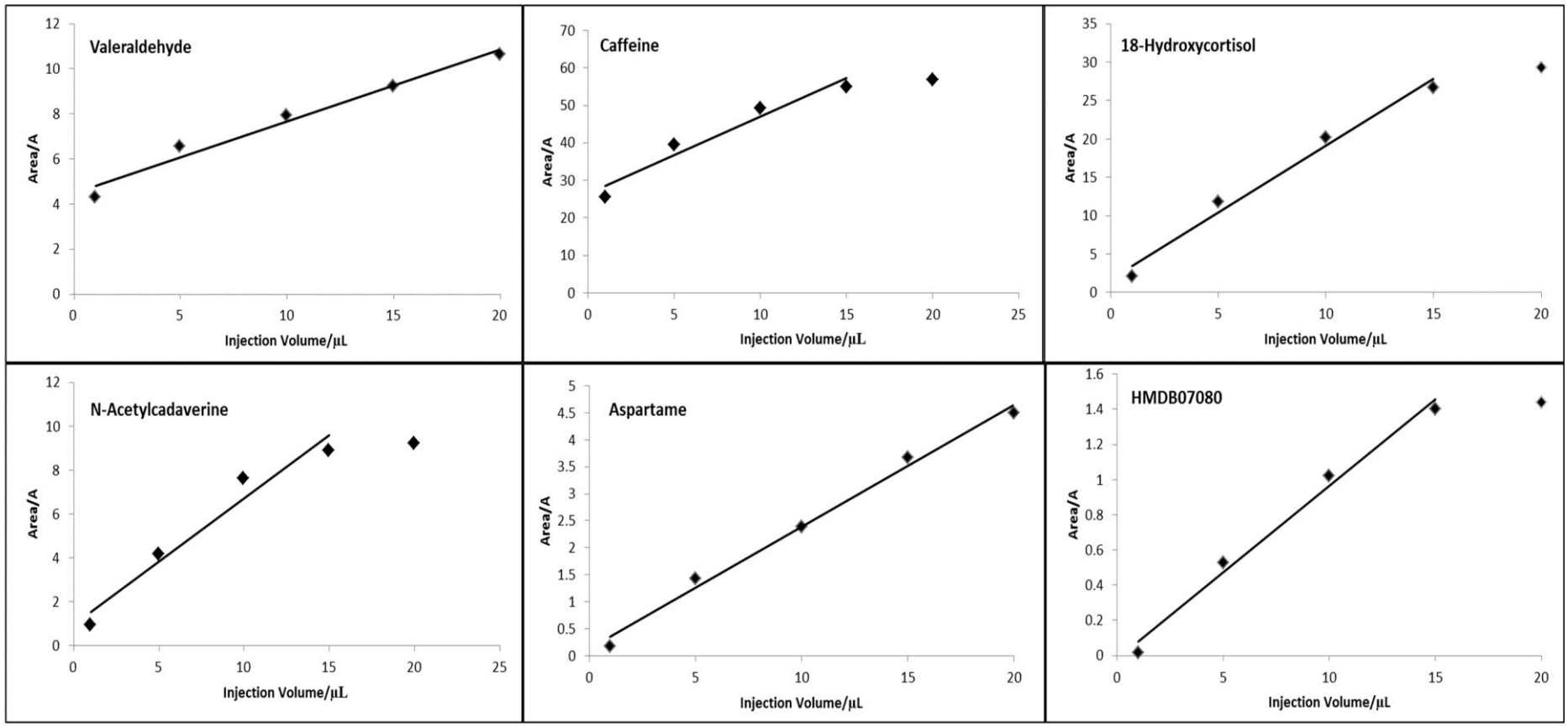



**Figure 4.6 Dynamic range analysis of the developed UPLC-IM-MS method. Injection volume vs chromatographic peak areas of six example endogenous metabolite ions with their HMDB assignments**

The results of a UHPLC reproducibility study (n=6) of peak area and retention time are summarised in **Table 4-1**. Data show relative standard deviation (RSD) less than 9% for chromatographic peak areas and less than 1% RSD for retention times across the range of the chromatogram. The data show that the method is consistent and reliable, which is required for measuring small differences in concentration of detectable metabolites in metabolic profiling studies.

**Table 4-1 Summary of reproducibility study of retention time and peak areas.**

| Tentative Assignment | m/z | Retention Time ($t_R$)/min | RSD% (Retention Time) | RSD % (Area) |
|---|---|---|---|---|
| **Valeraldehyde** | 104.1 | 0.57 | 0 | 8.78 |
| **N-Acetylcadaverine** | 125.1 | 1.92 | 0.85 | 4.13 |
| **Caffeine*** | 195.08 | 3.29 | 0.24 | 0.57 |
| **Aspartame** | 295.12 | 3.37 | 0 | 3.53 |
| **18-Hydroxycortisol*** | 379.21 | 3.6 | 0.22 | 8.71 |
| **HMDB07080** | 607.53 | 7.2 | 0 | 3.30 |

*Verification of IDs was carried out by comparison with analytical standards.

### 4.5.2 Targeted metabolite identification

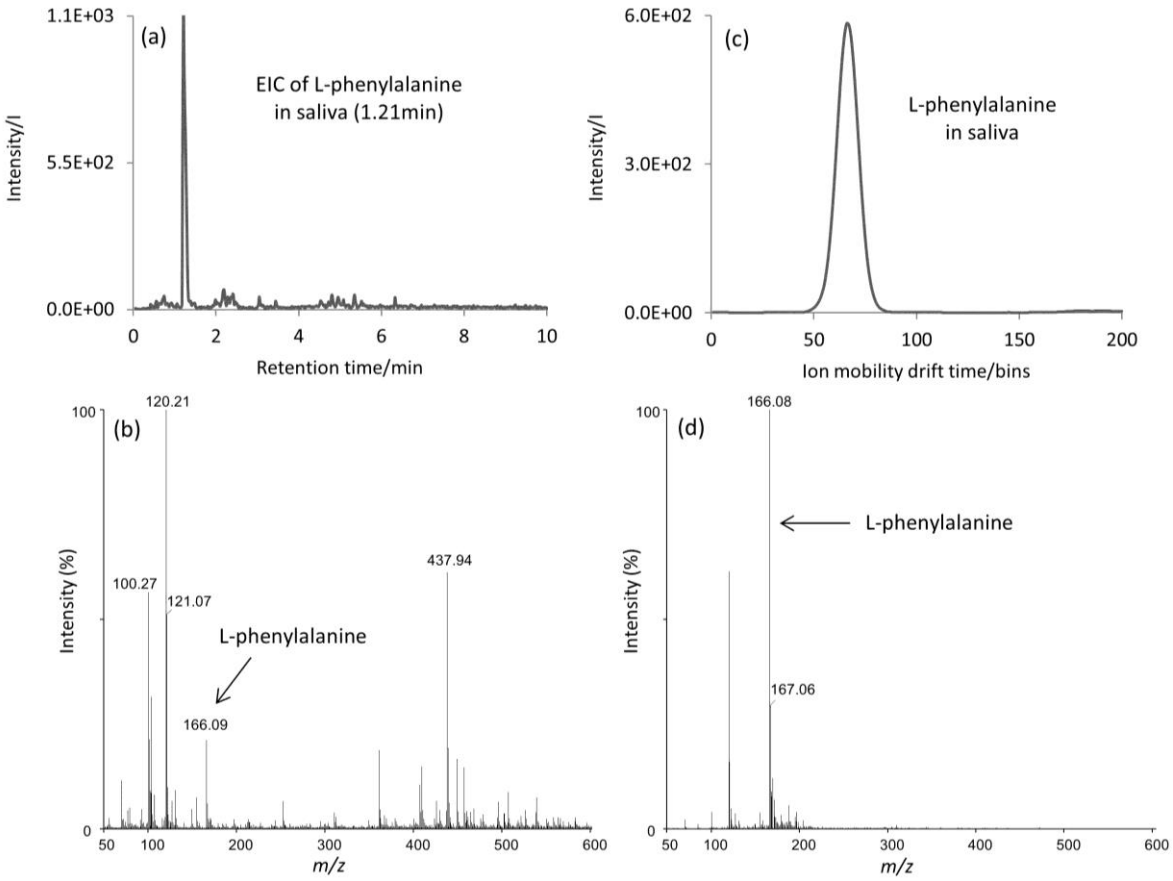



**Figure 4.7 UHPLC-IM-MS targeted analysis of L-phenylalanine (m/z 166.08) in a saliva extract (a) extracted ion chromatogram for L-phenylalanine (m/z 166.08±0.05), (b) mass spectrum at 1.21±0.03 min, (c) selected ion mobility spectrum for L-phenylalanine (*m/z* 166.08±0.05, 1.21±0.03 min), (d) mass spectrum at 1.21±0.03 min with ion mobility bins 64-68 selected, showing spectral quality improvement.** (Data produced on Synapt HDMS)

The nested UHPLC-IM-MS dataset collected from a saliva sample maybe interrogated for the presence of a targeted analyte. **Figure 4.7** shows an example of the extracted ion chromatogram (**Figure 4.7 a**) and ion mobility data (**Figure 4.7 c**) for a targeted analyte, L-phenylalanine present in saliva. Both ion mobility drift time and retention time showed a close match with a standard indicating positive identification of the L-phenylalanine and that matrix effects do not adversely affect retention times. The shape of the chromatographic and ion mobility peaks (**Figure 4.7 a and c**) indicates the presence of a

single species. Selection of the appropriate retention time and ion mobility region (retention time 1.21 min, ion mobility bin numbers 64-68) for L-phenylalanine enhances the relative mass spectral response of the protonated amino acid (**Figure 4.7 b and d**).

### 4.5.3 Global metabolite profiling analysis of before- and after-exercise saliva samples

The UHPLC-IM-MS analysis of saliva extracts generates a complex dataset consisting of mass-to-charge ratio, chromatographic retention time and ion mobility drift time. IM separation being orthogonal to that of liquid chromatographic separation and mass-to-charge ratio increases the peak capacity as a result of the separation of species of the same *m/z* or retention time in the mobility dimension. The resulting multidimensional UHPLC-IM-MS datasets present a challenge for the multivariate statistical techniques commonly employed for global metabolite profiling studies. The powerful software available for de-convoluting UHPLC-MS data is not equipped to account for ion-mobility dimension while compiling data matrix for multivariate analysis. The complexity of the data was therefore reduced in the initial stages of the analysis by collapsing the IM dimension.

Key factors in the analysis of complex biological metabolomic samples include chromatographic column conditioning, number of quality control samples and run order (Want et al. 2010). These factors were evaluated using pooled saliva samples collected from male volunteers before-exercise and after-exercise. **Figure 4.8** shows unsupervised principal component analysis (PCA) of all the randomised before-exercise and after-exercise saliva samples and replicates of quality control (QC) samples consisting of an extract of a pooled saliva sample, which was analysed with every batch of 5 samples. The model was based on a total of 4078 variables, which were normalised and Pareto scaled, such scaling being the preferred option while dealing with metabolomic data obtained by MS (van den Berg et al. 2006). Clustering of quality control (QC) samples (circled in **Figure 4.8**) provides evidence of consistent and reliable data quality adequate for metabolic profiling studies.

