Development of a multiple application LC-MS/MS method for targeted metabolic profiling of biological matrices

M Venter

orcid.org/ 0000-0002-5026-5336

Dissertation submitted in fulfilment of the requirements for the degree *Master of Science* in *Pharmaceutical Sciences* at the North-West University

Supervisor: Prof AF Grobler

Co-supervisor: Dr T Bretschneider

Graduation: July 2018

Student number: 23440112

“For I know the plans I have for you”, declares the Lord, “plans to prosper you and not to harm you, plans to give you hope and a future.”

-Jeremiah 29:11

TABLE OF CONTENT

ABSTRACT

Metabolomics is a growing field and a valuable instrument for the identification of dysregulation in the metabolome of a biological system. Different approaches and analytical platforms are available for metabolomics based studies. Although metabolomics is a promising diagnostic tool there are still obstacles to overcome. There is still no standardised totally comprehensive approach available to detect and quantify large numbers of metabolites. There is also no standardised sample preparation and metabolite extraction method established.

Targeted metabolic profiling is a feasible approach to metabolomics and allows investigation into the metabolome with high specificity. The establishment of a metabolic profiling method will be of great benefit in the characterisation of diseases whose pathogenesis still remains poorly understood. Idiopathic pulmonary fibrosis (IPF) is a lung disease with a prevalence of between 1.25 and 23.4 per 100 000 population in Europe and 1 in every 32 000 population is South Africa. IPF is one of many diseases whose pathogenesis still remains poorly understood and alternative investigation is required in order to understanding the onset and progression of the disease.

During this study the aim was to develop an LC-MS/MS based targeted metabolic profiling method that would be able to generate a metabolic profile for any disease state, together with a sample preparation and metabolite extraction method for various biological matrices. The aim of the study was achieved by developing an LC-MS/MS method using the Luna NH_2 column (2 mm x 150 mm, 5 µm, 100 Å), as well as developing a standardised protein precipitation sample preparation procedure. After a quality assessment was performed on all aspects of the analytical process, including the range, linearity, limits of detection and quantification, accuracy and precision, the performance of the method was considered stable and adequate for use in metabolic profiling.

As validation of the developed method, a targeted metabolic profile was generated for a fibrotic lung animal model (C57BL/6J bleomycin treated mouse model) resembling IPF. Since sampling lung tissue from IPF patients is an invasive approach, the alternative approach of using an animal model resembling the diseases state was used. A metabolic profile was generated for the C57BL/6J bleomycin treated animal model using the developed method and after univariate and multivariate statistical analysis was performed, several metabolites were identified as significant (p-values < 0.05).

The metabolic profile was compared to a metabolic profile of a lipopolysaccharide induced lung inflammation mouse model to identity any correlation to an inflammation induced lung disease state. The metabolic profile of the C57BL/6J bleomycin treated mouse model was also

compared to a transforming growth factor-β treated normal human lung fibroblast cellular model to identify any correlation to an *in vitro* IPF representation. The identified metabolites indicated a dysregulation in the glycolysis pathway as well as the methionine cycle, suggesting the key to understanding the pathogenesis of the disease may lie on an epigenetic level.

Keywords: Metabolomics; targeted metabolic profiling; biomarkers; LC-MS/MS; lung fibrosis; IPF.

DECLARATION

I, **Maryke Venter**, hereby declare that this dissertation and all research work done for the completion of this dissertation is a record of my own work (except where citations or acknowledgements indicate otherwise) and that the study, in part or as a whole, has not been submitted to any other university.

I would like to acknowledge the following individuals or organizations for their contributions to my study:

- The Bioanalysis group within the Drug Discovery Sciences department of Boehringer Ingelheim. Pharma GmbH&Co. KG, Biberach an der Riß, Germany, for the use of their equipment and for the financial support.
- Dr Marc Kästle from the Immunology and Respiratory department at Boehringer Ingelheim. Pharma GmbH&Co. KG, Biberach an der Riß, Germany, for providing all animal tissue from already euthanized animals.
- Dr Mirco Christoph Sobotta and his team from the Immunology and Respiratory department at Boehringer Ingelheim. Pharma GmbH&Co. KG, Biberach an der Riß, Germany, for providing all cell culture samples used during this study.

ACKNOWLEDGMENTS

Firstly I would like to express my deepest gratitude to God, for blessing me with all the opportunities, support structure and abilities He provided in abundance with regards to the completion of this degree. Without His grace none of this would have been possible.

I would like to express my sincere gratitude to the following people for their contribution to this study and my life:

Prof Anne Grobler, my supervisor, thank you very much for the opportunities you gave me and all the support throughout the course of the study.

Dr. Tom Bretchneider, my assistant supervisor, thank you for investing so much time into not only this project but also my development as a scientist. Thank you for believing in me and your continuous patience and guidance throughout my time as your student. It was a tremendous privilege to learn from you and to work under your proficient guidance.

Anna-Katharina Lachenmaier and **Chris Cantow**, thank you very much for your expert advice and guidance regarding the analytical instruments. Thank you for your friendship and support and making my stay in Germany treasurable. It was a privilege to learn from you and work with you.

My parents, **Leon** and **Hendriette Venter**, words can't express how grateful I am for all your support through the years and sacrifices you made to ensure that I had the best possible opportunities. I am truly grateful for all your love and support.

My grandparents, **Sampie Duvekot** and **Joey Venter**, thank you for all your support throughout the years. Thank you for believing in me and encouraging me to be the best I can be.

Peter Benz, thank you very much for all your love and support and thank you for making each day with you a big adventure. I am truly grateful to have you in my life.

Lastly, I would like to thank **Dr Andreas Luippold** and **Boehringer Ingelheim** for allowing me to perform this study at the Drug Discovery Sciences department of Boehringer Ingelheim in Biberach an der Riß, Germany. I would also like to thank **Dr Marc Kästle**, **Dr Mirco Sobotta** and their teams for their support in this study.

LIST OF ABBREVIATIONS, SYMBOLS AND UNITS

Symbols and Units

%	–	Percentage
°C	–	Degrees Celsius
Å	–	Angstrom
β	–	Beta
Da	–	Dalton
g/mol	–	Gram per mole
L/min	–	Litre per minute
min	–	Minute
mg	–	Milligram
mL	–	Millilitre
mM	–	Millimolar
mm	–	Millimetre
mM/s	–	Millimolar per second
m/z	–	Mass to charge ratio
ng	–	Nanogram
ng/mL	–	Nanogram per millilitre
pH	–	The negative log of the hydronium ion concentration within a solution.
psi	–	Pound-force per square inch
R^2	–	Correlation coefficient
rpm	–	Revolutions Per Minute
s	–	Seconds
µL	–	Microlitre
µM	–	Micromolar
µm	–	Micrometre
V	–	Volt
v/v	–	Volume/volume
w/v	–	Weight/volume

Abbreviations

ACN: Acetonitrile

ADP: Adenosine Diphosphate

AGAT: Arginase

AMP: Adenosine Monophosphate

ANOVA: Analysis of variance

ATP: Adenosine Triphosphate

AUC: Area Under the Curve

cAMP: Cyclic Adenosine Monophosphate

CBS: Cystathionine Beta Synthase

CE: Collision Energy

cGMP: Cyclic Guanosine Monophosphate

CMP: Cytidine monophosphate

CO_2: Carbon dioxide

CoA: Coenzyme A

COPD: Chronic Obstructive Pulmonary Disease

CUR: Curtain gas

CV: Coefficient of Variation

CXP: Collision cell Exit Potential

dAMP: Deoxyadenosine Monophosphate

dCMP: Deoxycytidine Monophosphate

DHF: Dihydrofolate

DHFR: Dihydrofolate Reductase

DMSO: Dimethyl sulfoxide

DP: Declustering Potential

dTMP: Deoxyribosylthymine Monophosphate

EI: Electron Impact

EMA: European Medicines Agency

EP: Entrance Potential

ESI: Electron Spray Ionisation

FAD: Flavin Adenine Dinucleotide

FBS: Foetal Bovine Serum

GAMT: Guanidinoasetate Methyltransferase

GC-MS: Gas Chromatography linked to Mass Spectrometer

GDP: Guanosine Diphosphate

GMP: Guanosine Monophosphate

GS1: Nebulizer gas

GS2: Drying gas

GTP: Guanosine Triphosphate

HCA: Hierarchical Cluster Analysis

H_2O: Water

HILIC: Hydrophilic Interaction Liquid Chromatography

HPLC: High-Performance Liquid Chromatography

IPF: Idiopathic Pulmonary Fibrosis

IMP: Inosine Monophosphate

IS: Internal Standard

ISV: Ionspray Voltage

KOH: Potassium Hydroxide

LC: Liquid Chromatography

LC-MS: Liquid Chromatography linked to Mass Spectrometry

LC-MS/MS: Liquid Chromatography linked to Tandem Mass Spectrometry

LLE: Liquid-Liquid Extraction

LLOQ: Lower Limit of Quantification

LPS: Lipopolysaccharide

LOD: Limit of Detection

MAT: Methionine Adenosyl Transferase
MeOH: Methanol
MRM: Multiple Reaction Monitoring
MS: Mass Spectrometry
MS: Methionine Synthases
MS1: First analyser
MS2: Second analyser
MS/MS: Tandem Mass Spectrometry
mTHF: Methyl-Tetrahydrofolate
MTHFR: Methylene Tetrahydrofolate Reductase

NAD: Nicotinamide Adenine Dinucleotide (oxidised)
NADH: Nicotinamide Adenine Dinucleotide (reduced)
NADP: Nicotinamide Adenine Dinucleotide Phosphate (oxidised)
NADPH: Nicotinamide Adenine Dinucleotide Phosphate (reduced)
NaOH: Sodium Hydroxide
NH_2: Amidogen
NHLF: Normal Human Lung Fibroblasts
NMR: Nuclear Magnetic Resonance
NO: Nitric Oxide
NOS: Nitric Oxide Synthase
NP: Normal Phase

OH: Hydroxide

PBS: Phosphate Buffered Saline
PCA: Principal Component Analyses
PLS: Partial Least Square
PLS-DA: Partial Least Square Discriminant Analysis
PLS-R: Partial Least Square Regression
QC: Quality Control
QQC: Spiked Quality Control

rh-FGFB: recombinant human Fibroblast Growth Factor Basic

RP: Reverse Phase

RSD: Relative Standard Deviation

RT: Retention Time

SHMT: Serine Hydroxyl Methyltranferase

SPE: Solid Phase Extraction

SRM: Selected Reaction Monitoring

TBA: Tributylamine

TCA: Tricarboxylic acid

TGF-β: Transforming Growth Factor Beta

THF: Tetrahydrofolate

thrA: Aspartokinase/homoserine dehydrogenase 1

thrB: Homoserine kinase

thrC: Threonine syntetase

TEM: Temperature

UDP: Uridine Diphosphate

ULOQ: Upper Limit of Quantification

UMP: Uridine Monophosphate

LIST OF EQUATIONS

Chapter 3:

LIST OF FIGURES

Chapter 2:

Chapter 3:

Chapter 4:

Chapter 5:

LIST OF TABLES

Chapter 3:

Chapter 4:

Chapter 5:

CHAPTER 1: INTRODUCTION

1.1 Problem statement

Metabolomics is a growing field and a valuable instrument for the identification of dysregulation in the metabolome of a biological system. Metabolic profiling is a popular approach used in metabolomics studies. Metabolic profiling of diseases provides the opportunity to identify biomarkers that can be used for diagnostic purposes, determining the stage of the disease or identifying dysregulated metabolic pathways that can be targeted as treatment options. Although there are various approaches and analytical platforms available there is still no standardised comprehensive analysis and sample preparation available. An incorporation of all analytical platforms to sample analysis could provide the desired comprehensive overview, but this approach will not be feasible, for extensive sample preparation, analytical time and data processing will be required as well as expensive analytical platform instrumentations and expertise. A well designed LC-MS/MS targeted metabolic profiling method will allow high sensitivity and specificity and will be suitable for labile, non-volatile polar and non-polar metabolites in their native form.

Complex diseases with unknown pathogenesis have benefitted greatly from metabolomics based studies. Metabolomics based studies provided new insight into the pathogenesis of such diseases as well as providing new information for characterisation of the diseases. Idiopathic pulmonary fibrosis is an example of such a complex disease, which pathogenesis still remains poorly understood and can benefit from such an analysis. Therefore a standardised comprehensive metabolic profiling analysis would be greatly beneficial for idiopathic pulmonary fibrosis as well as providing the protocol to be followed for the generation of metabolic profiles for other complex diseases.

1.2 Aim and objectives

The aim of the study was to establish a standardised LC-MS/MS method for targeted metabolic profiling of biological matrices. The method included the identification of as many metabolites as possible, producing the possibility for characterising multiple diseases. The method development included the establishment of a standardised sample preparation and metabolite extraction protocol of biological matrices. The developed method was validated by producing a metabolic profile for a fibrotic lung animal model (C57BL/6J bleomycin treated mouse model) representing IPF.

This study was divided into three phases:

- Phase one consisted of the LC-MS/MS metabolite analysis method development. Different methods were tested to find a method that allows the detection of as many metabolites as possible. The parameters of the method were determined by the characteristics of the metabolites.
- Phase two consisted of the development of a standardised method for sample preparation and metabolite extraction from different matrices. Different sample preparation techniques, including quenching and metabolite extraction were evaluated to find a sample preparation method suitable for the developed LC-MS/MS method.
- Phase three consisted of the validation of the developed LC-MS/MS method by generating a targeted metabolic profile for a C57BL/6J bleomycin treated mouse model. For the establishment of the metabolic profile for the C57BL/6J bleomycin treated mouse model, healthy and diseased lung tissue were compared. A comparison was also made between the metabolic profile of the C57BL/6J bleomycin treated mouse model and the metabolic profile of a lipopolysaccharide induce lung inflammation mouse model as well as the metabolic profile of a transforming growth factor-β treated normal human lung fibroblast cellular model.

1.3 Structure of study

1.3.1 Chapter 2: Literature review

In this chapter a literature review is provided with regards to metabolomics, the different approaches and analytical platforms that are available as well as the different applications in which metabolomics is used. The focus of the review was on the development of an LC-MS/MS based targeted metabolic profiling method, with emphasis on the importance of this approach and highlighting each aspect of the method development process.

1.3.2 Chapter 3: Method development

In this chapter the method development process, which was performed during this study, is described in detail as well as all challenges that was experienced. This includes all aspects of the LC system, optimisation of the MS system parameters, sample preparation, data analysis and statistical analysis. A quality assessment of the analytical aspects of the method is also described in this chapter with details regarding each experimental procedure that was performed.

1.3.3 Chapter 4: LC-MS/MS method for targeted metabolic profiling

This chapter consists of a summary of the final method with regards to sample preparation, LC method, optimised MS parameters for each metabolite, performing data analysis and statistical analysis.

1.3.4 Chapter 5: Metabolic profiling of a fibrotic lung animal model

As validation of the developed method a targeted metabolic profile was generated for a fibrotic lung animal model resembling IPF. All results with regards to the generation of the metabolic profile of the C57BL/6J bleomycin treated mouse model, using the method as described in Chapter 4, are discussed in this chapter as a full length article. This article has been written according to the guidelines provided by the journal and has been submitted to the *Respiratory Medicine* journal.

1.3.5 Chapter 6: Summary and future prospects

In this chapter a summary of the study is given together with a review on the developed method. The final conclusion of the metabolic profile of the fibrotic lung animal model is provided and recommendations for future research in this area are also proposed.

1.3.6 Chapter 7: Reference

All references used in this study are provided in this chapter. The references are listed according to the requirements as specified in the NWU's manual for post-graduate studies.

1.3.7 Appendix A: Author guidelines

The author guidelines provided by the *Respiratory Medicine* journal is given. These guidelines were followed to write the article provided in Chapter 5.

CHAPTER 2: LITERATURE REVIEW

2.1 Metabolomics

Metabolomics is a growing field and a valuable instrument that involves the identification of metabolites produced in a biological system (Álvarez-Sánchez *et al.*, 2010a; Bino *et al.*, 2004; Kang *et al.*, 2016; Kottmann *et al.*, 2012). Metabolites represent not only the downstream output of the genome but also the upstream input from the environment (Wishart, 2016). With the identification of metabolites, endogenous and exogenous, the ability to identify specific alterations in metabolic pathways arises. This creates the possibility to link dysregulated metabolic pathways to diseases (Kottmann *et al.*, 2012). Metabolomics have been used in various applications, including investigations into disease pathogenesis, toxicology, drug discovery, and nutrition (Cuperlovic-Culf & Culf, 2016; Lu & Chen, 2017). Metabolomics have been used to determine the cause and pathogenesis of complex diseases (Kottmann *et al.*, 2012), as well as distinguishing between diseases showing similar clinical presentations (Adamko *et al.*, 2015). Different metabolomics approaches can be followed with the use of various analytical platforms fulfilling the requirements of each of the different applications.

2.1.1 Application of metabolomics

Metabolomics can be used for various applications (Cuperlovic-Culf & Culf, 2016; Lu & Chen, 2017), but a very important application is the characterisation of complex diseases' pathogenesis. There are still a great number of complex diseases that have not been characterised and the onset and progression of these diseases is still unknown. Characterising the metabolic profile of these diseases could potentially provide new insight into the pathogenesis of the disease and provide new therapeutic approaches. New insight into the pathogenesis of a disease can lead to identifying new biomarkers that can be used for earlier diagnosis of the disease as well as establishing the state and progression of the disease. Idiopathic pulmonary fibrosis (IPF) is one of such complex diseases that are of interest since the onset and progression of the disease is still unknown (Costabel *et al.*, 2014; Kottmann *et al.*, 2012; LaBrecque *et al.*, 2014).

2.1.1.1 Idiopathic pulmonary fibrosis

IPF is a disorder characterized by progressive destruction of normal lung architecture by alveolar epithelial cell injury, proliferation of activated fibroblasts and myofibroblasts, and accumulation of the extracellular matrix that stiffens the lung and leads to respiratory failure (Kang *et al.*, 2016; Richeldi *et al.*, 2014; Sandbo, 2014). IPF is one of several lung diseases that are characterized by pulmonary fibrosis. Although the commencement of pulmonary diseases

such as the fibroproliferative phase of acute respiratory distress syndrome or fibrotic sarcoidosis has been characterized, the underlying etiology and pathogenesis of IPF still remains poorly understood (Sandbo, 2014). IPF has a prevalence of between 1.25 and 23.4 per 100 000 population in Europe, between 42.7 and 63 per 100 000 population in America and 1 in every 32 000 population in South Africa (Masekela *et al.*, 2016; Nalysnyk *et al.*, 2012). The survival duration from time of diagnosis for IPF patients are 2 to 3 year (Kottmann *et al.*, 2012). Although there is affective treatment available, Nintedanib (Boehringer Ingelheim Pharma GmbH & Co. KG, Germany) and Pirfenidone (Genentech Inc. member of the Roche Group, South San Francisco, CA, USA), that reduces the decline in lung function, the treatment still do not offer full recovery (Costabel *et al.*, 2014; King Jr *et al.*, 2014; Richeldi *et al.*, 2014). Therefore the need for further research into the pathogenesis of this disease is crucial.

2.1.1.2 Metabolomics and its application to respiratory diseases

Metabolomics have been used to establish metabolic profiles for several complex respiratory diseases including asthma, chronic obstructive pulmonary disease (COPD), cystic fibrosis and acute respiratory disease syndrome (ARDS) (Kang *et al.*, 2016; Stringer *et al.*, 2016). These metabolic profiles contribute to a better understanding of the pathogenesis and progression of these diseases. Biomarkers were also identified and may be used to distinguish between diseases that have similar clinical presentations (Adamko *et al.*, 2015; Stringer *et al.*, 2016). Although asthma and COPD have different pathogenesis and the degree of inflammation and cellular damage varies with the severity of the disease, it is still difficult to differentiate between them since these diseases have similar clinical presentations (Adamko *et al.*, 2015). Metabolic profiling of urine samples from patients with asthma and COPD, respectively, showed that 3-hydroxyisovalerate, taurine, histidine and succinate were identified as distinguishable biomarkers, since the changes in the levels of these metabolite concentration were significantly different between the patients with asthma and the patients with COPD (Adamko *et al.*, 2015).

Although metabolic profiles have been established for several respiratory diseases, metabolic profiling for IPF has not yet been fully investigated (Kang *et al.*, 2016; Rindlisbacher *et al.*, 2018). Previous biomarker identification in IPF lung tissue indicates increased levels of inosine, hypoxanthine and glycolytic intermediate metabolites including lactic acid (Kottmann *et al.*, 2012) and a decrease in adenosine triphosphate (ATP) and glucose (Kang *et al.*, 2016). These phenomena are also seen in cancer cells which portrays the Warburg effect (Cottrill & Chan, 2013). Tumours present a high energy and anabolic need to ensure rapid cell growth and proliferation. The serine biosynthesis pathway was recently identified as an important source for necessary metabolic intermediates for these dysregulated processes and it is of great interest to see in which other diseases a dysregulation in the glycolysis pathway and serine biosynthesis pathway can be seen (Cottrill & Chan, 2013).

2.1.2 Challenges for metabolomics

Although metabolomics is a powerful tool there are still shortcomings that need to be addressed. One of the major shortcomings is the lack of a total comprehensive approach allowing a global view on the metabolome of a biological system (Dettmer *et al.*, 2007). This is a difficult task for metabolites differ widely in characteristics and detectible concentration (Álvarez-Sánchez *et al.*, 2010a; Bino *et al.*, 2004). The most common strategy that is used to address this problem involves the integration of various analytical platforms. This improves metabolite coverage and increases the identification range, but feasibility is a great concern.

Another challenge that have to be addressed is the lack of a standardised sample preparation and metabolite extraction protocol (Álvarez-Sánchez *et al.*, 2010a; Bino *et al.*, 2004). The compilation of a standardised sample preparation protocol has been disregarded over the years. This is a difficult task as samples differ widely in matrix diversity and metabolite composition. The treatment and perturbations of the experiment can also influence the sample preparation procedure. A standardised samples preparation protocol for all biological samples is needed for comparable and reproducible results (Álvarez-Sánchez *et al.*, 2010a; Bino *et al.*, 2004).

Not only is the sample preparation and analytical approach of great importance but also the sample handling, storage and data handling. In metabolomics large volumes of data is generated and analysing such complex data sets has an impact on the quality of the identification and quantification of metabolites and interpretation of biological relevance (Dunn *et al.*, 2012). The analytical approach as well as the analytical platform greatly influences the volume of data handling that has to take place. With an untargeted approach the data processing increases tremendously since a lot more data clean-up and pre-processing are required, including peak identification for identification of each metabolite. This requires specialised programs, databases and experience (Dunn *et al.*, 2012; Godzien *et al.*, 2015).

2.1.3 Different approaches for metabolomics

Metabolomics joins genomics, transcriptomics, and proteomics in the field of omics and enables a greater understanding of a biological system (Rochfort, 2005; Wishart, 2016). Metabolites represent the final downstream products of genomic, transcriptomic and proteomic processes (see Figure 2.1). The number of metabolites that can be evaluated is much lower than the number of genes, transcriptomes and proteins. Genomics involves the study of about 25 000 genes, transcriptomics involves about 100 000 RNA transcripts and proteomics about 1 000 000 proteins (Solomon & Fischer, 2010; Theodoridis *et al.*, 2011). The advantage of metabolomics is that metabolites serves as a direct signature of biochemical activity and provides a better correlation with the phenotype (Patti *et al.*, 2012; Zhang *et al.*, 2012a).

The term “metabolomics” is an ‘umbrella term’ for many types of different approaches that involves the investigation into the metabolome of a biological system. The different approaches that can be taken are targeted metabolomics, untargeted metabolomics, metabolic footprinting, metabolic fingerprinting, fluxomics, lipidomics, metallomics and exposomics (Patti *et al.*, 2012; Wishart, 2016). The approach depends on the type of sample to be measured, requirements and aim of the study (Johnson & Gonzalez, 2012).

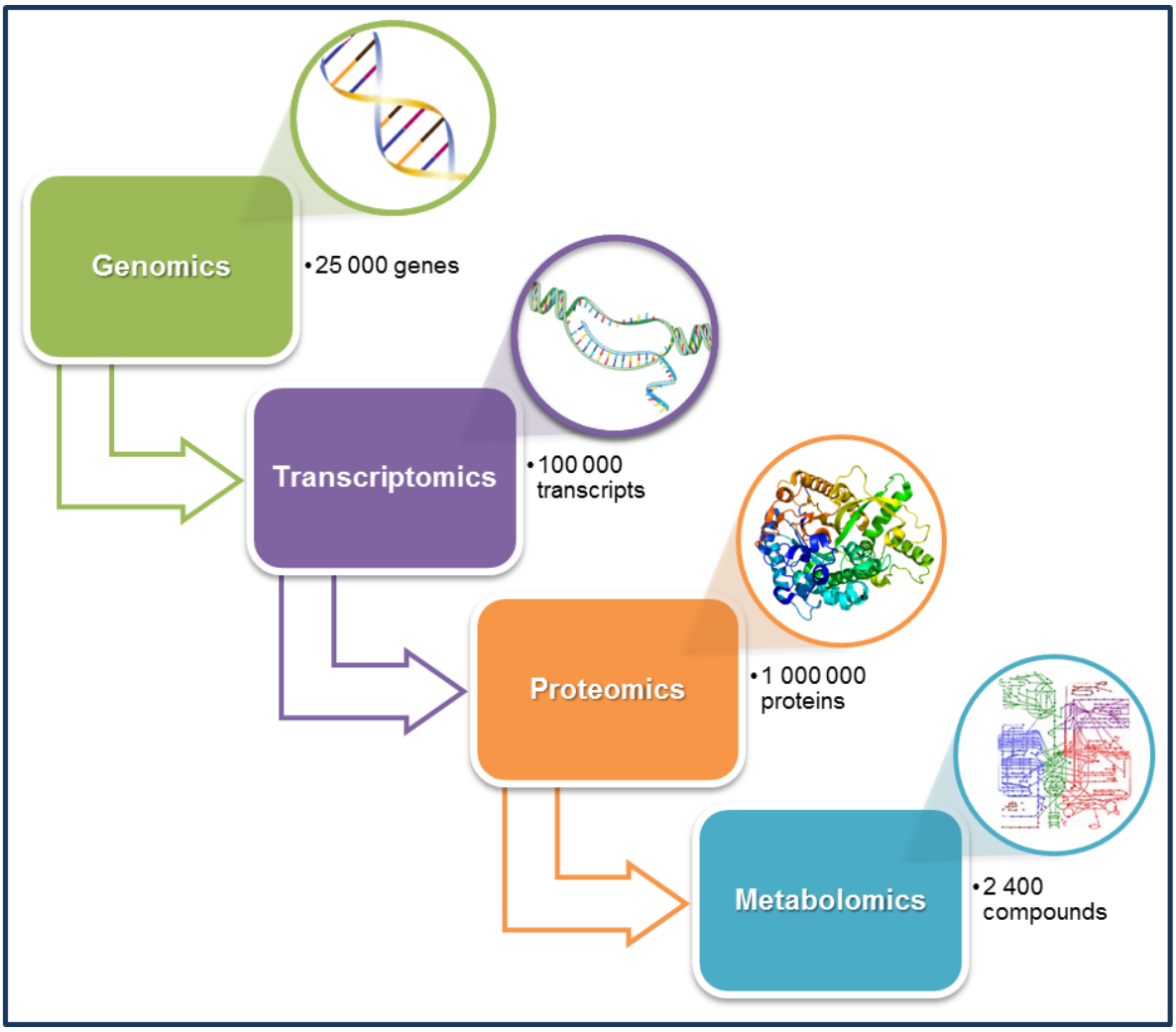

Figure 2.1: Central dogma illustrated by different ‘omics. Metabolomics represents the downstream product of genomics, transcriptomics and proteomic processes. The number of metabolites to be investigated is magnitudes less than the amount genes, transcriptomes and proteins from the other ‘omics (Johnson & Gonzalez, 2012; Roberts *et al.*, 2012; Solomon & Fischer, 2010; Theodoridis *et al.*, 2011)**.** Permission for the use of the metabolic profile map as graphics was gained from the Kyoto Encyclopedia of Genes and Genomes data base (Genome.jp, 2017). The remaining graphics used in this figure were gained from open-source websites (Art, 2017; En.wikipedia.org., 2017; Pixabay.com., 2017).

Targeted metabolomics is driven by a specific biological question and consists of a method measuring a specific list of metabolites. Usually the focus would be on metabolites of a specific pathway. By limiting a study to only metabolites of interest, a well-designed method can be created with optimized sample preparation and analytical parameters that will ensure high sensitivity and specificity (Griffiths *et al.*, 2010; Patti *et al.*, 2012; Wishart, 2016). Untargeted metabolomics consists of a method that is used to measure as many metabolites as possible

from a biological sample without bias. This approach provides a global scope of the biological system with minimal limitations. The limitations associated with untargeted metabolomics are related to the instrumentation used for analysis and identification of the different metabolites (Patti *et al.*, 2012; Wishart, 2016).

Metabolic footprinting refers to the analysis of extracellular metabolites, intended to define the pattern of extracellular metabolites. Metabolic fingerprinting refers to the analysis of intracellular metabolites. These approaches can provide a better understanding of cellular communication mechanisms (Mapelli *et al.*, 2008). Fluxomics is an approach that aims to define the genes involved in regulation by monitoring the flux of a single metabolite (Wiechert *et al.*, 2007). Lipidomics involves the large scale analysis of cellular lipids (Han, 2009). Metallomics refers to the analysis of elemental species and exposomics involves the study of the complete collection of environmental exposures (Szpunar, 2004).

Not only are there different approaches to metabolomics, but there are also diverse applications for it, including *in vivo* and *in vitro* studies of human and animal health, biomarker discovery, drug discovery and development, plant biology, microbiology, food chemistry and environmental monitoring (Wishart, 2016; Zhang *et al.*, 2012a). The diverse applications is due to the wide range of substrates that can be used, solids (tissue, biological waste and soil), liquids (biofluids, effluent and water) and gases (breath, fumes and scents) (Wishart, 2016).

2.2 Targeted metabolic profiling

With metabolomics there are several approaches that can be followed, which involves different analytical platforms. Incorporation of the different analytical platforms will provide a comprehensive approach for metabolic profiling but it requires extensive resources and experience and is not always feasible (Roberts *et al.*, 2012). Targeted metabolic profiling is limited to the identified metabolites for analysis but a well-designed analytical platform will minimise these limitations.

When setting up a well-designed targeted metabolic profiling method several factors have to be taken into account (see Figure 2.2). Since a targeted approach is limited to the identified metabolites of interest, care have to be taken when identifying these metabolites. After the metabolites of interests have been identified an analytical platform for identification has to be selected. The analytical platform has to be suitable for the analysis of the identified metabolites. It is important to consider all aspects of the analytical work flow to ensure a robust, and feasible method with the high sensitivity and specificity will be generated (Dunn *et al.*, 2005). Once the analytical platform has been chosen an appropriate sample preparation have to be selected to ensure the highest possible recovery of the identified metabolites (Dettmer *et al.*, 2007). The analysis of the samples is also important with regards to quality control, to ensure reliable and

repeatable data can be generated. Data handling and normalisations is very important especially for statistical analysis that follows (Boccard *et al.*, 2010).

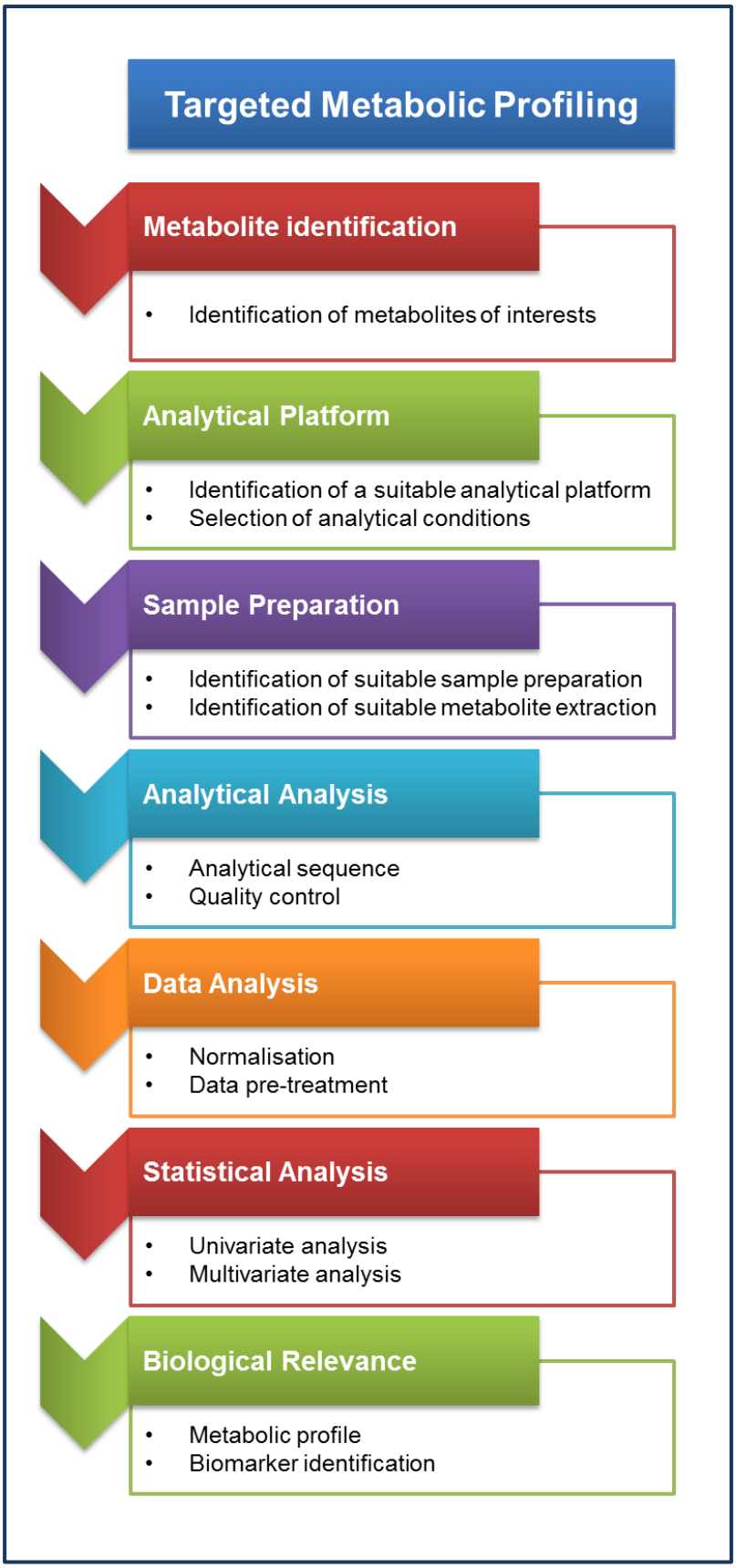

Figure 2.2: A flow chart summarising the important aspects of a targeted metabolic profiling method.

2.2.1 Metabolite identification

The metabolome of a biological system is of interest because of its direct correlation to the phenotype (Patti *et al.*, 2012; Zhang & Kaminski, 2012b). Characterising the metabolome of a biological system will provide greater understanding of any disease state. Many different metabolic pathways are present in the metabolome and most of these pathways are interlinked. When identifying metabolites to be included in a targeted metabolic profiling method these cross interactions of metabolites to various metabolic pathways have to be taken into account.

The central carbon metabolism includes the major energy production metabolic pathways and provides crucial information about the energy state of a biological system. Not only is the central carbon metabolism of great importance but also metabolic pathways such as nucleotide and protein biosynthesis, lipid and phospholipid turnover and redox stress (Armitage & Barbas, 2014). Metabolites in these metabolic pathways have been identified as biomarkers for dysregulation of cell growth and proliferation (Locasale, 2013). Figure 2.3 highlights important metabolites from the central carbon metabolism, serine biosynthesis pathway and methionine and folate cycle.

2.2.1.1 The central carbon system

In all biological systems the central carbon metabolism plays a key role in substrate degradation, energy and cofactor regeneration and biosynthetic precursor supply. The central carbon metabolism consists of the glycolysis, pentose-phosphate-pathway, tricarboxylic acid cycle (TCA) and the corresponding cofactors involved (see Figure 2.4). In the understanding of the central carbon metabolism has been of great importance in biotechnological production of fine chemicals, such as amino acids, vitamins, and antibiotics (Luo *et al.*, 2007). Determination of concentration and concentration dynamics of the central carbon metabolism provides key information of the metabolic state of a biological system (Luo *et al.*, 2007).

Investigations into the central carbon metabolism have been responsible for generating essential hypothesis in fields such as cancer research. During an investigation of the relationship between glycolysis, the TCA cycle and oxidative phosphorylation, a valuable hypothesis was generated, indicating that pyruvate was converted to lactate rather than fuelling the TCA cycle even in aerobic conditions. The hypothesis is known as the Warburg effect and have been identified in some cancer cell types and have provided essential information in characterising and understanding the metabolic state of these cancer cells (Armitage & Barbas, 2014).

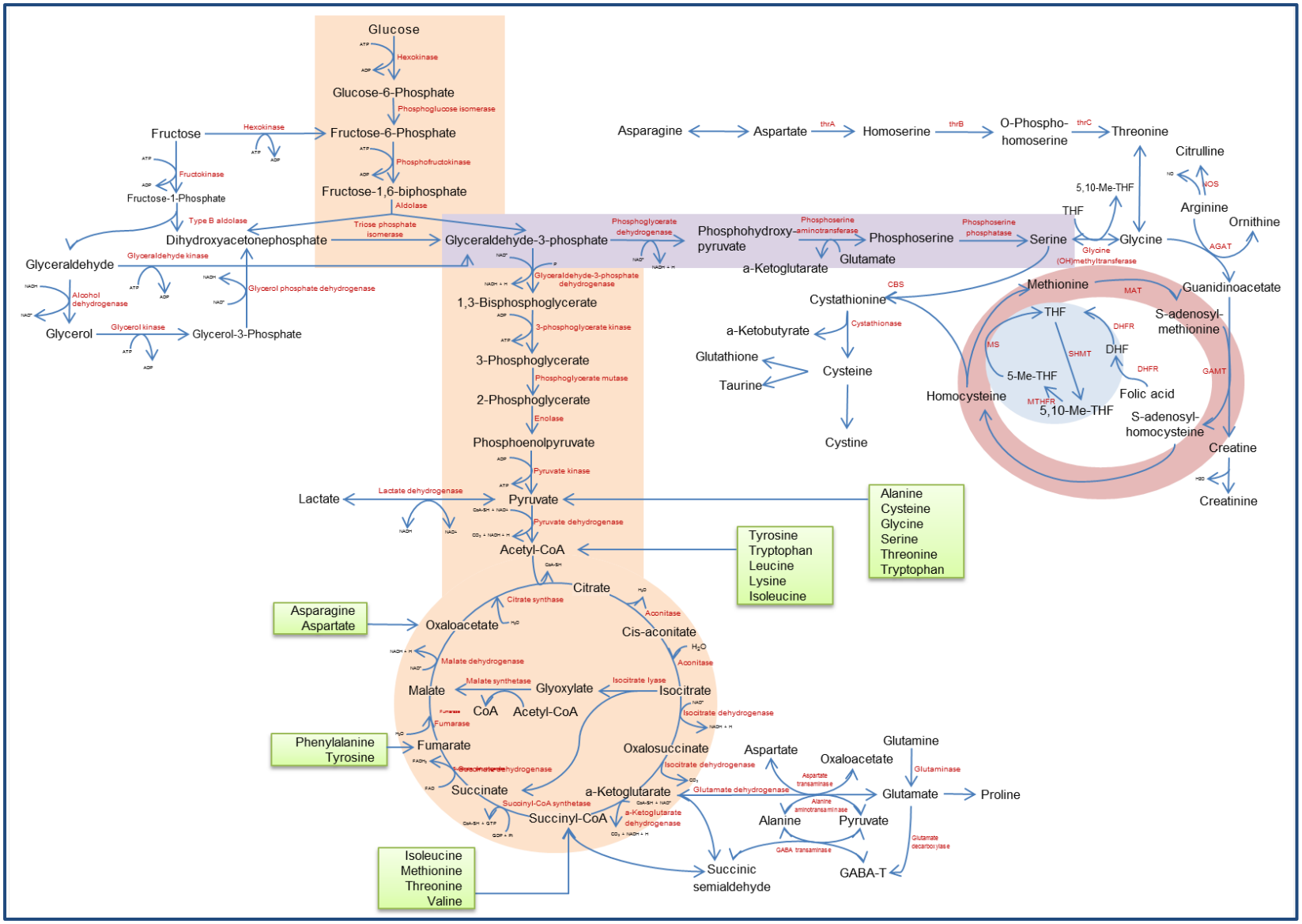

Figure 2.3: Extended central carbon metabolic pathway. Metabolic chart of the central carbon metabolism including the fructose metabolic pathway, amino acid entry point into the citric cycle, serine biosynthesis pathway and the methionine and folate cycle. This diagram was generated from information obtained from the Kyoto Encyclopedia of Genes and Genomes data base (Genome.jp, 2017).

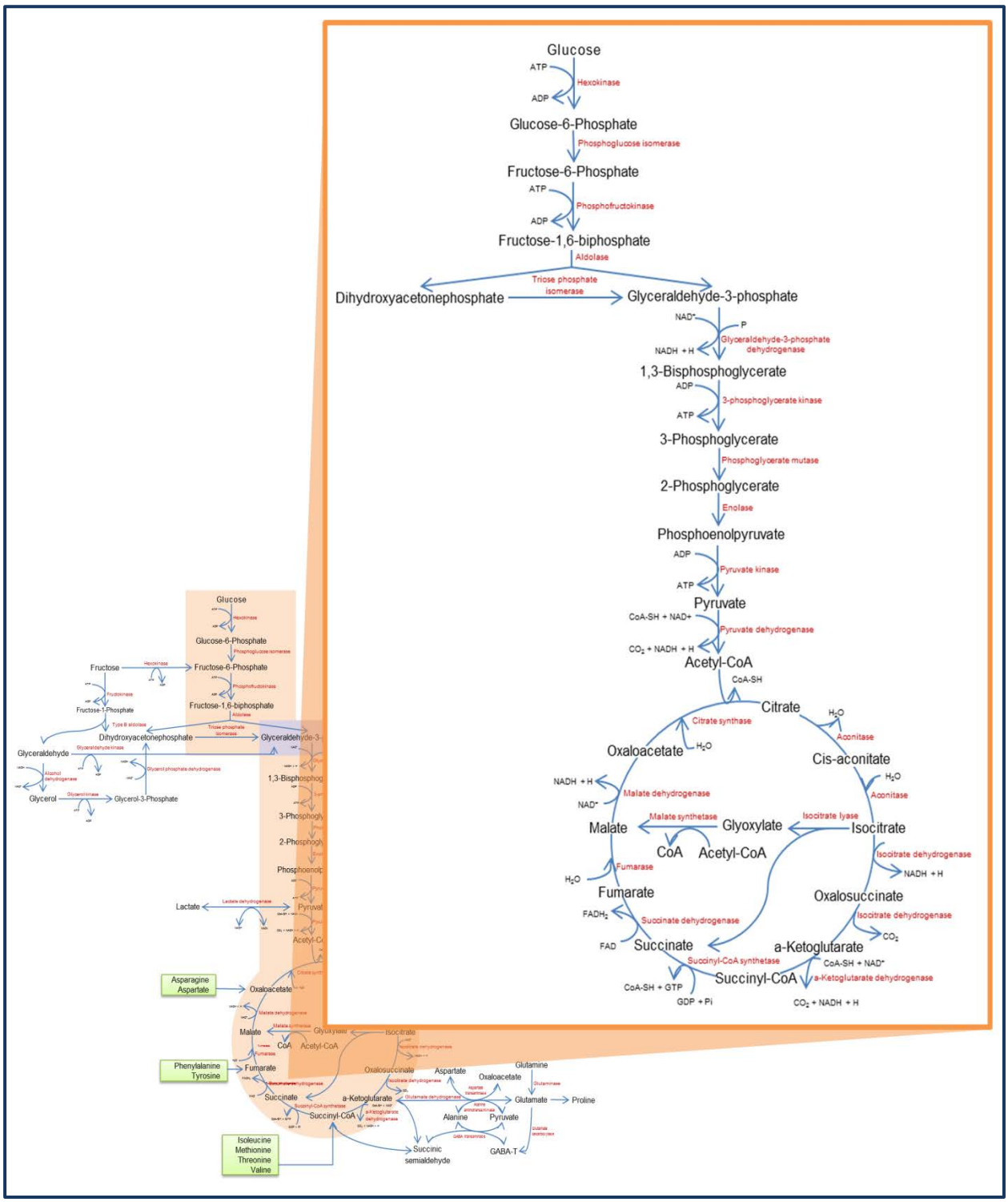

Figure 2.4: The central carbon metabolism. The metabolic chart highlights intermediates of the glycolysis and TCA cycle that provide key information of the energetic state of a biological system. This diagram is an enlargement of Figure 2.3 and was generated from information obtained from the Kyoto Encyclopedia of Genes and Genomes data base (Genome.jp, 2017).

