



**University of
Nottingham**
UK | CHINA | MALAYSIA

**Quantitative measurement of intracellular
metabolic changes in *Clostridium*
autoethanogenum using liquid chromatography
isotope dilution mass spectrometry**

Laudina Kyeiwah Safo, M.Chem (Hons.)

Thesis submitted to the University of Nottingham for the
degree of Doctor of Philosophy

January 2018

Abstract

Clostridium autoethanogenum is an important organism for biofuel production. Other 'omics' approaches have been used to understand the mode of operation of the organism but metabolomics gives information on the cellular activities in the cell. Metabolomics combined with other 'omics' data can provide a deeper understanding for pathway interpretation.

This project sets out to develop an analytical method that is suitable for analysis of highly charged polar compounds found in *C. autoethanogenum* metabolic pathways. Also investigate suitable isotope labelled internal standards to improve matrix effects to the metabolites as a result of the biological matrix.

A high-throughput hydrophilic interaction liquid chromatography isotope dilution mass spectrometry (HILIC-IDMS) was developed and validated using high resolution hybrid orbital trap MS for both targeted and untargeted metabolomics analysis of intracellular metabolic pathways of *Clostridium autoethanogenum*. Extraction of intracellular metabolites from *C. autoethanogenum* was achieved using a specifically developed sample preparation protocol using freeze thaw cycles (freeze in liquid nitrogen and thaw on ice repeated 3x). A total of 133 metabolites were monitored and validated. Limits of detection (LODs) ranged from 0.001 μM to 5 μM reported for compounds such as NADPH and NADH. Limits of quantification (LOQs) for all metabolites ranged from 0.001 μM to 10 μM for metabolites such as glucose-6-phosphate and glyceraldehyde-3 phosphate. Precision and accuracy were evaluated for all metabolites and found to be within the acceptable limits of $\pm 15\%$ with few exceptions for some nucleotides and organic acids. Stable isotopically labelled internal standards were generated from *C. pasteurianum* cells that provided coverage for about 100 metabolites. This enabled absolute intracellular concentrations to be obtained in combination with the estimated cell volume of *C. autoethanogenum* that was obtained from microscopy and flow cytometry measurements.

The developed HILIC-IDMS method was applied to various solvent production

optimisation experiments conducted using *C. autoethanogenum* and the main findings are reported below.

In chapter 4, the HILIC-IDMS method was applied to *C. autoethanogenum* in an experiment where the pH of the media was reduced to improve ethanol and solvent production. The metabolomics studies of this experiment gave intracellular concentrations that differentiated the acidogenic phase from the solventogenic phase. A total of 86 metabolites were quantified in this experiment. Intermediates in the tricarboxylic acid (TCA) cycle were the most affected during the acidogenic/solventogenic transition. Metabolites concentrations were used for metabolic pathways analysis to understand the pathways affected during the pH shift. The pathway analysis also confirmed the TCA cycle was the most affected pathway during the acidogenic/solventogenic transition.

In chapter 5, The HILIC-IDMS method was applied to a gas shift experiment to optimise ethanol and solvent production. Gas shift is another approach that can be used to optimise solvent production in similar as the pH shift experiment. The use of gas shift to induce solventogenic phase can be difficult as *C. autoethanogenum* has little tolerance for high levels of CO hence the increase in gas (CO) flow rate has to be done in a gradual fashion. Equally, TCA intermediates were observed to be the most affected as observed in the pH shift experiment.

In chapter 6, the method was applied to a study where pantothenate and phosphate concentrations in the growth media of *C. autoethanogenum* were reduced to increase ethanol production. Pantothenate is the precursor for coenzyme A (CoA) production and metabolomics study confirmed a decrease in CoA concentration when pantothenate concentration was reduced. Metabolomics also showed decrease in concentration of metabolites directly linked to CoA synthesis such as L- Aspartate. Metabolic pathway analysis also confirmed the pantothenate and CoA biosynthesis and its associated pathways were the most affected pathways during the pantothenate-limiting phase.

Both targeted and untargeted metabolomics analysis were performed on these

nutrient-limiting experiments and there were clear differences between the two different conditions before and after nutrient limitation. Supervised multivariate data analysis using OPLS-DA was used to compare higher pantothenate and low pantothenate concentration and there were clear separation and clustering between the two conditions. Cross validation obtained for R^2Y and Q^2 were 0.993 and 0.941 respectively. OPLS-DA plots for phosphate limitation also showed clustering and separation between the high phosphate concentration and reduced phosphate concentration with R^2Y and Q^2 0.981 and 0.837 respectively.

In conclusion, a novel high-throughput HILIC-IDMS method was developed and validated for analysis of different classes of polar compounds in bacteria. The method has the potential to be applied in other biological matrices for coverage of diverse range of polar compounds.

Dedication

This Thesis is dedicated to my lovely son

Tega Emonena Eserode Jedidiah Odjesa

Acknowledgements

I would like to acknowledge everyone who has made this PhD journey a success. Firstly, I would like to thank my supervisor Prof David Barrett for his supervision, guidance, patience and advice throughout the years. I would also like to thank my second supervisor Dr Dong-Hyun Kim and my third and fourth supervisors Dr Klaus Winzer and Prof Neil Thomas for their support. I would like to show gratitude to Dr Catherine Ortori for her invaluable technical and practical assistance. I would like to thank Mr Paul Cooling for HPLC training and technical support. I acknowledge past and present postdoctoral fellows on the GASCHEM project; Dr Sarah Schatschneider and Dr Salah Abdelrazig for their help and support during my studies. Dr Salah Abdelrazig especially for his patience and directions.

I express my thanks to the entire members of the Centre for Analytical Bioscience (CAB) for their support, help, advice and making the laboratory a lovely place to work. I wish to also thank the Synthetic Biology Research Centre based in CBS especially Dr Alex Grosse-Honebrink for helping the work on generating internal standards, Dr Anne Henstra for bioreactor experiments, Dr Thomas Millat and Rupert Norman for Escher Pathway analysis and Dr Bart Pander and Florence Annan for their interest in collaborating.

Finally, I would like to express my gratitude to my family, especially my mother Miss Catherine Arthur for her support in babysitting Tega during aspects of my studies. Also, to my fiancé, Enabulele Courage-Ogbebor, for his support and encouragement during the difficult times and my son Tega for being the motivation to succeed.

Contributions to conferences and workshops during the PhD

- Poster presentation at the American society of Mass Spectrometry (ASMS) annual conference held in Indianapolis, USA in June 2017
- Oral presentation of my research in seminar organised by the School of Pharmacy, University of Nottingham

Contents

Abstract	I
Dedication	IV
Acknowledgements	V
Contributions to conferences and workshops during the PhD	VI
Contents	VII
List of tables and figures.....	XIII
CHAPTER ONE.....	1
1. General introduction	1
1.1 Introduction.....	1
1.1.1 The Global energy crisis.....	1
1.1.2 The use of acetogens to produce fuels and chemicals	1
1.1.3 Fermentation of gas through the one carbon metabolic pathway (Wood-Ljungdahl pathway)	2
1.1.4 Metabolic pathway engineering of <i>C. autoethanogenum</i>	3
1.1.5 Metabolomics analysis of diverse polar compounds.....	6
1.1.6 Metabolomics analysis and its application in bacteria	6
1.1.7 Different analytical techniques used in metabolomics applications	9
1.1.8 LC-MS based metabolomics.....	11
1.1.9 Mass spectrometry techniques	12
1.1.10 Data processing in LC-MS based metabolomics	17
1.1.11 Multivariate analysis of MS-based metabolomics	18
1.1.12 General considerations for development of quantitative metabolomics analysis	20
1.1.13 General project aims and objectives	2
CHAPTER TWO.....	2
2. Method Development 1: investigation of LC-MS/MS methods for the quantitative analysis of intracellular metabolites in one carbon and associated pathways	25
2.1 Introduction.....	25
2.1.1 General background	25
2.1.2 Mass spectrometry applications in targeted metabolomics analysis..	26

2.1.3	Analytical challenges in method development for analysis of polar metabolites.....	27
2.1.4	Background of HILIC method.....	28
2.1.5	Project Aims and objectives.....	33
2.2	Materials and Methods	33
2.2.1	Experimental reagents and methods.....	33
2.2.2	Preparation of standards for calibration and direct MS infusion	33
2.2.3	LC Method development conditions	34
2.2.4	Optimised analysis of folates using HILIC-MS/MS	36
2.2.5	Preliminary method development for analysis of folates	37
2.2.6	Optimisation of MS Conditions.....	38
2.2.7	Data analysis for LC-MS/MS raw data for tetrahydrofolates and key metabolites.....	38
2.3	Results and discussion	39
2.3.1	Analytical strategies.....	39
2.3.2	Method development of folates and poly-glutamated folates	50
2.4	Conclusion	56
CHAPTER THREE		57
3.	Method development 2: Approaches for absolute quantification of intracellular metabolites present in <i>Clostridium autoethanogenum</i> metabolic pathways	57
3.1	Introduction.....	57
3.1.1	Limitations of LC-MS for quantitative analysis of complex biological matrices	57
3.1.2	Generation of multiple stable isotope IS from cells grown on ¹³ C source	58
3.1.3	Determining cell numbers and cell volume using flow cytometry	59
3.1.4	Evaluation of cell volume from microscopy images	59
3.1.5	Aims of this chapter	60
3.2	Materials and methods.....	60
3.2.1	Chemicals and reagent	60
3.2.2	Preparation of Calibration Samples from Stock Solutions.....	61
3.2.3	Preliminary HILIC method development	62
3.2.4	Extraction protocol optimisation for <i>C. autoethanogenum</i> cell lysis...	62
3.2.5	Exploring <i>E. coli</i> as a candidate organism for generation of U- ¹³ C-labelled internal standard	63

3.2.6	<i>C. pasteurianum</i> culture conditions to generate U- ¹³ C- labelled internal standards	64
3.2.7	Extraction of ¹³ C labelled isotope internal standards from <i>C. pasteurianum</i>	64
3.2.8	Growth media and bioreactor cultivation of <i>C. autoethanogenum</i>	65
3.2.9	Growth rate measurements using OD ₆₀₀	65
3.2.10	Sample preparation for microscopy	65
3.2.11	Microscopy Imaging.....	66
3.2.12	Flow Cytometry (FCM) Measurements for Cell Counting.....	67
3.2.13	LC-MS data analysis for identification of 133 metabolites	67
3.2.14	Method validation	68
3.3	Results and Discussion.....	69
3.3.1	Optimisation of extraction protocol for cell lysis of <i>C. autoethanogenum</i>	69
3.3.2	Stability test for extraction solvents.....	72
3.3.3	Evaluation of Matrix effect of key metabolites after extraction	73
3.3.4	HILIC-IDMS method development and validation	82
3.3.5	Selection of organism to generate uniformly ¹³ C-labelled internal standards	91
3.3.6	Optimisation of U- ¹³ C-labelled internal standard generation using <i>C. pasteurianum</i>	92
3.3.7	Cell number and viability	100
3.3.8	Determination of cell volume of <i>C. autoethanogenum</i>	103
3.4	Conclusion	105
CHAPTER FOUR.....		107
4.	Application of HILIC-IDMS method to determine the effect of pH on ethanol production in <i>C. autoethanogenum</i>	107
4.1	Introduction.....	107
4.1.1	The role of metabolomics in explaining the acidogenesis-solventogenesis transition in <i>C. autoethanogenum</i>	108
4.1.2	Aims and Objectives	109
4.2	Materials and methods.....	110
4.2.1	Reagents and materials	110
4.2.2	Culture Conditions for <i>Clostridium autoethanogenum</i> for Gas and pH Shift Experiments	110
4.2.3	Product shift by pH	111

4.2.4	Sample Collection and Extraction of Metabolites from <i>C. autoethanogenum</i>	111
4.2.5	Calculation of intracellular concentration using cell volume.....	112
4.2.6	LC-MS and data analysis	112
4.2.7	Transcriptomics analysis.....	113
4.3	Results and discussion	114
4.3.1	Targeted metabolomics approach for analysis of pH shift experiment in <i>C. autoethanogenum</i>	114
4.3.2	Metabolic pathway analysis and mapping	122
4.3.3	Effect of lowering pH in the media on conversion of acetate to ethanol 130	
4.4	Conclusion	132
CHAPTER FIVE.....		132
5.	Application of HILIC-IDMS method to determine the effect of increasing carbon monoxide flow on ethanol production in <i>C. autoethanogenum</i>	134
5.1	Introduction.....	134
5.1.1	Trends in omics studies in acidogenic-solventogenic transition in <i>Clostridia species</i>	135
5.1.2	Aims and Objectives	136
5.2	Materials and methods.....	136
5.2.1	Reagents and materials	136
5.2.2	Culture conditions for <i>C. autoethanogenum</i> for gas shift experiment.....	136
5.2.3	Sample collection and extraction of <i>C. autoethanogenum</i> cell pellets 137	
5.2.4	LC-MS and data analysis	137
5.2.5	Transcriptomics data analysis.....	138
5.2.6	Measurement of intracellular concentration of acetate and ethanol 138	
5.3 Results and discussion		139
5.3.1	Metabolomics analysis of cell pellets from gas shift experiment.....	139
5.3.2	Gas shift metabolomics data analysis using MetaboAnalyst 3.0.....	145
5.3.3	Metabolic pathway analysis of changes in the one carbon metabolic pathways during the gas shift experiment	149
5.3.4	Metabolic pathway analysis and mapping	152
5.3.5	Using metabolomics to understand the effect of increasing gas flow on ethanol production.....	154

5.3.6	Effects of gas shift in <i>C. autoethanogenum</i> on folates (Wood-Ljungdahl pathway intermediates)	159
5.3.7	Statistical and pathway analysis for intracellular folate concentration in <i>C. autoethanogenum</i> during gas shift experiment	161
5.4	Conclusion	164
CHAPTER SIX		164
6.	Application of HILIC-IDMS method to study the effects of pantothenate limitation on the metabolism of <i>Clostridium autoethanogenum</i>	165
6.1	Introduction	165
6.1.1	Pantothenate and Phosphate limitation in <i>C. autoethanogenum</i>	165
6.1.2	The use of metabolomics approaches in understanding the role of pantothenate in ethanol production	168
6.1.3	Aims and Objectives	168
6.2	Materials and Methods	169
6.2.1	Bioreactor operation parameters	169
6.2.2	Composition of B vitamin solution	170
6.2.3	Metabolomics sampling	171
6.2.4	Extraction of intracellular metabolites from <i>C. autoethanogenum</i>	171
6.2.5	LC-MS and data analysis	171
6.3	Results and discussion	175
6.3.1	Targeted metabolomics analysis of pantothenate limitation in <i>C. autoethanogenum</i>	175
6.3.2	Untargeted metabolomics data analysis	189
6.3.3	Pathway analysis using metabolite identification from Compound discoverer 2.1	195
6.3.4	Using metabolomics to understand the effect of pantothenate limitation on ethanol production	198
6.3.5	The use of HILIC-IDMS method to investigate phosphate limitation in <i>C. autoethanogenum</i>	199
6.3.6	Data analysis and pathway analysis using MetaboAnalyst 3.0	205
6.3.7	Multivariate data analysis using SIMCA-P-14 for phosphate limitation experiment	210
6.4	Conclusion	212
CHAPTER SEVEN		213
7.	Conclusions and Future work	213
7.1	Method development and validation for microbial metabolomics application	213

7.2	Considerations for quantitative metabolomics analysis.....	214
7.3	Applications of the developed HILIC-IDMS method to metabolomics analysis in <i>C. autoethanogenum</i>	216
7.4	Future Work.....	217
CHAPTER EIGHT		219
8	References.....	219
APPENDICES.....		235

List of tables and figures

Figure 1.1 The Wood-Ljungdahl pathway showing the Eastern and Western branches as depicted in acetogens. Carbon labelled red forming the methyl group of acetyl-CoA was generated from the CO ₂ from the eastern branch and the carbonyl group in acetyl-CoA from the CO ₂ from the western branch (Ragsdale and Pierce, 2008). H ₄ folate (tetrahydrofolate), HCO- H ₄ folate (formyl tetrahydrofolate), CH ₂ = H ₄ folate (methylene-tetrahydrofolate), CH ⁺ = H ₄ folate (Methenyl-tetrahydrofolate) and CH ₃ -H ₄ folate (Methyl-tetrahydrofolate). H ₄ folate has been proposed to be involved in CO ₂ fixation pathway (Ljungdahl et al., 1966).	3
Figure 1.2 One carbon metabolic pathway of C autoethanogenum starting from Wood-Ljungdahl pathway to the production of acetate, ethanol and 2,3-butanediol (23BD). The production of ethanol, acetate and 23BD are natural to the organism while production of isobutanol and succinate are non natural pathways that will be introduced by systems biology. Acetyl CoA is the precursor for ethanol and acetate production while pyruvate is the precursor for the production of 23BD.	5
Figure 1.3 A representation of the different approaches used in untargeted and targeted metabolomics analysis work flow (Heuberger et al., 2014).	9
Figure 1.4 Bar graphs showing the comparison between MS based metabolomics and NMR based metabolomics over the past 10 years. These data were generated using Web of science database.	11
Figure 1.5 A schematic showing ion formation in electrospray ionisation technique. (Challamalla et al., 2012).	13
Figure 1.6 Multiple reaction monitoring (MRM) by triple quadrupole (QqQ) mass spectrometer. Predetermined parent (precursor ions selected for MRM by QqQ mass analysis are shown as colored symbols in ion source. First (quadrupole, Q1 (MS ¹ level) selectively allows only the targeted parent ions to pass into the second quadrupole (Q2) for collision-induced dissociation (CID) fragmentation. Third quadrupole, (Q3) (MS ² level) selectively allows only specific pre-identified fragment ions to pass through to the detector. (Pisitkun et al., 2007).	14
Figure 1.7 This is a cross section of the C-trap and the Orbital trap analyser. The rf-only bent quadrupole of the C-trap stores ions; the rf is ramped down followed application of a high voltage pulse across the trap, with each m/z being emitted in a short packet. During voltage ramp the packets from the C-trap enter the analyser and spread into the oscillating rings inducing a current detected by the differential amplifier (Zubarev and Makarov, 2013)	17
Figure 2.1 An overview of a typical work flow for quantitative metabolic pathway profiling (Shen et al., 2016).	26
Figure 2.2 initial selections of metabolites for LC-MS/MS method optimisation. Acetaldehyde and isobutyraldehyde could not be optimised using LC-MS/MS due to their volatile nature.	39

Figure 2.3 Chromatograms showing separation obtained when 10 mM DMHA and TBA incorporated into the mobile phases using representative metabolites.[1-lactate, 2-malate,3-Fumarate, 4-citrate, 5- pyruvate]	42
Figure 2.4 Recovery analysis of folates in different buffers. (MW- methanol/water, M-EtOH- 2-mercaptoethanol, AA- ascorbic acid, DHF-dihydrofolate, THF-tetrahydrofolate) (n=2).	51
Figure 2.5 Analysis of polyglutamated folates present in <i>C. autoethanogenum</i> . In all cases the residues with three polyglutamate was the most abundant form found in <i>C. autoethanogenum</i> . (Glu-glutamic acid residues)	52
Figure 2.6 A representative chromatogram of tetrahydrofolate (H4PteGlu) and the polyglutamated forms up to four glutamic acid residues found in <i>C. autoethanogenum</i>	52
Figure 3.1 Microscopy images of <i>C. autoethanogenum</i> stained with syto9 and propidium iodide before (A) and after (B) freeze thaw cycle to confirm cell lysis. Live cells fluoresce green to indicate intact cells as shown in A was used as control. Cells were stained after extraction to confirm cell lysis. Dead cells fluoresce red due to compromised cell walls as shown in B indication complete cell lysis during freeze thaw extraction.	70
Figure 3.2 Percentage recovery of analytes using the homogenizer and freeze thaw extraction method analysed on the HILIC method. [FA- Fumaric acid, LA- Lactic acid, MA-Malic acid, CA- Citric acid, GA- Glyoxylic acid, 2KIA- 2-Ketoisovaleric acid, PA-Pyruvic acid, SA- Succinic acid, ISA- Isocitric acid].....	71
Figure 3.3 Recovery analysis of representative metabolites in different solvent combinations for optimisation of cell lysis of <i>C. autoethanogenum</i> (CMW-chloroform:methanol:water(2:1:1),MW-methanol:water(1:1), AMW- acetonitrile :methanol :water (2:2:1)). (n=2).	72
Figure 3.4 Stability analysis of representative metabolites in different solvent combinations for optimisation of cell lysis of <i>C. autoethanogenum</i> (CMW-chloroform:methanol:water(2:1:1),MW-methanol:water(1:1), AMW- acetonitrile :methanol :water (2:2:1)). (n=2).	73
Figure 3.5 Matrix effect of analytes using the homogenizer and freeze thaw extraction method analysed on the HILIC method. (A) represents matrix effect of compounds spiked at lower concentration of 125 μ M (B) represents matrix effect of compounds spiked at higher concentration of 250 μ M [FA- Fumaric acid, LA- Lactic acid, MA-Malic acid, CA- Citric acid, GA- Glyoxylic acid, 2KIA- 2-Ketoisovaleric acid, PA-Pyruvic acid, SA- Succinic acid, ISA- Isocitric acid].....	74
Figure 3.6 comparison of intracellular metabolites of <i>C. autoethanogenum</i> with <i>E. coli</i> . <i>A. platensis</i> and <i>C. pasteurianum</i> to select best organism for generation of U- ¹³ C-labelled internal standard. 424 metabolites were common between <i>E.coli</i> and <i>C. autoethanogenum</i> , 272 metabolites common between <i>A. platensis</i> and <i>C. autoethanogenum</i> and 873 metabolites common between <i>C. pasteurianum</i> and <i>C. autoethanogenum</i> . Identification of metabolites were carried out using Progenesis Q1 and using q-values to show significance.	92
Figure 3.7 A plot of OD measurements over 24 hour period of <i>C. pasteurianum</i> cell culture. The plot shows comparison between CaCO ₃ and MES as buffers for culturing of <i>C. pasteurianum</i> with and without the addition of biotin. The presence of biotin supports growth much better than omitting it and also CaCO ₃ was a better buffer compared to MES.	93

<i>Figure 3.8 Effect of glucose concentrations ranging from 0.5 % to 6 % on the growth of C. pasteurianum. Optical density measured at 600 nm (OD₆₀₀) was used to determine cell mass for each experiment.</i>	95
Figure 3.9 Flow cytometry measurements showing the presence of four different populations. Four Populations (Quadrants), R39 Low Side Scatter Low Forward Scatter (33%), R40 Low Side Scatter High Forward Scatter (38%), R37 High Side Scatter Low Forward Scatter (13%), R38 High Side Scatter High Forward Scatter (17%).	101
Figure 3.10 Bac Light Viability Gated on all Syto9 positive events (i.e. excluding debris and abiotic bacteria).	102
Figure 3.11 This histogram of cell volumes as determined by microscopy shows a skewed distribution where most cell volumes are 0.8-5 μm^3 . Frequencies of both unmodified pictures and deconvoluted pictures are given.	104
Figure 3.12 A: Calibration Curve of References Size Standards to Log 488nm Side Scatter. Log side scatter area measurement of Sphero Nano Fluorescent Particle Standard kit (Spherotech Inc) vs reported diameters. Line represents linear curve fit performed by Prism (V7.01 GraphPad). Relationship = $Y = 0.001376 * X + 5.789$. B: Correlation of devolution microscopy volume measurement and normalised side scatter area flow cytometry measurement from 4 independent reactors. Each dot represents the median 10-percentile value from an individual reactor measured by both techniques. Line represents exponential non-linear curve fit performed by Prism (V7.01 GraphPad). Relationship: $\text{Volume} = 0.3844 * \exp(0.002568 * \text{normalised SSC-A})$.	105
Figure 4.1 The timeline for sampling, extraction, LC-MS analysis and data analysis for pH shift experiment.	112
<i>Figure 4.2 Plots of representative metabolite concentration at different bioreactor conditions during the pH shift experiment using C. autoethanogenum.</i>	116
Figure 4.3 Heat map showing intracellular metabolic changes for each condition applied to the growth media of C. autoethanogenum. Colour code; Green -Low, Black Medium, Red -High.	121
Figure 4.4 Mapping of intracellular metabolites on the one carbon metabolic pathway during the pH shift experiment for the solventogenic phase using Escher online platform for pathway analysis. Colour code goes from light brown for lower concentration to darker red for higher concentrations. Circles represent metabolites and the concentration of metabolites is indicated by size and colour. Small plain circles are unmapped metabolites	124
Figure 4.5 Mapping of transcriptomics data on the one carbon metabolic pathway during the gas shift experiment for the solventogenic phase using Escher online platform for pathway analysis. Colour code goes from blue for down regulated genes to red for up regulated genes. The arrows represent Gene data and the colour code shows an up-regulation or down-regulation	125
Figure 4.6 Metabolic pathway visualisation of C. autoethanogenum generated using biocyc.org. Each box represents a condition as stated in the boxes above. The colour coding goes from blue to nearly red to indicate low and high levels respectively. A- Pyruvate fermentation to acetate IV and B-Mixed acid fermentation in C. autoethanogenum.	128
Figure 4.7 Optical Density measures at 600 nm (OD ₆₀₀) in correlation with Carbon monoxide (CO) levels. An increase in gas flow from 15 ml/min to 27 ml/min showed a	

proportional increase in biomass. (black triangle-CO flow rate, grey circle – OD measured at 600 nm).	131
Figure 4.8 Ethanol and acetate production. Effect of pH shift on the conversion of acetate to ethanol. Higher concentrations of acetate were observed at pH 5 compared to ethanol. At pH 4.7 concentrations of acetate and ethanol reached equilibrium. The conversion of acetate to ethanol was observed at pH 4.5. (Purple- pH, blue-acetate, grey-ethanol). (pH, n=1)	131
Figure 4.9 The effect of pH changes on extracellular concentrations of 2,3-butanediol and lactate. 23BD levels were at the highest during the solventogenic phase at pH 4.5 when the re-assimilation of acetate and ethanol occur. The levels of lactate seem to increase when the pH was increased back to when the re-assimilation of acetate had ceased. (red-pH, Purple-2,3-butanediol, blue-lactate)	132
Figure 5.1 The timeline for sampling, extraction, LC-MS analysis and data analysis for gas shift experiment.....	137
Figure 5.2 OPLSDA plots of acetate and ethanol phases of the gas shift experiment showing separation and clustering of groups. Red triangles represent acetate phase and green pluses represent ethanol phase. Each replicate is an average of samples collected over 24 days with three samples taken daily.	146
Figure 5.3 Variable Importance in Projection (VIP) scores of 20 metabolites represented in order of significance. On the top right is a mini heat map for all metabolites identified comparing the acetate and ethanol phases. Green colour indicates lower concentrations and high concentrations are coloured red. MetaboAnalyst only reports the first 14 characters of the compound name hence the loss in some of the details.	147
Figure 5.4 Heat map of 25 metabolites showing orderly progression of metabolites during the acidogenic-solventogenic transition in <i>C. autoethanogenum</i> obtained from statistical analysis done using MetaboAnalyst 3.0. Acetate phase is associated with the red class while green is for ethanol phase. Higher concentrations of metabolites are highlighted in red while lower concentrations are coloured in blue.	148
Figure 5.5 Mapping of intracellular metabolites on the one carbon metabolic pathway during the gas shift experiment for the solventogenic phase using Escher online platform for pathway analysis. Colour code goes from light brown for lower concentration to darker red for higher concentrations. Circles represent metabolites and the concentration of metabolites is indicated by size and colour. Small plain circles are unmapped metabolites.	150
Figure 5.6 Mapping of transcriptomics on the one carbon metabolic pathway during the gas shift experiment for the solventogenic phase using Escher online platform for pathway analysis. Colour code goes from blue for down regulated genes to red for up regulated genes. The arrows represent Gene data and the colour code shows an up-regulation or down-regulation.	151
Figure 5.7 Metabolic pathway analysis and visualisation using MetaboAnalyst 3.0. This is the metabolome view showing all metabolic pathways arranged according to the scores from enrichment analysis (y axis) and from topology analysis (x axis).	153
Figure 5.8 Optical density at steady state as a function of carbon monoxide supply rate. A shift from carbon-limited growth to non-carbon-limited growth occurred around a CO inflow of 1.2 mol/(d·L).	155
Figure 5.9 A- shows production of ethanol and acetate ad function of carbon monoxide (CO) uptake. B- shows the production of 2,3-butanediol (2,3BD) and lactate as a	

function of the CO uptake. Whereas acetate and ethanol production is increasing, with acetate as the major product, under carbon-limited conditions, formation of ethanol further rises as that of acetate drops during non-carbon limited conditions. Furthermore, the concentration of the minor products 2,3-butanediol and lactate increase with increasing carbon monoxide supply, however lactate could be detected under non-carbon-limited conditions only.	156
Figure 5.10 Intracellular concentration of CoA normalised to OD during the gas shift experiment using <i>C. autoethanogenum</i> . There is a decrease in CoA concentration during the non-carbon limiting phase.	157
Figure 5.11 The central metabolic network of gas fermentation in <i>C. autoethanogenum</i> . . The bar graphs represent the metabolite concentrations at 15, 34 and 44 (carbon limitation, transition, non-carbon limitation). (Designed by Dr Thomas Millat and Rupert Norman).....	158
Figure 5.12 Plots of intracellular concentrations of folates during the three steady states of the gas shift experiment. Intracellular concentration of monoglutamated folates and total folates were evaluated for the acetate, ethanol and 2,3-butanediol phases. (AP_mono- Acetate phase monoglutamated folate, AP_total - Acetate phase total folates, EP_mono- Ethanol phase monoglutamated, EP_total- Ethanol phase total folates, BD_mono- 2,3-butanediol phase monoglutamated, BD_total- 2,3-butanediol phase total folates).....	160
Figure 5.13 List of metabolites that contributed significantly to the changes observed in the folate analysis during the gas shift experiment. Metabolites with VIP scores>1 were chosen to determine metabolites of significant contribution. The metabolites are in descending order of importance. The coloured boxes on the right indicate absolute concentrations of the corresponding metabolites for each group, red means high and green means low. (AP_mono-glu- Acetate phase monoglutamated folate, EP_mono-glu - Ethanol phase monoglutamated,)	161
Figure 5.14 Glyoxylate and dicarboxylate metabolism pathway of <i>C. autoethanogenum</i> obtained from KEGG pathways showing direct links to important pathways such as TCA cycle.	163
Figure 6.1 A schematic of the biosynthesis of pantothenate and coenzyme A pathway for <i>C. autoethanogenum</i> and its associated pathways showing aspartic acid as precursor for synthesis of pantothenate. Image obtained from KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway databases.	167
Figure 6.2 A typical example of bioreactor setup. Orange indicates anaerobic media, pink indicated spent media. Blue indicated gases in, neon green represents waste gases. Black indicates agitator. Light green indicates fermentation vessel with water jacket. Red tubing indicates NaOH, yellow indicates Na ₂ S and purple indicates antifoam. Hot pink indicates media addition and the blue green indicates waste media tubing. Purple and green lines indicate the pH and redox probes. Light blue is the sampling port for the culture. The system is controlled by a BioFlo computer and data recorded on a computer. (Modified from diagram drawn by Florence Annan).	170
Figure 6.3 Representation of data before and after normalisation in Metaboanalyst 3.0. Data was normalised by the median of the data.	174
Figure 6.4 PCA (A) and PLS-DA (B) score plots of high pantothenate (HP) and low pantothenate (LP) showing separation and clustering of the two pantothenate conditions. Pink-HP and green-LP.	182

Figure 6.5 List of 40 metabolites that contributed significantly to the changes observed when pantothenate concentration was decreased. Metabolites with VIP scores >1 were chosen to determine metabolites of significant contribution. The metabolites are in descending order of importance. The coloured boxes on the right indicate absolute concentrations of the corresponding metabolites for each group.....	182
Figure 6.6 Pyruvate metabolism for <i>C. autoethanogenum</i> obtained from KEGG pathway database. This pathway produces oxaloacetate, which feeds directly into the TCA cycle. The production of acetate is also directly influenced through this pathway.....	186
Figure 6.7 Citric acid (TCA) cycle of <i>C. autoethanogenum</i> obtained from KEGG pathway database. Alanine, aspartate and glutamate metabolism (AAG) metabolism pathway is directly linked to the TCA cycle. AAG pathway produces L-aspartate that acts as precursor for pantothenate synthesis.....	187
Figure 6.8 Alanine, aspartate and glutamate metabolism of <i>C. autoethanogenum</i> obtained from KEGG pathway database. This pathway produces the L-aspartate for pantothenate synthesis for CoA production.....	188
Figure 6.9 PCA score plots for all samples obtained from the pantothenate limiting experiment. Pink represents low pantothenate concentration (LP), orange for high pantothenate concentration (HP) and brown represents pooled QC.....	190
Figure 6.10 OPLS-DA score plots for all samples obtained from the pantothenate limiting experiment. Pink represents low pantothenate concentration (LP) and orange for high pantothenate concentration (HP).....	190
Figure 6.11 Validation of OPLS-DA models of higher pantothenate and lower pantothenate conditions using permutation test (n=38). R ² is the model fit and Q ² is the predictive ability of the model. The Q ² intercept values of the regression line of randomly permuted Q ² value was -0.378, indicating that the model is statically sound and the high prediction ability of the model is not due to over fitting.....	191
Figure 6.12 A representative compound identification output obtained from compound discoverer 2.1 showing compound structure, name, retention time, chemical formula and molecular weight.....	192
Figure 6.13 Aminoacyl-tRNA biosynthesis pathway of <i>C. autoethanogenum</i> (KEGG Pathway database).....	197
Figure 6.14 Comparison of steady states for solvent production during pantothenate limitation. In steady state 3(SS), optimal concentrations of pantothenate was present in the media and in SS4 the pantothenate concentration was reduced to 10 % of the standard concentration. There is an increase in ethanol levels during this reduction in pantothenate concentration. (diagram from Florence Annan).	199
Figure 6.15 Overview of pathway analysis using MetaboAnalyst 3.0	207
Figure 6.16 Taurine and hypotaurine metabolism from KEGG pathways for <i>C. autoethanogenum</i>	209
Figure 6.17 PCA score plots for all samples obtained from the phosphate limiting experiment. Pink represents phosphate limiting phase (P_lim), orange for high phosphate concentration (HP) and blue represents pooled QC.(R ² X= 0.591 Q ² = 0.31)	210
Figure 6.18 OPLS-DA score plots for all samples obtained from the phosphate limiting experiment. . Pink represents phosphate limiting phase (P_lim), orange for high phosphate concentration (HP). (R ² X 0.326, R ² Y =0.981 Q ² = 0.837).....	211

Figure 6.19 Validation of OPLS-DA models of higher phosphate and lower phosphate concentration using permutation test (n=54). R2 is the model fit and Q2 is the predictive ability of the model. The Q2 intercept values of the regression line of randomly permuted Q2 value was -0.303, indicating that the model is statically sound and the high prediction ability of the model is not due to over fitting. 211

List of abbreviations

μM	micro-moles per litre
$^1\text{H-NMR}$	proton nuclear magnetic resonance
2,3BD	2,3-butanediol
AA	Ascorbic acid
AAG	
metabolism	Alanine, aspartate and glutamine metabolism
ABE	Acetone, butanol and ethanol
Acetyl CoA	Acetyl Coenzyme A
ACN	Acetonitrile
ADH	Alcohol dehydrogenase
AMC	Activated methyl cycle
AMW	Acetonitrile-methanol-water
AOR	Aldehyde oxidoreductase
APCI	Atmospheric pressure chemical ionisation
API	Atmospheric pressure ionisation
ATP	Adenosine triphosphate
C	Carbon
C1	One carbon
CD	Compound Discoverer
CE	Capillary electrophoresis or collision energy
CO	Carbon monoxide
CO ₂	Carbon dioxide
CoA	Coenzyme A
CXP	Collision exit potential
DHF	Dihydrofolate
DMHA	N,N-dimethylhexylamine
DP	Declustering potential
ESI	Electrospray Ionisation
EU	European Union
FA	Formic acid
FCM	Flow cytometry
FDA	Food and Drug Administration
FDR	False discovery rate
FT	Freeze thaw
FT-ICR-MS	Fourier transform ion cyclotron resonance mass spectrometer
GC-MS	Gas chromatography- mass spectrometry
H	Hydrogen
H ₂	Hydrogen gas
H ₂ O	Water
HILIC	Hydrophilic interaction liquid chromatography
HILIC-	Hydrophilic interaction liquid chromatography-isotope
IDMS	dilution mass spectrometry

HMDB	Human metabolome database
HMG	Homogemiser
HP	High Phosphate /Pantothenate
HPLC	High performance liquid chromatography
HPLC-MS	High performance liquid chromatography-mass spectrometry
HRMS	High resolution mass spectrometry
IDMS	Isotope dilution mass spectrometry
IPR	Ion pairing reagents
IPRP	ion pair reversed phase
kDa	kilo Dalton
LC-IDMS	Liquid chrmatography- isotope dilution mass spectrometry
LC-MS	Liquid chromatography- mass spectrometry
LC-MS/MS	Liquid chromatography- tandem mass spectrometry
LOD	Limit of detection
LOQ	Limit of quantification
LP	Low pantothenate
LTQ	Linear trap quadrupole
m/z	Mass to charge ratio
MeOH	Methanol
M-EtOH	2-mercaptoethanol
MRM	Multiple reaction monitoring
MVA	Multivariate analysis
Mw	Molecular weight
MW	Methanol-water
N	Nitrogen
NAD ⁺	Nicotinamide adenine dinucleotide (oxidising agent)
NADH	Nicotinamide adenine dinucleotide (reducing agent)
NH ₄ HCO ₃	Ammonium bicarbonate
nM	nano-moles per litre
nM	milli-moles per litre
O	Oxygen
OD ₆₀₀	Optical density at 600 nm
OPLS_DA	Orthogonal least square discrimination analysis
OSC	Orthoggonal signal correction
P_lim	Phosphate limitation
PC	Principal componenet
PCA	Principal componenet analysis
PLS-DA	Partial least square- discriminant analysis
ppm	Parts per million
Q ₂	Predictive ability
QC	Quality control
QqQ	Triple quadrupole mass spectrometry
Q-ToF-MS	Quadrupole time of flight mass spectrometer
Q-TRAP	Quadrupole- ion trap mass spectrometry
R ² Y	fitness of model
RP	Reversed phase

RSD	Relative standard deviation
RT	Retention time
SBRC	Synthetic Biology Research Centre
SD	standard deviation
SILIS	Stable isotope labelled internal standard
SIM	Selected ion monitoring
SRM	Selected reaction monitoring
Syngas	Synthesis gas
TBA	Tributylamine
TCA	Tricarboxylic acid cycle
THF	Tetrahydrofolate
TIC	Total ion chromatography
TLC	Thin layer chromatography
	Ultra-high-performance liquid chromatography-mass
UHPLC-MS	spectrometry
UHV	ultra-high vacuum
VIP	Variable importance for the projection
WLP	Wood-Ljungdahl Pathway

CHAPTER ONE

General Introduction

1. General introduction

1.1 Introduction

1.1.1 The Global energy crisis

Depletion of global oil reserves has increased interest in the use of biofuels as sustainable energy sources in recent years, alongside concerns on climate change. In order to promote the production of sustainable energy, it will be required of all EU member states to reach their assigned energy targets numbers and also a 20 % share of renewable energy by 2020 (van Groenestijn et al., 2013). By the same time, it will also be mandatory that 10 % of the renewable energy is used in transportation by member states (Latif et al., 2014). In 2012, bioethanol accounted for 28 % of total biofuels used in road transport in the EU while, on a volume basis, biodiesel accounted for about 70% of the biofuels market. Bioethanol production in EU was predicted to reach 5.38 billion litres in 2014 cf. to 4.6 billion litres estimated in in 2012 (Flach et al., 2013).