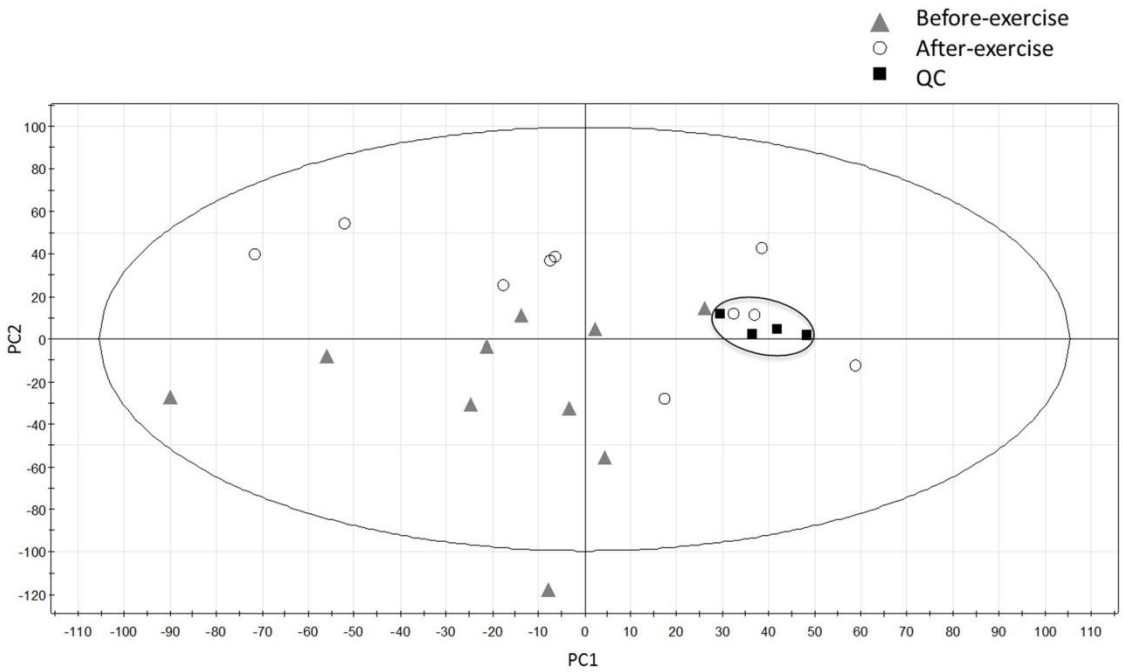



**Figure 4.8 Principal component analysis (PCA) of all saliva samples from the study (before-exercise, after-exercise and QC samples) Component 1 (14.17%) & Component 2 (12.68%).**

The unsupervised PCA approach indicates partial separation between the two classes of samples (**Figure 4.8**) with before-exercise samples generally having lower or negative principle component 1 (PC1) and 2 (PC2) scores while after-exercise samples showing higher PC1 and PC2 scores. Supervised approaches, such as partial least squares (PLS) and orthogonal partial least squares discriminant analysis (OPLS-DA) models, are often applied in metabonomics to uncover relationships buried in the background of larger inherent effects (Fonville et al. 2010). In biological systems, for instance, it is often the case that several factors contribute to the metabolic state of an organism e.g. age, gender, nutrition, environment etc. Hence, to focus on the biological question of interest, physiological stress in this case, and to disregard other sources of variation an OPLS-DA model was used. A total of 3067 normalised, Pareto scaled variables were used to construct the model, which was able to distinguish between the two classes of samples (**Figure 4.9**).

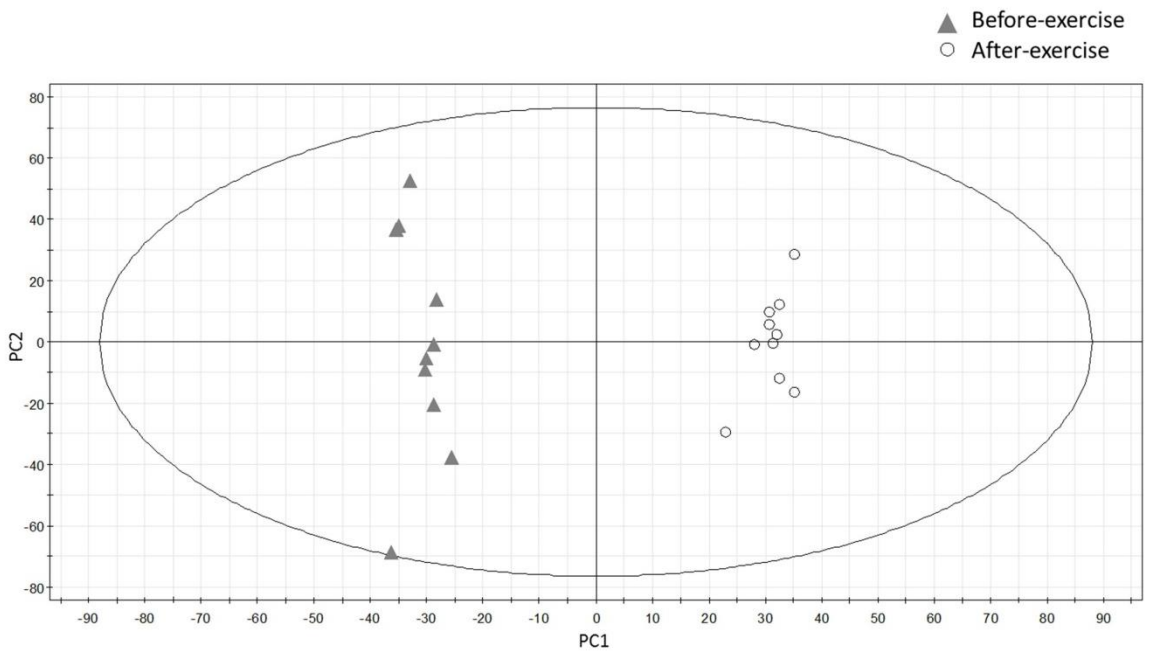



**Figure 4.9 A supervised orthogonal partial least squares – discriminant analysis (OPLS-DA) model, distinguishing between the two classes of samples (before-exercise and after-exercise saliva samples) Component 1 (14.44%) & Component 2 (8.26%).**

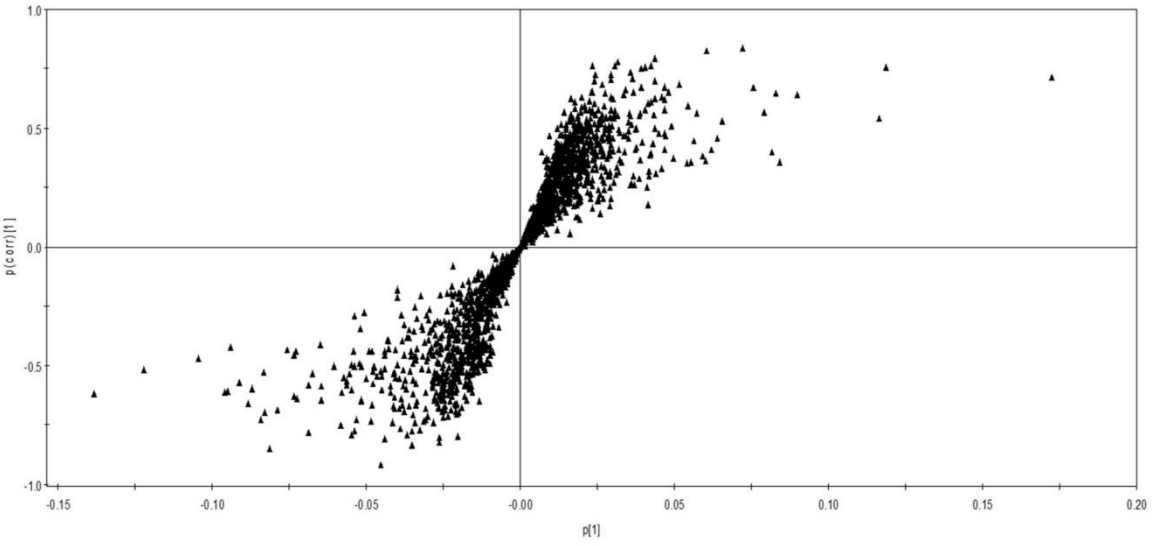



**Figure 4.10 The loadings s-plot obtained from the OPLS-DA model for the metabolomic comparison between pre-exercise and post-exercise saliva samples.**

A loadings S-plot was constructed using the OPLS-DA model to segregate discriminant metabolites (**Figure 4.10**). A number of variables were selected from extreme ends of the s-plot and their normalised trend plots were compared as initial screening for potential biomarker ions. Discriminant features were extracted from the S-plot and were verified by comparing their extracted ion chromatogram (EIC) from raw data. An overview of the top 6 discriminant metabolites obtained from the analysis is provided in **Table 4-2**.

**Table 4-2 A summary of top 6 discriminant metabolite ions.**

| **Retention time ($t_R$)/min** | ***m/z**** | **Regulation** | **Tentative Elemental Composition(s)*** | **ppm error** | **Fold change** | **RSD from QC samples (%)** |
|---|---|---|---|---|---|---|
| 0.8 | 86.0962 | Up-regulated | $C_5H_{11}N$ | 2.9 | 2 | 8.3 |
| 1.3 | 100.0755 | Up-regulated | $C_5H_9ON$ | 1.9 | 16 | 3.9 |
| 0.7 | 213.1233 | Up-regulated | $C_{10}H_{16}O_3N_2$ | 0.6 | 2 | 3.4 |
| 1.3 | 120.0033 | Down-regulated | $C_8H_9N$ | 2.3 | 46 | 1.2 |
| 1.8 | 185.0252 | Down-regulated | $C_9H_{16}O_2N_2$ | 0.1 | 2 | 5.1 |
| 1.4 | 217.1046 | Down-regulated | $C_6H_{12}O_3N_6$ | 1.3 | 6 | 12.9 |

**m/z* and tentative elemental compositions obtained from UHPLC-Orbitrap-MS data analysed by XCalibur Software (ThermoFisher Scientific)

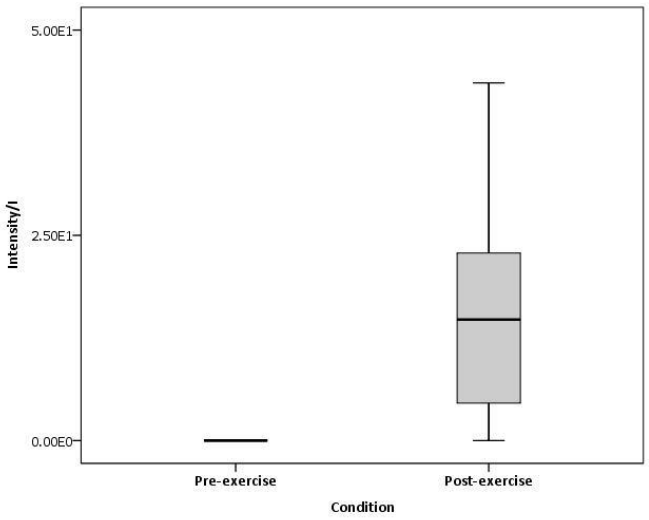



**Figure 4.11 Box and whisker plot for potential biomarker ion *m/z* 100.07±0.02**

The box and whisker plot for one of the potential biomarker ions 100.07±0.02 in **Figure 4.11** showing clear statistical difference in the normalised chromatographic peak areas from samples obtained before and after exercise.

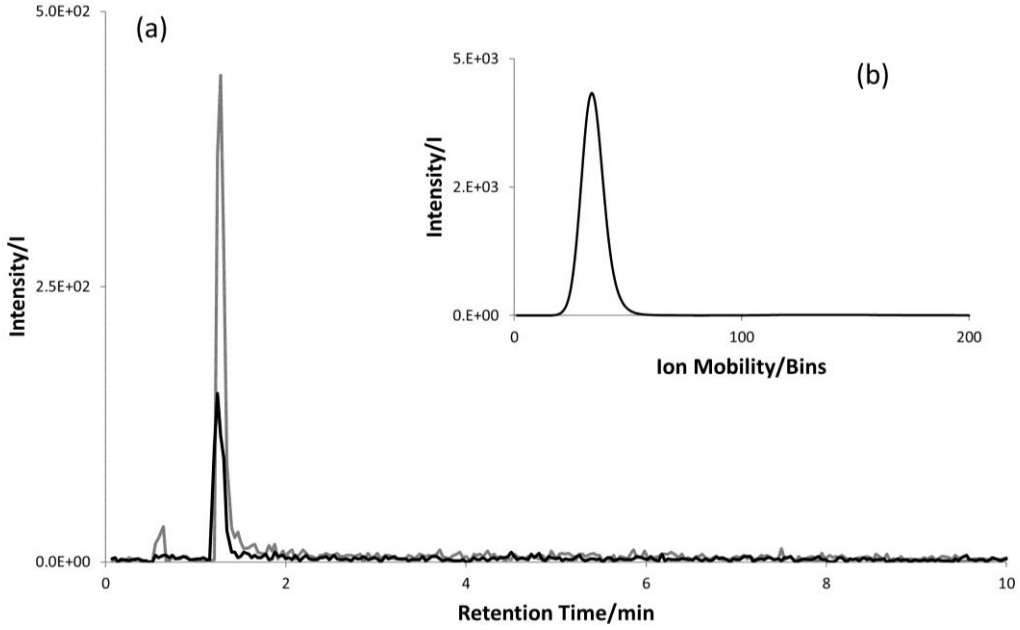



**Figure 4.12 Example raw data for selected potential biomarker ion *m/z* 100.07±0.02: (a) extracted ion chromatogram (EIC) showing change in intensity between before-exercise and after-exercise samples. Black trace denotes before-exercise and grey trace denotes after-exercise EICs. (b) selected ion mobility spectrum for *m/z* 100.07±0.02 showing a single sharp peak, indicating the presence of a single**

***m/z* 100.07 species at retention time of 1.31 minutes.** (Data produced on Synapt HDMS)

**Figure 4.12 (a)** shows an example of the raw data for selected potential biomarker ion *m/z* 100.07 ± 0.02 showing up-regulation of this ion with a retention time of 1.31 min. The retention time and *m/z* data for this molecule were used to extract the selected ion mobility spectrum shown in **Figure 4.12 (b)**. The single symmetrical IM spectrum peak indicates the presence of a single component. Identification of this component would therefore require the candidate molecule to match the ion mobility as well as the retention time and *m/z* characteristics.