2.2.1.2 Amino acids

Cell growth and proliferation requires proteins, lipids and nucleic acid for the construction of new cellular components as well as the maintenance of cellular redox, genetic and epigenetic status (Locasale, 2013). Metabolic pathways of amino acids such as glycine, serine, proline, histidine, methionine and phenylalanine should also be investigated since these amino acids have been identified as biomarkers in other disease states, including respiratory diseases such as asthma and COPD, and can provide an understanding of the metabolic state of a biological system (Armitage & Barbas, 2014).

2.2.1.3 Serine, glycine and one-carbon metabolism

The serine biosynthetic pathway has recently been identified as an important source of metabolic intermediates in assisting the high energetic and anabolic need for rapid cell growth and proliferation of tumours (DeNicola *et al.*, 2015). The one-carbon metabolism involves the folate and methionine cycles, integrates nutritional status from amino acids, glucose and vitamins, as well as generates biosynthesis of lipids, nucleotides and proteins and maintains the redox status of substrates for methylation reactions (see Figure 2.5) (Locasale, 2013). Therefore input metabolites, intermediates and metabolic products of the one-carbon metabolism are of great interest in metabolic profiling.

The one-carbon metabolism involves a complex metabolic network that is based on the chemical reaction of folate compounds. These reactions proceed in a cyclic manner during which a carbon unit is transferred to other metabolic pathways. Folic acid is a member of the vitamin B group and is reduced by a series of enzymes, leading to the generation of tetrahydrofolate (THF). THF participates in a number of metabolic reactions, which involves the movement of carbon atoms. The folate cycle is coupled to the methionine cycle through the generation of methyl-THF (mTHF) (see Figure 2.5). The trans-sulphuration pathway is coupled to the methionine cycle and serine is metabolised to glutathione via the trans-sulphuration pathway (see Figure 2.5). Serine and glycine serves as the main metabolites for the entry point into the one-carbon metabolism, but there are several entry points for both serine and glycine. Serine can be synthesised *de novo* via the serine biosynthesis pathway but can also be imported into the cell via amino acid transporters. The enzymatic cleavage of glycine can fuel the folate cycle by the generation of a carbon unit for the methylation of THF. Glycine can also be generated from many sources including choline, betaine, dimethylglycine, sarcosine and in some cells from threonine (Locasale, 2013).

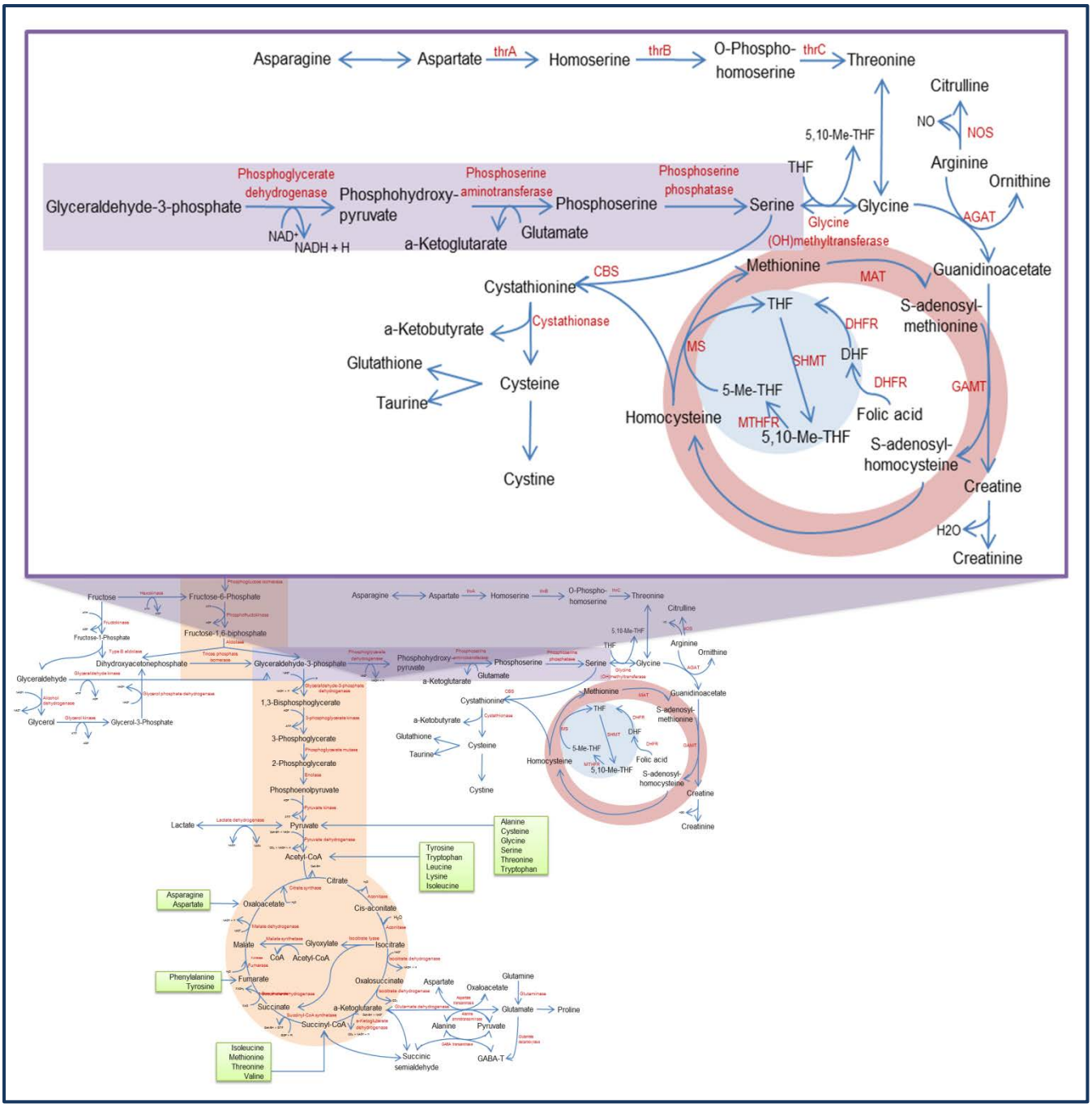

Figure 2.5: Serine biosynthesis pathway and the methionine and folate cycles. This diagram is an enlargement of Figure 2.3 and was generated by using information obtained from the Kyoto Encyclopedia of Genes and Genomes data base (Genome.jp, 2017).

2.2.2 Analytical platforms

The ultimate metabolic profiling platform would involve analysis directly on the sample, without sample preparation or storage, and provide unbiased results with respect to different metabolite classes. The analysis would have to be highly and equally sensitive to all the metabolites present in the sample, have a wide dynamic range and be robust and reproducible. Accurate and fast metabolite identification would also be needed (Theodoridis *et al.*, 2011). Unfortunately there is currently no analysis that can provide all these desired properties (Vuckovic, 2012).

Metabolites have a broad range of characteristics and abundance within a biological system. To be able to achieve the desired metabolite profile a wide range of instruments is used. The primary instrumentation used for metabolomics are: nuclear magnetic resonance (NMR) spectrometry, gas chromatography mass spectrometry (GC-MS), and liquid chromatography mass spectrometry (LC-MS) (Zhang *et al.*, 2012a). Every technique has its own advantages and disadvantages. Using multiple techniques will limit the shortcomings of a single-analysis technique (Zhang *et al.*, 2012a).

2.2.2.1 Nuclear magnetic resonance spectrometry

Nuclear magnetic resonance spectrometry (NMR) is a powerful and one of the most widely used techniques. This technique provides a wealth of structural information about analytes. Information such as chemical shift, spin-spin coupling and relaxation or diffusion enables fast identification of analytes in a sample. Sample preparation for NMR is straightforward with minimal preparation steps, samples are in a solution state with the addition of a deuterated solvent. The analysis is non-destructive and does not require the pre-selection of analytical conditions such as the ion source conditions or the selection of a stationary phase, mobile phase or temperature as in the case of chromatographic techniques. NMR provides many advantages but the sensitivity is poor and the concentration of potential biomarkers might be below the detection limit. A reverence library is also needed for identification purposes as well as a specialist operator of the instrument (Dunn *et al.*, 2005; Theodoridis *et al.*, 2011; Zhang *et al.*, 2012a).

2.2.2.2 Gas chromatography linked to mass spectrometry

Gas chromatography linked to mass spectrometry (GC-MS) has been used as a platform in non-targeted analysis and is especially used for hydrophilic metabolites. There are well defined spectral libraries available generated from GC-MS electron impact (EI). These libraries provide easy identification of unknown analytes by using well known and defined retention time or retention index. GC-MS requires sample derivatization, to be able to create volatile compounds. Compounds that are large, thermo-labile or are unable to be derivatized will not be detected by GC-MS analysis. Sample preparation is extensive and time consuming and a high-throughput technology is required to handle large volume of samples (Dunn *et al.*, 2005; Theodoridis *et al.*, 2011; Zhang *et al.*, 2012a).

2.2.2.3 Liquid chromatography linked to mass spectrometry

Liquid chromatography linked to mass spectrometry (LC-MS) is a well-known and the most widely used metabolomics platform (Lu & Chen, 2017). Sample preparation is minimal but depends on the type of chromatography. In most cases there is no need for derivatization of compounds prior to analysis. LC separation is better suited for the analysis of labile and non-

volatile polar and non-polar compounds in their native form. LC-MS provides high resolution and reproducible measurements that sets up the basis for subsequent data processing and multivariate data analysis. LC-MS analysis is limited by the characteristics and capacity of the column chosen for the analysis. Different groups of analytes can be measured on different types of columns.

Different columns for LC-MS

With reverse-phase (RP) LC the mobile phase is varied starting with a higher polarity solvent than the stationary phase and eluting analytes by increasing the organic solvent concentration in the eluent. The retention of non-polar compounds increases with an increase in polarity of the mobile phase. To be able to induce retention of polar compounds with reversed-phase chromatography, techniques such as solute derivatization or ion-pairing is used. With ion-pairing a reagent of opposite charge to an ionic compound is introduced into the mobile phase to form a non-covalent adduct with the ionic compound (Pesek & Matyska, 2007).

With normal-phase (NP) LC the stationary phase has a higher polarity than the solvent used in the mobile phase. Polar compounds are more strongly retained than non-polar compounds when the mobile phase polarity decreases. This type of chromatography is used enable retention, separation and detection of polar compounds (Pesek & Matyska, 2007). A widely used NPLC type column is the hydrophilic interaction liquid chromatography (HILIC) column.

A HILIC column is designed to retain and separate polar-ionic compounds from each other. This is achieved by the polar properties of the stationary phase of the column together with a high concentration non-polar (organic) solvent in the composition of the mobile phase (Pesek & Matyska, 2007). The disadvantage of a HILIC column is that typical hydrophobic compounds will have little or no retention. This disadvantage can be limited by the different functional groups that can be present on the stationary phase (Pesek & Matyska, 2007). Examples include aminopropyl ligands bound to silica, alkylamide packing phase and a mixed phase containing different types of ligands (-NH_2, -CN. –phenyl, -C_8, -C_{18}) (Buszewski & Noga, 2012).

2.2.3 Sample preparation

Sample preparation is dependent on the type of approach and analytical platform of choice and is an important aspect of method development since it is responsible for reproducibility. Different analytical platforms require different sample preparation procedures. With GC-MS, derivatization is needed to ensure all metabolites of interest are volatile. With LC-MS, metabolites have to be dissolved in solvent, water or organic phase. Quenching is an important step with regards to metabolomics. Quenching ensure representativeness of samples by efficiently interrupting the metabolism (Álvarez-Sánchez *et al.*, 2010b). Another important step in

sample preparation is metabolite extraction. Several extraction methods are available: dilute and shoot, protein precipitation, solid phase extraction, liquid-liquid extraction and ultrafiltration.

Sample preparation depends on the analytical method to be used and the purpose of the study. When a sample preparation protocol is set up a few steps have to be addresses. The selection of biological material (e.g. blood, urine, cells or tissue) together with the appropriate sampling technique is one of the limiting steps in metabolomics and its selection is based on the purpose and scientific question of the study. The sampling of biological material should ideally be non-invasive and repeatable. With most biological material used for metabolomics a quenching step is required directly after sample collection. Quenching ensures a rapid interruption of the metabolism. This is particularly imported when working with cell cultures and tissue since the metabolism can be altered by enzymatic interactions. Metabolite extraction is a very important step and care should be taken when selecting the appropriate extraction method. The purpose of the study should be clearly defined; whether or not all metabolites are of interest or only specific metabolites. Knowledge about the metabolites of interest is required to choose between liquid-liquid extraction, solid phase extraction or protein precipitation. It is also necessary to know the characteristics of the metabolites of interest to enable selection of the appropriate clean up method. The analytical method should also be defined prior to sample preparation to determine in which solvent the metabolite extraction should be prepared, with regards to LC-MS/MS or which derivatization has to be used for GC-MS/MS (Álvarez-Sánchez *et al.*, 2010a; Bino *et al.*, 2004).

2.2.3.1 Derivatization

Derivatization is a chemical modification of an analyte target structure. Derivatization is one of the most effective methods used to improve the detection of metabolites in GC- or LC-MS, by improving the binding characteristics of metabolites to LC columns, making metabolites volatile for GC-MS detection and stabilising metabolites (Aretz & Meierhofer, 2016). Although sensitivity is gained with derivatization of the metabolites there are also some drawbacks to derivatization. Not all metabolites can be derivatized with the same reagent and several extra preparation steps have to be included into the sample preparation. Furthermore, the mass spectra are different in terms of parent and fragment masses, which complicate the identification of the metabolites (Aretz & Meierhofer, 2016).

2.2.3.2 Quenching

Quenching is an important step in sample preparation in metabolomics studies. Quenching aims at stopping metabolism instantly by inhibiting endogenous enzymes. It ensures suppression of change in the metabolic profile during sample preparation and minimizes variability among samples (Álvarez-Sánchez *et al.*, 2010b).

There are some requirements that have to be met for quenching to be successful. Since the turnover rates of many primary metabolites is very fast (in the range of 1 mM/s) and the concentration of the different metabolites vary greatly, the inactivation of the metabolism should be faster than the metabolic changes occurring in a sample (Álvarez-Sánchez *et al.*, 2010b). Sample integrity is also very important particularly when working with cells. Care should be taken to preserve sample integrity during the sample preparation procedure. With regards to cells, leakages of intracellular metabolites should be minimized to ensure accurate representation of the sample. When quenching is performed care have to be taken to ensure that no significant variations are induced with regards to chemical and physical properties or concentration of the metabolites (Álvarez-Sánchez *et al.*, 2010b). This also applies to the storage of samples since incorrect storage of samples can influence the stability of metabolites.

Common strategies for quenching are based on rapid modification of sample conditions and usually include rapid change in pH or temperature. With regards to pH modification, quenching is achieved by instantly changing to extreme pH. The addition of potassium hydroxide (KOH) or sodium hydroxide (NaOH) will achieve a high alkaline pH. The addition of perchloric, hydrochloric or trichloroacetic acid will achieve high acidic pH. With regards to temperature modification, quenching is mainly carried out by cooling to lower than -20°C. One of the most popular methods is cold methanol quenching. This allows a rapid interruption of the metabolism in a sub-second time scale. This approach is used especially to discriminate between intracellular and extracellular metabolites (Álvarez-Sánchez *et al.*, 2010b).

2.2.3.3 Metabolite extraction

Metabolite extraction aims to efficiently release metabolites from the sample, removes impurities that can complicate the analysis (e.g. salts and proteins), concentrates trace metabolites before analysis as well as ensuring compatibility between the extract and the analytical technique (Álvarez-Sánchez *et al.*, 2010b). The metabolite extraction is an important step in the metabolomics analytical process and the effectiveness directly affects the quality of the final data. With the elimination of impurities that metabolite extraction provides, complications such as ionisation suppression is reduced (Vuckovic, 2012). Different extraction methods are available and the choice depends on the selectivity required by the chosen metabolomics approach. The extraction efficiency is limited by the solubility of the metabolites. Common extraction techniques include dilute and shoot, protein precipitation, liquid-liquid extraction (LLE), solid-phase extraction (SPE) or ultrafiltration extraction (Álvarez-Sánchez *et al.*, 2010b; Henion *et al.*, 1998; Vuckovic, 2012).

Dilute and shoot

Metabolomics studies require a sample preparation protocol that allows the analysis of all metabolites of interest. The dilute and shoot method exclude all other sample preparation steps

that can influence the abundance of metabolites (Henion *et al.*, 1998). A typical dilution factor that is used is between 1:1 and 1:10 with a solvent appropriate for the analysis. With this technique it is important that the metabolites are in a relatively high concentration and that the matrix components do not elute at the same time as the metabolites to ensure that the matrix compounds do not interfere with the ionisation of the metabolites (Henion *et al.*, 1998).

Protein precipitation

The most commonly used method of protein precipitation is the addition of an organic solvent to the sample homogenate. The addition of organic solvent not only removes proteins from the sample but also disrupts any binding between the metabolites and the proteins (Vuckovic, 2012). This allows the representation of total metabolite concentration. Acetonitrile, methanol, ethanol and acetone are some of the most effective organic solvent used in protein precipitation. Mixtures of these solvent are also used to accommodate the chosen approach and analytical platform as well as to increase metabolite coverage and robustness of the sample preparation technique. Acetonitrile and methanol have shown to result in the highest protein precipitation and allows a wide range of metabolites to be analysed (Gika & Theodoridis, 2011). A popular dilution ratio used with protein precipitation is 1:4, ensuring all protein is precipitated with the least dilution (Vuckovic, 2012).

Liquid-liquid extraction

LLE usually involves mixing an aqueous sample solution with an equal volume of immiscible organic solvent. The two immiscible liquid phases interact with the intent to extract metabolites from the aqueous layer into the organic layer. There are many factors that affect the recovery and selectivity of the metabolites from the aqueous solution. These factors include metabolite solubility and pK_a, as well as the pH and ionic strength of the solution (Henion *et al.*, 1998). Centrifugation is then used to separate the immiscible liquids, with the organic layer containing the metabolites of interest. The organic layer is then removed and concentrated by evaporation before reconstituting it in an appropriate solvent for LC/MS analysis. This extraction technique can provide high recovery of the metabolites of interest but the procedure is not amenable to automation and a great amount of metabolites are lost due to high selectivity (Henion *et al.*, 1998).

Solid-phase extraction

SPE is a less popular extraction method in terms of global metabolomics approaches. A large amount of sorbent is used, typically in cartridge format, to extract metabolites from a sample. The metabolites are subsequently removed from the sorbent by solvent elution (Vuckovic, 2012). With SPE metabolite pre-concentration can be achieved and matrix effect can be limited, increasing column lifetime. The main disadvantage of SPE is that it is highly selective and

decreases the metabolite coverage, making it unsuitable for global metabolomics studies (Vuckovic, 2012).

Ultrafiltration extraction

Ultrafiltration involves the filtration of a sample through a specific filter that only allows molecules of selected molecular masses to pass through. This is a simple technique where filtration is achieved by applying pressure through centrifugation. Typical molecular mass cut-offs are: 3000 Da, 10 000 Da, 30 000 Da. With the use of a 3000 Da cut off, protein and macromolecule elimination can be achieved. The main disadvantage of ultrafiltration is the significant loss of metabolites with hydrophobic properties (Vuckovic, 2012).

2.2.4 Analytical analysis

Prior to the analysis of samples, some factors have to be taken into account to ensure the analysis is of high quality. These factors include quality control samples (QC), spiked samples and the use of internal standards (Godzien *et al.*, 2015). The monitoring of these factors is an important indication of the quality of the data generated by the analysis. Other factors, such as the analytical run sequence, are also important to prevent any significant false variation among the experimental groups.

2.2.4.1 Quality control samples

QC samples are analysed at the start and end of an analytical run as well as at intermitted points throughout the analytical run. The function of QC samples is the monitoring of the performance of the method and instrumentation. In metabolomics studies, different approaches to QC sample preparation can be followed. A popular QC sample preparation is to pool equal aliquots of all samples to be measured in a batch. A less popular preparation approach involves only pooling and analysing a specific group (e.g. control group). The big disadvantage of this approach is that the QC samples do not provide an accurate representation of all the samples to be analysed (Godzien *et al.*, 2015).

In validated methods, a QC sample is usually spiked with a known concentration of the compounds being analysed. This allows confirmation of retention time and the reliability of quantitation in samples throughout an analytical run. This approach is not popular in metabolomics since untargeted approaches are usually used and involves hundreds to thousands of unknown compounds at unknown concentrations (Godzien *et al.*, 2015).

2.2.4.2 Internal standards

Another form of quality control that can be implemented is the use of an internal standard (IS). This allows monitoring of the sample preparation procedure as well as instrumentation functionality. ISs can be used to minimise individual variance between sample preparation and

matrix effect (Griffiths *et al.*, 2010). Ideally an internal standard should be used for each metabolite of interest, but it is experimentally difficult to apply an internal standard for each metabolite in a global metabolomics study. A single internal standard can be applied to correct for analytical variation for a group of metabolites that are relatable (e.g. all present in the same class of metabolites or have similar retention time) (Dunn *et al.*, 2012).

2.2.4.3 Analytical sequence

The order in which samples are run can have an influence on the results that is produced. Due to instrumentation and chromatography drifts that can occur, false variations among sample groups can be induced. Randomisation of the run order of the samples should ensure even distribution of the different experimental groups throughout the analytical run; this will prevent any biased variation (Dunn *et al.*, 2012; Godzien *et al.*, 2015).

2.2.5 Data analysis

Metabolomics studies usually produce a large amount of data. The aim of data analysis is to reduce the number of variables created by the analytical analysis and to normalise the data to prevent any bias from distorting the data (Boccard *et al.*, 2010). Data analysis can be further divided into data processing and data pre-treatment. The appropriate procedures for data handling depends greatly on the metabolomics approach, analytical platform, hypothesis or biological question, downstream data analysis method and the inherent properties of the data (e.g. dimensionality) (Boccard *et al.*, 2010).

2.2.5.1 Data processing

With a targeted approach the data processing is significantly simplified, since identification of metabolites is not necessary, while identification of metabolites is required in an untargeted approach. The data processing depends on the available software and type of raw data produced. With the start of data processing it is important to ensure that the correct peak for each metabolite is identified and correctly intergraded. After peak identification and integration, data normalisation con be performed by using the respective ISs (Godzien *et al.*, 2015; Walsh *et al.*, 2008).

2.2.5.2 Data pre-treatment

After data processing data pre-treatment is essential before statistical analysis can be performed on the data. There are different data pre-treatment processes available (Boccard *et al.*, 2010; Godzien *et al.*, 2015). Usually a filter is applied to the data and different filtering criteria can be applied. A common filtering approach that is used is the 50% presence criteria of a metabolite in all samples, where metabolites are excluded if it is presents are below 50%. Another filtering approach that is used is a 30% relative standard deviation (RSD) among the

QC samples, The RSD of each metabolite, within the QC samples is determined and if the RSD of the metabolite is above 30% it is excluded from the data set. Missing value (zero value) replacement is performed on metabolites that fit the criteria of 50% presence and is necessary for correct statistical analysis. Different approaches to missing value replacing are available, where replacement can be done by the median or ½ of the minimum of all samples involved (Godzien *et al.*, 2015).

2.2.6 Statistical analysis

The statistical analysis is employed to provide the final assessment and outcome of the study. Different types of analyses are available and the analysis depends on the requirements of the study as well as the generated data. Univariate and multivariate analysis are the most popular statistical analytical procedures. Although metabolomics experiments generate multivariate data, a univariate analysis method can be applied to identify an increase or decrease of a specific metabolite between different groups. Multiple parallel hypotheses are then needed when applying univariate tests to multivariate data (Boccard *et al.*, 2010; Goodacre *et al.*, 2007).

2.2.6.1 Univariate analysis

Popular univariate tests are the Student's *t*-test, z-test, ANOVA (analysis of variance) and the non-parametric equivalents. The predictive power of each variable is assessed by finding the probability of statistical significant differences between the defined groups and a straightforward indicator, p-value, is given. When multiple comparisons are made, false positives are likely to occur. Procedures such as the Bonferroni correction can be introduced to correct such errors (Boccard *et al.*, 2010; Goodacre *et al.*, 2007).

2.2.6.2 Multivariate analysis

Metabolomics studies produce multivariate data. This complicates the interpretation of the data since the data is dependent on the number of samples as well as the number of variables (e.g. metabolites) and this increases the dimensionality of the data. Different approaches in multivariate analysis can be used to reduce the dimensionality and compress the data into an easily understandable outcome (Goodacre *et al.*, 2007).

A popular unsupervised multivariate analysis is the principal component analysis (PCA). An unsupervised analysis attempts to find the natural partitions of patterns to enable an understanding of the relationship between samples as well as identifying the variables that are responsible for these relationships. PCA is an orthogonal transformation of multivariate data and provides a summary of the data as well as identifying outliers (Goodacre *et al.*, 2007; Putri & Fukusaki, 2014).

With multivariate analysis a supervised approach can also be followed and a partial least squares (PLS) analysis is a well-known supervised multivariate analysis. PLS has an intrinsic prediction power since bias is induced by identifying the different groups and this sharpens the partition between groups of observations so that maximum separation among different groups is obtained (Boccard *et al.*, 2010).

A hierarchical cluster analysis (HCA) can also be used for multivariate analysis and is employed to identify the similarities among samples and metabolites on a multidimensional space. This enables a visualisation of the biological features in two-dimensional space, thus simplifying the interpretation of the data by clustering the data (Putri & Fukusaki, 2014). The use of HCA is intended to group observations together that is similar to one another and produces a hierarchical structure that may reveal underlying patterns in the data set. Using HCA visualisation, interpretation of the data is possible as well as identification of the trends within a metabolic pathway, therefore an up or down regulated metabolic pathway can be identified.

2.2.7 Biological relevancy

After statistical relevancy is gained for the metabolites investigated, biomarkers can be identified. Biomarkers represent a characteristic that can be objectively measured and evaluated and serve as an indicator of normal biological processes, pathogenic processes or a pharmacological response to a therapeutic investigation (Ilyin *et al.*, 2004). Biomarkers are identified based on differences observed between experimental groups (healthy *versus* diseased) with statistical verification. These observations have to be correlated with biological relevancy and a metabolic map can be generated based on the interactions between metabolic pathways across a complete system. This allows the identification of up or down regulation of metabolic pathways. Before any identified metabolite can be assigned as a biomarker to a disease state and can be used of the characterisation of the pathogenesis of a disease or have an impact on health risk assessment, the biomarker has to be validated (Bonassi *et al.*, 2001). Biomarker validation can be divided into biomarker development, biomarker characterisation and longitudinal studies.

2.2.7.1 Biomarker development

The first step in biomarker validation is the validation of the analytical technique used for the analysis of the samples. All aspects of the analytical procedure, sample preparation, analytical platform and data handling, have to be validated. This is important to ensure the system is stable and comparable data will be generated without any differences induced by the analytical procedure. Guidelines can be followed with regards to method validation to ensure reliability, reproducibility, sensitivity and specificity is achieved with the developed method (Bonassi *et al.*, 2001; Ilyin *et al.*, 2004). During biomarker development, multiple metabolites can be identified

as significant but validation is required to determine the relevancy of the identified metabolites as well as its consistency.

2.2.7.2 Biomarker characterisation

All aspects of sampling and individual differences among participants have to be taken into account. There are several factors that can influence the metabolic profile of an individual. These factors include environmental impact, health state, lifestyle, diet, genetics, age and medicinal consumption. During biomarker validation all of these factors have to be normalised to ensure no differences are induced to experimental groups by these factors (Bonassi *et al.*, 2001). The identified metabolites have to be validated in terms of relevancy. The metabolites have to be characterised the metabolome of the biological system and the influences of these factors have to be determined.

2.2.7.3 Longitudinal studies

Finally an assessment has to be established between the relationship of the identified biomarkers and the associated disease. This can be done by using epidemiological studies and incorporating the different ‘omics platforms to determine if these observation can be identified within the other ‘omics platforms as well (Bonassi *et al.*, 2001; Ilyin *et al.*, 2004). This type of integrated approach is known as interactomics and can provide a much better comprehensive understanding of a disease state.

2.3 Summary

During this study a targeted metabolic profiling method was developed together with a sample preparation and metabolite extraction procedure to be used for the establishment of a metabolic profile for the bleomycin treated C57BL/6J mouse model, representing IPF. Throughout the method development stage the above mentioned method development steps, guidelines and requirements were taken into account. The following chapter (Chapter 3) describes the entire method development process used in this study, as well as descriptive detail on how each method development step, described above, were considered and how the requirements were met.

CHAPTER 3: METHOD DEVELOPMENT

3.1 Introduction

Metabolomics, the comprehensive study of metabolites, can provide a global view on a biological system (Bino *et al.*, 2004; Kang *et al.*, 2016; Kottmann *et al.*, 2012). During method development the aim is to optimise the parameters of the chosen analytical method to best suit its intended use and provide the optimal selectivity and sensitivity. Therefore method development entails the optimisation of the parameters needed for successful identification and chromatography separation. Not only is it necessary to optimise the analytical aspects of the method, but also the sample preparation and metabolite extraction procedures to ensure the best recovery of the metabolites of interest has been provided. In this chapter the method development process will be discussed for the establishment of a targeted metabolic profiling method to be used for generating a metabolic profile for complex diseases such as IPF. Details will be given with regards to the optimisation of the different parameters of the targeted metabolic profiling method. A quality assessment of the analytical parameters of the method was also performed and the detail with regards to the experiments performed for quality assessment is also discussed in this chapter.

3.2 Materials and instrumentation

3.2.1 Reagents

The following high purity high performance liquid chromatography (HPLC) graded solvents were used during this study; dimethyl sulfoxide (DMSO) (CAS: 67-68-5), acetonitrile (ACN) (CAS: 75-05-8), methanol (MeOH) (CAS: 67-56-1), ethanol (64-17-5), propanol (CAS: 67-36-0), formic acid (CAS: 64-18-6), ammonium hydroxide (CAS: 1336-21-6), ammonium acetate (CAS: 631-61-8) and hexane (CAS: 110-54-3). All solvents were purchased from Sigma-Aldrich Co., Taufkirchen; Germany.

3.2.3 Instrumentation

For this project a targeted approach was chosen, using a liquid chromatography linked tandem mass spectrometry (LC-MS/MS) triple quadrupole system as analytical platform for method development. An LC-MS/MS system provides the advantage that a broad range of metabolites can be measured without prior chemical alteration. An LC-MS/MS triple quadrupole system is also robust, reproducible and provides high sensitivity and selectivity (Lu & Chen, 2017; Vuckovic, 2012; Zhang *et al.*, 2012a).

3.2.3.1 LC-MS/MS

A targeted metabolic profiling method was developed on a Triple Quad™ 6500$^+$ and QTRAP® system from AB Sciex Pte., Ltd., Darmstadt; Germany. Together with the Triple Quad™, an Agilent 1290 Infinity II Multisampler and an Agilent 1290 Infinity II High Speed Pump system (Agilent Technologies Deutschland GmbH & Co., Waldbronn, Germany) was used to perform the analysis. The Triple Quad™ 6500$^+$ was utilised for these analyses due to its capability of switching between positive and negative mode in a single analysis and its capability of performing scheduled multiple reaction monitoring (MRM). Analyst 1.6.3 software was used for operating the LC-MS/MS platform as well as data interpretation.

3.3 Method development process

3.3.1 Metabolite identification

A literature investigation was done to identify metabolites of interest to be included into the method (Armitage & Barbas, 2014; DeNicola *et al.*, 2015; Locasale, 2013; Lu & Chen, 2017). The metabolites of interest are products, intermediates and cofactors present in the central carbon system, glycolysis, TCA cycle, the serine biosynthesis pathway, as well as all amino acids present in human metabolic pathways. One hundred (100) metabolites was selected and included in the development of the targeted metabolic profiling method. The selected metabolites are from various metabolic pathways and can provide significant information with regards to the metabolic state of a biological system. These metabolites were not selected just for metabolic profiling of an IPF model, but also for the generation of metabolic profiles for other disease states. A list of the metabolites can be seen in Table 3.1.

3.3.1.1 Standards

A targeted analysis was performed on 100 metabolites (listed in table 3.1). For establishing identification and determining retention times of each metabolite, standards of these metabolites were purchased and analysed. A stock solution of 2 mM was prepared of each standard. The standards were dissolved in either water or DMSO, depending on the metabolite's solubility and characteristics. The information on the metabolites analysed in this study is summarised in Table 3.1. High purity standard (≥ 95%) for each metabolite was purchased from either Sigma-Aldrich Co., Taufkirchen; Germany, eNovation Chemicals LLC, New Jersey; USA, TCI Deutchland GmbH, Eschborn; Germany or Cayman Chemical, Michigan; USA.

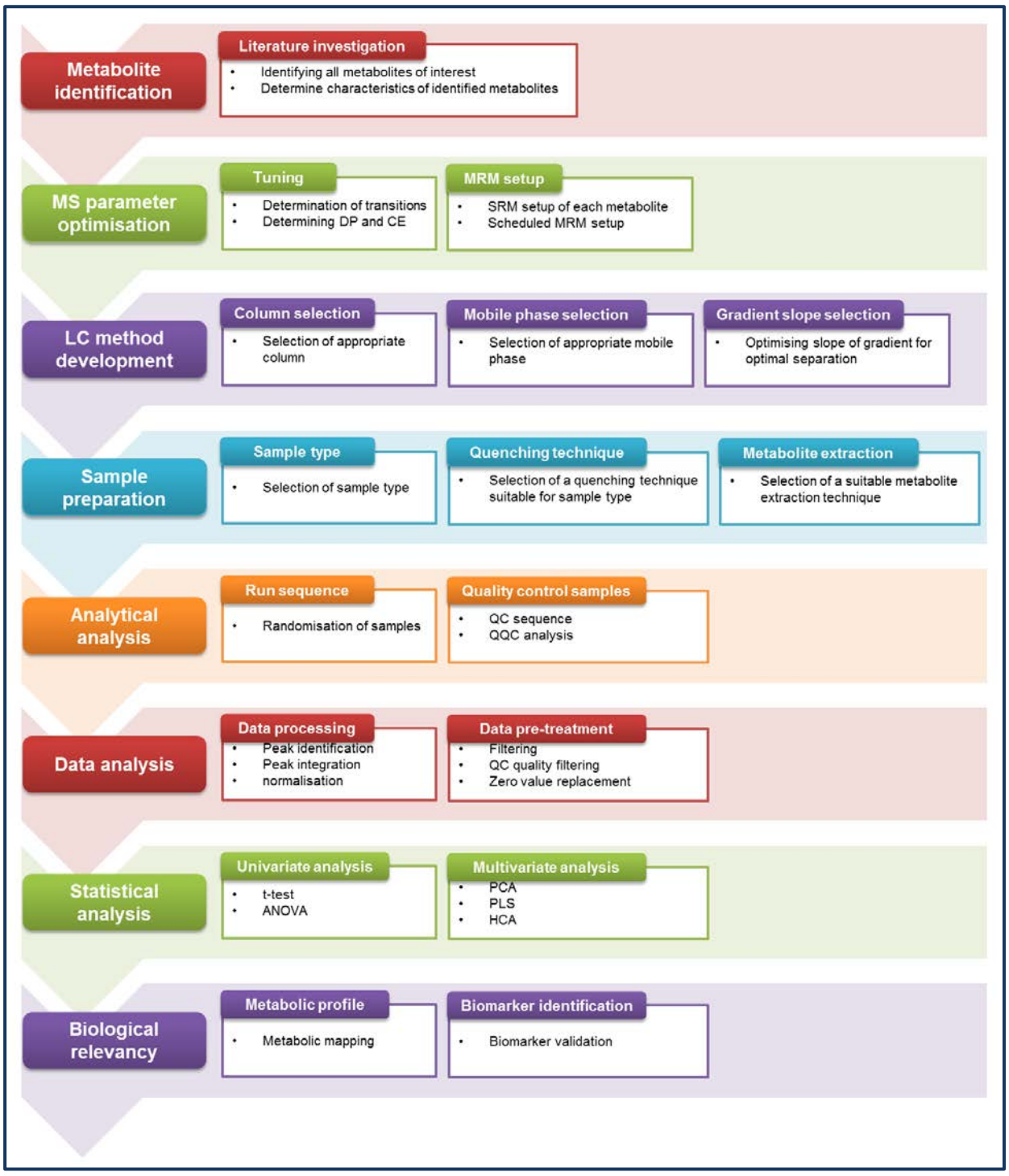

Figure 3.1: Flow chart representation of the method development process. This is an extended view of the flow chart provided in section 2.2 (Figure 2.2).

Table 3.1: Summary of all the metabolites of interest

Group	Metabolite	Formula	M_r	CAS	Company
Glycolysis	Glucose	$C_6H_{12}O_6$	180.063	50-99-7	Sigma
	Fructose-6-Phosphate	$C_6H_{13}O_9P$	260.029	26177-86-6	Sigma
	3-Phospho-D-Glycerate	$C_3H_7O_7P$	185.992	80731-10-8	Sigma
	2-Phosphoglyceric acid	$C_3H_7O_7P$	185.992	83418-48-8	Sigma
	Phosphoenolpyruvic acid	$C_3H_5O_6P$	167.982	4265-07-0	Sigma
	Pyruvate	$C_3H_4O_3$	88.016	113-24-6	Sigma
TCA	Oxaloacetic acid	$C_4H_4O_5$	132.006	328-42-7	Sigma
	Citric acid	$C_6H_8O_7$	192.027	77-92-9	Sigma
	cis-Aconitic acid	$C_6H_6O_6$	174.016	585-84-2	Sigma
	2-Ketoglutaric acid	$C_5H_6O_5$	146.021	328-50-7	Sigma
	Succinic acid	$C_4H_6O_4$	118.026	110-15-6	Sigma
	Fumaric acid	$C_4H_4O_4$	116.011	110-17-8	Sigma
	Malic acid	$C_4H_6O_5$	134.021	6915-15-7	Sigma
	Glyoxylic acid	$C_2H_2O_3$	74.000	563-96-2	Sigma
	Fructose	$C_6H_{12}O_6$	180.063	57-48-7	Sigma
	Lactic acid	$C_3H_6O_3$	90.031	50-21-5	Sigma
Serine bio-synthesis	3-Hydroxybutyrate	$C_4H_8O_3$	104.040	150-83-4	Sigma
	3-Phosphonooxypyruvate	$C_3H_5O_7P$	183.977	3913-50-6	eNC
	L-Serine-O-Phosphate	$C_3H_8NO_6P$	185.008	407-41-0	Sigma
	Serine	$C_3H_7NO_3$	105.042	56-45-1	Sigma
	Glycine	$C_2H_5NO_2$	75.032	56-40-6	Sigma
Amino acids	Acetylalanine	$C_5H_9NO_3$	131.058	97-69-8	Sigma
	Alanine	$C_3H_7NO_2$	89.0477	56-41-7	Sigma
	Arginine	$C_6H_{14}N_4O_2$	174.111	74-79-3	Sigma
	Asparagine	$C_4H_8N_2O_3$	132.053	70-47-3	Sigma
	Aspartic acid	$C_4H_7NO_4$	133.037	56-84-8	Sigma
	Citrulline	$C_6H_{13}N_3O_3$	175.095	372-75-8	Sigma
	Cysteine	$C_3H_7NO_2S$	121.019	52-90-4	Sigma
	Cystine	$C_6H_{12}N_2O_4S_2$	240.020	56-89-3	Sigma
	Glutamic acid	$C_5H_9NO_4$	147.053	56-86-0	Sigma
	Glutamine	$C_5H_{10}N_2O_3$	146.069	56-85-9	Sigma
	Glutathione Oxidized	$C_{20}H_{32}N_6O_{12}S_2$	612.152	27025-41-8	Sigma
	Glutathione Reduced	$C_{10}H_{17}N_3O_6S$	307.083	70-18-8	Sigma
	Histidine	$C_6H_9N_3O_2$	155.069	71-00-1	Sigma
	Hydroxyproline	$C_5H_9NO_3$	131.058	51-35-4	Sigma
	Isoleucine	$C_6H_{13}NO_2$	131.094	73-32-5	Sigma
	Leucine	$C_6H_{13}NO_2$	131.094	61-90-5	Sigma
	Lysine	$C_6H_{14}N_2O_2$	146.105	657-27-2	Sigma
	Phenylalanine	$C_9H_{11}NO_2$	165.079	63-91-2	Sigma
	Proline	$C_5H_9NO_2$	115.063	147-85-3	Sigma
	Taurine	$C_2H_7NO_3S$	125.014	107-35-7	Sigma
	Threonine	$C_4H_9NO_3$	119.058	72-19-5	Sigma
	Tryptophan	$C_{11}H_{12}N_2O_2$	204.089	73-22-3	Sigma
	Tyrosine	$C_9H_{11}NO_3$	181.073	60-18-4	Sigma
	Valine	$C_5H_{11}NO_2$	117.079	72-18-4	Sigma
Methionine cycle	Methionine	$C_5H_{11}NO_2S$	149.051	63-68-3	Sigma
	SAM	$C_{15}H_{23}N_6O_5S$	399.145	86867-01-8	Cayman
	SAH	$C_{14}H_{20}N_6O_5S$	384.120	979-92-0	Sigma
	Homocysteine	$C_4H_9NO_2S$	135.035	454-29-5	Sigma
	Creatine	$C_4H_9N_3O_2$	131.069	6020-87-7	TCI
	Creatinine	$C_4H_7N_3O$	113.059	60-27-5	TCI
	Folic acid	$C_{19}H_{19}N_7O_6$	441.139	59-30-3	Sigma
	Tetrahydrofolate	$C_{19}H_{23}N_7O_6$	445.430	135-16-0	Sigma

Table 3.1 (continued): Summary of all the metabolites of interest

Group	Metabolite	Formula	M_r	CAS	Company
Coenzyme	Acetyl-CoA	$C_{23}H_{38}N_7O_{17}P_3S$	809.120	102029-73-2	Sigma
	CoA	$C_{21}H_{36}N_7O_{16}P_3S$	767.115	85-61-0	Sigma
	FAD	$C_{27}H_{33}N_9O_{15}P_2$	785.157	84366-81-4	Sigma
	NAD	$C_{21}H_{27}N_7O_{14}P_2$	663.109	53-84-9	Sigma
	NADP	$C_{21}H_{28}N_7O_{17}P_3$	743.075	24292-60-2	Sigma
	NADH	$C_{21}H_{29}N_7O_{14}P_2$	665.441	606-68-8	Sigma
	NADPH	$C_{21}H_{30}N_7O_{17}P_3$	745.090	2646-71-1	Sigma
Nucleobase	Adenine	$C_5H_5N_5$	135.054	73-24-5	Sigma
	Guanine	$C_5H_5N_5O$	151.049	73-40-5	Sigma
	Hypoxanthine	$C_5H_4N_4O$	136.039	68-94-0	Sigma
	Thymine	$C_5H_6N_2O_2$	126.042	65-71-4	Sigma
	Uracil	$C_4H_4N_2O_2$	112.027	66-2-8	Sigma
	Cytosine	$C_4H_5N_3O$	111.000	71-30-7	Sigma
	Uridine	$C_9H_{12}N_2O_6$	244.069	58-96-8	Sigma
	Xanthine	$C_5H_4N_4O_2$	152.033	69-89-6	Sigma
Nucleoside	Adenosine	$C_{10}H_{13}N_5O_4$	267.097	58-61-7	Sigma
	Guanosine	$C_{10}H_{13}N_5O_5$	283.091	118-00-3	Sigma
	Inosine	$C_{10}H_{12}N_4O_5$	268.080	58-63-9	Sigma
Nucleotide	ADP	$C_{10}H_{15}N_5O_{10}P_2$	427.029	20398-34-9	Sigma
	GDP	$C_{10}H_{15}N_5O_{11}P_2$	443.024	43139-22-6	Sigma
	UDP	$C_9H_{14}N_2O_{12}P_2$	404.002	27821-45-0	Sigma
Nucleotide	ADP-Glucose	$C_{16}H_{25}N_5O_{15}P_2$	589.082	102129-65-7	Sigma
derivative	ADP-Ribose	$C_{15}H_{23}N_5O_{14}P_2$	559.071	68414-18-6	Sigma
Nucleotide	ATP	$C_{10}H_{16}N_5O_{13}P_3$	506.995	34369-07-8	Sigma
	GTP	$C_{10}H_{16}N_5O_{14}P_3$	522.990	36051-31-7	Sigma
	AMP	$C_{10}H_{14}N_5O_7P$	347.063	4578-31-8	Sigma
	CMP	$C_9H_{14}N_3O_8P$	323.051	6757-06-8	Sigma
	GMP	$C_{10}H_{14}N_5O_8P$	363.050	5550-12-9	Sigma
	IMP	$C_{10}H_{13}N_4O_8P$	348.047	352195-40-5	Sigma
	UMP	$C_9H_{13}N_2O_9P$	324.035	58-97-9	Sigma
	cAMP	$C_{10}H_{12}N_5O_6P$	329.052	60-92-4	Sigma
	cGMP	$C_{10}H_{12}N_5O_7P$	345.047	61093-23-0	Sigma
	dAMP	$C_{10}H_{14}N_5O_6P$	331.068	653-63-4	Sigma
	dCMP	$C_9H_{14}N_3O_7P$	307.056	1032-65-1	Sigma
	dTMP	$C_{10}H_{15}N_2O_8P$	322.056	3343-62-5	Sigma
Other	Carnitine	$C_7H_{15}NO_3$	161.105	6645-46-1	Sigma
	Orotic acid	$C_5H_4N_2O_4$	156.017	65-86-1	Sigma
	Salicylic acid	$C_7H_6O_3$	138.032	69-72-7	Sigma
	Ascorbic acid	$C_6H_8O_6$	176.032	50-81-7	Sigma
	Ketoisovalerate	$C_5H_8O_3$	116.047	3715-29-5	Sigma
	Kynurenine	$C_{10}H_{12}N_2O_3$	208.213	13441-51-5	Sigma
	3-OH-Kynurenine	$C_{10}H_{12}N_2O_4$	224.079	2147-61-7	Sigma
	Quinolinic acid	$C_7H_5NO_4$	167.021	89-00-9	Sigma
	Kynurenic acid	$C_{10}H_7NO_3$	189.042	492-27-3	Sigma
	Pantothenic acid	$C_9H_{17}NO_5$	219.110	137-08-6	Sigma
	Maleic acid	$C_4H_4O_4$	116.010	110-16-7	Sigma
	Ketoleucine	$C_6H_{10}O_3$	130.060	816-66-0	Sigma

3.3.1.2 Internal standards

Internal standards (IS) were acquired for quality control purposes. Information about the internal standards are listed in Table 3.2. These ISs were selected based on their retention time on the column and chemical characteristics. The ISs are well distributed within the elution profile. A stock solution was prepared of each IS, with a concentration of 2 mM. The ISs were dissolved in either water or DMSO, depending on the compound's solubility and characteristics. The ISs were purchase from ether Sigma-Aldrich Co., Taufkirchen; Germany, Omicrom Biochemicals, Inc. Indiana, USA or Buchem Chemie & Technik GmbH & Co. Köln, Germany.