Currently corn, wheat, rye and barley are the main feed stocks for production of bioethanol. This is undesirable because these feedstocks compete with availability of food. An alternative will be to use waste or readily available lignocellulosic biomass for production of bioethanol. However, the traditional approach of converting lignocellulosic biomass to bioethanol uses large amounts of energy (Balat and Balat, 2009) and hence is not efficient. One promising new approach to the sustainable generation of bioethanol is by converting biomass into gas to produce syngas which is utilised by anaerobic bacteria mainly *Clostridium sp* in a fermentor (Bengelsdorf et al., 2013, Mohammadi et al., 2011, Nalakath Abubackar et al., 2011).

1.1.2 The use of acetogens to produce fuels and chemicals

Fermentation of waste gases from cheap one carbon (C1) substrates such as carbon monoxide by acetogenic bacteria is a well-known approach for production of chemicals and low carbon fuels (Marcellin et al., 2016b). Acetogens can utilise carbon dioxide/hydrogen gas/ carbon monoxide (CO₂/H₂/CO) for the production of ethanol, lactate, acetate and 2,3-butanediol

(2,3BD) (Kopke et al., 2011b, Abrini et al., 1994, Koepke et al., 2014). Ethanol is a useful chemical in the biofuel industry, while 2,3BD is a precursor for the production of gamma-butyrolactone, methyl ethyl ketone and 1,3-butadiene. Nylon and rubber are derived from 1,3 butadiene making 2,3BD an industrially important compound (Marcellin et al., 2016a).

Clostridium autoethanogenum is a key organism for ethanol, acetate and 2,3-butanediol production (Abrini et al., 1994) using synthesis gas (syngas), which predominantly contains CO and H₂ (Abubackar et al., 2015a, Guo et al., 2010). Syngas is readily available and a cheap substrate. It is usually generated from natural gas, through gasification of oil and coal, and from greener sources like biomass, recycling of used plastics or from municipal waste. Syngas is already in use for non-bio synthesis in chemical industries, but the specific reactions demand a ratio of CO: H₂ with a purification step for these gases which has proven to be expensive, as contaminants tend to poison the catalysts (Koepke et al., 2010). This contamination is tolerated more in bacteria that ferment syngas (acetogens) (Ragauskas et al., 2006) and currently these have been used to produce additives of biofuel such as ethanol by INEOS Bio, LanzaTech and Coskata (Koepke et al., 2010).

1.1.3 Fermentation of gas through the one carbon metabolic pathway (Wood-Ljungdahl pathway)

Microorganisms are able to use CO mainly due to the presence of the Wood-Ljungdahl (Rahman et al.) metabolic pathway (Drake and Kusel, 2005, Drake et al., 2006, Wood, 1991). It has been reported that carbonyl is taken up in the carbonyl (Western) branch of the pathway whiles CO and/or carbon dioxide (CO₂) is taken up through the methyl (Eastern) branches of the pathway leading to the production of acetyl-CoA, a precursor for anabolic and catabolic processes in the cell (Kopke et al., 2011b). The methyl branch as the name suggests is responsible for the production of the methyl (CH₃) part of the acetyl-CoA through a series of reduction reactions using tetrahydrofolate as cofactors. The carbonyl group of the acetyl-CoA is generated from the carbonyl branch where CO₂ is reduced to CO by carbon monoxide dehydrogenase (CODH). Anaerobes can make use of the acetyl-CoA pathway to reduce CO₂ to form the acetyl portion of acetyl-CoA (Figure 1.1), to conserve energy, and to help assimilate

CO₂ into the cell as carbon (Drake and Kusel, 2005). According to Wood, ‘the acetyl-CoA Pathway plays a significant role in the carbon cycle. The direct combination of two CO₂ to form acetate may have been used by the earliest life forms rather than the more complicated cyclic mechanisms of autotrophism’ (Wood, 1991).

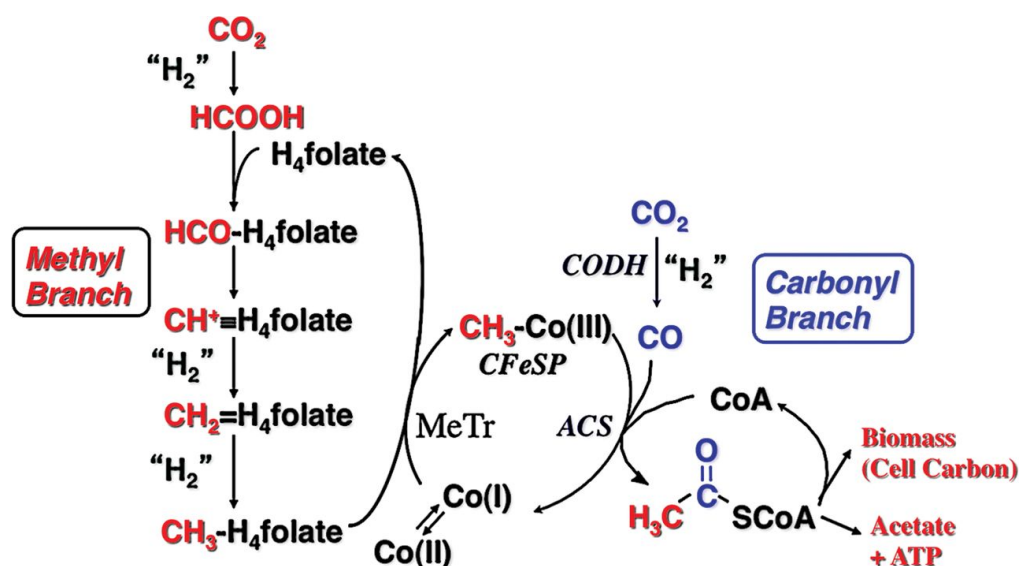


Figure 1.1 The Wood-Ljungdahl pathway showing the Eastern and Western branches as depicted in acetogens. Carbon labelled red forming the methyl group of acetyl-CoA was generated from the CO₂ from the eastern branch and the carbonyl group in acetyl-CoA from the CO₂ from the western branch (Ragsdale and Pierce, 2008). H₄folate (tetrahydrofolate), HCO- H₄folate (formyl tetrahydrofolate), CH₂= H₄folate (methylene-tetrahydrofolate), CH⁺= H₄folate (Methenyl-tetrahydrofolate) and CH₃- H₄folate (Methyl-tetrahydrofolate). H₄folate has been proposed to be involved in CO₂ fixation pathway (Ljungdahl et al., 1966).

1.1.4 Metabolic pathway engineering of *C. autoethanogenum*

Acetate, ethanol and 2,3-butanediol are the three main products produced naturally by *C. autoethanogenum*. Isobutanol and succinate pathways are non-natural and require introduction to the organism using bioengineering/synthetic biology approaches. Ethanol and acetate are derived from the acetyl CoA while

pyruvate is the precursor for the production of succinate, 2,3-butanediol and isobutanol (Figure 1.2). The ability of *C. autoethanogenum* and other acetogens in using one-carbon compounds (CO/CO₂) to make up to four carbon compounds (2,3-butanediol) is a useful approach in synthetic biology (Kopke et al., 2011b). The use of *C. autoethanogenum* on an industrial biotechnology application can be problematic if product yields are not efficiently managed. The accumulation of acetate in the media is detrimental to the organism and also expensive to dispose especially on a large scale. The amounts and ratios of metabolites produced depend on conditions applied to the bacteria medium. This could mean manipulating the flux of carbon through these pathways to minimise acetate production and maximise ethanol production. It would be desirable to produce more valuable products either by maximising the native product 2,3-butanediol or by introducing heterologous pathways leading the production of isobutanol and succinate.

To fully understand the effects of manipulating the carbon flux, the metabolic pathway intermediates can be monitored through metabolomics approaches. Due to the diverse nature of these intermediates, finding a suitable analytical method to monitor them can be challenging.

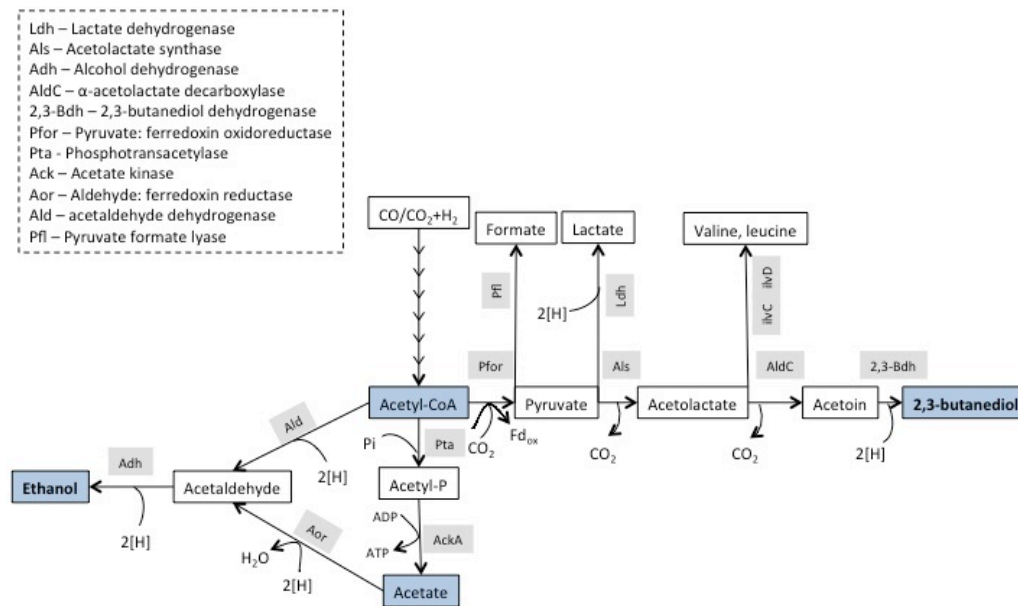


Figure 1.2 One carbon metabolic pathway of *C. autoethanogenum* starting from Wood-Ljungdahl pathway to the production of acetate ethanol and 2,3-butanediol (23BD). The production of ethanol, acetate and 23BD are natural to the organism while production of isobutanol and succinate are non-natural pathways that will be introduced by systems biology. Acetyl CoA is the precursor for ethanol and acetate production while pyruvate is the precursor for the production of 23BD.

1.1.5 Metabolomics analysis of diverse polar compounds

Choosing an analytical method for analysis of diverse ranges of metabolites can be challenging especially for highly charged and polar compounds. When analysis of a diverse range of metabolites is involved, usually two different methods are combined to improve the coverage. The two most common methods used are reversed-phase chromatography (RPC) and hydrophilic interaction liquid chromatography (HILIC) (Cui et al., 2013, Lu et al., 2008). RPC is usually not suitable for retention and separation of polar metabolites hence addition of ion pair reagent (Luo et al., 2007) to the mobile phase is a common approach to improve retention of polar compounds. Ion pair reagents are less favourable when mass spectrometry detection is used due to ion suppression/enhancement and problems with monitoring ions in the positive ionisation mode. Polar compounds are better retained on HILIC and require no ion pair reagent to be added to the mobile phases. Combinations of both methods have usually been used to monitor diverse classes of compounds. With the advancement in the field of metabolomics it is clear that a combination of methods rather than a single analytical method is needed to achieve a full coverage of metabolites as highlighted in this review (Haggarty and Burgess, 2017).

The main steps in metabolomics analysis are the analysis and detection of metabolites followed by comprehensive data analysis and interpretation. Despite the recent advances in metabolomics technologies, no single instrumental platform is known to perform a complete analysis of all metabolites present in an organism

1.1.6 Metabolomics analysis and its application in bacteria

Metabolomics covers qualitative and quantitative investigation of small molecules also referred to as metabolites within biological systems. Metabolites are organic molecules involved in metabolic reactions needed for maintenance of growth, homeostasis and normal cell function with molecular mass less than 3 kDa (Lokhov and Archakov, 2009).

Metabolomics analysis of bacteria has been well reported for a variety of different species of bacteria each method focusing on different aspects of metabolomics (Link et al., 2012). Metabolomics analysis in bacteria is typically

used to measure changes observed in metabolic pathways (intracellular or extracellular) when, for example, conditions in the growth conditions are changed (Paczia et al., 2015, Zhao et al., 2016). Metabolomics analysis in bacteria is analytically challenging, due to the low concentration levels of intracellular metabolites making detection of multiple metabolites difficult in most cases. Where convenient, more cells can be collected to boost the detection levels and the volume of extracting solvent can be reduced to concentrate the metabolites.

The metabolome constitutes a large set of compounds from diverse compound classes such as amino acids, nucleotides, lipids and organic acids. The difficulty in analysis is not only dependent on the differences in physical and chemical properties but equally the differing concentration levels in the organism. This poses a major challenge in choosing a suitable method for a specific experiment.

Two approaches are usually used in metabolomics studies, targeted and global (untargeted) metabolomics.

1.1.6.1 Targeted metabolomics

Metabolomics analysis where the focus is on detecting specific groups of metabolites is called Targeted metabolomics and is performed on anything from a few metabolites to several hundred. Targeted metabolomics tend to focus on (1) a single biological pathway (e.g TCA cycle) where the whole pathway or most of the metabolites in the pathway are determined, (2) a specific group or groups of related metabolites (e.g. lipids, amino acids) which may be found in different metabolic pathways, (3) a wide collection of known compounds covering a range of metabolite classes to reflect the changes in multiple metabolic pathways also referred to as widely targeted metabolomics (Sawada et al., 2009). Targeted metabolomics methods usually provide qualitative or semi quantitative fold change on the metabolites determined, and in some cases highly desirable absolute quantitative analysis. Hence targeted metabolomics methods are regarded as a comprehensive approach and have the advantage that changes observed are associated with known and well characterised metabolites (Begou et al., 2017).

In a targeted quantitative metabolomics method, it is important that method

validation is used. Method validation is a set of processes used to ensure that an analytical method is suitable for its intended purpose. The results obtained from this validation can be used to evaluate the reliability, reproducibility and quality of the analytical results. In a recent study where, amino acid profiles in the plasma of kidney transplant patients were compared using both targeted and untargeted LC-MS/MS assays showed that, though there were similarities between the results that were generated from both methods, there were equally significant differences observed (Klepacki et al., 2015). On the average, targeted methods are advantageous in detection, quantification and validation of predefined metabolites (Griffiths et al., 2010, Lu et al., 2008) and better controlled compared to untargeted methods that may be difficult to validate or control. Targeted metabolomics methods promise to provide a thorough investigative quantitative tool for the studies of microbial metabolism, where determination of changes in intracellular metabolite concentrations provide direct study of enzyme kinetics of metabolic pathways involved (Buchholz et al., 2002).

1.1.6.2 Untargeted metabolomics

In untargeted metabolomics (metabolic profiling, global metabolomics or metabolic fingerprinting) studies there is no prior knowledge of metabolites being studied and they provide a neutral outlook of the metabolic activities that occur in the entire organism or cell whereby every single metabolite present in the samples are detected. Analysis of samples in untargeted metabolomics involves a comparison between usually treated and control sets of samples with multivariate analysis (MVA) or univariate data analysis being used to compare the two conditions. Tentative identification of compounds of interest can usually be obtained but the formal identification can be achieved by a targeted metabolomics approach for metabolite confirmation (Dettmer et al., 2007).

There are very many applications of untargeted metabolomics in diverse studies (Bertini et al., 2014, Creek et al., 2011, Toyota et al., 2016) and all these different approaches require the use of appropriate metabolites databases to putatively identify metabolites. Depending on the type of samples investigated, databases such as Human Metabolome Database (HMDB) (Wishart et al., 2013) can be used and other data processing software that interact with databases such as

Compound Discoverer (Thermofisher) and Lipidsearch (Thermofisher) can be used.

Data analysis of untargeted metabolomics is usually shown as score plots through multivariate analysis to depict clustering and separation between groups.

In Figure 1.3 the schematic shows the differences between targeted and untargeted metabolomics approaches. As stated before, in targeted metabolomics the metabolites of interest are known prior to analysis while the opposite is the case for untargeted metabolomics approach. A combination of both methods on the same sample will give us an indication of which approach is better at giving detailed information on metabolic pathways specific for changes in environmental stimuli.

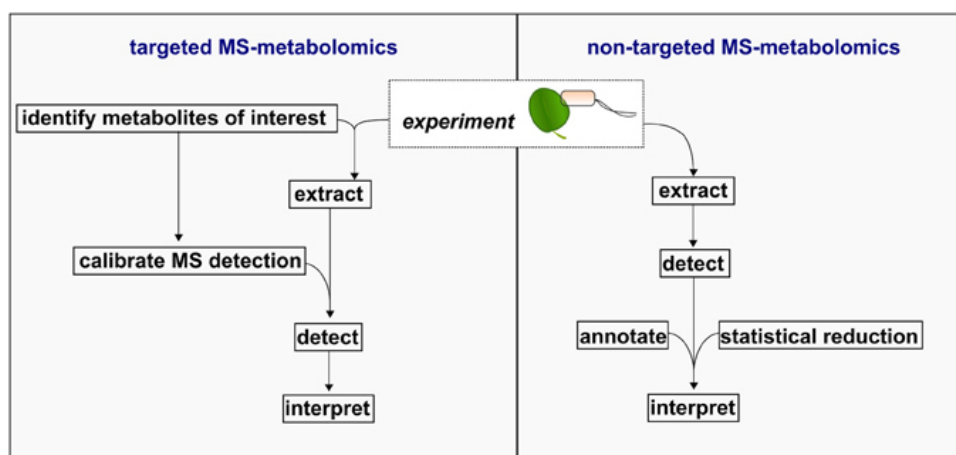


Figure 1.3 A representation of the different approaches used in untargeted and targeted metabolomics analysis work flow (Heuberger et al., 2014).

1.1.7 Different analytical techniques used in metabolomics applications

There are many separation techniques that are being used and currently been used for analysis of metabolites. Earlier methods included thin layer chromatography (TLC) (Brzezinka et al., 1999), capillary electrophoresis (CE), nuclear magnetic resonance (NMR), gas chromatography (GC), high performance liquid chromatography (HPLC), liquid chromatography (LC) and mass spectrometry (MS). Due to small sample volumes metabolomics analysis

requires a sensitive and selective method that can provide reliable measurement representative of the metabolome of interest. Therefore, analysis using NMR and MS are well known techniques and usually coupled to separation techniques such as chromatography and CE.

NMR can be advantageous when sample volume is crucial because it is a non-destructive technique making it possible to analyse the same sample using other techniques. However, due to the reduced sensitivity of proton (^1H) NMR in comparison to MS, the application of NMR to only high concentrations metabolites is limited. Also, there is difficulty with identification of large groups of compounds because of overlapping peaks making NMR less favourable compared to MS. The identification problems can be overcome with 2D NMR (Ward et al., 2007). Over the years the use of LC and GC separations coupled to MS (Chen et al., 2014) for metabolomics analysis have become well-established techniques because of their increased sensitivity and selectivity and ability to provide metabolite identification by chromatographic retention and a combination of accurate mass and fragmentation analysis (Figure 1.4).

Metabolomics analysis of non-polar and polar metabolites using LC-MS and GC-MS are reported in clinical studies (Maurer, 2005, Peng et al., 2014, Pfeiffer et al., 2004), microbial metabolic studies (Wu et al., 2005, Zakhartsev et al., 2015) and many more. GC-MS, though useful in identification and/or quantification of diverse range of metabolites, is limited to volatile compounds, with non-volatile compounds requiring a derivatisation step (Halliday et al., 2010). LC-MS has evolved to be the most preferred method for MS-based metabolomics due to increased sensitivity and selectivity of a broader range of metabolites including thermally labile and non-volatile metabolites without the need for derivatisation as required in GC-MS (Buescher et al., 2009).

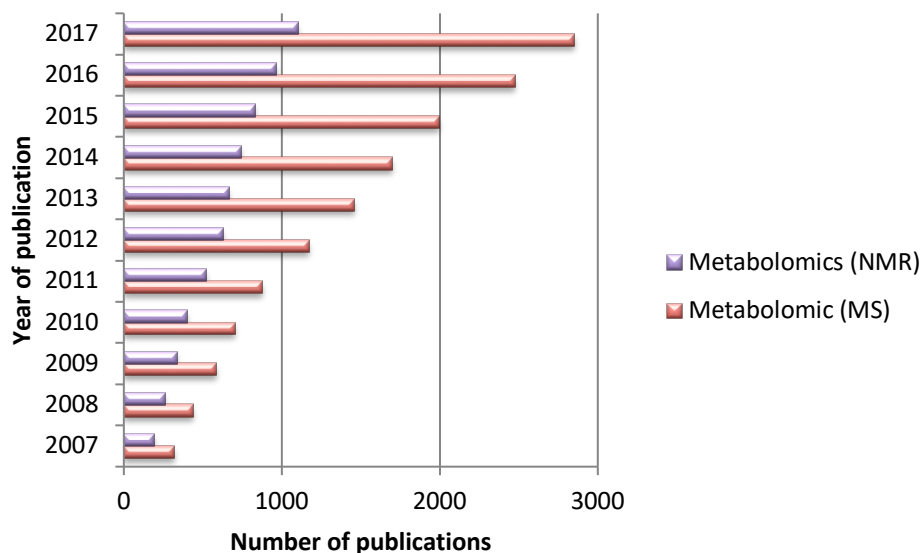


Figure 1.4 Bar graphs showing the comparison between MS based metabolomics and NMR based metabolomics over the past 10 years. These data were generated using Web of science database.

1.1.8 LC-MS based metabolomics

The rapid advancements over the last two decades in the field of MS has led to an increase in application of LC-MS for metabolomics studies. When LC is coupled to MS there are problems such as higher flow rate of liquid waste from the column and how to couple this atmospheric pressure liquid flow to the ultra-high vacuum (UHV) working environment of MS. The introduction of the atmospheric pressure ionisation interface (Tzur et al.) overcame such problems. There are different interfaces or soft ionisation techniques that produces less fragmentations and are routinely used such as atmospheric pressure photoionisation (APPI), (Robb et al., 2000) atmospheric pressure chemical ionisation (APCI) and electrospray ionisation (ESI). The use of ESI for ionisation with either HPLC-MS or ultrahigh performance liquid chromatography mass spectrometry (UHPLC-MS) is the most popular approach used in metabolomics. The use of UHPLC-MS has shown to improve chromatographic resolution, efficiency, and reduced analysis times requiring smaller sample volumes. More metabolites have been detected using UHPLC-MS compared to the conventional HPLC-MS (Han et al., 2009). UHPLC-MS

has many applications in metabolomics with either hydrophilic interaction liquid chromatography (HILIC) columns (Gika et al., 2012) or reversed-phase (RP) (Buescher et al., 2010, Siegel et al., 2014). The biological matrix investigated, and the analytes of interest will determine the type of column required. Polar or semi-polar metabolites are usually analysed using HILIC while non-polar metabolites are analysed using RP. MS coupled to RP and HILIC columns have been applied for the analysis of polar metabolites to improve metabolite coverage (Gika et al., 2008).

For quantitative or targeted metabolomics, tandem electrospray MS (MS/MS) or (MSⁿ) is widely used. Typical instrumentation for targeted metabolomics includes the tandem triple quadrupole MS (Xiao et al., 2012) or hybrid quadrupole-linear ion trap mass spectrometer (Q-Trap), due to accurate quantification of metabolites and improved sensitivity (Sawada et al., 2009). This approach requires the incorporation of an internal or external calibration for precise measurement. There are different ion modes for tandem MS which include multiple reaction monitoring (MRM), selected reaction monitoring (SRM), selected ion monitoring (SIM).

Despite the advantages in targeted analysis, the use of low-resolution mass analysers like quadrupoles and ion traps in untargeted metabolomics has significant disadvantages. Metabolomics applications using high resolution MS such as quadrupole time of flight (Q-ToF) and Orbital trap MS are usually desirable.

1.1.9 Mass spectrometry techniques

1.1.9.1 Electrospray ionisation

The most common ionisation method used in metabolomics studies is Electrospray ionisation (ESI) compared to atmospheric pressure chemical ionisation (APCI). The first application of ESI to produce gas phase ions was reported in the late 1960s by (Dole et al., 1968). The application of ESI in other fields such as large biomolecules was reported (Fenn et al., 1989) and where ESI was coupled to a mass spectrometer (MS) (Yamashita and Fenn, 1984). The continuous success of ESI application for analysis of biomolecules has made this technique very important in analytical chemistry.

In ESI, a high voltage is applied through a capillary to the liquid to produce charged droplets. Nitrogen gas is used as a counter current flow gas which helps evaporate the droplets. The ions in the gas phase are then separated and detected in the MS according to their mass to charge ratios (m/z). Due to the soft ionisation technique in ESI leading to little or no fragmentation makes it suitable for the analysis of fragile or unstable molecules. This can be disadvantageous for instances when fragmentation is important for detection or identification but can be rectified by using a tandem mass spectrometric technique such as MS/MS or MSⁿ. Detection of higher molecular mass compounds such as proteins is feasible when using ESI due to multiple charges that can be generated.

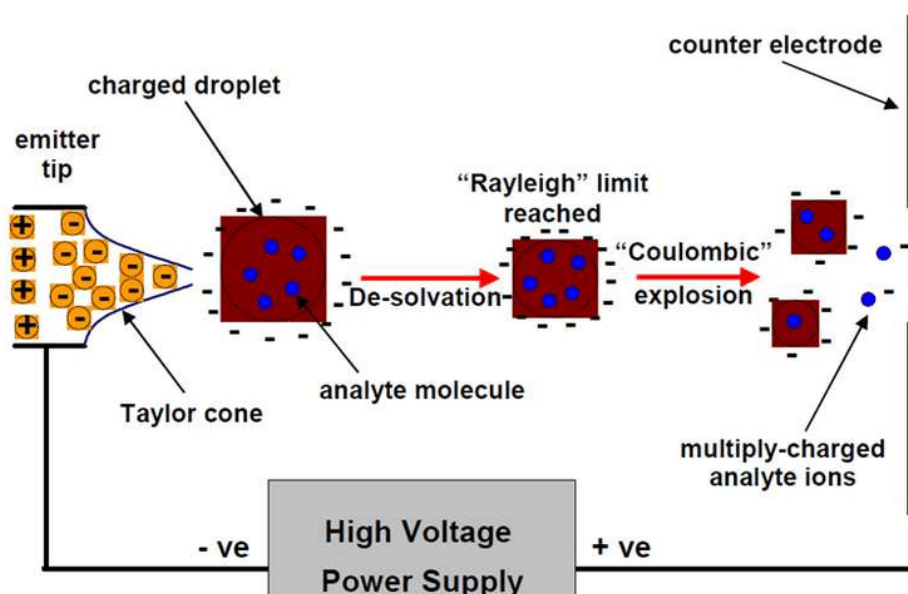


Figure 1.5A schematic showing ion formation in electrospray ionisation technique. (Challamalla et al., 2012).

1.1.9.2 Tandem mass spectrometry (MS/MS)

In tandem mass spectrometry (MS/MS), a quadrupole mass analyser is used to obtain quantitative data of analytes. In this mode Q1, the first mass analyser selects the ions of interest. The ions are fragmented in Q2, the collision cell and these fragments are analysed in a third quadrupole Q3 (Figure 1.6). Q2 has no mass filtering properties; it is a radiofrequency (RF) only quadrupole which uses

gases like helium, argon and nitrogen for collision induced dissociation of detected precursor ions (Q1 masses). Q1 and Q3 have mass filtering properties, which makes this fragmentation process crucial in structural identification (Pisitkun et al., 2007). Triple quadrupole mass spectrometry has been used in the quantification of small molecules and metabolites due to its sensitivity, broad dynamic range and reproducibility. In metabolomics analysis, the MRM is used to give the best signal for each analyte of interest by pre-optimisation of the collision energy (CE) and mass to charge ratio of the product ion (Lu et al., 2008).

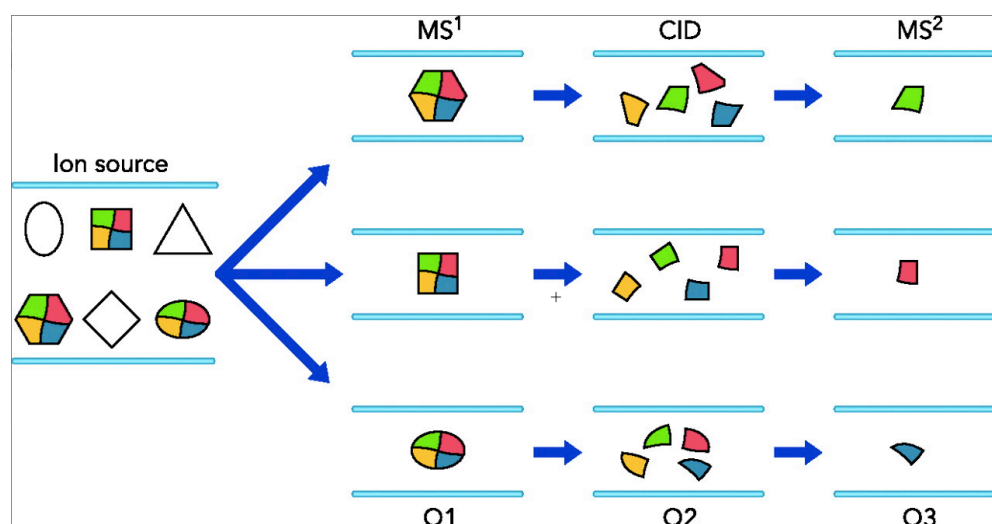


Figure 1.6 Multiple reaction monitoring (MRM) by triple quadrupole (QqQ) mass spectrometer. Predetermined parent (precursor ions selected for MRM by QqQ mass analysis are shown as colored symbols in ion source. First (quadrupole, Q1 (MS¹ level) selectively allows only the targeted parent ions to pass into the second quadrupole (Q2) for collision-induced dissociation (CID) fragmentation. Third quadrupole, (Q3) (MS² level) selectively allows only specific pre-identified fragment ions to pass through to the detector. (Pisitkun et al., 2007).

1.1.9.3 High-resolution accurate (Orbital trap) mass spectrometry (HRAMS)

The initial step in the series of inventions that brought about the orbital trap was its application of trapping ions, which in itself can be used for mass analysis. The orbital trap was invented by Makarov (Makarov, 2000) as illustrated in Figure 1.7. The Orbital trap can be seen as a modified version of Kingdon trap (Kingdon, 1923) with inner and outer electrodes specially shaped. The application of constant potential between the inner and outer electrodes produced a logarithmic radial potential. Pure electrostatic field traps ions with the absence of magnetic and dynamic (RF) fields in the Kingdon trap. The shape of the outer electrode was changed by Knight in order to improve bound ions on the axis of the trap to achieve an axial quadrupole (Knight, 1981). Configurations of the Kingdon and Knight's mass spectrometer model were not reported to be present in other mass spectrometer. In the late 1990s, Makarov invented the new orbital trap mass analyser (Makarov, 2000). The advancement in the principle associated with Makarov's model was incorporated into the hybrid LTQ Orbital trap instrument by Thermo Fisher Scientific 2005 (Makarov et al., 2006). The orbital trap, even though a completely new type of mass analyser, can be considered as an updated version Knight's model and Kingdon trap consisting of a coaxial spindle and barrel like inner and outer electrodes respectively trapping ions around the spindle in an orbital motion. The electrostatic fields used by Orbital trap and quadrupole ion trap analysers *i.e* static electrostatic fields and dynamic electric field respectively are different, with the Orbital trap considered as an altered form of the quadrupole ion trap. An image current produced by trapped ions using Fourier transform of the frequency signal is used to generate the mass spectrum (Zubarev and Makarov, 2013).

In an Orbital trap, centrifugal forces generated by the voltage-applied between both central and outer electrodes used in trapping ions control the electrostatic attraction of ions to the central electrode. Hence, ions move in an elliptical manner around the central electrode. Also, the ions produce a spiral type movement while they move front and back along the axis. The axial movement is dependent on the mass to charge (m/z) but not dependent on circular movement around the central electrode.

To start with, the field between electrodes is reduced to allow ions generated from an independent ion source and once after injecting ion packets, the voltage on the central electrode is adjusted to increase the electric field. The ions tend to migrate towards the central electrode and accumulate until they reach the right orbital route inside the trap. When the ramping has ceased, the field generated inside becomes static and detection begins. Ions varying in velocities that are present in each packet take up the required volume and move with uniform axial frequency but varying circular frequencies. Therefore, ions with specific m/z spread into rings that move back and forth along the central spindle. In commercial Orbital trap mass spectrometers, a curved linear trap (C-trap) is used for injection and also where the trap electrode returns ions. When trapping radio frequency voltages are decreased and direct current gradients are applied, ions can be generated into short packets in the C-trap. The C-trap is closely linked with the analyser, differential pumping injection optics (Zubarev and Makarov, 2013).

The RF waveforms enforced to the outer electrode for the excitation of coherent axial oscillations of ion rings is responsible for ion extraction. Packets introduced in to the Orbital trap at a specific offset position from its equator automatically begin axial oscillations without applying any extra excitation. When all ions of interest have finally entered the Orbital trap and moved away from the outer electrodes, the voltage is stabilised on the central electrode to initiate the detection process.

An image current is generated on the outer electrode from different coherent axial oscillations of different ion rings are divided into two equal sensors connected to an amplifier. All the ions trapped into the Orbital trap are detected simultaneous in a specified time and increasing the field strength and resolution can be improved depending on the time of detection. Comparatively, the Orbital trap produces comparatively less downfall of resolution with high m/z than Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) as the former has no magnetic field.

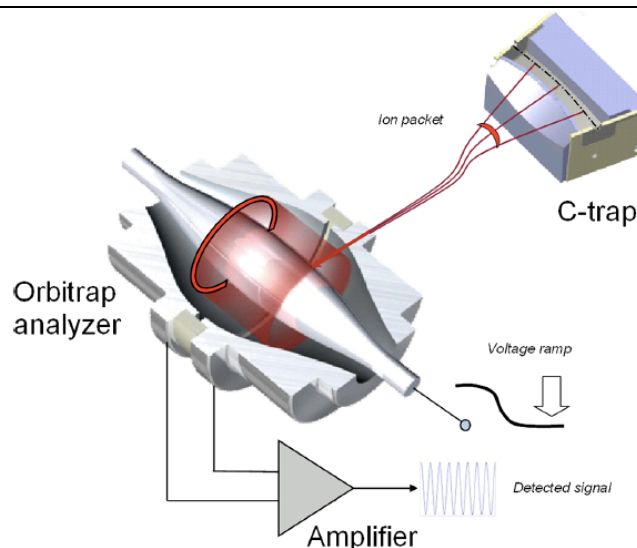


Figure 1.7 This is a cross section of the C-trap and the Orbital trap analyser. The rf-only bent quadrupole of the C-trap stores ions; the rf is ramped down followed application of a high voltage pulse across the trap, with each m/z being emitted in a short packet. During voltage ramp the packets from the C-trap enter the analyser and spread into the oscillating rings inducing a current detected by the differential amplifier (Zubarev and Makarov, 2013)

1.1.10 Data processing in LC-MS based metabolomics

In targeted metabolomics data analysis, there is specific software provided with the instrument used, which is acceptable for processing raw data of the limited number of metabolites of interest. Samples analysed on a Sciex triple quadrupole MS are usually analysed using Analyst software, which supports quantitative data analysis. In tandem MS based quantitative metabolomics, identification of metabolites is confirmed by their fragmentation pattern, which gives a higher level of confidence. Vendors such as Thermo-Fisher scientific have software capable of quantitative data analysis using retention time (RT) and accurate mass to identify metabolites. These include Xcalibur that is available on the Orbital trap (Exactive) and Tracefinder, which is new software that can process data from Orbital traps and hybrid quadrupole Orbital trap (Q-Exactive).

There are open source platforms that can be used for quantitative metabolic

pathway analysis such MetaboAnalyst (Xia et al., 2015) and Biocyc.org metabolic pathways. They can either be used for visualisation of metabolic data or detailed metabolic pathway analysis.

In MS based untargeted metabolomics, data analysis usually entails the comparison of control and treated set of samples as already mentioned. It involves generating large and complex datasets with analytical measurements. There are challenges with extracting useful analytical information from these types of datasets and poor quality of data extraction which could be due to:

1. Poor separation of isomers
2. The impurities and high matrix MS background noise and associated with the sample
3. The range of compounds in the sample
4. The different features associated with one analyte such as protonation/deprotonation adduct formation

The above points can hinder the development of reliable data processing approaches for high throughput metabolomics studies. Several software packages are available for untargeted MS-based metabolomics that use algorithms that correct some of the concerns listed above. Progenesis QI is commercial software used for peak picking and alignment to generate ion intensities for identification using a database of choice. Raw or normalised peak intensities from Progenesis QI can be imported into SIMCA-P for multivariate data analysis. Data can be shown as clustered and separation observed between groups. MetaboAnalyst can also be used for untargeted MS or NMR based metabolomics. Thermofisher has recently upgraded Compound Discoverer (Bauman and Ding, 2015) to incorporate multivariate analysis and direct pathway analysis using the data generated. Accurate data acquisition and data processing in high-throughput metabolomics using MS approaches can still be challenging although many data processing packages are available (Han et al., 2009).

1.1.11 Multivariate analysis of MS-based metabolomics

Multivariate analysis (MVA) can be used to develop predictive models for grouping of unknown samples. Since the numbers of samples is usually less than the number of variables, there is the likelihood of over fitting the data that can lead to inaccurate classification of study groups (Eriksson et al., 2006a). A validation process is required and recommended using independent datasets for all metabolomics analyses.