Accurate mass measurement data and isotopic ratio data on the discriminant metabolite ion were obtained by analysing a pooled saliva extract sample using high resolution mass spectrometry. The molecular formula, $C_5H_9NO$, was determined from the accurate mass and isotopic pattern. The generated molecular formula was used to search electronic sources including ChemSpider and the Human Metabolome Database (HMDB) (Wishart et al. 2013) for possible candidate metabolites. The number of candidate molecules was reduced by application of biological knowledge about the candidate metabolites, such as the biological availability of candidate metabolite in human saliva; δ-valerolactam and N-methyl-2-pyrrollidinone were shortlisted as possible matches. Authentic standards of the shortlist of candidate metabolites were analysed using UHPLC-IM-MS and by UHPLC-MS/MS.

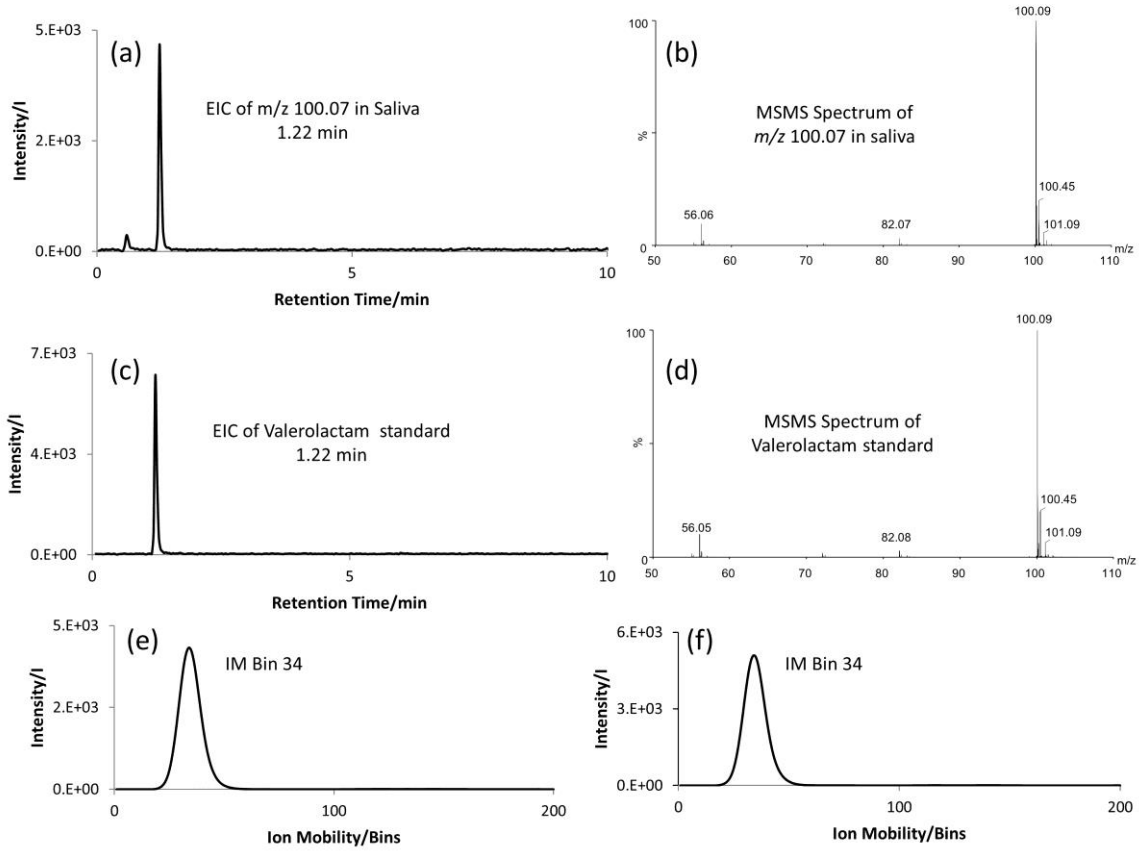



**Figure 4.13 Identification of discriminant metabolite (a) extracted ion chromatogram for *m/z* 100.07 in saliva, (b) MS/MS spectrum for *m/z* 100.07 in saliva, (c) extracted ion chromatogram for δ-valerolactam standard, (d) MS/MS spectrum of δ-valerolactam standard, (e) selected ion mobility spectrum for *m/z* 100.07, (f) selected ion mobility spectrum for δ-valerolactam.** (Data produced on Synapt HDMS)

The retention time, MS/MS spectrum and ion mobility drift time of the biomarker ion *m/z* 100.07±0.02 in the pooled saliva samples matched that of δ-valerolactam (**Figure 4.13**). This provides strong evidence for the assignment of this molecule as a potential biomarker of exercise induced physiological stress. δ-valerolactam is included in the human metabolome database (reference number HMDB11749) as a known metabolite present in saliva (Wishart et al. 2013). It has been shown in mouse models to result from the metabolism of cadaverine (Callery & Geelhaar 1984), which is known to be present at part per million levels in human saliva as a potential source of δ-valerolactam (Cooke et al. 2003). The physiological role of δ-valerolactam is not well understood and the effects of exercise and other stresses on its metabolism are at present unknown.

## 4.6 CONCLUSIONS

This study demonstrates proof of principle in a pilot study with a small number of participants that ion mobility data may be acquired within the timescale of an UHPLC-MS experiment for metabonomic analysis of saliva. The combination of UHPLC with IM-MS provides an added dimension enabling enhanced separation of co-eluting components of complex metabonomic samples on the basis of retention time, ion mobility and mass-to-charge ratio in a single 10-minute chromatographic run.

The results show the potential of the UHPLC-IM-MS technique for the targeted and the global metabolic profiling analysis of saliva and other biological fluids, and for the subsequent biomarker discovery process. In this study, δ-valerolactam has been identified as a potential biomarker in saliva of exercise-induced physiological stress on the basis of retention time, MS/MS spectrum and ion mobility drift time compared to a δ-valerolactam standard. The advantages of ion mobility spectrometry with UHPLC-MS for discovery and identification of metabolites from metabolite profiling studies have been demonstrated. This work shows the potential of hyphenated ion mobility mass spectrometry for non-hypothesis driven discovery-based metabolomic study.

## 4.7 CHAPTER FOUR REFERENCES


Álvarez-Sánchez, B.; Priego-Capote, F.; Luque de Castro, M.D. *J. Chromatogr. A* **2012,** 1248, 178–181.

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

Callery, P.S.; Geelhaar, L.A. *J. Neurochem.* **1984,** 43, 1631–1634.

Chiappin, S.; Antonelli, G.; Gatti, R.; De Palo, E.F. *Clin. Chim. Acta* **2007**, 383, 30–40.

Cooke, M.; Leeves, N.; White, C. *Arch. Oral Biol*. **2003**, 48, 323–327.

Creaser, C.S.; Griffiths, J.R.; Bramwell, C.J.; Noreen, S.; Hill, C. A.; Thomas, C.L.P. *Analyst* **2004**, 129, 984.

DeCaro, J. A *Am. J. Hum. Biol*. **2008**, 20, 617–619.

Dwivedi, P.; Wu, P.; Klopsch, S.J.; Puzon, G.J.; Xun, L.; Hill, H.H. *Metabolomics* **2007**, 4, 63–80.

Fonville, J.M.; Richards, S.E.; Barton, R.H.; Boulange, C.L.; Ebbels, T.M.D.; Nicholson, J.K.; Holmes, E.; Dumas, M.E. *J. Chemom*. **2010**, 24, 636–649.

González, D.; Marquina, R.; Norelis, R. *Res. Sport. Med.* **2008**, 16, 128–137.

Harry, E.L.; Weston, D.J.; Bristow, A.W.T.; Wilson, I.D.; Creaser, C.S. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **2008**, 871, 357–361.

Kanu, A.B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill H.H.Jr. *J. Mass Spectrom*. **2008**, 43, 1–22.

Lewis, G.D.; Farrell, L.; Wood, M.J.; Martinovic, M.; Rowe, G.C.; Souza, A.; Cheng, S.; Mccabe, E.L.; Shi, X.; Deo, R.; Roth, F.P.; Asnani, A.; Rhee, E.P.; Systrom, D.M.; Semigran, M.J.; Vasan, R.S.; Carr, S.A.; Wang, T.J.; Sabatine, M.S.; Clish, C.B.; Gerszten, R.E. *Sci Transl Med*. **2010**, 2, 33

Malkar, A.; Martin, H.J.; Patel, P.; Turner, M.A.; Watson, P.; Reid, H.J.; Maughan, R.J.; Thomas, C.L.P.; Creaser, C.S. **2012**. METABOLIC PROFILING OF SALIVA USING UPLC-IM-MS APPLIED TO MONITORING THE EFFECTS OF PHYSIOLOGICAL STRESS, in: ASMS 60th Conference on Mass Spectrometry and Allied Topics. Vancouver, Canada.

Papacosta, E.; Nassis, G.P. *J. Sci. Med. Sport* **2011**, 14, 424–434.

Perogamvros, I.; Owen, L.J.; Newell-Price, J.; Ray, D.W.; Trainer, P.J.; Keevil, B.G. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **2009**, 877, 3771–3775.

Pollard, T.M. *Am. J. Hum. Biol*. **1995**, 7, 265–274.

Polson, C.; Sarkar, P.; Incledon, B.; Raguvaran, V.; Grant, R. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci*. **2003**, 785, 263–275.

Shvartsburg, A. A.; Smith, R.D. *Anal. Chem*. **2008**, 80, 9689–9699.

Swartz, M.E. *J. Liq. Chromatogr. Relat. Technol*. **2005**, 28, 1253–1263.

Want, E.J.; Wilson, I.D.; Gik, H.; Theodoridis, G.; Plumb, R.S.; Shockcor, J.; Holmes, E.; Nicholson, J.K.; Gika, H. *Nat. Protoc*. **2010**, 5, 1005–1018.

Wishart, D.S.; Jewison, T.; Guo, A.C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.,; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. *Nucleic Acids Res.* **2013**, 41, D801–D807.

# CHAPTER FIVE:

## CONCLUSIONS AND FURTHER WORK

## 5.1 THESIS OVERVIEW

This thesis describes analytical methods for the detection of metabolic biomarkers of disease and physiological change. The research was directed towards the development and application of novel analytical techniques, based on mass spectrometry and ion mobility spectrometry, and bioinformatics strategies for the identification of diagnostic/prognostic biomarkers. Methods have been developed based on UHPLC-MS alone and in combination with ion mobility spectrometry (UHPLC-IM-MS) for untargeted metabolite profiling of biofluids such as saliva and mouse plasma. The applications of these methods include investigations on biomarkers of the efficacy of colorectal cancer preventative treatment, diagnostic biomarkers of asthma and physiological stress.

The individual chapters are summarised in the following sections and directions for further work are proposed.

### 5.1.1 Summary of chapter one

Chapter one covers the theoretical principles of ion mobility spectrometry (IMS), mass spectrometry (MS) and multivariate statistical analysis (MVA) techniques. An account of modern day metabolomics, physiology and the application of saliva as a diagnostic medium is presented. The chapter also provides a review of analytical techniques used in metabolite profiling of biofluids with particular emphasis on salivary metabolite profiling.

### 5.1.2 Summary of chapter two

Chapter two demonstrates the potential of tri-wave ion mobility spectrometry in combination with ultra-high performance liquid chromatography and mass spectrometry for the metabolic analysis of mouse plasma. Untargeted metabolite profiling analysis of mouse plasma was carried out with minimal sample pre-treatment, involving only protein precipitation, to assess the chemopreventive efficacy of rice bran in colorectal cancer. Plasma derived from $APC^{min}$ mice, a well-established model for colon carcinogenesis, was used for untargeted metabolite profiling and potential biomarker discovery. The UHPLC-IM-MS method was shown to yield improvements compared to conventional metabolite profiling methods based on HPLC-MS in the following manner:

- The UHPLC-IM-MS method offered increased analytical space for untargeted metabolite profiling of mice plasma. This was achieved by a multi-dimensional separation of analytes based on UHPLC retention times, ion mobility and mass-to-charge ratios in a single 10 minute chromatographic run.
- The cycle time for the analysis of samples was 15 minutes with column clean-up and conditioning phases built into the UHPLC gradient method, which is ideal for high throughput analysis of metabolomic samples.
- The improvement of mass spectral data quality obtained by the removal of polyethylene glycol (PEG) contamination from the samples using post-acquisition processing was successfully demonstrated.
- Potential biomarker ions of efficacy of rice bran supplementation in colorectal cancer were identified.