Table 3.2: Summary of all the IS used in this study

Group	Metabolite	M_r	CAS	Company	Solvent
Glycolysis	D-Glucose-13C6	186.11	110187-42-3	Sigma	Water
	D-Fructose-13C6	186.11	201595-65-5	Sigma	Water
Amino acids	L-Serine-13C3,15N,2,3,3-d3	112.08		Sigma	Water
	L-Glutamic acid-13C5	152.09		Sigma	Water
	Glycine-C13	76.06		Sigma	Water
	L-Isoleucine-13C6,15N	138.12		Sigma	Water
	L-Leucine-5,5,5-d3	134.19	87828-86-2	Sigma	Water
	L-Lysine-4,4,5,5-d4	186.67		Sigma	Water
	L-Phenyl-d5-alanine	170.22		Sigma	Water
	Thymine-d4 (methyl-d3,6-d1)	130.14	156054-85-2	Sigma	Water
	Tryptophan D5	209.26		Sigma	DMSO
Coenzyme	Acetyl-1,2-13C2 coenzyme A	811.56		Sigma	Water
Nucleobase	2-(Methyl-13C,d3-thio)adenine	185.23	1216721-76-4	Sigma	DMSO
Nucleoside	[1',2',3',4',5'-13C5]adenosine	272.21	159496-13-6	Omicron	Water
Nucleotide	Adenosine-15N5 5′-monophosphate	352.19		Sigma	Water
	Adenosine-15N5 5′-triphosphate	512.15		Sigma	Water
Other	Quinolinic acid-4,5,6-D3	170.14	138946-42-6	Buchem	DMSO
	D4-Kynurenine	212.24	194546-33-3	Buchem	DMSO
	Kynurenic acid-3,5,6,7,8-d5	194.20		Sigma	DMSO

3.3.2 MS parameter optimisation

To be able to identify and differentiate between metabolites, a unique transition for each metabolite is required. This is done by identifying the precursor ion and product ion of each metabolite. The precursor ion is analysed by the first analyser (MS1) and the product ion, which is a fragment of the precursor ion, is analysed by the second analyser (MS2) (Griffiths *et al.*, 2010). Parameters such as the declustering potential (DP), collision energy (CE), entrance potential (EP) and collision cell exit potential (CXP) can be optimised for each metabolite to ensure high sensitivity.

The identification of the precursor ion and product ion together with the optimised parameters for each metabolite was done by tuning (please see 3.3.2.1, below). Other parameters such as the source temperature (TEM), ionspray voltage (ISV), curtain gas (CUR), nebulizing gas (GS1) and drying gas (GS2) was used as recommended by the operator guide (see Table 3.3) (SCIEX, 2014).

Table 3.3: MS parameters

Parameter	Value
TEM	500 °C
ISV	4500 V
CUR	40 psi
GS1	50 psi
GS2	50 psi

3.3.2.1 Tuning

Tuning is done for the determination of ionisation mode, positive or negative, as well as the identification of the precursor and product ion of each metabolite with its corresponding DP and CE value. Each metabolite is tuned separately to ensure the correct product ion is identified. Different tuning approaches are available. With the use of the Analyst 1.6.3 software (AB Sciex Pte., Ltd., Darmstadt; Germany) it is possible to perform a compound optimisation tuning or manual tuning. The compound optimisation approach is very convenient since an automated tuning is done by identifying the most intense product ion of the specified precursor ion together with the optimal DP, CE and CXP values.

An aliquot of the metabolite stock solution (2 mM) was diluted with an 80% acetonitrile + 0.1% formic acid solution to a detectable concentration. This dilution was then used with a direct infusion to perform the tuning. For metabolites with high sensitivity the compound optimisation approach was used and for metabolites with low sensitivity manual tuning was performed. The ionisation mode, precursor and product ion as well as the optimised parameters for each metabolite can be seen in Chapter 4, Table 4.3.

3.3.2.2 MRM setup

After the identification of each metabolite's precursor and product ion a selected reaction monitoring (SRM) profile can be created for each metabolite with the optimised MS parameters. This allows the highest selectivity for each metabolite. An MRM scan can then be generated by combining all the SRM profiles of the metabolites (Griffiths *et al.*, 2010). During this scan the analyser jumps from one profile to the other. A scheduled MRM can also be generated were the scan, for a specific metabolite or group of metabolites, is limited to an approximated retention time. This increases selectivity by reducing unnecessary dwell time (Griffiths *et al.*, 2010).

An SRM profile was generated for each metabolite by the data generated from tuning. A scheduled MRM was generated to be able to include all metabolites' SRM profiles in one run. A positive-negative switching was also included since many metabolites were only detectable in a negative ionisation mode. The retention times of the metabolites can be seen in Table 4.3 and is based on the optimised HILIC method described in Chapter 4.

3.3.3 LC method development

The LC method is a core part that determines the success of the method. With LC method development there are several factors that have to be optimised to achieve the best separation of the metabolites. Factors include selection of a column, mobile phase, buffer, pH, gradient slope, flowrate and injection volume.

3.3.3.1 Column selection

The selection of a column and mobile phase depend upon each other as well as the characteristics of the metabolites of interest. Knowledge of the characteristics of all metabolites of interest is vital prior to the selection of the column. Different columns are available to provide separation of different groups of metabolites. The selection of a column is usually based on the polarity of the metabolites. Hydrophobic metabolites are usually separated on an RP column and there are different RP columns available containing different stationary phase packing material (C_1, C_3, C_4, C_8, C_{18}, phenyl, CN, TFA, etc.) (Sandi *et al.*, 1997). Hydrophilic metabolites are usually separated on a NP column and a HILIC column is a popular column used for such metabolites.

The metabolites of interest in this study have a broad range with regards to polarity. Since the aim of the study was to develop a method that would allow the analysis of all the metabolites of interest in one analysis, different columns (Phenomenex Kinetex C_{18}, Atlantis T3, ProteCol C_{18} and Luna NH_2 (HILIC)) were tested to find an appropriate column to separate all the metabolites of interest.

Reversed phased columns

The first column tested was a RP C_{18} column (Phenomenex Kinetex C_{18}, 2.1 mm x 30 mm, 5 micron particle size). A short summary of the method used is as follow; a 4 min method using 0.1% formic acid in water (solvent A) and 0.1% formic acid in ACN (solvent B) was used with a flow rate of 400 µL/min. The gradient was as follows: starting with a 95% solvent A, hold for 0.1 min followed by a linear gradient of 95% - 5% solvent A for 2.5 min, followed by an isocratic gradient of 5% solvent A for 0.7 min before returning to starting conditions at 4 min. The polar metabolites were tested first to establish whether or not this column would be able to retain and separate these polar compounds. 3-Phospho-D-glycerate was not retained on the column, therefore reliable detection was not possible (see Figure 3.2).

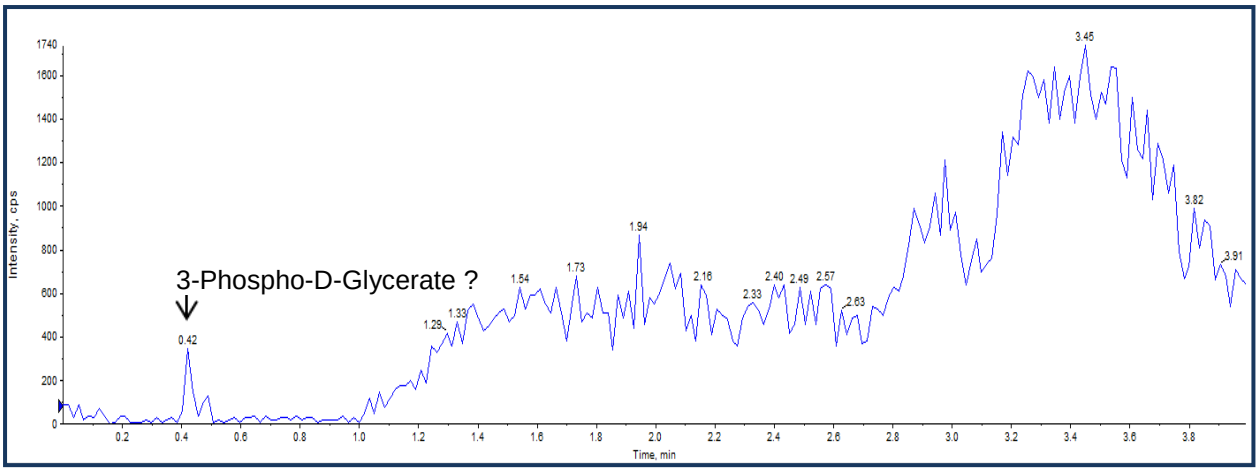

Figure 3.2: Phenomenex Kinetex C_{18} chromatography of 3-Phospho-D-glycerate. A Phenomenex Kinetex C_{18}, 2.1 mm x 30 mm, 5 micron particle size column was used with a 4 min method. Due to the high hydrophilicity of 3-phospho-D-glycerate retention on the RP column was not possible. The figure was generated using the Analyst 1.6.3 software.

An Atlantis T3 column containing trifunctional C_{18} alkyl stationary phase (2.1 mm x 100 mm, 5 micron particle size) was tested using 98% aqueous phase as starting condition. A short summary of the method used is as follow: a 10 min method using 0.1% formic acid in water (solvent A) and 0.1% formic acid in ACN (solvent B) with a flow rate of 400 µL/min. The gradient was as follows: starting with a 98% solvent A, hold for 0.5 min followed by a linear gradient of 98% - 5% solvent A for 5 min, followed by an isocratic gradient of 5% solvent A for 1 min before returning to starting conditions at 10 min. Although retention of 2-phosphoglyceric acid was possible with this column, the retention is not optimal since the elution time is very early and peak shape does not show a Gaussian distribution. Obtaining any retention of 3-phospho-D-glycerate was still unsuccessful (see Figure 3.3).

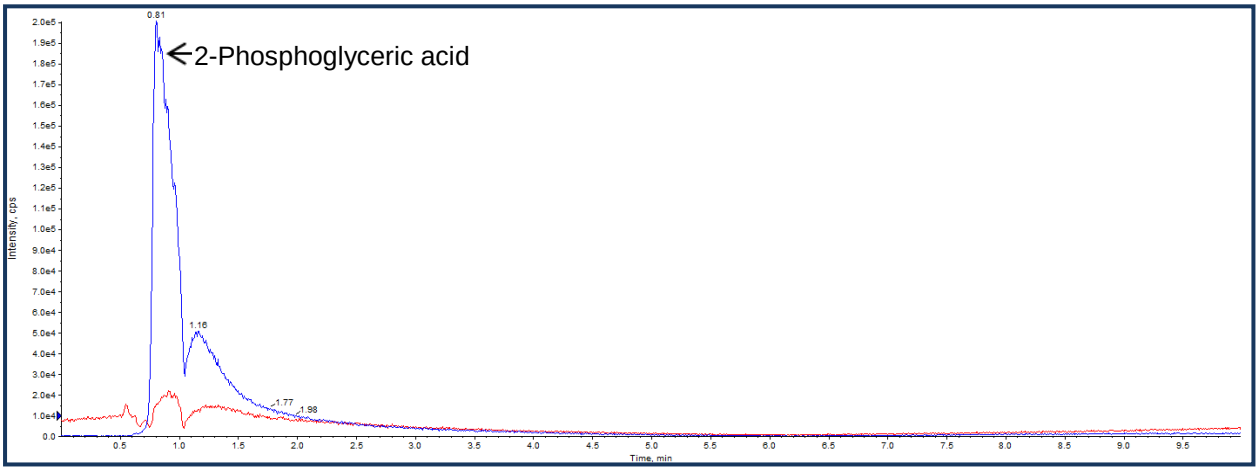

Figure 3.3: Atlantis T3 C_{18} chromatography of 2-Phosphoglyceric and 3-Phospho-D-Glycerate. An Atlantis T3 column with trifunctional C_{18} alkyl stationary phase 2.1 mm x 100 mm, 5 micron particle size was used with a 10 min method. 2-Phosphoglyceric acid (blue chromatogram) was able to be retained with poor peak shape and 3-Phospho-D-Glycerate (red chromatogram) was not able to be retained with this column. The figure was generated using the Analyst 1.6.3 software.

A ProteCol C_{18} Q103 (2.1 mm x 150 mm, 3 µm, 100 Å) was tested with an ion-paring reagent, tributylamine (TBA), as described by Luo *et al.* (2007). 3-Phospho-D-glycerate as well as another highly polar metabolite, L-serine-O-phosphate, were retained on this column (see Figure 3.4). Prolonged use of an ion-paring reagent has several drawbacks including the limitation of column use once an ion-paring reagent was used (Watson, 2017). A decline in detector sensitivity was seen with prolonged used of an ion-paring reagent, therefore alternative methods were investigated.

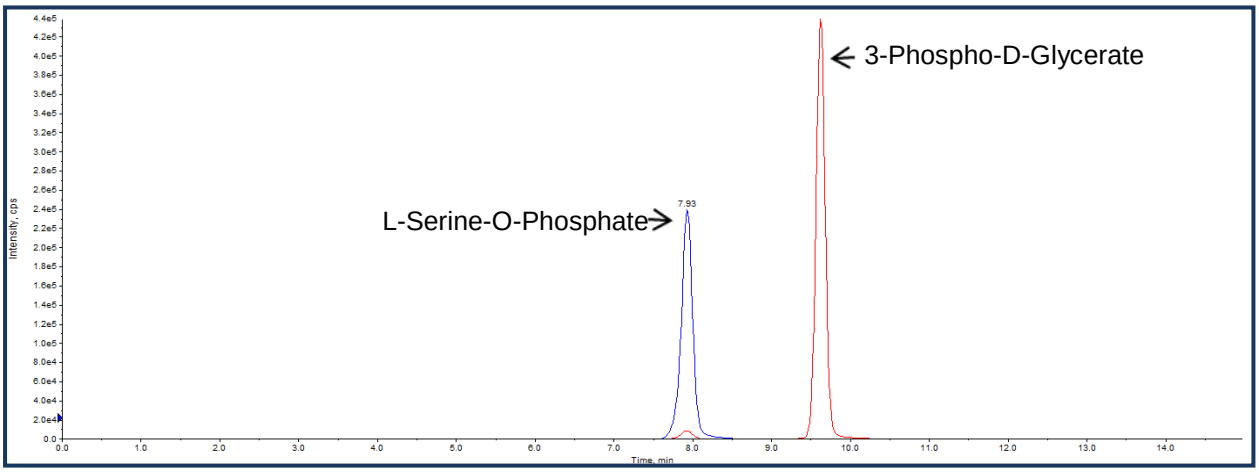

Figure 3.4: ProteCol C18 Q103 chromatography of 3-Phospho-D-Glycerate and L-Serine-O-Phosphate. A ProteCol C_{18} Q103 column (2.1 mm x 150 mm, 3 µm, 100 Å) was used with a 15 min method. 3-Phospho-D-glycerate (red chromatogram) and L-serine-O-phosphate (blue chromatogram) were retained by the column using the ion-paring reagent. The figure was generated using the Analyst 1.6.3 software.

Normal phased columns

A Luna NH_2 (HILIC) column (2 mm x 150 mm, 5 µm, 100 Å) was tested and retention of the polar metabolites was archived (see Figure 3.5). The method described by Bajad *et al.* (2006), with solvent A: 20 mM ammonium acetate + 20 mM ammonium hydroxide in 95:5 H_2O: ACN, pH 9 and solvent B: ACN was tested and adjusted for optimisation. Different lengths of the Luna NH_2 (HILIC) column were tested, 150 mm and 250 mm in length, to find the optimal separation of the metabolites with the shortest run time. The Luna NH_2 150 mm (HILIC) column provided the best separation at a run time of 20 min.

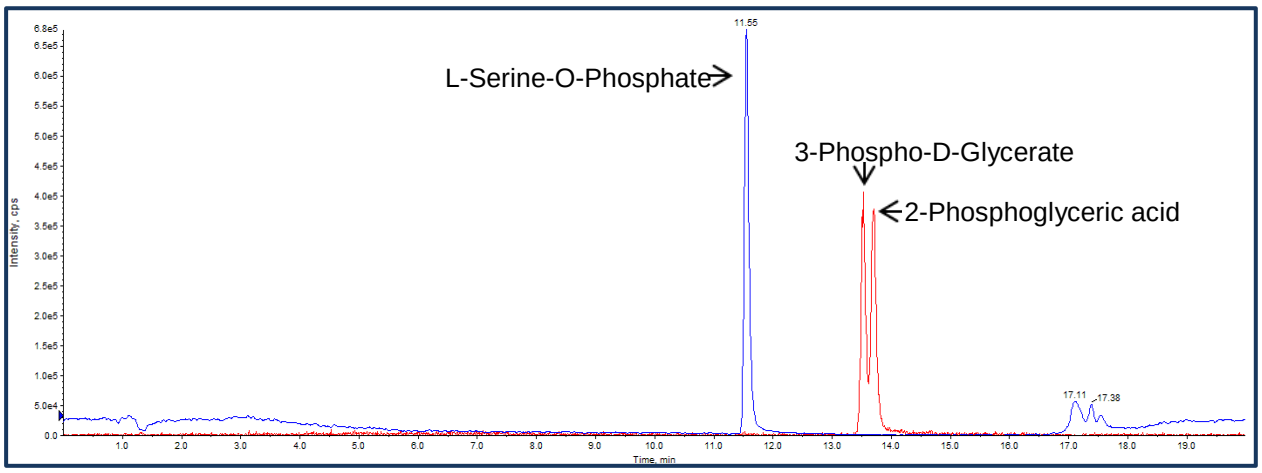

Figure 3.5: Luna NH_2 (HILIC) chromatography of L-Serine-O-Phosphate, 3-Phospho-D-Glycerate and 2-Phosphoglyceric acid. A Luna NH_2 column, 2 mm x 150 mm, 5 µm, 100 Å was used with a 20 min method. L-Serine-O-Phosphate (blue chromatogram) was able to be retained by the column with a good peak shape. 3-Phospho-D-glycerate (red chromatogram) and 2-phosphoglyceric acid (red chromatogram) have the same transitions, but separation is still possible with this column. The figure was generated using the Analyst 1.6.3 software.

3.3.3.2 Mobile phase selection

The selection of the mobile phase is determined by the column since stability of the column determines the reproducibility of the method (Kirkwood *et al.*, 2013; Wu, 2009). The pH of the mobile phase is an important factor determining the stability of the column, it is important to know the optimal working conditions and limitations of the column (Wu, 2009). The Luna NH_2 (HILIC) column has a wide range of pH stability (pH 1.5 till pH 11) (Phenomenex, 2015). The basic conditions described by Bajad *et al.* (2006) gave the desired retention and ionisation of the metabolites of interest, and was used in the final method.

3.3.3.3 Gradient slope selection

The slope of the gradient is important with regards to separation. Many of metabolites have isobaric masses and it is crucial to separate these metabolites for accurate identification. Some metabolite's product ion is the precursor ion of a different metabolite and separation of these metabolites are also crucial for accurate identification. Not only is separation necessary but the peak shape as well, a Gaussian distribution peak shape is ideal and the gradient and flow rate plays an important role in achieving this.

Metabolites of interest in this study that is isobaric in mass is; glucose and fructose, 3-Phospho-D-Glycerate and 2-Phosphoglyceric acid, isoleucine and leucine and fumaric acid maleic acid. All these metabolites could be separated except fumaric acid and maleic acid (see Figure 3.6).

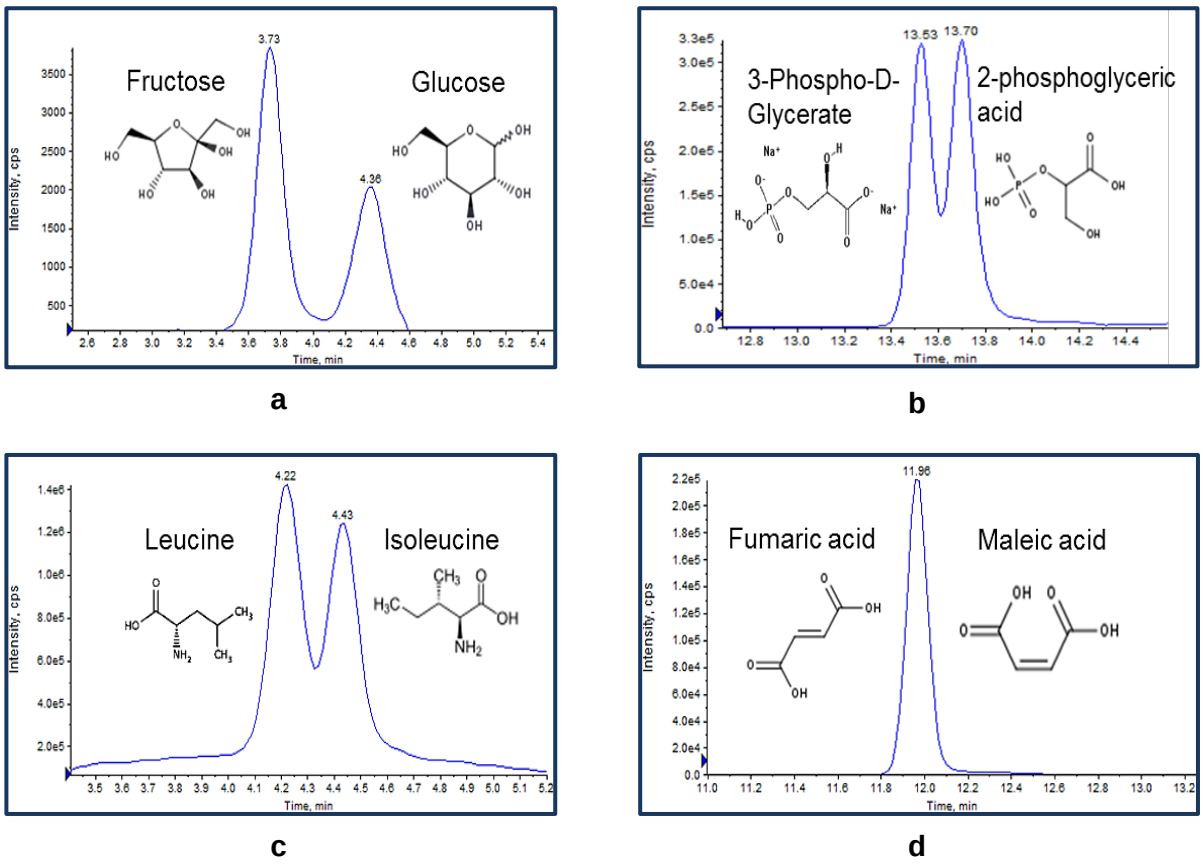

Figure 3.6: Chromatographic separation of a) Glucose and Fructose, b) 3-Phospho-D-Glycerate and 2-phosphoglyceric acid, c) Isoleucine and Leucine and d) Fumaric acid and Maleic. In graph a, b and c sufficient separation were achieved. In figure d no resolution was achieved for the separation of fumaric acid and maleic acid. These graphs were generated using the Analyst 1.6.3 software.

Metabolites such as adenosine, ADP, ADP-Glucose, ADP-Ribose, ATP, AMP, cAMP and dAMP share the same product ion. This product ion is the precursor ion of adenine and therefore all these metabolites had to be separated from adenine (see Figure 3.7). During method development these analytes were able to be separated chromatographically (see Table 4.3 for retention time).

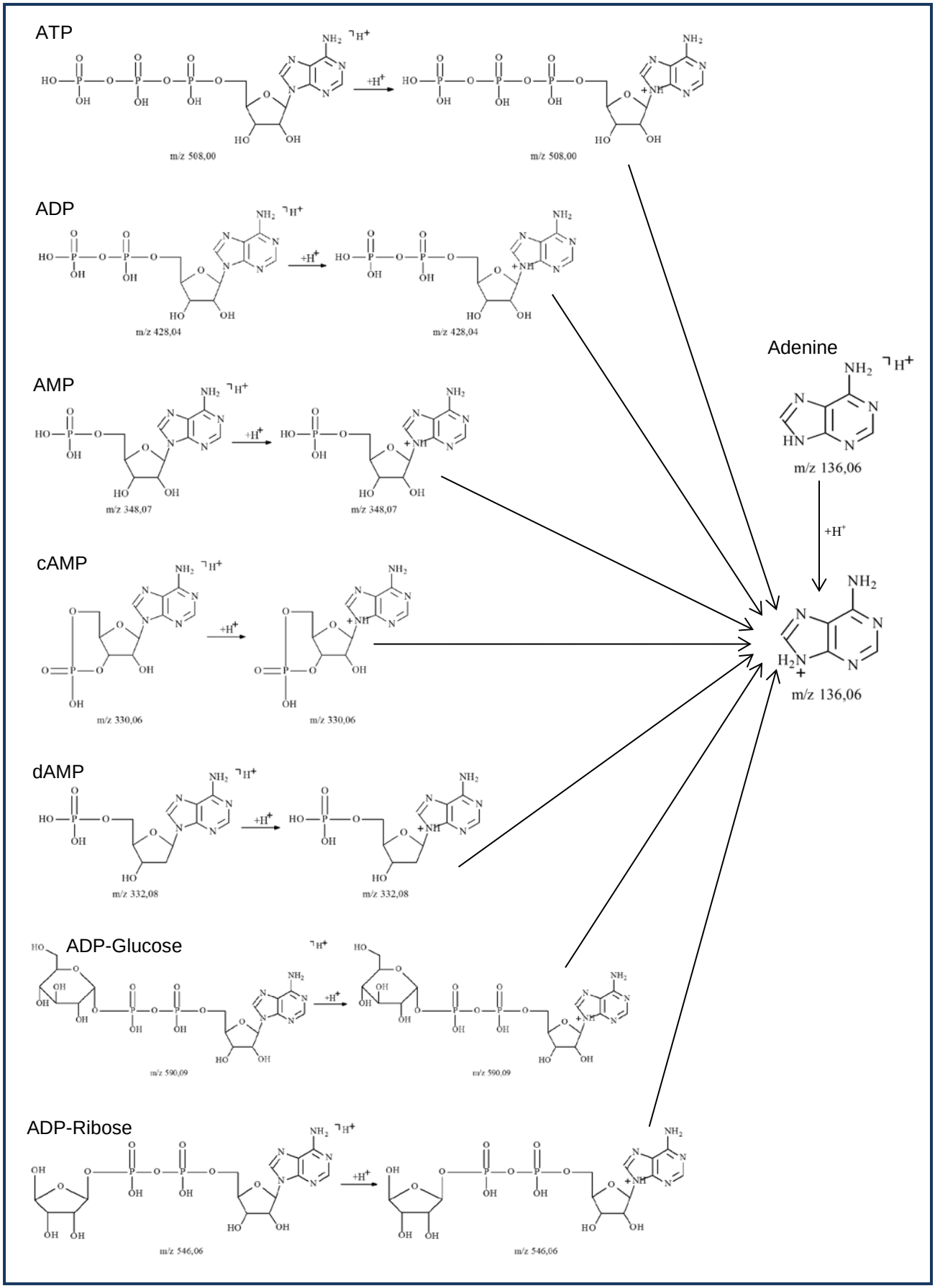

Figure 3.7: MS/MS Fragmentation patterns of ATP, ADP, AMP, cAMP, dAMP, ADP-Glucose, ADP-Ribose. All the metabolites shared the same product ion the adenine precursor ion. This figure was generated by using the MarvinSketch 14.12.15.0 software.

3.3.4 Sample preparation

Sample preparation is dependent on the approach and analytical platform and it is vital that the metabolic state of the biological system stay intact throughout the entire procedure. For a targeted metabolic profiling method, the sample preparation should be non-selective, simple and fast to ensure no metabolites of interest will be lost and/or degraded. It should also be reproducible and contain a quenching step to ensure the results represent the true metabolic state of the sample at the time of sampling (Vuckovic, 2012).

For the optimisation of sample preparation for the developed method, different sample preparation approaches with regards to metabolite extraction was investigated. A literature investigation was done to determine which sample preparation approaches will accommodate the developed method the best (Dettmer *et al.*, 2007; Gika & Theodoridis, 2011; Henion *et al.*, 1998; Vuckovic, 2012). Since the metabolites of interest have a broad range with regards to polarity, the best approach for metabolite extraction was protein precipitation. An acetonitrile and methanol mixture (1:1) was used as organic solvent in a 1:4 (*v/v* or *w/v*) ratio.

3.3.5 Analytical analysis

There are several factors that can influence the results produced during the analytical analysis of the samples. Some of these factors can introduce bias into the study, which will result in false differences recorded among experimental groups (Dunn *et al.*, 2012). Instrumentation and chromatographical drift are among these factors and are the main causes of bias inducement. There are several procedures that can be followed to ensure the quality of the analysis and system is accurate (Dunn *et al.*, 2012). Randomisation of the order of sample preparation and analysis is recommended: this ensures that no bias is introduced by the instrumentation and analysis. Another procedure that can be performed to evaluate the quality of the analysis is the use of quality control (QC) samples and spiked samples, where a QC sample is spiked with known concentration of the metabolites.

3.3.5.1 Run sequence

During this study, randomisation of the run order of the samples was done for all sample analysis. A Microsoft Excel randomisation equation was used to re-order the run sequence of the samples. QC samples was prepared and included in the run, at the beginning, middle and end of the analytical run. When sample size was large a QC sample was run after every 5 samples. Five (5) QC samples were also run at the beginning of the analysis, before the samples were analysed, to ensure the system is calibrated. A QQC (spiked QC) sample was also prepared and run prior to the start of sample analysis, ensuring accurate peak identification.

3.3.5.2 Quality control samples

During this study, QC samples were used for evaluation of the method, instrumentation and system. The following QC preparation procedure was used: QC samples were generated by pooling equal aliquots of all samples to be analysed, after which, several QC samples were prepared using the same sample preparation protocol for all samples. The multiple QC preparation provides the opportunity for the sample preparation procedure to be evaluated as well.

A QQC sample was generated by spiking a known concentration of all metabolites of interest into a QC sample. Since the metabolites of interest are within a low mass range and impurities can have similar transitions, the correct peak identification is crucial. The QQC sample can be used to verify peak identification by evaluating the retention time of the metabolites within the QQC sample.

3.3.6 Data handling

Metabolomics studies usually produce a large amount of data. Care should be taken when handling such complex data sets, since the analysis and handling of the data can impact the quality of the identification and quantification of the results. Data handling can be further divided into data processing and data pre-treatment (Boccard *et al.*, 2010). Since a targeted LC-MS/MS analytical platform was used during this study the data handling is simplified.

3.3.6.1 Data processing

Data processing was done using Analyst 1.6.3. A quantitation method was designed for peak identification, peak integration and IS normalisation. With low abundant metabolites care was taken with peak integration to obtain an accurate peak area. Normalisation is an important factor in data processing and internal standards were used for this purpose. The number of metabolites analysed in this method is large and it is not feasible to use an internal standard for each metabolite. Nineteen (19) ISs were chosen for normalisation and the selection of these ISs were based on their retention time. The ISs are well distributed throughout the entire run. Metabolites were grouped according to their retention time and normalised to an IS with a similar retention time (see Table 4.3).

3.3.6.2 Data pre-treatment

After the data has been extracted, data pre-treatment and data clean-up were performed to prepare the data in such a way that statistical analysis can be performed on the data. Data clean-up was performed by excluding all metabolites that are below the detection limit (see table 3.4 for limit of detection (LOD) values) and have less than 50% presence in all samples. The QC samples were then assessed and all metabolites that have a RSD value above 30%

were excluded, for it indicates that the metabolite was unstable during analysis and any differences between the experimental groups cannot be trusted. Furthermore a zero value replacement was performed by replacing zero values with the half of the LOD for the metabolite (Croghan & Egeghy, 2003). The data was then arranged in the format compatible with the Excel macro file that was used for statistical analysis.

3.3.7 Statistical analysis

Various software and online tools are available to perform data analysis. For this study an Excel-based statistical analysis method, which is freely available as a macro file (developed by Dr. Tsugawa, and available at http://prime.psc.riken.jp), was used. With this Excel-based method it is possible to generate statistical relevant data by using the multi t-test, graph analysis, PCA, correlation analysis, partial least square regression (PLS-R) and partial least square discrimination analysis (PLS-DA) (Putri & Fukusaki, 2014).

3.4 Quality assessment

The validation of any bioanalysis method is mandatory to ensure reliable results are provided in routine application. Method validation is responsible for demonstrating that a biological method can accurately quantify analytes within a specific matrix. Although there are no guidelines for validating metabolomics-based studies, the “Guideline on bioanalytical method validation” by the European Medicines Agency (EMA) was used for assessing the quality of all aspects of the method by assessing the parameters suggested by the EMA. These parameters include range and linearity, limits of quantification and detection, accuracy, precision and carry-over (González *et al.*, 2014; Peters *et al.*, 2007; Rozet *et al.*, 2011).

The aim of this study was to establish a standardised LC-MS/MS method for targeted metabolic profiling of biological matrices, to be used for metabolic profiling of complex diseases such as IPF. The developed method will only be used to identify differences in metabolite abundances between experimental groups. Quantification of the identified metabolites is out of the scope for this study, but to ensure the method is reliable, reproducible and robust the parameters above have been tested to evaluate the quality of the method.

3.4.1 Range and linearity

Although a literature investigation was done to define the average concentration of all the metabolites in lung matrix, it was not possible to define the average concentration for each metabolite due to the lack of information available on the subject and not all these metabolites have been quantified in animal lung tissue. Therefore, a preliminary analysis was done prior to method validation to obtain observed concentrations of the metabolites. The metabolites’

abundance differed substantially as well as instrumentation sensitivity towards each metabolite. Due to this substantial difference among the metabolites of interest, a detection range and linear range was determined for each metabolite (Table 3.4). The detection range is defined by the LOD and upper limit of quantification (ULOQ). The linear range is defined by the lower limit of quantification (LLOQ) and ULOQ. This was created by generating an area ratio between the metabolite peak area to the assigned IS area ratio. Linearity of each metabolite was assessed by the correlation coefficient (R^2) obtained by plotting the area ratio of the analyte and IS to the concentration of each calibration point, normalised to the IS concentration. An example of this plotting, the calibration curve of adenosine, can be seen in Figure 3.8 and a summary of the R^2 values for each metabolite can be seen in Table 3.4 .

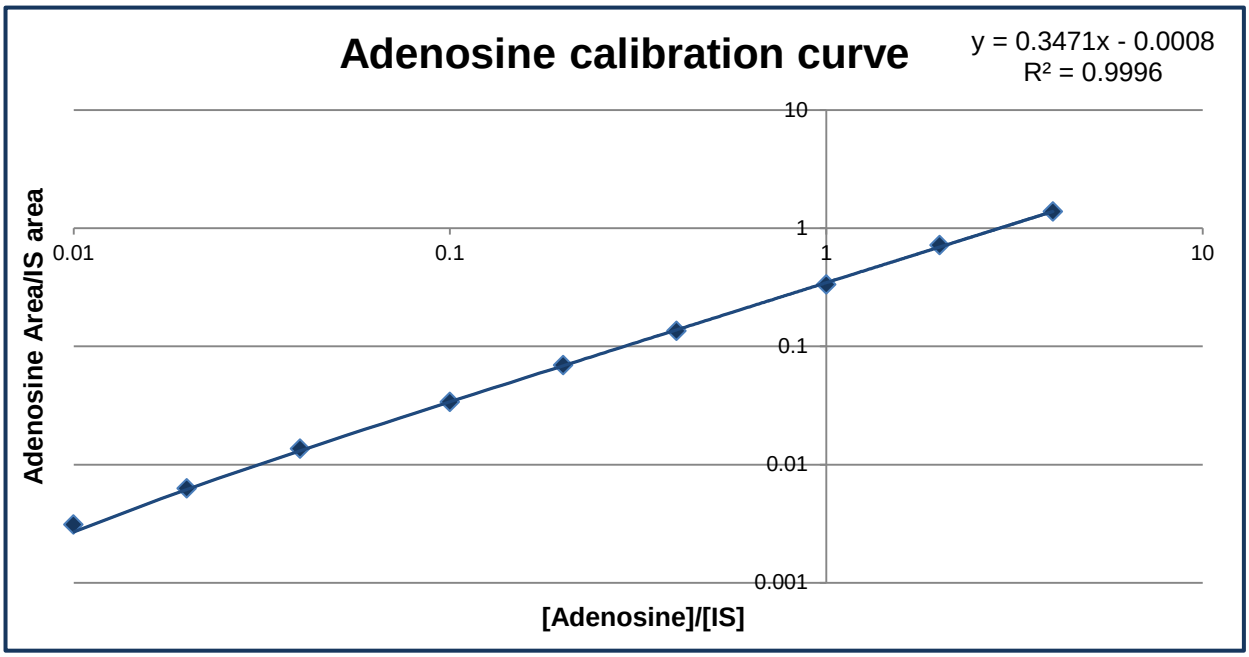

Figure 3.8: A double log plotted calibration curve of adenosine. Nine calibration points were included in the analysis and each concentration was analysed in triplicate with a standard deviation below 10%.

3.4.2 Limits of detection and quantification

The LOD refers to the lowest concentration value of the metabolites that can be detected, not necessarily quantified, with acceptable accuracy and precision. The LLOQ and ULOQ refer to the lowest and highest concentration values of a metabolite that can be used for quantification purposes. There are several approaches that can be followed for the determination of the LOD, LLOQ and ULOQ. These approaches include methodological approach based on standard deviation at the LLOQ, this is in turn based on the signal to noise (S/N) ratio approach, linear regression, standard deviation of the response and the slope, etc. (González *et al.*, 2014; Peters *et al.*, 2007; Rozet *et al.*, 2011).

3.4.2.1 Calculation of LOD, LLOQ, ULOQ

Due to the low molecular masses of the metabolites of interest some impurities may have the same transition as the metabolites. This can increase the signal to noise in the detection of some analytes. To ensure an appropriate LOD and LLOQ is achieved the standard deviation of response and slope approach was used. The formulas used to calculate the LOD and LLOQ were Equation 3.1 and Equation 3.2 respectively, where σ represents the standard deviation of the response and *S* represents the slope of the calibration curve (Rozet *et al.*, 2011). With regards to the LLOQ, the calculated value was visually compared with the calibration curves to ensure the value is within the linear regression of the calibration curve. The ULOQ of each metabolite was determined by the highest concentration value within the linear regression of the calibration range. The final LOD, LLOQ and ULOQ values of all metabolites are summarised in Table 3.4.

Equation 3.1: Limit of detection

$$\mathrm{LOD} = \frac{3\sigma}{S}$$

Where σ represents the standard deviation of the response and *S* represents the slope of the calibration curve.

Equation 3.2: Lower limit of quantification

$$\mathrm{LLOQ} = \frac{10\sigma}{S}$$

Where σ represents the standard deviation of the response and *S* represents the slope of the calibration curve.