Pre-processing of the data before MVA has a large impact on the subsequent models generated and can significantly affect the metabolites extracted and identified.

In metabolomics studies, there are two main ways of grouping data into different classes or conditions. There are supervised, and unsupervised methods used for untargeted metabolomics data analysis. The most popular approach in unsupervised metabolomics data analysis is principal component analysis (PCA). PCA uses an algorithm that explains variations in the dataset as much as possible by reducing high dimensional set of data to a minimal number of dimensions. The original variables are combined in a linear fashion for each principal component (PC). The PCs are uncorrelated (orthogonal) with the first few contributing the most variation with subsequent PCs contributing less (Nicholson et al., 1999). When PCs have been identified, PCA maps showing plots of each sample can be plotted to give a visual representation of the results. Because most of the variability is encompassed in the PCs, clustering of data points in the PCA map is typically observed, with each cluster representing a particular metabolic fingerprint. A loadings plot helps interpret PCA maps better by determining which metabolites differed the most between classes. The loadings are plotted against the original variables making it unique for each PC. The loadings plots give a graphical interpretation to what metabolites contributed most significantly to each PC, giving an indication of the metabolites that differed the most between classes (Coen and Kuchel, 2004). Even though PCA detects clusters and outliers very well, the results obtained are only used in directing future analysis since the accuracy can be improved by switching to

supervised methods. Although the largest contribution of the variation is from the PCs, this variation may not correspond closely to the separation between classes (Bylesjo et al., 2006). Hence PCA is usually used as a first step to help in development of the model using supervised methods for better classification. Supervised methods can be used to improve separation between classes and identification of metabolites that differ significantly between classes (Mao et al., 2009). Supervised methods are better when it comes to development of classifiers and ability to predict where samples fall with respect to the classes already classified because they use information on class membership (Nicholson et al., 2002). The commonly used supervised methods are partial least squares with discriminate analysis (PLS-DA) (Mao et al., 2009), orthogonal partial least squares (Cloarec et al., 2005), and orthogonal PLS-DA (Maher et al., 2007). PLS-DA plots uses both the class (descriptor matrix) and the spectra to define a surface in a set dimensional space to separate data into classes (Nicholson and Lindon, 2008). PLS-DA models can be less accurate and difficult to interpret, in the presence of noise. Since noise is a common interference in biological data sets, there is the need to filter out noisy and unrelated variation that has no correlation with class identification (Gavaghan et al., 2002). The commonly used approach to improve the integrity of data sets is orthogonal signal correction, OSC (Ferreira et al.) and has been shown to increase model accuracy in many applications (Brindle et al., 2002, Mao et al., 2009). Filtering out the variation in the descriptor matrix X using OSC reduces noise that is orthogonal to the variation in the results matrix Y (Bylesjo et al., 2006). Also, OPLS-DA is a continuation of PLS-DA that comes with combined OSC filter and models the differences common to X and Y separately from the uncorrelated variation. The advantage of this approach is not only the fact that models are easier to interpret due to the noise been modelled separately, but equally provides the freedom to investigate the results from the noise (Cloarec et al., 2005).

1.1.12 General considerations for development of quantitative metabolomics analysis

Interest in extracting key intracellular metabolites poses the question of how best

to extract these metabolites without contamination or damaging the cell. A process of inhibiting the metabolic process in order to extract the relevant endometabolomes will be ideal. Quenching is a rapid metabolic inhibition process that enables extraction and quantitative analysis of intracellular metabolites (Dunn and Winder, 2011).

Quenching of the metabolic process could mean rapidly transferring a sample from the culture to a quenching solution at either low or high temperatures. Washing of the bacteria medium of any extracellular metabolites is recommended to eliminate contamination that will affect total metabolite concentration in targeted or non-targeted analysis. Although this washing step is important, it is recommended to check if it is required, due to cell wall permeability in the initial quenching process, if organic solvents and low temperatures are used. This can lead to a leakage of metabolites during the wash process and result in lower concentrations of intracellular metabolites. Extraction using cold organic solvent is well known but the ratios or percentages needed will depend on the type of microorganism used and which ratio works best in the extraction process (Li et al., 2013).

As stated earlier on, *C. autoethanogenum* is gram-positive and this means it has thick peptidoglycan layer in its cell wall. This could potentially pose problems with lysing of the cell wall with ease to extract the intracellular metabolites.

1.1.12.1 The use of Stable isotope labelled internal standards to overcome the matrix effect

The use of internal standards in quantitative bioanalysis is a common and accepted procedure. The purpose of internal standard is to correct for variability in sample preparation and instrumental response. Compounds used as stable isotope labelled internal standards (SIL ISs) have the atoms present in the analyte substituted with their corresponding stable isotopes. Examples include ^2H , ^{13}C , ^{18}O and ^{15}N . Ideal properties of SIL ISs are that they should behave similarly to the compound of interest in chromatographic separation and mass spectrometry

detection. It is therefore crucial that SIL ISs have the same retention time as its unlabelled (eg ^{12}C) analyte. Deuterium (^2H) labelled internal standards have different retention times to their hydrogen (^1H) analyte because the bond formed between ^2H isotopes and carbon is stronger than that between ^1H isotopes and carbon (Wieling, 2002). The increase in mass between ^1H and ^2H is 100% while that of ^{12}C and ^{13}C accounts for only 8% increase. However, ^{13}C , ^{18}O and ^{15}N labelled internal standards have physicochemical properties very similar to their corresponding analyte than ^2H labelled internal standards. Hence, the use of ^{13}C , ^{18}O and ^{15}N labelled internal standards are recommended instead of ^2H labelled internal standards where possible (Couchman and Morgan, 2011, Guo et al., 2007, Hewavitharana, 2011, Stokvis et al., 2005, Van Eeckhaut et al., 2009, Wang et al., 2007). The technique of incorporating isotope labelled internal standard into calibration mixtures or samples is termed isotope dilution mass spectrometry (IDMS), which has been reported to improve accuracy and quantitative linearity (Seifar et al., 2009). It has been shown that ^{13}C labelled internal standards are capable of improving ion suppression effects by at least 70 -80% (Berg and Strand, 2011).

1.1.12.2 Limitations of conventional normalisation methods

In order to report intracellular concentration per cell, the number of cells collected is required. There are techniques such as flow cytometry that can count the number of cells and can be used to determine cell volume. Cell volume is essential for determination of absolute intracellular metabolite concentrations and metabolic pathway modelling, which is the ultimate aim for systems biology (Usuda et al., 2010, Yizhak et al., 2010). The measurement of intracellular metabolite concentration is important for reporting (absolute) quantitative metabolic changes associated in microorganisms with perturbations of environmental or genetic conditions.

To report concentration per cell, the number of cells per specified volume is required to normalise the data. Direct cell counting of microorganisms in

samples is difficult due to their small sizes. Optical density measured at 600 nm (OD₆₀₀) is a well-known approach used to estimate the number of cells present in a culture. The use of OD₆₀₀ values to normalise bacteria cells in a quantitative analysis have been reported to be robust (Marcinowska et al., 2011) but this approach is applied to homogeneous cell cultures. Measurement of cell dry weight is also another conventional approach in determination of cell amounts and quantifying protein contents (Smart et al., 2010). The process of using cell dry weight can be time consuming and is not preferred for metabolomics studies involving large sample sets. These traditional approaches are prone to errors (Huege et al., 2011), for example OD₆₀₀ measurement does not give an indication of cell viability or cell volume. These disadvantages associated with the use of OD₆₀₀ and cell dry weight makes normalisation of metabolomics data using these parameters inadequate.

1.1.13 Summary

Choosing an analytical method to analyse a wide range of compounds can be challenging due to the diversity in physical and chemical properties of the compounds. There are two main stationary phases used in LC-MS metabolomics analysis, a hydrophobic stationary phase (C₁₈) or hydrophilic stationary phase (HILIC). The choice of stationary phase is dependent on the metabolites of interest. Metabolomics analysis of biological samples poses challenges such as matrix effects. The use of stable isotope labelled internal standards (SILISs) is known to rectify this matrix effect to good extent. Due to the cost of SILISs it is difficult to cover a wide range of metabolites, hence approaches such as generating SILISs in microbes by using isotope labelled substrate. In order to report the concentration of metabolites in bacteria, the cell volume of the bacteria is required to make this possible.

1.1.14 General project aims and objectives

This PhD research aims to develop a liquid chromatography isotope dilution mass spectrometry (LC-IDMS) method for metabolomics analysis in *C. autoethanogenum*. Below are the specific objectives:

-
- To develop LC-MS method for the analysis of polar intracellular metabolites in *C. autoethanogenum*.
 - To investigate suitable SILISs for absolute quantification of intracellular metabolites detected in *C. autoethanogenum*.
 - To report intracellular concentration by determining the cell volume of *C. autoethanogenum*
 - Apply the developed LC-IDMS method to study the metabolic effects of changing the bioreactor conditions on the production of key products of *C. autoethanogenum*. These studies include the effects of changes in the pH of the culture, change in gas (CO) flow rate, and the effect of limiting pantothenate and phosphate in the growth medium of *C. autoethanogenum*.

CHAPTER TWO

Method Development 1: investigation of LC-MS/MS methods for the quantitative analysis of intracellular metabolites in one carbon and associated pathways

2. Method Development 1: investigation of LC-MS/MS methods for the quantitative analysis of intracellular metabolites in one carbon and associated pathways

2.1 Introduction

2.1.1 General background

Systems biology aims to study the properties of metabolic network and its regulation within biological systems for quantitative mathematical modelling that integrate ‘omics’ data. Metabolomics data combined with transcriptomics and proteomics have become instrumental in optimisation of strains (Bolten et al., 2007, Douma et al., 2010, Oldiges et al., 2007, van Gulik, 2010). Metabolomics plays a very significant role in bridging the gap between phenotype and genotype since it reinforces the changes found in proteome, genome and metabolome to give a superior interpretation of the phenotype of an organism compared to other omics methods (Cascante and Marin, 2008). Since metabolome is as a result of the integration and amplification of the levels of other ‘omics’, it gives details on immediate snapshot of cell physiology and activity (Heux et al., 2015). Additionally, metabolites are functional entities in the cells whose levels are directly affected by genetic modifications and changes in the cellular environment (Carnicer et al., 2016). This explains why metabolomics has become an essential tool in metabolic engineering for improvement of microbial strain.

Advancement in methods based on MS have boosted the metabolomics field over recent years (Dettmer et al., 2007, Mashego et al., 2004, van Dam et al., 2002) with many applications in microorganisms (Carnicer et al., 2012, Kim et al., 2015) Advanced microbial metabolic flux analysis using ^{13}C isotope labelling (Tang et al., 2009) has greatly enhanced the possibilities for quantitative analysis of metabolic pathways (Oldiges et al., 2007). There is no universal method for the analysis of metabolome of bacterial cells due to the huge diversity of cell and chemical properties, hence protocols have to be

evaluated and adapted for each organism (Bolten et al., 2007, Canelas et al., 2008, Carnicer et al., 2012, Rabinowitz and Kimball, 2007).

The two main aspects of targeted metabolomics analysis are to find a suitable method (instrumental, extraction) and parameters for absolute quantification (internal standards, quenching and cell volume) (Figure 2.1). For this reason, the methodology chapter will be divided into two, one to cover instrumental and extraction methods (Chapter 2) and the second chapter to cover the quantitative aspect (Chapter 3).

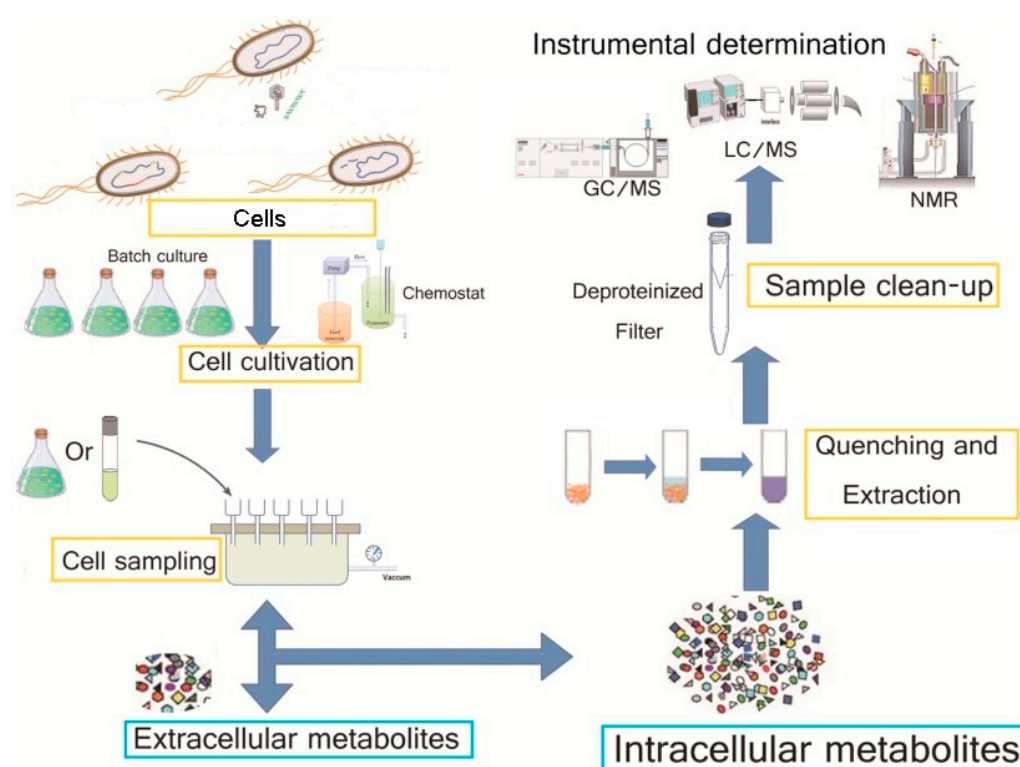


Figure 2.1 An overview of a typical work flow for quantitative metabolic pathway profiling (Shen et al., 2016)

2.1.2 Mass spectrometry applications in targeted metabolomics analysis

The metabolic pathways of microorganisms consist of a range of chemically diverse, low molecular weight and often unstable metabolites that pose a

significant challenge for quantitative analysis in an intracellular environment (Garratt et al., 2005). The hydrophilic nature and low concentrations (Schiesel et al., 2010) of most of these metabolites poses limitations with frequently used analytical methods such as gas chromatography mass spectrometry (GC-MS) where these compounds would require derivatisation and nuclear magnetic resonance (NMR) which is less sensitive. Liquid chromatography mass spectrometry (LC-MS) addresses this low sensitivity issue and provides a suitable combination of robustness and versatility and has been used most widely for analysis of these types of polar metabolites (Buescher et al., 2009).

Previous analyses of multiple water-soluble metabolites have been reported using different approaches such as ion pair reversed phase (IPRP) (Cordell et al., 2008, Guo et al., 2011, Luo et al., 2007, Seifar et al., 2009), chemical derivatisation reversed phase liquid chromatography (Halliday et al., 2010), ion exchange liquid chromatography (Edwards et al., 2007, Shi et al., 2002) and HILIC (Cubbon et al., 2007, Gika et al., 2012, Huang et al., 2012, Kiefer et al., 2008, Li et al., 2012, Bajad et al., 2006, Fairchild et al., 2010). The afore mentioned methods were used to successfully monitor a wide range of metabolites but there are drawbacks to the methods. The use of ion pair is less favourable when mass spectrometry detection is used due to increased ion suppression /enhancement. HILIC presents a better approach to identifying compounds without the use of ion pair reagents, but the methods used here covered few metabolite range and acidic mobile phases were used that hindered the detection of certain metabolites.

2.1.3 Analytical challenges in method development for analysis of polar metabolites

Metabolites vary widely in terms of chemical structure and polarity (Fairchild et al., 2010), hence many published metabolomics MS methods analyse in both positive and negative modes separately which have been found to provide wider coverage of metabolites (Lu et al., 2008). Normally HILIC (polar retention) is complementary with reversed phase liquid chromatography (RPLC) (nonpolar retention) due to their combined ability to provide coverage of a wide range of

metabolites. (Nunez et al., 2008, Sato et al., 2010, Wang et al., 2008). Examples of methods used were LC-MS by (Amador-Noguez et al., 2011, Luo et al., 2007), LC-MS/MS by (Bylund et al., 2007) and HILIC-MS/MS by (Huang et al., 2012) (Table 2.1). The use of ZIC-HILIC for analysis of polar metabolites for global metabolomics approach was also reported by (Creek et al., 2011).

Analyses of metabolites present in the activated methyl cycle (AMC) have been reported by (Halliday et al., 2010, Bajad et al., 2006) and folate cycle by (Garratt et al., 2005, Lu et al., 2007) but no single LC-MS method has been reported that measures all metabolites in these pathways simultaneously. The possible approaches to provide simultaneous measurements of these pathways are either IPRP chromatography or HILIC coupled with either mass spectrometry or tandem mass spectrometry.

Folates are important in one carbon unit transfers within the cell (Hanson and Roje, 2001). They are crucial for the synthesis of coenzyme A, nucleic acids and methionine (Rebeille et al., 2006). Folates are structurally made up of pteronic acid which is covalently bound to poly-glutamate chain. The pteronic acid can undergo oxidation/ reduction and one carbon addition/ elimination to produce diverse related compounds. The analysis of folates in cells can be challenging due to the different folate modifications that could occur: methylation, poly-glutamation and oxidation/reduction. It is therefore recommended that extraction buffers with antioxidants such ascorbic acid and 2-mercaptoethanol are used to reduce the interconversion of the folates during extraction (Lu et al., 2007).

2.1.4 Background of HILIC method

HILIC provides an advanced technique in separation and retention of polar compounds and is also suitable for various groups of compounds such as sugars, nucleotides, amides, organic acids and amino acids (Bajad et al., 2006, Creek et al., 2011, Gika et al., 2012). HILIC is a form of normal phase liquid chromatography which uses hydrophilic stationary phase with eluents similar to ones used in reversed-phase (Cubbon et al., 2007, Guo and Gaiki, 2005). Analyte retention in HILIC is believed to be as a result of partitioning of the analyte

between a water enriched layer of stagnant eluent on a hydrophilic stationary phase and a hydrophobic bulk eluent, with the main eluting components being 5-40% acetonitrile (Alpert, 1990). In HILIC the strongly eluting solvent is water and this gives a number advantages over conventional normal phase chromatography;

1. Eluent preparation is less complicated since the need of total control over the solvent water content is omitted.
2. Normal phase chromatography eluents are non-polar (mostly in hexane), hence polar analytes have low solubility in these eluents.
3. The mobile phases for C18 are still used for HILIC, which are more MS compatible than non-polar solvents.

The HILIC method was optimised from a previously reported method (Kim et al., 2015). HILIC does not require the use of ion pair reagents and this makes it more suitable for mass spectrometry platforms in comparison to ion pair chromatography

Table 2.1 An overview of published methods for analysis of polar and highly charged metabolites

Compounds/Samples	Method	Column	Mobile Phases	Flow rate	Run time(mins)	Reference
Pyruvate and Lactate	LC-MS/MS	21mmX 50mm sunfire C18 column with 5µm particle size	A (50% deionised water) B (50% ACN,0.1% FA)	200µL/min	3.5	(Chuang et al., 2009)
Pyruvate and lactate	RPLC-MS	C18,(2.1mm X75mm) 3µm, C18,(2mm x 4mm) (3µm)guard	A(0.1% v/v FA/H2O), B (0.1% v/v FA/ACN)	200µL/min	25	(Uran et al., 2007)
Pyruvate, Lactate, Fumarate Succinate, Citrate and Isocitrate	LC-MS/MS	C610-H (300mm X7.8mm)	methanol 0.01% formic acid (10:90)	400µL/min		(Bylund et al., 2007)
Succinate, malate,	RPIPLC-MS/MS	Synergi Hydro Column (4µm particle size, 150mm x 2mm)	A (10mM TBA + 15mM AA in 97:3 Water: methanol B (100% methanol)	200µL/min	50 mins	(Lu et al., 2008)
Citrate, fumarate, malate, succinate, oxaloacetate	HILIC	Luna Aminopropyl (NH ₂) (250mm x2mm, 5µm particle size)	A (10mM NH ₄ Ac in water with AA, pH 5.8), B (ACN)	150µL/min	50 mins	(Bajad et al., 2006)
Citrate,fumarate,succinate and malate	LC-ESI-TOFMS	C18, (2.1 X150mm) 5µm particle size	A (H2O, 1% ACN, 0.1% FA), B (ACN,1% H2O, 0.1% FA)	250µL/min		(Jaitz et al., 2011)

Pyruvate, citrate, fumarate, succinate and malate	HPLC-MS/MS	C18 (250mm X 3mm) 3µm particle size	0.1% v/v formic acid	400µL/min		(Flores et al., 2012)
Pyruvate and 3-methyl-2-oxobutyrate	LC/MS/MS	agilent zorbax eclipse XDB-C18 (150x4.6mm,5µm)	A (30:70 mix of 0.2% FA in ACN and 4mM ammonium acetate),B(0.2% FA in ACN)	500µL/min	60	(Hidalgo et al., 2013)
Folates	RPIP LC-MS/MS	Luna C18(2) 150X 2.0mm with C18 Securityguard column	A (MeOH/H ₂ O: 5:95, v/v with 5mM DMHA, pH 8.1 B, (MeOH with 5mM DMHA	200µL/min		(Garratt et al., 2005)
AMC metabolites	LC-ESI-MS/MS	Luna ODS C18(2) 5µm, 150 X 2.0mm with guard	A (aqueous 10mM ammonium formate), B (10mM ammonium formate in metanol)	200µL/min	30	(Halliday et al., 2010)
Methotrexate polyglutamates	HPLC	C18 (25cm X 4.6mm) 5µm particle size	A (10mM ammonium acetate pH 6.5/2mL/LH ₂ O ₂ , B (acetonitrile)	1000µL/min	20	(Dervieux et al., 2003)
Isobutraldehyde	GC-MS	A DB-WAXETR capillary column (60 m × 0.25 mm I.D., film thickness 0.5 µm)		1000µL/min		(Ferreira et al., 2006)

<i>E. Coli</i> metabolites (acetyl CoA)	RPIP LC-MS/MS	LC C18 column 5µm ODS- cartridge (100X3mm)	A (5mM hexylamine, pH 6.3) B (9:1 methanol: 10mM NH ₄ Ac, pH 8.3)	400µL/min		(Coulier et al., 2006)
Citrate, lactate, pyruvate, succinate	HILIC MS/MS	ZIC-HILIC column (250mm x4.6mm, 5µm)	A (ACN),B (10mM NH ₄ Ac)	500µL/min	14	(Huang et al., 2012)
Small polar molecules	HILIC MS/MS	BEH amide column (2.1 mm x 150 mm)	A(ACN/water,95:5,0.1 % f.a, 0.075 % NH ₄ OH), B(ACN/water,2:98,0.2 % f.a, 0.1 % NH ₄ OH)	500µL/min	42	(Gika et al., 2012)

2.1.5 Project Aims and objectives

- Identify key metabolites from one carbon metabolic pathways and evaluate reversed-phase ion pair chromatography and HILIC methods
- Develop and validate the most suitable method from the two stated above for identification of multiple classes of metabolites present in other key metabolic pathways.
- Adapt a method suitable for the analysis of folates in *C. autoethanogenum*

2.2 Materials and Methods

2.2.1 Experimental reagents and methods

LC-MS grade solvents acetonitrile and methanol were obtained from Fisher Scientific (Bishop Meadow Road, Loughborough). *N,N*-dimethylhexylamine (DMHA) and oxaloacetic acid were purchased from Alfa Aesar (Port of Heysham Industrial park, Lancashire). Tributylamine (TBA), citric acid monohydrate (95% - 102%) and DL-lactic acid (-90% T) were purchased from (Sigma Aldrich UK). Succinic acid (99%), glyoxylic acid 50% in water, 3-methyl-2-oxobutanoic acid sodium salt (98+ %), DL-isocitric acid trisodium salt hydrate (95%), DL- malic acid (99+ %), fumaric acid (99+ %) and pyruvic acid sodium salt (99+ %) were purchased from Fisher Scientific (Bishop Meadow Road, Loughborough).

2.2.2 Preparation of standards for calibration and direct MS infusion

A stock solution of all target metabolites (28 in total) (Table 2.8) obtained from the one carbon and associated pathways were prepared by dissolving standard stocks in methanol or water to a final concentration of 1 mM. Eight calibration mixtures were prepared by serial dilution ranging from 0.0001 μ M to 150 μ M. Individual dilutions were made to 10 μ M concentration for direct infusion to optimise precursor and product ions.

Table 2.2 optimised Precursor and product ions for the set of compounds monitored in negative ionisation mode.

COMPOUNDS	PRECURSOR ION m/z	PRODUCT IONS
	[M-H] ⁻	m/z
Pyruvic acid	86.98	43.1, 59.0, 58.0
Lactic acid	89.041	43.1, 71.2
Fumaric acid	114.985	71, 58.8
Succinic acid	116.982	73, 98.9, 72.3
Malic acid	132.998	114.8, 71.1, 73
Isocitric acid	191.002	110.9, 173, 73
Citric acid	190.998	110.9, 87.1, 85.1
Glyoxylic acid	72.876	45, 30.2
Acetylphosphate	136.983	90.8, 45.1, 73.1
Acetoin	86.915	86.915
Oxaloacetic acid	130.987	86.9, 85.78.9
2-Ketoisovaleric acid	114.948	70.6, 87.1

2.2.3 LC Method development conditions

2.2.3.1 Preliminary method development on C₁₈ columns

LC-MS/MS method was initially developed on Agilent Poroshell 120 SB-C₁₈ column 3 µm particle size (2.1 mm x 50 mm) using acidic mobile phases A and B containing 0.1 % formic acid in water and Acetonitrile respectively. The gradient of elution started with 15 % mobile phase B to 100 % of mobile phase B in 10 min. This was held for 2 min and gradient returned back to 15 %B and equilibrated for 2 min. Further optimisation of the LC-MS/MS method was investigated using an Agilent Zorbax SB aq C₁₈ column (2.1 mm x 100 mm). Mobile phase A consisted of 30:70 acetonitrile:water in 4 mM NH₄OAc and 0.2 % formic acid; mobile phase B was 0.2 % formic acid in ACN (Hidalgo et al., 2013).

Table 2.3 Final gradient of elution for reversed phase chromatography using Agilent Zorbax SB aq C₁₈ column (2.1mmx100mm) at a flow rate of 300 µL/min

Time (min)	Mobile phase A (%)	Mobile phase B (%)
0.0	95	5
5.0	50	50
10.0	0	100
14.0	0	100
17.0	50	50
20.0	95	5
23.0	95	5

2.2.3.2 Preliminary method development using reversed phase ion pair (RPIP) chromatography

LC-MS/MS method using IPRP chromatography was investigated. Dimethylhexylamine (DMHA) was used as an ion-pairing reagent (IPR) to improve chromatographic separation. Initial concentration of 5 mM DMHA was added to both mobile phases but was later increased to 10 mM. The preliminary LC method was further developed by replacing DMHA in the mobile phases with tributylamine (TBA) at the final concentration of 10 mM (Lu et al., 2008, Luo et al., 2007). The column used initially was Phenomenex Luna 5µm particle size C₁₈ (2.0 mm X 150 mm) which was later changed to ACE 3 C₁₈ 3 µm particle size (2.1 mmx 150 mm).

Table 2.4 Final gradient of elution for RPIP chromatography using ACE 3 C₁₈ column, 3 µm particle size (2.1mm x 150mm) at a flow rate of 300 µL/min

Time (min)	Mobile phase A (%)	Mobile phase B (%)
0.0	100	0
2.0	100	0

6.0	85	15
10.0	70	30
12.0	70	30
18.0	100	0
25.0	100	0

*Mobile phase A – 10 mM TBA in 1:9 v/v ACN:water pH 6.3, Mobile phase B-
Mobile phase A – 10 mM TBA in 9:1 v/v ACN:water.*

2.2.3.3 Preliminary HILIC method development

The development of the HILIC method involved optimisation of a method that had already been reported (Kim et al., 2015). The column used was ZIC – pHILIC, 5µm PEEK (2.1 mm x4.6 mm) with guard column PEEK (20x 2.1 mm) (Merck Sequant, Roxley Green Business Park, Watford, UK). A linear gradient starting at 80 % B to 5 % B for 10 min, 7 min from 5 %B to 80 %B and 7 min re-equilibration at 80 % B was used to elute the analytes at a constant flow rate of 300 µL/min maintaining column temperature at 45 °C. Mobile phase A was 20 mM ammonium carbonate in water (pH 9.1) and 100 % acetonitrile was used as mobile phase B. Samples were kept at 4 °C during analysis with an injection volume of 10 µL.

2.2.4 Optimised analysis of folates using HILIC-MS/MS

The LC conditions are similar to the HILIC conditions except for the reduction in runtime from 24 mins to 17 mins. Chromatographic separations were carried out using ZIC-pHILIC 150 mm × 4.6 mm, 5 µm particle size column (Merck Sequant, Roxley Green Business Park, Watford, UK). Column temperature was maintained at 45 °C. Mobile phase A was 20 mM ammonium carbonate in water (pH 9.1) and 100 % acetonitrile was used as mobile phase B. An injection volume of 10 µL was used and samples were kept at 4 °C in the auto sampler prior to analysis.

Positive mode ionisation was used for analysis of the tetrahydrofolates with HILIC. The source conditions used were; ion spray voltage of 4200 V, gas temperature of 500°C, curtain gas (nitrogen) set at 10, ion source gas 1 at 20 and ion source gas 2 at 0, collision gas (CAD) at 5, Entrance potential at 10 V, declustering potential at 86 V, collision energy at 21 V and collision exit potential at 8 V.

Table 2.5 Final gradient of elution for analysis of folates on HILIC method using ZIC –pHILIC, 5µm PEEK (2.1mm x4.6 mm) at a flow rate of 300 µL/min.

Time (min)	Mobile phase A (%)	Mobile phase B (%)
0.0	20	80
8.0	95	5
10.0	20	80
17.0	20	80

2.2.5 Preliminary method development for analysis of folates

The method development started with choosing a suitable buffer for extraction of tetrahydrofolates. Four different buffers were tested for stability and recovery.

1. 20 mM NH₄OAc in 50:50 v/v methanol:water + 0.1 % ascorbic acid
2. 20 mM NH₄OAc in 80:20 v/v methanol:water + 0.1 % ascorbic acid
3. 20 mM NH₄OAc in 50:50 v/v methanol:water + 0.1 % 2-mercaptoethanol
4. 20 mM NH₄OAc in 80:20 v/v methanol:water + 0.1 % 2-mercaptoethanol.

2.2.6 Optimisation of MS Conditions

2.2.6.1 Final optimised LC MS/MS conditions for IPRP chromatography for analysis of folate and set of metabolites

All analyses were done using an optimised LC-MS/MS method. Ion-pair reversed-phase chromatography and tetrahydrofolates analyses were carried out on a hybrid triple quadrupole mass spectrometer (4000 QTRAP) from Applied Biosystems (Manchester, UK) coupled with Shimadzu LC system (Wolverton Mills South, Milton Keynes). Electrospray ionisation operated in negative mode with source conditions as follows: ion spray voltage of -4500V, gas temperature of 500°C, curtain gas (nitrogen) set at 10, ion source gas 1 at 20 and ion source gas 2 at 10, collision gas (CAD) at 6, Entrance potential at -10 V, declustering potential at -45 V, collision energy at -12 V and collision exit potential at -3 V. Q1 and Q3 masses were determined by direct infusion of 10 µM standards with the results obtained listed in (Table 2.6). Collision energy (CE), de-clustering potential (DP) and collision exit potential (CXP) were optimized for each product ion and these values used in the LC-MS/MS analysis. An ACE 3 C₁₈ column with 3 µm particle size, 100 Å pore size, a length of 150 mm and a diameter of 2.1 mm was used. A binary gradient starting at 95 % of A to 10 % A for 7 min, 7-9 min from 10 % A to 95 %A and then re-equilibrating for 6 min at a flow rate of 200 µl/min. Mobile phase A was 10 mM TBA in 1:9 v/v acetonitrile: water (pH =6) and mobile phase B was 10 mM TBA in 1:9 v/v water: acetonitrile. 5 µL injection volume was used and column temperature maintained at 45 °C.

2.2.7 Data analysis for LC-MS/MS raw data for tetrahydrofolates and key metabolites

Raw data obtained for LC-MS/MS analysis were processed using Analyst software (AB SCIEX). Compound identification was achieved by selecting the precursor ions and their corresponding product ions.

2.3 Results and discussion

2.3.1 Analytical strategies

A thorough search of literature for a suitable method as represented in Table 2.1 did not provide a method fit for this particular analysis. A new method was then developed to for the analysis of these polar intermediates in the one carbon metabolic pathway of *C. autoethanogenum*. Two methods, HILIC and IPRP chromatography were investigated to choose the method most suited for this analysis.

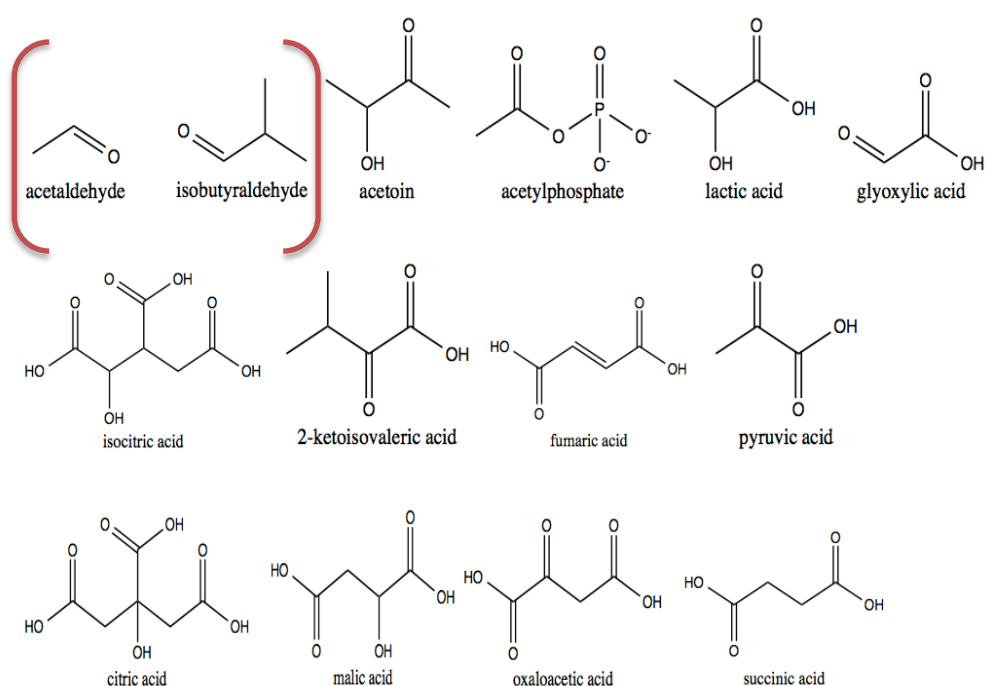


Figure 2.2 initial selections of metabolites for LC-MS/MS method optimisation. Acetaldehyde and isobutyraldehyde could not be optimised using LC-MS/MS due to their volatile nature.

2.3.1.1 Mass spectrometry optimisation

In Table 2.2 the MS parameters such as product and precursor ions used in the multiple reaction monitoring (MRM) and optimised voltages of the key metabolites investigated were shown. In all instances the $[M-H]^-$ was the

monitored ion due to the ion suppression effects when monitoring $[M+H]^+$ when ion pair reagents are used. In the case of the majority of the organic acids the most intense product ions were due to decarboxylation or elimination of water. Analogous compounds tend to produce similar fragmentation patterns due to similarity in structure that could make the non-specific MRM transitions to give a false signal if the collision-induced dissociation (CID) in the interface is low. An example is the insource fragmentation of malate ($m/z=133$ $[M-H]^-$), which gives 115 ($[M-H-18]^-$) as the main product ion as a result of water loss. It then decomposed in the collision cell to generate $m/z=71$ as product ion ($[M-H-18-44]^-$ because of decarboxylation. This same MRM transition was observed for fumarate ($115>71$) (Luo et al., 2007). This interference was noticeable when selecting fumarate ion around the retention time for malate. Further optimisation was done to differentiate fumarate and malate through retention times. Multiple transitions were obtained for each compound in negative ionisation mode but only the most abundant transition was reported in Table 2.6. Due to smaller molecular mass of some of the compounds monitored, the associated product ions were below the calibrated mass of $m/z=50$. Overall, it was possible to assign unique precursor/products ion pairs to all of the selected analytes.

Table 2.6 Precursor and product ions of metabolites used in IPRP chromatography and HILIC analysis in negative ionisation mode.

Compound	Precursor ion	Product ion	CE (V)
Fumaric acid	115.0	71.0	-12
Lactic acid	89	43	-18
2-Ketoisovalerate	115	70.6	-12
Malic acid	133	115.1	-16
Citric acid	191	110.9	-18
Isocitric acid	191	173	-14
Oxaloacetic acid	131	86.9	-12
Pyruvic acid	87	43	-12
Succinic acid	117	73	-16

α -Ketoglutaric acid	145	101	-20
ADP	426	159	-30
Acetyl CoA	808.1	408	-30
AMP	346	79.1	-30
ADP	426	159	-30
ATP	506	159	-30
NAD	662	540.1	-30
NADH	664.1	408	-20
NADP	742.1	620	-20
NADPH	744.1	408	-30
Adenosine	266	134.2	-46
Adenine	134.1	106.8	-24
Homoserine	118.1	99	-14
Homocysteine	134.1	91.7	-16
SAM	398		-30
SAH	383.2	134.1	-30

2.3.1.2 Optimisation of ion pair method

Tetra-alkylammonium salts have been used as ion pair reagents in the past to improve separation of anionic compounds. However, due to the many drawbacks associated with the use of non-volatile IPRs such as increased background contribution and ion suppression, when MS detection is used we sought an alternative approach. In order to avoid these drawbacks but ascertain resolution and sufficient retention of the anions, the option was to use volatile alkylamines (Garratt et al., 2005, Haandel et al., 2009, McGinnis et al., 2013) as IPRs. The addition of alkylamines in the mobile phases causes interferences with full scan modes in the orbitrap but this is much reduced when using the MRM function on the triple quadrupole mass spectrometry. Two IPRs with different alkyl chains were investigated for separation of metabolites. Initial concentration of 5 mM of dimethylhexylamine (DMHA) (Qian et al., 2004) and tributylamine

(TBA) (Buescher et al., 2009) were used with pH adjusted to 6.2 using acetic acid. A final concentration of 10 mM gave better peak separation with both DMHA and TBA, but TBA gave better peak shapes at this concentration Figure 2.3. Also, the use of organic solvents for the mobile phases was investigated. In this instance acetonitrile was a better organic solvent than methanol due to improved chromatographic separation and peak shapes. The chromatogram run time was also optimised to 25 min (Table 2.4) compared to previously reported method in excess of 42 min (Luo et al., 2007). Due to the increase in ion suppression/enhancement with IPR when monitoring in positive mode of ionisation, an alternative method was investigated which did not require the use of IPR to enable the detection of more metabolites better detected in positive ionisation mode.