### 5.1.3 Summary of chapter three

Chapter three presents the development and application of a robust methodology based on UHPLC-MS for salivary metabolite profiling and the discovery of potential biomarkers of asthma. The specific outcomes from this study are:

- A method has been developed and established for metabolite profiling of saliva using UHPLC-MS.
- The analytical method has been characterised (i.e. validated) on the basis of the reproducibility and dynamic range of the method.
- Unsupervised multivariate data analysis methods have been utilised after logical data reduction procedures to reduce the complexity of the very large dataset obtained from untargeted metabolite profiling studies.
- Ten potential metabolite biomarker ions of asthma have been identified from the moderate asthmatics in the study with their tentative elemental composition assignments.
- A predictive model based on partial least squares – discriminant analysis (PLS-DA) has been constructed using the ten discriminant ions, which shows good predictive capability for moderate asthmatics and controls. The predictive model was used to test its ability to classify mild asthmatic population.

- Current methods for diagnosis of asthma - FEV1%, methacholine challenge and blood eosinophil level - all use invasive sample collection methods and thus cause patient discomfort. The developed method has been proposed to be used in conjunction with existing methods or as a preliminary screen for the diagnosis of asthma as passive drool sampling of saliva is carried out non-invasively.

### 5.1.4 Summary of chapter four

The potential of ion mobility in combination with existing ultra-high performance liquid chromatography and mass spectrometry methods for the untargeted metabolite profiling of saliva is demonstrated in Chapter four. The method developed is based on protein precipitation and UHPLC-IM-MS, along with multivariate data processing strategies for global metabolite profiling of saliva. The specific outcomes from the study were:

- The developed method utilises <1 mL volume of saliva (i.e. 500 µL), which is easily obtainable by passive drool method. The sample pre-treatment involves only protein precipitation as a clean-up stage.
- The chromatographic gradient for UHPLC-IM-MS analysis of saliva was optimised while maintaining the analysis cycle time at 15 minutes, which is advantageous for high throughput metabolite profiling analyses.
- Ion mobility (IM) separation was optimised for untargeted salivary metabolite profiling by UHPLC-IM-MS.
- The developed UHPLC-IM-MS method was characterised (i.e. validated) by assessing the reproducibility and dynamic range of the analytical method using endogenous metabolite ions thus accounting for matrix effects. These factors are important considerations for identification of up-regulated or down-regulated metabolite ions.
- The advantages of using ion mobility for targeted analysis and its advantages for enhancing the selectivity of target metabolites have been illustrated.
- The developed method has been successfully applied to a metabolite profiling study to discover potential biomarker ions of physiological stress.
- Valerolactam has been identified as a potential biomarker of physiological stress from saliva by comparison of retention time ($t_R$), ion mobility drift time and MS/MS spectra with purchased standard of δ-valerolactam.

## 5.2 FUTURE DIRECTIONS

### 5.2.1 Validation of biomarkers

Studies described in this thesis have been of pilot study nature with relatively small number of participants/patients as compared to large scale metabolomic studies. This has been the case as the research described focuses more on method development for profiling of biofluids. Although the analytical methods described in this research have been validated, the output, which is the identified potential biomarker ions, is yet to be validated.

Many studies in metabolomics suffer similar problems with relatively small number of samples and thus await validation of the outputs, the accepted strategy for validation being to analyse a larger sample set and to use quantitative MS methods. All but one of the potential biomarker ions identified from the three studies (colorectal cancer, asthma and physiological stress) described in this thesis need to be identified by comparison with analytical standards using criteria such as UHPLC retention time, accurate mass, MS/MS spectrum and ion mobility drift time as illustrated by the example of valerolactam in Chapter Four.

It is proposed that the quantification of the identified metabolites such as valerolactam be carried out using a multiple reaction monitoring (MRM) method by using UHPLC-MS/MS with an internal standard with a larger cohort of samples. The internal standard in this case ideally would be deuterated-valerolactam or a molecule/derivative from piperidine family such as Methimazole (Anon 2013).

### 5.2.2 Combined metabolomics and proteomics sample pre-treatment

The integration of all '–omics' sciences under the umbrella of systems biology is gaining interest worldwide. Metabolomic and proteomic studies, however, typically involve different sample preparation steps. This is a significant roadblock in combined metabolomic and proteomic studies especially where sample quantity is limited or precious.

The sample pre-treatment by protein precipitation employed in this research is a straightforward process which removes proteins from saliva by a simple acetonitrile crash method. The interference or contamination of the sample with external reagents or compounds is minimal as only an organic solvent – acetonitrile is used during sample pre-treatment. A major potential advantage offered by this method is, it can facilitate designing of metabolomic and proteomic studies from the same sample, as the metabolomics part of the project will utilise the supernatant obtained from the sample pre-treatment whereas the precipitated proteins pellet remaining behind can be digested and utilised for proteomic studies. Protein precipitation employed during this research is capable of extracting polar and non-polar metabolites with minimal degradation and efficiently extract proteins from complex biofluid such as saliva for both metabolomic and proteomic studies. This presents a viable opportunity to combine sample preparation for integrated metabolomic and proteomic studies.

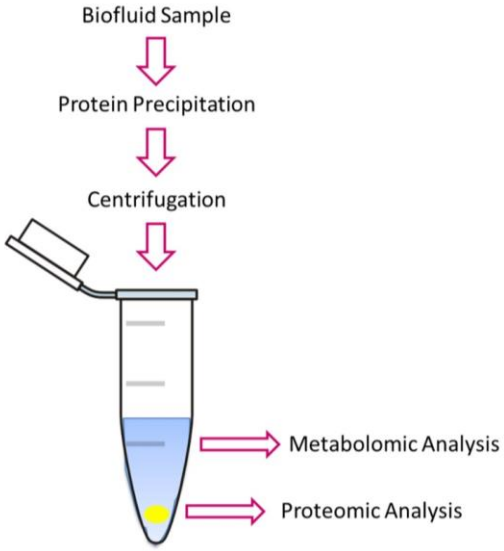



**Figure 5.1 Combined sample pre-treatment strategy for metabolomics and proteomics.**

### 5.2.3 FAIMS for metabolomics

Field asymmetric waveform ion-mobility spectrometry (FAIMS), which is also known as differential mobility spectrometry (DMS), is an ion mobility based electrophoretic technique (Kolakowski & Mester 2007; Nazarov 2012). High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) is a gas phase separation technique that exploits the difference in the mobility of an ion in low and high electric fields to separate ions in mixture. FAIMS has previously been demonstrated to be a highly effective separation step prior to mass spectral analysis (Brown et al. 2010). This technique is yet to be used for global metabolic profiling analysis. With instrumental developments and data collection strategies in FAIMS, there is a potential for this technology to be used for metabolite profiling analysis.

Preliminary work was carried out on method development for metabolite profiling of saliva by UHPLC-FAIMS-MS. Sample collection and pre-treatment of saliva was carried out according to the established methods described in this thesis. Analysis was carried out using a prototype miniaturised ultra-FAIMS device (Owlstone, Cambridge, UK) combined with a UHPLC-MS setup. The UHPLC-MS setup consisted of an Agilent 1200 series HPLC system interfaced with an Agilent 6230 time-of-flight mass spectrometer (Agilent Technologies, Santa Clara, CA, USA).

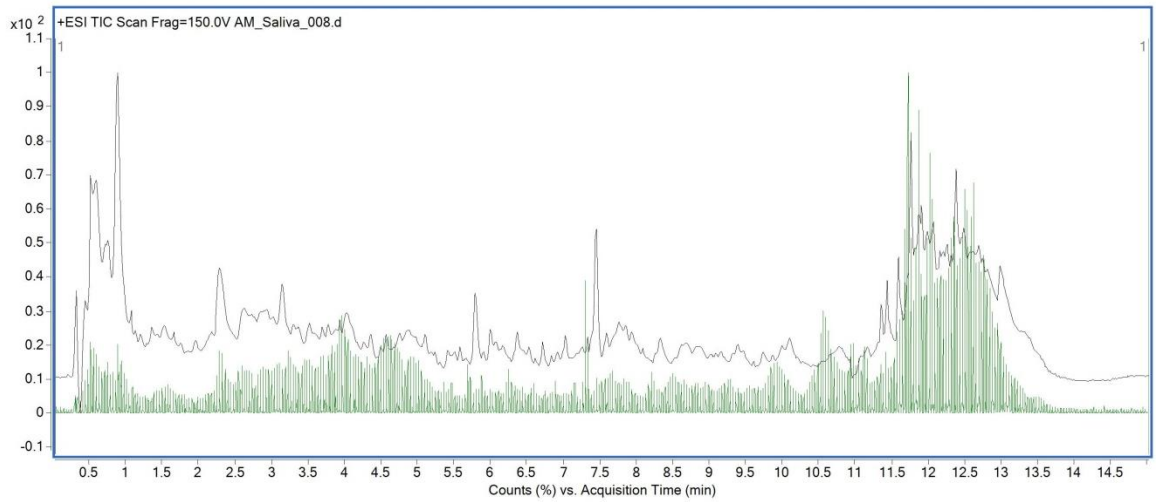



**Figure 5.2 UHPLC-FAIMS-MS analysis of saliva extract. Black trace denotes UHPLC-MS analysis of saliva extract and green trace denotes the analysis of saliva using UHPLC-FAIMS-MS.**

**Figure 5.2** shows the comparison between analysis of saliva extract using UHPLC-MS and UHPLC-FAIMS-MS. The miniaturised ultra-FAIMS device used enables FAIMS scan rates of <2 seconds which enable c.a. 4 scans within a UHPLC chromatographic peak (assuming ~8 second peak width). However, since the miniaturised ultra-FAIMS device and mass spectrometric data acquisition software were controlled by separate software, the start points of the analysis by the FAIMS module and the mass spectrometer module could not be synchronised over multiple runs. As a result due to these limitations of the software, the analysis of saliva extract by UHPLC-FAIMS-MS could not be further investigated. However, with future software developments, the utility of FAIMS in conjunction with UHPLC-MS in metabolic profiling studies could be explored.

## 5.3 SUMMARY OF THESIS

The research presented in this thesis has shown the potential of the use of a non-invasively obtained biofluid – saliva- in disease diagnosis by discovery of potential biomarkers from metabolite profiling studies.

The research has also sought to explore the analytical potential of the use of ion mobility spectrometry in combination with existing UHPLC-MS based techniques for targeted and untargeted metabolite profiling studies on biofluids. The additional dimension of separation offered by IM has demonstrated improvement in mass spectral data quality, enhancement in selectivity for targeted metabolite analysis as well as has aided in identification of potential biomarker molecules.

The field of metabolomics and untargeted metabolite profiling in particular is expected to continue to grow in the upcoming decade with applications in fields as varied as cancer research, food science and nutritional research, personalised medicine, drug toxicity, drug discovery, drug metabolism and pharmacokinetic (DMPK) studies. These studies will provide new insights into integrated systems biology. Metabolomics being a technologically driven science, developments in instrumentation such as ion mobility and field asymmetric waveform ion mobility spectrometry along with new data analysis strategies are expected become valuable tools in the future development of metabolic profiling.

## 5.4 REFERENCES

Brown, L.J.; Toutoungi, D.E.; Devenport, N. A.; Reynolds, J.C.; Kaur-Atwal, G.; Boyle, P.; Creaser, C.S. *Anal. Chem.* **2010**, 82, 9827–9834.

Kolakowski, B.M.; Mester, Z. *Analyst* **2007**, 132, 842–864.

Methimazole - HMDB14901 [WWW Document], **2013**. . Human Metabolome Database. URL http://www.hmdb.ca/metabolites/HMDB14901 (accessed 21.9.13).

Nazarov, E.G. *Int. J. Ion Mobil. Spectrom*. **2012**, 15, 83–84.

# Appendix 1 – List of publications

## 6.1 LIST OF JOURNAL ARTICLES

- Malkar, A., Devenport, N.A., Martin, H.J., Patel, P., Turner, M.A., Watson, P., Maughan, R.J., Reid, H.J., Sharp, B.L., Thomas, C.L.P., Reynolds, J.C., Creaser, C.S., **2013**. Metabolic profiling of human saliva before and after induced physiological stress by ultra-high performance liquid chromatography–ion mobility–mass spectrometry. Metabolomics DOI:10.1007/s11306-013-0541-x.
- Huo, R., Agapiou, a, Bocos-Bintintan, V., Brown, L.J., Burns, C., Creaser, C.S., Devenport, N. a, Gao-Lau, B., Guallar-Hoyas, C., Hildebrand, L., Malkar, a, Martin, H.J., Moll, V.H., Patel, P., Ratiu, a, Reynolds, J.C., Sielemann, S., Slodzynski, R., Statheropoulos, M., Turner, M. a, Vautz, W., Wright, V.E., Thomas, C.L.P., **2011**. The trapped human experiment. Journal of breath research 5, 046006.