Table 3.4: Summary of detection and quantitation limits, linear ranges and corresponding correlation coefficients of metabolites of interest

Group	Metabolite	LOD (ng/mL)	LLOQ (ng/mL)	ULOQ (ng/mL)	Linear Range (ng/mL)	R^2
Glycolysis	Glucose	180	900	3601	900 – 3601	0.993
	Fructose-6-Phosphate	203	406	26003	406– 26003	0.998
	3-Phospho-D-Glycerate	186	372	3720	372 – 3720	0.997
	2-Phosphoglyceric acid	19	37	1860	37 – 1860	0.997
	Phosphoenolpyruvic acid	168	336	8399	336 – 8399	0.988
	Pyruvate	550	1100	8802	1100 – 8802	0.992
TCA	Oxaloacetic acid	6600	13201	26401	13201 – 26401	0.987
	Citric acid	300	601	9601	601 – 9601	0.997
	cis-Aconitic acid	8701	17402	34803	17402 – 34803	0.976
	2-Ketoglutaric acid	913	1825	29204	1825 – 29204	0.993
	Succinic acid	738	1475	2320	1475 – 2320	0.998
	Fumaric acid	232	580	2320	58 – 2320	0.999
	Malic acid	420	838	13402	838 – 13402	0.990
	Glyoxylic acid	463	925	3700	3700 – 925	0.998
	Fructose	180	360	3601	360 – 3601	0.999
	Lactic acid	4502	9003	180062	9003 – 180062	0.993
Serine bio-synthesis	3-Hydroxybutyrate	52	104	2081	104 – 2081	0.998
	3-Phosphonooxypyruvate	4599	9199	36795	9199 – 36795	0.997
	L-Serine-O-Phosphate	93	185	3700	185 – 3700	0.992
	Serine	82	164	10504	164 – 10504	0.998
	Glycine	60	150	7503	150 – 7503	0.993
Amino acids	Acetylalanine	13	26	2621	26 – 261	0.998
	Alanine	1113	2226	8905	2226 – 8905	0.984
	Arginine	136	272	17411	272 – 17411	0.997
	Asparagine	52	103	26411	103 – 26411	0.999
	Aspartic acid	3326	6652	26607	6652 – 26607	0.988
	Citrulline	69	137	17510	137 – 17510	0.997
	Cysteine	379	756	12102	756 – 12102	0.999
	Cystine	469	939	120010	939 – 120010	0.997
	Glutamic acid	58	115	14705	115 – 14705	0.997
	Glutamine	86	171	43821	171 – 43821	0.998
	Glutathione Oxidized	239	1224	61215	1224 – 61215	0.991
	Glutathione Reduced	307	614	3071	614 – 3071	0.998
	Histidine	121	242	15507	242 – 1507	0.993
	Hydroxyproline	204	410	13106	410 – 13106	0.995
	Isoleucine	26	51	13109	51 – 13109	0.999
	Leucine	102	205	13109	205 – 13109	0.999
	Lysine	23	457	14611	457 – 14611	0.998
	Phenylalanine	83	165	3302	165 – 3302	0.998
	Proline	230	575	2301	575 – 2301	0.995
	Taurine	1950	3913	250028	3913 – 250028	0.998
	Threonine	93	186	11906	186 – 11906	0.996
	Tryptophan	41	102	4082	102 – 4082	0.998
	Tyrosine	141	283	18107	283 – 18107	0.998
	Valine	59	117	2342	117 – 2342	0.999
Methionine cycle	Methionine	30	75	2981	75 – 2981	0.999
	SAM	200	399	7983	399 – 7983	0.995
	SAH	77	192	7682	192 – 7682	0.996
	Homocysteine	423	844	27007	844 – 27007	0.999
	Creatine	6.6	13.1	2621.4	13.1 – 2621.4	0.996

Table 3.4 (continued): Summary of detection and quantitation limits, linear ranges and corresponding correlation coefficients of metabolites of interest

Group	Metabolite	LOD (ng/mL)	LLOQ (ng/mL)	ULOQ (ng/mL)	Linear Range (ng/mL)	R^2
Methionine cycle	Creatinine	6	11	2261	11 – 2261	0.998
	Folic acid	22	44	8823	44 – 8823	0.998
	Tetrahydrofolate	5566	11132	44526	11132 – 44526	0.998
Coenzyme	Acetyl-CoA	405	809	16182	809 – 16182	0.997
	CoA	767	1534	15342	1534 – 15342	0.999
	FAD	157	393	15703	393 – 15703	0.991
	NAD	133	332	13262	332 – 13262	0.991
	NADP	149	372	14862	372 – 14862	0.995
	NADH	1331	3327	13309	3327 – 13309	0.998
	NADPH	4657	9314	37255	9314 – 37255	0.995
Nucleobase	Adenine	7	14	2701	14 – 2701	0.999
	Guanine	30	76	3021	76 – 3021	0.998
	Hypoxanthine	27	68	2721	68 – 2721	0.999
	Thymine	6	25	2521	25 – 2521	0.999
	Uracil	6	11	11203	11 – 11203	0.999
	Cytosine	5.6	11	555	11 – 555	0.999
	Uridine	12	24	4881	24 – 4881	0.999
	Xanthine	237	476	15203	476 – 15203	0.995
Nucleoside	Adenosine	1	13	5342	13 – 5342	0.999
	Guanosine	14	28	5662	28 – 5662	0.999
	Inosine	27	105	26808	105 – 26808	0.998
Nucleotide	ADP	427	854	8540	854 – 8541	0.999
	GDP	222	443	8861	443 – 8861	0.994
	UDP	202	404	8080	404 – 8080	0.997
Nucleotide derivative	ADP-Glucose	295	589	11782	589 – 11782	0.989
	ADP-Ribose	295	589	11782	589 – 11782	0.989
Nucleotide	ATP	101	254	10140	254 – 10140	0.986
	GTP	6537	13075	52299	13075 – 52299	0.998
	AMP	347	694	6941	694 – 6941	0.996
	CMP	162	323	6461	323 – 6461	0.998
	GMP	73	182	7261	182 – 7261	0.997
	IMP	348	696	6961	696 – 6961	0.996
	UMP	65	162	6481	162 – 6481	0.999
	cAMP	33	66	32905	66 – 32905	0.999
	cGMP	35	69	6901	69 – 6901	0.999
	dAMP	17	33	3311	33 – 3311	0.999
	dCMP	61	154	6141	154 – 6141	0.992
	dTMP	32	64	6441	64 – 6441	0.996
Other	Carnitine	16	32	16111	32 – 16111	0.998
	Orotic acid	4	8	3120	8 – 3120	0.999
	Salicylic acid	69	138	2761	138 – 2761	0.992
	Ascorbic acid	3521	8802	35206	8802 – 35206	0.978
	Ketoisovalerate	58	116	2321	116 – 2321	0.997
	Kynurenine	104	208	4164	208 – 4164	0.998
	3-OH-Kynurenine	22	45	4482	45 – 4482	0.999
	Quinolinic acid	33	84	3340	84 – 3340	0.999
	Kynurenic acid	95	189	3781	189 – 3781	0.998
	Pantothenic acid	22	44	4382	44 – 4382	0.999
	Maleic acid	232	580	2320	580 – 2320	0.999
	Ketoleucine	2029	4071	32515	4071 – 32515	0.988

3.4.3 Accuracy and Precision

Accuracy and precision assessments ensure that the data generated by the developed method is correct and reproducible. Accuracy describes the closeness of the measured concentration of an analyte within the matrix and the true concentration of the analyte within the matrix. Precision describes the closeness of the multiple measurements of samples that have the same concentration. Inter-day and intra-day accuracy and precision should be determined to evaluate the stability of the method.

For the assessment of accuracy and precision of the method a stock solution was prepared of all metabolites with a concentration 10 times higher than the highest QC concentration. The stock solution was stored at -20°C. This stock solution was used to generate a high, middle and low QC in acetonitrile containing the different ISs. For 3 consecutive days, 3 samples were generated for the high-, middle- and low-QC. Accuracy was determined from the data generated by an inter- and intra-day assay, using Equation 3.3. Precision of the inter- and intra-day data was determined as a percentage (RSD, %), using Equation 3.4. The criteria for the inter- and intra-day accuracy are within 15% of the nominal concentration (100 ± 15%) for all QC samples (González *et al.*, 2014): the criteria for inter- and intra-day precision are below a 15% RSD margin (Peters *et al.*, 2007). The results of the inter- and intra-day accuracy and precision assessment are summarised in Table 3.5.

During the evaluation of the inter- and intra-day accuracy and precision, 33 metabolites did not fit the criteria for inter-day accuracy and 40 metabolites did not fit the criteria for intra-day accuracy. With regards to precision, 24 metabolites did not fit the criteria set for inter- and intra-day precision. The metabolites that did not fit the evaluation criteria were flagged but were still included in the analysis during this study. A quality control procedure was implemented for evaluation of the analysis and to ensure the data that is generated by this method is reliable. The quality control procedure is described in Chapter 4.

Equation 3.3: Accuracy (%)

$$\text{Accuracy} = \frac{\text{mean observed concentration}}{\text{nominal concentration}} \times 100\%$$

Equation 3.4: Precision (RSD, %)

$$\text{RSD} = \frac{\text{standard deviation}}{\text{mean}} \times 100\%$$

Table 3.5: Summary of the inter-day and intra-day accuracy and precision of the metabolites of interest

Metabolite	Low QC					Medium QC					High QC				
	Conc (ng/mL)	Inter-day		Intra-day		Conc (ng/mL)	Inter-day		Intra-day		Conc (ng/mL)	Inter-day		Intra-day	
		RSD %	Acc (%)	RSD %	Acc (%)		RSD %	Acc (%)	RSD %	Acc (%)		RSD %	Acc (%)	RSD %	Acc (%)
Glucose	346	20	73	23	85	693	35	77	35	77	1385	11	78	5	107
Fructose 6-Phosphate	2500	8	96	5	118	5001	4	88	10	88	10001	8	98	4	89
3-phospho-D-glycerate	358	20	120	18	118	715	30	115	30	112	1431	16	120	17	108
2-Phosphoglyceric acid	358	29	87	35	94	715	15	84	10	105	1431	21	110	21	110
Phosphoenolpyruvic acid	808	35	71	23	75	1615	28	69	31	69	3230	30	84	6	111
Pyruvate	846	27	71	17	73	1693	26	76	18	129	6770	31	112	31	112
Oxaloacetic acid	2539	21	75	20	95	5077	19	69	31	81	10154	26	70	21	120
Citric acid	923	6	70	4	97	1846	3	84	6	80	3693	2	79	9	87
cis-Aconitic acid	3346	20	70	6	94	6693	10	99	25	123	13386	8	82	30	129
2-Ketoglutaric acid	2808	13	95	3	102	5616	28	73	28	73	11232	21	69	6	100
Succinic acid	2270	8	110	3	138	4539	4	107	4	107	9079	16	88	16	86
Fumaric acid	223	21	79	20	90	446	26	98	28	98	892	44	97	27	95
Malic acid	1289	12	102	6	101	2577	16	88	16	88	5155	37	71	37	71
Glyoxylic acid	356	6	129	7	132	712	10	90	15	90	1423	24	122	19	75
Fructose	346	5	74	17	79	693	9	76	9	76	1385	11	83	8	110
Lactic acid	17314	11	91	14	126	34627	18	91	35	91	69255	29	80	17	99
3-Hydroxybutyrate	200	12	106	20	94	400	15	91	15	91	800	21	84	10	100
3-Phosphonooxypyruvate	3538	11	122	7	122	7076	22	86	22	86	14152	17	79	24	100
L-Serine-O-Phosphate	712	7	98	7	98	1423	4	107	2	109	2846	16	79	16	79
Serine	1010	25	125	33	135	2020	29	122	15	128	4040	12	118	16	101
Glycine	721	7	71	4	72	1443	8	77	10	75	2886	8	75	3	73
Acetylalanine	504	3	117	3	117	1008	4	85	5	86	2016	8	78	8	78
Alanine	856	7	79	7	95	1712	3	90	3	90	3425	11	84	7	100
Arginine	1674	21	103	8	101	3348	17	121	17	121	6697	12	99	6	103
Asparagine	2539	9	83	16	70	5079	2	91	2	91	10158	5	92	6	106

Table 3.5 (continued): Summary of the inter-day and intra-day accuracy and precision of the metabolites of interest

Metabolite	Low QC					Medium QC					High QC				
	Conc (ng/mL)	Inter-day		Intra-day		Conc (ng/mL)	Inter-day		Intra-day		Conc (ng/mL)	Inter-day		Intra-day	
		RSD %	Acc (%)	RSD %	Acc (%)		RSD %	Acc (%)	RSD %	Acc (%)		RSD %	Acc (%)	RSD %	Acc (%)
Aspartic acid	2558	29	112	20	120	5117	4	80	4	68	10234	9	79	3	114
Citrulline	1684	13	107	14	116	3367	13	95	13	95	6734	10	86	7	89
Cysteine	1164	13	87	5	107	2327	7	90	7	90	4655	9	89	8	96
Cystine	11539	9	70	14	80	23079	8	72	8	72	46158	6	91	6	107
Glutamic acid	1414	11	95	16	103	2828	6	81	6	81	5656	7	97	5	119
Glutamine	4214	5	81	8	90	8427	6	86	6	86	16854	8	84	5	99
Glutathione Oxidized	5886	10	98	6	136	11772	7	98	7	98	23544	13	92	11	92
Glutathione Reduced	295	1	125	1	134	591	3	97	3	97	1181	7	125	7	70
Histidine	2982	5	122	13	83	5964	13	83	30	86	11928	33	75	30	129
Hydroxyproline	1260	17	73	34	69	2520	4	88	4	88	5041	6	86	3	104
Isoleucine	1261	5	114	4	111	2521	5	76	5	76	5042	4	107	3	116
Leucine	1261	6	80	7	86	2521	2	89	2	89	5042	5	91	3	103
Lysine	1405	7	80	12	86	2810	2	89	2	89	5619	5	92	4	105
Phenylalanine	32	9	92	14	64	63	17	86	10	70	127	20	74	7	109
Proline	221	18	81	22	80	443	4	89	4	89	885	8	95	6	105
Taurine	24041	11	103	14	88	48082	5	96	5	96	96165	8	87	9	97
Threonine	1145	11	82	10	79	2290	15	83	15	83	4579	11	90	12	108
Tryptophan	392	16	81	5	114	785	9	90	9	90	1570	11	91	3	105
Tyrosine	1741	9	80	12	76	3482	4	87	4	87	6964	6	89	8	106
Valine	225	3	76	7	71	450	6	78	6	102	901	10	97	14	119
Methionine	287	29	92	35	80	573	9	91	9	91	1147	12	91	8	104
SAM	768	9	74	6	71	1535	4	82	4	89	3070	7	79	10	102
SAH	739	6	85	3	75	1477	4	89	10	108	2955	6	87	20	75
Homocysteine	2597	10	124	8	135	5194	12	108	2	88	10387	11	81	3	110
Creatine	252	7	86	7	96	504	2	88	3	92	1008	3	96	5	99

Table 3.5 (continued): Summary of the inter-day and intra-day accuracy and precision of the metabolites of interest

Metabolite	Low QC					Medium QC					High QC				
	Conc (ng/mL)	Inter-day		Intra-day		Conc (ng/mL)	Inter-day		Intra-day		Conc (ng/mL)	Inter-day		Intra-day	
		RSD %	Acc (%)	RSD %	Acc (%)		RSD %	Acc (%)	RSD %	Acc (%)		RSD %	Acc (%)	RSD %	Acc (%)
Creatinine	217	4	88	3	77	435	3	92	1	90	870	5	93	5	103
Folic acid	848	2	84	15	73	1697	1	90	2	99	3393	4	90	5	109
Tetrahydrofolate	4281	6	130	15	128	8563	2	99	6	103	17125	4	103	10	93
Acetyl-CoA	1556	6	108	7	96	3112	6	103	10	90	6224	15	83	4	116
CoA	1475	25	94	17	106	2950	14	89	25	130	5901	26	88	22	97
FAD	1510	30	128	9	131	3020	25	126	9	78	6040	20	110	26	100
NAD	1275	12	103	23	108	2550	9	78	18	70	5101	15	72	25	100
NADP	1429	14	95	3	99	2858	18	70	18	86	5716	18	73	38	100
NADH	2559	28	97	26	70	5119	18	86	25	69	10238	30	74	29	76
NADPH	3582	24	101	18	108	7164	21	113	20	98	14329	14	104	26	96
Adenine	260	2	75	5	85	519	8	72	2	82	1039	6	70	2	106
Guanine	290	4	72	8	87	581	2	82	2	93	1162	7	84	5	95
Hypoxanthine	262	5	96	16	79	523	2	93	5	96	1046	3	84	9	98
Thymine	242	10	95	6	85	485	5	96	2	94	970	7	86	4	105
Uracil	1077	3	85	4	80	2154	2	94	2	93	4309	3	98	5	108
Cytosine	53	4	77	4	109	107	2	93	2	87	213	3	99	2	101
Uridine	469	3	94	10	91	939	2	87	3	81	1877	7	82	4	103
Xanthine	1462	6	80	34	114	2924	3	81	28	124	5847	7	83	29	127
Adenosine	514	6	90	8	81	1027	6	94	2	82	2055	2	89	3	109
Guanosine	544	4	89	8	91	1089	2	82	3	93	2178	8	86	9	101
Inosine	2578	5	90	8	127	5155	3	93	9	72	10311	6	88	14	115
ADP	821	23	126	18	138	1642	9	72	26	129	3285	9	103	32	128
GDP	852	28	103	24	96	1704	21	91	10	69	3408	20	72	13	99
UDP	777	5	95	2	99	1554	6	106	10	120	3108	17	73	14	104
ADP-Glucose	1133	27	84	31	73	2266	28	102	36	69	4531	25	92	29	100
ADP-Ribose	1075	28	89	38	125	2150	36	114	29	122	4301	26	72	28	128

Table 3.5 (continued): Summary of the inter-day and intra-day accuracy and precision of the metabolites of interest

Metabolite	Low QC					Medium QC					High QC				
	Conc (ng/mL)	Inter-day		Intra-day		Conc (ng/mL)	Inter-day		Intra-day		Conc (ng/mL)	Inter-day		Intra-day	
		RSD %	Acc (%)	RSD %	Acc (%)		RSD %	Acc (%)	RSD %	Acc (%)		RSD %	Acc (%)	RSD %	Acc (%)
ATP	975	24	72	28	69	1950	24	70	27	109	3900	21	69	22	74
GTP	5029	6	126	9	131	10058	12	130	20	125	20115	7	128	21	129
AMP	667	4	124	7	93	1335	8	104	5	102	2670	8	113	2	110
CMP	1243	3	83	5	98	2485	11	84	5	108	4970	15	79	15	79
GMP	698	16	95	22	88	1396	5	98	9	85	2793	6	107	2	116
IMP	669	19	71	18	108	1339	9	85	10	106	2677	8	102	4	102
UMP	1246	10	106	9	94	2493	3	109	2	114	4985	15	84	15	84
cAMP	3164	22	95	33	104	6328	9	94	17	70	12656	4	110	10	127
cGMP	664	5	87	6	92	1327	17	79	3	83	2654	14	97	4	106
dAMP	637	3	83	6	106	1273	7	86	5	104	2547	13	76	13	76
dCMP	590	9	94	11	94	1181	6	106	15	98	2362	6	107	7	104
dTMP	619	20	103	18	111	1239	21	89	6	103	2477	19	120	5	108
Carnitine	1549	12	96	9	75	3098	6	103	3	91	6196	6	110	9	100
Orotic acid	300	5	86	3	97	600	3	91	8	81	1200	6	85	5	103
Salicylic acid	265	7	74	17	73	531	8	81	3	86	1062	11	72	1	104
Ascorbic acid	3385	11	72	20	94	6770	3	86	26	69	27082	20	111	20	111
Ketoisovalerate	223	12	73	16	88	446	14	74	4	85	893	17	116	5	100
Kynurenine	400	17	79	5	89	801	4	85	3	88	1602	9	82	4	104
3-OH-Kynurenine	431	6	82	4	81	862	3	88	10	74	1724	3	92	4	105
Quinolinic acid	321	5	69	9	75	642	11	74	15	74	1285	3	104	6	105
Kynurenic acid	364	29	105	39	96	727	24	74	12	87	1454	28	79	9	106
Pantothenic acid	421	12	73	10	69	843	12	87	3	89	1685	7	99	3	111
Maleic acid	223	21	79	20	90	446	26	98	28	98	892	44	97	27	95
Ketoleucine	3126	7	72	15	72	6253	3	89	10	80	12506	4	102	3	98

*Concentration (Conc) and Accuracy (Acc)

3.4.4 Carryover

Carryover is a phenomenon caused by residual amounts of analytes present in the analytical instrument after an injection. This can affect the accuracy and precision of results (González *et al.*, 2014). During the validation of the analytical method the carryover was assessed by injecting a blank sample (only ACN) after a high calibration sample throughout the entire validation process. The detectability of the metabolites in the blank samples was evaluated by comparing the detectability to the LOD of the metabolites. Most metabolites were undetectable in the blank samples. Although the detection of some metabolites is highly sensitive and requires a broad calibration range, a residual amount was detectible in some of the blank samples, but it was below the LOD of the metabolite.

3.5 Results and discussion

Since the aim of the study was to establish a standardised LC-MS/MS method for targeted metabolic profiling of biological matrices, different biological matrices were analysed to evaluate the compatibility of the method. During the evaluation of the compatibility of the method for various biological matrices the detectability of the identified metabolites, listed in Table 3.1, were assessed by implementing the developed targeted metabolic profiling method. The following matrices were used for the evaluation of the compatibility of the method: lung, liver, plasma, hypothalamus and liver lysate. One tissue sample per animal was available and originated from already euthanized animals. The tissue samples were collected from storage at the Drug Discovery Sciences Department of Boehringer Ingelheim, Germany. Three lung, liver, hypothalamus, liver lysate and plasma samples were used for the compatibility test and these samples originated from three different healthy C57BL/6J mice. The use of these samples were ethically approved by the Regierungspräsidium in Tübingen, Germany (TVV 12-012) and additional ethical approval for the use of the matrices was obtained from the Ethics committee (AnimCare) of the North-West University (NWU-00275-17-A5). Fibroblasts samples were also analysed to evaluate the compatibility of the method. Three fibroblast samples from normal human lung fibroblasts, obtained from healthy, non-smoking donors (Lonza Rockland Inc, Rockland, ME, USA) were collected from the Immunology and Respiratory department at Boehringer Ingelheim, Germany for the analysis. A summary of the detectability of the different metabolites within the various matrices are given in Table 3.6. All metabolites identified in the various matrices were detectible above their respective LODs after the analysis of the three samples per matrix.

Table 3.6: Summary of the detectability of the metabolites of interest in various matrices

Group	Metabolite	Matrix (n=3)					
		Lung	Fibroblast	Liver	Liver lysate	Plasma	Hypo-thalamus
Glycolysis	Glucose	X	X	X	X	X	X
	Fructose 6-Phosphate		X	X	X		X
	3-phospho-D-glycerate	X	X	X			
	2-Phosphoglyceric acid	X		X			
	Phosphoenolpyruvic acid	X			X		
	Pyruvate						
TCA	Oxaloacetic acid	X			X		
	Citric acid						
	cis-Aconitic acid	X			X		
	2-Ketoglutaric acid	X	X	X		X	
	Succinic acid						X
	Fumaric acid	X	X		X	X	X
	Malic acid	X					X
	Glyoxylic acid	X	X	X	X		X
	Fructose	X	X	X	X	X	X
	Lactic acid	X	X	X	X	X	X
Serine bio-synthesis	3-Hydroxybutyrate		X			X	X
	3-Phosphonooxypyruvate						
	L-Serine-O-Phosphate	X		X			
	Serine	X	X	X	X	X	X
	Glycine	X	X	X	X	X	X
Amino acids	Acetylalanine	X		X	X	X	X
	Alanine	X	X	X	X	X	X
	Arginine	X	X	X		X	X
	Asparagine	X	X	X	X	X	X
	Aspartic acid	X		X	X		
	Citrulline	X		X	X	X	
	Cysteine						
	Cystine						
	Glutamic acid	X	X	X	X	X	X
	Glutamine	X	X	X	X	X	X
	Glutathione Oxidized	X	X	X	X		X
	Glutathione Reduced	X	X	X	X		X
	Histidine	X	X	X	X	X	X
	Hydroxyproline					X	
	Isoleucine	X	X	X	X	X	X
	Leucine	X	X	X	X	X	X
	Lysine	X	X	X	X	X	
	Phenylalanine	X	X	X	X	X	X
	Proline	X	X	X	X	X	X
	Taurine	X	X	X	X	X	X
	Threonine	X	X	X	X	X	X
	Tryptophan	X	X	X	X	X	X
	Tyrosine	X	X	X	X	X	X
	Valine	X	X	X	X	X	X
Methionine cycle	Methionine	X	X	X	X	X	X
	SAM	X				X	X
	SAH	X					X
	Homocysteine	X				X	X
	Creatine	X	X			X	X

Table 3.6 (continued): Summary of the detectability of the metabolites of interest in various matrices.

Group	Metabolite	Matrix (n=3)					
		Lung	Fibroblast	Liver	Liver lysate	Plasma	Hypo-thalamus
Methionine cycle	Creatinine	X	X			X	X
	Folic acid	X	X				
	Tetrahydrofolate						
Coenzyme	Acetyl-CoA		X				X
	CoA		X		X		X
	FAD	X	X	X	X		X
	NAD		X	X	X		X
	NADP		X	X	X		
	NADH		X	X			X
	NADPH						X
Nucleobase	Adenine	X	X	X	X	X	X
	Guanine	X	X	X	X	X	
	Hypoxanthine	X	X	X	X		X
	Thymine	X		X	X		X
	Uracil	X	X		X	X	X
	Cytosine	X		X	X	X	X
	Uridine	X	X	X	X		X
	Xanthine	X	X	X	X	X	X
Nucleoside	Adenosine	X	X	X	X	X	X
	Guanosine	X	X	X	X	X	X
	Inosine	X	X	X	X		X
Nucleotide	ADP	X	X	X			X
	GDP		X	X	X		X
	UDP		X	X	X		X
Nucleotide derivative	ADP-Glucose			X			
	ADP-Ribose			X			
Nucleotide	ATP	X	X	X	X		X
	GTP						
	AMP	X	X	X	X		X
	CMP		X	X	X		X
	GMP	X	X	X	X		X
	IMP	X	X	X	X		X
	UMP	X	X	X	X		X
	cAMP						
	cGMP				X		
	dAMP		X	X	X		
	dCMP		X				
	dTMP		X		X		
Other	Carnitine	X	X	X	X	X	X
	Orotic acid	X			X	X	X
	Salicylic acid	X			X	X	X
	Ascorbic acid						
	Ketoisovalerate					X	
	Kynurenine		X				
	3-OH-Kynurenine						
	Quinolinic acid						
	Kynurenic acid						
	Pantothenic acid	X	X	X	X	X	X
	Maleic acid	X	X		X	X	X
	Ketoleucine						

The comparison of the detectability of the metabolites in the various matrices showed that most of the metabolites can be detected in all the matrices that were tested but a few metabolites still could not be detected in all the matrices. The plasma analysis yielded the lowest number of detectible metabolites. Most of the metabolites not detectible in plasma were intracellular metabolites from the glycolysis pathway and TCA cycle. The endogenous levels of the different metabolites differ greatly between the different matrices and for some matrices the endogenous levels were found to be below the LOD of this method. Some intermediate metabolites such as kynurenine, 3-hydroxykynurenine and quinolinic acid were not detectible in any of the analysed matrices. This may be due to rapid downstream metabolism of the metabolites (Fuertig *et al.*, 2016) and therefore detectible concentrations are low. Cysteine and cystine were also not detected in any of the matrices. It is known that these two metabolites are prone to oxidation (Johnson *et al.*, 2008) which may be the reason why it is not detectible without a reduction agent present in the sample preparation.

3.6 Conclusion

In this chapter the method development of a standardised LC-MS/MS method for targeted metabolic profiling of biological matrices was described. After identifying the metabolites to be included in the targeted approach, all parameters of the MS were optimised for each metabolite to ensure the highest sensitivity is achieved for the detection of the metabolites. Once both the sample preparation procedure that is non-selective and the HILIC LC separation method were developed and optimised, the established workflow was evaluated by performing a quality assessment.

During the quality assessment the detectible range and linearity of each metabolite included in the targeted metabolic profiling method were determined, as well as the limits of quantification and detection for each metabolite. The established targeted metabolic profiling method was also validated by performing an inter- and intra-day accuracy and precision analysis. Some metabolites did not fit the evaluation criteria for the inter- and intra-day accuracy and precision. These metabolites include homocysteine and the reduced form of glutathione, which is prone to oxidation and consequently interfere with accuracy analysis. To ensure that the method can be used to generate reliable results a quality control procedure was implemented for the evaluation of the analysis. This quality control procedure included the analysis of a QC sample, with known concentrations of all the metabolites, at the beginning, middle and end of the analysis. The multiple analysis of the QC sample was then used to evaluate the intra-day accuracy and precision and if a metabolite did not fit the quality control criteria, the results of this metabolite were excluded from the data set. An in-depth description of the quality control procedure together with the quality control criteria is described in Chapter 4.

As part of the quality assessment of the developed method, various matrices from different origin, mouse and human, were tested to determine the compatibility of the method. With the use of the developed method the baseline levels of most of the metabolites could be detected for the healthy C57BL/6J mouse model in the different tissue matrices. The final description of the optimised targeted metabolic profiling method is described in Chapter 4. The method was used to generate a metabolic profile for a fibrotic lung animal model, which is described in Chapter 5 in a full length article prepared for submission to the *Respiratory Medicine* journal.

CHAPTER 4: LC-MS/MS METHOD FOR TARGETED METABOLIC PROFILING

4.1 Metabolites of interest

The metabolites listed in Table 3.1 are all compatible in terms of detectability with the developed method. The metabolites are products, intermediates and cofactors present in the central carbon system, glycolysis, TCA cycle, the serine biosynthesis pathway as well as all amino acids present in human metabolic pathways. A visual representation of the compatible metabolites (highlighted in green) can be seen in Figure 4.1.

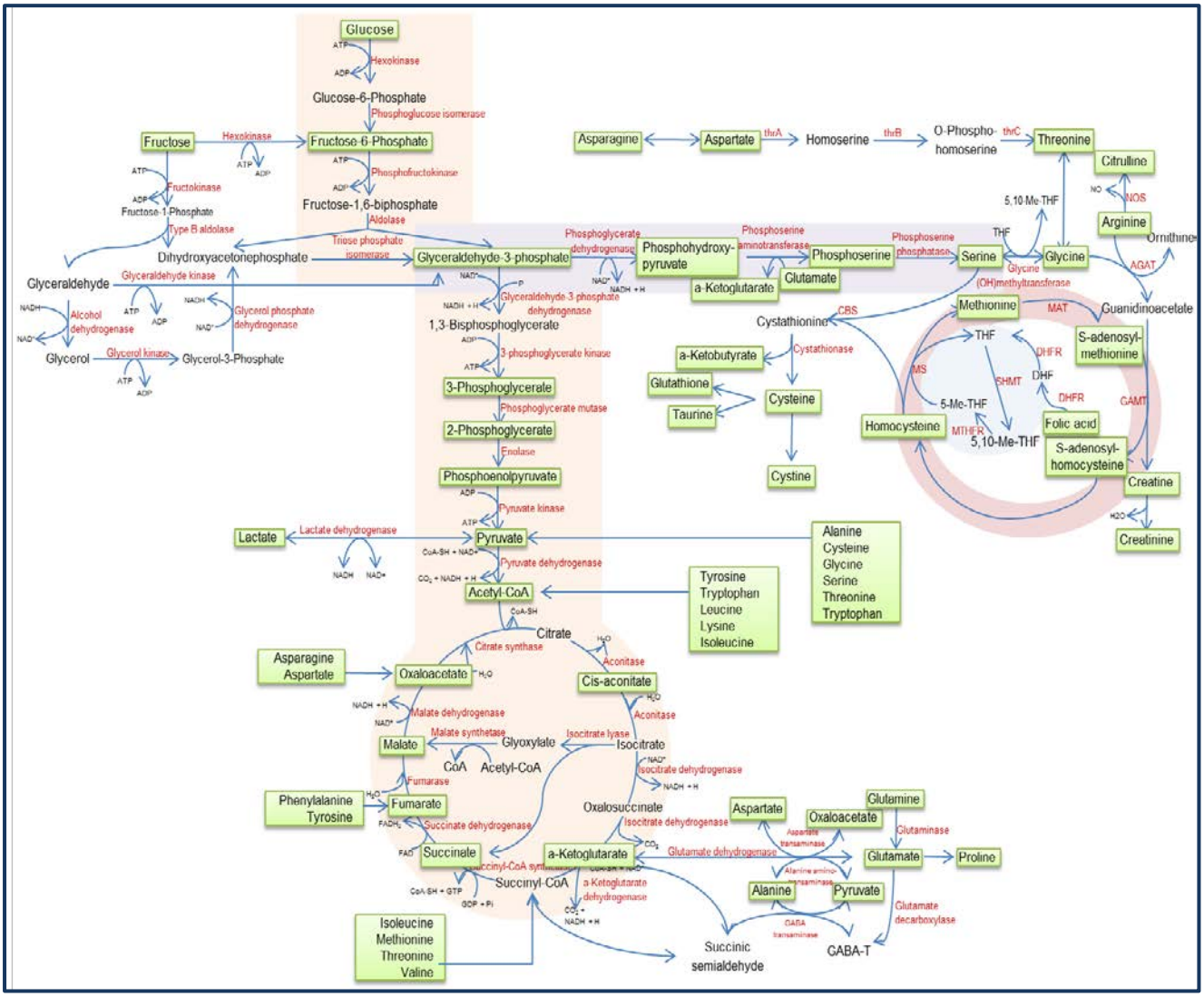

Figure 4.1: Visual representation of metabolites compatible for detection by the developed method. All metabolites compatible with the method is highlighted in green. This diagram was generated using information obtained from the Kyoto Encyclopaedia of Genes and Genomes data base (Genome.jp, 2017).

4.2 MS parameters

The precursor and product ion of each metabolite was determined with optimal ionisation conditions including the optimal DP and CE values (see Table 4.3). The optimal MS parameters for the internal standards used, were also determined (see Table 4.4). Other MS parameters that were kept at a fixed value can be seen in Table 4.1.

Table 4.1: Fixed MS parameters

Parameter	Value
TEM	500 °C
ISV	4500 V
CUR	40 psi
GS1	50 psi
GS2	50 psi

4.3 LC conditions

A 20 min HILIC method was developed using the Luna NH_2 column (2 mm x 150 mm, 5 µm, 100 Å). Solvent A consisted of 20 mM ammonium acetate + 20 mM ammonium hydroxide in 95:5 H_2O: ACN ratio with a pH 9 and solvent B consisted of only ACN. A 0.4 mL/min flowrate was used together with the gradient described in Table 4.2.

Table 4.2: Gradient used for the 20 min HILIC method

Time (min)	Solvent A (%)	Solvent B (%)
0	15	85
1.5	15	85
14	98	2
17	98	2
18.5	15	85
20	15	85

Table 4.3: Summary of all the MS parameters for the different metabolites included in the method

Group	Metabolite	Formula	M_r	$[M+H]^+$ (*m/z*)	$[M+H]^-$ (*m/z*)	Product ion (*m/z*)	DP (V)	CE (V)	RT (min)	Internal Standard
Glycolysis	Glucose	$C_6H_{12}O_6$	180.063		178.957	89.000	-55	-12	4.38	C13-Glucose
	Fructose-6-Phosphate	$C_6H_{13}O_9P$	260.029	261.036		109.028	20	20	11.98	D3-Quinolinic acid POS
	3-Phospho-D-Glycerate	$C_3H_7O_7P$	185.992	187.000		140.900	20	15	14.40	D3-Quinolinic acid POS
	2-Phosphoglyceric acid	$C_3H_7O_7P$	185.992	187.001		98.984	20	40	15.20	D3-Quinolinic acid POS
	Phosphoenolpyruvic acid	$C_3H_5O_6P$	167.982		166.974	78.958	-20	-20	12.18	D5-Glutamic acid NEG
	Pyruvate	$C_3H_4O_3$	88.016		86.787	42.970	-25	-10	11.47	D5-Glutamic acid NEG
TCA	Oxaloacetic acid	$C_4H_4O_5$	132.006		131.000	86.975	-65	-16	11.44	D5-Glutamic acid NEG
	Citric acid	$C_6H_8O_7$	192.027		191.019	111.008	-20	-20	13.8	D3-Quinolinic acid NEG
	cis-Aconitic acid	$C_6H_6O_6$	174.016		173.009	85.029	-20	-20	11.46	D5-Glutamic acid NEG
	2-Ketoglutaric acid	$C_5H_6O_5$	146.021		145.014	100.900	-42	-11	11.97	D5-Glutamic acid NEG
	Succinic acid	$C_4H_6O_4$	118.026		117.019	73.029	-20	-40	12.00	D5-Glutamic acid NEG
	Fumaric acid	$C_4H_4O_4$	116.011		115.000	71.000	-20	-20	11.97	D5-Glutamic acid NEG
	Malic acid	$C_4H_6O_5$	134.021		133.013	71.014	-20	-40	11.91	D5-Glutamic acid NEG
	Glyoxylic acid	$C_2H_2O_3$	74.000		72.800	45.000	-30	-12	12.52	D5-Glutamic acid NEG
	Fructose	$C_6H_{12}O_6$	180.063		178.957	89.000	-55	-12	3.76	C13-Fructose
	Lactic acid	$C_3H_6O_3$	90.031		88.900	42.800	-50	-15	6.52	C13-Serine NEG
Serine bio-synthesis	3-Hydroxybutyrate	$C_4H_8O_3$	104.040		102.895	59.042	-10	-12	6.20	C13-Serine NEG
	3-Phosphonooxypyruvate	$C_3H_5O_7P$	183.977		182.900	78.869	-55	-40	13.90	D3-Quinolinic acid NEG
	L-Serine-O-Phosphate	$C_3H_8NO_6P$	185.008	186.000		88.200	20	20	12.11	D3-Quinolinic acid POS
	Serine	$C_3H_7NO_3$	105.042	105.700		60.000	35	15	6.40	C13-Serine POS
	Glycine	$C_2H_5NO_2$	75.032	76.040		29.700	20	20	6.02	C13-Glycine
Amino acids	Acetylalanine	$C_5H_9NO_3$	131.058		130.050	88.040	-20	-20	6.61	C13-Serine NEG
	Alanine	$C_3H_7NO_2$	89.0477	89.800		62.200	20	9	5.58	D5-Tryptophan
	Arginine	$C_6H_{14}N_4O_2$	174.111	175.119		70.067	20	20	6.95	D5-Kynurenic acid
	Asparagine	$C_4H_8N_2O_3$	132.053	133.061		74.020	20	20	6.41	C13-Serine POS
	Aspartic acid	$C_4H_7NO_4$	133.037	133.986		43.010	91	61	9.19	D5-Glutamic acid POS
	Citrulline	$C_6H_{13}N_3O_3$	175.095	176.104		70.066	20	20	6.14	C13-Glycine
	Cysteine	$C_3H_7NO_2S$	121.019	122.028		58.996	20	40	6.95	D5-Kynurenic acid
	Cystine	$C_6H_{12}N_2O_4S_2$	240.020	241.235		74.000	66	31	10.17	D5-Glutamic acid POS
	Glutamic acid	$C_5H_9NO_4$	147.053	148.100		84.000	65	20	9.28	D5-Glutamic acid POS
	Glutamine	$C_5H_{10}N_2O_3$	146.069	147.077		84.045	20	20	6.15	C13-Glycine
	Glutathione Oxidized	$C_{20}H_{32}N_6O_{12}S_2$	612.152		611.144	306.075	-20	-20	12.04	D5-Glutamic acid NEG
	Glutathione Reduced	$C_{10}H_{17}N_3O_6S$	307.083		306.076	143.046	-20	-20	11.63	D5-Glutamic acid NEG

Table 4.3 (continued): Summary of all the MS parameters for the different metabolites included in the method

Group	Metabolite	Formula	M_r	$[M+H]^+$ (*m/z*)	$[M+H]^-$ (*m/z*)	Product ion (*m/z*)	DP (V)	CE (V)	RT (min)	Internal Standard
Amino acids	Histidine	$C_6H_9N_3O_2$	155.069	156.077		110.071	20	20	6.60	D5-Kynurenic acid
	Hydroxyproline	$C_5H_9NO_3$	131.058	131.955		86.000	76	9	5.42	D5-Phenylalanine
	Isoleucine	$C_6H_{13}NO_2$	131.094	132.102		86.097	20	10	4.30	N15-Isoleucine
	Leucine	$C_6H_{13}NO_2$	131.094	132.100		44.049	20	40	4.23	C13-Adenosine
	Lysine	$C_6H_{14}N_2O_2$	146.105	147.113		84.081	20	20	7.72	D3-Lysine
	Phenylalanine	$C_9H_{11}NO_2$	165.079	166.087		120.081	20	20	4.81	D5-Phenylalanine
	Proline	$C_5H_9NO_2$	115.063	116.071		70.066	20	40	5.30	D5-Tryptophan
	Taurine	$C_2H_7NO_3S$	125.014	126.023		64.969	20	40	5.84	C13-Glycine
	Threonine	$C_4H_9NO_3$	119.058	120.066		56.050	20	40	6.05	C13-Glycine
	Tryptophan	$C_{11}H_{12}N_2O_2$	204.089	205.098		146.059	20	20	4.96	D5-Tryptophan
	Tyrosine	$C_9H_{11}NO_3$	181.073	182.082		91.055	20	40	5.76	D5-Tryptophan
	Valine	$C_5H_{11}NO_2$	117.079	118.087		72.081	20	10	4.89	D5-Tryptophan
Methionine cycle	Methionine	$C_5H_{11}NO_2S$	149.051	150.059		104.053	20	10	5.02	D5-Tryptophan
	SAM	$C_{15}H_{23}N_6O_5S$	399.145	400.196		251.085	51	21	6.27	C13-Glycine
	SAH	$C_{14}H_{20}N_6O_5S$	384.120	385.074		136.062	20	21	6.69	D5-Kynurenic acid
	Homocysteine	$C_4H_9NO_2S$	135.035	136.040		90.030	10	15	4.22	N15-Isoleucine
	Creatine	$C_4H_9N_3O_2$	131.069	132.056		90.050	10	17	5.10	D5-Phenylalanine
	Creatinine	$C_4H_7N_3O$	113.059	114.069		44.100	10	21	2.29	D3- Adenine
	Folic acid	$C_{19}H_{19}N_7O_6$	441.139	442.180		295.093	20	21	16.00	C13-Acetyl-CoA POS
	Tetrahydrofolate	$C_{19}H_{23}N_7O_6$	445.430	446.262		299.139	10	27	11.97	D3-Quinolinic acid POS
Coenzyme	Acetyl-CoA	$C_{23}H_{38}N_7O_{17}P_3S$	809.120		808.117	408.000	-200	-50	16.23	C13-Acetyl-CoA NEG
	CoA	$C_{21}H_{36}N_7O_{16}P_3S$	767.115		766.107	407.800	-200	-48	15.72	C13-Acetyl-CoA NEG
	FAD	$C_{27}H_{33}N_9O_{15}P_2$	785.157	786.165		348.069	20	20	12.02	D3-Quinolinic acid POS
	NAD	$C_{21}H_{27}N_7O_{14}P_2$	663.109	664.117		136.061	20	40	9.53	D5-Glutamic acid POS
	NADP	$C_{21}H_{28}N_7O_{17}P_3$	743.075	744.083		136.061	20	40	14.41	D3-Quinolinic acid POS
	NADH	$C_{21}H_{29}N_7O_{14}P_2$	665.441	666.128		136.060	20	40	11.62	D3-Quinolinic acid POS
	NADPH	$C_{21}H_{30}N_7O_{17}P_3$	745.090	746.099		729.018	200	25	15.97	N15-ATP
Nucleobase	Adenine	$C_5H_5N_5$	135.054	136.062		119.035	20	20	3.34	C13-Adenosine
	Guanine	$C_5H_5N_5O$	151.049	152.057		135.030	20	20	4.56	D4-Kynurenine
	Hypoxanthine	$C_5H_4N_4O$	136.039	137.047		55.029	20	40	5.08	D5-Tryptophan
	Thymine	$C_5H_6N_2O_2$	126.042		124.800	42.000	-35	-36	1.65	D3-Tymine
	Uracil	$C_4H_4N_2O_2$	112.027		111.019	41.998	-20	-20	1.98	D3-Tymine
	Cytosine	$C_4H_5N_3O$	111.000	112.030		52.010	61	41	2.95	C13-Adenosine
	Uridine	$C_9H_{12}N_2O_6$	244.069	245.077		113.035	20	20	2.75	C13-Adenosine
	Xanthine	$C_5H_4N_4O_2$	152.033	153.041		110.035	20	20	9.20	D5-Glutamic acid POS

Table 4.3 (continued): Summary of all the MS parameters for the different metabolites included in the method