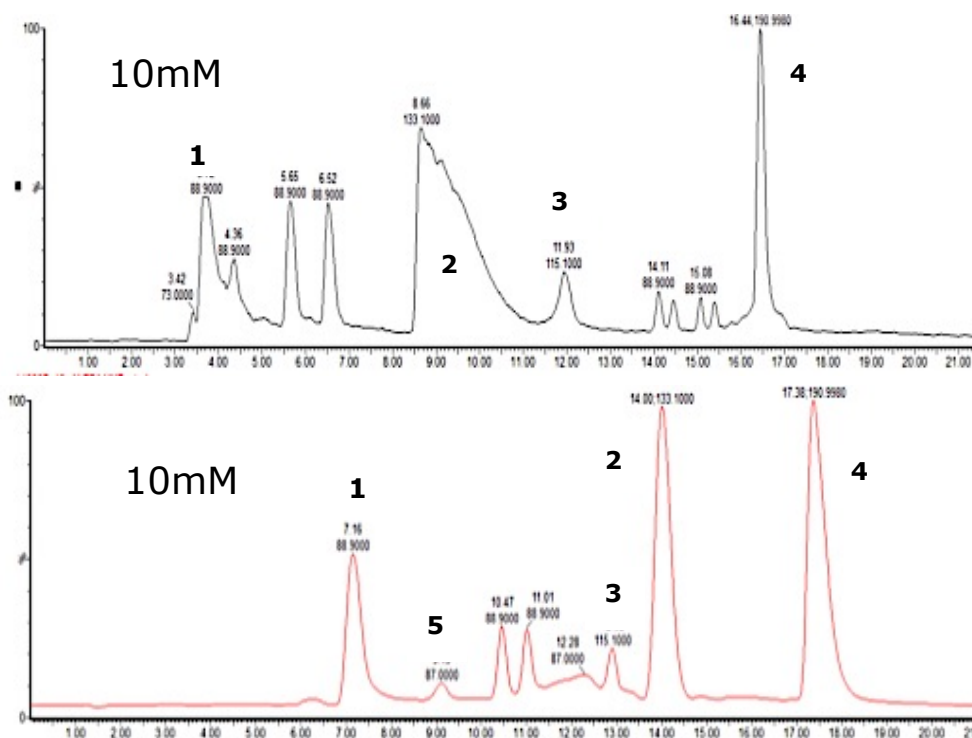


Figure 2.3 Chromatograms ahowing separation obtained when 10 mM DMHA and TBA incorporated into the mobile phases using representative *metabolites*. [1-lactate, 2- malate,3-Fumarate, 4-citrate, 5- pyruvate]

2.3.1.3 Validation and comparison of the two methods

The validation results for IPRP and HILIC methods are reported in Table 2.7 and Table 2.8 respectively. The calibration graphs for RPIP method showed excellent linearity with correlation coefficients (R^2) over three significant figures greater than 0.9877. Calibration graphs for HILIC method gave $R^2 > 0.9870$. The limit of detection (LOD), μM and limit of quantification (LOQ), μM were reported in the tables stated above by evaluating signal to noise ratios (S/N) of 3:1 and 10:1 respectively.

2.3.1.3.1 Ion pair reversed phase chromatography

In Table 2.7 LODs varied from 0.0001 to 25 μM and LOQs 0.0005 to 150 μM . Pyruvate showed lower sensitivity, which could be attributed to the decreased intensity of the product ion (m/z -43). On the whole, LODs of organic acids using the IPRP method were more sensitive than reported in (Bajad et al., 2006, Cui et al., 2013, Huang et al., 2012, Lu et al., 2006, Luo et al., 2007). Comparison of LODs ATP and NADP with similar methods showed that our LODs were ten times lower than reported by (Bajad et al., 2006, Cordell et al., 2008, Cui et al., 2013). Other metabolites such as adenosine and adenine had comparable LODs with those reported by Cordell (Cordell et al., 2008). All metabolites were monitored in negative mode ionisation which is not usually a problem for detecting amino acids but in this instance amino acids were not detected probably due to the presence of the ion pair reagent; hence no validation was reported for them with this method. To evaluate the IPRP method for repeatability, three concentration levels were chosen for the standard mixture which were injected six times and relative standard deviation (RSD) better than 21% for 2 μM , 9 % for 10 μM and 12 % for 50 μM were reported in Table 2.7.

2.3.1.3.2 Hydrophilic interaction liquid chromatography (HILIC)

In Table 2.8 validation of the HILIC method was shown. LODs reported ranged from 0.001 μM to 2 μM . LOQs varied from 0.005 μM to 10 μM . LODs reported by our HILIC method for organic acids were ten times lower than the LODs reported for the HILIC- MS/MS by (Bajad, 2006; Huang, 2012; Lu, 2006). Comparison of LODs of our HILIC method for eight out of the nine cofactors reported were ten times higher than those reported by (Cui et al., 2013, Luo et al., 2007) but two times lower compared to IPRP method by (Cordell et al., 2008) and similar to HILIC- MS/MS methods by (Bajad et al., 2006). The LODs reported for amino acids using our HILIC method was more sensitive than reported here (Lu et al., 2006). Repeatability of the HILIC method was evaluated as above and RSD better than 15 % for 2 μM , 13 % for 10 μM and 13 % for 50 μM were reported.

Ion pairing chromatography (IPC) and ion exchange chromatography are usually incompatible with mass spectrometry detections due to high concentration levels of salt present in the eluting mobile phases (Lu et al., 2008). The application of ion pairing reagents (IPR) could be limited due to long equilibrating times and strong bond it forms with the column. HILIC provides an advanced technique in separation and retention of polar compounds and also suitable for various groups of compounds including sugars, nucleotides, amides, organic acids and amino acids (Bajad et al., 2006, Gika et al., 2012).

In summary 19 metabolites were measured using the RPIP method and 28 were measure using the HILIC method. The HILIC method was more sensitive than the IPRP method with lower LODs reported.

Table 2.7 Results from initial ion pair reversed phase method validation. The validation results show the different linear ranges, linearity, limit of detection (LOD), limit of quantification (LOQ), precision and accuracy for each metabolite. The validation was conducted to conform to the FDA guidelines. According to the FDA guidelines, precision must be below 15 % and accuracy 100 ± 15 %. LOD and LOQ are used to test how sensitive a method is, the lower the concentration, the more sensitive the method for that particular compound.

Compound	Linear range (μM)	R^2	LOD (μM)	LOQ (μM)	L=2 μM	Precision (%)		Accuracy (%)		
						M=10 μM	H=50 μM	L=2 μM	M=10 μM	H=50 μM
2-Ketoisovaleric acid	0.1-150	0.9999	0.5	1	6.1	3.0	2.1	105.7	102.6	99.6
Pyruvic acid	2-150	0.9993	25	>150	5.4	2.2	0.5	96.9	106.3	100.9
Lactic acid	0.1-150	0.9998	10	100	6.7	2.3	6.7	101.0	103.9	103.7
Succinic acid	0.1-150	0.9958	1	2	2.9	2.8	5.9	111.3	113.6	111.2
Fumaric acid	0.1 -150	0.9989	0.5	2	4.5	1.2	0.9	116.1	99.5	108.6
Malic acid	0.1-150	0.9994	0.01	0.1	5.1	2.9	2.2	117.0	101.8	104.0
Citric acid	0.1-150	0.9974	0.0001	0.0005	4.9	3.4	3.4	99.2	104.9	101.9
Isocitric acid	0.1-150	0.9985	0.0001	0.0005	0.4	1.4	2.3	105.6	105.7	106.5
α -Ketoglutaric acid	0.5-150	0.9992	1	2	14.9	7.8	6.2	104.3	102.4	107.2
Oxaloacetic acid	0.1-150	0.9991	0.5	1	9.4	3.5	2.6	105.8	101	107.9

Compound	Linear range (μM)	R ²	LOD (μM)	LOQ (μM)	Precision (%)			Accuracy (%)		
					L=2μM	M=10μM	H=50μM	L=2μM	M=10μM	H=50μM
Adenosine	1-150	0.9976	5	10	9.6	8.2	7.5	115.9	102.5	104.4
Adenine	2-150	0.9877	10	25	6.9	3.7	4.4	112.0	99.7	91.0
AMP	0.5-150	0.9981	0.5	1	20.8	3.9	1.7	91.8	100.3	105.5
ADP	0.1-150	1.0000	0.1	1	5.6	4.3	1.1	91.4	103.9	99.9
ATP	0.1-150	0.9999	0.001	0.05	3.0	3.9	3.1	109.8	98.4	99.8
NAD	0.1-150	0.9998	0.5	2	5.1	3.1	0.8	90.9	100.1	100.3
NADP	0.1-150	0.9992	0.0001	0.0005	13.6	7.0	0.9	110.5	99.3	96.8
NADPH	0.1-150	0.9963	0.5	2	8.7	5.0	5.8	114.7	95.3	94.3
Acetyl CoA	1-150	0.9950	1	2	4.5	2.6	11.8	99.1	98.9	108.5

L – Low, M – Medium, H – High

Table 2.8 Results from initial HILIC method validation. The validation results show the different linear ranges, linearity, limit of detection (LOD), limit of quantification (LOQ), precision and accuracy for each metabolite. The validation was conducted to conform to the FDA guidelines. According to the FDA guidelines, precision must be below 15 % and accuracy 100 ± 15 %. LOD and LOQ are used to test how sensitive a method is, the lower the concentration, the more sensitive the method for that particular compound

Compound	Linear range (μM)	R^2	LOD (μM)	LOQ (μM)	L=2 μM	Precision (%)		L=2 μM	Accuracy (%)	
						M=10 μM	H=50 μM		M=10 μM	H=50 μM
2-Ketoisovaleric acid	0.0001-150	0.9994	0.05	0.25	3.1	4.6	2.6	98.4	100.6	98.9
Pyruvic acid	0.0001-150	0.9945	0.05	0.5	5.1	2.5	4.0	101.8	96.4	93.6
Lactic acid	1-150	0.9986	0.1	0.5	12.0	9.8	2.0	139.5	109.2	95.5
Succinic acid	0.5-150	0.9992	0.1	0.5	2.5	2.3	1.6	100.7	102.8	97.7
Fumaric acid	0.0001-150	0.9986	0.05	0.5	3.5	2.5	2.8	101.5	98.5	97
Malic acid	0.1-150	0.9998	0.02	0.25	7.5	3.3	0.9	79	105.2	100.3
Citric acid	0.0001-150	0.9987	0.001	0.005	2.6	3.7	6.1	79.3	101.1	104.1
α -Ketoglutaric acid	0.0001-150	0.9991	0.001	0.01	4.9	5.2	3.1	107.4	94.2	93
Oxaloacetic acid	5-150	1.0000	1	2	-	1.2	1.3	-	99.0	100.5

Compound	Linear range (μM)	R ²	LOD (μM)	LOQ (μM)	L=2μM	Precision (%)			Accuracy (%)	
						M=10μM	H=50μM	L=2μM	M=10μM	H=50μM
Acetyl phosphate	0.5-150	0.9948	1	5	1.4	12.1	4.3	96.0	88.0	105.7
Adenosine	0.001-150	0.9950	0.01	0.1	9.9	7.7	3.1	93.6	94.4	102.3
Adenine	0.0001-150	0.9956	0.01	0.1	5.0	10.9	5.1	76.7	99.3	114.5
Homocysteine	1-150	0.9907	0.1	1	-	4.8	10.2	-	88	95.9
Leucine	0.0001-150	0.9951	0.001	0.01	8.9	3.2	8.8	77	102.5	105.6
Methionine	0.0001-150	0.9982	0.05	0.5	5.2	5.4	6.6	79.0	104.7	104.6
Valine	1-150	0.9986	0.001	0.01	15.0	7.9	4	62.7	101.7	100.4
Homoserine	0.1-150	0.9914	0.1	0.5	3.5	8.6	12.7	107.1	78.7	106.9
S-Adenosyl-L-methionine	0.1-150	0.9963	0.1	0.5	0.7	6.3	9.5	83.1	92.5	100.5
S-Adenosyl-L-homocysteine	0.1-150	0.9927	0.001	0.005	8.2	2.4	5.2	102.7	104.0	106.2
AMP	0.5-150	0.9957	0.5	2	-	12.5	7.2	-	134.4	102.1
ADP	2-150	0.9986	0.1	1	11.7	3.5	2.2	134.2	98.9	103.5
ATP	2-150	0.9957	2	5	3.3	9.2	6.4	214	97.3	94.2

Compound	Linear range (μM)	R ²	LOD (μM)	LOQ (μM)	L=2μM	Precision (%)			Accuracy (%)		
						M=10μM	H=50μM	L=2μM	M=10μM	H=50μM	
NAD	0.5-150	0.9971	0.1	0.5	14.6	9.4	4.5	81.0	108.9	105	
NADH	5-150	0.9998	2	5	-	2.2	1.0	-	98.0	101.3	
NADP	1-150	0.9949	1	2	3.3	5.2	5.8	176	98.5	96.3	
NADPH	1-150	0.9941	2	10	3.5	6.2	7.3	168	103.3	97.1	
Acetyl CoA	5-150	0.9870	0.5	2	-	3.7	1.7	-	103.4	98.9	
Acetoacetyl CoA	5-150	0.9996	1	5	-	1.8	1.2	-	100.1	99.7	

L – Low, M – Medium, H – High

2.3.2 Method development of folates and poly-glutamated folates

Precursor and daughter ions were optimised by direct infusion of pre-dissolved standards of folates at 25 nM concentrations. The optimised parameters are listed in Table 2.9 for negative ionisation mode. These product and precursor ions were used in the initial optimisation of IPRP and HILIC methods.

Table 2.9 Precursor and product ions of the folates monitored

Compound	Polarity	Precursor ion	Product ion	CE (V)
5,10-Methenyl THF acid	Negative	454.2	324.9	-28
5-Methyl THF acid	Negative	458.2	328.9	-34
5,10-Methylene THF acid	Negative	456.2	135.0	-44
10-Formyl THF acid	Negative	472.2	315.1	-38
DHF acid	Negative	442.2	175.8	-38
THF acid	Negative	444.2	314.9	-29
Folic acid	Negative	440.2	174.8	-50
Methotrexate (IS)	Negative	453.1	324.1	-30

IS – Internal standard

Extraction buffers were evaluated for the extraction of folates. Due to the unstable nature of folates and their ability to interconvert, the presence of an antioxidant and the right pH was crucial. The antioxidants investigated here were ascorbic acid (AA) and 2-mercaptoethanol (M-EtOH) at a concentration of 0.1 %. The pH was adjusted by adding 20 mM of NH₄OAc to all extraction buffers (Lu et al., 2007). The combinations of solvents used were 50:50 methanol: water (MW) and 80:20 MW. Recovery analysis was conducted by spiking folate standard mixtures to *C. autoethanogenum* cell pellets prior to extraction (B) and after extraction (C). The recovery showed that 20 mM NHOAc in 50 % v/v methanol: water +0.1 % ascorbic acid (pH 6.1) was the ideal buffer for folate extraction. The percentage recoveries were calculated by a ratio of B/C multiplied by 100 (Matuszewski et al., 2003). The recovery results shown in Figure 2.4 confirms that the buffer containing 0.1% ascorbic acid in 50 %

methanol was best for extraction of the folates. The extraction method used here was the freeze thaw approach which has not been reported for microbial folate extraction (Lu et al., 2007) or plants (Garratt et al., 2005).

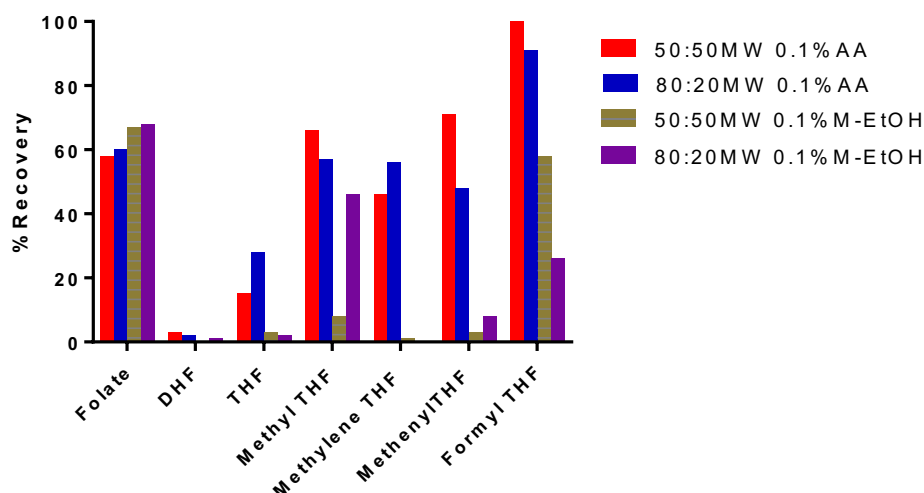


Figure 2.4 Recovery analysis of folates in different buffers. (MW- methanol/water, M-EtOH- 2-mercaptoethanol, AA- ascorbic acid, DHF- dihydrofolate, THF- tetrahydrofolate) (n=2).

The presence of Poly-glutamated folates were investigated in *C. autoethanogeunum* using the HILIC method and precursor and product ions monitored in positive ionisation mode. Since there no standards to optimise the MS/MS parameters, these were taken from an already published method (Lu et al., 2007) as shown in Appendix 1. Analysis of up to eight glutamate residues were investigate in this organism. As the number of glutamic acid residues increased there was a corresponding increase in the retention times as reported by (Garratt et al., 2005). Figure 2.5 showed that in all seven different folates, the polyglutamate folate with three glutamic acid residues were the most abundant which was also reported in *E. coli* for the 5-Methyltetrahydrofolate (Lu et al., 2007).

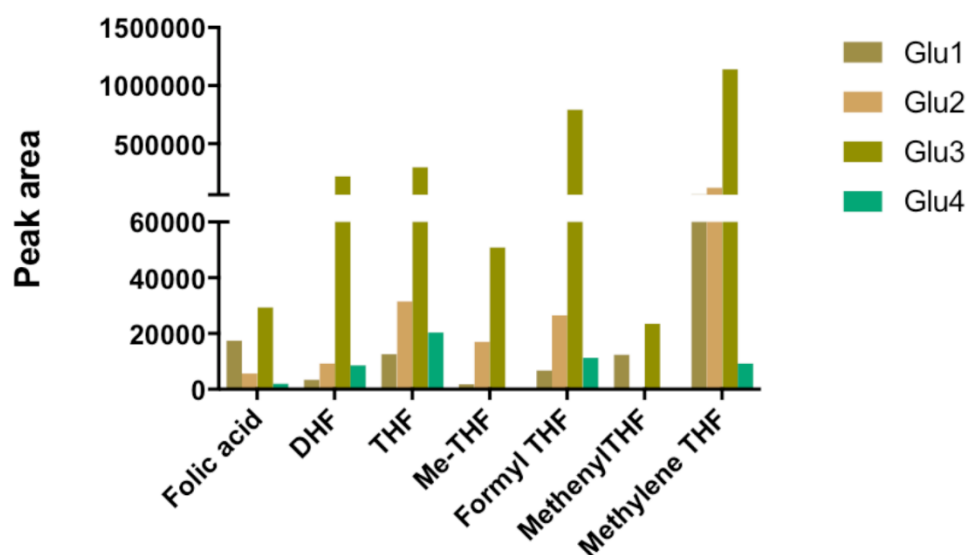


Figure 2.5 Analysis of polyglutamated folates present in *C. autoethanogenum*. In all cases the residues with three polyglutamate was the most abundant form found in *C. autoethanogenum*. (Glu-glutamic acid residues)

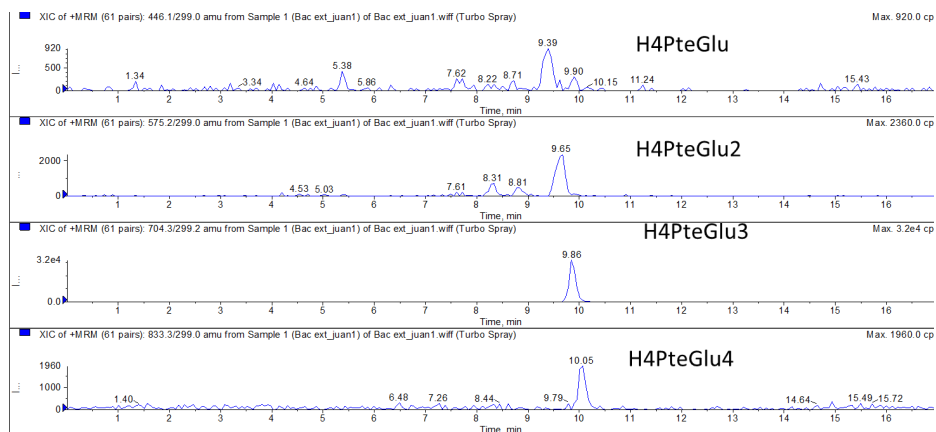


Figure 2.6 A representative chromatogram of tetrahydrofolate (H4PteGlu) and the poly-glutamated forms up to four glutamic acid residues found in *C. autoethanogenum*.

2.3.2.1 IPRP and HILIC methods validation for folates

The LOD, LOQ, linearity, precision and accuracy were evaluated for folates using both IPRP chromatography (Table 2.10) and HILIC (Table 2.11). Detection of DHF, methylene THF and methenyl THF were not successful on the RPIP chromatography method. LODs for RPIP method ranged from 0.01 to 5 nM and that of HILIC method ranged from 0.01 to 2 nM. Evaluation of repeatability of RPIP method after 6 injections of standard mixture at three concentration levels gave RSDs better than 19 % at 2 nM, 16 % at 10 nM and 16% at 50 nM. Similar evaluation of the HILIC method gave RSDs better than 13 % at 2 nM, 22 % at 10 nM and 11% at 50 nM. LODs of folates were compared with LC-MS/MS method reported by (Biagi et al., 2012) and our LODs were more than ten times lower. LODs for DHF acid and 5-methyl THF acid obtained were 1000x and 5x respectively lower than LODs reported in (Bajad et al., 2006).

Table 2.10 Initial reversed phase ion pair method validation of folates. Precision and accuracy are reported for three concentration points, a low (L) of 2 nM, medium (M) of 10 nM and a higher oncentration (H) of 50 nM.

Precision (%)							Accuracy (%)				
Compound		Linear range (nM)	R ²	LOD (nM)	LOQ (nM)	L=2nM	M=10nM	H=50nM	L=2nM	M=10nM	H=50nM
5-Methyl acid	THF	0.01-150	0.9850	0.1	0.5	18.9	15.1	15.8	106.3	79.9	96.9
10-Formyl acid	THF	1-150	0.9871	0.1	0.5	8.6	12.6	3.6	102.2	94.9	109.4
THF acid		1-150	0.9922	5	10	9.8	13.7	13.5	113.0	93.8	100.5
Folic acid		0.001-150	0.9926	0.01	0.05	15.0	9.71	6.4	106.8	120.9	107.5

Table 2.11 Initial HILIC method validation of folates. Precision and accuracy are reported for three concentration points, a low (L) of 2 nM, medium (M) of 10 nM and a higher oncentration (H) of 50 nM.

Precision (%)											Accuracy (%)
Compound		Linear range (nM)	R ²	LOD (nM)	LOQ (nM)	L=2nM	M=10nM	H=50nM	L=2nM	M=10nM	H=50nM
5,10-Methenyl THF acid		0.01-150	0.9904	<0.01	0.01	12.4	2.8	10.1	103.2	107.3	93.3
5-Methyl acid	THF	1-150	0.9964	2	10	1.7	2.5	1.5	102.5	100.7	94.3
5,10-Methylene THF acid		0.01-150	0.9994	0.01	0.05	6.7	1.1	0.9	88.8	101.1	101.4
10-Formyl acid	THF	0.01-150	0.9983	0.01	0.05	11.1	3.2	1.6	103.4	105.9	91
DHF acid		0.01-150	0.9975	0.05	0.1	8.5	5.0	0.9	134.7	104.7	109.3
THF acid		1-150	0.9988	0.05	0.5	3.7	19.4	2.9	131.9	99.9	100.0
Folic acid		0.1-150	0.9975	0.01	0.05	10.2	22	7.4	122	101.2	96.02

2.4 Conclusion

Careful and systematic method optimisation is necessary when multiple metabolites are analysed in complex matrices. Two methods, ion pair chromatography and HILIC were evaluated and validated. Method validation is important to establish that the method is suitable for its intended purpose and that it meets internationally accepted criteria. The HILIC method was better than the ion pair method when validation parameters were compared. Polar metabolites represent an analytical challenge and the results here show that HILIC can provide a better solution than either reversed phase or ion-pair chromatography. The HILIC method also provided a good detection and validation than the ion pair method. There was more inter-conversion of folates when ion pair reagent was used due to the unstable nature of folates. Although the ion pair method developed would not be suitable for this purpose, it is a novel method with reasonable run time compared to other methods.

In the chapters that follow the HILIC method will be optimised and applied to measure intracellular metabolic changes that occur in *C. autoethanogenum* during external pH optimisation and gas (carbon monoxide) dependent experiments.

CHAPTER THREE

Method development 2: Approaches for absolute quantification of intracellular metabolites present in *Clostridium autoethanogenum* metabolic pathways

3. Method development 2: Approaches for absolute quantification of intracellular metabolites present in *Clostridium autoethanogenum* metabolic pathways

3.1 Introduction

3.1.1 Limitations of LC-MS for quantitative analysis of complex biological matrices

Biological matrices such as bacterial cells typically contain thousands of different molecules. Intracellular small molecular weight metabolites have to be measured against a background of proteins, carbohydrates, membrane components and other subcellular material.

Mass spectrometry detection presents unparalleled selectivity and sensitivity and is therefore a superior candidate for LC based quantitative analysis (Stokvis et al., 2005). Despite the improved sensitivity and specificity of LC-MS, co-elution of metabolites, impurities, matrix components and degradation products cause nonlinear response from the detector and ion suppression / enhancement, which raise concerns in qualitative and quantitative LC-MS detections (Chambers et al., 2007, Gosetti et al., 2010, Tang and Kebarle, 1993, Van Eeckhaut et al., 2009) especially in electrospray ionisation (ESI) (Berg and Strand, 2011). Atmospheric pressure chemical ionisation (APCI) is less prone to ion suppression compared to the commonly used positive mode Electrospray ionisation (ESI⁺) (Dams et al., 2003, Matuszewski et al., 2003, Souverain et al., 2004). Matrix effects can therefore lead to compromise in precision and accuracy in quantifying metabolites by LC-MS. There is therefore the need to correct for the interferences from biological matrices, variation in instrument response during analysis and degradation of metabolites during sample preparation. There are some recommendations on reducing ion suppression. These include using negative ionisation mode, APCI instead of ESI, diluting samples, reducing injection volume of samples, splitting the flow after column separation and before Mass Spectrometer analysis and including stable isotope labelled internal standards (SIL ISs) (Chambers et al., 2007, Couchman and Morgan, 2011, Dams et al., 2003, Gangl et al., 2001, Heller, 2007, Jemal et al., 2010, Jessome and

Volmer, 2006, Matuszewski, 2006, Maurer, 2005, Van Hoof et al., 2005). Suggestions from most of these journals states that, a combination of two or more of the recommendations listed above, could help minimise effects from ion suppression. APCI is usually used for non-polar compounds and has limited mass range compared to ESI. Among the recommendations the most suitable approach will be the generation of SIL ISs to overcome the limitations associated with MS analysis.

3.1.2 Generation of multiple stable isotope IS from cells grown on ^{13}C source

Due to the high cost of most commercially available ^{13}C labelled internal standards, some laboratories have developed methods to extract ^{13}C labelled metabolites from *E. coli* cells cultured on a ^{13}C labelled carbon source (Berg and Strand, 2011, Cui et al., 2013, Kim et al., 2015, Lu et al., 2007). *Saccharomyces cerevisiae* is also widely used to generate internal standard for quantitative analysis (Magdenoska et al., 2015) by growing on isotope labelled carbon source. The SILISs generated from these microorganisms can have many applications in parasites (Kim et al., 2015) and unrelated microorganism (Cui et al., 2013). A problem with this approach is that the internal standard metabolites may not be representative of those present in other organisms or tissues. In an ideal world, generation of SILISs should be from the same organism of interest but this is not always the case due to complexity of metabolism resulting in requiring multiple carbon sources. The unavailability or cost of suitable isotopically labelled substrate for the organism is usually the case. Organisms such as *E. coli* and *S. cerevisiae* grow readily on glucose and glucose is a cheaper isotope labelled substrate hence, being widely used for generation of SILISs. There are usually some metabolites in common between cells and organisms so the use of *E. coli* and *S. cerevisiae* to generate SILISs give some degree of coverage of metabolites but not always to a great extent.

A commercially available U- ^{13}C labelled extract of *Arthrospira platensis* grown on a ^{13}C source provides multiple stable isotope internal standards to enable valid measurements and requiring no microbiology laboratory to culture cells

(Schatschneider et al., 2017). Since most metabolites are similar among different microorganisms, metabolites within the *A. platensis* metabolic pathway should potentially be useful as multiple internal standards for IDMS metabolite analysis from the *C. autoethanogenum* metabolic pathway.

3.1.3 Selection of organism to generate uniformly ^{13}C -labelled internal standards

To generate U- ^{13}C -labelled internal standards for absolute quantification of intracellular metabolites in *C. autoethanogenum*, a commercially available U- ^{13}C -*Arthrospira platensis* extract was previously used in our laboratory, which gave coverage of 55 metabolites but as our list of target metabolites increased it was necessary to look at alternative approaches. The use of commercial U- ^{13}C -*Arthrospira platensis* extract was convenient since it was readily available and did not require any microbiology facilities to culture cells. In the quest to increase coverage of the U- ^{13}C -labelled internal standard, metabolites from *C. autoethanogenum* were compared to *E. coli* and *C. pasteurianum* cell extracts. There are reported instances where *E. coli* was used to generate U- ^{13}C -labelled internal standard (Cui et al., 2013, Dong-Hyun Kim, 2015, Lu et al., 2007) with one application in *C. thermocellum* (Cui et al., 2013). With this in mind, metabolites from *C. autoethanogenum* were compared to metabolites in *E. coli* and 424 metabolites were found to be common between them compared 272 metabolites found in *Arthrospira*. A comparison made between metabolites found in *C. autoethanogenum* and *C. pasteurianum* showed 873 metabolites common between them (Figure 3.6). Strict guidelines were used in Progenesis QI using *q* and *p* values to only assign peaks with similar detection and intensity and reject peaks below noise levels. The number of metabolites similar to *C. autoethanogenum* metabolites showed that *C. pasteurianum* was the best organism for generation of U- ^{13}C -labelled internal standard.

The new method is best suited for the purpose of monitoring these metabolites in *C. autoethanogenum*

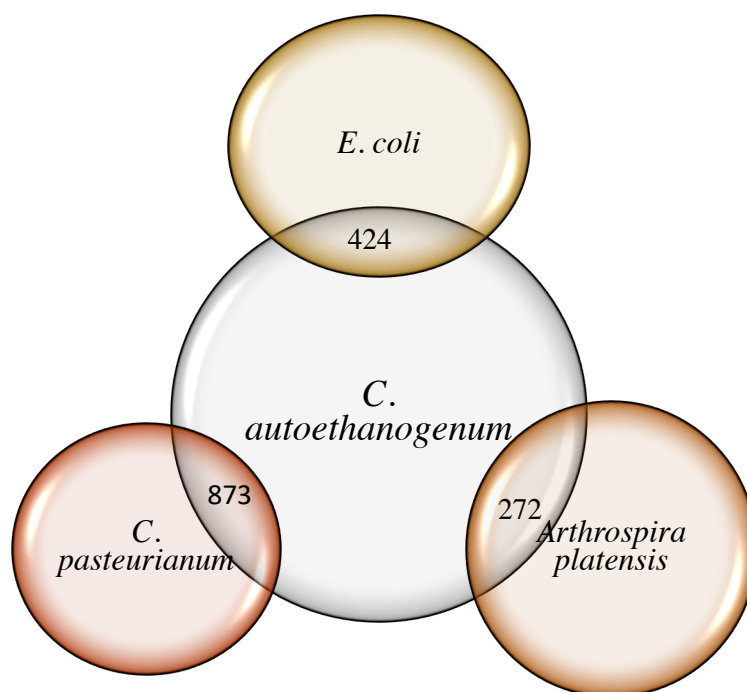


Figure 3.1 comparison of intracellular metabolites of C. autoethanogenum with E. coli. A. platensis and C. pasteurianum to select best organism for generation of U-¹³C-labelled internal standard. 424 metabolites were common between E.coli and C. autoethanogenum, 272 metabolites common between A. platensis and C. autoethanogenum and 873 metabolites common between C. pasteurianum and C. autoethanogenum. Identification of metabolites were carried out using Progenesis QI and using q-values to show significance.

3.1.4 Determining cell numbers and cell volume using flow cytometry

Cell volume and cell numbers are important for reporting intracellular metabolite concentrations. Conventional methods such as microscopy can introduce some bias in the measure due to extensive sample preparation techniques. Measurements of OD specific cell concentration using flow cytometry (FCM) and cell volume by fluorescence microscopy for *Escherichia coli* has been reported (Volkmer and Heinemann, 2011). Flow cytometry has, over the years, become a practical and accurate alternative to microscopy for the measurements of microbial cell counts. This is mainly because thousands of cells can be analysed on the timescale of a few seconds (Bouvier et al., 2001). Applications

of FCM for measurements of cell volume, cell size, surface area and cell diameter are well reported (Akerlund et al., 1995, Bouvier et al., 2001, Schwartz et al., 1983) buttressing the benefits of FCM measurements. FCM uses cell sorters to measure scattered lights at different angles (Delcher et al.) into an illuminating laser beam. Forward scatter (FSC) is usually used for cell volume determination, since both internal and surface irregularities and roughness of membranes could contribute to side scatter (SSC) signals (Heldal et al., 1994, Tzur et al., 2011).

3.1.5 Evaluation of cell volume from microscopy images

Microscopy images of bacteria can provide useful information such as cell volumes, cell numbers, width and length of the bacteria (Bloem et al., 1995). Techniques such as confocal and epifluorescence microscopy are usually used for microscopy imaging (Bloem et al., 1995, Heldal et al., 1994). Sample preparation of microorganisms for microscopy imaging can be demanding, due to addition of fluorescence dye to bind DNA, proteins, nucleic acids or cell wall (Bejerano-Sagie et al., 2006, Berlatzky et al., 2008, Bloem et al., 1992, Pogliano et al., 1999). The extensive sample preparation step can sometimes introduce some degree of bias in the measurements (Heldal et al., 1985).

The cell volume of bacteria cells can be estimated from microscopy images using the length and width of cells. Cell volume (V) calculations can be obtained from length (l) and width (w) by assuming the cells are cylindrically shaped with two half sphere ends and then the volume can be calculated with the formula (Heldal et al., 1985, Loferer-Krossbacher et al., 1998, Volkmer and Heinemann, 2011) $V = (\pi w^2 (l - w/3))/4$. Bacteria bio-volume can be estimated from the product of mean cell volume and total count. Multiplying the bio-volume with specific carbon content will give the biomass (Fry, 1990).

3.1.6 Aims of this chapter

In this chapter,

- Commercially available U-¹³C-labelled *Arthrospira platensis* and U-¹³C-labelled *C. pasteurianum* are compared as suitable organisms for

generation of U-¹³C-labelled internal standard for absolute quantification of intracellular metabolites in *C. autoethanogenum*.

- To report intracellular concentration for modelling purposes using cell volume and cell numbers of *C. autoethanogenum*, which have not been reported before.

3.2 Materials and methods

3.2.1 Chemicals and reagent

HPLC grade ammonium carbonate was purchased from Sigma Aldrich, St. Louis, MO, USA. Uniformly ¹³C-labelled glucose (U-¹³C, 97 %) was purchased from Cambridge Isotopes Laboratory Andover, MA, USA. LC-MS grade acetonitrile was bought from VWR Lutterworth, Leicestershire, UK. Ultra-pure water (18.2 MΩ) was used for this work (ELGA Purelab, High Wycombe). HPLC grade methanol and acetonitrile were obtained from Fisher Scientific (Bishop Meadow Road, Loughborough).

Ultra-pure water (18.2 MΩ) was used for this work (ELGA Purelab, High Wycombe). citric acid monohydrate (95 % - 102 %), oxaloacetic acid, DL-lactic acid, adenine, adenosine, acetyl CoA, adenosine triphosphate (ATP), adenosine diphosphate (ADP), adenosine monophosphate (AMP), nicotinamide adenine dinucleotide, nicotinamide adenine dinucleotide phosphate (NADP), nicotinamide adenine dinucleotide phosphate reduced (NADPH), S-adenosyl-L-methionine, D-sedoheptulose-7-phosphate, 2-isopropylmalic acid, D-gluconic acid sodium salt, L-(-)-3-phenyllactic acid, dihydroxyacetone phosphate, dihydroxyacetic acid, D-(-)-3-phosphoglyceric acid and L-glyceric acid sodium salt were purchased from Sigma Aldrich UK. succinic acid (99 %), glyoxylic acid 50 % in water, 3-methyl-2-oxobutanoic acid sodium salt (98+ %), DL-isocitric acid trisodium salt hydrate (95 %), DL- malic acid (99+ %), fumaric acid (99+ %), 4-hydroxyphenylpyruvic acid, citraconic acid (99+%) and pyruvic acid sodium salt (99+ %) were purchased from Acros Organics supplied by Fisher Scientific (Bishop Meadow Road, Loughborough). UDP-D-glucose and

cytidine-5'-diphosphate were bought from VWR Lutterworth, Leicestershire, UK.

3.2.2 Preparation of Calibration Samples from Stock Solutions

A stock solution of 133 target metabolites (Table 3.1) obtained from the one carbon, fermentation, tricarboxylic acid cycle and amino acid biosynthesis pathways were prepared by dissolving standard stocks in methanol, ethanol or water depending on solubility to a final concentration of 1 mM. Eight calibration mixtures were prepared by serial dilution using water ranging from 0.5 μ M to 200 μ M. 40 μ L of ^{13}C labelled *C. pasteurianum* extract (multiple internal standards) was added to 40 μ L of each calibration standard mixture with relevant dilutions. Both were thoroughly mixed and kept at 4°C in the auto-sampler prior to LC-MS analysis.