## 6.2 ORAL PRESENTATIONS

- "Analytical methods for metabolite profiling of saliva."
  *A. Malkar and C. S. Creaser.*
  Annual Meeting for the British Mass Spectrometry Society (BMSS): MS Evolution meets Intelligent design
  Eastbourne Winter Gardens, Eastbourne, UK **(2013)**.
- "Monitoring the effects of physiological stress by metabolic profiling of saliva using UPLC-IM-MS."
  *A. Malkar, H. Martin, P. Patel, M.A. Turner, P. Watson, H.J. Reid, R. Maughan, C.L.P. Thomas & C.S. Creaser.*
  Annual Meeting of the British Mass Spectrometry Society (BMSS): Mass Spectrometry for a Changing World
  Cardiff, UK. **(2012)**.
- "Metabonomic analysis of the chemopreventive efficacy of rice bran in colorectal cancer by UPLC-IM-MS."
  *A.Malkar, J.C. Reynolds, D.J.L. Jones, S. Sale & C.S.Creaser.*
  The Annual Meeting of the British Mass Spectrometry Society (BMSS): From Atoms to Biomolecules
  Cardiff, UK. **(2011)**.

## 6.3 POSTER PRESENTATIONS

- "Metabolic profiling of saliva using UPLC-IM-MS applied to monitoring the effects of physiological stress."
  *A. Malkar, H. Martin, P. Patel, M.A. Turner, P. Watson, H.J. Reid, R. Maughan, C.L.P. Thomas & C.S. Creaser.*
  The 60th Conference on Mass Spectrometry and Allied Topics, American Society of Mass Spectrometry (ASMS)
  Vancouver, Canada. **(2012)**.
- "Efficacy of dietary fibre for the prevention of colorectal cancer – using systems biology approach to identify potential plasma biomarkers of efficacy."
  *S. Sale, D.J.L. Jones, A. Malkar & C.S. Creaser.*
  Hope Researchers Evening
  De Montford University, Leicester, UK. **(May 2011)**.
- "Profiling of the salivary metabonome by ultra-performance liquid chromatography-ion mobility-spectrometry."
  *A.Malkar, J.C. Reynolds & C.S.Creaser.*
  Health and Life Sciences Conference
  Loughborough University, Loughborough, UK. **(March 2011)**.
- "Ultra-performance liquid chromatography-ion mobility-mass spectrometry profiling of the salivary metabonome."
  *A.Malkar, J.C.Reynolds & C.S.Creaser.*
  Annual Meeting of the British Mass Spectrometry Society (BMSS): Mass Spectrometry for a Changing World
  Cardiff, UK. **(2010)**.
- "Salivary metabolite profiling using ultra-performance liquid chromatography-ion mobility-mass spectrometry."
  *A. Malkar, J.C. Reynolds & C.S. Creaser.*
  Analytical Research Forum
  Loughborough, UK. **(2010)**.

# Appendix 2 – Workflows

## 7.1 SALIVA SAMPLING PROTOCOL

Requirements:

- Chromacol® Universal glass vials with screw caps (30mL) will be used for the collection of whole saliva from the participant.
- Eppendorf® LoBind microcentrifuge tubes 2.0mL will be used for aliquoting and storage of saliva samples.
- Micropipette (1000µL volume) and pipette tips for aliquoting the saliva and plasma samples.
- BD Vacutainer® LH (10.0mL) blood collection vial for blood plasma.

Setup:

- Centrifuge set up at 4$^{o}$C and 1500g, timed for required duration (15 minutes) for processing blood into plasma.
- Ice box with dry ice for storing of aliquoted samples prior to transport to -80$^{o}$C freezer.

Things to Avoid Prior to Sampling:

The participant should be asked NOT to:

- Consume alcohol 24 hours prior to the sample collection.
- Brush their teeth 1 hour prior to sampling.
- Have a large meal, dairy products, tea/coffee or salivary stimulants such as lemon-drops, chewing gum etc. within 1 hour prior to sample collection.

Things to Document Prior to Sampling:

- Date and time of the sample collection should be noted.
- Prescription and over-the-counter medication taken should be documented.
- Caffeine and Nicotine consumption/habits should be noted.
- Participants should be scanned for oral health problems or injuries.

Protocol:

The participant will be briefed about the sampling procedures.

The participant will be asked to drink approximately half a glass of unflavoured drinking water 10 minutes prior to the beginning of saliva sampling by passive drool method. This aids in stimulating saliva production.

Blood sampling will be carried out during these 10 minutes before the start of passive drool sampling. 10mL of blood will be collected which corresponds to approximately 5mL of plasma.

BD Vacutainer® LH vial with the collected blood, will be placed in centrifuge for processing blood into plasma.

At the end of 10 minutes, participant will be issued with a saliva collection vial i.e. universal glass vial (30mL)

The participant will be asked to sit comfortably with their head tilted forward; this should allow saliva to pool in front of the mouth.

The participant should be encouraged to avoid the instinct to swallow as well as roll the tongue around in mouth as this stimulates saliva production.

After about 1 minute, there should be enough saliva pooled in front of the mouth such that the participant will be able to dribble it in the collection vial. This process should be continued for a total of 10 minutes.

At the end of which the collection vial will be handed back to the investigator for aliquoting and storage.

Sputum sampling will be carried out next for participants willing to continue.

**Aliquoting and Storage:**

Whole saliva will be kept cold ($4^{o}$C i.e. on ice) immediately after collection.

The investigator will then aliquot the saliva from collection vial in 600µL aliquots in Eppendorf® LoBind microcentrifuge tubes (2.0 mL) with the help of micropipette.

Care should be taken to avoid the foam part of the saliva while aliquoting.

Saliva being a viscous fluid, slow aspiring with the micropipette is important for accurately transferring the desired amount for aliquoting.

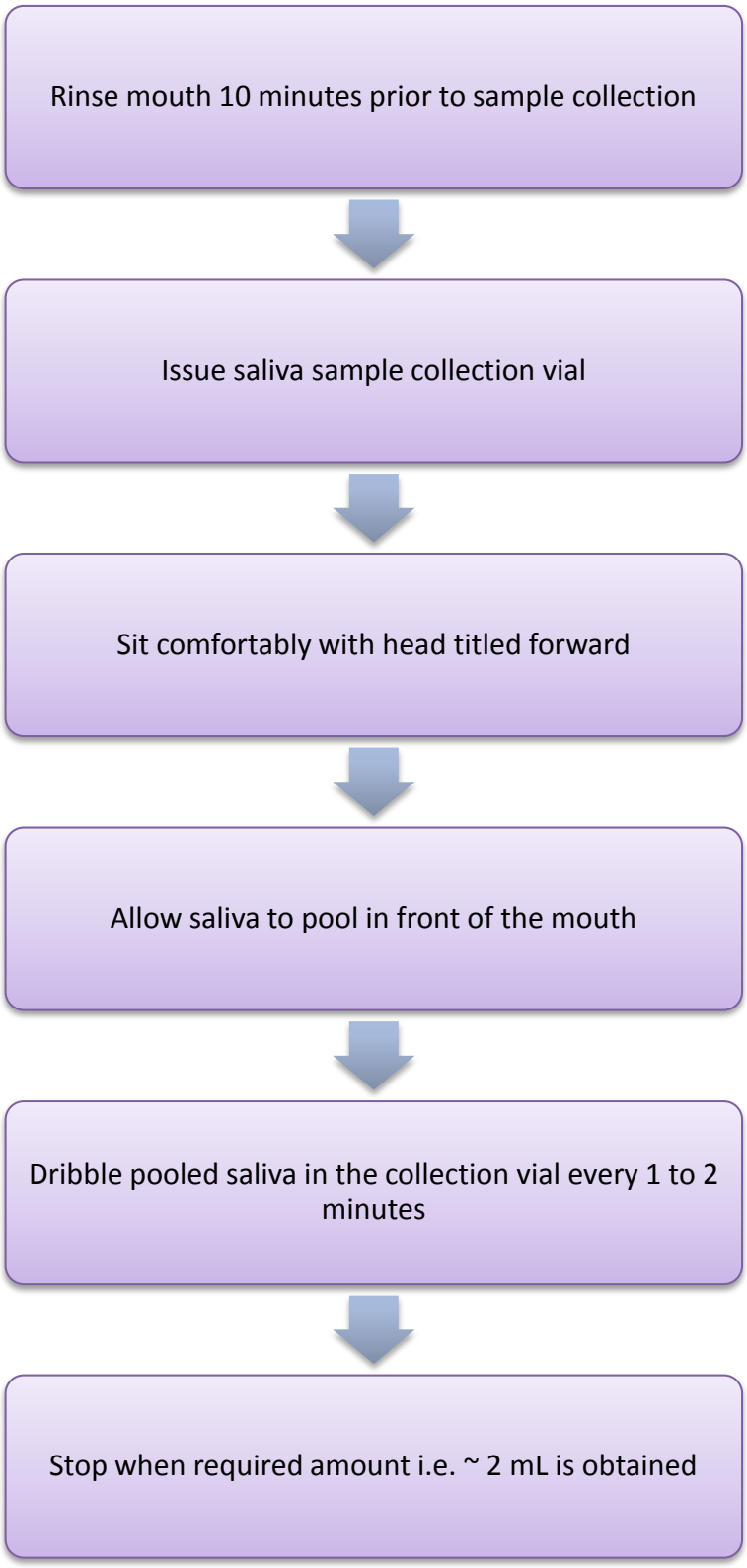
Aliquoted saliva samples should be immediately transferred to the ice box containing dry ice.

Plasma should be aliquoted as 1mL aliquots in 3 Eppendord LoBind® microcentrifuge tubes (total 3mL plasma).

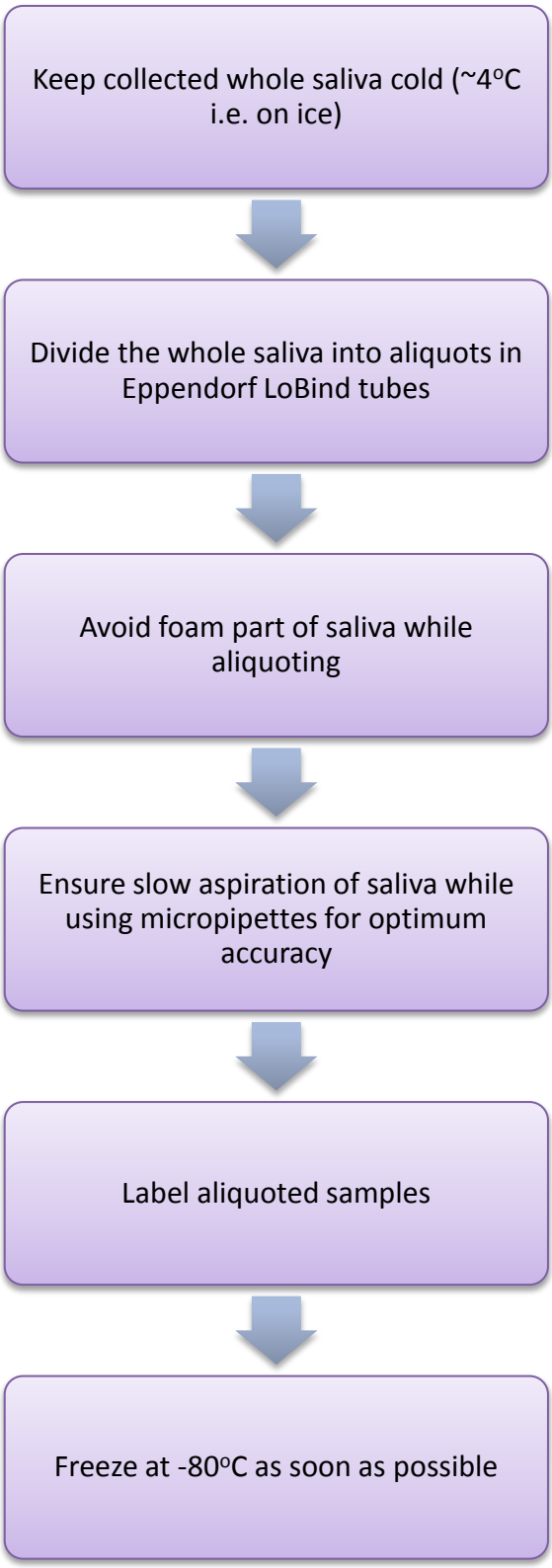
Plasma samples should be transferred to the ice box containing dry ice immediately after aliquoting.

At the end of sampling, all the samples should be transferred to the -80$^{o}$C freezer as soon as possible for storage.