Group	Metabolite	Formula	M_r	$[M+H]^+$ (*m/z*)	$[M+H]^-$ (*m/z*)	Product ion (*m/z*)	DP (V)	CE (V)	RT (min)	Internal Standard
Nucleoside	Adenosine	$C_{10}H_{13}N_5O_4$	267.097	268.105		136.200	20	23	2.69	C13-Adenosine
	Guanosine	$C_{10}H_{13}N_5O_5$	283.091	284.099		152.050	20	20	5.14	D5-Tryptophan
	Inosine	$C_{10}H_{12}N_4O_5$	268.080	269.100		137.100	70	17	5.65	D5-Tryptophan
Nucleotide	ADP	$C_{10}H_{15}N_5O_{10}P_2$	427.029	428.037		136.060	20	40	14.75	N15-ATP
	GDP	$C_{10}H_{15}N_5O_{11}P_2$	443.024		442.017	78.959	-20	-40	16.36	C13-Acetyl-CoA NEG
	UDP	$C_9H_{14}N_2O_{12}P_2$	404.002		402.994	158.920	-20	-40	15.04	D3-Quinolinic acid NEG
Nucleotide derivative	ADP-Glucose	$C_{16}H_{25}N_5O_{15}P_2$	589.082	590.090		136.060	20	20	11.96	D3-Quinolinic acid POS
	ADP-Ribose	$C_{15}H_{23}N_5O_{14}P_2$	559.071	560.079		136.060	20	40	11.45	D3-Quinolinic acid POS
Nucleotide	ATP	$C_{10}H_{16}N_5O_{13}P_3$	506.995	508.004		136.060	20	40	18.24	N15-ATP
	GTP	$C_{10}H_{16}N_5O_{14}P_3$	522.990		521.983	158.920	-155	-50	13.59	D3-Quinolinic acid NEG
	AMP	$C_{10}H_{14}N_5O_7P$	347.063	347.900		136.300	40	25	12.24	N15-AMP
	CMP	$C_9H_{14}N_3O_8P$	323.051	324.059		112.050	20	20	12.08	D3-Quinolinic acid POS
	GMP	$C_{10}H_{14}N_5O_8P$	363.050	364.066		152.050	20	20	13.33	D3-Quinolinic acid POS
	IMP	$C_{10}H_{13}N_4O_8P$	348.047	349.055		137.046	20	20	12.29	D3-Quinolinic acid POS
	UMP	$C_9H_{13}N_2O_9P$	324.035	325.044		97.020	20	20	12.25	D3-Quinolinic acid POS
	cAMP	$C_{10}H_{12}N_5O_6P$	329.052	330.060		136.060	20	20	9.19	D5-Glutamic acid POS
	cGMP	$C_{10}H_{12}N_5O_7P$	345.047	346.055		152.050	20	20	10.07	D5-Glutamic acid POS
	dAMP	$C_{10}H_{14}N_5O_6P$	331.068	332.076		136.060	20	20	12.28	N15-AMP
	dCMP	$C_9H_{14}N_3O_7P$	307.056	308.065		112.050	20	20	12.06	D3-Quinolinic acid POS
	dTMP	$C_{10}H_{15}N_2O_8P$	322.056	323.064		81.000	20	25	12.02	D3-Quinolinic acid POS
Other	Carnitine	$C_7H_{15}NO_3$	161.105	162.310		102.909	31	23	5.43	D5-Tryptophan
	Orotic acid	$C_5H_4N_2O_4$	156.017		155.009	111.010	-20	-20	7.27	C13-Serine NEG
	Salicylic acid	$C_7H_6O_3$	138.032		137.024	93.030	-20	-20	4.66	Glucose C13
	Ascorbic acid	$C_6H_8O_6$	176.032		174.830	86.900	-200	-28	9.28	D5-Glutamic acid NEG
	Ketoisovalerate	$C_5H_8O_3$	116.047		114.847	70.924	-5	-10	4.93	C13-Glucose
	Kynurenine	$C_{10}H_{12}N_2O_3$	208.213	209.100		192.000	20	24	4.73	D4-Kynurenine
	3-OH-Kynurenine	$C_{10}H_{12}N_2O_4$	224.079	225.100		208.000	40	13	6.06	C13-Glycine
	Quinolinic acid	$C_7H_5NO_4$	167.021	168.000		78.000	20	28	12.2	D3-Quinolinic acid POS
	Kynurenic acid	$C_{10}H_7NO_3$	189.042	190.100		144.100	40	25	6.51	D5-Kynurenic acid
	Pantothenic acid	$C_9H_{17}NO_5$	219.110	220.200		90.100	90	21	6.99	D5-Kynurenic acid
	Maleic acid	$C_4H_4O_4$	116.010		115.000	71.000	-20	-20	11.97	C13-Glutamic acid NEG
	Ketoleucine	$C_6H_{10}O_3$	130.060		128.986	85.079	-10	-12	15.00	D3-Quinolinic acid NEG

Table 4.4: MS parameters of all the internal standards included in the method

Group	Metabolite	M_r	$[M+H]^+$ (*m/z*)	$[M+H]^-$ (*m/z*)	Product ion (*m/z*)	DP (V)	CE (V)	RT (min)	Conc. Used (ng/mL)
Glycolysis	D-Glucose-13C6	186.11		185.024	91.673	-80	-12	4.38	991.97
	D-Fructose-13C6	186.11		185.024	91.673	-80	-12	3.76	1116.66
Amino acids	L-Serine-13C3,15N,2,3,3-d3 POS	112.08	113.000		66.080	10	15	6.40	1494.03
	L-Serine-13C3,15N,2,3,3-d3 NEG	112.08		110.870	77.990	-20	-16	6.40	1494.03
	L-Glutamic acid-13C5 POS	152.09	153.000		87.900	10	21	9.28	2027.36
	L-Glutamic acid-13C5 NEG	152.09		151.000	107.100	-15	-18	9.28	2027.36
	Glycine C13	76.06	77.200		31.000	20	20	6.02	50655.96
	L-Isoleucine-13C6,15N	138.12	139.034		92.098	31	13	4.30	921.26
	L-Leucine-5,5,5-d3	134.19	135.044		89.127	46	13	4.23	895.05
	L-Lysine-4,4,5,5-d4	186.67	151.068		88.100	36	21	7.72	12445.29
	L-Phenyl-d5-alanine	170.22	171.045		125.115	31	17	4.81	1135.37
	Thymine-d4 (methyl-d3,6-d1)	130.14		128.900	42.010	-10	-30	1.65	868.03
	Tryptophan-D5	209.26	210.100		122.100	86	21	4.96	1395.76
Coenzyme	Acetyl-1,2-13C2 Coenzyme A POS	811.56	812.130		305.130	200	50	16.23	54078.70
	Acetyl-1,2-13C2 Coenzyme A NEG	811.56		810.125	463.050	-200	-50	16.23	54078.70
Nucleobase	2-(Methyl-13C,d3-thio) adenine	185.23	186.200		134.000	26	15	1.93	1235.48
Nucleoside	[1',2',3',4',5'-13C5]adenosine	272.21	114.069		44.100	64	27	2.69	182.38
Nucleotide	Adenosine-15N5 5′-monophosphate	352.19	352.948		141.079	56	23	12.24	23480.51
	Adenosine-15N5 5′-triphosphate	512.15	513.000		141.000	20	40	18.24	3416.04
Other	Quinolinic acid-4,5,6-D3 POS	170.14	171.000		81.000	20	28	12.20	1134.83
	Quinolinic acid-4,5,6-D3 NEG	170.14		169.026	125.031	-10	-27	12.20	1134.83
	D4-Kynurenine	212.24	213.100		140.100	20	13	4.73	1415.64
	Kynurenic acid-3,5,6,7,8-d5	194.20	195.100		149.100	40	25	6.51	1295.31

4.4 Sample preparation

A standardised sample preparation protocol was developed for the following matrices: plasma, lung, liver, hypothalamus and cultured cells. The following steps were performed as part of the protocol: homogenisation, protein precipitation and transfer.

4.4.1 Homogenisation

Homogenisation was performed in a Precellys homogeniser (Bertin Corp. Maryland: USA). Lung and liver samples were homogenised in a 1:4 (*w/v*) ratio with an ACN: MeOH 1:1 ratio mixture. The hypothalamus samples were homogenised in 600 µL ACN: MeOH. Since the hypothalamus is so small, 600 µL was used to ensure sufficient liquid is available for homogenisation. After homogenisation the samples were centrifuged for 5 min at 13 000 rpm.

4.4.2 Protein precipitation

The following steps were the same for all matrices: 5 µL of the supernatant/plasma/QQC/QC pool was added to 70 µL of ACN: MeOH 1:1 in a 96-well plate and placed in a -20°C freezer for at least 15 min for optimal protein precipitation. The samples were then centrifuged for 4 min at 4 000 rpm.

4.4.3 Transfer

After centrifugation 30 µL of supernatant was transferred to 170 µL ACN containing the different internal standards with appropriate concentrations (see Table 4.4). The plate was shaken of 30 s and centrifuged for 1 min at 2 000 rpm to ensure no air bubbles are present in the mixture.

4.4.4 QC sample preparation

Equal amounts of aliquots from all samples of a specific matrix were pooled. This generated the QC sample. For tissue matrices, lung, liver and hypothalamus, equal amounts of the supernatant were used. Multiple QC samples were prepared for quality control evaluation. Five (5) µL of the pooled QC sample was used for protein precipitation and all other steps were performed as described above.

4.4.5 QQC sample preparation

A spiked QC (QQC) sample was generated by spiking a known concentration of all metabolites of interest into a QC sample. An equal amount of pooled QC sample and standard mix was used to generate the QQC sample. After mixing, 5 µL of the QQC sample was used for protein precipitation and all other sample preparation steps were followed as described above.

4.5 Analytical analysis

A randomised run order of the samples was used for sample analysis and a Microsoft Excel randomisation equation was used to re-order the run sequence of the samples. QC samples were analysed at the beginning, middle and end of the analytical run or after every 5 samples. Five (5) QC samples were analysed at the beginning of analysis before any analyses of samples were performed, to ensure the system is calibrated. A QQC sample was also analysed prior to the start of sample analysis, ensuring accurate peak identification.

4.6 Data handling

A large amount of data was generated while using this method and care was taken with handling the complex data set. Since a targeted LC-MS/MS analytical platform was used during this study, the data handling was simplified. The data handling process is divided into data processing and data pre-treatment.

4.6.1 Data processing

Data processing was done using Analyst 1.6.3. A quantitation method was designed for peak identification, peak integration and IS normalisation. This includes chromatogram peak integration and normalisation to IS by generation of an area ratio value between the observed metabolite's peak area to the assigned spiked IS area. The data was then extracted to an Excel file.

4.6.2 Data pre-treatment

After data extraction, data pre-treatment and data clean-up were done before statistical analysis was performed on the data. Data clean-up was performed as describes in section 3.3.6.1 by excluding all metabolites that were below the detection limit (see table 3.4 for LOD values) and were below 50% presence in all samples. A quality control procedure was implemented to assess the quality of the analysis and to eliminate any metabolites that may be responsible for introducing bias into the study and distorting the data (see section 3.6). Therefore after the data processing of the QC samples, the metabolites that had an RSD value above 30% were excluded from the data set. A zero value replacement was performed and the data was then arranged in the format compatible with the Excel macro file that was used for statistical analysis.

4.7 Statistical analysis

As described in section 3.3.7 an Excel-based statistical analysis macro file was used. Univariate and multivariate statistical analysis were performed including the multi t-test with Bonferroni correction, graph analysis and PCA. An HCA analysis was also performed for the identification of patterns in the data set.

4.8 Biological relevancy

After statistical analysis, the metabolites identified as significant between experimental groups can be assessed in terms of biological relevancy. A metabolic profile can be established, and biomarkers can be identified. With regards to this study, a metabolic profile for a fibrotic lung animal model was established and is described in Chapter 5 in a full length article. Biomarker identification requires validation with large samples sizes and cross validation. This is an important step with the investigation of diseases such as IPF, but biomarker validation was out of the scope for this study.

4.9 Discussion

This chapter serves as a summary of the developed LC-MS/MS based targeted metabolic profiling method containing all details related to sample preparation, analytical parameters as well as data handling and statistical analysis, to provided adequate results and answer biological relevant questions. The description of the development, optimisation and quality assessment of the developed method are provided in Chapter 3. The LC-MS/MS based metabolic profiling method was used, as described in this chapter, to generate a metabolic profile for a fibrotic lung animal model. The results of the metabolic profiling of a fibrotic lung animal model are discussed in Chapter 5 in a full length article prepared for submission to the *Respiratory Medicine* journal.

CHAPTER 5: METABOLIC PROFILING OF A FIBROTIC LUNG ANIMAL MODEL

This chapter is presented in a full length text article that has been submitted to the *Respiratory Medicine* journal. In this chapter the letter of submission of manuscript is provided together with the manuscript written according to the guidelines specified by the journal. For comprehensible reading purposes the table and figures, referred to in this chapter has been enlarged and placed in chronological order together with the text. The author guidelines can be reviewed in appendix A of the dissertation or at:
https://www.elsevier.com/journals/respiratory-medicine/0954-6111/guide-for-authors

Title of manuscript:

Metabolic profiling of the C57BL/6J bleomycin induced lung fibrosis mouse model

Authors: Venter Maryke [a,b],
Grobler Anne Frederica [b],
Kästle Marc [a],
Sobotta Mirco Christoph [a]
Bretschneider Tom [a],

[a] Boehringer Ingelheim Pharma GmbH &Co. KG, Biberach an der Riß, Germany

[b] DST/NWU Preclinical Drug Development Platform, North-West University, Potchefstroom, South Africa

Letter of proof of submission

Gmail Maryke Venter <maryke100venter@gmail.com>

Thank you for your submission to Respiratory Medicine

Respiratory Medicine Editorial Office" <eesserver@eesmail.elsevier.com> Thu, Nov 9, 2017 at 12:13 PM
Reply-To: "Respiratory Medicine Editorial Office\"" <respiratorymedicine@elsevier.com>
To: maryke100venter@gmail.com

Dear Ms. venter,

Thank you for sending your manuscript Metabolic profiling of the C57BL/6J bleomycin induced lung fibrosis mouse model for consideration to Respiratory Medicine. Please accept this message as confirmation of your submission.

When should I expect to receive the Editor's decision?

For Respiratory Medicine, the average editorial time (in weeks) from submission to final decision is: 9.

What happens next?

Here are the steps that you can expect as your manuscript progresses through the editorial process in the Elsevier Editorial System (EES).

1. First, your manuscript will be assigned to an Editor and you will be sent a unique reference number that you can use to track it throughout the process. During this stage, the status in EES will be "With Editor".

2. If your manuscript matches the scope and satisfies the criteria of Respiratory Medicine, the Editor will identify and contact reviewers who are acknowledged experts in the field. Since peer-review is a voluntary service, it can take some time but please be assured that the Editor will regularly remind reviewers if they do not reply in a timely manner. During this stage, the status will appear as "Under Review".

Once the Editor has received the minimum number of expert reviews, the status will change to "Required Reviews Complete".

3. It is also possible that the Editor may decide that your manuscript does not meet the journal criteria or scope and that it should not be considered further. In this case, the Editor will immediately notify you that the manuscript has been rejected and may recommend a more suitable journal.

For a more detailed description of the editorial process, please see Paper Lifecycle from Submission to Publication: http://help.elsevier.com/app/answers/detail/a_id/160/p/8045/

How can I track the progress of my submission?
You can track the status of your submission at any time at http://ees.elsevier.com/YRMED

Once there, simply:
1. Enter your username: Your username is: maryke100venter@gmail.com

If you need to retrieve password details, please go to: http://ees.elsevier.com/YRMED/automail_query.asp

2. Click on [Author Login]. This will take you to the Author Main Menu
3. Click on [Submissions Being Processed]

Many thanks again for your interest in Respiratory Medicine.

Kind regards,

Professor J Virchow

If you require further assistance, you are welcome to contact our Researcher Support team 24/7 by live chat and email or 24/5 by phone: http://support.elsevier.com

Metabolic profiling of the C57BL/6J bleomycin induced lung fibrosis mouse model

Venter Maryke [a,b], Grobler Anne Frederica [b], Kästle Marc [a], Sobotta Mirco Christoph[a], Bretschneider Tom [a]

[a] *Boehringer Ingelheim Pharma GmbH &Co. KG, Biberach an der Riß, Germany*
[b] *DST/NWU Preclinical Drug Development Platform, North-West University, Potchefstroom, South Africa*

ABSTRACT

Background: IPF is characterized by progressive destruction of normal lung architecture. The disease has a prevalence of between 1.25 and 23.4 per 100 000 population in Europe and 1 in every 32 000 population in South Africa. The underlying pathogenesis of IPF still remains poorly understood and alternative investigations for this disease is needed. Metabolic profiling has provided valuable information about the pathogenesis as well as treatment options for respiratory diseases such as asthma, COPD and cystic fibrosis. The establishment of a metabolic profiling method would be greatly beneficial in the characterisation of diseases, such as IPF.

Methods: During this study an LC-MS/MS based targeted metabolic profiling method was developed using the Luna NH_2 column (2 mm x 150 mm, 5 µm, 100 Å) operating with solvents at pH 9. A standardised sample preparation procedure was developed for various matrices.

Results: A metabolic profile was generated for a bleomycin treated mouse model resembling IPF. After statistical analysis, 26 metabolites were identified as significant (p-value < 0.05). These metabolites were crosschecked against an inflammation inducing lipopolysaccharide treated mouse model. In addition, a metabolic profile was also generated for a TGF-β stimulated fibroblasts cellular model.

Conclusion: Metabolites such as lactic acid, inosine, hypoxanthine, proline and intermediates from the methionine cycle were identified as significant in both the bleomycin treated model and the cellular model. A substantial correlation between the bleomycin treated mouse model and the TGF-β treated NHLF cellular model is present as well as to literature. These metabolites can be potential biomarkers for lung fibrosis.

Keywords: Metabolomics; targeted metabolic profiling; biomarkers; LC-MS/MS; lung fibrosis; IPF

1. Introduction

Metabolomics is a growing field and a valuable instrument for the identification of dysregulation in the metabolome of a biological system [1-4]. Metabolomics provides a vast range of applications and have been used to determine the cause and pathogenesis of complex diseases [4], as well as distinguishing between diseases showing similar clinical presentations [5]. Within metabolomics based studies there are different approaches that can be followed, including targeted metabolomics, untargeted metabolomics, metabolic footprinting, metabolic fingerprinting, fluxomics, lipidomics, metallomics and exposomics [6,7]. Various analytical platforms are available for the execution of the different approaches.

Metabolic profiling, a metabolomics based approach, have been used in the characterisation of complex diseases of which the onset and progression is still unknown. The use of a metabolic profiling approach to characterise diseases could potentially provide new insight into the pathogenesis of the disease and provide new therapeutic approaches [4,8,9].

Idiopathic Pulmonary Fibrosis (IPF) is a complex lung disease characterised by progressive destruction of normal lung architecture and accumulation of extracellular matrix that stiffens the lung and leads to respiratory failure [3,10,11]. IPF has a high prevalence of between 1.25 and 23.4 per 100 000 population in Europe, between 42.7 and 63 per 100 000 population in America and 1 in every 32 000 population in South Africa [12,13]. The survival duration from time of diagnosis for IPF patients are 2 to 3 year after diagnosis [4]. The current treatments that are available, Nintedanib (Boehringer Ingelheim Pharma GmbH & Co. KG, Germany) and Pirfenidone (Genentech Inc. member of the Roche Group, South San Francisco, CA, USA), are effective by reducing the decline in lung function but neither of these treatments offer full recovery [8,10,14]. There is a growing need for further research into the pathogenesis of this disease. Since metabolic profiling has provided beneficial insight into the pathogenesis of respiratory diseases such as asthma, COPD and cystic fibrosis [3], the establishment of the metabolic profile for IPF can provide essential insight into the onset and progression of the disease.

The metabolic profiling approach is a promising diagnostic tool to be used but there are still shortcomings that have to be addressed. There is still no standardised totally comprehensive approach available to detect and quantify large number of metabolites and there is also no standardised sample preparation and metabolite extraction method have been established [1,2,15]. Establishing a targeted metabolic profiling method is a feasible approach to metabolomics and allows investigation into the metabolome with high specificity [16].

The establishment of a targeted metabolic profiling method requires a well-designed method. Several factors have to be taken into account, such as metabolite identification, selection of the analytical platform, sample preparation, analytical analysis, data analysis, statistical analysis and biological relevance [15,16,17]. Since the metabolome of a biological system provides a direct correlation to the phenotype, characterising of the metabolome of a biological system could provide significant understanding of any disease state [6,18]. The selection of the analytical platform for metabolomics studies is important. Several analytical platforms are available, including nuclear magnetic resonance spectrometry (NMR), gas chromatography mass spectrometry (GC-MS) and liquid chromatography mass spectrometry (LC-MS), each with its own advantages and disadvantages [19]. The choice of an analytical platform influences the sample preparation procedure, an optimal sample preparation procedure should be compatible with the analytical platform [20]. Furthermore the analytical analysis and data handling is also important, since it can influence the quality of the results of the study [17].

During this study the aim was to develop an LC-MS/MS based targeted metabolic profiling method that would be able to generate a metabolic profile for any disease state, together with a sample preparation and metabolite extraction method suitable for various matrices. The method was validated by using the method for the generation of a metabolic profile for a fibrotic lung animal model.

2. Materials and methods

The targeted metabolic profiling method include metabolites, intermediates and cofactors present in the central carbon system, glycolysis, tricarboxylic acid (TCA) cycle, the serine biosynthesis pathway as well as all amino acids present in human metabolic pathways. A visual representation of the metabolites included in the method (highlighted in green) can be seen in Figure 1.

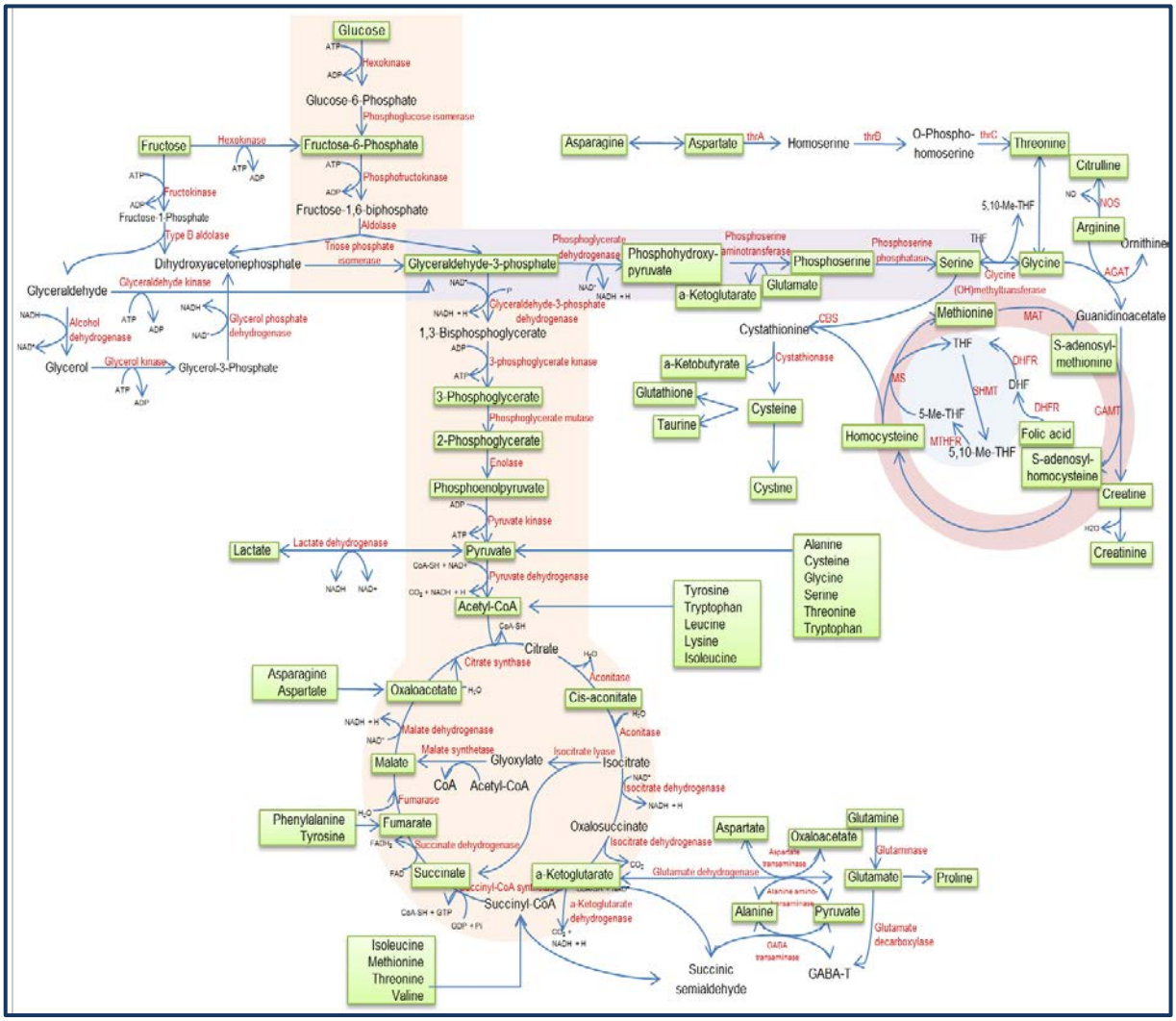

Figure 1. A visual representation of the metabolites in the developed method. All metabolites highlighted in green are compatible with the developed method. This diagram was generated from information obtained from the Kyoto Encyclopaedia of Genes and Genomes data base [25].

2.1 Reagents and standards

During this study the following high purity high-performance liquid chromatography (HPLC) graded solvents were used, dimethyl sulfoxide (DMSO) (CAS: 67-68-5), acetonitrile (ACN) (CAS: 75-05-8), methanol (MeOH) (CAS: 67-56-1), formic acid (CAS: 64-18-6), ammonium hydroxide (CAS: 1336-21-6) and ammonium acetate (CAS: 631-61-8). All solvents were purchased from Sigma-Aldrich Co. (Taufkirchen, Germany).

After a literature review was performed 100 metabolites and 19 Internal Standards (IS) were selected for targeted analysis [21-24]. The ISs were selected based on their retention time on the column and characteristics. Standards of these metabolites and ISs were purchased and a stock solution was prepared for each metabolite and IS, either in water or DMSO depending on the metabolite's solubility and characteristics. The stock solutions were used for identification, by generating a unique multiple reaction monitoring (MRM) transition for each metabolite as well as determining the retention time of each metabolite. The information of the metabolites analysed in this study is summarised in Table S1 and all information of the ISs used during this study are listed in Table S2. High purity standards (≥ 95%) for each metabolite and IS were purchased from either Sigma-Aldrich Co., eNovation Chemicals LLC (New Jersey, USA), TCI Deutschland GmbH (Eschborn, Germany), Cayman Chemical (Michigan, USA), Omicrom Biochemicals, Inc. (Indiana, USA) or Buchem Chemie & Technik GmbH & Co. (Köln, Germany).

2.2 Sample selection

To be able to fully understand the pathogenesis of IPF, human lung samples would have to be examined. This is an invasive approach with no benefit to the participant. An alternative to study the disease is to use animal models and cell cultures that closely resemble the disease state. Therefore during this study lung tissue from healthy and diseased C57BL/6J mouse models that resembles IPF was used. The diseased lung tissue used during the study was collected from storage and originated from C57BL/6J mice that were treated with bleomycin.

2.3 Ethical aspects

All tissue samples, healthy and diseased, from the bleomycin treated mouse study and the lipopolysaccharide induced lung inflammation study used during this study were from already euthanized animals and were collected from storage at the Drug Discovery Sciences Department of Boehringer Ingelheim, Germany. This study was ethically approved by the Regierungspräsidium in Tübingen, Germany (TVV 12-012). Additional ethical approval was obtained from the Ethics committee (AnimCare) of the North-West University (NWU-00275-17-A5).

2.4 Sample Preparation

A standardised sample preparation protocol, developed during this study, was used for the preparation of all samples. The sample preparation procedure is suitable for the following matrices: plasma, lung, liver, hypothalamus and cell culture. The sample preparation protocol consists of a homogenisation, protein precipitation and transfer step.

2.4.1 Homogenisation

Homogenisation was performed using the Precellys homogeniser (Bertin Corp. Maryland, USA). The lung samples were homogenised in a 1:4 (*w/v*) ratio with an ACN:MeOH 1:1 ratio mixture. After homogenisation the samples were centrifuged for 5 min at 13 000 rpm.

2.4.2 Protein Precipitation

Five (5) µL of the supernatant of the lung homogenate was added to 70 µL of ACN:MeOH 1:1 in a 96-well plate and placed in a -20 °C freezer for at least 15 min, to achieve optimal protein precipitation. The samples were then centrifuged for 4 min at 4 000 rpm.

2.4.3 Transfer

After centrifugation, 30 µL of supernatant was transferred to 170 µL ACN containing the different internal standards with appropriate concentrations (see Table S4). The plate was shaken for 30 s and centrifuged for 1 min at 2 000 rpm to ensure no air bubble is present in the mixture.

2.4.4 Quality control

Quality control (QC) samples were generated for the assessment of the sample preparation procedure, analytical method, instrumentation and system. A master QC sample was generated by pooling equal amounts of aliquots of the supernatant from the lung homogenates. The master QC sample was used to generate multiple QC samples by using 5 µL of the master QC sample and following all the steps described above.

A spiked QC sample (a QQC sample) was generated by spiking a known concentration of all metabolites of interest into a QC sample. An equal amount of master QC sample and standard mix was used to generate the QQC sample. After mixing, 5 µL of the QQC sample was used and all other sample preparation steps were followed as described above. To ensure that no bias is introduced into the results of the study a randomised run order of the samples was used for sample analysis and a Microsoft Excel randomisation equation was used to re-order the run sequence of the samples. QC samples were analysed at the beginning, middle and end of the analytical run. Five (5) QC samples were analysed at the start of analysis before any analyses of samples were performed, to ensure the system is calibrated and stable. A QQC sample was also analysed prior to the start of sample analysis, ensuring accurate peak identification.

2.5 LC Analysis

A modified version of the method described by Bajad *et al*. [26] was used. A 20 min HILIC method was developed using the Luna NH_2 column (2 mm x 150 mm, 5 µm, 100 Å). Solvent A consisted of 20 mM ammonium acetate and 20 mM ammonium hydroxide in 95:5 H_2O:ACN ratio (*v/v*) with pH 9; solvent B consisted of ACN. A 0.4 mL/min flowrate was used throughout the analysis. The LC gradient started with an isocratic gradient with 85% solvent B for 1.5 min, followed by a linier gradient to 2% solvent B within 12.5 min, a hold of 2% solvent B for 3 min, which was then followed by a 1.5 min gradient to starting conditions and was held constant for 1.5 min.

2.6 MS Parameters

The targeted metabolic profiling method was developed on a Triple Quad™ 6500^+ and QTRAP® system from AB Sciex Pte., Ltd. (Darmstadt, Germany). Together with the Triple Quad™, an Agilent 1290 Infinity II Multisampler and an Agilent 1290 Infinity II High Speed Pump system (Agilent Technologies Deutschland GmbH & Co., Waldbronn, Germany) was used to perform the analysis. The Triple Quad™ system was operated in the positive and negative switching mode with a source temperature of 500°C, ionspray voltage of 4500V, curtain gas of 40, gas 1 of 50 and gas 2 of 50. Transitions and MS parameters (DP: declustering potential and CE: collision energy) for all metabolites and ISs were determined and are listed in Table S3 and S4.

2.7 Data Processing

Data processing was done by using the Analyst 1.6.3 software (AB Sciex Pte., Ltd. Darmstadt, Germany). A quantitation method was designed for peak identification, peak integration and IS normalisation. This includes chromatogram peak integration and normalisation to IS by generating an area ratio value between the observed metabolite's peak area to the assigned spiked IS area. The data was then extracted to an Excel file.

2.8 Data Pre-treatment

After data extraction, data pre-treatment and clean-up was performed before statistical data analysis. Data clean-up was performed by excluding all metabolites that are below the detection limit and have a lower that 50% presence in all samples. The QC samples were assessed and all metabolites that have a RSD value above 30% were excluded. A zero value replacement was performed by replacing zero values with half of the limit of detection (LOD) for the metabolite. The data was then arranged in the format compatible with the Excel macro file that was used for statistical analysis.

2.9 Statistical Analysis

For the statistical analysis an Excel-based statistical analysis method, which is freely available as a macro file (developed by Dr Tsugawa, available at http://prime.psc.riken.jp), was used [27]. With this Excel-based method univariate and multivariate analyses were performed, including the multi t-test with Bonferroni correction, graph analysis and principal component analysis (PCA). A hierarchical cluster analysis (HCA) analysis was also performed for the identification of patterns in the data set.

3. Results and Discussion

3.1 C57BL/6J bleomycin treated mouse model

For the establishment of a metabolic profile for the C57BL/6J bleomycin treated mouse model, resembling IPF, metabolites have to be identified that differ in their abundance between healthy and diseased groups. Bleomycin (see Figure 1 for structure) is a chemotherapeutic antibiotic and causes an inflammatory and fibrotic reaction within a short period of time. An initial elevation of pro-inflammatory cytokines (interleukin-1, tumour necrosis factor-α, interleukin-6 and interferon-γ) is followed by an increased expression of pro-fibrotic markers (transforming growth factor-β1, fibronectin and procollagen-1) [28]. In the mouse model, the pro-fibrotic stage peaks at 14 days post exposure and the turnover between inflammation and fibrosis appears to occur around day 9 after bleomycin treatment [28].

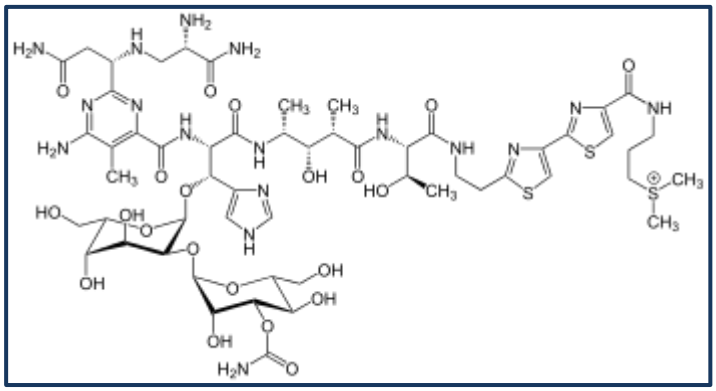

Figure 2. The structure of bleomycin. This figure was generated by using the MarvinSketch 14.12.15.0 software.

To generate a metabolic profile for the fibrotic lung animal model for this study, one lung lobe of 5 healthy C57BL/6J and 5 bleomycin treated C57BL/6J mice were available for analysis. Samples were obtained from 9-15 week old mice, treated with a 0.5 mg/kg bleomycin dosing (inhalation anaesthesia isoflurane (3-4%) bleomycin dissolved in 0.9% NaCl) by intratracheal application. The application volume was 2 mL/kg. Bleomycin treatment was administered for two consecutive days, the mice were euthanized 20 days post bleomycin treatment and the lungs were collected.

A metabolic profile was generated for the C57BL/6J bleomycin treated mouse model, resembling lung fibrosis, by identifying metabolites that demonstrated a difference in abundance between the healthy and diseased groups. After statistical analysis was performed, 26 metabolites were identified as significant (p-value < 0.05) between the healthy and treated group. Table 1 summarises the identified metabolites and contains the corresponding quality control, with regards to the QC RSD%, corresponding p-value generated by the t-test, average normalised values (area ratio between metabolite and corresponding IS), RSD% within each group and percentage increase/decrease. An RSD% of the QCs was generated for each metabolite to assess whether or not the metabolites were stable throughout the analysis. The metabolites that had a QC RSD value above 30% were identified as unstable and were excluded from the data set. A p-value was generated for each metabolite by using the Excel macro file. All the metabolites listed in Table 1 have a p-value below 0.05, indicating significant difference is present between the healthy and diseased groups. The average of the normalised area ratio of each metabolite to the assigned IS is provided as well as the RSD% within each group. The RSD% of each group demonstrates the variance within the groups.

During the quality control assessment of the data two metabolites, S-Adenosyl-L-homocysteine (SAH) and adenine, had a higher than 30% RSD within a group and some metabolites (e.g. threonine and valine) demonstrates high difference in variance between the two groups' RSD%. An explanation for this may be that differences occur due to uneven distribution of the bleomycin treatment, resulting in uneven development of fibrosis throughout the entire lung. Alternatively, differences in response to bleomycin treatment by the animals may occur, since not all animals respond to bleomycin treatment in the same manner [28]. This phenomenon is difficult to assess since not all the metabolites demonstrated the same phenomenon within the same range. Currently histological examination of lung tissue is the standard protocol for the assessment of the degree of fibrosis and the response to bleomycin [28]. Therefore the response to bleomycin is difficult to assess before metabolite analysis without sample loss. All positive values indicate the percentage increase and all negative values indicate the percentage decrease.

Table 1. Summary of statistical analysis of bleomycin treated C57BL/6J mice and healthy C56BL/6J mice.

Metabolite	QC RSD%	p-value	Control group		Bleomycin treated		% increase/ decrease
			Average	RSD%	Average	RSD%	
Phosphoenolpyruvic acid	16	0.000143	0.012	8	0.018	9	46
Glyoxylic acid	9	0.000658	0.241	7	0.346	12	43
Lactic acid	14	0.004602	3.682	29	7.442	25	102
Glycine	19	0.033382	7.156	10	5.972	12	-17
Acetylalanine	14	0.002854	0.163	16	0.420	23	157
Asparagine	6	0.000042	0.628	18	1.202	10	91
Glutamic acid	4	0.001563	29.360	4	37.540	10	28
Glutamine	15	0.000003	225.800	10	358.000	3	59
Isoleucine	4	0.000118	0.564	13	0.917	9	63
Leucine	8	0.000037	0.030	13	0.060	13	106
Phenylalanine	7	0.000254	0.767	11	1.221	12	59
Proline	8	0.000068	26.300	15	43.820	8	67
Threonine	25	0.000148	1.886	20	3.462	10	84
Tryptophan	9	0.000024	17.000	8	25.660	7	51
Tyrosine	10	0.000630	7.480	17	11.760	11	57
Valine	10	0.002892	90.020	16	122.600	8	36
Methionine	9	0.000098	11.554	14	18.960	9	64
SAH	14	0.000258	0.0161	17	0.006	37	-62
Creatinine	7	0.028479	0.243	7	0.195	19	-20
Adenine	8	0.010175	0.001	31	0.002	20	-60
Hypoxanthine	9	0.001497	22.500	14	32.960	11	46
Uracil	7	0.000020	1.358	8	2.730	12	101
Uridine	7	0.002965	0.006	12	0.012	20	118
Adenosine	5	0.016183	0.437	9	0.340	17	-22
Guanosine	8	0.000320	67.220	14	35.740	20	-47
Inosine	8	0.004120	140.200	13	167.000	14	19

The PCA score plot seen in Figure 3 was generated from the data presented in Table 1 and reveals that a substantial distinction is present between the control group (blue) and the bleomycin treated group (red) and therefore a substantial difference is present between the metabolic profiles of the healthy and bleomycin treated groups. This supports the observation provided by the univariate statistical analysis. The PCA score plot also reveals that a variation is present within the groups and the variation within the bleomycin treated group is much higher. This finding may indicate that the degree of fibrosis of the treated animals was uneven, but does not influence the substantial difference present between the two groups.

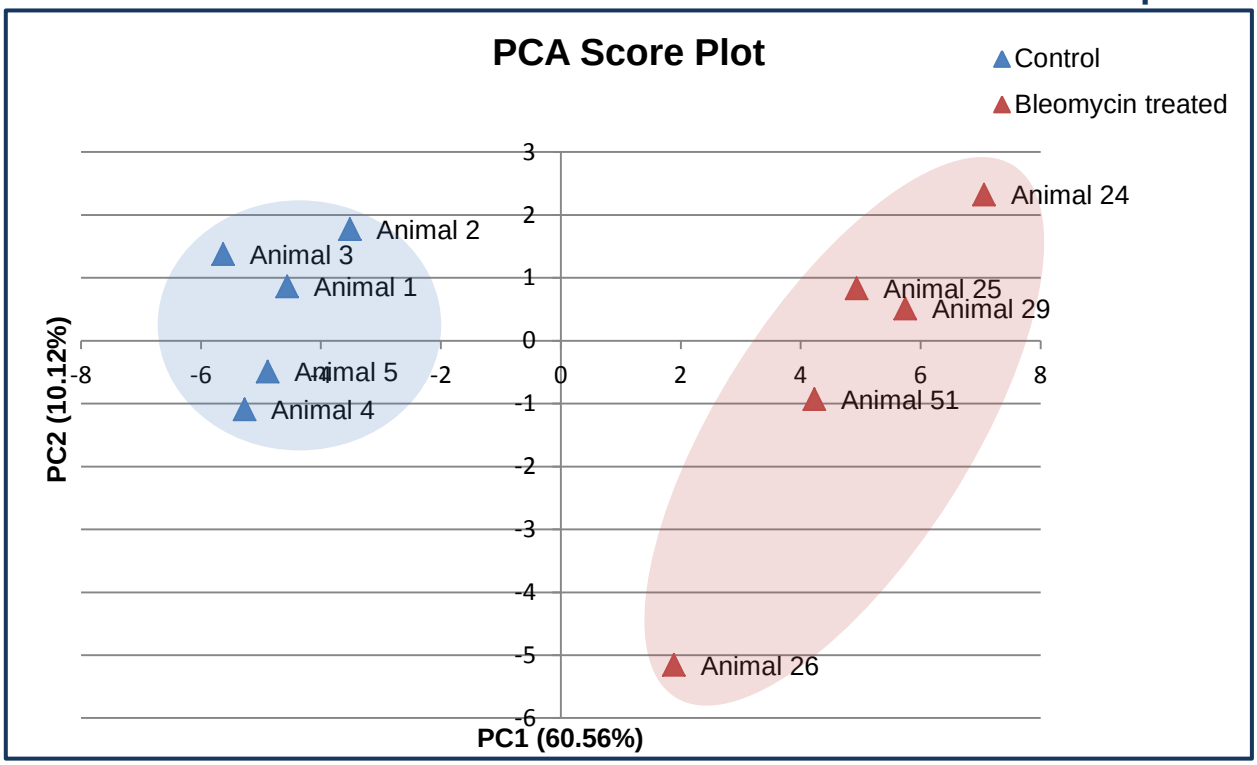

Figure 3. A principle component analysis score plot of the data of 5 control lung samples and 5 bleomycin treated lung samples. This figure was generated by using the Excel-based statistical analysis macro file as described in section 2.9.

An HCA was subsequently performed and the heat map can be seen in Figure 4. This heat map provides a simplified visualisation of all the data. Up and down regulation can be identified by the colour code. The HCA also allows the identification of outliers, with regards to metabolites as well as animals. An example of a metabolite that can be identified as an outlier is ketoleucine. Within the HCA heat map substantial variance can be identified among the different animals in a group. The QC RSD% of ketoleucine was higher that 30%, therefore this can indicate that the metabolite is not stable throughout the analysis. Animal outliers can also be identified. Using the HCA heat map, animal 26 in the bleomycin treated group can be considered as an outlier, since great variation in numerous metabolites are present when compared to the other animals of the treated group.