3.2.3 Preliminary HILIC method development

LC-MS was performed on Dionex LC system combined with Hybrid Quadrupole Orbitrap mass spectrometer (Q-Exactive, Thermo Fisher Scientific, Hemel Hempstead, UK).

Heated electrospray ionisation (HESI-II) probe was used for all MS analyses. Polarity-switching mode of ionisation and the parameters for the MS analysis were optimised using previously reported settings by Creek (Creek et al., 2011). Resolution was set at 140000, automatic gain control was 1×10^6 , m/z ranging from 70 - 1400, sweep gas 1, auxiliary gas 20, sheath gas 40, capillary temperature 250°C and probe temperature 275 °C. Source voltage +4 kV, tube voltage +70 kV, capillary voltage +50 V and skimmer voltage +20 V was used for positive mode ionisation. Negative mode ionisation was performed at source voltage of -3.5 kV, tube voltage -70 V, capillary voltage -50 V and skimmer voltage -20 V. Both negative and positive polarities were calibrated using at least 5 mass ranges for each polarity prior to analysis.

3.2.4 Extraction protocol optimisation for *C. autoethanogenum* cell lysis

Initial extraction methods investigated were homogenising technique (HMG) and the freeze thaw approach (FT) (Winder et al., 2008). Cell pellets obtained from *C. autoethanogenum* bioreactor were subjected to both methods with spiked standard mixture of ^{12}C metabolites at concentrations of 125 μM and 250 μM .

3.2.4.1 Precellys/Homogeniser method

1mL of 80% aqueous methanol was added to 10 mg of cell pellet containing 0.5g of acid washed glass beads (425-600 μm dia.). This was then spiked with 50 μL of 125 μM and 250 μM standard solution three of each making six samples. The cells were disrupted three times in a Precellys 24 (PEQLAB Biotechnologie, Erlangen, Germany) mill at 6,500 rpm with 1 min cycles and 15 s break. Cell debris was separated by centrifugation at 20,000 g rpm at RT for 20 min 700 μL of cell extracts were transferred into a 1.5 mL tube, evaporated to dryness using a centrifugal evaporator. This was then re-suspended in 200 μL of water, filtered with Ultrafree MC centrifugal device. The extraction method was repeated for another six sets of samples but without the prior spiking of the standards. 700 μL of the extracts was then spiked with 50 μL of 125 μM and 250 μM (3 samples each), evaporated to dryness as previously done and reconstituted in 200 μL of water.

3.2.4.2 Freeze thaw method

10 mg of cell pellets were re-suspended in 500 μL of ACN: MeOH: H_2O (2:2:1) pre-cooled at -20°C . 50 μL of 125 μM and 250 μM standards were spiked prior

to extraction (3 samples each). The freeze thaw extraction was performed by flash freeze in liquid nitrogen for 1 minute and briefly thaws on ice, then vortexed for 30 seconds (4X). Centrifugation at a maximum speed (13,000 rpm) for 5 min at -4 °C and the supernatant transferred into fresh cooled tube. The sample was evaporated to dryness and reconstituted in 200 µL of water. The freeze thaw extraction was repeated as before but without prior spiking of the standard mixture. 700 µL of the extracts was spiked with 50 µL of 125 µM and 250 µM standards after extraction, evaporated to dryness and reconstituted in 200 µL of water.

3.2.5 Exploring *E. coli* as a candidate organism for generation of U-¹³C-labelled internal standard

E. coli MG1655 was cultured in M9 media with similar salts and glucose concentration as previously reported (Kim et al., 2015). Intracellular metabolites were extracted (cold methanol) using existing procedures (Rabinowitz and Kimball 2007; Winder et al. 2008). The cells were harvested at optical density at 600 nm (OD₆₀₀) of 1.0 and centrifuged at 7000g for 5 min. Then, cells were extracted by freeze–thaw cycles (i.e., freeze in liquid N₂ for 30 s and thawed at on ice, then vortexed for 30 s) and centrifuged at 7000g for 10 min. The supernatant was collected and stored at -80 °C prior to analysis.

3.2.6 *C. pasteurianum* culture conditions to generate U- ¹³C- labelled internal standards

Clostridium pasteurianum (DSM 525) was used to generate uniformly ¹³C-labelled multiple internal standards. The organism was purchased from Leibniz-Institut DSMZ- Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH freeze dried. The organism was activated on a CBM agar prepared from MgSO₄.7H₂O, 300 mg; MnSO₄.H₂O, 11.37 mL of a 1mg/mL stock; FeSO₄.7H₂O, 15 mL of a 1mg/mL stock, p-aminobenzoic acid, 1.5 mL of a 1mg/ml stock; Biotin, 30 µL of a 0.1mg/mL stock; thiamine.HCl, 1.5 mL of a 1mg/mL stock; casein hydrolysate, 6 g (enzymatic); glucose, 150 mL of a 50 % stock; K₂HPO₄, 15 mL of a 50 mg/mL stock; and KH₂PO₄, 15 ml of a 50 mg/ml

stock made up to a litre. Frozen (-80 °C) was streaked out on a pre-warmed (37 °C) CBM agar plate and incubated in anaerobic cabinet at 37 °C for 24 h. The cells were inoculated from plate to Biebl media (Biebl, 2001) containing KH₂PO₄, 0.5 g ; K₂HPO₄, 0.5 g ; (NH₄)₂SO₄, 5 g ; MgSO₄ x 7H₂O, 0.2 g ; CaCl₂ x 2H₂O, 0.02 g ; FeSO₄ x 7H₂O, 9.14 mg. The media was made up to 897.75 mL, pH adjusted to 6.2 using HCl and autoclaved. This media was then made up to 1 L by adding 2 mL of trace elements solution, biotin, 25 µg, 20 mL CaCO₃ (250 g/l stock) and 80 mL of 50 % uniformly ¹³C-glucose (final concentration of 4 %). Trace element solution consists of MnCl₂ x 4 H₂O, 100 mg; ZnCl₂, 70 mg; H₃BO₃, 60 mg; NaMoO₄ x 2H₂O, 40 mg; CoCl₂ x 6H₂O, 28.7 mg; CuCl₂ x 2H₂O, 20 mg; NiCl₂ x 6H₂O, 20 mg; HCl (25 %) , 1 ml; per 1 L solution.

3.2.7 Extraction of ¹³C labelled isotope internal standards from *C. pasteurianum*

A precooled 1 mL of methanol (-20 °C) was added to 1 mL cell pellet obtained from *C. pasteurianum* culture and vortexed for 1 min. The re-suspended cells were lysed by freezing in liquid nitrogen, thawing on ice and vortexing for 30 s. The freeze-thaw approach was repeated three times. The mixture was centrifuged (10000 g for 5 min) to remove cell debris and the supernatant was removed and stored at -80 °C to be used as uniformly ¹³C labelled isotope internal standard.

3.2.8 Growth media and bioreactor cultivation of *C. autoethanogenum*

In this experiment *Clostridium autoethanogenum* JA1-1 DSM10061 was used and obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ). *C. autoethanogenum* was grown in continuous gas fed bioreactor with conditions already published (Humphreys et al., 2015).

3.2.9 Growth rate measurements using OD₆₀₀

Samples were collected at different steady states from the Carbon monoxide fed bioreactor and growth rate measured by OD using spectrophotometer at 600 nm (Jenway 7305). The samples were diluted with the media in a 1:10 ratio.

3.2.10 Sample preparation for microscopy

Cells were harvested at an OD₆₀₀ of 5. Equal volumes of Syto9 and propidium iodide (Molecular Probes-Thermofisher,UK) were mixed thoroughly in a tube. 3 µL of the dye mixture was added to 1 mL of bacteria suspension. This dilution contributes 0.3% DMSO to the staining solution because higher concentrations of DMSO could adversely affect the staining. The mixture was mixed thoroughly and incubated in the dark for 15 minutes.

Propidium iodide is red fluorescent stain whiles Syto9 is green fluorescent nucleic acid stain. Syto9 stains all bacteria in the population either damaged or intact membranes. Propidium iodide only stains bacteria cells with damaged membranes reducing the intensity of Syto9 staining when both are combined. When mixed in equal volumes, propidium iodide and Syto9 stains bacteria with damaged cell membrane fluorescent red and those with intact cell membranes fluorescent green. Propidium iodide has an excitation/emission maxima of 490/635 nm while Syto9 has an excitation/emission maxima of 480/500 nm

3.2.11 Microscopy Imaging

Images were captured in laser wide field mode on the Elyra PS1. Super resolution microscope, using C-Apochromat 63x/1.2 W Korr M27 objective with water as immersion media. Samples were mounted with high precision coverslips 1.5H (Paul Marienfeld GmbH, Germany, no. 0107032) and sealed with nail polish. 488 nm and 561 nm lasers were used, in combination with the LBF -488/561 nm filters. Z-stacks in multiple random positions were recorded.

Images were analysed using Volocity software (PerkinElmer, Waltham, MA 02451 USA). A point-spread function (PSF) was calculated based on the parameters used for image acquisition. This PSF was used to deconvolve the images with 30 iterations to minimise the out of focus haze and noise throughout the image volume. Once the images were processed, the cells in the Syto9 channel were automatically detected using a 'auto-threshold' building block,

‘separate touching objects’ building block to split overlapping cells and an ‘exclude objects by size’ building block to remove any objects smaller than $0.6 \mu\text{m}^3$ from the analysis. Next the ‘volume, intensity, skeletal length diameter and longest axis measurements were generated from the image overlay. Following this, a visual inspection was carried out and thresholds were further modified to get a similar range of false positives and false negatives giving special attention to the 3 largest, 3 median and 3 smallest objects detected. In half of the samples Tetraspeck microspheres (Molecular Probes by Life Technologies, USA, cat. no: T7282) $1 \mu\text{m}$ fluorescent beads were used to calibrate the measurements. In total 4363 objects were measured.

3.2.12 Flow Cytometry (FCM) Measurements for Cell Counting

Samples were analysed on an Astrios EQ flow cytometer (Beckman Coulter) with dual 488nm laser forward scatter (488nm laser) PMTs (PMT1 unmasked, PMT mask M2). Biotic bacteria were identified by positive Syto-9 fluorescence (488nm laser, 513/26nm band pass filter) and dead cells identified by positive staining for propidium iodide (561nm laser, 614/20nm band pass filter). Counting beads were identified by positive red fluorescent staining (642nm laser, 671/40nm band pass filter). Data was analysed using Kaluza (V1.5, Beckman Coulter). For each experiment 488nm side scatter area signals of bacteria were normalised to size standard references (Sphero Nano Fluorescent Particle Standard kit, Spherotech Inc) to eliminate any small machine performance variation. To determine the relationship between normalised 488nm side scatter area measurements and bacterial volume matched samples from four reactors were analysed by flow cytometry and deconvolution microscopy. Flow cytometry measurements from 1 million biotic bacteria were stratified into median 10 percentiles and correlated to median 10 percentiles of microscopy volume measurements. Curve fit analysis was performed in Prism (V7.01 GraphPad). To determine cell by cell volume in subsequent experiments volume was interpolated from normalised side scatter area signals of live

bacteria (Syto-9/PI-) using the previous correlation. Using 488nm Side scatter height vs width analysis cells were also classified as singlets or elongated/doublets (greater than 2X SSC-width measurement). For each sample the number of counting beads (Flow Set 675, Beckman Coulter) analysed was converted to volume analysed and viable Syto9+ counts were classified as PI+ (dead) or PI- (live).

3.2.13 LC-MS data analysis for identification of 133 metabolites

All LC-MS data were processed using Xcalibur and Tracefinder (Thermo Fisher Scientific, Hemel Hempstead, UK). Metabolites identification using authentic standards and U-¹³C-labelled internal standard from microorganisms were carried out using accurate mass to 5 decimal places (error<5ppm) and further confirmation using retention time. The peak area ratios obtained from peak integration were transferred into GraphPad PRISM 7.0 for calibration lines to be plotted.

3.2.14 Method validation

Calibration lines were plotted using concentration of analytes on X-axis and peak area ratios of analyte to U-¹³C-labelled *C. pasteurianum* cell extract (¹²C/¹³C) on the Y-axis. The limit of detection (LOD) and limit of quantification (LOQ) were obtained by using their signal to noise (S/N) ratios of 3:1 and 10:1 respectively. Accuracy and precision were determined from 6 replicate injections of each calibration point. Precision was shown as the relative standard deviation (%RSD) while accuracy of the method was determined as [(mean measured concentration) / (theoretical concentration)] X 100.

3.3 Results and Discussion

3.3.1 Optimisation of extraction protocol for cell lysis of *C.*

autoethanogenum

Initial optimisations of extraction protocols for the cell lysis of *C. autoethanogenum* were done by comparing two methods, homogenising (HMG) (Shin et al., 2010) and freeze thaw (FT) (Winder and Dunn, 2011). FT cycles are well reported for intracellular metabolite extraction of microorganisms such as *E. coli* (Winder et al., 2008). Representative compounds were chosen and spiked into cell pellet of *C. autoethanogenum* prior and after extraction to evaluate recovery. Recovery analysis was conducted in accordance with previous protocol (Matuszewski et al., 2003). The samples were spiked at two different concentrations 125 μ M and 250 μ M before and after extraction of the cell pellets. FT extraction protocol proved superior to the HMG method at both spiked concentrations Figure 3.2.

A further optimisation of the extraction protocol was to investigate combination of solvents that will give us the optimum results using the freeze thaw approach. The plots in Figure 3.3 show the recovery of each analyte was comparable and the solvent compositions did not affect the recovery. It is well reported that cold methanol used for quenching causes cell leakage in both gram positive and negative bacteria (Bolten et al., 2007, Shin et al., 2010). Hence, the extraction of *C. autoethanogenum* cell pellet was carried out with cold methanol to lyse the cells followed by the addition of water and chloroform. Figure 3.1 shows the microscopy images of *C. autoethanogenum* cells before and after extraction. Syto9 and propidium iodide fluorescence stains were used to indicate live/dead cells. Before extraction green fluorescence was observed with cell wall of the cells intact but after extraction cell debris were observed and no intact cells were present.

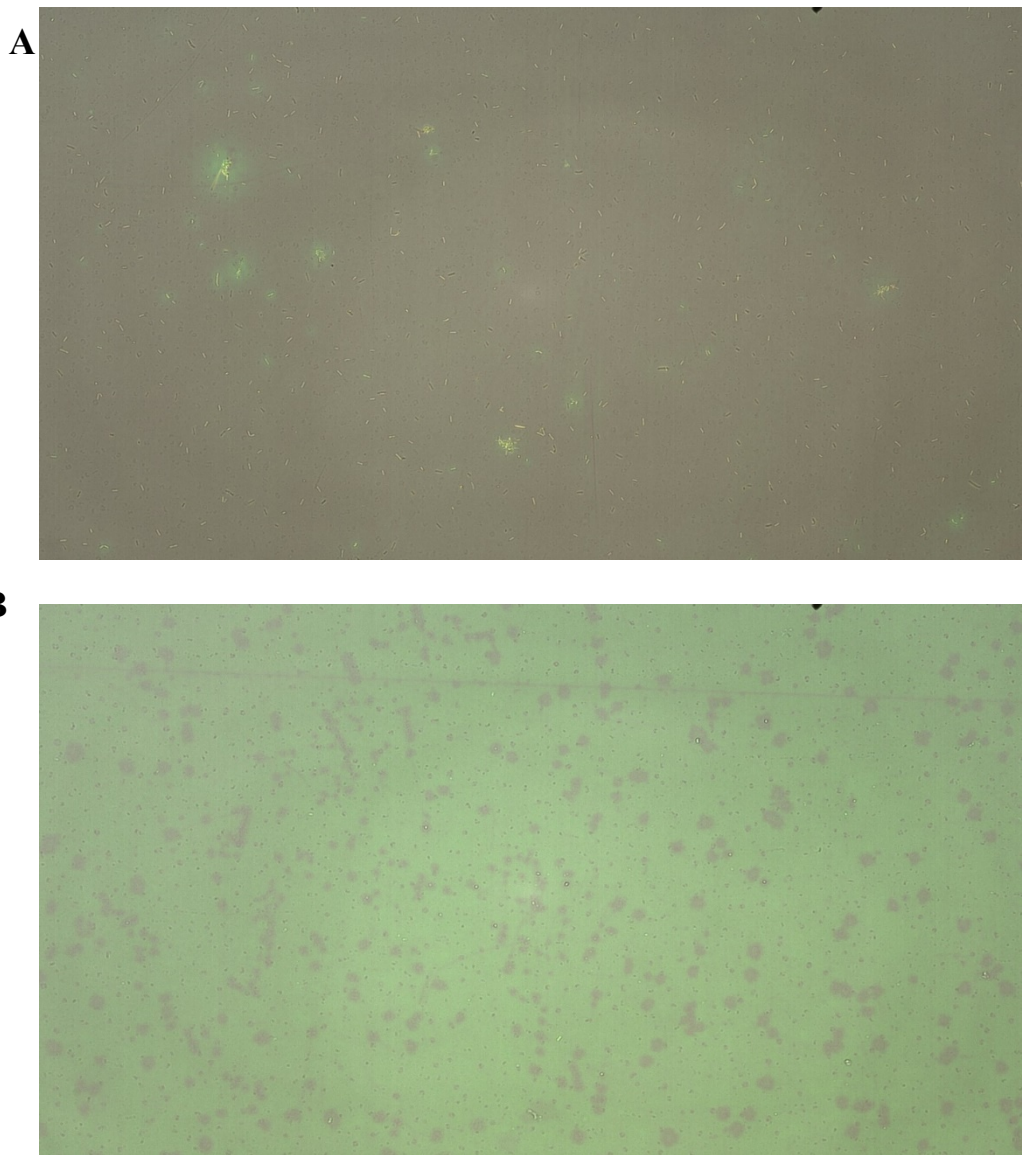


Figure 3.2 Microscopy images of C. autoethanogenum stained with syto9 and propidium iodide before (A) and after (B) freeze thaw cycle to confirm cell lysis. Live cells fluoresce green to indicate intact cells as shown in A was used as control. Cells were stained after extraction to confirm cell lysis. Dead cells fluoresce red due to compromised cell walls as shown in B indication complete cell lysis during freeze thaw extraction.

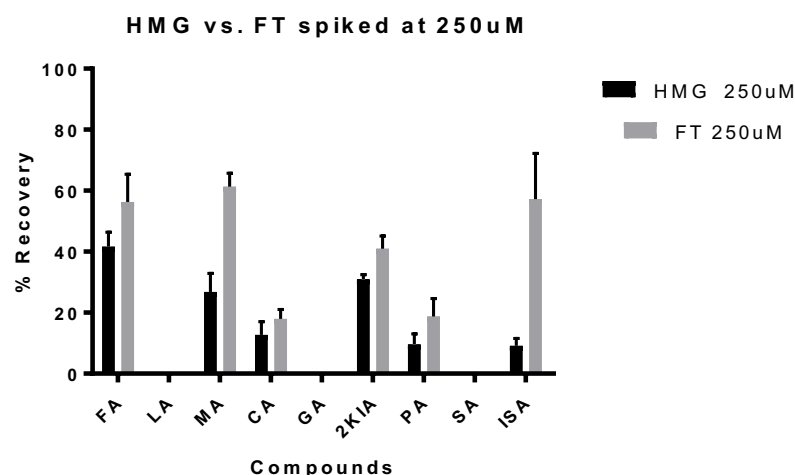
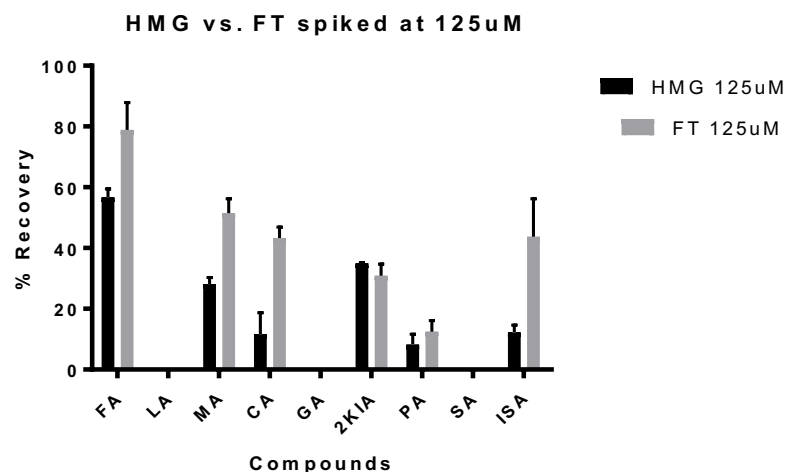


Figure 3.3 Percentage recovery of analytes using the homogenizer and freeze thaw extraction method analysed on the HILIC method. Recovery of metabolites were evaluated at two different concentration points, 125 μ M and 250 μ M. The recovery was calculated by subtracting the peak area of the spiked prior to extraction from spiked post extraction then divide it by the peak area of the standard mixture multiplied by 100. Error bars are from the standard deviation (SD) [FA- Fumaric acid, LA- Lactic acid, MA- Malic acid, CA- Citric acid, GA- Glyoxylic acid, 2KIA- 2-Ketoisovaleric acid, PA- Pyruvic acid, SA- Succinic acid, ISA- Isocitric acid]

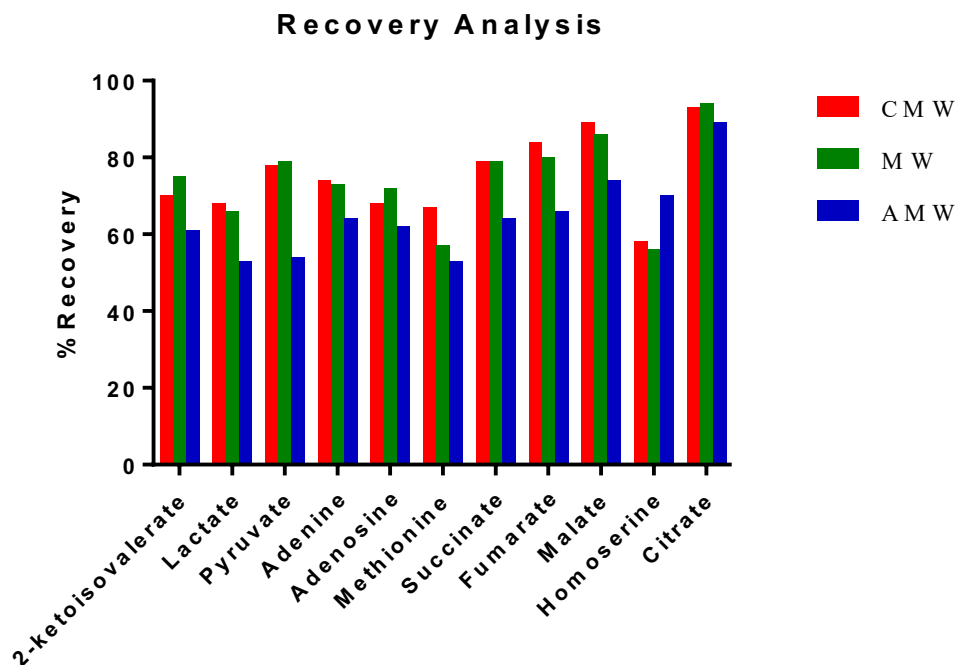


Figure 3.4 Recovery analysis of representative metabolites in different solvent combinations for optimisation of cell lysis of *C. autoethanogenum* (CMW-chloroform:methanol:water(2:1:1), MW-methanol:water(1:1), AMW-acetonitrile :methanol :water (2:2:1)). (n=2). The recovery was calculated by subtracting the peak area of the spiked prior to extraction from spiked post extraction then divide it by the peak area of the standard mixture multiplied by 100

3.3.2 Stability test for extraction solvents

Stability of the selected metabolites were evaluated in the extraction solvents and after 36 hours (Figure 3.4) the recovery was the same showing stability of the metabolites in the solvent combinations. The best option will be chloroform: methanol: water (CMW) because it has the added advantage to retain lipids in the chloroform layer giving less interference and cleaner chromatography. The optimised extraction protocol was applied to cell pellet of *C. autoethanogenum* (refer to Section 3.2.4.2).

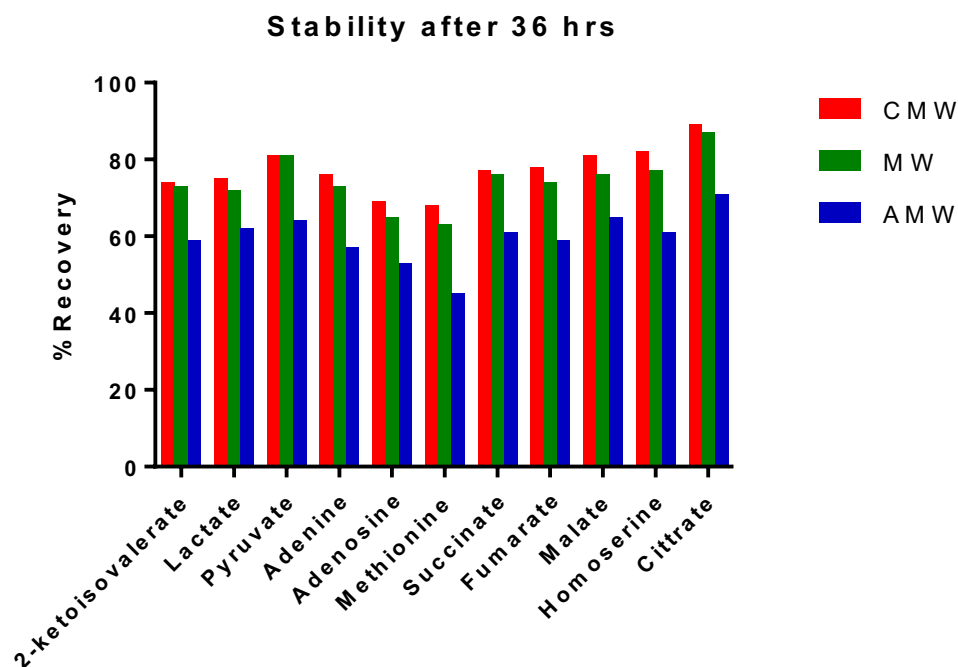


Figure 3.5 Stability analysis of representative metabolites in different solvent combinations for optimisation of cell lysis of *C. autoethanogenum* (CMW-chloroform:methanol:water(2:1:1), MW-methanol:water(1:1), AMW-acetonitrile :methanol :water (2:2:1)). (n=2).

3.3.3 Evaluation of Matrix effect of key metabolites after extraction

Representative metabolites were selected, and ion suppression/enhancement was evaluated as shown in Figure 3.5. Values of matrix effect >100% indicates ionisation enhancement while values <100% indicates ionisation suppression (Matuszewski et al., 2003). Equally, the FT method experiences less ion suppression compared to metabolites extracted using HMG approach. Overall assessment of both methods showed the FT extraction method was a better method to use in future extraction of *C. autoethanogenum* cell pellets.

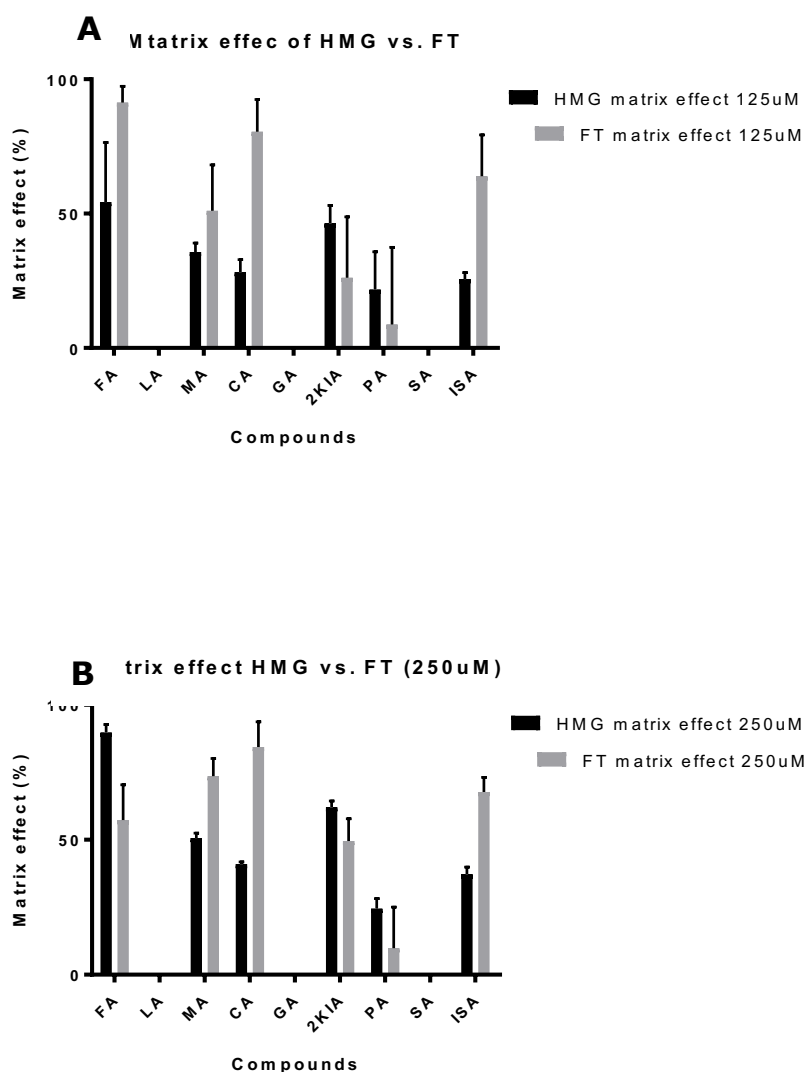


Figure 3.6 Matrix effect of analytes using the homogenizer and freeze thaw extraction method analysed on the HILIC method. (A) represents matrix effect of compounds spiked at lower concentration of 125 μ M (B) represents matrix effect of compounds spiked at higher concentration of 250 μ M. Error bars are from standard deviation (SD). [FA- Fumaric acid, LA- Lactic acid, MA- Malic acid, CA- Citric acid, GA- Glyoxylic acid, 2KIA- 2-Ketoisovaleric acid, PA- Pyruvic acid, SA- Succinic acid, ISA- Isocitric acid]

Table 3.1 Full list of 133 key metabolites monitored with retention times (RT), ionisation modes, LODs, LOQs and both ^{12}C and ^{13}C masses

Compounds	RT (min)	Mode	M+H (12C)	M+H (U-13C)	M-H (12C)	M-H (U-13C)	LOD (uM)	LOQ (uM)
2-Hydroxyglutaric acid	8.58	M-H	149.0444	154.0612	147.0299	152.0467	1	5
2-Isopropylmalic acid	7.64	M-H	177.0757	184.0992	175.0612	182.0847	<0.01	0.01
2-Ketoisovaleric acid	5.02	M-H	117.0546	122.0714	115.0401	120.0568	0.1	0.25
3-Hydroxybutyryl-CoA	7.47	M-H	854.1593	879.2431	852.1447	877.2286	0.05	0.1
3-Phosphoglycerate	9.05	M-H	187.0002	190.0103	184.9857	187.9957	0.05	0.1
Acetoacetic acid	6.93	M-H	103.0390	107.0524	101.0244	105.0378	1	5
Acetoacetyl CoA	7.39	M+H	852.1436	877.2275	850.1290	875.2129	0.25	0.5
Acetyl phosphate	8.29	M-H	140.9947	143.0014	138.9802	140.9869	0.5	1
Acetyl-CoA	7.24	M-H	810.1330	833.2102	808.1185	831.1957	0.05	0.1
Acetyllysine	8.16	M+H	189.1234	197.1502	187.1088	195.1357	0.01	0.05
Adenine	7.86	M+H	136.0618	141.0785	134.0472	139.0640	<0.001	0.001
Adenosine	7.5	M+H	268.1040	278.1376	266.0895	276.1230	0.1	0.25
Adenosine diphosphate (ADP)	8.44	M-H	428.0367	438.0702	426.0221	436.0557	0.01	0.05
Adenosine diphosphate glucose (ADP-Glucose)	8.33	M-H	590.0895	606.1432	588.0750	604.1286	0.01	0.05
Adenosine monophosphate (AMP)	7.83	M-H	348.0704	358.1039	346.0558	356.0894	0.01	0.05
Adenosine triphosphate (ATP)	8.89	M-H	508.0030	518.0366	505.9885	516.0220	0.1	0.5
Butyryl-CoA	7.02	M+H	838.1644	863.2483	836.1498	861.2337	0.01	0.05
Cis aconitic acid	9.33	M-H	175.0237	181.0438	173.0092	179.0293	0.01	0.05
Citraconic acid	6.78	M-H	131.0339	136.0507	129.0193	134.0361	0.01	0.05
Citric acid	9.44	M-H	193.0343	199.0544	191.0197	197.0399	0.05	0.1
Coenzyme A (CoA)	7.58	M+H	768.1225	789.1929	766.1079	787.1784	0.5	1

Compounds	RT (min)	Mode	M+H (¹² C)	M+H (U- ¹³ C)	M-H (¹² C)	M-H (U- ¹³ C)	LOD (uM)	LOQ (uM)
Cystathionine	9.01	M+H	223.0747	230.0982	221.0602	228.0836	0.05	0.1
Cysteine	9.03	M+H	122.0270	125.0371	120.0125	123.0225	1	5
Cytidine	8.39	M+H	244.0928	253.1230	242.0782	251.1084	0.001	0.01
Cytidine monophosphate (CMP)	8.55	M-H	324.0591	333.0893	322.0446	331.0748	0.01	0.05
Cytidine triphosphate (CTP)	9.34	M-H	483.9918	493.0220	481.9772	491.0074	0.25	0.5
Cytidine-diphosphate (CDP)	8.97	M-H	404.0255	413.0557	402.0109	411.0411	0.1	0.25
Cytosine	8.32	M-H	112.0505	116.0640	110.0360	114.0494	0.25	0.5
Deoxyadenosine	7.18	M+H	252.1091	262.1427	250.0946	260.1281	0.001	0.01
Deoxyadenosine diphosphate (dADP)	8.17	M-H	412.0418	422.0753	410.0272	420.0608	0.01	0.05
Deoxyadenosine triphosphate (dATP)	8.64	M-H	492.0081	502.0417	489.9936	500.0271	0.05	0.1
Deoxycytidine diphosphate (dCDP)	8.53	M-H	388.0305	397.0607	386.0160	395.0462	0.001	0.01
Deoxycytidine monophosphate (dCMP)	8.23	M-H	308.0642	317.0944	306.0497	315.0799	0.01	0.05
Deoxycytidine triphosphate (dCTP)	9.07	M-H	467.9969	477.0271	465.9823	475.0125	0.1	0.25
Dephospho-CoA	5.59	M+H	688.1562	709.2266	686.1416	707.2121	0.25	0.5
D-Fructose-6-phosphate	8.70	M-H	261.0370	267.0571	259.0224	265.0426	0.01	0.05
D-Gluconic acid	8.17	M-H	197.0656	203.0857	195.0510	201.0712	0.001	0.01
D-glucose 6-phosphate	8.49	M-H	261.0370	267.0571	259.0224	265.0426	1	10
Dihydroorotic acid	7.60	M-H	159.0400	164.0568	157.0255	162.0423	0.001	0.01
Dihydroxyacetone phosphate (DHAP)	8.52	M-H	171.0053	174.0154	168.9907	172.0008	0.01	0.05
Erythrose-4-phosphate	8.74	M-H	201.0159	205.0293	199.0013	203.0147	0.05	0.1
Flavine mononucleotide (FMN)	7.19	M+H	457.1119	474.1689	455.0973	472.1544	0.01	0.05
Flavine adenosine diphosphate (FAD)	7.24	M-H	786.1644	813.2550	784.1499	811.2404	0.01	0.05

Compounds	RT (min)	Mode	M+H (¹² C)	M+H (U- ¹³ C)	M-H (¹² C)	M-H (U- ¹³ C)	LOD (uM)	LOQ (uM)
Fructose 1,6-bisphosphate	9.15	M-H	341.0033	347.0235	338.9888	345.0089	0.1	0.25
Fumaric acid	8.89	M-H	117.0182	121.0317	115.0037	119.0171	0.01	0.05
Glucosamine-6-phosphate	8.59	M-H	260.0530	266.0731	258.0384	264.0586	<0.001	0.001
Glutathione	8.21	M-H	308.0911	318.1246	306.0765	316.1101	<0.50	0.5
Glyceraldehyde-3-phosphate	8.71	M-H	171.0053	174.0154	168.9907	172.0008	5	10
Glyceric acid	7.60	M-H	107.0339	110.0439	105.0193	108.0294	0.001	0.05
Glycine	9.22	M+H	76.0393	78.0460	74.0248	76.0315	0.05	0.1
Glyoxylic acid	8.35	M-H	75.0077	77.0144	72.9931	74.9998	0.5	1
Guanine	8.85	M+H	152.0567	157.0735	150.0421	155.0589	0.05	0.1
Guanosine	8.58	M+H	284.0989	294.1325	282.0844	292.1179	<0.001	0.001
Guanosine diphosphate (GDP)	9.30	M-H	444.0316	454.0652	442.0171	452.0506	0.05	0.1
Guanosine monophosphate (GMP)	8.75	M-H	364.0653	374.0988	362.0507	372.0843	<0.001	0.001
Guanosine triphosphate (GTP)	9.58	M-H	523.9979	534.0315	521.9834	532.0169	0.25	0.5
homocysteine	8.67	M+H	136.0427	140.0561	134.0281	138.0415	1	5
Hydroxyisocaproic acid	6.82	M-H	133.0859	139.1061	131.0714	137.0915	0.01	0.05
Hydroxyphenyllactic acid	7.01	M-H	183.0652	192.0954	181.0506	190.0808	0.001	0.01
Hydroxyphenylpyruvate	6.90	M-H	181.0495	190.0797	179.0350	188.0652	0.01	0.05
Hypoxanthine	7.99	M+H	137.0458	142.0626	135.0312	140.0480	0.001	0.01
Inosine	8.02	M+H	269.0880	279.1216	267.0735	277.1070	0.01	0.1
Inosine monophosphate (IMP)	8.55	M-H	349.0544	359.0879	347.0398	357.0734	<0.05	0.05
Isocitric acid	9.23	M-H	193.0343	199.0544	191.0197	197.0399	0.05	0.1
Ketoglutaric acid (2-Oxoglutaric acid)	8.55	M-H	147.0288	152.0456	145.0142	150.0310	0.001	0.01