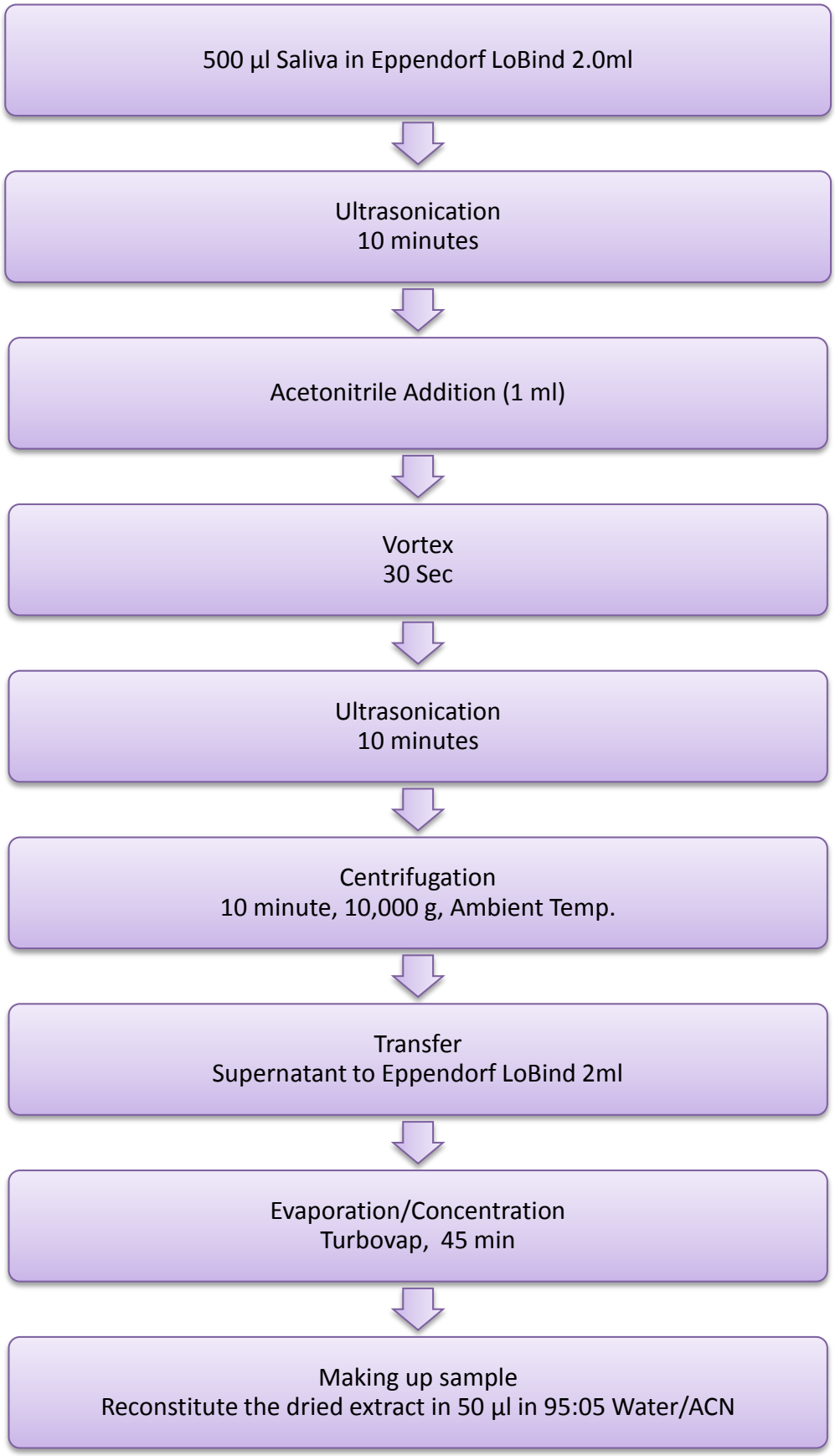
## 7.2 SALIVA SAMPLE COLLECTION WORKFLOW



## 7.3 ALIQUOTING AND STORAGE PROTOCOL



## 7.4 SALIVA SAMPLE PRE-TREATMENT WORKFLOW



## 7.5 COLORECTAL CANCER – RUN ORDER

| Run No. | Sample No. | Description | | Run No. | Sample No. | Description |
|---|---|---|---|---|---|---|
| 2 | | CONDITIONING | | 49 | 19 | C3 |
| 3 | | CONDITIONING | | 50 | | Blank |
| 4 | | CONDITIONING | | 51 | 20 | B2 |
| 5 | | CONDITIONING | | 52 | | Blank |
| 6 | | Blank | | 53 | | QC |
| 7 | 1 | A5 | | 54 | | Blank |
| 8 | | Blank | | 55 | 21 | C4 |
| 9 | 2 | D2 | | 56 | | Blank |
| 10 | | Blank | | 57 | 22 | B1 |
| 11 | 3 | B4 | | 58 | | Blank |
| 12 | | Blank | | 59 | 23 | C2 |
| 13 | 4 | C15 | | 60 | | Blank |
| 14 | | Blank | | 61 | 24 | A4 |
| 15 | 5 | C7 | | 62 | | Blank |
| 16 | | Blank | | 63 | 25 | C6 |
| 17 | | QC | | 64 | | Blank |
| 18 | | Blank | | 65 | | QC |
| 19 | 6 | C1 | | 66 | | Blank |
| 20 | | Blank | | 67 | 26 | C13 |
| 21 | 7 | C10 | | 68 | | Blank |
| 22 | | Blank | | 69 | 27 | C11 |
| 23 | 8 | D7 | | 70 | | Blank |
| 24 | | Blank | | 71 | 28 | D8 |
| 25 | 9 | D1 | | 72 | | Blank |
| 26 | | Blank | | 73 | 29 | C16 |
| 27 | 10 | A7 | | 74 | | Blank |
| 28 | | Blank | | 75 | 30 | C9 |
| 29 | | QC | | 76 | | Blank |
| 30 | | Blank | | 77 | | QC |
| 31 | 11 | C12 | | 78 | | Blank |
| 32 | | Blank | | 79 | 31 | D4 |
| 33 | 12 | D3 | | 80 | | Blank |
| 34 | | Blank | | 81 | 32 | C14 |
| 35 | 13 | A6 | | 82 | | Blank |
| 36 | | Blank | | 83 | 33 | A2 |
| 37 | 14 | C8 | | 84 | | Blank |
| 38 | | Blank | | 85 | 34 | C5 |
| 39 | 15 | D6 | | 86 | | Blank |
| 40 | | Blank | | 87 | 35 | A1 |
| 41 | | QC | | 88 | | Blank |
| 42 | | Blank | | 89 | | QC |
| 43 | 16 | B3 | | | | |
| 44 | | Blank | | | | |
| 45 | 17 | A3 | | | | |
| 46 | | Blank | | | | |
| 47 | 18 | D5 | | | | |
| 48 | | Blank | | | | |

## 7.6 ASTHMA PATIENT INFORMATION

Column Explanations:

- Patient ID
- BDP equivalent: How much inhaler steroid a patient is taking each day (e.g. 200mcg perday)
- Gender: 0=male 1=female
- When diagnosed: 0=<1 year 1= 1-5 years
- Rhinitis 0=no 1=yes
- GO reflux 0=no 1=yes
- Oral contraceptive pill 0=no 1=yes
- Ezcema 0=no 1=yes
- Non-steroidal anti-inflammatory allergy 0=no 1=yes
- Family history of allergy 0=no 1=yes
- JACQ: symptom questionnaire: higher the number the worse the symptoms
- V1IGE: Blood IGE level
- V1SRMEOS: Blood eosinophil level
- NO values: 10,30,50,100,200
- FEV1
- FEV1% pred: >80% = less likely to have asthma <80% more likely to have asthma (all patients with a value <80% are highlighted in red).
- FVC
- FVC%pred
- FEV1/FVC ratio
- V1eos: sputum eosinophil count: >3% =more likely to have asthma (all patients with a value >3% are highlighted in red).
- Neutrophils
- Macrophages
- Lymphocytes
- Epithelial
- SPT: skin prick test 0=no 1=yes
- Aspergillus
- Number of + spt each patient had
- PC20 = methacholine challenge (<8mg/ml = more likely to have asthma all patients with a value of less than 8 are highlighted in red).

| PatientID | BDP Equiv | V1 BDP Daily | Gender | DOB | WhenDiagnosed | Height (cm) | Weight (kg) | BMI | Rhinitis | GOReflux | OralCon | Eczema | NSAID | Family |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | |
| 1 | | | Female | 19/04/1966 | | 167 | 47.6 | | | | | | | |
| 2 | | | Male | 10/06/1986 | | 181 | 75.4 | | | | | | | |
| 3 | | | Female | 26/12/1948 | | NR | NR | | | | | | | |
| 4 | 200 | 200 | Male | 28/12/1941 | 1 | 171 | 78.1 | 26.70907288 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 1000 | 1000 | Female | 18/04/1953 | 1 | 162 | 58 | 22.10028959 | 0 | 0 | 0 | 0 | 0 | 1 |
| 6 | 800 | 800 | Male | 15/03/1937 | 1 | 172 | 74.9 | 25.31773932 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 400 | 400 | Female | 08/02/1943 | 1 | 153 | 62 | 26.48553975 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 400 | 400 | Female | 05/03/1947 | 1 | 158 | 62.3 | 24.95593655 | 0 | 0 | 0 | 1 | 0 | 1 |
| 9 | 400 | 400 | Male | 26/02/1952 | 1 | 181 | 86 | 26.25072495 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 200 | 200 | Male | 21/09/1953 | 1 | 174.5 | 90.5 | 29.72060985 | 1 | 0 | 0 | 0 | 0 | 0 |
| 11 | 800 | 800 | Male | 04/11/1959 | 1 | 179 | 96 | 29.96161169 | 0 | 0 | 0 | 0 | 0 | 1 |
| 12 | 400 | 400 | Male | 27/09/1942 | 1 | 168 | 76 | 26.92743764 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 400 | 400 | Female | 04/08/1985 | 1 | 171 | 81 | 27.70083102 | 0 | 0 | 1 | 1 | 0 | 1 |
| 14 | | | Male | 17/05/1992 | | 184 | 69 | | | | | | | |
| 15 | | | Female | 04/07/1992 | | 166 | 87.3 | | | | | | | |
| 16 | 400 | 400 | Female | 25/07/1954 | 1 | 170 | 65.5 | 22.66435986 | 0 | 0 | 0 | 1 | 0 | 0 |
| 17 | | | Female | 11/04/1989 | | 158 | 67.9 | | | | | | | |
| 18 | 400 | 400 | Male | 25/04/1961 | 1 | 163 | 63 | 23.71184463 | 0 | 0 | 0 | 1 | 0 | 0 |
| 19 | 800 | 800 | Female | 16/05/1940 | 1 | 170 | 88 | 30.44982699 | 0 | 1 | 0 | 1 | 0 | 1 |
| 20 | 400 | 400 | Female | 08/11/1971 | 1 | 149 | 69.2 | 31.16976713 | 0 | 0 | 0 | 1 | 1 | 1 |
| 21 | 400 | 400 | Male | 03/05/1938 | 1 | 170 | 80 | 27.6816609 | 1 | 1 | 0 | 0 | 0 | 0 |
| 22 | 400 | 400 | Female | 14/02/1951 | 1 | 173 | 80.6 | 26.93040195 | 0 | 0 | 0 | 1 | 0 | 1 |
| 23 | 800 | 800 | Male | 17/10/1960 | 1 | 178 | 107.6 | 33.96035854 | 1 | 1 | 0 | 0 | 0 | 0 |
| 24 | 800 | 800 | Female | 26/07/1959 | 1 | 177 | 89.3 | 28.50394203 | 0 | 1 | 0 | 0 | 0 | 0 |
| 25 | 400 | 400 | Female | 11/01/1945 | 1 | 156 | 55.2 | 22.68244576 | 0 | 1 | 0 | 0 | 0 | 0 |
| 28 | 800 | 800 | Female | 12/06/1972 | 1 | 171 | 146.4 | 50.06668719 | 0 | 1 | 0 | 1 | 0 | 0 |
| 29 | 400 | 400 | Female | 04/01/1945 | 1 | 162 | 73 | 27.81588173 | 1 | 1 | 0 | 1 | 1 | 0 |
| 30 | 800 | 800 | Female | 24/06/1943 | 1 | 155 | 70 | 29.13631634 | 1 | 1 | 0 | 1 | 0 | 0 |
| 31 | 400 | 400 | Male | 28/06/1970 | 1 | 168 | 90.6 | 32.10034014 | 0 | 0 | 0 | 0 | 0 | 0 |
| 32 | 200 | 200 | Male | 01/07/1943 | 1 | 182 | 72.9 | 22.00821157 | 0 | 0 | 0 | 0 | 0 | 0 |
| 33 | 2000 | 2000 | Male | 03/05/1947 | 1 | 173 | 101 | 33.74653346 | 0 | 0 | 0 | 0 | 0 | 0 |
| 34 | | | Female | 10/03/1976 | | NR | NR | | | | | | | |
| 35 | 50 | 50 | Male | 22/02/1974 | 1 | 183 | 99.7 | 29.77096957 | 1 | 1 | 0 | 0 | 0 | 0 |
| 36 | 800 | 800 | Male | 03/10/1956 | 0 | 173 | 82.3 | 27.49841291 | 0 | 0 | 0 | 0 | 0 | 0 |
| 37 | 2400 | 2400 | Male | 17/12/1949 | 1 | 174.5 | 91.7 | 30.11469528 | 0 | 0 | 0 | 0 | 0 | 1 |
| 38 | | | Female | 26/04/1959 | | 158 | 48 | | | | | | | |
| 39 | | | Female | NR | | NR | NR | | | | | | | |
| 40 | | | Female | 22/07/1990 | | 170 | 78.2 | | | | | | | |
| 41 | | | Female | 17/02/1992 | | 167 | 62.4 | | | | | | | |
| 42 | | | Female | 03/12/1990 | | 159 | 55.7 | | | | | | | |
| 43 | | | Female | 05/02/1992 | | 172 | 53.3 | | | | | | | |
| 44 | 800 | 800 | Female | 15/06/1945 | 1 | 166 | 57.1 | 20.72143998 | 0 | 0 | 0 | 0 | 0 | 0 |
| 45 | 400 | 400 | Female | 23/03/1941 | 1 | 161 | 66 | 25.46198063 | 0 | 0 | 0 | 1 | 0 | 1 |
| 46 | 100 | 100 | Female | 31/03/1978 | 0 | 167 | 56 | 20.07960128 | 0 | 0 | 0 | 1 | 0 | 1 |
| 47 | | | Female | NR | | NR | NR | | | | | | | |
| 48 | | | Male | 05/05/1985 | | 163 | 60 | | | | | | | |
| 49 | | | Male | NR | | NR | NR | | | | | | | |
| 50 | | | Female | 12/12/1987 | | 178 | 63 | | | | | | | |
| 51 | | | Male | 17/01/1952 | | 185 | 95 | | | | | | | |
| 52 | | | Male | 31/10/1988 | | 173 | 95 | | | | | | | |
| 53 | | | Female | 31/12/1952 | | 165 | 90 | | | | | | | |
| 54 | | | Male | 13/09/1984 | | 185 | 79 | | | | | | | |
| 55 | | | Male | 12/08/1992 | | NR | NR | | | | | | | |
| 56 | 400 | 400 | Male | 11/06/1943 | 1 | 173 | 76.4 | 25.52708076 | 0 | 0 | 0 | 0 | 0 | 0 |
| 57 | 400 | 400 | Female | 16/05/1980 | 1 | 170 | 63.6 | 22.00692042 | 0 | 0 | 1 | 0 | 1 | 1 |
| 58 | 1000 | 1000 | Female | 24/06/1973 | 1 | 167 | 87 | 31.19509484 | 0 | 0 | 0 | 0 | 0 | 1 |
| 59 | 400 | 400 | Female | 18/01/1963 | 1 | 169 | 67 | 23.45856238 | 0 | 0 | 0 | 1 | 0 | 1 |