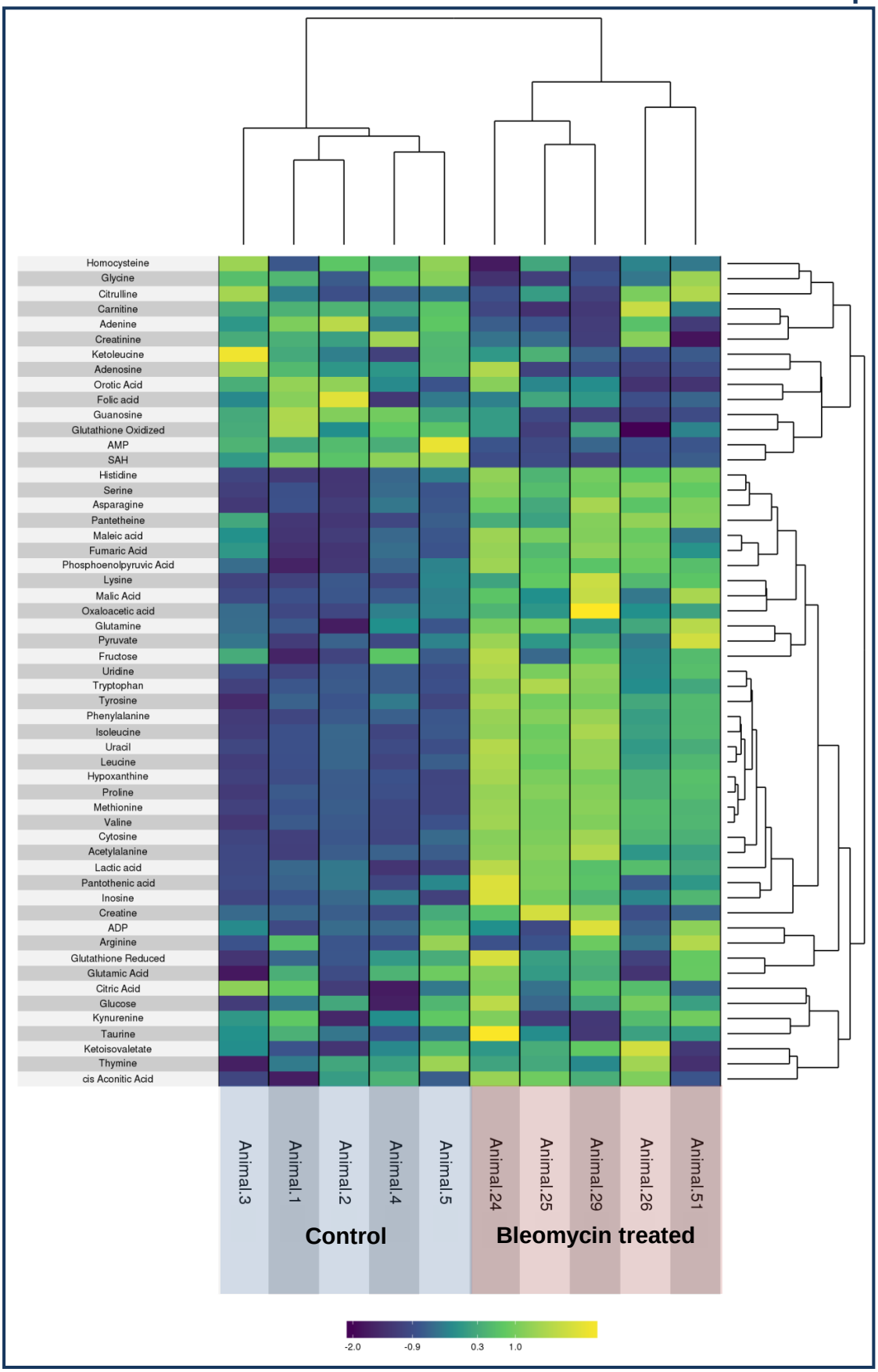

Figure 4. A hierarchical cluster analysis of the data of 5 control lung samples and 5 bleomycin treated lung samples. This figure was generated using the R 3.2.5 software.

With the identification of these 26 significant metabolites a metabolic profile for the fibrotic animal model can be generated, but since bleomycin treatment induces not only a fibrotic response but also an initial inflammatory response it is important to validate the metabolic profile. An approach for the validation is to compare the metabolic profile of the fibrotic animal model to an inflammation induced animal model. This will ensure that the identified metabolites are not a result of inflammation but a result of fibrosis. Therefore during this study healthy and bleomycin treated C57BL/6J mice were compared as well as the bleomycin treated C57BL/6J mice and lipopolysaccharide treated C57BL/6J mice.

3.2 LPS treated mouse model

Lipopolysaccharide (LPS)-induced lung inflammation is commonly used in acute lung injury studies, especially with regards to chronic obstructive pulmonary diseases and acute exacerbation [29]. LPS provokes the intrinsic immune response and initiates a cascade of inflammatory responses. Samples were available from C57BL/6J mice resulting from an acute exposure to LPS. These samples were used for validation purposes.

Firstly a metabolic profile was generated for the LPS treated mouse model by identifying metabolites that showed significant difference (p-value < 0.05) in abundance between the healthy and LPS treated group. The metabolic profile was generated by using the analytical method described above. One lung lobe each from 4 healthy C57BL/6J mice and 4 C57BL/6J LPS treated mice were available for analysis. After statistical analysis was performed on the generated data, 13 metabolites were identified as significant (see Table 2). Table 2 provides a summary of the 13 metabolites that were identified as significant with their corresponding QC RSD%, p-value, percentage increase/decrease as well as the RSD% within each group.

Table 2. Summary of the statistical analysis of the lung samples from healthy and LPS treated mice.

Metabolite	QC RSD%	p-value	Control group		LPS treated		% increase/ decrease
			Average	RSD%	Average	RSD%	
Glucose	3	0.041154	98.525	7	88.175	5	-11
Fructose 6-Phosphate	2	0.026789	0.077	29	0.031	73	-60
S-Adenosyl-L-homocysteine	15	0.003285	0.017	10	0.011	18	-35
Adenine	9	0.002849	0.006	5	0.004	15	-29
Guanine	15	0.000829	0.095	16	0.025	64	-73
Hypoxanthine	19	0.005220	6.738	12	2.578	68	-62
Thymine	6	0.015373	0.004	4	0.003	20	-26
Guanosine	15	0.000896	23.900	20	5.104	77	-79
Inosine	16	0.002461	129.500	9	45.625	69	-65
AMP	4	0.005856	0.190	28	0.436	24	131
GMP	10	0.010669	1.695	10	2.373	14	40
IMP	8	0.009252	1.330	26	2.868	26	116
UMP	11	0.002666	1.860	24	3.023	5	63

The PCA score plot analysis revealed that substantial variation exists between the healthy and LPS treated group. The PCA score plot also revealed that there is a difference present amongst the LPS treated group (see Figure 5). The variation present amongst the treated group is supported by the high RSD% of the metabolites within the treated group. Due to the small sample size an outlier analysis was not performed.

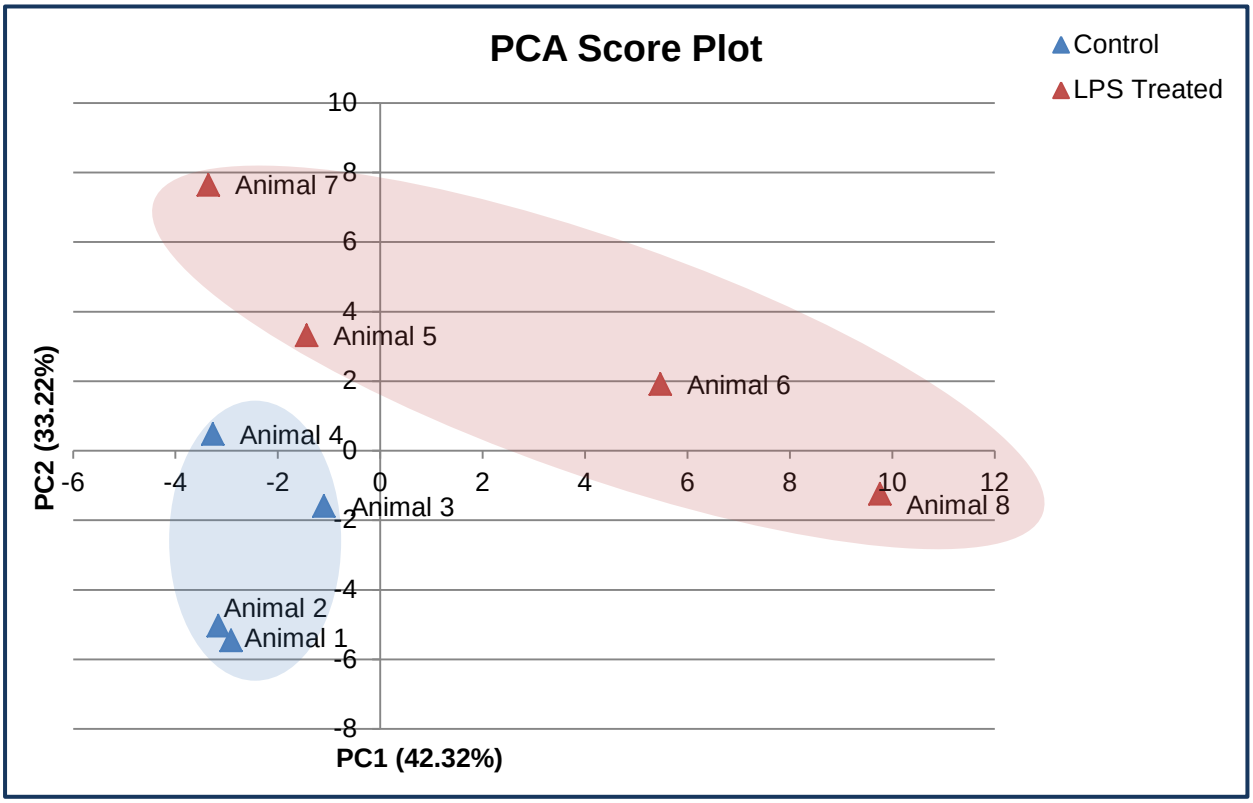

Figure 5. A principle component analysis score plot from the data of 4 control lung samples and 4 LPS treated lung samples. This figure was generated by using the Excel-based statistical analysis macro file described on section 2.9.

The 13 metabolites identified as significant in the C57BL/6J LPS treated mouse model were cross checked against the metabolites that were identified as significant in the C57BL/6J bleomycin treated mouse model. Five metabolites were identified as significant in both groups. Although these 5 metabolites were identified in both groups, the metabolites did not demonstrate the same trend with regards to percentage increase/decrease. For the evaluation of the trend, a 30% RSD margin was used, evaluating whether or not the metabolites demonstrated an increase/decrease within an RSD value of 30% (see Table 3).

Table 3. Summary of the 5 metabolites identified in both the C57BL/6J bleomycin treated mouse model and the LPS treated mouse model.

Metabolite	RSD% bleomycin and LPS	C57BL/6J bleomycin treated mouse model			LPS treated mouse model		
		QC RSD%	p-Value	% Change	QC RSD%	p-Value	% Change
S-Adenosyl-L-Homocysteine	39	14	0.000258	-62	15	0.003285	-35
Adenine	49	8	0.010175	-60	9	0.002849	-29
Hypoxanthine	968	9	0.001497	46	19	0.005220	-62
Guanosine	36	8	0.000320	-47	15	0.000896	-79
Inosine	259	8	0.004120	19	16	0.002461	-65

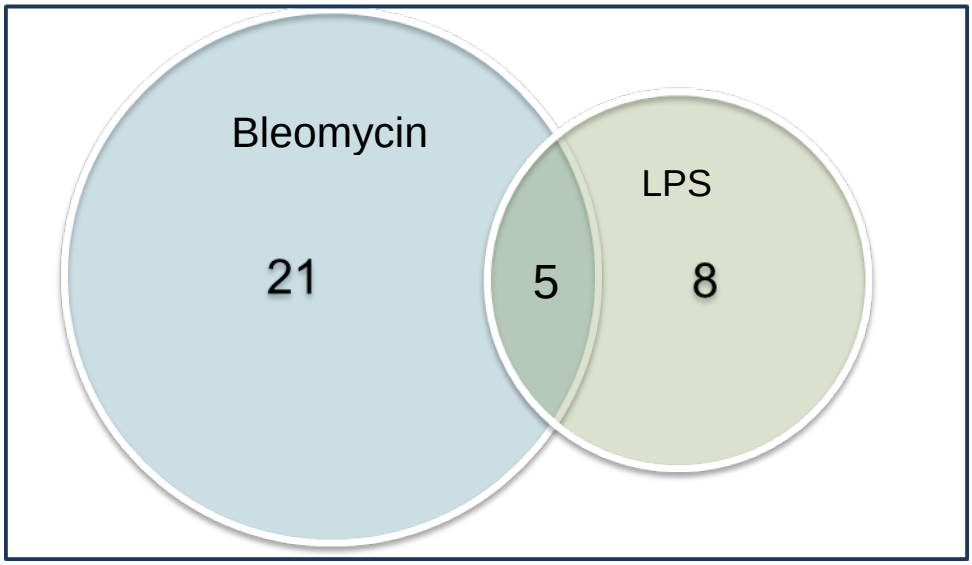

Figure 6. Venn diagram of the identified metabolites within the C57BL/6J bleomycin treated mouse model and the LPS treated mouse model.

The identified metabolites from the LPS treated group represent a metabolic profile for an inflammatory lung model. These metabolites either did not correspond to the C57BL/6J bleomycin treated mouse model's metabolic profile or did not demonstrate the same change. Therefore the metabolic profile of the C57BL/6J bleomycin treated mouse model can be assigned to a fibrotic metabolic profile.

Although the C57BL/6J bleomycin treated mouse model's metabolic profile can be assigned to a fibrotic state, it is uncertain whether or not this metabolic profile correlates with that of an IPF fibrotic state. To be able to validate that the metabolic profile of the C57BL/6J bleomycin treated mouse model correlates with the metabolic profile of IPF in human patients, biopsy samples from human participants would have to be assessed. This falls outside the scope of this study but another non-invasive investigation approach can be used. A transforming growth factor-beta (TGF-β) treated normal human lung fibroblast cellular model was used for this purpose.

3.3 TGF-β treated normal human lung fibroblasts

Since IPF is characterized by progressive destruction of normal lung architecture because of proliferation of activated fibroblasts and myofibroblasts [3,10,11], TGF-β stimulated normal human lung fibroblasts (NHLF) is frequently used in studies investigating IPF. TGF-β is a key cytokine responsible for the transformation of fibroblasts to myofibroblasts. Myofibroblasts are responsible for generating excess collagen and other extracellular matrix protein, leading to scar tissue formation in the lung [30].

As part as the validation of the metabolic profile of the C57BL/6J mouse model a metabolic profile was generated for a TGF-β treated NHLF cellular model. The metabolic profile of the TGF-β treated NHLF cellular model was compared to the metabolic profile generated for the C57BL/6J bleomycin treated mouse model. All cell samples were provided by the Immunology and Respiratory department at Boehringer Ingelheim, Germany.

NHLFs were obtained from healthy, non-smoking donors (Lonza Rockland Inc, Rockland, ME, USA). The NHLF were seeded in a 6-well plastic culture dish (Thermo Scientific, Massachusetts, USA) at a cell number of 200 000 cells per well. The cells were then grown in fibroblast growth medium (FGM-2, Lonza Rockland Inc, Rockland, ME, USA) and supplemented with insulin (0.5 ml), recombinant human fibroblast growth factor-B (rh-FGFB) (0.5 ml) and 2% foetal bovine serum (FBS) (10 ml), at 37°C in a humidified (5% CO_2) atmosphere. After 24 hours, half of the NHLF cells were stimulated with TGF-β (5 ng/mL) for 24 hours. The media was then removed and the cells were washed with phosphate buffered saline (PBS) (500 µL). The cells were scraped from the culture dish and pelleted by centrifugation for 5 min at 10 000 rpm. The PBS was discarded and the cell pellets were sonicated in ice cold ACN:MeOH 1:1

(500 µL) to disrupt all cell membranes. The supernatant was used for LC-MS/MS analysis using the method as described.

After analysis of the NHLF cell samples, 24 metabolites were identified as significant (p-value < 0.05). The summary of the identified metabolites together with the corresponding QC RSD%, p-value, percentage increase/decrease, average and RSD% of each group can be seen in Table 4. After the generation of the PCA score plot, two samples from the TGF-β group was identified as outliers, since these two samples present the same trend as the untreated samples (treated sample 5 and 6, see Figure 7). Therefore these samples were removed from the data set before further statistical analysis was performed.

Table 4. Summary of statistical analysis of untreated NHLF cells and TGF-β treated NHLF cells.

Metabolite	QC RSD%	p-value	Control group		TGF-β stimulated		% increase/ decrease
			Average	RSD%	Average	RSD%	
Glucose	13	0.026328	6.162	13	15.650	30	154
Fructose	15	0.025110	0.379	12	0.926	29	144
Lactic acid	3	0.014829	0.061	13	0.119	21	93
Arginine	11	0.044501	0.013	38	0.051	45	284
Asparagine	5	0.028454	0.011	11	0.022	25	95
Glutamic Acid	10	0.006794	0.080	8	0.155	15	93
Glutamine	12	0.012541	3.390	10	6.675	19	97
Histidine	18	0.033616	0.020	8	0.036	24	76
Isoleucine	10	0.032980	0.012	13	0.023	27	95
Lysine	24	0.012199	0.009	19	0.013	19	47
Phenylalanine	19	0.043239	0.014	14	0.028	30	99
Proline	11	0.015208	1.567	10	3.568	23	128
Tryptophan	12	0.014021	0.443	22	0.752	29	70
Tyrosine	21	0.004536	0.388	23	0.588	10	52
Valine	5	0.012023	2.492	7	5.280	20	112
Methionine	14	0.024844	0.258	12	0.640	29	148
Creatine	17	0.011033	0.008	13	0.021	23	158
Creatinine	14	0.023444	0.037	12	0.094	29	152
Uracil	14	0.000595	0.001	28	0.001	22	140
Uridine	12	0.000087	0.0002	20	0.0005	7	82
Adenosine	9	0.000076	0.006	16	0.011	13	95
Guanosine	24	0.006080	0.193	25	0.295	11	53
Inosine	22	0.045526	0.191	35	0.270	16	41
Carnitine	9	0.000367	1.358	19	2.345	11	73

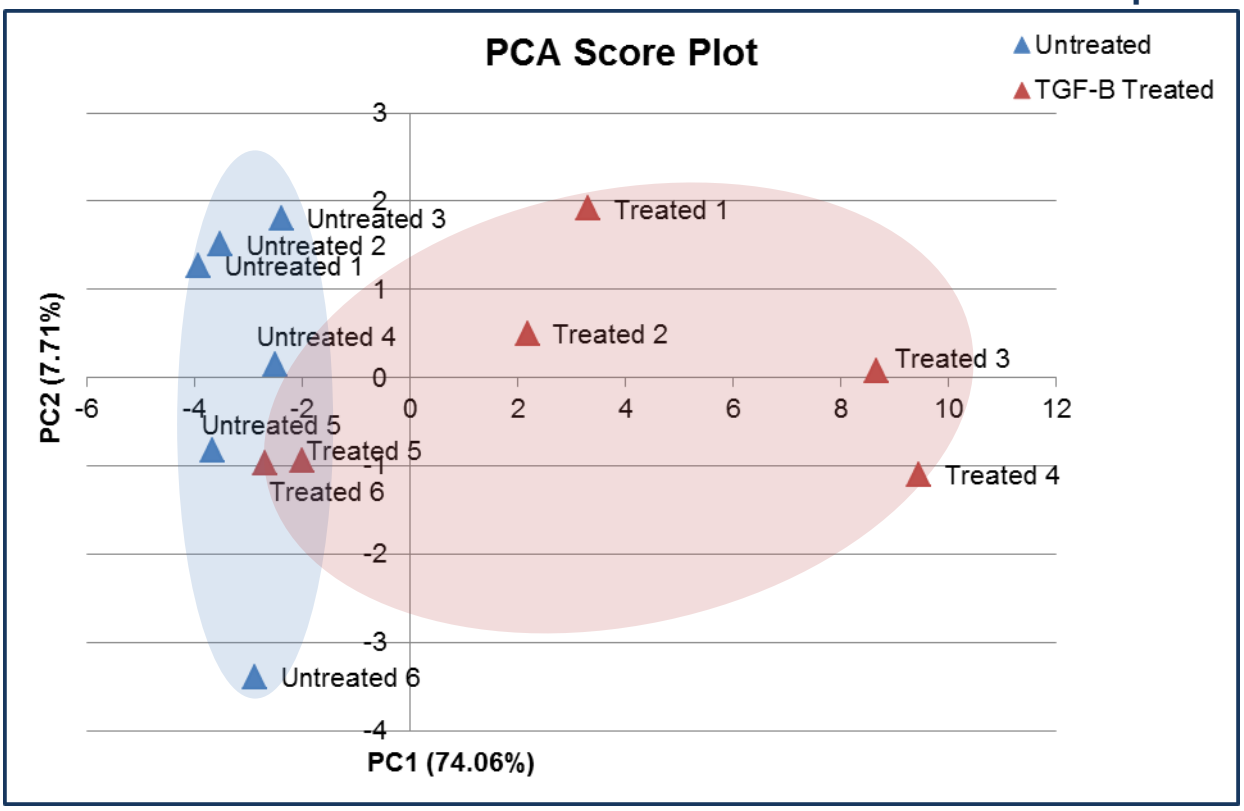

Figure 7. A principle component analysis score plot of the data of 6 control NHLF cell samples (untreated) and 6 TGF-β NHLF treated samples. Illustrating great variance present within each group (untreated and treated). Since two of the samples from the treated group show the same trend as the untreated group, these two samples were identified as outliers and were excluded from statistical analysis (treated sample 5 and 6). This figure was generated by using the Excel-based statistical analysis macro file described on section 2.9.

The 24 identified metabolites were cross checked against the identified metabolites from the C57BL/6J bleomycin treated mouse model metabolic profile. 17 metabolites were identified as significant in both the bleomycin treated mouse model and the NHLF TGF-β treated cell samples. A summary of these 17 metabolites can be seen in Table 5.5. Of the 17 metabolites only 7 metabolites (highlighted in Table 5) demonstrated the same trend with regards to percentage increase/decrease, below a 30% RSD margin.

Table 5. Summary of the 17 metabolites identified in both the C57BL/6J Bleomycin treated mouse model and the NHLF TGF-β treated cell samples.

Metabolite	RSD% Bleomycin and TGF-β	C57BL/6J Bleomycin treated mouse model			NHLF TGF-β treated cell samples		
		QC RSD%	p-Value	% Change	QC RSD%	p-Value	% Change
Lactic acid	**6**	**14**	**0.004602**	**102**	**3**	**0.014829**	**93**
Asparagine	**3**	**6**	**0.000042**	**91**	**5**	**0.028454**	**95**
Glutamic Acid	76	4	0.001563	28	10	0.006794	93
Glutamine	34	15	0.000003	59	12	0.012541	97
Isoleucine	**29**	**4**	**0.000118**	**63**	**10**	**0.032980**	**95**
Phenylalanine	36	7	0.000254	59	19	0.043239	99
Proline	44	8	0.000068	67	11	0.015208	128
Tryptophan	**22**	**9**	**0.000024**	**51**	**12**	**0.014021**	**70**
Tyrosine	**7**	**10**	**0.000630**	**57**	**21**	**0.004536**	**52**
Valine	73	10	0.002892	36	5	0.012023	112
Methionine	56	9	0.000098	64	14	0.024844	148
Creatinine	184	7	0.028479	-20	14	0.023444	152
Uracil	**23**	**7**	**0.000020**	**101**	**14**	**0.000595**	**140**
Uridine	**25**	**7**	**0.002965**	**118**	**12**	**0.000087**	**82**
Adenosine	227	5	0.016183	-22	9	0.000076	95
Guanosine	2520	8	0.000320	-47	24	0.006080	53
Inosine	52	8	0.004120	19	22	0.045526	41

* The 7 metabolites demonstrating the same trend for percentage change (increase or decrease) between the C57BL/6J bleomycin treated mouse model and the TGF-β treated cell samples are indicated in bold.

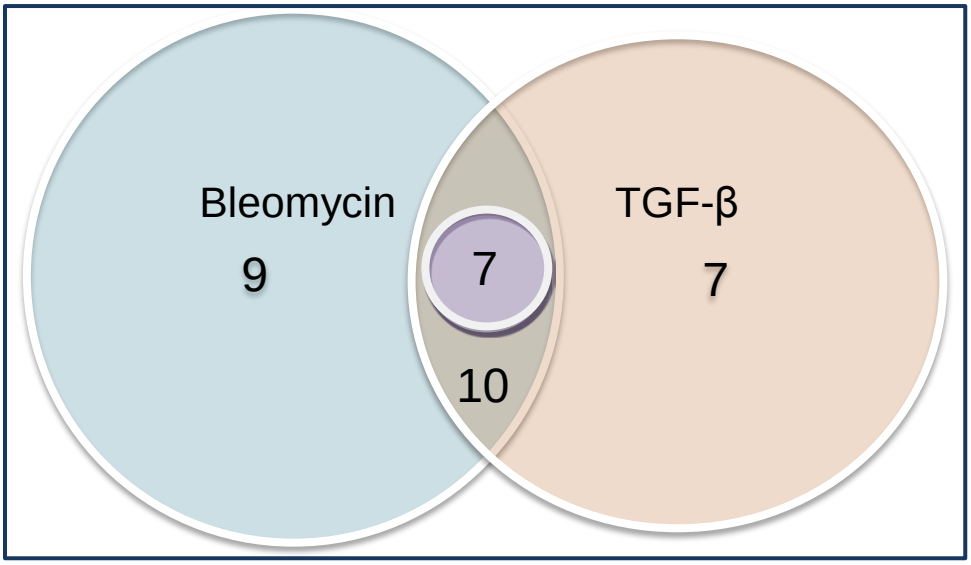

Figure 8. Venn diagram of the identified metabolites within the bleomycin induced fibrotic lung mouse model and the TGF-β treated NHLF cell samples. Twenty-six (26) metabolites were identified as significant in the C57BL/6J bleomycin treated mouse model and 24 metabolites were identified as significant in the NHLF TGF-β treated cell samples. 17 metabolites overlaps within these two groups and 7 metabolites demonstrate the same trend for percentage increase/decrease.

The seven metabolites that demonstrate the same trend for percentage increase/decrease between the metabolic profiles of NHLF TGF-β treated cellular cultures and the C57BL/6J bleomycin treated mouse model are lactic acid, asparagine, isoleucine, tryptophan, tyrosine, uracil and uridine, as shown in Figure 9.

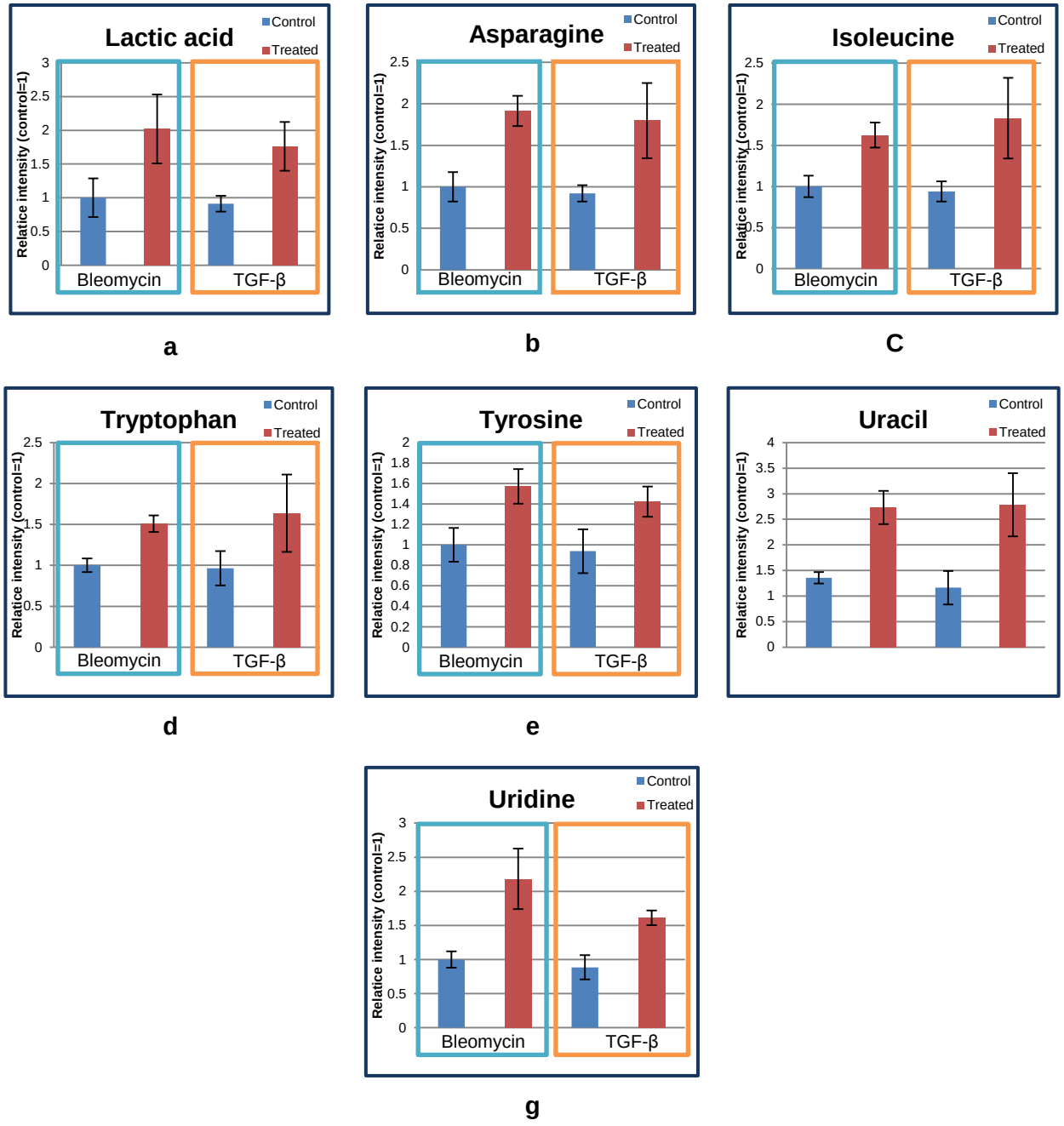

Figure 9. The seven metabolites identified as significant in both the C57BL/6J bleomycin treated mouse model and the NHLF TGF-β treated cellular model. The following metabolites a) lactic acid, b) asparagine, c) isoleucine, d) tryptophan, e) tyrosine, f) uracil and g) uridine have been identified as significant in the C57BL/6J bleomycin treated mouse model as well as in the NHLF TGF-β treated cellular model's metabolic profiles. These 7 metabolites demonstrate the same trend with regards to percentage increase between these two groups.

Lactic acid has previously been identified as a metabolite of interest after the analysis of human lung tissue from IPF patients [3,4]. Kottmann et al. [4] identified a connection between lactic acid levels and myofibroblast differentiation. Increased concentration of lactic acid was detected in a TGF-β treated fibroblast study and lactic acid is associated with TGF-β activation. Lactic acid is responsible for lowering the pH of the intracellular environment, generating favourable conditions for TGF-β to differentiate to myofibroblasts from fibroblasts [3,4]. This type of metabolic dysregulation has been identified in cancer cells and is known as the Warburg effect, where a high energy production demand result in the up-regulation of the glycolysis pathway and pyruvate is diverted to lactic acid production instead of entering the TCA cycle [21].

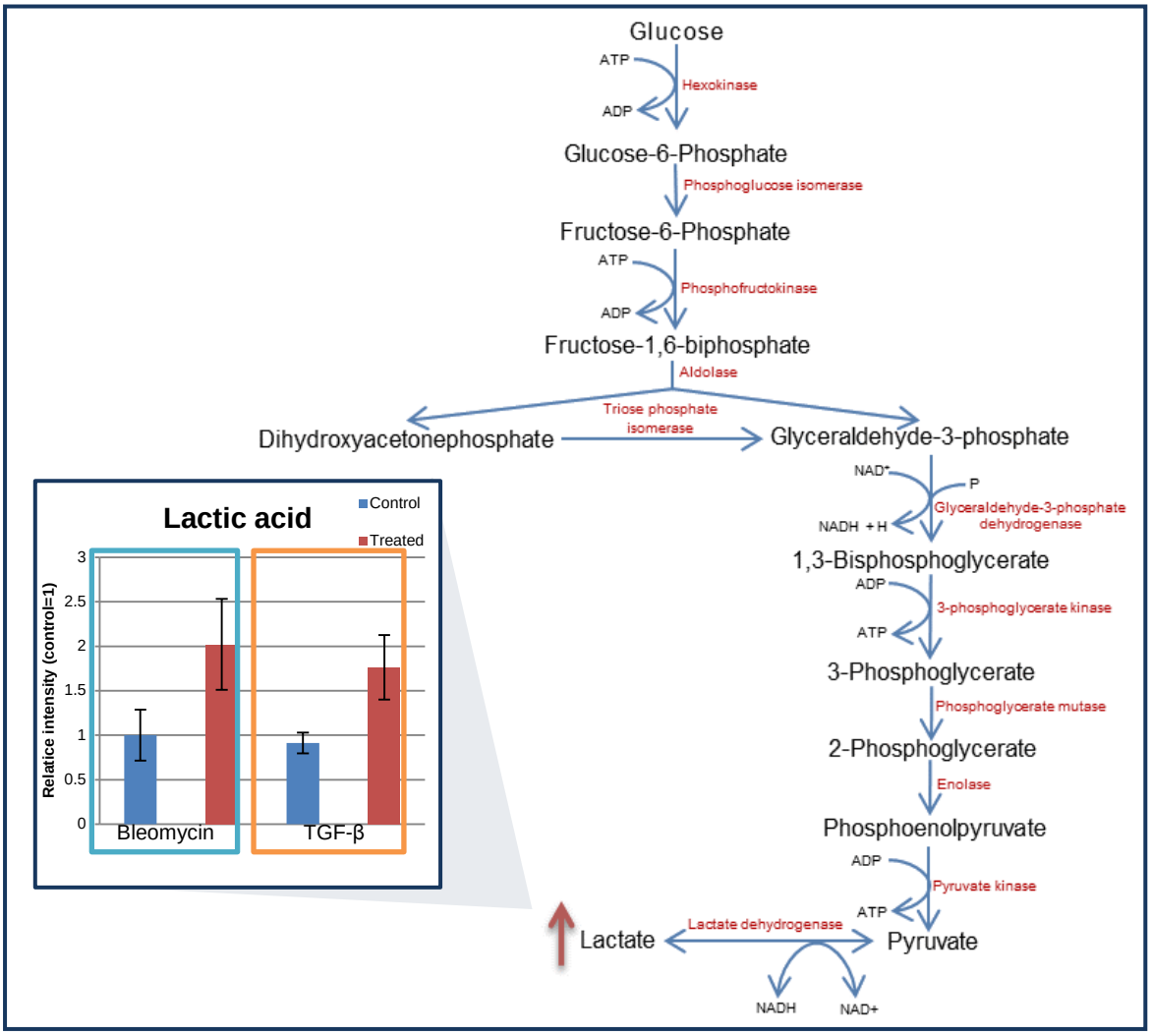

Figure 10. Up-regulation identified in the C57BL/6B bleomycin treated mouse model and TGF-β treated NHLF cellular model. This diagram was generated by using information obtained from the Kyoto Encyclopaedia of Genes and Genomes data base [25].

Kang *et al.* [3] identified elevation of other metabolites in IPF human lung tissue, including ATP, inosine and hypoxanthine. During this study the detection of ATP was below the LOD. Inosine was one of the metabolites identified as significant in both the C57BL/6J bleomycin treated mouse model and the NHLF TGF-β treated cellular model but demonstrated a trend of percentage increase above the 30% RSD margin. Hypoxanthine was only identified as a metabolite of significant in the C57BL/6J bleomycin treated mouse model. The elevation of inosine and hypoxanthine has been suggested to be linked to intracellular ATP depletion since inosine and hypoxanthine are breakdown products of ATP (see Figure 4) [3]. The depletion of intracellular ATP serves as an indication for oxidative stress and hypoxia, resulting in mitochondrial damage and induction of apoptosis pathways.

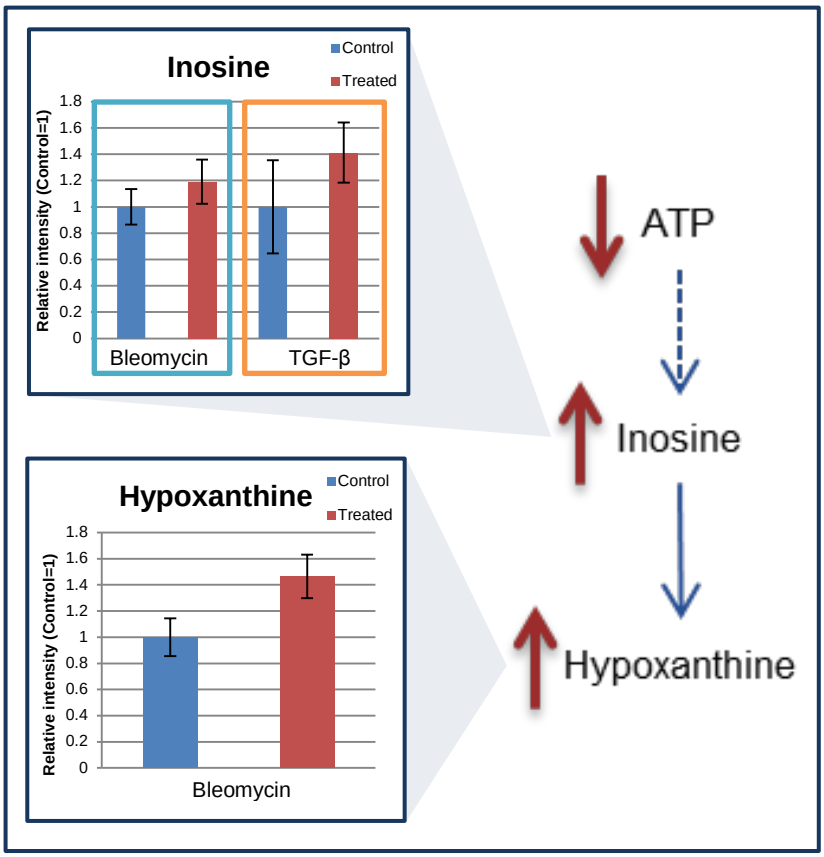

Figure 11. Up-regulation of inosine and hypoxanthine. This representation of ATP depletion and increase of inosine and hypoxanthine is similar to the results demonstrated by Kang et al. [3].

Glycine, methionine and SAH, important intermediates in the one carbon metabolism pathway and methionine cycle [31,32], have also been identified as significant in the C57BL/6J bleomycin treated mouse model. The identification of these metabolites as possible biomarkers may indicate that a dysregulation in the methionine cycle is present, which in turn can result in dysregulation of methylation of proteins, DNA, RNA and lipids and consequently in dysregulated cell proliferation [31,33].

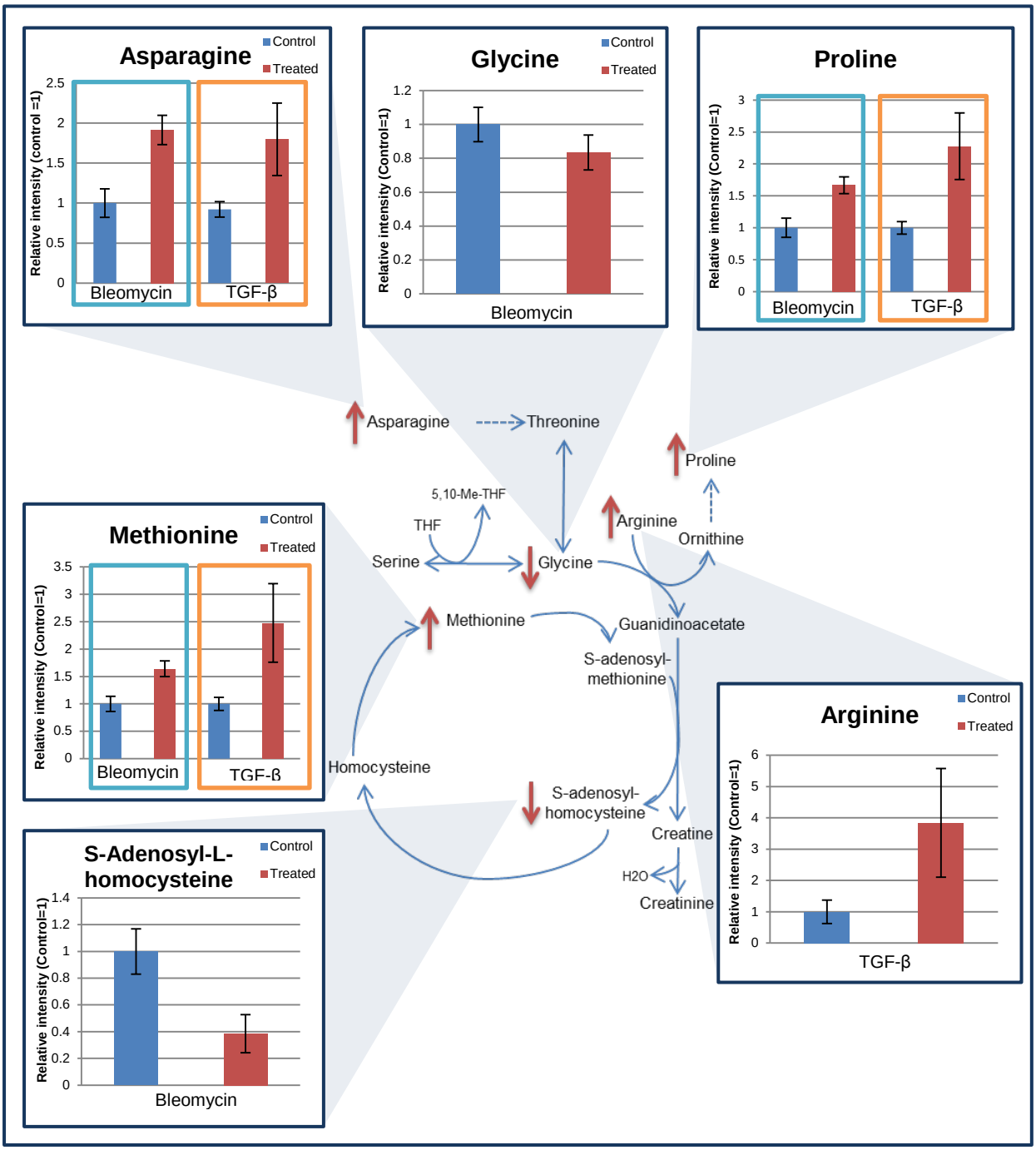

Figure 12. Dysregulation of asparagine, glycine, proline, arginine, methionine and S-adenosyl-L-homocysteine. This diagram was generated by using information obtained from the Kyoto Encyclopaedia of Genes and Genomes data base [25].

Proline have been identified as significant in both the C57BL/6J bleomycin treated mouse model and the TGF-β NHLF cellular model. Proline was identified by Kang et al. [3] as a possible biomarker for IPF after lung samples from IPF patients were analysed. Collagen fibrils are the most abundant protein in extracellular matrix and excess collagen deposit is associated with the progression of IPF. Collagen synthesis is a proline-dependant process, with proline being a major amino acid component for collagen. The total hydroxyproline content have previously been used to assess the state of lung fibrosis, but an acid hydrolysis of the lung tissue is required. Proline can be considered a biomarker of lung fibrosis since an increase in proline concentration would suggest an up-regulation in collagen synthesis. The assessment of the total hydroxyproline content in the lung tissue was outside of the scope of this study.

The following amino acids were also identified as significant in the C57BL/6J bleomycin treated mouse model: asparagine, glutamic acid, glutamine, isoleucine, leucine, phenylalanine, tryptophan, tyrosine and valine. Elevated levels of amino acid in lung tissue have been associated with dysregulated cell proliferation [34], since proliferating cells require amino acids for both protein synthesis and as a nitrogen source for non-essential amino acid synthesis. Elevated levels of glutamine have been identified as a major free amino acid substrate for nitrogen and carbon conception [34].

4. Conclusion

Several metabolites were identified as significant in the C57BL/6J bleomycin treated mouse model. After comparing the metabolic profile of the C57BL/6J to the metabolic profile of the LPS treated mouse model, it can be concluded that the C57BL/6J bleomycin treated mouse model's metabolic profile resembles a fibrotic state. The comparison of the C57BL/6J bleomycin treated mouse model's metabolic profile with the metabolic profile of the TGF-β NHLF cellular model revealed substantial correlation. Substantial correlation is also present between the C57BL/6J bleomycin treated mouse model's metabolic profile and recent clinical findings, especially to that of Kang *et al.* [3]. The identified metabolites suggest a dysregulation in the glycolysis pathway as well as the methionine cycle. Since IPF is such a complex disease and has been shown to be influenced by genetic and environmental factors [32], the key to understanding the etiology of IPF might lie at the epigenetic level of lung fibrosis. Epigenetic modifications subjugate all aspects of the central dogma and a dysregulation at the epigenetic level can explain the rapid progression of lung fibrosis in IPF, since a dysregulation of epigenetic modifications will initiate a cascade of dysregulation on all levels of the central dogma.