Compounds	RT (min)	Mode	M+H (¹² C)	M+H (U- ¹³ C)	M-H (¹² C)	M-H (U- ¹³ C)	LOD (uM)	LOQ (uM)
Lactic acid	6.98	M-H	91.0390	94.0490	89.0244	92.0345	1	5
L-Alanine	8.97	M+H	90.0550	93.0650	88.0404	91.0505	0.1	0.5
L-Arginine	12.32	M+H	175.1190	181.1391	173.1044	179.1245	<0.001	0.001
L-Asparagine	8.99	M+H	133.0608	137.0742	131.0462	135.0596	0.1	0.5
L-Aspartic acid	8.55	M-H	134.0448	138.0582	132.0302	136.0437	0.1	0.25
L-Citrulline	9.10	M+H	176.1030	182.1231	174.0884	180.1085	0.001	0.01
L-Glutamic acid	8.41	M-H	148.0604	153.0772	146.0459	151.0627	0.05	0.1
L-Glutamine	8.88	M+H	147.0764	152.0932	145.0619	150.0786	0.01	0.05
L-Histidine	9.02	M+H	156.0768	162.0969	154.0622	160.0823	0.5	1
L-Homoserine	8.96	M+H	120.0655	124.0789	118.0510	122.0644	0.05	0.1
Lipoic acid	4.62	M-H	207.0508	215.0776	205.0362	213.0631	<0.01	0.01
L-Isoleucine	7.99	M+H	132.1019	138.1220	130.0874	136.1075	0.05	0.1
L-Leucine	7.99	M+H	132.1019	138.1220	130.0874	136.1075	0.05	0.1
L-Lysine	11.57	M+H	147.1128	153.1329	145.0983	151.1184	0.1	1
L-Methionine	7.93	M+H	150.0583	155.0751	148.0438	153.0605	0.001	0.01
L-Ornithine	11.07	M-H	133.0972	138.1139	131.0826	136.0994	0.01	0.05
L-Phenylalanine	7.45	M-H	166.0863	175.1164	164.0717	173.1019	0.001	0.01
L-Proline	8.50	M-H	116.0706	121.0874	114.0561	119.0728	0.05	0.25
L-Serine	9.15	M-H	106.0499	109.0599	104.0353	107.0454	1	5
L-Threonine	8.67	M-H	120.0655	124.0789	118.0510	122.0644	0.001	0.05
L-Tryptophan	8.29	M-H	205.0972	216.1341	203.0826	214.1195	0.01	0.05
L-Tyrosine	8.61	M-H	182.0812	191.1114	180.0666	189.0968	1	5

Compounds	RT (min)	Mode	M+H (¹² C)	M+H (U- ¹³ C)	M-H (¹² C)	M-H (U- ¹³ C)	LOD (uM)	LOQ (uM)
L-Valine	8.33	M+H	118.0863	123.1030	116.0717	121.0885	1	5
Malic acid	8.87	M-H	135.0288	139.0422	133.0142	137.0277	0.001	0.01
Malonyl CoA	8.38	M-H	854.1229	878.2034	852.1083	876.1888	0.05	0.1
Mesaconic acid	9.33	M-H	131.0339	136.0507	129.0193	134.0361	0.01	0.05
N-Acetylalanine	6.94	M-H	132.0655	137.0823	130.0510	135.0677	<0.05	0.05
N-Acetylcysteine	6.83	M-H	164.0376	169.0544	162.0230	167.0398	0.05	0.1
N-Acetylglucosamine	8.13	M-H	222.0972	230.1241	220.0827	228.1095	0.01	0.05
N-Acetyl-glucosamine-6-phosphate	8.29	M-H	302.0635	310.0904	300.0490	308.0758	1	5
N-Acetylglutamic acid	8.13	M-H	190.0710	197.0945	188.0564	195.0799	<0.001	0.001
N-Acetylglutamine	7.09	M-H	189.0870	196.1105	187.0724	194.0959	0.001	0.01
N-Acetyloronithine	8.79	M-H	175.1077	182.1312	173.0932	180.1167	0.001	0.01
Nicotinamide adenine dinucleotide	8.06	M-H	665.1242	686.1947	663.1097	684.1801	0.05	0.1
Nicotinamide adenine dinucleotide phosphate - reduced (NADPH)	8.30	M-H	746.0984	767.1688	744.0838	765.1543	5	10
Nicotinamide adenine dinucleotide phosphate (NADP)	8.59	M-H	745.0906	766.1610	743.0760	764.1465	0.1	0.25
Nicotinamide adenine dinucleotide reduced (NADH)	7.39	M+H	666.1320	687.2025	664.1175	685.1880	5	10
O-Acetyl-L-serine	7.25	M+H	148.0604	153.0772	146.0459	151.0627	0.1	0.25
Orotic acid	7.41	M-H	157.0244	162.0412	155.0098	160.0266	0.001	0.01
Oxalic acid	9.47	M-H	91.0026	93.0093	88.9880	90.9947	0.25	0.5
Oxaloacetic acid	9.47	M-H	133.0131	137.0266	130.9986	135.0120	0.25	0.5
Phenyllactic acid	4.68	M-H	167.0703	176.1005	165.0557	174.0859	0.001	0.01

Compounds	RT (min)	Mode	M+H (¹² C)	M+H (U- ¹³ C)	M-H (¹² C)	M-H (U- ¹³ C)	LOD (uM)	LOQ (uM)
Phenylpyruvic acid	4.41	M-H	165.0546	174.0848	163.0401	172.0703	0.01	0.05
Phosphoenolpyruvate	9.24	M-H	168.9897	171.9997	166.9751	169.9852	0.01	0.05
Pyroglutamic acid	7.27	M-H	130.0499	135.0666	128.0353	133.0521	0.05	0.25
Pyruvic acid	6.85	M-H	89.0233	92.0334	87.0088	90.0188	0.5	1
Ribose-5-phosphate	8.64	M-H	231.0264	236.0432	229.0119	234.0287	0.001	0.05
S-Adenosyl homocysteine	8.31	M+H	385.1289	399.1758	383.1143	397.1613	0.001	0.01
S-Adenosylmethionine (SAM)	9.07	M+H	399.1445	414.1948	397.1300	412.1803	<0.01	0.01
Sedoheptulose-7-phosphate	8.72	M-H	291.0476	298.0710	289.0330	296.0565	0.01	0.05
sn-Glycerol-3-phosphate	8.22	M-H	173.0210	176.0310	171.0064	174.0165	0.01	0.05
Succinic acid	8.61	M-H	119.0339	123.0473	117.0193	121.0328	0.5	1
Taurodeoxycholic acid	4.08	M-H	500.3040	526.3913	498.2895	524.3767	0.01	0.05
Thiamine	7.04	M+H	266.1196	278.1598	264.1050	276.1453	NF	NF
Thiamine diphosphate	8.99	M-H	426.0522	438.0925	424.0377	436.0780	5	10
Thymidine	6.86	M+H	243.0975	253.1311	241.0830	251.1165	0.001	0.01
Thymine	7.15	M+H	127.0502	132.0670	125.0357	130.0524	0.05	0.5
Trans-Aconitic acid	9.35	M+H	175.0237	181.0438	173.0092	179.0293	1	5
Trehalose	8.95	M-H	343.1235	355.1637	341.1089	353.1492	0.1	0.25
Uracil	7.54	M-H	113.0346	117.0480	111.0200	115.0334	0.5	1
Uridine	7.70	M+H	245.0768	254.1070	243.0623	252.0925	<0.001	0.001
Uridine 5'-diphosphate (UDP)	8.66	M-H	405.0095	414.0397	402.9949	412.0251	<0.001	0.001
Uridine 5'-monophosphate (UMP)	8.39	M-H	325.0431	334.0733	323.0286	332.0588	0.001	0.05
Uridine diphosphate glucose (UDP-glucose)	8.54	M-H	567.0623	582.1126	565.0477	580.0981		

Compounds	RT (min)	Mode	M+H (¹²C)	M+H (U- ¹³C)	M-H (¹²C)	M-H (U-¹³C)	LOD (uM)	LOQ (uM)
Uridine diphosphate-N-acetylglucosamine (UDP-N-acetylglucosamine)	8.35	M-H	608.0888	625.1459	606.0743	623.1313	<0.01	0.01
Uridine triphosphate (UTP)	9.20	M-H	484.9758	494.0060	482.9613	491.9914	0.05	0.1
Xanthine	7.86	M-H	153.0407	158.0575	151.0261	156.0429	0.001	0.01
Xanthosine	7.91	M-H	285.0830	295.1165	283.0684	293.1020	0.01	0.05

NF- Not found

3.3.4 HILIC-IDMS method development and validation

LC-MS conditions used in this study gave a satisfactory separation and detection of the metabolites selected from one carbon and associated metabolic pathways. Retention times of compounds ranged from 4.9 to 12.5 minutes and good peak shapes were obtained for majority of the compounds. A few examples are presented in Appendix 4 to Appendix 6 to give an indication of separation with this method. The metabolites were grouped according to their accurate mass and retention times. Each group was then analysed at different concentrations for validation purposes using switching polarity mode to obtain chromatograms in both negative and positive ionisation modes for the metabolites.

Uniformly ^{13}C -labelled internal standards were generated using *C. pasteurianum* grown on U- ^{13}C -labelled glucose. The internal standard was added to calibration standards prior to analysis in a 1:1 v/v ratio. In validating the analytical method LOD, LOQ (Table 3.2), linearity, accuracy and precision were evaluated (Table 3.3). Precision and accuracy for all metabolites were found to be within acceptable limits of 15 % of nominal concentrations (Hartmann et al., 1998). Three concentration points were chosen, low concentration of 5 μM , medium of 25 μM and a high concentration of 100 μM to reflect intracellular concentrations at the different bacteria biomass when cells are collected. At lower concentrations, precision and accuracy tend to fail validation limits due to an increase in the contribution of noise, problems with peak integration as observed by (Cordell et al., 2008) but validation presented in this work shows most metabolites passed validation at lower concentration.

LODs and LOQs ranged from 0.001 -5 μM and these are lower than those reported in similar methods (Bajad et al., 2006, Cui et al., 2013, Huang et al., 2012, Lu et al., 2006, Luo et al., 2007) by unit conversions. Analysis of some tricarboxylic acid (TCA) cycle intermediates such as pyruvate, succinate, malate, α -ketoglutarate, citrate and fumarate using Q-Exactive have been reported (Hu et al., 2015) but no validation of the method was reported. A LC-MS based metabolomics reported by Bajad et al covered a wide range of metabolites using HILIC-MS/MS. Our range of metabolites fit within their list

but also the paper mentions a validated method and there is no indication of this in the paper (Bajad and Shulaev, 2011). LODs reported for carboxylic acids succinate, citrate, pyruvate, Alpha-ketoglutarate malate and fumarate were much lower than those reported in similar HILIC MS/MS method by Huang (Bajad et al., 2006, Huang et al., 2012, Lu et al., 2006) ranging from 2-8 μM after unit conversion, 0.05 -8.5 μM reported by Bajad (Bajad et al., 2006), 0.03 – 0.2 μM reported by Luo (Luo et al., 2007) and 0.1-0.3 μM reported by Cui (Cui et al., 2013). LODs reported for ATP, ADP and AMP by Luo (Luo et al., 2007) were 0.06 μM , 0.03 μM and 0.01 μM respectively, which is higher than reported here. LOD values reported by Lu (Lu et al., 2006) for valine, homoserine, leucine, adenine, methionine and S-adenosyl methionine ranged between 0.004 to 0.08 μM which were higher than reported here. Cordell (Cordell et al., 2008) also reported LODs for nucleotides such as UTP, GMP, CMP, CDP, AMP, using a reversed phase ion pair chromatography. These LODs ranged between 0.25 μM to 0.5 μM . Other nucleotides such as NAD^+ , NADH, NADP and NADPH also had LODs within this range, which were still higher compared to the results obtained with this method.

The LOQ of all target metabolites were determined by using their signal to noise ratio of (10:1). The LOQs were much lower than reported with similar methods or different methods (Cui et al., 2013, Huang et al., 2012) due to improved sensitivity of the instrument and improved chromatographic parameters to reduce background noise, . Overall, a more sensitive method was developed for the range of metabolites analysed compared to other published methods.

Precision and accuracy were evaluated for detected metabolite categories (*Table 3.2*). Precisions of amino acids, nucleosides, nucleotides and sugars were with the 15 % limit, but some organic acids exceeded this limit. Accuracy for all metabolite categories were within the acceptable limits except for few organic acids, nucleotides and sugars. Validation of the method showed that it was suitable for its intended purpose in analysing intracellular metabolites.

Table 3.2 Precision and accuracy of metabolite categories among the 133 metabolites monitored

Metabolite Category	Precision range (%)	Accuracy range (%)
Amino acids (26)	2.98-12.72	80-109
Nucleosides (11)	0.58-14.00	79-103
Nucleotides (18)	1.00-11.00	68-109
Organic acids (31)	0.43 – 25.00	75-125
Sugar phosphates (12)	4.19-12.25	75-115.5

Table 3.3 Full validation table for detected metabolites

Compound	Linear range (μM)	R^2	Precision (%)			Accuracy (%)		
			L(5 μM)	M (25 μM)	H (100 μM)	L(5 μM)	M (25 μM)	H (100 μM)
Hydroxyisocaproic acid	0.5-200	0.99	5.51	12.82	10.22	98.48	90.15	107.95
2-Hydroxyglutaric acid	0.5-200	0.98	36.92	10.25	9.62	23.05	110.78	116.93
2-Isopropylmalic acid	0.5-200	0.99	5.06	12.91	10.97	105.11	92.15	106.25
2-ketoisovaleric acid	0.5-100	1.00	25.06	33.04		95.40	101.94	0.00
Citraconic acid	0.5-100	0.99	4.68	11.75	10.00	115.88	92.46	103.12
N-acetylalanine	0.5-100	0.97	3.86	14.97	10.80	122.41	87.43	105.09
Hydroxyphenyllactic acid	0.5-200	0.98	3.38	11.34	14.64	136.75	80.98	104.08
Lactic acid	0.5-200	0.99	9.20	13.37	10.56	77.33	97.29	106.37
Pyroglutamic acid	0.5-50	0.98	6.55	18.42		104.92	90.97	0.00
Orotic acid	0.5-50	0.99	5.40	13.59		90.29	94.99	0.00
Adenosine	0.5-200	0.98	5.43	14.18	9.02	103.64	90.06	108.80
Uracil	0.5-200	0.96	8.12	48.28	12.34	97.77	79.40	119.66
Adenine	0.5-200	0.99	7.08	15.25	9.00	83.22	90.74	110.22
N-acetylglutamic acid	0.5-200	0.99	4.97	12.43	9.37	101.02	89.77	108.46
Uridine diphosphate-N-acetylglucosamine	0.5-200	0.99	7.54	13.81	9.34	90.70	92.08	109.18
Dihydroxyacetone phosphate (DHAP)	0.5-200	0.95	28.93	11.59	24.22	145.43	87.05	96.09
Inosine monophosphate (IMP)	0.5-25	0.96	11.01	14.52		47.96	101.17	0.00
2-Hydroxyglutaric acid	0.5-200	0.98	37.08	10.24	9.62	23.04	110.78	116.93
Acetoacetic acid	0.5-100	0.95	23.95	5.53	18.36	65.47	141.01	98.41
L-Tyrosine	0.5-100	0.96	2.18	12.39	22.13	49.14	100.15	101.34
			Precision (%)			Accuracy (%)		

Compound	Linear range (μM)	R ²	L(5 μM)	M (25 μM)	H (100 μM)	L(5 μM)	M (25 μM)	H (100 μM)
D-Fructose-6-phosphate	0.5-200	0.99	7.69	13.14	9.10	78.84	94.05	111.34
Adenosine triphosphate (ATP)	0.5-50	0.97	11.37	22.09		86.53	105.00	0.00
Succinic acid	0.5-200	0.99	8.96	12.12	11.02	91.19	99.40	91.50
L-Glutamine	0.5-200	0.99	8.03	13.56	9.32	69.67	97.47	109.84
L-Asparagine	0.5-200	0.99	8.00	12.93	9.77	73.23	89.73	109.10
Cytidine-diphosphate (CDP)	0.5-200	0.99	8.96	12.51	6.40	62.08	97.27	111.00
3-Phosphoglycerate	0.5-200	0.99	4.19	12.25	9.73	107.73	89.70	106.60
Fructose 1,6-bisphosphate	0.5-200	0.99	5.74	13.20	8.91	82.35	90.50	111.50
L-Citrulline	0.5-200	0.99	12.90	12.62	9.71	64.98	100.15	113.21
S-Adenosylmethionine (SAM)	0.5-200	0.99	9.98	13.19	7.74	51.78	98.51	113.81
L-Serine	0.5-200	0.99	4.26	13.68	3.41	117.86	85.36	103.40
Phosphoenolpyruvate	0.5-50	0.97	5.15	21.04		86.39	89.69	0.00
Uridine triphosphate (UTP)	0.5-200	0.99	7.46	10.74	12.39	117.02	88.20	109.02
Guanosine diphosphate (GDP)	0.5-200	0.99	27.13	12.73	9.97	38.28	100.04	111.12
trans-Aconitic acid	0.5-200	0.99	4.15	10.01	13.60	139.65	91.13	98.82
Citric acid	0.5-200	0.99	7.64	11.87	12.51	85.12	91.66	112.04
Guanosine triphosphate (GTP)	0.5-200	0.99	7.01	10.48	65.84	98.88	64.39	108.50
L-Lysine	0.5-200	0.99	10.22	12.72	8.82	65.88	96.88	111.87
Hydroxyphenylpyruvate	0.5-200	1.00	1.53	7.49	20.43	137.28	92.47	101.34
Butyryl-CoA	0.5-200	0.95	3.51	7.18	20.63	147.69	83.83	104.16
Deoxyadenosine	0.5-50	0.96	5.82	33.50	0.00	139.81	80.24	23.54
Dihydroorotic acid	0.5-200	0.99	2.66	6.51	7.96	113.61	91.91	109.68
					Precision (%)		Accuracy (%)	

Compound	Linear range (μM)	R ²	L(5 μM)	M (25 μM)	H (100 μM)	L(5 μM)	M (25 μM)	H (100 μM)
xanthosine	0.5-25	0.86	8.21	41.71		54.63	80.05	0.00
L-Leucine	0.5-200	0.99	0.71	5.72	1.79	97.47	97.71	109.95
Hypoxanthine	0.5-200	0.98	0.05	3.86	15.73	277.52	80.50	69.27
Deoxycytidine monophosphate (dCMP)	0.5-200	0.97	43.41	69.21	62.41	9.95	75.72	61.02
Dihydroorotic acid	0.5-200	0.99	2.66	6.51	7.96	113.61	91.91	109.68
Erythrose-4-phosphate	0.5-50	0.92	4.31	17.26		107.26	78.83	0.00
Fumaric acid	0.5-200	0.99	0.43	4.82	4.49	117.39	96.32	107.75
Glyceraldehyde-3-phosphate	0.5-200	0.99	6.83	8.24	10.89	163.94	75.47	98.11
Guanine	0.5-25	1.00	0.00	31.92	66.24	162.51	294.66	60.24
Hydroxyphenylpyruvate	0.5-500	1.00	1.53	7.49	20.43	137.28	92.47	101.34
Hypoxanthine	0.5-50	0.99	0.56	10.33		92.72	108.91	0.00
Inosine	0.5-200	1.00	2.82	19.64	9.50	53.13	44.38	165.81
Nicotinamide adenine dinucleotide (Khoury et al.)	0.5-200	0.99	1.01	6.17	4.52	125.73	95.10	102.40
Deoxycytidine monophosphate (dCMP)	0.5-200	0.99	9.42	36.96	58.86	0.00	103.70	23.24
Acetyl phosphate	0.5-200	0.98	0.91	2.34	18.51	309.81	81.42	57.76
N-acetyl-glucosamine-6-phosphate	0.5-200	0.99	1.67	7.62	8.68	146.64	92.02	89.43
Glyoxylic acid	0.5-25	0.98	11.47	20.07		58.44	89.61	0.00
L-Valine	0.5-200	1.00	2.88	4.74	1.45	105.57	94.06	109.12
Uridine 5'-monophosphate (UMP)	0.5-200	0.98	0.37	3.42	1.80	183.79	91.10	87.02
L-Glutamic acid	0.5-50	1.00	5.76	5.83	6.57	15.38	45.40	193.30
Cytidine	0.5-200	0.97	1.47	7.09	17.26	137.15	88.07	104.07
L-Proline	0.5-200	0.99	1.76	4.02	3.57	126.89	88.02	107.23
					Precision (%)		Accuracy (%)	

Compound	Linear range (μM)	R ²	L(5 μM)	M (25 μM)	H (100 μM)	L(5 μM)	M (25 μM)	H (100 μM)
Cytidine monophosphate (CMP)	0.5-200	0.99	2.15	8.95	3.18	119.13	93.70	107.17
L-Aspartic acid	0.5-200	0.99	1.29	3.77	2.19	187.89	85.34	92.71
Mesaconic acid	0.5-200	0.99	2.35	9.11	7.02	138.19	84.16	107.10
Erythrose-4-phosphate	0.5-200	0.97	0.50	5.61	61.08	222.91	58.39	105.03
Glyceraldehyde-3-phosphate	0.5-200	0.99	6.83	8.21	10.90	163.93	75.44	98.13
Sedoheptulose-7-phosphate	0.5-200	0.99	1.68	5.90	9.86	145.83	89.59	100.97
Malic acid	0.5-200	1.00	1.31	5.11	4.06	125.42	93.10	105.51
Fumaric acid	0.5-200	0.99	0.42	4.82	4.49	117.39	96.32	107.75
trehalose	0.5-200	1.00	2.83	7.17	2.45	97.54	99.98	108.83
L-Homoserine	0.5-200	0.98	9.69	6.04	3.84	41.01	115.94	117.29
Cytidine triphosphate (CTP)	0.5-50	0.99	13.83	62.67	37.61	98.84	100.13	
cis Aconitic acid	0.5-200	0.99	1.31	3.90	5.91	122.11	87.94	110.08
L-Ornithine	0.5-200	1.00	3.14	4.48	2.29	116.32	95.91	104.76
Phenylpyruvic acid	0.5-200	1.00	0.12	0.42	35.52	361.16	76.92	44.54
phenyllactic acid	0.5-200	1.00	1.27	3.71	7.17	130.75	87.02	96.22
N-Acetylcysteine	0.5-200	0.97	4.57	10.60	14.28	114.05	82.59	115.82
Pyruvic acid	0.5-200	0.97	0.73	4.66	8.44	218.47	63.53	105.53
Thymidine	0.5-50	1.00	3.17	29.51		109.50	66.84	0.00
N-acetylglutamine	0.5-200	0.99		3.62	7.52	0.00	92.39	107.57
Acetyl-CoA	0.5-200	0.97	2.16	13.77	26.74	117.25	88.42	108.16
O-Acetyl-L-serine	0.5-200	0.95	0.35	2.48	5.75	196.90	85.48	83.41
L-Phenylalanine	0.5-200	1.00	1.16	3.03	3.90	93.98	94.62	103.93
					Precision (%)		Accuracy (%)	

Compound	Linear range (μM)	R ²	L(5 μM)	M (25 μM)	H (100 μM)	L(5 μM)	M (25 μM)	H (100 μM)
Glyceric acid	0.5-200	1.00	3.06	4.23	3.49	108.75	89.65	100.92
Uridine	0.5-200	0.95	9.16	10.47	16.77	161.73	90.62	93.18
L-Isoleucine	0.5-200	1.00	21.40	5.47	5.51	127.99	90.59	105.55
Adenosine monophosphate (AMP)	0.5-200	1.00	0.27	2.76	4.69	205.48	90.92	75.14
L-Methionine	0.5-200	0.98	0.73	2.59	2.96	178.98	74.23	106.95
Acetyllysine	0.5-200	1.00	3.29	7.40	3.95	53.28	92.24	118.73
D-Gluconic acid	0.5-200	0.99	0.50	3.62	41.67	165.46	80.36	103.08
sn-Glycerol-3-phosphate	0.5-200	1.00	3.35	3.81	7.12	96.60	91.25	103.26
L-Tryptophan	0.5-200	0.96	1.13	1.74	6.16	223.78	63.87	103.67
Cytosine	0.5-200	0.99	10.22	55.70	45.73	132.93	86.94	108.15
Deoxycytidine diphosphate (dCDP)	0.5-200	0.99	1.38	7.33	37.20	247.05	78.44	67.03
D-glucose 6-phosphate	10-200	0.97		17.19	13.64		65.42	123.76
Guanosine	0.5-50	0.99	6.36	5.58		103.82	86.14	0.00
Glucosamine-6-phosphate	0.5-200	1.00	1.87	4.56	6.45	97.21	92.33	103.71
L-Threonine	0.5-200	1.00	4.64	2.16	2.49	75.59	99.16	105.91
Ketoglutaric acid (2-Oxoglutaric acid)	0.5-50	0.99	4.02	7.26		157.41	87.53	0.00
Uridine 5'-diphosphate (UDP)	0.5-200	1.00	1.54	4.12	8.04	92.86	95.39	101.97
N-Acetylornithine	0.5-200	0.99	2.87	4.40	6.57	171.00	85.12	86.92
Guanosine monophosphate (GMP)	0.5-200	1.00	11.07	3.76	4.31	20.02	105.09	114.92
L-Histidine	0.5-200	1.00	2.92	2.80	4.86	96.92	95.47	102.86
Cystathionine	0.5-25	1.00	4.29	4.33		104.92	99.82	0.00
Isocitric acid	0.5-200	1.00	10.24	13.80	3.34	87.58	90.98	118.37
					Precision (%)		Accuracy (%)	

Compound	Linear range (μM)	R²	L(5 μM)	M (25 μM)	H (100 μM)	L(5 μM)	M (25 μM)	H (100 μM)
L-Arginine	0.5-200	1.00	0.59	2.98	4.63	83.67	97.04	105.36

3.3.5 Optimisation of U-¹³C-labelled internal standard generation using *C. pasteurianum*

Clostridium pasteurianum (DSM 525) was the organism selected to generate U-¹³C-labelled internal standard for absolute quantification of intracellular metabolites in *C. autoethanogenum*. The growth media conditions were optimised to give the best results and the most economical approach. The initial growth media optimisation was done using ¹²C glucose at a starting concentration of 6 %. The culturing media for *C. pasteurianum* contained calcium carbonate (CaCO₃) and biotin. An experiment was carried where CaCO₃ and biotin were omitted from the media to observe their effects on cell growth. In another experimental set up, CaCO₃ was replaced with 2-(N-morpholino)ethanesulfonic acid (MES) (Lehmann and Lutke-Eversloh, 2011) with the addition and omission of biotin. The results are shown in Figure 3.7 and it was concluded that biotin was essential for the growth of the organism. There was no growth observed with the media containing MES. The media containing both CaCO₃ and biotin supported growth of *C. pasteurianum* much better than the media without biotin.

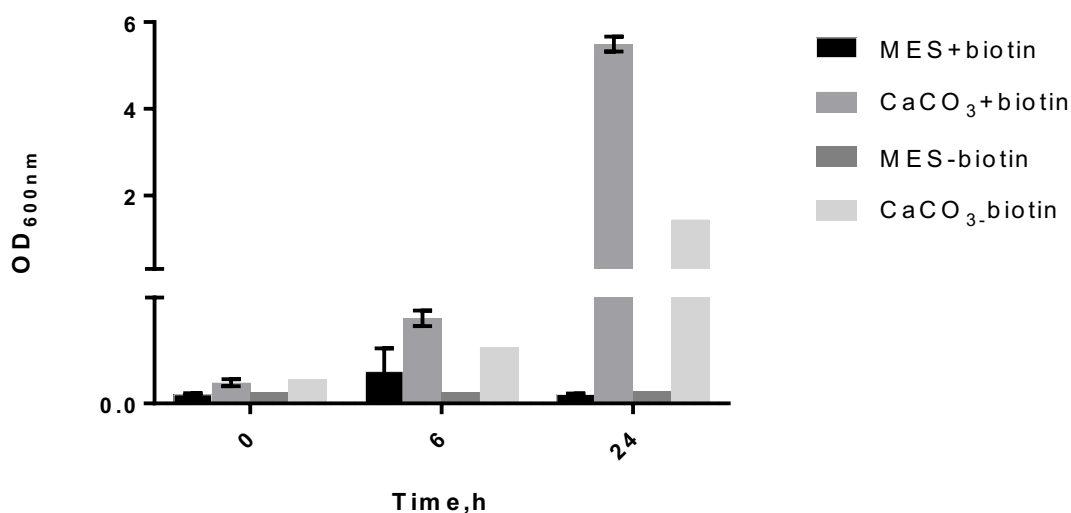


Figure 3.7 A plot of OD measurements over 24 hour period of *C. pasteurianum* cell culture. The plot shows comparison between CaCO_3 and MES as buffers for culturing of *C. pasteurianum* with and without the addition of biotin. The presence of biotin supports growth much better than omitting it and also CaCO_3 was a better buffer compared to MES.

To confirm that CaCO_3 was not a source of carbon (energy), ^{13}C -labelled $\text{Ca}^{13}\text{CO}_3$ was added to the media with the same concentration of unlabelled glucose used before. If the CaCO_3 was utilised by the organism, then we will expect some metabolites to be ^{13}C labelled. Since only uniformly ^{13}C -labelled metabolites were of interest, ToxID (ThermoFisher) was used to identify any metabolites that could have been U- ^{13}C -labelled and there were none. ToxID uses the accurate mass and expected retention times to detect compounds in their corresponding polarities. Accurate U- ^{13}C -masses of all metabolites were used to find if any metabolite was fully labelled, but no metabolites were detected in this form. This concludes that CaCO_3 only played the role as a pH buffer but not a source of carbon.

The concentration of glucose was also optimised to make it a practical compromise between ^{13}C yield and cost. An initial concentration of 6 % was used but other lower concentrations were evaluated as shown in Figure 3.8. The lowest plausible concentration for culturing of *C. pasteurianum* was selected to 1.5 %. Comparison of intracellular metabolites of *C. pasteurianum* grown on 1.5 % glucose gave 50 metabolites in common with the 133 metabolites of interest. Increasing the glucose concentration to 4% gave 100 metabolites in common with our metabolites of interest (Table 3.4). A final glucose concentration of 4 % was used throughout to generate U- ^{13}C -labelled internal standard for absolute quantification of intracellular metabolites of *C. autoethanogenum*.

The HILIC-IDMS method used here gave better coverage for quantification of intracellular metabolites found in *C. autoethanogenum* compared to similar methods reported for absolute quantification of intracellular metabolites from *T. brucei* (Kim et al., 2015) and *C. thermocellum* (Cui et al., 2013) using *E. coli* to generate the U- ^{13}C -labelled internal standards.

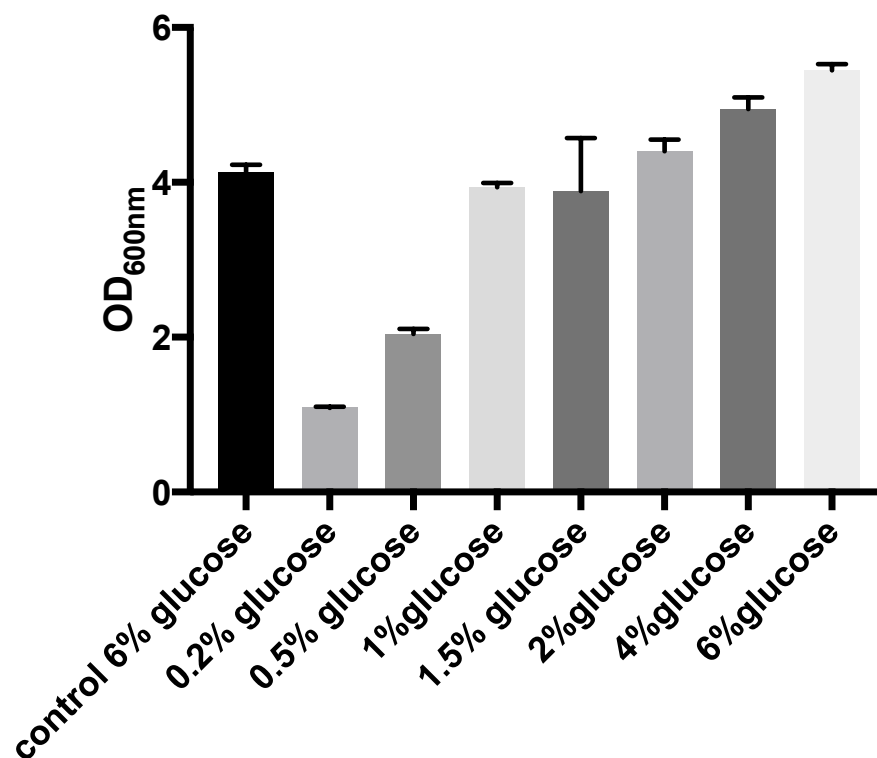


Figure 3.8 Effect of glucose concentrations ranging from 0.5 % to 6 % on the growth of *C. pasteurianum*. Optical density measured at 600 nm (OD_{600}) was used to determine cell mass for each experiment.

Table 3.4 List of metabolites among the 133 standard mixture found in *C. pasteurianum* depending on glucose concentration used as a means of generating U-13C-labelled internal standard. ToXID was used to identify compounds using their accurate mass and expected retention times.

1% glucose	2% glucose	4% glucose	m/z	Expected RT	Actual RT
		Glycine	74.025	9.29	9.74
		Pyruvic acid	87.009	7	7.19
		Oxalic acid	88.988	9.64	9.87
Lactic acid	Lactic acid	Lactic acid	89.024	7.19	7.22
Alanine	Alanine	Alanine	90.055	9.01	9.28
		Acetoacetate	101.025	6.99	7.17
	Glycerate	Glycerate	105.019	7.73	7.65
Serine	Serine	Serine	106.05	9.16	9.55
Uracil	Uracil	Uracil	111.02	7.68	7.85
Cytosine	Cytosine	Cytosine	112.051	8.43	8.14
Fumaric acid	Fumaric acid	Fumaric acid	115.004	9.3	9.32
		2-ketoisovaleric acid	115.04	5.46	5.27
Proline	Proline	Proline	116.071	8.67	8.78
	Succinic acid	Succinic acid	117.019	8.8	8.93
L-valine	L-valine	L-valine	118.086	8.47	8.59
homoserine	homoserine	homoserine	120.066	9.18	9.33
Threonine	Threonine	Threonine	120.066	8.73	8.24
		Pyroglutamic acid	128.035	7.44	7.22
Citraconic acid	Citraconic acid	Citraconic acid	129.019	7.75	7.7
Mesaconic acid	Mesaconic acid	Mesaconic acid	129.019	9.45	9.79
	N-acetylalanine	N-acetylalanine	130.05	6.88	7.18
Hydroxyisocaproic acid	Hydroxyisocaproic acid	Hydroxyisocaproic acid	131.071	5.17	5.26
	Ornithine	Ornithine	131.082	11.45	11.4
	Aspartic acid	Aspartic acid	132.03	8.81	8.91

1% glucose	2% glucose	4% glucose	m/z	Expected RT	Actual RT
L-leucine		L-leucine	132.102	7.84	8.12
Malic acid	Malic acid	Malic acid	133.014	9.17	9.25
Asparagine	Asparagine	Asparagine	133.061	9.15	9.33
Adenine	Adenine	Adenine	136.062	8.02	8.14
Hypoxanthine	Hypoxanthine	Hypoxanthine	137.046	8.13	8.27
		Acetyl phosphate	138.98	8.54	8.38
	Ketoglutaric acid	Ketoglutaric acid	145.014	8.85	9.22
Glutamic acid	Glutamic acid	Glutamic acid	146.045	8.65	8.53
O-Acetyl-L-Serine	O-Acetyl-L-Serine	O-Acetyl-L-Serine	146.046	7.1	7.55
2-hydroxyglutaric acid	2-hydroxyglutaric acid	2-hydroxyglutaric acid	147.029	8.7	8.6
Glutamine	Glutamine	Glutamine	147.077	9.05	9.19
L-Lysine	L-Lysine	L-Lysine	147.113	11.87	12.1
L-methionine	L-methionine	L-methionine	150.058	7.99	8.29
Xanthine	Xanthine	Xanthine	151.026	7.98	8.36
Orotate	Orotate	Orotate	155.009	7.64	7.77
histidine	histidine	histidine	156.077	9	9.33
Dihydroorotate	Dihydroorotate	Dihydroorotate	157.025	7.84	7.93
N-AcetylCysteine	N-AcetylCysteine	Phenylpyruvate	163.04	4.62	5.07
phenyllactic acid	phenyllactic acid	phenyllactic acid	165.056	4.9	5.22
Phenylalanine	Phenylalanine	Phenylalanine	166.086	7.54	7.81
	Phosphoenolpyruvate	Phosphoenolpyruvate	166.975	9.52	9.66
		Dihydroxyacetone phosphate	168.99	8.76	9.25
		Glyceraldehyde-3-phosphate	168.99	8.98	9.25
		sn-Glycerol--phosphate	171.006	8.63	8.46
trans-aonitic acid	trans-aconitic acid	trans-aonitic acid	173.009	9.74	9.79
2-Isopropylmalic acid	2-Isopropylmalic acid	2-Isopropylmalic acid	175.061	7.86	7.7
N-acetylornithine	N-acetylornithine	N-acetylornithine	175.108	8.89	9.16

1% glucose	2% glucose	4% glucose	m/z	Expected RT	Actual RT
Arginine	Arginine	Arginine	175.12	12.68	12.82
citrulline	citrulline	citrulline	176.104	9.31	9.52
	Hydroxyphenylpyruvate	Hydroxyphenylpyruvate	179.034	7.04	7.17
Hydroxyphenyllactic acid	Hydroxyphenyllactic acid	Hydroxyphenyllactic acid	181.05	7.24	7.34
Tyrosine	Tyrosine	Tyrosine	182.082	8.73	8.9
	Phosphoglycerate	Phosphoglycerate	184.985	9.29	9.66
N-acetylglutamic acid	N-acetylglutamic acid	N-acetylglutamic acid	188.056	8.38	8.2
N-acetylglutamine	N-acetylglutamine	N-acetylglutamine	189.087	7.21	7.18
Acetyllysine	Acetyllysine	Acetyllysine	189.123	8.2	8.42
Citric acid	Citric acid	Citric acid	191.019	9.7	9.69
		D-gluconate	195.051	8.29	8.5
		Erythrose-4-phosphate	199.001	8.97	9.29
Tryptophan	Tryptophan	Tryptophan	205.097	8.36	8.66
	Cystathionine	Cystathionine	221.06	9.1	9.44
N-acetylglucosamine	N-acetylglucosamine	N-acetylglucosamine	222.097	9.1	9.45
	Ribose-5-phosphate	Ribose-5-phosphate	229.011	8.2	8.42
	uridine	uridine	243.062	8.92	8.77
Cytidine	Cytidine	Cytidine	244.093	7.81	8.08
Deoxyadenosine	Deoxyadenosine	Deoxyadenosine	252.109	8.56	8.66
Glucosamine-6-phosphate	Glucosamine-6-phosphate	Glucosamine-6-phosphate	258.038	7.23	7.33
	D-Fructose-6-phosphate	D-Fructose-6-phosphate	259.022	8.78	9.09
	D-Glucose-6-Phosphate	D-Glucose-6-Phosphate	259.022	8.73	9.09
adenosine	adenosine	adenosine	268.103	8.73	9.08
Inosine	Inosine	Inosine	269.089	7.64	7.68
	Guanosine	Guanosine	284.099	8.15	8.24
		Sedoheptulose-7-phosphate	289.032	8.66	8.97
		N-acetyl-glucosamine-6-phosphate	300.049	8.98	9.08

1% glucose	2% glucose	4% glucose	m/z	Expected RT	Actual RT
CMP	CMP	CMP	322.044	8.3	8.28
UMP	UMP	UMP	323.028	8.81	8.89
trehalose/sucrose	trehalose/sucrose	trehalose/sucrose	341.108	8.64	8.48
Adenosine monophosphate	Adenosine monophosphate	Adenosine monophosphate	346.056	9.11	9.32
GMP	GMP	GMP	362.05	7.92	7.93
S-adenosyl methionine	S-adenosyl methionine	S-adenosyl methionine	399.145	9.17	9.32
	UDP-D-glucose	UDP-D-glucose	565.048	8.78	9.09
UDP-N-acetyl-glucosamine	UDP-N-acetyl-glucosamine	UDP-N-acetyl-glucosamine	606.074	8.37	8.29
		CDP	402.01	9.08	9.38
		UDP	402.995	9.17	9.32
		Adenosine diphosphate (ADP)	426.022	8.87	9.25
		GDP	442.017	8.56	8.46
		Flavin mononucleotide (FMN)	455.097	9.51	9.76
		UDP-D-glucose	565.048	7.25	7.6
		UDP-N-acetyl-glucosamine	606.074	8.78	9.1
NAD ⁺	NAD ⁺	NAD ⁺	664.116	8.37	8.38
dephospho-CoA	dephospho-CoA	dephospho-CoA	688.157	6.99	7.19
	FAD	FAD	784.149	8.06	8.01
		Acetyl CoA	808.119	7.36	7.52
	Butyryl-CoA	Butyryl-CoA	836.149	7.64	7.7

3.3.6 Cell number and viability

Cell numbers per 1 mL of sample were evaluated for four bioreactors to range between 1.1 E9 (1.1 billion) to 1.2 E9 (1.2 billion). Cell viability was also evaluated and was found to be between 79-82 %. An OD₆₀₀ was also measured at the time of sampling and there was no observed correlation between cell numbers per ml and OD₆₀₀.