| PatientID | V1JACQ1-5 | V1JACQ1-6 | V1JACQ1-7 | V1IGE | V1SRMEOS | MEANV1NO10 | MEANV1NO30 | MEANV1NO50 | MEANV1NO100 | MEANV1NO200 | V1Alv | V1Bronch | V1DNO | V1FEV1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| 1 | | | | | | | | | | | | | | |
| 2 | | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | | | | |
| 4 | 0.8 | 0.67 | 1 | 222 | 0.2 | | 33.6 | 23.7 | 11.9 | 7.25 | 2.1 | 61.2 | 0.8 | 2.35 |
| 5 | 1.4 | 1.33 | 1.8 | 1160 | 0.4 | | 163.3 | 121.9 | 57.8 | 29.4 | 0.9 | 340.2 | 0 | 1.5 |
| 6 | 0.2 | 0.17 | 0.43 | 324 | 0.4 | 135.7 | 56.25 | 38.6 | 21.7 | 13.25 | 6.2 | 89.1 | 0.9 | 2.27 |
| 7 | 0.6 | 0.67 | 0.57 | 4 | 0.1 | | 19.8 | 14.35 | 7.15 | 5.45 | 2.8 | 31.3 | 0.7 | 1.85 |
| 8 | 0.2 | 0.17 | 0.71 | 184 | 0.1 | 73.1 | 24.4 | 15.65 | 10.2 | 5.9 | 2.5 | 41.5 | 0.8 | 1.38 |
| 9 | 0.2 | 0.3 | 1 | 91 | 0.1 | | 18.55 | 12.55 | 7.8 | 5.45 | 3.1 | 28.3 | 1 | 1.9 |
| 10 | 1 | 0.83 | 1.29 | 91 | 0.6 | | 56.25 | 51.55 | 26.05 | 15.55 | 6.5 | 112.9 | 0.7 | 2.26 |
| 11 | 1 | 0.8 | 1 | 98 | 0.5 | | 30 | 20.75 | 12.7 | 7.7 | 3.6 | 51.1 | 0.9 | 3.08 |
| 12 | 0.5 | 0.33 | 0.57 | 3 | 0.5 | | 59.95 | 39.75 | 21.85 | | 5.2 | 101.1 | 0.8 | 2.4 |
| 13 | 0 | 0 | 0.57 | 109 | 0.4 | 115.65 | 44.15 | 29.2 | 15.2 | 8.3 | 2.3 | 74.9 | 0.8 | 2.32 |
| 14 | | | | | | | | | | | | | | |
| 15 | | | | | | | | | | | | | | |
| 16 | 0 | 0.17 | 0.71 | 10200 | 0.2 | 204.8 | 91.3 | 64.35 | 36.35 | 18.65 | 6.4 | 160.5 | 0.6 | 1.97 |
| 17 | | | | | | | | | | | | | | |
| 18 | 0 | 0 | 0.57 | 22610 | 0.5 | 77.15 | 34.85 | 26.3 | 13.95 | 8.65 | 4.4 | 54.5 | 0.8 | 2.02 |
| 19 | 0.6 | 0.5 | 1.14 | 382 | 0.1 | | 36 | 33.65 | 17.45 | 7.7 | 1.1 | 89.1 | 0.1 | 1.35 |
| 20 | 0 | 0 | 0.29 | 114 | 0.3 | | 26.8 | 17.8 | 10.35 | 7.75 | 4.3 | 39.6 | 0.9 | 2.07 |
| 21 | 0.8 | 0.7 | 0.6 | 133 | 0 | | 26.1 | 21.95 | 11.8 | 7.95 | 4.2 | 46 | 0.8 | 3.15 |
| 22 | 0.2 | 0.17 | 0.71 | 303 | 0.3 | 82.4 | 32.6 | 25.2 | 15.1 | 7.7 | 3.5 | 57.1 | 0.6 | 1.65 |
| 23 | 1.6 | 1.67 | 1.43 | 93 | 0.2 | 44.15 | 27.35 | 15.2 | 9.05 | 6.5 | 3.8 | 32.9 | 0.8 | 3.58 |
| 24 | 0.8 | 1 | 1.14 | | 0.4 | | 34.65 | 19 | 12.95 | 8.05 | 3.7 | 52.6 | 0.7 | 2.4 |
| 25 | 0 | 0 | 0.29 | 28 | 0.3 | 75.6 | 29.8 | 15.55 | 10 | 5.25 | 1.4 | 47.3 | 0.4 | 1.72 |
| 28 | 0.4 | 0.33 | 0.86 | 75 | 0.2 | | 13.3 | 9.35 | 6.2 | 2.35 | -0.2 | 32.6 | 0 | 2.3 |
| 29 | 0 | 0 | 0.43 | 101 | 0.3 | 88.7 | 39.3 | 20.95 | 12.5 | 8.5 | 3.6 | 56.7 | 0.6 | 1.63 |
| 30 | 1 | 1 | 1.57 | 126 | 0.6 | | | 20.7 | 13.4 | 8.05 | 3.7 | 54.3 | 0.8 | 1.32 |
| 31 | 0 | 0 | 0 | 38 | 0.1 | | 22.7 | 17.1 | 9.8 | 4.65 | 1.1 | 45.4 | 0.2 | 4.25 |
| 32 | 0 | 0.16 | 0.57 | 315 | 0.1 | 116.6 | 59.65 | 42.15 | 23.15 | 12.75 | 5.1 | 98.6 | 0.7 | 2.53 |
| 33 | 1 | 1 | 1.57 | 223 | 0.7 | | 89.2 | 61.6 | 39.25 | 20.9 | 8.4 | 160.8 | 0.8 | 1.82 |
| 34 | | | | | | | | | | | | | | |
| 35 | 0.8 | 0.83 | 1.29 | 154 | 0.1 | 93.1 | 38.65 | 21.95 | 13.25 | 8.35 | 3.6 | 57.9 | 0.9 | 2.6 |
| 36 | 0.2 | 0.33 | 0.29 | 139 | 0.4 | | 37.6 | 26.9 | 14.35 | 8.45 | 2.9 | 67.6 | 0.5 | 3.45 |
| 37 | 1 | 1 | 0.86 | 40 | 0.5 | | 134.05 | 95.05 | 55.9 | 33.55 | 15 | 231.1 | 0.9 | 3.28 |
| 38 | | | | | | | | | | | | | | |
| 39 | | | | | | | | | | | | | | |
| 40 | | | | | | | | | | | | | | |
| 41 | | | | | | | | | | | | | | |
| 42 | | | | | | | | | | | | | | |
| 43 | | | | | | | | | | | | | | |
| 44 | 1.2 | 1.17 | 1 | 77 | 0.4 | | | 17.3 | | | 57.1 | 119 | 1 | 2.55 |
| 45 | 0.2 | 0.17 | 0.14 | 3 | 0.1 | | 30.45 | 17.25 | 11 | 6.95 | 3 | 46.8 | 0.8 | 2.01 |
| 46 | 0 | 0 | 0 | 22 | 0.1 | | 12.2 | 10.25 | 6.1 | 2.9 | 1 | 25.4 | 0.4 | 2.6 |
| 47 | | | | | | | | | | | | | | |
| 48 | | | | | | | | | | | | | | |
| 49 | | | | | | | | | | | | | | |
| 50 | | | | | | | | | | | | | | |
| 51 | | | | | | | | | | | | | | |
| 52 | | | | | | | | | | | | | | |
| 53 | | | | | | | | | | | | | | |
| 54 | | | | | | | | | | | | | | |
| 55 | | | | | | | | | | | | | | |
| 56 | 0.4 | 0.33 | 0.57 | | | 62.1 | 36.25 | 28.95 | 15.35 | 8.65 | 4.7 | 55.2 | 0.6 | 2.8 |
| 57 | 1 | 0.83 | 0.71 | | 0.1 | | 12.25 | 7.45 | 2.75 | 3.8 | 2.4 | 13.3 | 0.5 | 3.4 |
| 58 | 1 | 1 | 1.14 | 149 | 0.3 | | 22.2 | 15.6 | 9.5 | 5.45 | 2.4 | 39.1 | 0.9 | 2.45 |
| 59 | 0.8 | 0.67 | 0.57 | 36 | 0.2 | | 7.8 | 4.3 | 1.95 | 0.5 | -0.8 | 15.6 | 0.5 | 3.4 |