To our knowledge, this study is a first with regards to metabolic profiling of the C57BL/6J bleomycin treated mouse model for fibrotic lung inducement, to date no comparison has been made between the metabolic profiles of the C57BL/6J bleomycin mouse model and a cellular model resembling IPF or to human IPF lung tissue. Therefore we hypothesise that the

C57BL/6J bleomycin treated mouse model for lung inflammation induction is a suitable animal model for the investigation of lung fibrosis and that the key to understanding lung fibrosis lies at the epigenetic level and the dysregulation of epigenetic modifications. This hypothesis is supported by a recent study done by Guiot *et al*. [35], who investigated circulating nucleosomes and observed a significantly reduced level of cell free nucleosome associated methylated DNA (5-methylcytosine) in IPF patients compared to healthy subjects, indicating that a dysregulation of epigenetic modification is present [35].

Declaration of Conflicting Interests

The authors have no conflict of interests to disclose with regards to the research, authorship or publication of this article.

Funding

Funding for the study was provided by Boehringer Ingelheim Pharma GmbH &Co. KG and all other expenses was covered by the DST/NWU Preclinical Drug Development Platform of the North-West University.

References

[1] Álvarez-Sánchez, B., Priego-Capote, F. & de Castro, M.L. 2010. Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation. TrAC Trends in Analytical Chemistry, 29 (2):111-119.

[2] Bino, R.J., Hall, R.D., Fiehn, O., Kopka, J., Saito, K., Draper, J*., et al.* 2004. Potential of metabolomics as a functional genomics tool. Trends in plant science, 9 (9):418-425.

[3] Kang, Y.P., Lee, S.B., Lee, J.m., Kim, H.M., Hong, J.Y., Lee, W.J*., et al*. 2016. Metabolic profiling regarding pathogenesis of Idiopathic Pulmonary Fibrosis. Journal of proteome research:1717-1724.;

[4] Kottmann, R.M., Kulkarni, A.A., Smolnycki, K.A., Lyda, E., Dahanayake, T., Salibi, R., *et al*. 2012. Lactic acid is elevated in idiopathic pulmonary fibrosis and induces myofibroblast differentiation via pH-dependent activation of transforming growth factor-β. American journal of respiratory and critical care medicine:740-751.

[5] Adamko, D.J., Nair, P., Mayers, I., Tsuyuki, R.T., Regush, S. & Rowe, B.H. 2015. Metabolomic profiling of asthma and chronic obstructive pulmonary disease: A pilot study differentiating diseases. Journal of Allergy and Clinical Immunology, 136 (3):571-580. e573.

[6] Patti, G.J., Yanes, O. & Siuzdak, G. 2012. Innovation: Metabolomics: the apogee of the omics trilogy. Nature reviews Molecular cell biology, 13 (4):263-269.

[7] Wishart, D.S. 2016. Emerging applications of metabolomics in drug discovery and precision medicine. Nature Reviews Drug Discovery:473-484.

[8] Costabel, U., Bendstrup, E., Cottin, V., Dewint, P., Egan, J.J., Ferguson, J., *et al.* 2014. Pirfenidone in idiopathic pulmonary fibrosis: expert panel discussion on the management of drug-related adverse events. Advances in therapy, 31 (4):375-391.

[9] LaBrecque, D.R., Abbas, Z., Anania, F., Ferenci, P., Khan, A.G., Goh, K.-L., *et al*. 2014. World Gastroenterology Organisation global guidelines: Nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Journal of clinical gastroenterology, 48 (6):467-473.

[10] Richeldi, L., du Bois, R.M., Raghu, G., Azuma, A., Brown, K.K., Costabel, U., *et al*. 2014. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. New England Journal of Medicine, 370 (22):2071-2082.

[11] Sandbo, N. 2014. Mechanisms of fibrosis in IPF. (In Idiopathic Pulmonary Fibrosis. Springer. p. 161-205).

[12] Masekela, R., Olorunju, S., Green, R.J. & Magidimisa, N. 2016. Lung function decline is accelerated in South Africans with cystic fibrosis. South African Family Practice, 58 (1):24-27.

[13] Nalysnyk, L., Cid-Ruzafa, J., Rotella, P. & Esser, D. 2012. Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. European Respiratory Review, 21 (126):355-361.

[14] King Jr, T.E., Bradford, W.Z., Castro-Bernardini, S., Fagan, E.A., Glaspole, I., Glassberg, M.K., *et al.* 2014. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. New England Journal of Medicine, 370 (22):2083-2092.

[15] Dettmer, K., Aronov, P.A. & Hammock, B.D. 2007. Mass spectrometry-based metabolomics. Mass spectrometry reviews, 26 (1):51-78.

[16] Dunn, W.B., Bailey, N.J. & Johnson, H.E. 2005. Measuring the metabolome: current analytical technologies. Analyst, 130 (5):606-625.

[17] Boccard, J., Veuthey, J.L. & Rudaz, S. 2010. Knowledge discovery in metabolomics: an overview of MS data handling. Journal of separation science, 33 (3):290-304.

[18] Zhang, Y. & Kaminski, N. 2012. Biomarkers in idiopathic pulmonary fibrosis. Current opinion in pulmonary medicine, 18 (5):441.

[19] Zhang, A., Sun, H., Wang, P., Han, Y. & Wang, X. 2012. Modern analytical techniques in metabolomics analysis. Analyst, 137 (2):293-300.

[20] Álvarez-Sánchez, B., Priego-Capote, F. & de Castro, M.L. 2010. Metabolomics analysis II. Preparation of biological samples prior to detection. TrAC Trends in Analytical Chemistry, 29 (2):120-127.

[21] Armitage, E.G. & Barbas, C. 2014. Metabolomics in cancer biomarker discovery: current trends and future perspectives. Journal of pharmaceutical and biomedical analysis, 87:1-11.

[22] DeNicola, G.M., Chen, P.-H., Mullarky, E., Sudderth, J.A., Hu, Z., Wu, D., *et al.* 2015. NRF2 regulates serine biosynthesis in non-small cell lung cancer. Nature genetics, 47 (12):1475-1481.

[23] Locasale, J.W. 2013. Serine, glycine and one-carbon units: cancer metabolism in full circle. Nature Reviews Cancer, 13 (8):572-583.

[24] Lu, Y. & Chen, C. 2017. Metabolomics: Bridging Chemistry and Biology in Drug Discovery and Development. Current Pharmacology Reports, 3 (1):16-25.

[25] Genome.jp. 2017. KEGG: Kyoto Encyclopedia of Genes and Genomes. [online] Available at: http://www.genome.jp/kegg/ [Accessed 18 May. 2017].

[26] Bajad, S.U., Lu, W., Kimball, E.H., Yuan, J., Peterson, C. & Rabinowitz, J.D. 2006. Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. Journal of chromatography A, 1125 (1):76-88.

[27] Putri, S.P. & Fukusaki, E. 2014. Mass spectrometry-based metabolomics: a practical guide. CRC Press.1-269.

[28] Moeller, A., Ask, K., Warburton, D., Gauldie, J. & Kolb, M. 2008. The bleomycin animal model: a useful tool to investigate treatment options for idiopathic pulmonary fibrosis? The international journal of biochemistry & cell biology, 40 (3):362-382.

[29] Håkansson, H.F., Smailagic, A., Brunmark, C., Miller-Larsson, A. & Lal, H. 2012. Altered lung function relates to inflammation in an acute LPS mouse model. Pulmonary pharmacology & therapeutics, 25 (5):399-406.

[30] Pegorier, S., Campbell, G.A., Kay, A.B. & Lloyd, C.M. 2010. Bone Morphogenetic Protein (BMP)-4 and BMP-7 regulate differentially Transforming Growth Factor (TGF)-β1 in normal human lung fibroblasts (NHLF). Respiratory research, 11 (1):85.

[31] Maddocks, O.D., Labuschagne, C.F., Adams, P.D. & Vousden, K.H. 2016. Serine metabolism supports the methionine cycle and DNA/RNA methylation through de novo ATP synthesis in cancer cells. Molecular cell, 61 (2):210-221.

[32] Yang, I.V. & Schwartz, D.A. 2015. Epigenetics of idiopathic pulmonary fibrosis. Translational Research, 165 (1):48-60.

[33] Melnyk, S., Fuchs, G.J., Schulz, E., Lopez, M., Kahler, S.G., Fussell, J.J., *et al.* 2012. Metabolic imbalance associated with methylation dysregulation and oxidative damage in children with autism. Journal of autism and developmental disorders, 42 (3):367-377.

[34] Mayers, J.R., Torrence, M.E., Danai, L.V., Papagiannakopoulos, T., Davidson, S.M., Bauer, M.R., *et al.* 2016. Tissue of origin dictates branched-chain amino acid metabolism in mutant Kras-driven cancers. Science, 353 (6304):1161-1165.

[35] Guiot, J., Struman, I., Chavez, V., Henket, M., Herzog, M., Scoubeau, K., Hardat, N., Bondue, B., Corhay, J.L., Moermans, C. and Louis, R., 2017. Altered epigenetic features in circulating nucleosomes in idiopathic pulmonary fibrosis. Clinical Epigenetics, 9(1), p.84.

Supplementary data

Table S1. Summary of all the metabolites of interest

Group	Metabolite	Formula	M_r	CAS	Company
Glycolysis	Glucose	$C_6H_{12}O_6$	180.063	50-99-7	Sigma
	Fructose-6-Phosphate	$C_6H_{13}O_9P$	260.029	26177-86-6	Sigma
	3-Phospho-D-Glycerate	$C_3H_7O_7P$	185.992	80731-10-8	Sigma
	2-Phosphoglyceric acid	$C_3H_7O_7P$	185.992	83418-48-8	Sigma
	Phosphoenolpyruvic acid	$C_3H_5O_6P$	167.982	4265-07-0	Sigma
	Pyruvate	$C_3H_4O_3$	88.016	113-24-6	Sigma
TCA	Oxaloacetic acid	$C_4H_4O_5$	132.006	328-42-7	Sigma
	Citric acid	$C_6H_8O_7$	192.027	77-92-9	Sigma
	cis-Aconitic acid	$C_6H_6O_6$	174.016	585-84-2	Sigma
	2-Ketoglutaric acid	$C_5H_6O_5$	146.021	328-50-7	Sigma
	Succinic acid	$C_4H_6O_4$	118.026	110-15-6	Sigma
	Fumaric acid	$C_4H_4O_4$	116.011	110-17-8	Sigma
	Malic acid	$C_4H_6O_5$	134.021	6915-15-7	Sigma
	Glyoxylic acid	$C_2H_2O_3$	74.000	563-96-2	Sigma
	Fructose	$C_6H_{12}O_6$	180.063	57-48-7	Sigma
	Lactic acid	$C_3H_6O_3$	90.031	50-21-5	Sigma
Serine bio-synthesis	3-Hydroxybutyrate	$C_4H_8O_3$	104.040	150-83-4	Sigma
	3-Phosphonooxypyruvate	$C_3H_5O_7P$	183.977	3913-50-6	eNC
	L-Serine-O-Phosphate	$C_3H_8NO_6P$	185.008	407-41-0	Sigma
	Serine	$C_3H_7NO_3$	105.042	56-45-1	Sigma
	Glycine	$C_2H_5NO_2$	75.032	56-40-6	Sigma
Amino acids	Acetylalanine	$C_5H_9NO_3$	131.058	97-69-8	Sigma
	Alanine	$C_3H_7NO_2$	89.0477	56-41-7	Sigma
	Arginine	$C_6H_{14}N_4O_2$	174.111	74-79-3	Sigma
	Asparagine	$C_4H_8N_2O_3$	132.053	70-47-3	Sigma
	Aspartic acid	$C_4H_7NO_4$	133.037	56-84-8	Sigma
	Citrulline	$C_6H_{13}N_3O_3$	175.095	372-75-8	Sigma
	Cysteine	$C_3H_7NO_2S$	121.019	52-90-4	Sigma
	Cystine	$C_6H_{12}N_2O_4S_2$	240.020	56-89-3	Sigma
	Glutamic acid	$C_5H_9NO_4$	147.053	56-86-0	Sigma
	Glutamine	$C_5H_{10}N_2O_3$	146.069	56-85-9	Sigma
	Glutathione Oxidized	$C_{20}H_{32}N_6O_{12}S_2$	612.152	27025-41-8	Sigma
	Glutathione Reduced	$C_{10}H_{17}N_3O_6S$	307.083	70-18-8	Sigma
	Histidine	$C_6H_9N_3O_2$	155.069	71-00-1	Sigma
	Hydroxyproline	$C_5H_9NO_3$	131.058	51-35-4	Sigma
	Isoleucine	$C_6H_{13}NO_2$	131.094	73-32-5	Sigma
	Leucine	$C_6H_{13}NO_2$	131.094	61-90-5	Sigma
	Lysine	$C_6H_{14}N_2O_2$	146.105	657-27-2	Sigma
	Phenylalanine	$C_9H_{11}NO_2$	165.079	63-91-2	Sigma
	Proline	$C_5H_9NO_2$	115.063	147-85-3	Sigma
	Taurine	$C_2H_7NO_3S$	125.014	107-35-7	Sigma
	Threonine	$C_4H_9NO_3$	119.058	72-19-5	Sigma
	Tryptophan	$C_{11}H_{12}N_2O_2$	204.089	73-22-3	Sigma
	Tyrosine	$C_9H_{11}NO_3$	181.073	60-18-4	Sigma
	Valine	$C_5H_{11}NO_2$	117.079	72-18-4	Sigma
Methionine cycle	Methionine	$C_5H_{11}NO_2S$	149.051	63-68-3	Sigma
	SAM	$C_{15}H_{23}N_6O_5S$	399.145	86867-01-8	Cayman
	SAH	$C_{14}H_{20}N_6O_5S$	384.120	979-92-0	Sigma
	Homocysteine	$C_4H_9NO_2S$	135.035	454-29-5	Sigma
	Creatine	$C_4H_9N_3O_2$	131.069	6020-87-7	TCI
	Creatinine	$C_4H_7N_3O$	113.059	60-27-5	TCI
	Folic acid	$C_{19}H_{19}N_7O_6$	441.139	59-30-3	Sigma
	Tetrahydrofolate	$C_{19}H_{23}N_7O_6$	445.430	135-16-0	Sigma

Table S1 (continued). Summary of all the metabolites of interest

Group	Metabolite	Formula	M_r	CAS	Company
Coenzyme	Acetyl-CoA	$C_{23}H_{38}N_7O_{17}P_3S$	809.120	102029-73-2	Sigma
	CoA	$C_{21}H_{36}N_7O_{16}P_3S$	767.115	85-61-0	Sigma
	FAD	$C_{27}H_{33}N_9O_{15}P_2$	785.157	84366-81-4	Sigma
	NAD	$C_{21}H_{27}N_7O_{14}P_2$	663.109	53-84-9	Sigma
	NADP	$C_{21}H_{28}N_7O_{17}P_3$	743.075	24292-60-2	Sigma
	NADH	$C_{21}H_{29}N_7O_{14}P_2$	665.441	606-68-8	Sigma
	NADPH	$C_{21}H_{30}N_7O_{17}P_3$	745.090	2646-71-1	Sigma
Nucleobase	Adenine	$C_5H_5N_5$	135.054	73-24-5	Sigma
	Guanine	$C_5H_5N_5O$	151.049	73-40-5	Sigma
	Hypoxanthine	$C_5H_4N_4O$	136.039	68-94-0	Sigma
	Thymine	$C_5H_6N_2O_2$	126.042	65-71-4	Sigma
	Uracil	$C_4H_4N_2O_2$	112.027	66-2-8	Sigma
	Cytosine	$C_4H_5N_3O$	111.000	71-30-7	Sigma
	Uridine	$C_9H_{12}N_2O_6$	244.069	58-96-8	Sigma
	Xanthine	$C_5H_4N_4O_2$	152.033	69-89-6	Sigma
Nucleoside	Adenosine	$C_{10}H_{13}N_5O_4$	267.097	58-61-7	Sigma
	Guanosine	$C_{10}H_{13}N_5O_5$	283.091	118-00-3	Sigma
	Inosine	$C_{10}H_{12}N_4O_5$	268.080	58-63-9	Sigma
Nucleotide	ADP	$C_{10}H_{15}N_5O_{10}P_2$	427.029	20398-34-9	Sigma
	GDP	$C_{10}H_{15}N_5O_{11}P_2$	443.024	43139-22-6	Sigma
	UDP	$C_9H_{14}N_2O_{12}P_2$	404.002	27821-45-0	Sigma
Nucleotide derivative	ADP-Glucose	$C_{16}H_{25}N_5O_{15}P_2$	589.082	102129-65-7	Sigma
	ADP-Ribose	$C_{15}H_{23}N_5O_{14}P_2$	559.071	68414-18-6	Sigma
Nucleotide	ATP	$C_{10}H_{16}N_5O_{13}P_3$	506.995	34369-07-8	Sigma
	GTP	$C_{10}H_{16}N_5O_{14}P_3$	522.990	36051-31-7	Sigma
	AMP	$C_{10}H_{14}N_5O_7P$	347.063	4578-31-8	Sigma
	CMP	$C_9H_{14}N_3O_8P$	323.051	6757-06-8	Sigma
	GMP	$C_{10}H_{14}N_5O_8P$	363.050	5550-12-9	Sigma
	IMP	$C_{10}H_{13}N_4O_8P$	348.047	352195-40-5	Sigma
	UMP	$C_9H_{13}N_2O_9P$	324.035	58-97-9	Sigma
	cAMP	$C_{10}H_{12}N_5O_6P$	329.052	60-92-4	Sigma
	cGMP	$C_{10}H_{12}N_5O_7P$	345.047	61093-23-0	Sigma
	dAMP	$C_{10}H_{14}N_5O_6P$	331.068	653-63-4	Sigma
	dCMP	$C_9H_{14}N_3O_7P$	307.056	1032-65-1	Sigma
	dTMP	$C_{10}H_{15}N_2O_8P$	322.056	3343-62-5	Sigma
Other	Carnitine	$C_7H_{15}NO_3$	161.105	6645-46-1	Sigma
	Orotic acid	$C_5H_4N_2O_4$	156.017	65-86-1	Sigma
	Salicylic acid	$C_7H_6O_3$	138.032	69-72-7	Sigma
	Ascorbic acid	$C_6H_8O_6$	176.032	50-81-7	Sigma
	Ketoisovalerate	$C_5H_8O_3$	116.047	3715-29-5	Sigma
	Kynurenine	$C_{10}H_{12}N_2O_3$	208.213	13441-51-5	Sigma
	3-OH-Kynurenine	$C_{10}H_{12}N_2O_4$	224.079	2147-61-7	Sigma
	Quinolinic acid	$C_7H_5NO_4$	167.021	89-00-9	Sigma
	Kynurenic acid	$C_{10}H_7NO_3$	189.042	492-27-3	Sigma
	Pantothenic acid	$C_9H_{17}NO_5$	219.110	137-08-6	Sigma
	Maleic acid	$C_4H_4O_4$	116.010	110-16-7	Sigma
	Ketoleucine	$C_6H_{10}O_3$	130.060	816-66-0	Sigma

Table S2. Summary of all the IS used

Group	Metabolite	M_r	CAS	Company	Solvent
Glycolysis	D-Glucose-13C6	186.11	110187-42-3	Sigma	Water
	D-Fructose-13C6	186.11	201595-65-5	Sigma	Water
Amino acids	L-Serine-13C3,15N,2,3,3-d3	112.08		Sigma	Water
	L-Glutamic acid-13C5	152.09		Sigma	Water
	Glycine-C13	76.06		Sigma	Water
	L-Isoleucine-13C6,15N	138.12		Sigma	Water
	L-Leucine-5,5,5-d3	134.19	87828-86-2	Sigma	Water
	L-Lysine-4,4,5,5-d4	186.67		Sigma	Water
	L-Phenyl-d5-alanine	170.22		Sigma	Water
	Thymine-d4 (methyl-d3,6-d1)	130.14	156054-85-2	Sigma	Water
	Tryptophan D5	209.26		Sigma	DMSO
Coenzyme	Acetyl-1,2-13C2 coenzyme A	811.56		Sigma	Water
Nucleobase	2-(Methyl-13C,d3-thio)adenine	185.23	1216721-76-4	Sigma	DMSO
Nucleoside	[1',2',3',4',5'-13C5]adenosine	272.21	159496-13-6	Omicron	Water
Nucleotide	Adenosine-15N5 5′-monophosphate	352.19		Sigma	Water
	Adenosine-15N5 5′-triphosphate	512.15		Sigma	Water
Other	Quinolinic acid-4,5,6-D3	170.14	138946-42-6	Buchem	DMSO
	D4-Kynurenine	212.24	194546-33-3	Buchem	DMSO
	Kynurenic acid-3,5,6,7,8-d5	194.20		Sigma	DMSO

Table S3. Summary of all the MS parameters for the different metabolites included in the method

Group	Metabolite	Formula	M_r	$[M+H]^+$ (*m/z*)	$[M+H]^-$ (*m/z*)	Product ion (*m/z*)	DP (V)	CE (V)	RT (min)	Internal Standard
Glycolysis	Glucose	$C_6H_{12}O_6$	180.063		178.957	89.000	-55	-12	4.38	C13-Glucose
	Fructose-6-Phosphate	$C_6H_{13}O_9P$	260.029	261.036		109.028	20	20	11.98	D3-Quinolinic acid POS
	3-Phospho-D-Glycerate	$C_3H_7O_7P$	185.992	187.000		140.900	20	15	14.40	D3-Quinolinic acid POS
	2-Phosphoglyceric acid	$C_3H_7O_7P$	185.992	187.001		98.984	20	40	15.20	D3-Quinolinic acid POS
	Phosphoenolpyruvic acid	$C_3H_5O_6P$	167.982		166.974	78.958	-20	-20	12.18	D5-Glutamic acid NEG
	Pyruvate	$C_3H_4O_3$	88.016		86.787	42.970	-25	-10	11.47	D5-Glutamic acid NEG
TCA	Oxaloacetic acid	$C_4H_4O_5$	132.006		131.000	86.975	-65	-16	11.44	D5-Glutamic acid NEG
	Citric acid	$C_6H_8O_7$	192.027		191.019	111.008	-20	-20	13.8	D3-Quinolinic acid NEG
	cis-Aconitic acid	$C_6H_6O_6$	174.016		173.009	85.029	-20	-20	11.46	D5-Glutamic acid NEG
	2-Ketoglutaric acid	$C_5H_6O_5$	146.021		145.014	100.900	-42	-11	11.97	D5-Glutamic acid NEG
	Succinic acid	$C_4H_6O_4$	118.026		117.019	73.029	-20	-40	12.00	D5-Glutamic acid NEG
	Fumaric acid	$C_4H_4O_4$	116.011		115.000	71.000	-20	-20	11.97	D5-Glutamic acid NEG
	Malic acid	$C_4H_6O_5$	134.021		133.013	71.014	-20	-40	11.91	D5-Glutamic acid NEG
	Glyoxylic acid	$C_2H_2O_3$	74.000		72.800	45.000	-30	-12	12.52	D5-Glutamic acid NEG
	Fructose	$C_6H_{12}O_6$	180.063		178.957	89.000	-55	-12	3.76	C13-Fructose
	Lactic acid	$C_3H_6O_3$	90.031		88.900	42.800	-50	-15	6.52	C13-Serine NEG
Serine bio-synthesis	3-Hydroxybutyrate	$C_4H_8O_3$	104.040		102.895	59.042	-10	-12	6.20	C13-Serine NEG
	3-Phosphonooxypyruvate	$C_3H_5O_7P$	183.977		182.900	78.869	-55	-40	13.90	D3-Quinolinic acid NEG
	L-Serine-O-Phosphate	$C_3H_8NO_6P$	185.008	186.000		88.200	20	20	12.11	D3-Quinolinic acid POS
	Serine	$C_3H_7NO_3$	105.042	105.700		60.000	35	15	6.40	C13-Serine POS
	Glycine	$C_2H_5NO_2$	75.032	76.040		29.700	20	20	6.02	C13-Glycine
Amino acids	Acetylalanine	$C_5H_9NO_3$	131.058		130.050	88.040	-20	-20	6.61	C13-Serine NEG
	Alanine	$C_3H_7NO_2$	89.0477	89.800		62.200	20	9	5.58	D5-Tryptophan
	Arginine	$C_6H_{14}N_4O_2$	174.111	175.119		70.067	20	20	6.95	D5-Kynurenic acid
	Asparagine	$C_4H_8N_2O_3$	132.053	133.061		74.020	20	20	6.41	C13-Serine POS
	Aspartic acid	$C_4H_7NO_4$	133.037	133.986		43.010	91	61	9.19	D5-Glutamic acid POS
	Citrulline	$C_6H_{13}N_3O_3$	175.095	176.104		70.066	20	20	6.14	C13-Glycine
	Cysteine	$C_3H_7NO_2S$	121.019	122.028		58.996	20	40	6.95	D5-Kynurenic acid
	Cystine	$C_6H_{12}N_2O_4S_2$	240.020	241.235		74.000	66	31	10.17	D5-Glutamic acid POS
	Glutamic acid	$C_5H_9NO_4$	147.053	148.100		84.000	65	20	9.28	D5-Glutamic acid POS
	Glutamine	$C_5H_{10}N_2O_3$	146.069	147.077		84.045	20	20	6.15	C13-Glycine
	Glutathione Oxidized	$C_{20}H_{32}N_6O_{12}S_2$	612.152		611.144	306.075	-20	-20	12.04	D5-Glutamic acid NEG
	Glutathione Reduced	$C_{10}H_{17}N_3O_6S$	307.083		306.076	143.046	-20	-20	11.63	D5-Glutamic acid NEG

Table S3 (continued). Summary of all the MS parameters for the different metabolites included in the method

Group	Metabolite	Formula	M_r	$[M+H]^+$ (*m/z*)	$[M+H]^-$ (*m/z*)	Product ion (*m/z*)	DP (V)	CE (V)	RT (min)	Internal Standard
Amino acids	Histidine	$C_6H_9N_3O_2$	155.069	156.077		110.071	20	20	6.60	D5-Kynurenic acid
	Hydroxyproline	$C_5H_9NO_3$	131.058	131.955		86.000	76	9	5.42	D5-Phenylalanine
	Isoleucine	$C_6H_{13}NO_2$	131.094	132.102		86.097	20	10	4.30	N15-Isoleucine
	Leucine	$C_6H_{13}NO_2$	131.094	132.100		44.049	20	40	4.23	C13-Adenosine
	Lysine	$C_6H_{14}N_2O_2$	146.105	147.113		84.081	20	20	7.72	D3-Lysine
	Phenylalanine	$C_9H_{11}NO_2$	165.079	166.087		120.081	20	20	4.81	D5-Phenylalanine
	Proline	$C_5H_9NO_2$	115.063	116.071		70.066	20	40	5.30	D5-Tryptophan
	Taurine	$C_2H_7NO_3S$	125.014	126.023		64.969	20	40	5.84	C13-Glycine
	Threonine	$C_4H_9NO_3$	119.058	120.066		56.050	20	40	6.05	C13-Glycine
	Tryptophan	$C_{11}H_{12}N_2O_2$	204.089	205.098		146.059	20	20	4.96	D5-Tryptophan
	Tyrosine	$C_9H_{11}NO_3$	181.073	182.082		91.055	20	40	5.76	D5-Tryptophan
	Valine	$C_5H_{11}NO_2$	117.079	118.087		72.081	20	10	4.89	D5-Tryptophan
Methionine cycle	Methionine	$C_5H_{11}NO_2S$	149.051	150.059		104.053	20	10	5.02	D5-Tryptophan
	SAM	$C_{15}H_{23}N_6O_5S$	399.145	400.196		251.085	51	21	6.27	C13-Glycine
	SAH	$C_{14}H_{20}N_6O_5S$	384.120	385.074		136.062	20	21	6.69	D5-Kynurenic acid
	Homocysteine	$C_4H_9NO_2S$	135.035	136.040		90.030	10	15	4.22	N15-Isoleucine
	Creatine	$C_4H_9N_3O_2$	131.069	132.056		90.050	10	17	5.10	D5-Phenylalanine
	Creatinine	$C_4H_7N_3O$	113.059	114.069		44.100	10	21	2.29	D3- Adenine
	Folic acid	$C_{19}H_{19}N_7O_6$	441.139	442.180		295.093	20	21	16.00	C13-Acetyl-CoA POS
	Tetrahydrofolate	$C_{19}H_{23}N_7O_6$	445.430	446.262		299.139	10	27	11.97	D3-Quinolinic acid POS
Coenzyme	Acetyl-CoA	$C_{23}H_{38}N_7O_{17}P_3S$	809.120		808.117	408.000	-200	-50	16.23	C13-Acetyl-CoA NEG
	CoA	$C_{21}H_{36}N_7O_{16}P_3S$	767.115		766.107	407.800	-200	-48	15.72	C13-Acetyl-CoA NEG
	FAD	$C_{27}H_{33}N_9O_{15}P_2$	785.157	786.165		348.069	20	20	12.02	D3-Quinolinic acid POS
	NAD	$C_{21}H_{27}N_7O_{14}P_2$	663.109	664.117		136.061	20	40	9.53	D5-Glutamic acid POS
	NADP	$C_{21}H_{28}N_7O_{17}P_3$	743.075	744.083		136.061	20	40	14.41	D3-Quinolinic acid POS
	NADH	$C_{21}H_{29}N_7O_{14}P_2$	665.441	666.128		136.060	20	40	11.62	D3-Quinolinic acid POS
	NADPH	$C_{21}H_{30}N_7O_{17}P_3$	745.090	746.099		729.018	200	25	15.97	N15-ATP
Nucleobase	Adenine	$C_5H_5N_5$	135.054	136.062		119.035	20	20	3.34	C13-Adenosine
	Guanine	$C_5H_5N_5O$	151.049	152.057		135.030	20	20	4.56	D4-Kynurenine
	Hypoxanthine	$C_5H_4N_4O$	136.039	137.047		55.029	20	40	5.08	D5-Tryptophan
	Thymine	$C_5H_6N_2O_2$	126.042		124.800	42.000	-35	-36	1.65	D3-Tymine
	Uracil	$C_4H_4N_2O_2$	112.027		111.019	41.998	-20	-20	1.98	D3-Tymine
	Cytosine	$C_4H_5N_3O$	111.000	112.030		52.010	61	41	2.95	C13-Adenosine
	Uridine	$C_9H_{12}N_2O_6$	244.069	245.077		113.035	20	20	2.75	C13-Adenosine
	Xanthine	$C_5H_4N_4O_2$	152.033	153.041		110.035	20	20	9.20	D5-Glutamic acid POS

Table S3 (continued). Summary of all the MS parameters for the different metabolites included in the method

Group	Metabolite	Formula	M_r	$[M+H]^+$ (*m/z*)	$[M+H]^-$ (*m/z*)	Product ion (*m/z*)	DP (V)	CE (V)	RT (min)	Internal Standard
Nucleoside	Adenosine	$C_{10}H_{13}N_5O_4$	267.097	268.105		136.200	20	23	2.69	C13-Adenosine
	Guanosine	$C_{10}H_{13}N_5O_5$	283.091	284.099		152.050	20	20	5.14	D5-Tryptophan
	Inosine	$C_{10}H_{12}N_4O_5$	268.080	269.100		137.100	70	17	5.65	D5-Tryptophan
Nucleotide	ADP	$C_{10}H_{15}N_5O_{10}P_2$	427.029	428.037		136.060	20	40	14.75	N15-ATP
	GDP	$C_{10}H_{15}N_5O_{11}P_2$	443.024		442.017	78.959	-20	-40	16.36	C13-Acetyl-CoA NEG
	UDP	$C_9H_{14}N_2O_{12}P_2$	404.002		402.994	158.920	-20	-40	15.04	D3-Quinolinic acid NEG
Nucleotide	ADP-Glucose	$C_{16}H_{25}N_5O_{15}P_2$	589.082	590.090		136.060	20	20	11.96	D3-Quinolinic acid POS
derivative	ADP-Ribose	$C_{15}H_{23}N_5O_{14}P_2$	559.071	560.079		136.060	20	40	11.45	D3-Quinolinic acid POS
Nucleotide	ATP	$C_{10}H_{16}N_5O_{13}P_3$	506.995	508.004		136.060	20	40	18.24	N15-ATP
	GTP	$C_{10}H_{16}N_5O_{14}P_3$	522.990		521.983	158.920	-155	-50	13.59	D3-Quinolinic acid NEG
	AMP	$C_{10}H_{14}N_5O_7P$	347.063	347.900		136.300	40	25	12.24	N15-AMP
	CMP	$C_9H_{14}N_3O_8P$	323.051	324.059		112.050	20	20	12.08	D3-Quinolinic acid POS
	GMP	$C_{10}H_{14}N_5O_8P$	363.050	364.066		152.050	20	20	13.33	D3-Quinolinic acid POS
	IMP	$C_{10}H_{13}N_4O_8P$	348.047	349.055		137.046	20	20	12.29	D3-Quinolinic acid POS
	UMP	$C_9H_{13}N_2O_9P$	324.035	325.044		97.020	20	20	12.25	D3-Quinolinic acid POS
	cAMP	$C_{10}H_{12}N_5O_6P$	329.052	330.060		136.060	20	20	9.19	D5-Glutamic acid POS
	cGMP	$C_{10}H_{12}N_5O_7P$	345.047	346.055		152.050	20	20	10.07	D5-Glutamic acid POS
	dAMP	$C_{10}H_{14}N_5O_6P$	331.068	332.076		136.060	20	20	12.28	N15-AMP
	dCMP	$C_9H_{14}N_3O_7P$	307.056	308.065		112.050	20	20	12.06	D3-Quinolinic acid POS
	dTMP	$C_{10}H_{15}N_2O_8P$	322.056	323.064		81.000	20	25	12.02	D3-Quinolinic acid POS
Other	Carnitine	$C_7H_{15}NO_3$	161.105	162.310		102.909	31	23	5.43	D5-Tryptophan
	Orotic acid	$C_5H_4N_2O_4$	156.017		155.009	111.010	-20	-20	7.27	C13-Serine NEG
	Salicylic acid	$C_7H_6O_3$	138.032		137.024	93.030	-20	-20	4.66	Glucose C13
	Ascorbic acid	$C_6H_8O_6$	176.032		174.830	86.900	-200	-28	9.28	D5-Glutamic acid NEG
	Ketoisovalerate	$C_5H_8O_3$	116.047		114.847	70.924	-5	-10	4.93	C13-Glucose
	Kynurenine	$C_{10}H_{12}N_2O_3$	208.213	209.100		192.000	20	24	4.73	D4-Kynurenine
	3-OH-Kynurenine	$C_{10}H_{12}N_2O_4$	224.079	225.100		208.000	40	13	6.06	C13-Glycine
	Quinolinic acid	$C_7H_5NO_4$	167.021	168.000		78.000	20	28	12.2	D3-Quinolinic acid POS
	Kynurenic acid	$C_{10}H_7NO_3$	189.042	190.100		144.100	40	25	6.51	D5-Kynurenic acid
	Pantothenic acid	$C_9H_{17}NO_5$	219.110	220.200		90.100	90	21	6.99	D5-Kynurenic acid
	Maleic acid	$C_4H_4O_4$	116.010		115.000	71.000	-20	-20	11.97	C13-Glutamic acid NEG
	Ketoleucine	$C_6H_{10}O_3$	130.060		128.986	85.079	-10	-12	15.00	D3-Quinolinic acid NEG

Table S4. MS parameters of all the internal standards included in the method

Group	Metabolite	M_r	$[M+H]^+$ (*m/z*)	$[M+H]^-$ (*m/z*)	Product ion (*m/z*)	DP (V)	CE (V)	RT (min)	Conc. Used (ng/mL)
Glycolysis	D-Glucose-13C6	186.11		185.024	91.673	-80	-12	4.38	991.97
	D-Fructose-13C6	186.11		185.024	91.673	-80	-12	3.76	1116.66
Amino acids	L-Serine-13C3,15N,2,3,3-d3 POS	112.08	113.000		66.080	10	15	6.40	1494.03
	L-Serine-13C3,15N,2,3,3-d3 NEG	112.08		110.870	77.990	-20	-16	6.40	1494.03
	L-Glutamic acid-13C5 POS	152.09	153.000		87.900	10	21	9.28	2027.36
	L-Glutamic acid-13C5 NEG	152.09		151.000	107.100	-15	-18	9.28	2027.36
	Glycine C13	76.06	77.200		31.000	20	20	6.02	50655.96
	L-Isoleucine-13C6,15N	138.12	139.034		92.098	31	13	4.30	921.26
	L-Leucine-5,5,5-d3	134.19	135.044		89.127	46	13	4.23	895.05
	L-Lysine-4,4,5,5-d4	186.67	151.068		88.100	36	21	7.72	12445.29
	L-Phenyl-d5-alanine	170.22	171.045		125.115	31	17	4.81	1135.37
	Thymine-d4 (methyl-d3,6-d1)	130.14		128.900	42.010	-10	-30	1.65	868.03
	Tryptophan-D5	209.26	210.100		122.100	86	21	4.96	1395.76
Coenzyme	Acetyl-1,2-13C2 Coenzyme A POS	811.56	812.130		305.130	200	50	16.23	54078.70
	Acetyl-1,2-13C2 Coenzyme A NEG	811.56		810.125	463.050	-200	-50	16.23	54078.70
Nucleobase	2-(Methyl-13C,d3-thio) adenine	185.23	186.200		134.000	26	15	1.93	1235.48
Nucleoside	[1',2',3',4',5'-13C5]adenosine	272.21	114.069		44.100	64	27	2.69	182.38
Nucleotide	Adenosine-15N5 5′-monophosphate	352.19	352.948		141.079	56	23	12.24	23480.51
	Adenosine-15N5 5′-triphosphate	512.15	513.000		141.000	20	40	18.24	3416.04
Other	Quinolinic acid-4,5,6-D3 POS	170.14	171.000		81.000	20	28	12.20	1134.83
	Quinolinic acid-4,5,6-D3 NEG	170.14		169.026	125.031	-10	-27	12.20	1134.83
	D4-Kynurenine	212.24	213.100		140.100	20	13	4.73	1415.64
	Kynurenic acid-3,5,6,7,8-d5	194.20	195.100		149.100	40	25	6.51	1295.31

CHAPTER 6: SUMMARY AND FUTURE PROSPECTS

6.1 Summary

A need for an investigation into complex diseases such as IPF is growing. Due to the unknown nature of the pathogenesis of IPF, the short survival duration (2 to 3 years after diagnosis), and the limited effective treatment available, alternative approaches to characterising IPF is vital. The application of metabolomics have effected great advances in such investigations (Cuperlovic-Culf & Culf, 2016; Lu & Chen, 2017), especially when metabolic profiling approaches are used. Metabolic profiling has provided new insight into the pathogenesis of diseases such as asthma and COPD, as well as the identification of new therapeutic approaches. Although metabolomics is a great tool, there are still shortcomings to address, such as the lack of a total comprehensive approach and a standardised sample preparation procedure.

Therefore during this study, a targeted LC-MS/MS based method for metabolic profiling was established together with a sample preparation method suitable for various matrices.

In summary the aim and objectives of the study were:

1. The development of a standardised LC-MS/MS method for targeted metabolic profiling.
2. Establishment of a standardised sample preparation and metabolite extraction procedure suitable for various matrices.
3. Validation of the developed method by generating a metabolic profile for a bleomycin induced lung fibrosis C57BL/6L mouse model.

6.1.1 Method development

The first aim of the study was addressed through the development of the LC-MS/MS method. A HILIC method was developed with the use of the Luna NH_2 column (2 mm x 150 mm, 5 µm, 100 Å). A hundred (100) metabolites can be identified using this method in a scheduled MRM mode, with adequate accuracy and precision. The quality assessment that was performed on the analytical aspects of the LC-MS/MS method demonstrated acceptable linearity, precision and accuracy without carryover. From the quality assessment it was clear that the method could be employed for the generation of metabolic profiles for any disease state.

6.1.2 Standardised sample preparation

As part of the study a standardised sample preparation and metabolites extraction procedure that is non-selective, simple and robust was established. Various matrices, originating from three healthy C57BL/6J mice and fibroblasts from human donors were tested to determine the detectability of the different metabolites. Since the presence and detectability of the metabolites differ within the different matrices, not all metabolites were detected in all matrices. After the inter- and intra-day accuracy and precision quality assessment of the analytical aspects of the developed method, it was clear that some metabolites may be responsible for bias in the results. This issue was addressed with the implementation of an adequate quality control procedure performed with each analysis, eliminating any metabolites that are found to be unstable during the analysis.

6.1.3 Metabolic profile for a fibrotic lung C57BL/6L mouse model

As validation of the developed method, a metabolic profile for a fibrotic lung animal model was generated. This was achieved by comparing the abundance of the detectable metabolites in healthy lung samples against bleomycin induced fibrotic lung samples. After statistical analysis was performed 26 metabolites were identified as significant (p-values < 0.05). Since bleomycin exposure involves a fibrotic response as well as an inflammatory response, a cross analysis to a LPS mouse model was performed to establish whether or not the identified metabolites were a result of fibrosis or a consequence of inflammation. After cross analysis five metabolites overlapped between these two groups but none of the five metabolites demonstrated the same trend with regards to percentage increase/decrease. A metabolic profile was also generated for a TGF-β treated NHLF cellular model. The 24 significant metabolites within the cellular model were cross analysed with the bleomycin treated mouse model and 17 metabolites overlapped between these two groups. Of the 17 metabolites, seven metabolites demonstrated the same trend with regards to percentage increase/decrease with a 30% RSD margin. The seven metabolites identified as significant demonstrated an up regulation towards the healthy group and are: lactic acid, asparagine, isoleucine, tryptophan, tyrosine, uracil and uridine. With the substantial correlation between the C57BL/6J bleomycin treated mouse model's metabolic profile and that of the TGF-β NHLF cellular model, as well as to literature, the identified metabolites can be potential biomarkers for IPF.

6.2 Conclusion

With the identification of several metabolites, including lactic acid, inosine, hypoxanthine, methionine, SAH and non-essential amino acids proline and asparagine, in the C57BL/6J bleomycin treated mouse model and the substantial correlation to the TGF-β treated NHLF cellular model as well as to literature (Kang *et al.*, 2016; Kottmann *et al.*, 2012), these

metabolites can potentially be assigned as biomarkers for lung fibrosis. Validation of these metabolites would have to be done to determine whether or not these metabolites are biomarkers of IPF. The identification of the dysregulation in the methionine cycle suggests that an understanding of the pathogenesis of lung fibrosis may lie at the epigenetic level.

This study is a first with regards to metabolic profiling of the C57BL/6J bleomycin induced fibrotic lung mouse model. To our knowledge, there has to date not been a comparison made between the metabolic profiles of the C57BL/6J bleomycin mouse model and a cellular model resembling IPF or to human IPF lung tissue. With the results generated by this study we hypothesise that the C57BL/6J bleomycin treated mouse model for lung inflammation induction is a suitable animal model for the investigation of lung fibrosis and that the key to understanding lung fibrosis lies at an epigenetic level and the dysregulation of epigenetic modifications. Our hypothesis is supported by a recent study done by Guiot *et al.* (2017). The study done by Guiot *et al.* (2017) involved an investigation of circulating nucleosomes. These authors observed a significant reduction in cell free nucleosome level associated with methylated DNA (5-methylcytosine) in IPF patients compared to healthy subjects, indicating that a dysregulation of epigenetic modification is present.

During this study an LC-MS/MS targeted metabolic profiling method was established together with a sample preparation procedure that is non-selective. The quality assessment of the developed method demonstrated that the method is sensitive, robust, reproducible and compatible for various matrices from different origins. This method can be used in future metabolic profiling studies for characterising any disease state. The metabolic profile of a fibrotic lung animal model was established during this study and yielded important information, contributing to a better understanding of the fibrotic lung animal model and how the model correlates to the human fibrotic lung condition. Potential biomarkers were identified for lung fibrosis and the identification of the dysregulation in the methionine cycle suggested that lung fibrosis may lie at the epigenetic level. Not only was a comparison made between the metabolic profile of the C57BL/6J bleomycin treated mouse model and the TGF-β treated normal human lung fibroblasts model but also to an LPS treated lung inflammation mouse model. This contributes to a better understanding of the differences and similarities that are present between the metabolic profiles of the bleomycin treated fibrotic lung mouse model and the LPS treated lung inflammation model. From the knowledge gained in this study about the metabolic profile of the fibrotic lung animal model and potential biomarkers for lung fibrosis, further studies can be performed especially in the field of epigenetics and epigenetic modification to determine the cause of IPF and identify a treatment plan.