Flow cytometry gave a good indication of the different populations that were present (Figure 3.9). Based on forward 2 and side scatter analysis four populations are apparent. 71% of the cells are in the two low side scatter populations. Viability determined by Syto9 and propidium iodide staining showed >87% of cells to be viable. Dead cells were evenly distributed through the four-population identified based on scatter analysis. Forward Scatter 1 (no mask) Pulse Height vs Area analysis suggests the populations with high FS1 (Mask M2) are either doublets/or elongated cells.

Application of these fluorescent staining to *C. autoethanogenum* cells and analysed using FCM gave indication of cell viability (Figure 3.10).

Table 3.5 Flow cytometry measurements of cell numbers per ml of sample and cell viability after propidium iodide/SYTO9 fluorescence staining for four biological replica bioreactors.

Measurements	Mean	RSD (%)
Cell number/ml	1.15E+09	5.02
Cell viability (%)	80.5	1.60
OD ₆₀₀	3.15	5.5

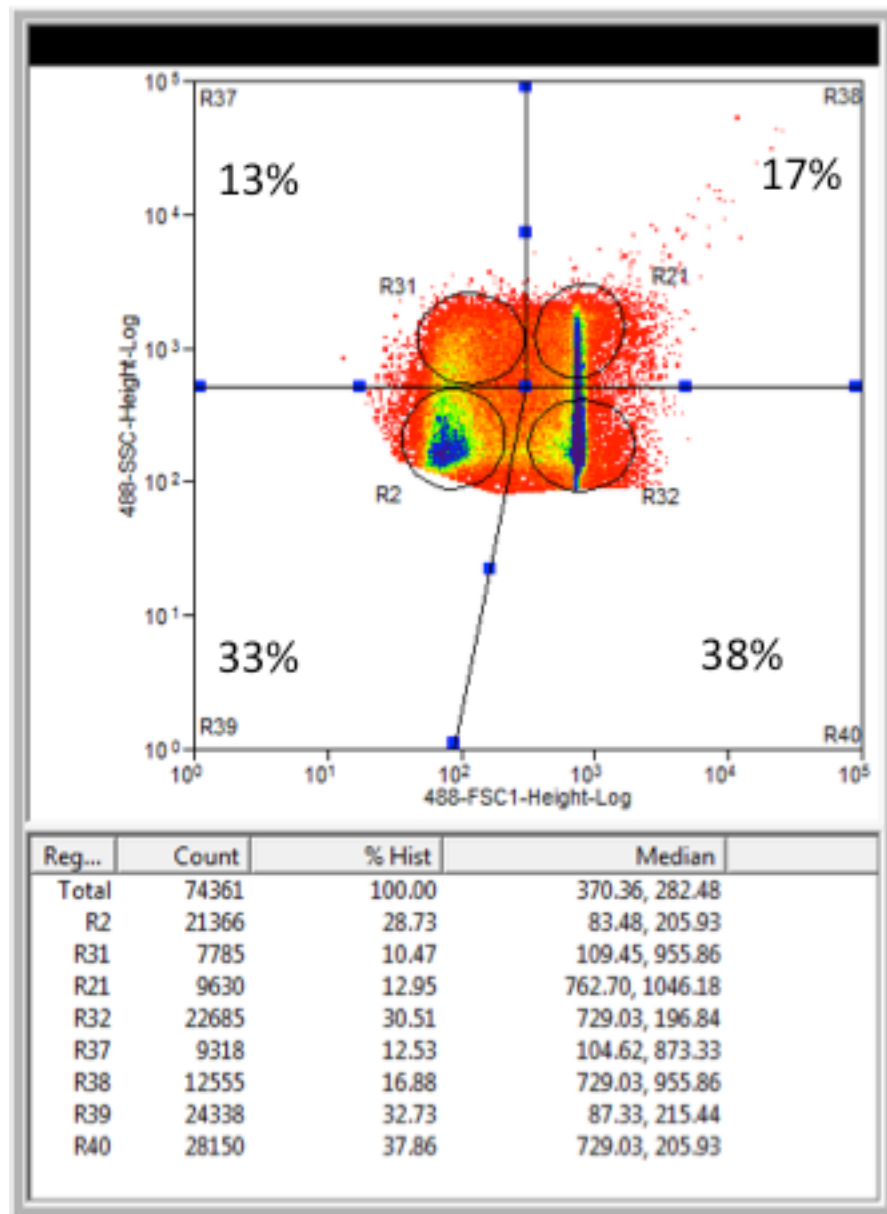


Figure 3.9 Flow cytometry measurements showing the presence of four different populations. Four Populations (Quadrants), R39 Low Side Scatter Low Forward Scatter (33%), R40 Low Side Scatter High Forward Scatter (38%), R37 High Side Scatter Low Forward Scatter (13%), R38 High Side Scatter High Forward Scatter (17%).

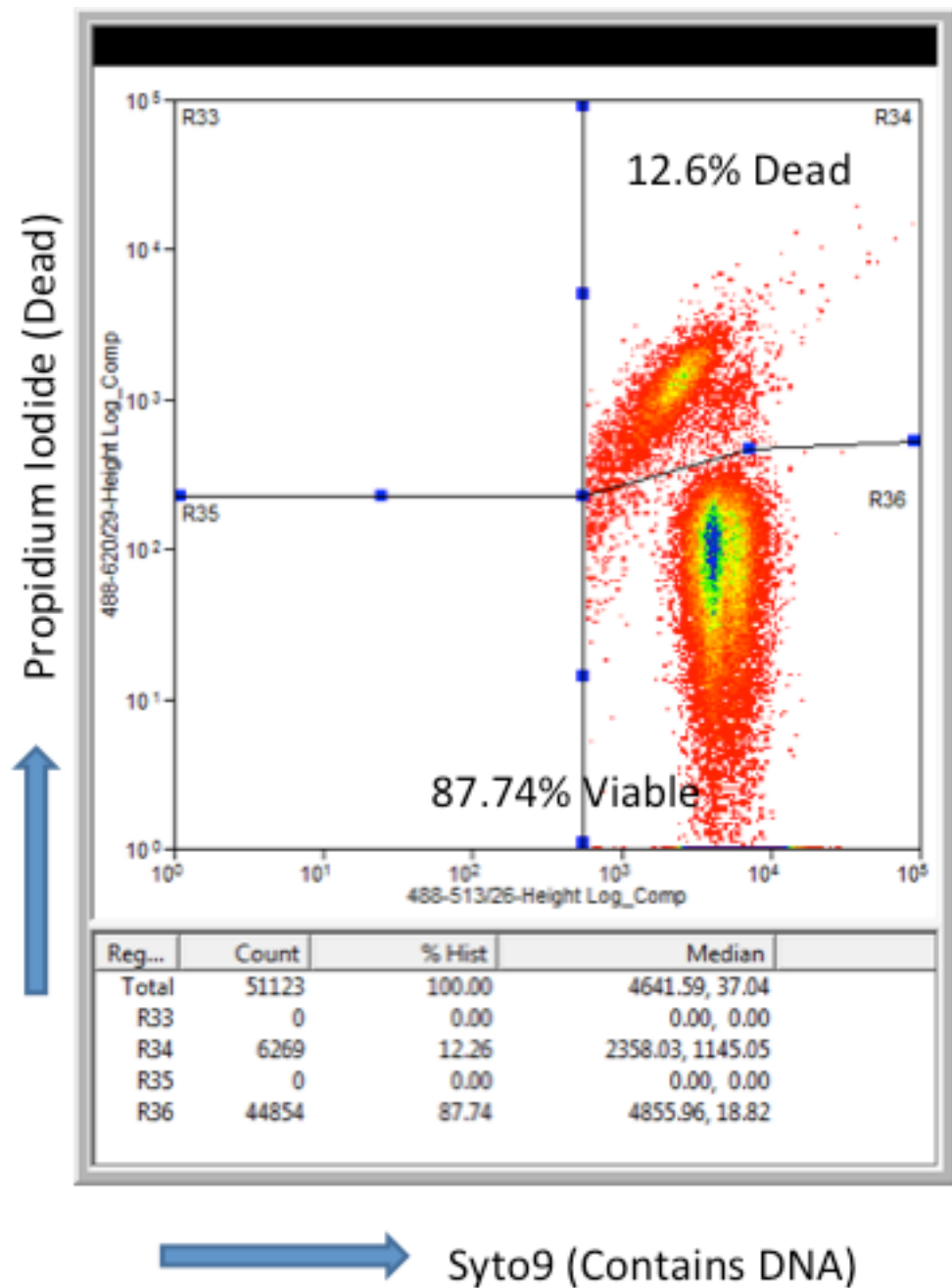


Figure 3.10 Bac Light Viability Gated on all Syto9 positive events (i.e. excluding debris and abiotic bacteria).

3.3.7 Determination of cell volume of *C. autoethanogenum*

The cell volume is important for absolute quantification of metabolites, hence making a significant contribution to quantitative metabolomics studies. With the cell volume of *C. autoethanogenum*, intracellular metabolites concentrations can be generated for metabolic pathway modelling. The reporting of concentration using the cell volume is widely applied (Tycko et al., 1985).

Microscopy images were used to determine cell size of *C. autoethanogenum* taken from 4 bioreactors. Cells were stained with SYTO9/propidium iodide as detailed above before microscopy imaging to determine cell viability. Cell volume of *C. autoethanogenum* was determined by using the formula already reported (Heldal et al., 1985, Loferer-Krossbacher et al., 1998, Volkmer and Heinemann, 2011) assuming cells are cylindrically shaped with two sphere-capped ends.

Average cell volume of *C. autoethanogenum* was determined to be between 2.6-2.8 fL (1 fL = $1\mu\text{m}^3$). A median cell volume was determined to be 2 – 2.1 fL as shown in the histograms (Figure 3.11). Regular shaped and elongated cells were observed. Regular shaped cells had a greater contribution to the total viable cells with little contribution from elongated cells (Table 3.6). For confidence in our cell volume measurements, we set up a control experiment where we added fluorescent beads (known size) to confirm the accuracy of our measurements. Information of cell volume of *Clostridium autoethanogenum* has not been reported but reported for other *Clostridium sp.* such as *Clostridium perfringens* (Guerlava et al., 1998).

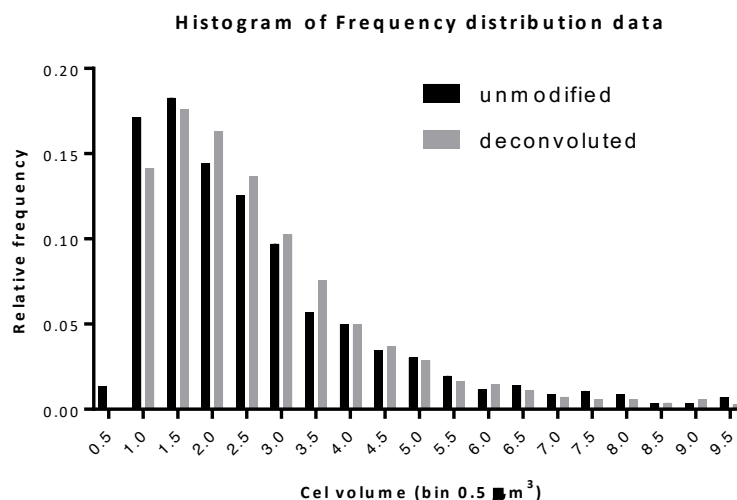


Figure 3.11 This histogram of cell volumes as determined by microscopy shows a skewed distribution where most cell volumes are 0.8-5 μm^3 . Frequencies of both unmodified pictures and deconvoluted pictures are given.

After obtaining cell volume of *C. autoethanogenum* from microscopy, the next step was to investigate if cell volume could be obtained from flow cytometry measurements as previously reported (Bouvier et al., 2001). A relationship has been observed between FSC and bacteria cell volume (Akerlund et al., 1995, Bouvier et al., 2001, Boye et al., 1983). Also, cell volume obtained from microscopy were plotted against SSC and we observe a correlation. Although cell volume measurements of *C. autoethanogenum* can be estimated using both SSC and FSC, the flow cytometry needs to be calibrated in relation to the scattering signals after which the data shown in Figure 3.12 can be used.

Table 3.6 Cell volume of *C. autoethanogenum* from flow cytometry measurements after calibration of scattering signal using microscopy data. Cell viability and cell population distribution can be reported for flow cytometry measurements. Cell population distribution showed majority of viable cells were regularly shaped and 10-15% were elongated. Smaller cell volume reported for the elongated cells were due to observed smaller widths compared to regular shaped cells.

	Total viable cells	Regular shaped cells	
--	---------------------------	-----------------------------	--

	Mean Size (um3)	Median Size (um3)	Mean Size (um3)	Median Size (um3)	% Viable
Mean	2.70	2.05	2.74	2.12	88.98
SD	0.06	0.04	0.07	0.05	2.72
RSD %	2.34	1.71	2.46	2.20	3.06

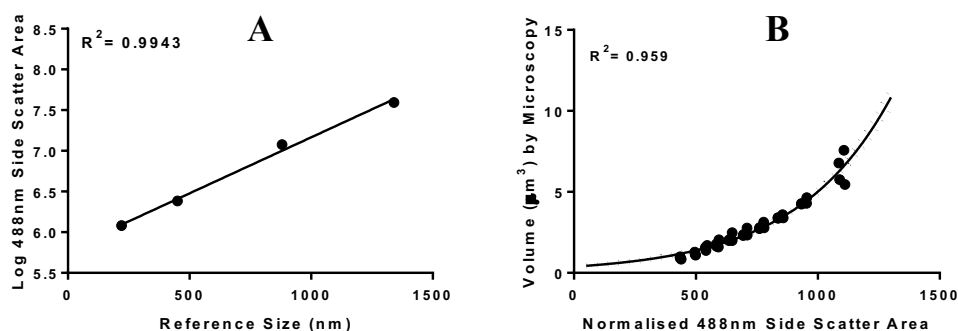


Figure 3.12 A: Calibration Curve of References Size Standards to Log 488nm Side Scatter. Log side scatter area measurement of Sphero Nano Fluorescent Particle Standard kit (Spherotech Inc) vs reported diameters. Line represents linear curve fit performed by Prism (V7.01 GraphPad). Relationship = $Y = 0.001376 * X + 5.789$. B: Correlation of devolution microscopy volume measurement and normalised side scatter area flow cytometry measurement from 4 independent reactors. Each dot represents the median 10-percentile value from an individual reactor measured by both techniques. Line represents exponential non-linear curve fit performed by Prism (V7.01 GraphPad). Relationship: $Volume = 0.3844 * \exp(0.002568 * \text{normalised SSC-A})$.

3.4 Conclusion

The use of *C. pasteurianum* for generation of U-¹³C-labelled internal standard provided better coverage for metabolites of *C. autoethanogenum* than *A. platensis*.

This is the first time that ¹³C uniformly labelled multiple internal standards have been used to develop a quantitative method for intracellular metabolite measurements in *C. autoethanogenum*. The method has wider applicability to other organisms where high-quality analytical data are required for pathway modelling purposes.

In most quantitative methods where metabolite concentrations are reported, the cell volume of the organism are usually not accounted for, but absolute intracellular concentration of metabolites present in *C. autoethanogenum* have been reported here using cell volume which has never been reported before.

CHAPTER FOUR

**Application of HILIC-IDMS method to
determine the effect of pH on ethanol production
in *C. autoethanogenum***

4. Application of HILIC-IDMS method to determine the effect of pH on ethanol production in *C. autoethanogenum*

4.1 Introduction

The industrial application of gas fermentation using *C. autoethanogenum* is flexible and robust due to the high selectivity for target end products and is economically feasible even with a small volume of gas and there is good tolerance for contaminants from waste gases (Daniell et al., 2012, Griffin and Schultz, 2012). *C. autoethanogenum* metabolism is characterised by an acidogenesis phase and a solventogenic phase. Exponential growth occurs during the acidogenesis phase after which solvent production and re-assimilation of acid follows. The acidogenic phase involves increased production of acids (acetate) while the solventogenic phase involves increased solvent production (ethanol, 2,3BD, lactate) as cell growth slows down. The acidogenic/solventogenic phase transition is common between *Clostridia sp.* but poorly understood (Alsaker and Papoutsakis, 2005, Amador-Noguez et al., 2011). A broad understanding of the mechanism that facilitates the transition to the solventogenic phase would lead to improved success rates for commercial production of solvent by *C. autoethanogenum*.

In *C. autoethanogenum*, levels of acetate have been reported to be higher than ethanol in many studies (Abubackar et al., 2016, Gupta and Gao, 2009). These high levels of acetate limit the application of *C. autoethanogenum* in biofuel industries as the carbon flows more in the direction of acetate production than the desired ethanol product. It has been reported that factors like low pH and absence or presence of specific trace elements might enhance production of ethanol (Abubackar et al., 2016, Abubackar et al., 2012b, Abubackar et al., 2015c). At pH values below 7 acetate is found in the free acetic acid form and it has been reported that the acetic acid produced at lower pH may play inhibitory role in homoacetic fermentation (Schiel-Bengelsdorf and Durre, 2012). Recently, the production of unwanted acetate by *C. autoethanogenum* has been minimised by optimising the fermentation conditions (Abubackar et al., 2015c). The reported pH known to support growth in *C. autoethanogenum* is 5.8 (Abrini

et al., 1994). The reduction of pH from 5.75 to 4.75 has been shown to enhance ethanol production to about 200% in *C. autoethanogenum* from batch bottles (Abubackar et al., 2012b). In a similar experiment using continuous gas fed bioreactors at pH 5.75, lower ethanol concentrations were observed but ethanol concentrations increased when pH was reduced to 4.75 (Abubackar et al., 2015c). The reason for the higher levels of ethanol at acidic pH is that *C. autoethanogenum* can utilise acetate accumulated over time during growth on CO gases (Abubackar et al., 2016). Acetaldehyde:ferredoxin oxidoreductase (Li et al.) in *C. autoethanogenum* is responsible for the conversion of the accumulated acetate to acetaldehyde and alcohol dehydrogenase (Gupta and Gao) converts acetaldehyde into ethanol (Isom et al., 2015). So far experiments have shown that starting growth of *C. autoethanogenum* at the optimum pH resulted in an increase in the biomass and, when followed by a reduction of pH the production of ethanol was enhanced. Hence this pH shift can prevent the unwanted accumulation of acetate (Abubackar et al., 2016).

Although reduction in pH increases solvent production, very little is reported on the effects these changes have on intracellular intermediates that lead to the production of ethanol. The use of advanced analytical methods provides high quality metabolite data for pathway modelling. This type of metabolite information is critically important in generating good mathematical models of key pathways involved in solventogenesis. The HILIC-IDMS method developed in chapter 3 will be applied to evaluate its performance in a challenging application to study the effect pH on continuous culture in a bioreactor.

4.1.1 The role of metabolomics in explaining the acidogenesis-solventogenesis transition in *C. autoethanogenum*

Studies of acidogenic-solventogenic transitions in *Clostridia* sp. are usually focussed on the investigation of the biochemical and transcriptional regulation of pathways that are directly associated with production of acid and solvent (Green et al., 1996). However, the acidogenic/solventogenic pathways only constitute a small fraction of the whole metabolomics network, whose

functionality is dependent on the availability of substrates such as acetyl CoA, NADH, ATP and NADPH from central metabolic pathways. Hence, it is important to note that the solventogenic transition is dependent on the broad metabolic changes occurring in the organism. The acidogenic/solventogenic transition is well reported in *C. acetobutylicum* using transcriptomics (Alsaker and Papoutsakis, 2005, Janssen et al., 2010, Jones et al., 2008) and metabolomics approaches (Amador-Noguez et al., 2011). However, *C. autoethanogenum* has not been investigated in detail and the closest attempt to understand the acidogenic/solventogenic transition was done by (Marcellin et al., 2016b) where they combined proteomics, transcriptomics and metabolomics approaches to understand energy metabolism in the organism but did not use metabolomics data in detail. Due to the complex nature of the changes and controlled metabolism at many levels, transcriptional data are usually insufficient in determining the overall metabolic changes. The use of metabolomics provides a cellular level of data that represents the effects of acidogenic/solventogenic transition on individual metabolites, which can improve the understanding of metabolic pathways affected. Metabolites selected cover one carbon pathway, TCA cycle and central metabolic pathway of *C. autoethanogenum*. Metabolites from these pathways will aid in understanding the cellular events that occur during the acidogenic/solventogenic transition in *C. autoethanogenum*.

The HILIC-IDMS method with ¹³C-labelled internal standards from *Clostridium pasteurianum* contributes to the generation of improved quality data. This is the first time a detailed metabolomics study of pH-induced solventogenesis in *C. autoethanogenum* has been reported.

4.1.2 Aims and Objectives

The objective of this chapter is to use targeted metabolomics approach to understand the changes that occur to intracellular metabolites during the acidogenic/solventogenic transition in *C. autoethanogenum*. The list of 133 metabolites chosen covers a wide range of the metabolic pathways of *C. autoethanogenum*, hence provides the potential for a deeper interpretation of the metabolic pathways affected during the acidogenic/solventogenic transition.

4.2 Materials and methods

4.2.1 Reagents and materials

All chemicals and reagents used are listed in Chapter 3, section 3.2.1.

4.2.2 Culture Conditions for *Clostridium autoethanogenum* for Gas and pH Shift Experiments

These pH shift bioreactor experiments were performed by Dr Anne Henstra and Louise Sewell (Synthetic Biology Research Centre (SBRC), Centre for Biomolecular Science, University of Nottingham). *C. autoethanogenum* DSM 10061 was purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH. The medium was maintained at pH 5 with composition of salts listed below per litre of distilled water: MgCl₂·6H₂O, 0.5 g; CaCl₂·2H₂O, 0.37 g; KCl, 0.15 g; NaCl, 0.12 g; NaH₂PO₄·2H₂O, 0.866 g; NH₄Cl, 1.0 g; resazurin, 1 mg; thiamine HCl, 0.5 mg; riboflavin, 0.5 mg; nicotinic acid, 0.5 mg; pantothenic acid, 0.5 mg; pyridoxine HCl, 0.1 mg; biotin, 0.2 mg; folic acid, 0.2 mg; 4-aminobenzoic acid, 0.5 mg; cyanocobalamin, 0.5 mg; lipoic acid, 0.5 mg; FeCl₂·4H₂O, 19.35 µg; NiCl₂·6H₂O, 1.19 µg; ZnCl₂, 0.69 µg; CoCl₂, 0.476 µg; HBO₄, 0.124 µg; MnCl₂·4H₂O, 0.396 µg; NaMoO₄·2H₂O, 0.484 µg; Na₂SeO₃, 0.346 µg; and Na₂WO₄·2H₂O, 0.658 µg. Experiments were conducted in a 3 L bioreactor (New Brunswick Scientific BIOFLO 115) containing 1.5 L culture volume and 100 % CO as the gaseous C1 substrate. pH was controlled by the addition of 2 M NaOH. Foaming was controlled by addition of antifoam (Sigma Aldrich, St. Louis, MO, USA) and sodium sulphide (100 mM). Agitation of the culture was at 600 rpm with a media dilution rate of 1 L/day (Masterflex L/S pump). CO was continuously fed into the bioreactor at a constant flow rate controlled through thermal mass flow controllers (TMFC) (Sierra SmartTrak100) at a rate of 15 and 27 mL/min and Nitrogen (N₂) as a balance gas to make up a total of 100 mL/min.

4.2.3 Product shift by pH

All bioreactor experiments and preparation of growth media in this chapter were performed by Dr Anne Henstra and Louise Sewell of the Synthetic Biology Research Centre (SBRC), Centre for Biomolecular Science, University of Nottingham. A seed reactor (B041) which will be the main focus in this chapter, was started from PETC-MES with CO starter cultures and served to inoculate three bioreactors that were used to perform the pH-shift experiment (B042, B043, B044). The seed reactor was then used to pilot the pH-shift. For the experiment all reactors were operated over the same four conditions (15mL/min CO, pH 5.0; 27mL/min CO, pH 5.0; 27mL/min CO, pH 4.5; 27 mL/min CO, pH 5.0), except for the pilot reactor which had an additional pH step (27mL/min CO, pH 4.7) inserted before the pH 4.5 step. The total gas flow was kept constant at 100 mL/min by balancing the N₂ flow. Reactors were run with the intent to obtain each condition in steady-state, with steady-state defined as a minimum of 3 volumes dilution over which OD₆₀₀, acetate and ethanol concentrations varied less than 10%.

4.2.4 Sample Collection and Extraction of Metabolites from *C. autoethanogenum*

C. autoethanogenum cells were obtained from anaerobic gas fermentation cultures at 37°C. Cells were harvested when optical densities measured at 600 nm (OD₆₀₀) ranged from 3 to 5. 1 mL (3 technical replicates) of culture was immediately cooled on ice, and then centrifuged (10000 g for 5 min) at 4°C. The supernatant was removed, snap freeze cell pellets and kept at -80°C. Cell pellets were re-suspended in 100 µL ¹³C labelled *C. pasteurianum* extract in methanol. The re-suspended cells were lysed by freezing in liquid nitrogen, thawing on ice and vortexing for 30 s. The freeze-thaw approach was repeated three times. The lysed cells were then diluted with 100 µL with precooled water and 200 µL of chloroform. The mixture was centrifuged at 10000 g for 5 min to remove cell debris. Aqueous layer was transferred into HPLC vials and stored 4 °C prior to

LC-MS analysis. Timelines from sample collection to data analysis are shown in Figure 4.1 below.

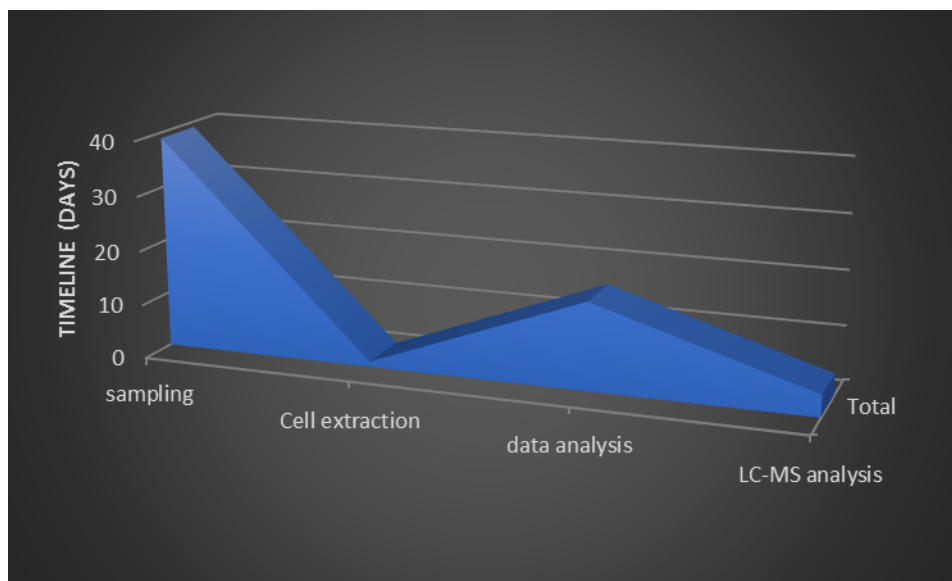


Figure 4.1 The timeline for sampling, extraction, LC-MS analysis and data analysis for pH shift experiment.

4.2.5 Calculation of intracellular concentration using cell volume

Intracellular concentrations were evaluated using the cell volume of *C. autoethanogenum*. Concentrations were first normalised to the actual volume of cells collected (for eg. 0.1 mL to 1 mL). Intracellular concentrations were then normalised using number of cells per mL (1.15×10^9) and the cell volume of 2.1 fL.

4.2.6 LC-MS and data analysis

Analysis of raw files containing mass spectrometry data was done using Tracefinder (ThermoFischer Scientific). Compounds were identified using the accurate mass and retention time from authentic standards. Calibration curves for quantification of metabolites were plotted using Graphpad. Metabolic

pathway analysis was done using web-based platforms Escher (<https://escher.github.io/>), MetaboAnalyst 3.0 and biocyc.org.

4.2.7 Transcriptomics analysis

Transcriptomics analysis by RNAseq was performed by the Deep Seq next generation sequencing facility (The University of Nottingham). RNA for transcriptomics analysis with RNA integrity number (RIN) of 8 or better was isolated using the following kits following manufacturers recommendations; FastRNA PRO Blue (MP Biomedicals) for cell disruption in a FastPrep 24 homogenizer (MP Biomedicals); DirectZol RNA miniprep (Zymo Research Products) for RNA isolation; Ambion Turbo DNase (Thermo-Fisher Scientific) for DNA removal; RNA clean & concentrator (Zymo Research Products) for RNA clean up. An on-column DNA digestion at room temperature (RT) and a post column DNA digestion at 37 °C for 30 min were performed. Only small batches of up to 4 cell pellets were processed through the homogenisation steps at a time. Presence of remaining DNA was tested by PCR with primers F and R. Quality and quantity of the obtained RNA was assessed by Nanodrop (Thermo Scientific), Qubit (Thermo-Fisher Scientific) and RNA-nano gel-electrophoresis (2100 Bioanalyser, Agilent).

4.3 Results and discussion

4.3.1 Targeted metabolomics approach for analysis of pH shift experiment in *C. autoethanogenum*

A total of 86 metabolites were quantified out of the 133 using U-¹³C-labelled internal standard from *C. pasteurianum*. A full list of metabolite concentrations measured during the experiment are provided in Table 4.2. During exponential growth AMP, ADP, cytosine and NAD⁺ were at higher levels compared to solventogenic phase except in the case of AMP and ADP where levels continued to increase during solventogenic phase (Figure 4.2). An increase in the levels of ADP and AMP during solventogenic phase might be as a result of ATP utilisation during acid re-assimilation to produce solvents (Amador-Noguez et al., 2011). Increase of CO flow rate to 27 mL/min instantly increased levels of certain metabolites. Metabolites such as L-citrulline, L-homoserine, L-proline, L-threonine, L-arginine, phenylacetylglutamine, Sn-glycerol-3-phosphate, UDP, and uracil. On the contrary, levels of some metabolites decreased during the CO shift. Examples include acetylcysteine, cytosine, cis/trans-aconitate, D-gluconate, inosine, guanosine, L-methionine, L-glutamine, L-tryptophan, N-acetyl-D-glucosamine, NAD⁺, phenyllactate, phosphoenolpyruvate, pyruvate, 2-methylmaleate, 2-oxoglutarate, UTP, 3-(4-hydroxyphenyl)lactate, 3-(4-hydroxyphenyl)pyruvate and 3-methyl-2-oxobutanoic acid.

The heat maps shown in Figure 4.3 showed an orderly progression of changes observed in intracellular metabolites during the gas and pH shift phases. As stated above, change in gas flow rate had less significant changes compared to pH shift. The most significant changes to intracellular metabolites were observed during pH shifts where the organism was mainly in the solventogenic phase. In Table 4.1 the two steady states of acetate and ethanol production were compared and generally there was an increase in the intracellular concentrations of most metabolites from acidogenesis to solventogenesis. The colour code red indicates a lower concentration while green denotes a higher concentration. Acetoacetate and acetylphosphate were among the few metabolites that had concentrations lower in solventogenesis than acidogenesis.

Statistical analysis was conducted to confirm if the changes observed were significant. P-values were evaluated using t-test using $\alpha=0.05$) (Table 4.1). The p-values were adjusted using Bonferroni multiple test correction to correct the errors associated with multiple testing. Initial t-test conducted showed 64 metabolites were significant with p-values <0.05 but after Bonferroni correction only 52 metabolites had changes that were significant. This shows the importance of multiple testing to avoid reporting changes that were not significant (Milroy, 2016).

Intracellular metabolic levels have been reported by Amador-Noguez (Amador-Noguez et al., 2011) where they characterised the metabolic changes associated with the transition from exponential phase to solventogenic phase in *C. acetobutylicum*. 114 metabolites were monitored, and 72 metabolites were quantified for both the exponential and solventogenic phases. They also observed an increase in arginine, AMP during the transition from acidogenic phase to solventogenic phase. They also report intracellular metabolite concentrations with no indication if these concentrations were normalised to cell volume of *C. acetobutylicum*. Marcellin et al (2016) report the presence of overflow products like lactate which is of biotechnological interest and succinate from the TCA cycle (incomplete) of *C. autoethanogenum*, which was also observed in this experiment. The method used here is a superior method for absolute quantification of intracellular metabolites with the use of U- ^{13}C -labelled internal standard and a known cell volume for *C. autoethanogenum*.

Further interpretation of the results was done by pathway analysis using metaboanalyst3.0 and Escher (https://escher.github.io/builder/index.html?enable_editing=true&js_source=local#

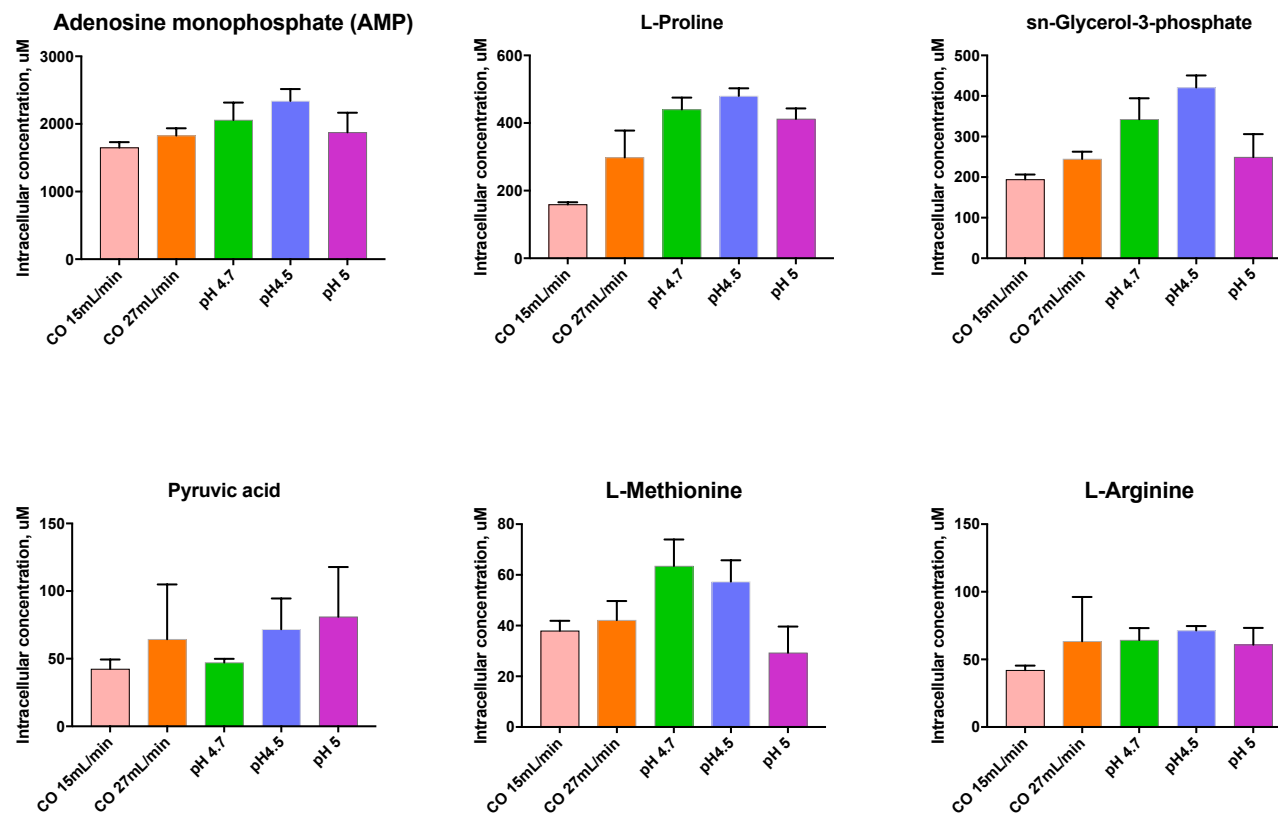


Figure 4.2 Plots of representative metabolite concentration at different bioreactor conditions during the pH shift experiment using *C. autoethanogenum*.