| PatientID | V1FEV1%PRED | V1FVC | V1FVC%PRED | V1FEV1/FVC | V1EOS | V1NEUTS | V1MACROS | V1LYMPH | V1EPITH | V1SPTDONE | V1HIST | V1SAL | V1PC20 | NO category |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| 1 | | | | | | | | | | | | | | |
| 2 | | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | | | | |
| 4 | 79 | 3.05 | 75 | 0.77 | 0.5 | 88.5 | 10.75 | 0 | 0.25 | 1 | 1 | 0 | 2.99 | 0 |
| 5 | 59 | 2.8 | 100 | 0.5 | 10.75 | 16 | 68.5 | 0 | 4.75 | 1 | 1 | 0 | 0.03 | 2 |
| 6 | 81 | 3.35 | 91 | 0.68 | 20.75 | 67.75 | 11 | 0 | 0.5 | 1 | 1 | 0 | 32 | 1 |
| 7 | 105 | 2.4 | 112 | 0.77 | 0 | 47 | 40.75 | 0.25 | 12 | 1 | 1 | 0 | 32 | 0 |
| 8 | 67 | 2.55 | 103 | 0.54 | | | | | | 1 | 1 | 0 | 2.3 | 0 |
| 9 | 54 | 3.5 | 74 | 0.54 | 0.5 | 63.5 | 34.75 | 0 | 1.25 | 1 | 1 | 0 | 7.56 | 0 |
| 10 | 67 | 3.45 | 82 | 0.66 | 6 | 26.5 | 40 | 1 | 26.5 | 1 | 1 | 0 | 2.52 | 2 |
| 11 | 82 | 4.48 | 96 | 0.69 | 1.75 | 72.82 | 22.44 | 0.25 | 2.74 | 1 | 1 | 0 | 6.63 | 0 |
| 12 | 87 | 4.15 | 116 | 0.58 | 19 | 49.75 | 24.75 | 0.25 | 6.25 | 1 | 1 | 0 | 32 | 1 |
| 13 | 66 | 3.15 | 78 | 0.74 | | | | | | 1 | 1 | 0 | 0.5 | 1 |
| 14 | | | | | | | | | | | | | | |
| 15 | | | | | | | | | | | | | | |
| 16 | 67 | 3.4 | 90 | 0.58 | 0.25 | 46 | 51.75 | 0.25 | 1.75 | 1 | 1 | 0 | 1.17 | 2 |
| 17 | | | | | | | | | | | | | | |
| 18 | 65 | 3.63 | 96 | 0.56 | 4.75 | 41.75 | 53.5 | 0 | 0 | 0 | 0 | 0 | 0.59 | 1 |
| 19 | 53 | 2 | 63 | 0.68 | 0.75 | 63.5 | 24.5 | 1 | 10.25 | 1 | 1 | 0 | 8 | 1 |
| 20 | 88 | 2.65 | 97 | 0.78 | 1.5 | 46.25 | 51 | 0.25 | 1 | 1 | 1 | 0 | 2 | 0 |
| 21 | 99 | 4.45 | 106 | 0.71 | 0 | 38.5 | 60.75 | 0.25 | 0.5 | 1 | 1 | 0 | 32 | 0 |
| 22 | 63 | 2.2 | 66 | 0.75 | 0.75 | 41 | 57 | 0.25 | 1 | 1 | 1 | 0 | 1 | 1 |
| 23 | 97 | 5 | 107 | 0.72 | 0.5 | 78 | 17.75 | 1.75 | 2 | 1 | 1 | 0 | 0.03 | 0 |
| 24 | 80 | 3.2 | 88 | 0.75 | 0.25 | 99.5 | 0 | 0 | 0.25 | 1 | 1 | 0 | 0 | 0 |
| 25 | 81.1 | 3 | 107.5 | 0.57 | | | | | | 1 | 1 | 0 | 7.16 | 0 |
| 28 | 68 | 3.15 | 76 | 0.73 | 0.8 | 91.4 | 7.4 | 0.4 | 0.2 | 1 | 1 | 0 | 6.35 | 0 |
| 29 | 76 | 2.6 | 101 | 0.63 | 0.75 | 91.25 | 7.75 | 0 | 0.25 | 1 | 1 | 0 | 32 | 0 |
| 30 | 57 | 2 | 64 | 0.66 | 9.5 | 32 | 54.25 | 0 | 4.25 | 1 | 1 | 0 | 8 | 0 |
| 31 | 107 | 5.3 | 109 | 0.8 | 1 | 64 | 31 | 2 | 2.5 | 1 | 1 | 0 | 32 | 0 |
| 32 | 73 | 4.6 | 103 | 0.55 | 0 | 100 | 0 | 0 | 0 | 1 | 1 | 0 | 32 | 1 |
| 33 | 58 | 2.85 | 71 | 0.64 | 19.5 | 29.75 | 29.5 | 0 | 21.25 | 1 | 1 | 0 | 2 | 2 |
| 34 | | | | | | | | | | | | | | |
| 35 | 60 | 3.95 | 75 | 0.66 | | | | | | 1 | 1 | 0 | 16 | 0 |
| 36 | 97 | 4.5 | 97 | 0.77 | | | | | | 1 | 1 | 0 | 32 | 1 |
| 37 | 101 | 4.25 | 102 | 0.77 | 11 | 74.5 | 13.75 | 0 | 0.25 | 1 | 1 | 0 | 15.79 | 2 |
| 38 | | | | | | | | | | | | | | |
| 39 | | | | | | | | | | | | | | |
| 40 | | | | | | | | | | | | | | |
| 41 | | | | | | | | | | | | | | |
| 42 | | | | | | | | | | | | | | |
| 43 | | | | | | | | | | | | | | |
| 44 | 116 | 3.33 | 115 | 0.77 | | | | | | 1 | 0 | 0 | 32 | 0 |
| 45 | 100 | 2.65 | 110 | 0.76 | 0.25 | 83.75 | 13.5 | 2.5 | 0.25 | 0 | 0 | 0 | 1 | 0 |
| 46 | 82 | 3 | 82 | 0.87 | | | | | | 1 | 1 | 0 | | 0 |
| 47 | | | | | | | | | | | | | | |
| 48 | | | | | | | | | | | | | | |
| 49 | | | | | | | | | | | | | | |
| 50 | | | | | | | | | | | | | | |
| 51 | | | | | | | | | | | | | | |
| 52 | | | | | | | | | | | | | | |
| 53 | | | | | | | | | | | | | | |
| 54 | | | | | | | | | | | | | | |
| 55 | | | | | | | | | | | | | | |
| 56 | 87 | 3.8 | 93 | 0.74 | | | | | | 1 | 1 | 0 | 32 | 1 |
| 57 | 102 | 4 | 104 | 0.85 | | | | | | 1 | 1 | 0 | 32 | 0 |
| 58 | 88 | 3.1 | 88 | 0.79 | | | | | | 1 | 1 | 0 | | 0 |
| 59 | 118 | 4.6 | 137 | 0.76 | | | | | | 1 | 1 | 0 | 32 | 0 |

## 7.7 ASTHMA SCHEME OF ANALYSIS (RUN ORDER)

| Serial No. | Participant | Sample Code | Vial Code | Asthmatic | Control | Gender | Run Number |
|---|---|---|---|---|---|---|---|
| 1 | Blank | - | - | - | - | - | AM_Saliva_001 |
| 2 | Standard | - | - | - | - | - | AM_Saliva_002 |
| 3 | Blank | - | - | - | - | - | AM_Saliva_003 |
| 4 | CR1 | QC 1 | - | - | - | - | AM_Saliva_004 |
| 5 | CR2 | QC1 | - | - | - | - | AM_Saliva_005 |
| 6 | CR3 | QC1 | - | - | - | - | AM_Saliva_006 |
| 7 | CR4 | QC1 | - | - | - | - | AM_Saliva_007 |
| 8 | CR5 | QC1 | - | - | - | - | AM_Saliva_008 |
| 9 | QC1 | QC1 | - | - | - | - | AM_Saliva_009 |
| 10 | 54 | MAA054 | 70406 | 1 | | M | AM_Saliva_010 |
| 11 | 28 | MAA028 | 65609 | 1 | | F | AM_Saliva_011 |
| 12 | 55 | MAA055 | 70773 | | 1 | M | AM_Saliva_012 |
| 13 | 3 | MAA003 | 54386 | | 1 | F | AM_Saliva_013 |
| 14 | 56 | MAA056 | 70788 | 1 | | M | AM_Saliva_014 |
| 15 | QC2 | QC2 | - | - | - | - | AM_Saliva_015 |
| 16 | 51 | MAA051 | 70376 | | 1 | M | AM_Saliva_016 |
| 17 | 49 | MAA049 | 70170 | | 1 | M | AM_Saliva_017 |
| 18 | 1 | MAA001 | 52276 | | 1 | F | AM_Saliva_018 |
| 19 | 21 | MAA021 | 58279 | 1 | | M | AM_Saliva_019 |
| 20 | 10 | MAA010 | 55501 | 1 | | M | AM_Saliva_020 |
| 21 | QC3 | QC3 | - | - | | - | AM_Saliva_021 |
| 22 | 58 | MAA058 | 71210 | 1 | | F | AM_Saliva_022 |
| 23 | 16 | MAA016 | 56563 | 1 | | F | AM_Saliva_023 |
| 24 | 36 | MAA036 | 67893 | 1 | | M | AM_Saliva_024 |
| 25 | 7 | MAA007 | 55231 | 1 | | F | AM_Saliva_025 |
| 26 | 47 | MAA047 | 70153 | | 1 | F | AM_Saliva_026 |
| 27 | QC4 | QC4 | - | - | - | - | AM_Saliva_027 |
| 28 | 38 | MAA038 | 67911 | | 1 | F | AM_Saliva_028 |
| 29 | 11 | MAA011 | 55510 | 1 | | M | AM_Saliva_029 |
| 30 | 20 | MAA020 | 58219 | 1 | | F | AM_Saliva_030 |
| 31 | 44 | MAA044 | 68771 | 1 | | F | AM_Saliva_031 |
| 32 | 18 | MAA018 | 57366 | 1 | | M | AM_Saliva_032 |
| 33 | QC5 | QC5 | - | - | - | - | AM_Saliva_033 |
| 34 | 30 | MAA030 | 65801 | 1 | | F | AM_Saliva_034 |
| 35 | 12 | MAA012 | 55606 | 1 | | M | AM_Saliva_035 |
| 36 | 48 | MAA048 | 70161 | | 1 | M | AM_Saliva_036 |
| 37 | 15 | MAA015 | 56191 | | 1 | F | AM_Saliva_037 |
| 38 | 19 | MAA019 | 57613 | 1 | | F | AM_Saliva_038 |
| 39 | QC6 | QC6 | - | - | - | - | AM_Saliva_039 |
| 1 | QC 6 DAY 2 | QC 6 DAY 2 | - | - | - | - | AM_Saliva_040 |
| 2 | 6 | MAA006 | 55141 | 1 | | M | AM_Saliva_041 |
| 3 | 32 | MAA032 | 66163 | 1 | | M | AM_Saliva_042 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 4 | 13 | MAA013 | 56113 | 1 | | F | AM_Saliva_043 |
| 5 | 35 | MAA035 | 67776 | 1 | | M | AM_Saliva_044 |
| 6 | 25 | MAA025 | 64761 | 1 | | F | AM_Saliva_045 |
| 7 | QC7 | QC7 | - | - | - | - | AM_Saliva_046 |
| 8 | 41 | MAA041 | 68019 | | 1 | F | AM_Saliva_047 |
| 9 | 17 | MAA017 | 57025 | | 1 | F | AM_Saliva_048 |
| 10 | 9 | MAA009 | 55492 | 1 | | M | AM_Saliva_049 |
| 11 | 33 | MAA033 | 67074 | 1 | | M | AM_Saliva_050 |
| 12 | 53 | MAA053 | 70394 | | 1 | F | AM_Saliva_051 |
| 13 | QC8 | QC8 | - | - | - | - | AM_Saliva_052 |
| 14 | 24 | MAA024 | 64753 | 1 | | F | AM_Saliva_053 |
| 15 | 37 | MAA037 | 67903 | 1 | | M | AM_Saliva_054 |
| 16 | 23 | MAA023 | 64768 | 1 | | M | AM_Saliva_055 |
| 17 | 8 | MAA008 | 55421 | 1 | | F | AM_Saliva_056 |
| 18 | 34 | MAA034 | 67637 | | 1 | F | AM_Saliva_057 |
| 19 | QC9 | QC9 | - | - | - | - | AM_Saliva_058 |
| 20 | 57 | MAA057 | 71207 | 1 | | F | AM_Saliva_059 |
| 21 | 5 | MAA005 | 54872 | 1 | | F | AM_Saliva_060 |
| 22 | 29 | MAA029 | 65675 | 1 | | F | AM_Saliva_061 |
| 23 | 4 | MAA004 | 54737 | 1 | | M | AM_Saliva_062 |
| 24 | 43 | MAA043 | 68038 | | 1 | F | AM_Saliva_063 |
| 25 | QC10 | QC10 | - | - | - | - | AM_Saliva_064 |
| 26 | 2 | MAA002 | 52294 | | 1 | M | AM_Saliva_065 |
| 27 | 39 | MAA039 | 67920 | | 1 | F | AM_Saliva_066 |
| 28 | 22 | MAA022 | 60258 | 1 | | F | AM_Saliva_067 |
| 29 | 40 | MAA040 | 68010 | | 1 | F | AM_Saliva_068 |
| 30 | 45 | MAA045 | 68997 | 1 | | F | AM_Saliva_069 |
| 31 | QC11 | QC11 | - | - | - | - | AM_Saliva_070 |
| 32 | 46 | MAA046 | 70055 | 1 | | F | AM_Saliva_071 |
| 33 | 31 | MAA031 | 65811 | 1 | | M | AM_Saliva_072 |
| 34 | 52 | MAA052 | 70385 | | 1 | M | AM_Saliva_073 |
| 35 | 42 | MAA042 | 68028 | | 1 | F | AM_Saliva_074 |
| 36 | 50 | MAA050 | 70182 | | 1 | F | AM_Saliva_075 |
| 37 | QC12 | QC12 | - | - | - | - | AM_Saliva_076 |
| 38 | 59 | MAA059 | 71230 | 1 | | F | AM_Saliva_077 |
| 39 | 14 | MAA014 | 56181 | | 1 | M | AM_Saliva_078 |
| 40 | QC13 | QC13 | - | - | - | - | AM_Saliva_079 |
| 41 | BLANK | - | - | - | - | - | AM_Saliva_080 |
| 42 | STANDARD | - | - | - | - | - | AM_Saliva_081 |
| 43 | BLANK | - | - | - | - | - | AM_Saliva_082 |