6.3 Future prospects

As contribution to future developments for IPF characterisation, the following suggestions are made:

- Firstly, the findings in this study should be validated since sample size was limited during the study.
- Together with the validation of the findings of this study, the identified metabolites should be validated to determine if one or more these metabolites can be assigned as biomarker(s) for a fibrotic lung animal model. It must also be investigated if these metabolites are also biomarkers for IPF by analysing lung samples from human IPF patients. The metabolites identified as significant in this study is based on the observed increased levels of the metabolites in the treated group. Therefore, a comparison should be made between lung samples from IPF patients and the findings in this study to evaluate whether or not this up regulation can be seen in IPF.
- During this study the methionine cycle was identified as dysregulated and future studies could incorporate other -omics platforms including epigenetics to determine the influence of the dysregulated methionine cycle and how it contributes to the pathogenesis of IPF.
- This study provided key information for the characterisation of the C57BL/6J bleomycin treated mouse model for fibrotic lung induction. This information can be used for optimisation of the bleomycin model to ensure accurate representation of IPF in the animal model.
- A future study should include the investigation of plasma biomarker identification for the C57BL/6J bleomycin model. It would be greatly beneficial if the identified biomarkers, with the dysregulated levels, can be identified in plasma as well. This would provide the opportunity to study the progression of lung fibrosis and get a better understanding of the pathogenesis of the disease.
- Since IPF is characterised by multiple aspects contributing to the progression of the disease such as alveolar epithelial cell injury, proliferation of activated fibroblasts and myofibroblasts, and accumulation of the extracellular matrix that stiffens the lung and leads to respiratory failure, it would be beneficial to gain an -omics overview of each aspect. Biomarker identification (metabolites, lipids, proteins, etc.) should be performed, using all omics platforms, for each aspect separately (epithelial cell injury, proliferation, myofibroblasts differentiation and the extracellular matrix accumulation). This could provide new insight into therapeutic approaches.

CHAPTER 7: REFERENCES

Adamko, D.J., Nair, P., Mayers, I., Tsuyuki, R.T., Regush, S. & Rowe, B.H. 2015. Metabolomic profiling of asthma and chronic obstructive pulmonary disease: A pilot study differentiating diseases. *Journal of Allergy and Clinical Immunology,* 136 (3):571-580. e573.

Álvarez-Sánchez, B., Priego-Capote, F. & de Castro, M.L. 2010a. Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation. *TrAC Trends in Analytical Chemistry,* 29 (2):111-119.

Álvarez-Sánchez, B., Priego-Capote, F. & de Castro, M.L. 2010b. Metabolomics analysis II. Preparation of biological samples prior to detection. *TrAC Trends in Analytical Chemistry,* 29 (2):120-127.

Aretz, I. & Meierhofer, D. 2016. Advantages and pitfalls of mass spectrometry based metabolome profiling in systems biology. *International journal of molecular sciences,* 17 (5):632.

Armitage, E.G. & Barbas, C. 2014. Metabolomics in cancer biomarker discovery: current trends and future perspectives. *Journal of pharmaceutical and biomedical analysis,* 87:1-11.

Art, S.M. 2017. Transcription - Servier Medical Art. [online] Available at: http://smart.servier.com/smart_image/transcription/ [Accessed 3 Nov. 2017].

Bajad, S.U., Lu, W., Kimball, E.H., Yuan, J., Peterson, C. & Rabinowitz, J.D. 2006. Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *Journal of chromatography A,* 1125 (1):76-88.

Bino, R.J., Hall, R.D., Fiehn, O., Kopka, J., Saito, K., Draper, J., et al. 2004. Potential of metabolomics as a functional genomics tool. *Trends in plant science,* 9 (9):418-425.

Boccard, J., Veuthey, J.L. & Rudaz, S. 2010. Knowledge discovery in metabolomics: an overview of MS data handling. *Journal of separation science,* 33 (3):290-304.

Bonassi, S., Neri, M. & Puntoni, R. 2001. Validation of biomarkers as early predictors of disease. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 480:349-358.

Buszewski, B. & Noga, S. 2012. Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Analytical and bioanalytical chemistry,* 402 (1):231-247.

Costabel, U., Bendstrup, E., Cottin, V., Dewint, P., Egan, J.J., Ferguson, J., et al. 2014. Pirfenidone in idiopathic pulmonary fibrosis: expert panel discussion on the management of drug-related adverse events. *Advances in therapy,* 31 (4):375-391.

Cottrill, K.A. & Chan, S.Y. 2013. Metabolic dysfunction in pulmonary hypertension: the expanding relevance of the Warburg effect. *European journal of clinical investigation,* 43 (8):855-865.

Croghan, C. & Egeghy, P. 2003. Methods of dealing with values below the limit of detection using SAS. *Southern SAS User Group*:22-24.

Cuperlovic-Culf, M. & Culf, A. 2016. Applied metabolomics in drug discovery. *Expert opinion on drug discovery,* 11 (8):759-770.

DeNicola, G.M., Chen, P.-H., Mullarky, E., Sudderth, J.A., Hu, Z., Wu, D., et al. 2015. NRF2 regulates serine biosynthesis in non-small cell lung cancer. *Nature genetics,* 47 (12):1475-1481.

Dettmer, K., Aronov, P.A. & Hammock, B.D. 2007. Mass spectrometry-based metabolomics. *Mass spectrometry reviews,* 26 (1):51-78.

Dunn, W.B., Bailey, N.J. & Johnson, H.E. 2005. Measuring the metabolome: current analytical technologies. *Analyst,* 130 (5):606-625.

Dunn, W.B., Wilson, I.D., Nicholls, A.W. & Broadhurst, D. 2012. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans.2249-2260.

En.wikipedia.org. 2017. Protein GBA3 PDB 2E9L.png. [online] Available at: https://en.wikipedia.org/wiki/File:Protein_GBA3_PDB_2E9L.png [Accessed 3 Nov. 2017].

Fuertig, R., Ceci, A., Camus, S.M., Bezard, E., Luippold, A.H. & Hengerer, B. 2016. LC–MS/MS-based quantification of kynurenine metabolites, tryptophan, monoamines and neopterin in plasma, cerebrospinal fluid and brain. *Bioanalysis,* 8 (18):1903-1917.

Genome.jp. 2017. KEGG: Kyoto Encyclopedia of Genes and Genomes. [online] Available at: http://www.genome.jp/kegg/ [Accessed 18 May. 2017].

Gika, H. & Theodoridis, G. 2011. Sample preparation prior to the LC–MS-based metabolomics/metabonomics of blood-derived samples.1647-1652.

Godzien, J., Alonso-Herranz, V., Barbas, C. & Armitage, E.G. 2015. Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample. *Metabolomics,* 11 (3):518-528.

González, O., Blanco, M.E., Iriarte, G., Bartolomé, L., Maguregui, M.I. & Alonso, R.M. 2014. Bioanalytical chromatographic method validation according to current regulations, with a special focus on the non-well defined parameters limit of quantification, robustness and matrix effect. *Journal of chromatography A,* 1353:10-27.

Goodacre, R., Broadhurst, D., Smilde, A.K., Kristal, B.S., Baker, J.D., Beger, R., et al. 2007. Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics,* 3 (3):231-241.

Griffiths, W.J., Koal, T., Wang, Y., Kohl, M., Enot, D.P. & Deigner, H.P. 2010. Targeted metabolomics for biomarker discovery. *Angewandte Chemie International Edition,* 49 (32):5426-5445.

Guiot, J., Struman, I., Chavez, V., Henket, M., Herzog, M., Scoubeau, K., et al. 2017. Altered epigenetic features in circulating nucleosomes in idiopathic pulmonary fibrosis. *Clinical epigenetics,* 9 (1):84.

Han, X. 2009. Lipidomics: developments and applications. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences,* 877 (26):2663.

Henion, J., Brewer, E. & Rule, G. 1998. Peer reviewed: Sample preparation for LC/MS/MS: Analyzing biological and environmental samples. *Analytical chemistry,* 70 (19):650A-656A.

Ilyin, S.E., Belkowski, S.M. & Plata-Salamán, C.R. 2004. Biomarker discovery and validation: technologies and integrative approaches. *Trends in biotechnology,* 22 (8):411-416.

Johnson, C.H. & Gonzalez, F.J. 2012. Challenges and opportunities of metabolomics. *Journal of cellular physiology,* 227 (8):2975-2981.

Johnson, J.M., Strobel, F.H., Reed, M., Pohl, J. & Jones, D.P. 2008. A rapid LC-FTMS method for the analysis of cysteine, cystine and cysteine/cystine steady-state redox potential in human plasma. *Clinica chimica acta,* 396 (1-2):43-48.

Kang, Y.P., Lee, S.B., Lee, J.m., Kim, H.M., Hong, J.Y., Lee, W.J., et al. 2016. Metabolic profiling regarding pathogenesis of Idiopathic Pulmonary Fibrosis. *Journal of proteome research*:1717-1724.

King Jr, T.E., Bradford, W.Z., Castro-Bernardini, S., Fagan, E.A., Glaspole, I., Glassberg, M.K., et al. 2014. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *New England Journal of Medicine,* 370 (22):2083-2092.

Kirkwood, J.S., Maier, C. & Stevens, J.F. 2013. Simultaneous, Untargeted Metabolic Profiling of Polar and Nonpolar Metabolites by LC-Q-TOF Mass Spectrometry. *Current protocols in toxicology*:4.39. 31-34.39. 12.

Kottmann, R.M., Kulkarni, A.A., Smolnycki, K.A., Lyda, E., Dahanayake, T., Salibi, R., et al. 2012. Lactic acid is elevated in idiopathic pulmonary fibrosis and induces myofibroblast differentiation via pH-dependent activation of transforming growth factor-β. *American journal of respiratory and critical care medicine*:740-751.

LaBrecque, D.R., Abbas, Z., Anania, F., Ferenci, P., Khan, A.G., Goh, K.-L., et al. 2014. World Gastroenterology Organisation global guidelines: Nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Journal of clinical gastroenterology,* 48 (6):467-473.

Locasale, J.W. 2013. Serine, glycine and one-carbon units: cancer metabolism in full circle. *Nature Reviews Cancer,* 13 (8):572-583.

Lu, Y. & Chen, C. 2017. Metabolomics: Bridging Chemistry and Biology in Drug Discovery and Development. *Current Pharmacology Reports,* 3 (1):16-25.

Luo, B., Groenke, K., Takors, R., Wandrey, C. & Oldiges, M. 2007. Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography–mass spectrometry. *Journal of chromatography A,* 1147 (2):153-164.

Mapelli, V., Olsson, L. & Nielsen, J. 2008. Metabolic footprinting in microbiology: methods and applications in functional genomics and biotechnology. *Trends in biotechnology,* 26 (9):490-497.

Masekela, R., Olorunju, S., Green, R.J. & Magidimisa, N. 2016. Lung function decline is accelerated in South Africans with cystic fibrosis. *South African Family Practice,* 58 (1):24-27.

Nalysnyk, L., Cid-Ruzafa, J., Rotella, P. & Esser, D. 2012. Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. *European Respiratory Review,* 21 (126):355-361.

Patti, G.J., Yanes, O. & Siuzdak, G. 2012. Innovation: Metabolomics: the apogee of the omics trilogy. *Nature reviews Molecular cell biology,* 13 (4):263-269.

Pesek, J. & Matyska, M.T. 2007. A comparison of two separation modes: HILIC and aqueous normal phase chromatography. *LC GC NORTH AMERICA,* 25 (5):480.

Peters, F.T., Drummer, O.H. & Musshoff, F. 2007. Validation of new methods. *Forensic science international,* 165 (2):216-224.

Phenomenex, I. 2015. LUNA Brochure. In Phenomenex (Ed.) (pp. 1-31).

Pixabay.com. 2017. Free Image on Pixabay - Dna, Dna Strand, Dna Helix, Helix. [online] Available at: https://pixabay.com/en/dna-dna-strand-dna-helix-helix-1116336/ [Accessed 3 Nov. 2017].

Putri, S.P. & Fukusaki, E. 2014. Mass spectrometry-based metabolomics: a practical guide. CRC Press:50-115.

Richeldi, L., du Bois, R.M., Raghu, G., Azuma, A., Brown, K.K., Costabel, U., et al. 2014. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *New England Journal of Medicine,* 370 (22):2071-2082.

Rindlisbacher, B., Schmid, C., Geiser, T., Bovet, C. & Funke-Chambour, M. 2018. Serum metabolic profiling identified a distinct metabolic signature in patients with idiopathic pulmonary fibrosis–a potential biomarker role for LysoPC. *Respiratory research,* 19 (1):7.

Roberts, L.D., Souza, A.L., Gerszten, R.E. & Clish, C.B. 2012. Targeted metabolomics. *Current protocols in molecular biology*:30.32. 31-30.32. 24.

Rochfort, S. 2005. Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. *Journal of natural products,* 68 (12):1813-1820.

Rozet, E., Marini, R., Ziemons, E., Boulanger, B. & Hubert, P. 2011. Advances in validation, risk and uncertainty assessment of bioanalytical methods. *Journal of pharmaceutical and biomedical analysis,* 55 (4):848-858.

Sandbo, N. 2014. Mechanisms of fibrosis in IPF. (*In* Idiopathic Pulmonary Fibrosis. Springer. p. 161-205).

Sandi, A., Bede, A., Szepesy, L. & Rippel, G. 1997. Characterization of different RP-HPLC columns by a gradient elution technique. *Chromatographia,* 45 (1):206-214.

SCIEX, A. 2014. Turbo V Ion Source Operation Giude (pp. 17-54): AB Sciex Pte. Ltd.

Solomon, D.J. & Fischer, S.M. 2010. Metabolomics at Agilent: driving value with comprehensive solutions. *Metabolomics research applications*:1-13.

Stringer, K.A., McKay, R.T., Karnovsky, A., Quémerais, B. & Lacy, P. 2016. Metabolomics and its application to acute lung diseases. *Frontiers in immunology,* 7:44.

Szpunar, J. 2004. Metallomics: a new frontier in analytical chemistry. *Analytical and bioanalytical chemistry,* 378 (1):54-56.

Theodoridis, G., Gika, H.G. & Wilson, I.D. 2011. Mass spectrometry-based holistic analytical approaches for metabolite profiling in systems biology studies. *Mass spectrometry reviews,* 30 (5):884-906.

Vuckovic, D. 2012. Current trends and challenges in sample preparation for global metabolomics using liquid chromatography–mass spectrometry. *Analytical and bioanalytical chemistry,* 403 (6):1523-1548.

Walsh, M., Nugent, A., Brennan, L. & Gibney, M. 2008. Understanding the metabolome–challenges for metabolomics. *Nutrition Bulletin,* 33 (4):316-323.

Watson, D.W. 2017. Critical Evaluation of HPLC Methods: Working with Ionizable Analytes. *Advances in Pharmaceutical Analysis,* 35 (6):406.

Wiechert, W., Schweissgut, O., Takanaga, H. & Frommer, W.B. 2007. Fluxomics: mass spectrometry versus quantitative imaging. *Current opinion in plant biology,* 10 (3):323-330.

Wishart, D.S. 2016. Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery*:473-484.

Wu, J. 2009. Development and evaluation of new silica based polar and/or mixed modal stationary phases for HPLC. uniwien) 2-25 p.

Zhang, A., Sun, H., Wang, P., Han, Y. & Wang, X. 2012a. Modern analytical techniques in metabolomics analysis. *Analyst,* 137 (2):293-300.

Zhang, Y. & Kaminski, N. 2012b. Biomarkers in idiopathic pulmonary fibrosis. *Current opinion in pulmonary medicine,* 18 (5):441.

APPENDIX A: AUTHOR GUIDLINES

RESPIRATORY MEDICINE

AUTHOR INFORMATION PACK

TABLE OF CONTENTS

ISSN: 0954-6111

DESCRIPTION

Contact the Editorial Office respiratorymedicine@elsevier.com

Respiratory Medicine is an internationally-renowned journal devoted to the rapid publication of clinically-relevant **respiratory medicine** research. It combines cutting-edge original research with state-of-the-art reviews dealing with all aspects of **respiratory diseases** and **therapeutic interventions**. Topics include adult and paediatric medicine, epidemiology, immunology and cell biology, physiology, occupational disorders, and the role of **allergens** and **pollutants**.

Respiratory Medicine is increasingly the journal of choice for publication of phased trial work, commenting on effectiveness, dosage and methods of action.

To order this journal online, visit http://www.journals.elsevierhealth.com/periodicals/yrmed/pricing.

IMPACT FACTOR

2016: 3.217 © Thomson Reuters Journal Citation Reports 2017

ABSTRACTING AND INDEXING

AIDS Abstracts
Current Contents/Life Sciences and Clinical Medicine
MEDLINE®
EMBASE
Science Citation Index
Excerpta Medica
Current Awareness in Biological Sciences
SIIC Data Bases
Scopus

EDITORIAL BOARD

GUIDE FOR AUTHORS

Your Paper Your Way

We now differentiate between the requirements for new and revised submissions. You may choose to submit your manuscript as a single Word or PDF file to be used in the refereeing process. Only when your paper is at the revision stage, will you be requested to put your paper in to a 'correct format' for acceptance and provide the items required for the publication of your article.
To find out more, please visit the Preparation section below.

Respiratory Medicine is an internationally-renowned, clinically-oriented journal, combining cutting-edge original research with state-of-the-art reviews dealing with all aspects of respiratory diseases and therapeutic interventions, but with a clear clinical relevance. The journal is an established forum for the publication of phased clinical trial work at the forefront of interventive research. As well as full-length original research papers, the journal publishes reviews, correspondence, and short reports. The Journal also publishes regular supplements on areas of special interest.

Submission checklist

You can use this list to carry out a final check of your submission before you send it to the journal for review. Please check the relevant section in this Guide for Authors for more details.

Ensure that the following items are present:

One author has been designated as the corresponding author with contact details:
• E-mail address
• Full postal address

All necessary files have been uploaded:
Manuscript:
• Include keywords
• All figures (include relevant captions)
• All tables (including titles, description, footnotes)
• Ensure all figure and table citations in the text match the files provided
• Indicate clearly if color should be used for any figures in print
Graphical Abstracts / Highlights files (where applicable)
Supplemental files (where applicable)

Further considerations
• Manuscript has been 'spell checked' and 'grammar checked'
• All references mentioned in the Reference List are cited in the text, and vice versa
• Permission has been obtained for use of copyrighted material from other sources (including the Internet)
• A competing interests statement is provided, even if the authors have no competing interests to declare
• Journal policies detailed in this guide have been reviewed
• Referee suggestions and contact details provided, based on journal requirements

For further information, visit our Support Center.

BEFORE YOU BEGIN

Ethics in publishing

Please see our information pages on Ethics in publishing and Ethical guidelines for journal publication.

Declaration of interest

All authors must disclose any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work. Examples of potential conflicts of interest include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/ registrations, and grants or other funding. If there are no conflicts of interest then please state this: 'Conflicts of interest: none'. More information.

Submission declaration and verification

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see 'Multiple, redundant or concurrent publication' section of our ethics policy for more information), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. To verify originality, your article may be checked by the originality detection service CrossCheck.

Authorship

All authors should have made substantial contributions to all of the following: (1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, (3) final approval of the version to be submitted.

Changes to authorship

Authors are expected to consider carefully the list and order of authors **before** submitting their manuscript and provide the definitive list of authors at the time of the original submission. Any addition, deletion or rearrangement of author names in the authorship list should be made only **before** the manuscript has been accepted and only if approved by the journal Editor. To request such a change, the Editor must receive the following from the **corresponding author**: (a) the reason for the change in author list and (b) written confirmation (e-mail, letter) from all authors that they agree with the addition, removal or rearrangement. In the case of addition or removal of authors, this includes confirmation from the author being added or removed.

Only in exceptional circumstances will the Editor consider the addition, deletion or rearrangement of authors **after** the manuscript has been accepted. While the Editor considers the request, publication of the manuscript will be suspended. If the manuscript has already been published in an online issue, any requests approved by the Editor will result in a corrigendum.

Clinical trial results

In line with the position of the International Committee of Medical Journal Editors, the journal will not consider results posted in the same clinical trials registry in which primary registration resides to be prior publication if the results posted are presented in the form of a brief structured (less than 500 words) abstract or table. However, divulging results in other circumstances (e.g., investors' meetings) is discouraged and may jeopardise consideration of the manuscript. Authors should fully disclose all posting in registries of results of the same or closely related work.

When submitting a Clinical Trial paper to the journal via the online submission system please select Clinical Trial Paper as an article type.

Reporting clinical trials

Randomized controlled trials should be presented according to the CONSORT guidelines. At manuscript submission, authors must provide the CONSORT checklist accompanied by a flow diagram that illustrates the progress of patients through the trial, including recruitment, enrollment, randomization, withdrawal and completion, and a detailed description of the randomization procedure. The CONSORT checklist and template flow diagram are available online.

Registration of clinical trials

Registration in a public trials registry is a condition for publication of clinical trials in this journal in accordance with International Committee of Medical Journal Editors recommendations. Trials must register at or before the onset of patient enrolment. The clinical trial registration number should be included at the end of the abstract of the article. A clinical trial is defined as any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects of health outcomes. Health-related interventions include any intervention used to modify a biomedical or health-related outcome (for example drugs, surgical procedures, devices, behavioural treatments, dietary interventions, and process-of-care changes). Health outcomes include any biomedical or health-related measures obtained in patients or participants, including pharmacokinetic measures and adverse events. Purely observational studies (those in which the assignment of the medical intervention is not at the discretion of the investigator) will not require registration.

Article transfer service
This journal is part of our Article Transfer Service. This means that if the Editor feels your article is more suitable in one of our other participating journals, then you may be asked to consider transferring the article to one of those. If you agree, your article will be transferred automatically on your behalf with no need to reformat. Please note that your article will be reviewed again by the new journal. More information.

Copyright

Upon acceptance of an article, authors will be asked to complete a 'Journal Publishing Agreement' (see more information on this). An e-mail will be sent to the corresponding author confirming receipt of the manuscript together with a 'Journal Publishing Agreement' form or a link to the online version of this agreement.

Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within their institutions. Permission of the Publisher is required for resale or distribution outside the institution and for all other derivative works, including compilations and translations. If excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by authors in these cases.

For open access articles: Upon acceptance of an article, authors will be asked to complete an 'Exclusive License Agreement' (more information). Permitted third party reuse of open access articles is determined by the author's choice of user license.

Author rights
As an author you (or your employer or institution) have certain rights to reuse your work. More information.

Elsevier supports responsible sharing
Find out how you can share your research published in Elsevier journals.

Role of the funding source

You are requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement then this should be stated.

Funding body agreements and policies
Elsevier has established a number of agreements with funding bodies which allow authors to comply with their funder's open access policies. Some funding bodies will reimburse the author for the Open Access Publication Fee. Details of existing agreements are available online.
After acceptance, open access papers will be published under a noncommercial license. For authors requiring a commercial CC BY license, you can apply after your manuscript is accepted for publication.

Open access

This journal offers authors a choice in publishing their research:

Subscription
• Articles are made available to subscribers as well as developing countries and patient groups through our universal access programs.
• No open access publication fee payable by authors.
Open access
• Articles are freely available to both subscribers and the wider public with permitted reuse.
• An open access publication fee is payable by authors or on their behalf, e.g. by their research funder or institution.

Regardless of how you choose to publish your article, the journal will apply the same peer review criteria and acceptance standards.

For open access articles, permitted third party (re)use is defined by the following Creative Commons user licenses:

Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)
For non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.

The open access publication fee for this journal is **USD 3300**, excluding taxes. Learn more about Elsevier's pricing policy: https://www.elsevier.com/openaccesspricing.

Green open access
Authors can share their research in a variety of different ways and Elsevier has a number of green open access options available. We recommend authors see our green open access page for further information. Authors can also self-archive their manuscripts immediately and enable public access from their institution's repository after an embargo period. This is the version that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and in editor-author communications. Embargo period: For subscription articles, an appropriate amount of time is needed for journals to deliver value to subscribing customers before an article becomes freely available to the public. This is the embargo period and it begins from the date the article is formally published online in its final and fully citable form. Find out more.

This journal has an embargo period of 12 months.

Elsevier Publishing Campus
The Elsevier Publishing Campus (www.publishingcampus.com) is an online platform offering free lectures, interactive training and professional advice to support you in publishing your research. The College of Skills training offers modules on how to prepare, write and structure your article and explains how editors will look at your paper when it is submitted for publication. Use these resources, and more, to ensure that your submission will be the best that you can make it.

Language (usage and editing services)
Please write your text in good English (American or British usage is accepted, but not a mixture of these). Authors who feel their English language manuscript may require editing to eliminate possible grammatical or spelling errors and to conform to correct scientific English may wish to use the English Language Editing service available from Elsevier's WebShop.

Submission

Our online submission system guides you stepwise through the process of entering your article details and uploading your files. The system converts your article files to a single PDF file used in the peer-review process. Editable files (e.g., Word, LaTeX) are required to typeset your article for final publication. All correspondence, including notification of the Editor's decision and requests for revision, is sent by e-mail.

Submit your article
Please submit your article via http://ees.elsevier.com/yrmed.

Submissions are allocated to a handling editor, typically an Associate Editor. Should the paper be considered suitable for peer review, appropriate reviewers will be recruited. Authors are required to provide the name and full contact details of 2 potential reviewers, though choice of reviewers is at the discretion of the handling editor.
The final decision-making responsibility lies with the handling editor, who reserves the right to reject the paper despite favourable reviews depending on the priorities of the journal.

Reviews
The journal welcomes submission of state-of-the-art reviews on important topics with a clinical relevance. Potential review authors are encouraged to contact the Deputy Editor Dr N. Hanania hanania@bcm.tmc.edu in advance with their review proposals.

Case Reports
Case reports will no longer be considered for publication in Respiratory Medicine, but instead should be directed to the sister publication Respiratory Medicine Case Reports. Please note that this is a separate publication. Case reports should be submitted for consideration by Respiratory Medicine Case Reports via http://ees.elsevier.com/rmcr/. Respiratory Medicine Case Reports is an open access journal and all authors will be required to pay a £250 processing fee to cover the costs of publishing the article, which authors will be required to pay once an article has passed peer review.

<Brief Communications

These should be submitted as detailed above but should not exceed 1000 words, and may normally contain only one illustration or table. Brief communications containing new information may be selected for rapid peer review and publication at the discretion of the editor and editorial board.

PREPARATION

NEW SUBMISSIONS

Submission to this journal proceeds totally online and you will be guided stepwise through the creation and uploading of your files. The system automatically converts your files to a single PDF file, which is used in the peer-review process.
As part of the Your Paper Your Way service, you may choose to submit your manuscript as a single file to be used in the refereeing process. This can be a PDF file or a Word document, in any format or lay-out that can be used by referees to evaluate your manuscript. It should contain high enough quality figures for refereeing. If you prefer to do so, you may still provide all or some of the source files at the initial submission. Please note that individual figure files larger than 10 MB must be uploaded separately.

References
There are no strict requirements on reference formatting at submission. References can be in any style or format as long as the style is consistent. Where applicable, author(s) name(s), journal title/book title, chapter title/article title, year of publication, volume number/book chapter and the pagination must be present. Use of DOI is highly encouraged. The reference style used by the journal will be applied to the accepted article by Elsevier at the proof stage. Note that missing data will be highlighted at proof stage for the author to correct.

Formatting requirements
There are no strict formatting requirements but all manuscripts must contain the essential elements needed to convey your manuscript, for example Abstract, Keywords, Introduction, Materials and Methods, Results, Conclusions, Artwork and Tables with Captions.
If your article includes any Videos and/or other Supplementary material, this should be included in your initial submission for peer review purposes.
Divide the article into clearly defined sections.

Figures and tables embedded in text
Please ensure the figures and the tables included in the single file are placed next to the relevant text in the manuscript, rather than at the bottom or the top of the file. The corresponding caption should be placed directly below the figure or table.

Peer review

This journal operates a single blind review process. All contributions are typically sent to a minimum of two independent expert reviewers to assess the scientific quality of the paper. The Editor is responsible for the final decision regarding acceptance or rejection of articles. The Editor's decision is final. More information on types of peer review.

REVISED SUBMISSIONS

Use of word processing software
Regardless of the file format of the original submission, at revision you must provide us with an editable file of the entire article. Keep the layout of the text as simple as possible. Most formatting codes will be removed and replaced on processing the article. The electronic text should be prepared in a way very similar to that of conventional manuscripts (see also the Guide to Publishing with Elsevier). See also the section on Electronic artwork.
To avoid unnecessary errors you are strongly advised to use the 'spell-check' and 'grammar-check' functions of your word processor.

Article structure

Subdivision - unnumbered sections
Divide your article into clearly defined sections. Each subsection is given a brief heading. Each heading should appear on its own separate line. Subsections should be used as much as possible when cross-referencing text: refer to the subsection by heading as opposed to simply 'the text'.

Introduction
State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.

Material and methods
Provide sufficient detail to allow the work to be reproduced. Methods already published should be indicated by a reference: only relevant modifications should be described.

Results
Results should be clear and concise.

Discussion
This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

Conclusions
The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section.

Appendices
If there is more than one appendix, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Fig. A.1, etc.

Essential title page information

• ***Title.*** Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible.
• ***Author names and affiliations.*** Please clearly indicate the given name(s) and family name(s) of each author and check that all names are accurately spelled. You can add your name between parentheses in your own script behind the English transliteration, e.g. "Y. Wang (###)". Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower-case superscript letter immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name and, if available, the e-mail address of each author.
• ***Corresponding author.*** Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. **Ensure that the e-mail address is given and that contact details are kept up to date by the corresponding author.**
• ***Present/permanent address.*** If an author has moved since the work described in the article was done, or was visiting at the time, a 'Present address' (or 'Permanent address') may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

Structured abstract

A structured abstract, by means of appropriate headings, should provide the context or background for the research and should state its purpose, basic procedures (selection of study subjects or laboratory animals, observational and analytical methods), main findings (giving specific effect sizes and their statistical significance, if possible), and principal conclusions. It should emphasize new and important aspects of the study or observations.

A list of three to six keywords should be supplied: full instructions are provided when submitting the article online.

Abbreviations
Define abbreviations that are not standard in this field in a footnote to be placed on the first page of the article. Such abbreviations that are unavoidable in the abstract must be defined at their first mention there, as well as in the footnote. Ensure consistency of abbreviations throughout the article.

Acknowledgements
Collate acknowledgements in a separate section at the end of the article before the references and do not, therefore, include them on the title page, as a footnote to the title or otherwise. List here those individuals who provided help during the research (e.g., providing language help, writing assistance or proof reading the article, etc.).

Formatting of funding sources
List funding sources in this standard way to facilitate compliance to funder's requirements:

Funding: This work was supported by the National Institutes of Health [grant numbers xxxx, yyyy]; the Bill & Melinda Gates Foundation, Seattle, WA [grant number zzzz]; and the United States Institutes of Peace [grant number aaaa].

It is not necessary to include detailed descriptions on the program or type of grants and awards. When funding is from a block grant or other resources available to a university, college, or other research institution, submit the name of the institute or organization that provided the funding.

If no funding has been provided for the research, please include the following sentence:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Units
Follow internationally accepted rules and conventions: use the international system of units (SI). If other units are mentioned, please give their equivalent in SI.

Math formulae
Please submit math equations as editable text and not as images. Present simple formulae in line with normal text where possible and use the solidus (/) instead of a horizontal line for small fractional terms, e.g., X/Y. In principle, variables are to be presented in italics. Powers of e are often more conveniently denoted by exp. Number consecutively any equations that have to be displayed separately from the text (if referred to explicitly in the text).

Footnotes
Footnotes should be used sparingly. Number them consecutively throughout the article. Many word processors build footnotes into the text, and this feature may be used. Should this not be the case, indicate the position of footnotes in the text and present the footnotes themselves separately at the end of the article.

Artwork

Electronic artwork
General points
• Make sure you use uniform lettering and sizing of your original artwork.
• Preferred fonts: Arial (or Helvetica), Times New Roman (or Times), Symbol, Courier.
• Number the illustrations according to their sequence in the text.
• Use a logical naming convention for your artwork files.
• Indicate per figure if it is a single, 1.5 or 2-column fitting image.
• For Word submissions only, you may still provide figures and their captions, and tables within a single file at the revision stage.
• Please note that individual figure files larger than 10 MB must be provided in separate source files. A detailed guide on electronic artwork is available.
You are urged to visit this site; some excerpts from the detailed information are given here.
Formats
Regardless of the application used, when your electronic artwork is finalized, please 'save as' or convert the images to one of the following formats (note the resolution requirements for line drawings, halftones, and line/halftone combinations given below):
EPS (or PDF): Vector drawings. Embed the font or save the text as 'graphics'.
TIFF (or JPG): Color or grayscale photographs (halftones): always use a minimum of 300 dpi.
TIFF (or JPG): Bitmapped line drawings: use a minimum of 1000 dpi.
TIFF (or JPG): Combinations bitmapped line/half-tone (color or grayscale): a minimum of 500 dpi is required.
Please do not:
• Supply files that are optimized for screen use (e.g., GIF, BMP, PICT, WPG); the resolution is too low.
• Supply files that are too low in resolution.
• Submit graphics that are disproportionately large for the content.

Color artwork
Please make sure that artwork files are in an acceptable format (TIFF (or JPEG), EPS (or PDF), or MS Office files) and with the correct resolution. If, together with your accepted article, you submit usable color figures then Elsevier will ensure, at no additional charge, that these figures will appear in color online (e.g., ScienceDirect and other sites) regardless of whether or not these illustrations are reproduced in color in the printed version. **For color reproduction in print, you will receive**

information regarding the costs from Elsevier after receipt of your accepted article. Please indicate your preference for color: in print or online only. Further information on the preparation of electronic artwork.

Illustration services
Elsevier's WebShop offers Illustration Services to authors preparing to submit a manuscript but concerned about the quality of the images accompanying their article. Elsevier's expert illustrators can produce scientific, technical and medical-style images, as well as a full range of charts, tables and graphs. Image 'polishing' is also available, where our illustrators take your image(s) and improve them to a professional standard. Please visit the website to find out more.

Figure captions
Ensure that each illustration has a caption. A caption should comprise a brief title (**not** on the figure itself) and a description of the illustration. Keep text in the illustrations themselves to a minimum but explain all symbols and abbreviations used.

Tables

Please submit tables as editable text and not as images. Tables can be placed either next to the relevant text in the article, or on separate page(s) at the end. Number tables consecutively in accordance with their appearance in the text and place any table notes below the table body. Be sparing in the use of tables and ensure that the data presented in them do not duplicate results described elsewhere in the article. Please avoid using vertical rules and shading in table cells.

References

Citation in text
Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Any references cited in the abstract must be given in full. Unpublished results and personal communications are not recommended in the reference list, but may be mentioned in the text. If these references are included in the reference list they should follow the standard reference style of the journal and should include a substitution of the publication date with either 'Unpublished results' or 'Personal communication'. Citation of a reference as 'in press' implies that the item has been accepted for publication.

Reference links
Increased discoverability of research and high quality peer review are ensured by online links to the sources cited. In order to allow us to create links to abstracting and indexing services, such as Scopus, CrossRef and PubMed, please ensure that data provided in the references are correct. Please note that incorrect surnames, journal/book titles, publication year and pagination may prevent link creation. When copying references, please be careful as they may already contain errors. Use of the DOI is encouraged.

A DOI can be used to cite and link to electronic articles where an article is in-press and full citation details are not yet known, but the article is available online. A DOI is guaranteed never to change, so you can use it as a permanent link to any electronic article. An example of a citation using DOI for an article not yet in an issue is: VanDecar J.C., Russo R.M., James D.E., Ambeh W.B., Franke M. (2003). Aseismic continuation of the Lesser Antilles slab beneath northeastern Venezuela. Journal of Geophysical Research, https://doi.org/10.1029/2001JB000884. Please note the format of such citations should be in the same style as all other references in the paper.

Web references
As a minimum, the full URL should be given and the date when the reference was last accessed. Any further information, if known (DOI, author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list.

Data references
This journal encourages you to cite underlying or relevant datasets in your manuscript by citing them in your text and including a data reference in your Reference List. Data references should include the following elements: author name(s), dataset title, data repository, version (where available), year, and global persistent identifier. Add [dataset] immediately before the reference so we can properly identify it as a data reference. The [dataset] identifier will not appear in your published article.

Reference management software
Most Elsevier journals have their reference template available in many of the most popular reference management software products. These include all products that support Citation Style Language styles, such as Mendeley and Zotero, as well as EndNote. Using the word processor plug-ins from these products, authors only need to select the appropriate journal template when preparing their article, after which citations and bibliographies will be automatically formatted in the journal's style. If no template is yet available for this journal, please follow the format of the sample references and citations as shown in this Guide.

Users of Mendeley Desktop can easily install the reference style for this journal by clicking the following link:
http://open.mendeley.com/use-citation-style/respiratory-medicine
When preparing your manuscript, you will then be able to select this style using the Mendeley plug-ins for Microsoft Word or LibreOffice.

Reference formatting
There are no strict requirements on reference formatting at submission. References can be in any style or format as long as the style is consistent. Where applicable, author(s) name(s), journal title/book title, chapter title/article title, year of publication, volume number/book chapter and the pagination must be present. Use of DOI is highly encouraged. The reference style used by the journal will be applied to the accepted article by Elsevier at the proof stage. Note that missing data will be highlighted at proof stage for the author to correct. If you do wish to format the references yourself they should be arranged according to the following examples:

Reference style
Text: Indicate references by number(s) in square brackets in line with the text. The actual authors can be referred to, but the reference number(s) must always be given.
Example: '..... as demonstrated [3,6]. Barnaby and Jones [8] obtained a different result ....'
List: Number the references (numbers in square brackets) in the list in the order in which they appear in the text.
Examples:
Reference to a journal publication:
[1] J. van der Geer, J.A.J. Hanraads, R.A. Lupton, The art of writing a scientific article, J. Sci. Commun. 163 (2010) 51–59.
Reference to a book:
[2] W. Strunk Jr., E.B. White, The Elements of Style, fourth ed., Longman, New York, 2000.
Reference to a chapter in an edited book:
[3] G.R. Mettam, L.B. Adams, How to prepare an electronic version of your article, in: B.S. Jones, R.Z. Smith (Eds.), Introduction to the Electronic Age, E-Publishing Inc., New York, 2009, pp. 281–304.
Reference to a website:
[4] Cancer Research UK, Cancer statistics reports for the UK. http://www.cancerresearchuk.org/aboutcancer/statistics/cancerstatsreport/, 2003 (accessed 13 March 2003).
Reference to a dataset:
[dataset] [5] M. Oguro, S. Imahiro, S. Saito, T. Nakashizuka, Mortality data for Japanese oak wilt disease and surrounding forest compositions, Mendeley Data, v1, 2015. https://doi.org/10.17632/xwj98nb39r.1.

Journal abbreviations source
Journal names should be abbreviated according to the List of Title Word Abbreviations.

Video
Elsevier accepts video material and animation sequences to support and enhance your scientific research. Authors who have video or animation files that they wish to submit with their article are strongly encouraged to include links to these within the body of the article. This can be done in the same way as a figure or table by referring to the video or animation content and noting in the body text where it should be placed. All submitted files should be properly labeled so that they directly relate to the video file's content. In order to ensure that your video or animation material is directly usable, please provide the files in one of our recommended file formats with a preferred maximum size of 150 MB in total. Any single file should not exceed 50 MB. Video and animation files supplied will be published online in the electronic version of your article in Elsevier Web products, including ScienceDirect. Please supply 'stills' with your files: you can choose any frame from the video or animation or make a separate image. These will be used instead of standard icons and will personalize the link to your video data. For more detailed instructions please visit our video instruction pages.

Note: since video and animation cannot be embedded in the print version of the journal, please provide text for both the electronic and the print version for the portions of the article that refer to this content.

Supplementary material

Supplementary material such as applications, images and sound clips, can be published with your article to enhance it. Submitted supplementary items are published exactly as they are received (Excel or PowerPoint files will appear as such online). Please submit your material together with the article and supply a concise, descriptive caption for each supplementary file. If you wish to make changes to supplementary material during any stage of the process, please make sure to provide an updated file. Do not annotate any corrections on a previous version. Please switch off the 'Track Changes' option in Microsoft Office files as these will appear in the published version.

RESEARCH DATA

This journal encourages and enables you to share data that supports your research publication where appropriate, and enables you to interlink the data with your published articles. Research data refers to the results of observations or experimentation that validate research findings. To facilitate reproducibility and data reuse, this journal also encourages you to share your software, code, models, algorithms, protocols, methods and other useful materials related to the project.

Below are a number of ways in which you can associate data with your article or make a statement about the availability of your data when submitting your manuscript. If you are sharing data in one of these ways, you are encouraged to cite the data in your manuscript and reference list. Please refer to the "References" section for more information about data citation. For more information on depositing, sharing and using research data and other relevant research materials, visit the research data page.

Data linking

If you have made your research data available in a data repository, you can link your article directly to the dataset. Elsevier collaborates with a number of repositories to link articles on ScienceDirect with relevant repositories, giving readers access to underlying data that gives them a better understanding of the research described.

There are different ways to link your datasets to your article. When available, you can directly link your dataset to your article by providing the relevant information in the submission system. For more information, visit the database linking page.

For supported data repositories a repository banner will automatically appear next to your published article on ScienceDirect.

In addition, you can link to relevant data or entities through identifiers within the text of your manuscript, using the following format: Database: xxxx (e.g., TAIR: AT1G01020; CCDC: 734053; PDB: 1XFN).

Mendeley Data

This journal supports Mendeley Data, enabling you to deposit any research data (including raw and processed data, video, code, software, algorithms, protocols, and methods) associated with your manuscript in a free-to-use, open access repository. Before submitting your article, you can deposit the relevant datasets to *Mendeley Data*. Please include the DOI of the deposited dataset(s) in your main manuscript file. The datasets will be listed and directly accessible to readers next to your published article online.

For more information, visit the Mendeley Data for journals page.

Data statement

To foster transparency, we encourage you to state the availability of your data in your submission. This may be a requirement of your funding body or institution. If your data is unavailable to access or unsuitable to post, you will have the opportunity to indicate why during the submission process, for example by stating that the research data is confidential. The statement will appear with your published article on ScienceDirect. For more information, visit the Data Statement page.

AudioSlides

The journal encourages authors to create an AudioSlides presentation with their published article. AudioSlides are brief, webinar-style presentations that are shown next to the online article on ScienceDirect. This gives authors the opportunity to summarize their research in their own words and to help readers understand what the paper is about. More information and examples are available. Authors of this journal will automatically receive an invitation e-mail to create an AudioSlides presentation after acceptance of their paper.

AFTER ACCEPTANCE

Online proof correction

Corresponding authors will receive an e-mail with a link to our online proofing system, allowing annotation and correction of proofs online. The environment is similar to MS Word: in addition to editing text, you can also comment on figures/tables and answer questions from the Copy Editor. Web-based proofing provides a faster and less error-prone process by allowing you to directly type your corrections, eliminating the potential introduction of errors.

If preferred, you can still choose to annotate and upload your edits on the PDF version. All instructions for proofing will be given in the e-mail we send to authors, including alternative methods to the online version and PDF.

We will do everything possible to get your article published quickly and accurately. Please use this proof only for checking the typesetting, editing, completeness and correctness of the text, tables and figures. Significant changes to the article as accepted for publication will only be considered at this stage with permission from the Editor. It is important to ensure that all corrections are sent back to us in one communication. Please check carefully before replying, as inclusion of any subsequent corrections cannot be guaranteed. Proofreading is solely your responsibility.

Offprints

The corresponding author will, at no cost, receive a customized Share Link providing 50 days free access to the final published version of the article on ScienceDirect. The Share Link can be used for sharing the article via any communication channel, including email and social media. For an extra charge, paper offprints can be ordered via the offprint order form which is sent once the article is accepted for publication. Both corresponding and co-authors may order offprints at any time via Elsevier's Webshop. Corresponding authors who have published their article open access do not receive a Share Link as their final published version of the article is available open access on ScienceDirect and can be shared through the article DOI link.

AUTHOR INQUIRIES

Visit the Elsevier Support Center to find the answers you need. Here you will find everything from Frequently Asked Questions to ways to get in touch.

You can also check the status of your submitted article or find out when your accepted article will be published.