Table 4.1 Heat map of intracellular concentrations of metabolites in mM extracted from C. autoethanogenum to represent intracellular metabolite concentration during the acetate and ethanol phases. Red colour indicates lower concentration while green represents higher concentration. Acetate phase is an average of concentration from days 12-18 (n=21) and ethanol phase is an average of intracellular metabolites from days 19-26 (n=24). Standard deviation (SD) of both phases are shown. T-test was performed on the dataset and p-values are shown to be significant when p-value<0.05. Adjusted p-value evaluated using Bonferroni correction for multiple comparisons.

Metabolites	Acetate phase (mM)	Ethanol phase (mM)	SD acetate	SD ethanol	p-values	Adjusted p-value
N-acetylalanine	0.080	0.165	0.024	0.010	2.567E-14	2.207E-12
L-Lysine	0.364	0.669	0.113	0.061	8.328E-12	7.162E-10
L-Aspartic acid	1.041	2.219	0.297	0.222	1.387E-10	1.193E-08
S-Adenosylmethionine (SAM)	0.586	0.859	0.143	0.069	1.746E-10	1.502E-08
Cytidine monophosphate (CMP)	0.196	0.308	0.039	0.027	3.906E-10	3.359E-08
Adenosine	0.108	0.150	0.028	0.011	7.901E-10	6.795E-08
L-Proline	0.636	0.988	0.241	0.051	2.018E-09	1.736E-07
L-Tyrosine	1.169	1.618	0.254	0.147	3.758E-09	3.232E-07
Uridine 5'-monophosphate (UMP)	0.383	0.493	0.099	0.031	9.525E-09	8.191E-07
sn-Glycerol-3-phosphate	0.516	0.831	0.130	0.100	1.187E-08	1.021E-06
L-Threonine	0.784	1.084	0.182	0.109	1.205E-08	1.036E-06
L-Homoserine	0.498	0.651	0.120	0.056	1.409E-08	1.212E-06
Uridine diphosphate-N-acetylglucosamine	0.601	0.786	0.177	0.053	1.911E-08	1.643E-06
L-Citrulline	0.645	0.906	0.258	0.070	3.519E-08	3.026E-06
Uridine 5'-diphosphate (UDP)	0.780	1.012	0.208	0.101	5.705E-08	4.907E-06
Adenosine monophosphate (AMP)	3.858	4.744	0.850	0.415	8.275E-08	7.117E-06
Glucosamine-6-phosphate	0.283	0.439	0.083	0.064	1.605E-07	1.380E-05
Adenine	0.028	0.035	0.008	0.003	1.678E-07	1.443E-05

Metabolites	Acetate phase (mM)	Ethanol phase (mM)	SD acetate	SD ethanol	p-values	Adjusted p- value
Acetyllysine	0.013	0.020	0.004	0.003	3.470E-07	2.984E-05
Orotic acid	0.017	0.029	0.006	0.005	4.721E-07	4.060E-05
Guanosine monophosphate (GMP)	0.045	0.080	0.013	0.013	6.555E-07	5.637E-05
L-Asparagine	0.457	0.899	0.104	0.137	7.469E-07	6.423E-05
N-Acetylornithine	0.051	0.103	0.016	0.019	1.001E-06	8.609E-05
L-Leucine	0.131	0.152	0.033	0.013	1.015E-06	8.725E-05
Cytidine-diphosphate (CDP)	0.017	0.030	0.003	0.004	1.215E-06	1.045E-04
Mesaconic acid	0.738	0.979	0.327	0.122	1.501E-06	1.291E-04
L-Ornithine	0.596	0.871	0.221	0.160	1.791E-06	1.541E-04
L-Methionine	0.089	0.127	0.021	0.019	1.796E-06	1.544E-04
D-Fructose-6-phosphate	2.169	2.636	0.601	0.335	1.853E-06	1.594E-04
N-acetylglutamic acid	0.466	0.572	0.060	0.054	3.580E-06	3.079E-04
Uracil	0.190	0.222	0.058	0.028	3.695E-06	3.178E-04
L-Glutamic acid	6.360	7.015	1.625	0.467	4.438E-06	3.817E-04
Hydroxyphenyllactic acid	0.004	0.006	0.001	0.001	4.998E-06	4.298E-04
2-Hydroxyglutaric acid	1.509	1.879	0.425	0.282	5.494E-06	4.725E-04
N-acetyl-glucosamine-6-phosphate	1.096	1.263	0.303	0.151	6.368E-06	5.477E-04
Sedoheptulose-7-phosphate	0.386	0.549	0.120	0.108	1.249E-05	1.074E-03
3-Phosphoglycerate	0.397	0.600	0.102	0.110	1.249E-05	1.074E-03
L-Glutamine	1.866	3.653	0.529	0.633	1.287E-05	1.107E-03
Hydroxyisocaproic acid	0.041	0.077	0.009	0.015	1.325E-05	1.140E-03
L-Phenylalanine	0.264	0.317	0.109	0.031	1.645E-05	1.415E-03
phenyllactic acid	0.005	0.007	0.002	0.001	2.0861E-05	1.794E-03

Metabolites	Acetate phase (mM)	Ethanol phase (mM)	SD acetate	SD ethanol	p-values	Adjusted p- value
Lactic acid	0.821	1.717	0.156	0.339	2.592E-05	2.229E-03
Guanosine diphosphate (GDP)	1.007	1.394	0.270	0.262	2.695E-05	2.318E-03
Ribose-5-phosphate	1.533	2.292	0.345	0.443	5.417E-05	4.658E-03
Malic acid	0.023	0.034	0.007	0.008	6.302E-05	5.420E-03
Glyceric acid	0.159	0.267	0.062	0.066	6.810E-05	5.857E-03
2-ketoisovaleric acid	0.002	0.004	0.000	0.001	7.736E-05	6.653E-03
Succinic acid	4.640	7.200	1.271	1.752	1.610E-04	1.384E-02
Hydroxyphenylpyruvate	0.001	0.001	0.000	0.000	3.787E-04	3.257E-02
Xanthine	1.515	1.537	0.421	0.212	4.194E-04	3.607E-02
Uridine triphosphate (UTP)	0.026	0.039	0.008	0.011	4.265E-04	3.668E-02
L-Arginine	0.135	0.146	0.059	0.011	5.676E-04	4.881E-02
trans-Aconitic acid	0.004	0.013	0.002	0.005	6.797E-04	5.845E-02
L-Histidine	0.103	0.169	0.026	0.050	7.025E-04	6.041E-02
Erythrose-4-phosphate	0.016	0.026	0.006	0.009	1.521E-03	1.308E-01
Phosphoenolpyruvate	0.092	0.121	0.027	0.033	4.416E-03	3.798E-01
Deoxycytidine diphosphate (dCDP)	0.011	0.018	0.004	0.007	4.858E-03	4.178E-01
Guanine	0.149	0.289	0.090	0.122	5.829E-03	5.013E-01
Flavine adenosine diphosphate (FAD)	1.638	1.972	0.467	0.545	6.251E-03	5.376E-01
Glyceraldehyde-3-phosphate	38.494	45.115	7.056	10.780	7.584E-03	6.523E-01
Guanosine	0.026	0.035	0.004	0.009	8.097E-03	6.964E-01
L-Serine	0.033	0.051	0.005	0.017	8.744E-03	7.520E-01
O-Acetyl-L-serine	0.012	0.014	0.004	0.006	1.815E-02	1.561E+00
Inosine monophosphate (IMP)	1.964	1.883	0.864	0.498	2.253E-02	1.937E+00

Metabolites	Acetate phase (mM)	Ethanol phase (mM)	SD acetate	SD ethanol	p-values	Adjusted p- value
Pyruvic acid	0.137	0.132	0.067	0.042	5.472E-02	4.706E+00
Acetyl-CoA	6.287	10.832	3.436	7.902	8.020E-02	6.897E+00
L-Valine	0.600	0.892	0.138	0.577	9.033E-02	7.768E+00
L-Tryptophan	0.039	0.038	0.008	0.010	1.229E-01	1.057E+01
Inosine	0.027	0.019	0.009	0.006	1.419E-01	1.220E+01
Fumaric acid	0.012	0.015	0.004	0.012	1.693E-01	1.456E+01
Nicotinamide adenine dinucleotide	3.647	3.307	0.610	0.206	1.932E-01	1.662E+01
Oxalic acid	0.223	0.222	0.119	0.157	2.068E-01	1.778E+01
N-Acetylcysteine	0.016	0.662	0.006	1.709	3.167E-01	2.724E+01
Citraconic acid	0.004	0.057	0.002	0.147	3.261E-01	2.804E+01
Adenosine triphosphate (ATP)	0.246	0.202	0.129	0.168	3.572E-01	3.072E+01
Ketoglutaric acid (2-Oxoglutaric acid)	0.018	0.014	0.005	0.002	3.928E-01	3.378E+01
trehalose	0.006	0.009	0.003	0.008	4.949E-01	4.256E+01
Hypoxanthine	0.018	0.012	0.005	0.003	4.980E-01	4.283E+01
2-Isopropylmalic acid	0.020	0.018	0.004	0.004	5.745E-01	4.941E+01
NADP	1.497	1.643	1.503	0.815	6.270E-01	5.392E+01
Fructose 1,6-bisphosphate	0.342	0.228	0.098	0.032	6.398E-01	5.502E+01
Citric acid	0.058	0.066	0.045	0.018	6.801E-01	5.849E+01
Acetyl phosphate	1.777	1.342	0.965	1.573	8.026E-01	6.902E+01
Isocitric acid	0.087	0.112	0.146	0.081	8.237E-01	7.084E+01
Dihydrooorotic acid	0.286	0.093	0.455	0.218	8.426E-01	7.246E+01
Acetoacetic acid	0.534	0.409	0.142	0.064	8.998E-01	7.738E+01

4.3.2 Metabolic pathway analysis and mapping

4.3.2.1 Using Escher pathway analysis

Mapping the metabolites data on the metabolic pathway give in depth explanation to the above results. During acidogenesis the levels of citrate and cis-aconitate are high. This could be as a result of activation of the reversible reaction converting malate to oxaloacetate, going on to increase levels of cis-aconitate through isocitrate to produce carbon dioxide as seen in tricarboxylic acid (TCA) cycle. This explains why there are higher levels of citrate, malate and cis aconitate but no production of fumarate. The conversion of malate to oxaloacetate is highly feasible in the presence of NAD^+ which is reported higher in acidogenesis than solventogenesis making this reversible reaction feasible during acidogenesis. On the contrary, during solventogenesis when levels of NAD^+ reduces, conversion of oxaloacetic to malate acid is favoured and further converted to fumarate. Since the conversion of oxaloacetate to citrate is a non-reversible reaction, the citrate accumulated is utilised for the production of CO_2 , hence the observed lower levels of citrate during solventogenesis.

Further pathway analysis of the one carbon metabolic pathway (Figure 4.4) during solventogenesis, showed how the intermediates that led to ethanol, acetate and 23BD were affected. The main observation was that the concentration of pyruvate was comparatively high during solventogenesis. The concentration of acetylphosphate was lower than pyruvate but more than intermediates such as acetyl CoA. In order to confirm results from metabolomics data, data from transcriptomics analysis were mapped unto the one carbon pathway for comparison (Figure 4.5). Transcriptomic data confirmed the up-regulation of the enzyme responsible for the conversion of acetyl CoA to pyruvate; pyruvate ferredoxin oxidoreductase (PFOR). The enzymes responsible for the conversion of pyruvate to 23BD and lactate were down regulated, probably accounting for the accumulation of pyruvate. Transcriptomic data confirms the activation of the enzyme responsible for the conversion of acetate to acetaldehyde (acetaldehyde oxidoreductase, AOR), the acetaldehyde was then converted to ethanol by alcohol dehydrogenase (ADH)

enzyme. This means there are at least two pathways for the production of ethanol which were activated, hence the observed increase in ethanol concentration. The upregulation of AOR and ADH during autotrophic growth of *C. autoethanogenum* was also reported by (Marcellin et al., 2016b).

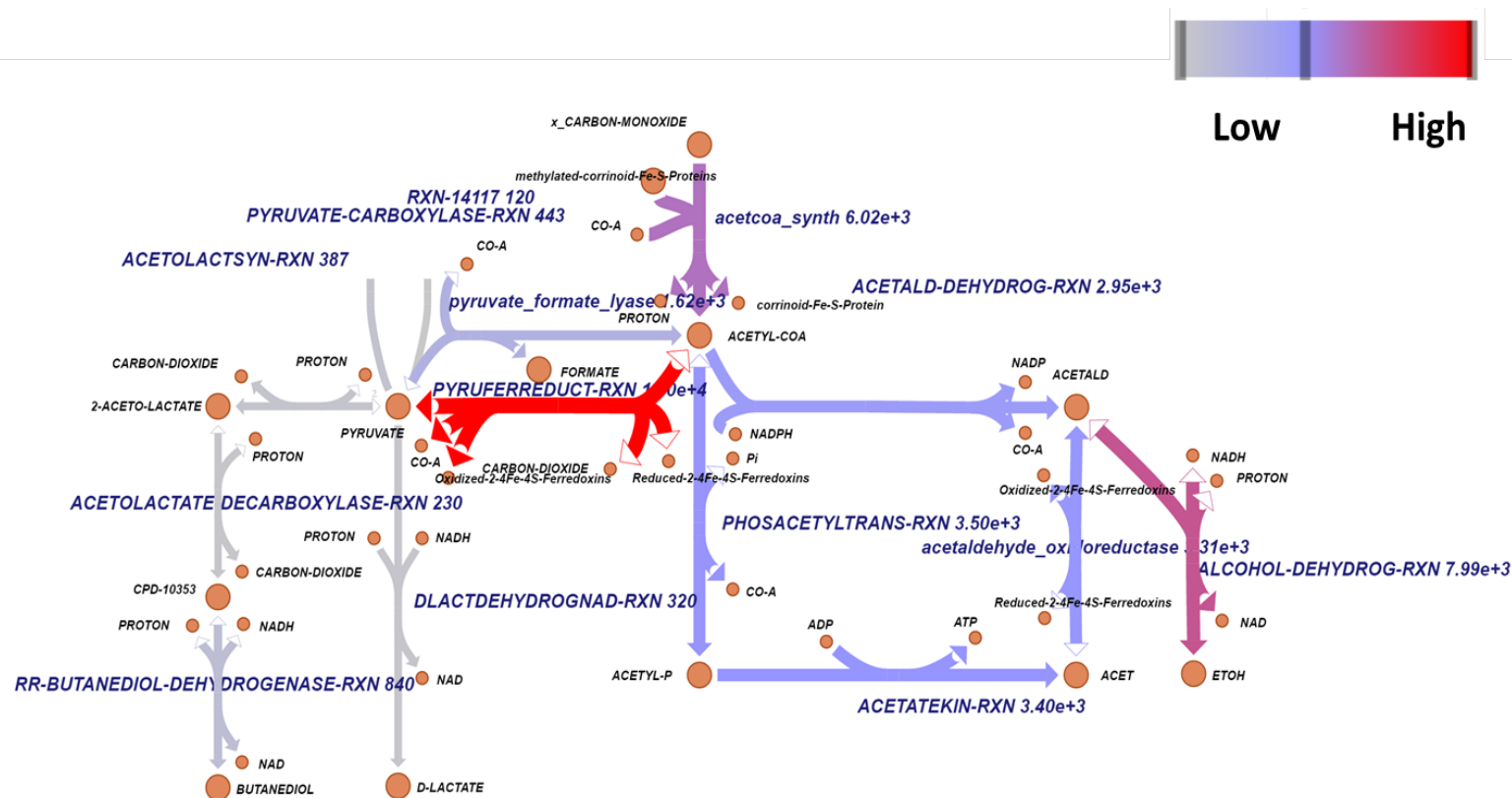


Figure 4.5 Mapping of transcriptomics data on the one carbon metabolic pathway during the gas shift experiment for the solventogenic phase using Escher online platform for pathway analysis. Colour code goes from blue for down regulated genes to red for up regulated genes. The arrows represent Gene data and the colour code shows an up-regulation or down-regulation

4.3.2.2 Using MetaboAnalyst 3.0

Metabolic pathway analysis of the identified metabolites using Metaboanalyst 3.0 was conducted to provide an in depth understanding of the effects of the acidogenesis-solventogenesis shift. The list of metabolic pathways of significance was listed in Table 4.2. Citric acid (TCA) cycle and alanine, aspartate and glutamate metabolism were two most significant pathways. The metabolic pathways of significance are indicated by their p-values <0.05 and false discovery rate (FDR) corrected p-values <0.05 . A total of 17 metabolic pathways were significant but reduced to 11 after the FDR correction.

It is well reported that the central metabolic pathways of clostridia remain unresolved fully. The TCA cycle in *C. acetobutylicum* is incomplete (Amador-Noguez et al., 2010) and the oxidative branch of the pentose phosphate pathway is absent (Crown et al., 2011). This is also the case for *C. autoethanogenum* and this makes it difficult to identify the enzymes responsible for the conversion of succinate to fumarate, oxaloacetate to citrate (citrate synthase) and succinyl-CoA to succinate. The enzyme responsible for the biosynthesis of α -ketoglutarate, which is a precursor for glutamate, proline and glutamine remains unknown. In an experiment conducted on *C. acetobutylicum* where the metabolites were monitored by growing the organism on ^{13}C -labelled glucose showed that at the solventogenic phase, the electron consuming TCA is shut down because there was no detection of labelling of fumarate or malate from the carbon source, but the oxidative part of the TCA cycle remained active. They also reported the excretion of malate and fumarate during acidogenesis and their re-assimilation during the transition into solventogenesis. This was confirmed by introducing ^{13}C -labelled fumarate during the transition and this was quickly converted into malate, which was metabolised through the TCA cycle and converted to aspartate or glutamate (Amador-Noguez et al., 2011). This was equally observed in this data presented for *C. autoethanogenum* with increased concentration of malate, aspartate and glutamate.

The comparison between acidogenesis-solventogenesis transition of *C. acetobutylicum* and *C. autoethanogenum* has showed that the TCA cycle plays

a major role in this transition. Although *C. acetobutylicum* was grown on glucose and *C. autoethanogenum* was grown on CO, the results are comparable because of the similarities between the two organisms although not completely the same in every respect.

The use of Escher provides mapping data for two data points. Escher has the limitation of mapping more than two data points. The use of Metaboanalyst 3.0 for metabolic pathway analysis gives list of pathways affected or of significance but no mapping details. Metabolic pathway analysis using biocyc database gives another level of visualisation of metabolomics data on specific pathways. Biocyc databases give direct access to pathways that are specific to the organism. The pathways are generated as collages and there is no limit on the data to map. In

Figure 4.6 shows pyruvate fermentation to acetate in (A) and mixed acid fermentation to acetate in (B). The pathway visualisation is depicted with five data points from the pH experiment with each box representing a specific condition. Visualisation of pyruvate fermentation to acetate and mixed fermentation pathways shows how the levels of pyruvate directly influences acetate production and the different way ethanol can be produced.

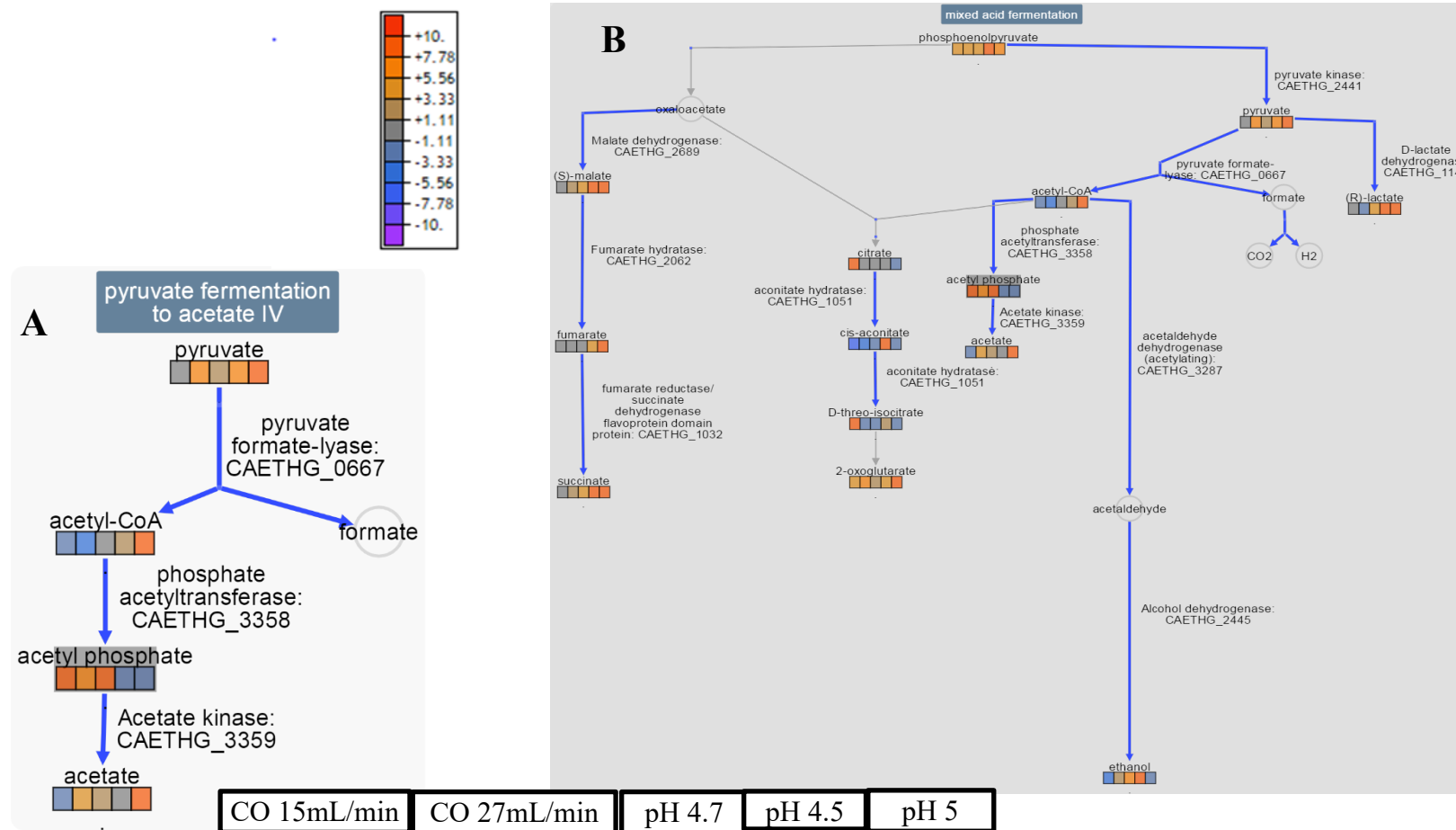


Figure 4.6 Metabolic pathway visualisation of *C. autoethanogenum* generated using biocyc.org. Each box represents a condition as stated in the boxes above. The colour coding goes from blue to red to indicate low and high levels respectively. A-Pyruvate fermentation to acetate IV and B-Mixed acid fermentation in *C. autoethanogenum*.

Table 4.2 List of metabolic pathways from analysis of list of intracellular metabolites using Metaboanalyst 3.0 including false discovery rate (FDR) and p values. Metabolic pathways of significance determined by using p values <0.05 together with their node of importance score. P values calculated by t-test. FDR-false discovery rate. Holm adjust- Holm-Bonferroni adjusted p-values (used to prevent errors associated with multiple comparisons).

Metabolic pathways	p-values	FDR	Impact
Citrate cycle (TCA cycle)	3.47E-05	0.00139	0.36183
Alanine, aspartate and glutamate metabolism	3.47E-05	0.00139	0.975
Arginine and proline metabolism	0.0001883	0.00495	0.51216
Aminoacyl-tRNA biosynthesis	0.0002474	0.00495	0.17857
Pyrimidine metabolism	0.0003927	0.00536	0.49364
Pyruvate metabolism	0.000402	0.00536	0.64528
Purine metabolism	0.0008467	0.00968	0.38866
Nitrogen metabolism	0.0011074	0.01107	0
Valine, leucine and isoleucine biosynthesis	0.002048	0.01821	0.1907
Phenylalanine, tyrosine and tryptophan biosynthesis	0.0032083	0.02333	0.07761
Glyoxylate and dicarboxylate metabolism	0.0032083	0.02333	0.17391
Glycine, serine and threonine metabolism	0.0137	0.09134	0.36
Taurine and hypotaurine metabolism	0.023024	0.13896	1
Phenylalanine metabolism	0.026055	0.13896	1
Novobiocin biosynthesis	0.026055	0.13896	1
Nicotinate and nicotinamide metabolism	0.029672	0.14836	0.15315
Cyanoamino acid metabolism	0.034291	0.16137	0
Butanoate metabolism	0.053104	0.23602	0.26829

4.3.3 Effect of lowering pH in the media on conversion of acetate to ethanol

The study involved pH shift of continuous batch cultures using CO as the sole C1 substrate. The culture was started CO flow rate of 15 mL/min and pH of 5. At these conditions, the levels of acetate were higher than ethanol levels. The accumulation of acetate was observed till day 11. An attempt to shift to solventogenic phase the CO flow rate was increased to 27 mL/min. This CO shift increased levels of both ethanol and acetate the same rate with further acetate accumulation. This can be characterised as early solventogenic phase. Optical density measured at 600 nm (OD_{600}) showed an increase of gas flow from 15 mL/min to 27 mL/min directly increased cell mass at pH 5 (Figure 4.7). The explanation for this is that at gas flow rate of 15 mL/min, carbon was the limiting nutrient, and this is rectified when the gas flow was increased to 27 mL/min at this point nitrogen was the limiting nutrient. The pH was decreased from 5 to 4.7 to increase ethanol production. As the pH was reduced, acid production decreased with increase in solvent (ethanol and 2,3BD) production due to re-assimilation of acids into solvents (mid solventogenic phase) (Amador-Noguez et al., 2011). A further drop in pH from 4.7 to 4.5 showed complete conversion of acetate to ethanol (Abubackar et al., 2011) as shown in Figure 4.8. The reduction of pH of the media also improved the yields of solvents such as 2,3-BD and lactate (Figure 4.9).

The conversion of acetate to ethanol observed by measuring extracellular content was equally backed by analysis of intracellular metabolites with in-depth interpretation to metabolic pathways affected.

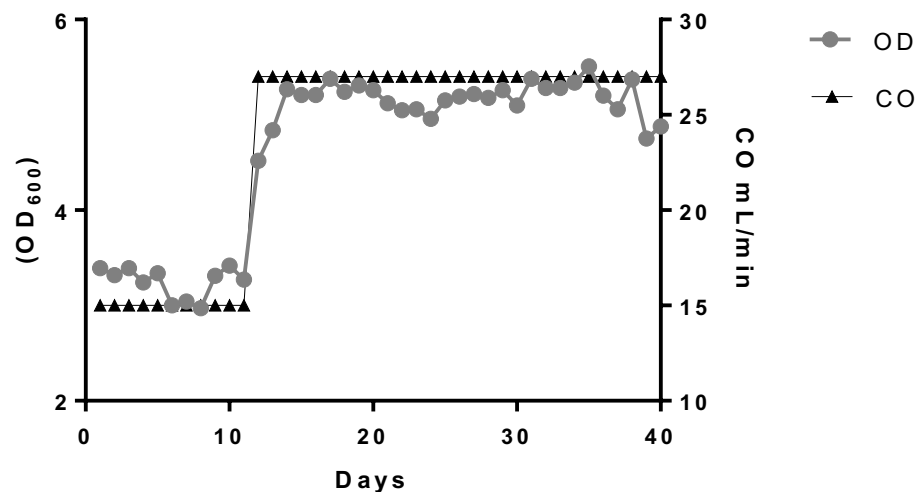


Figure 4.7 Optical Density measures at 600 nm (OD_{600}) in correlation with Carbon monoxide (CO) levels. An increase in gas flow from 15 ml/min to 27 ml/min showed a proportional increase in biomass. (black triangle-CO flow rate, grey circle – OD measured at 600 nm).

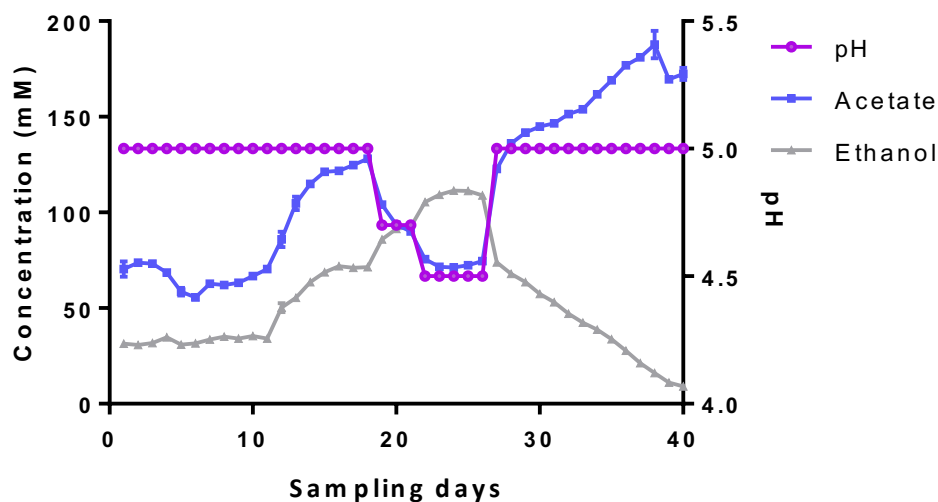


Figure 4.8 Ethanol and acetate production. Effect of pH shift on the conversion of acetate to ethanol. Higher concentrations of acetate were observed at pH 5 compared to ethanol. At pH 4.7 concentrations of acetate and ethanol reached equilibrium. The conversion of acetate to ethanol was observed at pH 4.5. (Purple- pH, blue-acetate, grey-ethanol). (pH, $n=1$)

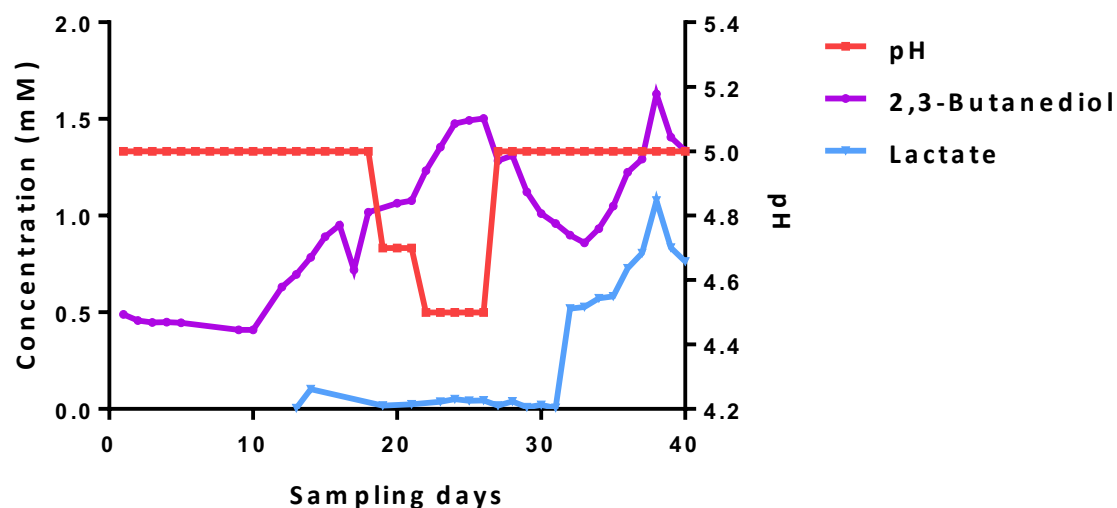


Figure 4.9 The effect of pH changes on extracellular concentrations of 2,3-butanediol and lactate. 23BD levels were at the highest during the solventogenic phase at pH 4.5 when the re-assimilation of acetate and ethanol occur. The levels of lactate seem to increase when the pH was increased back to when the re-assimilation of acetate had ceased. (red-pH, Purple-2,3-butanediol, blue-lactate)

4.4 Conclusion

The transition from acidogenic phase to solventogenic phase in *C. autoethanogenum* is an important aspect to be investigated to improve industrial application of the organism. In this experiment, metabolomics data obtained from pH shift experiment of *C. autoethanogenum* was used to understand the changes that occur during this transition. The results in this chapter have shown that reduction of the pH of the medium from 5 to 4.5 led to the re-assimilation of acetate for conversion into ethanol. Metabolomics data analysis showed intermediates from the TCA cycle changed the most during the acidogenic/solventogenic transition. Metabolic pathway analysis using list of metabolites identified and quantified, also confirmed the TCA cycle as the most

significant pathway affected including alanine, aspartate and glutamate metabolism pathway which is directly linked to TCA cycle.

Metabolomics analysis in this chapter has also been shown to give an insight into ways of enhancing acetyl-CoA production and the percentage of acetyl-CoA produced that gets converted into solvents. This chapter demonstrates the advantages of using targeted metabolomics analysis. There is currently substantial interest in improving solvent production by *C. autoethanogenum* by genetically engineering the organism to modify genes directly involved the acidogenic-solventogenic pathways. The research in this chapter suggests that the use of pathway-focussed metabolomics analysis in combination with gene data can provide the improved understanding required to increase solvent production in *C. autoethanogenum*.

I have successfully demonstrated that the HILIC -IDMS method is suitable for the analysis of a complex bioreactor samples. The fully validated method using IDMS used can provide the high-quality data needed to do pathway mapping. A fully quantitative method with intracellular concentrations is the best way forward for metabolomics analysis in targeted metabolomics studies. Also, the use of Clostridium-generated ^{13}C multiple IS metabolites is a major step forward in generating multiple internal standards for the range of metabolites investigated.

CHAPTER FIVE

**Application of HILIC-IDMS method to
determine the effect of increasing carbon
monoxide flow on ethanol production in *C.*
*autoethanogenum***

5. Application of HILIC-IDMS method to determine the effect of increasing carbon monoxide flow on ethanol production in *C. autoethanogenum*

5.1 Introduction

A new technological solution on a commercial scale that promises to reduce our dependence on fossil fuels, reduce the emission of greenhouse gases and route to capture CO₂ is gas fermentation (Daniell et al., 2012, Durre, 2017). *Clostridium autoethanogenum* is a well investigated organism in applications involving commercial gas fermentation for ethanol production derived from CO, H₂ and CO₂.

Although majority of microbes are capable of growth on syngas, the ethanologenic clostridial acetogens are of the main commercial interest. They include *C. autoethanogenum*, *C. ljungdahlii*, and *C. ragsdalei*. This is likely to be due to the long-standing industrial applications of these solventogenic *Clostridia species* for acetone, butanol and ethanol (ABE) from sugars (Bramono et al., 2011). *Clostridia species* are capable of using a wide spectrum of substrates for growth making them increasingly important in biotechnological applications. *C. ljungdahlii* and *C. autoethanogenum* are the most easily manipulated for synthetic biology approaches (Kopke et al., 2011b).

The biological conversion of waste gases containing carbon monoxide (CO) by acetogens provides a means by which waste gases can be efficiently utilised to produce valuable fuels such as ethanol, hydrogen and butanol (Mohammadi et al., 2011, Abubackar et al., 2011, Daniell et al., 2012, Durre, 2017, Molitor et al., 2016). One major bottleneck for the commercial application of this technique is the poor solubility of CO gas in aqueous medium. Therefore, in systems where CO is the only substrate, the mass transfer of CO gas and liquid at higher cell concentration during the bioconversion process is the main limitation. The process is kinetically limited when either the CO consumption rate or cell concentration is too low (Abubackar et al., 2011). The overall process yield is decreased by these rate limiting conditions which are often encountered at some point during the bioconversion process.

Homoacetogens capable of producing ethanol from CO include *Clostridium ragsdalei*, *Clostridium ljungdahlii* and *Clostridium autoethanogenum* among others (Liu et al., 2012, Mohammadi et al., 2011, Mohammadi et al., 2012a). In recent work involving biological conversion of CO was done using *C. autoethanogenum* as the biocatalyst (Abubackar et al., 2015c). The pathway used by *C. autoethanogenum* to produce ethanol, is the Wood-Ljungdahl pathway (WLP) or acetyl-CoA biochemical pathway (Kopke et al., 2011b). This pathway comprises of the eastern and western branches (Ragsdale, 1997). The eastern branch involves the reduction of carbon dioxide (CO₂) to form the methyl of acetyl-CoA. The western branch, which is peculiar to anaerobic bacteria generates either CO from CO₂ or directly utilise CO from the media that serves as the carbonyl group for the synthesis of acetyl-CoA (Abubackar et al., 2011). Acetogens use the WLP for acetyl-CoA production through the reduction of CO₂ since the WLP serves as an electron sink for electron donors and also linked to energy conservation needed for growth (Schuchmann and Muller, 2016).

5.1.1 Trends in omics studies in acidogenic-solventogenic transition in *Clostridia* species

The use of different omics approaches such as transcriptomics, proteomics and metabolomics are well reported for many clostridium species (Amador-Noguez et al., 2011, Marcellin et al., 2016b, Zhao et al., 2016). Metabolomics is a good complementary technique to support transcriptomics and proteomics studies since metabolites are representative of cellular activities that directly influence metabolite levels. There is little reported on the use of targeted metabolomics approaches to improve the understanding of metabolic shift in *C. autoethanogenum*. It is well reported that manipulation of the CO (gas) flow during syngas fermentation using *C. autoethanogenum* can increase the production of ethanol (Abubackar et al., 2015a, Cotter et al., 2009a). Analysis of the end products such as ethanol, acetate and 2,3BD are well reported (Abubackar et al., 2015c, Kopke et al., 2011a, Kopke et al., 2011b) but little is reported of the effects of gas flow as a means of inducing solventogenesis on

intracellular metabolites using metabolomics studies. Analysis of diverse range of metabolites using a single analytical method has advantages for application to the study of the metabolic effects of changes in gas flow. In this chapter a the novel HILIC-IDMS method developed in Chapter 3 using U-¹³C-labelled internal standard generated from *C. pasteurianum* will be used to report the effects of increasing gas flow on intracellular metabolites that lead to the formation acetate, ethanol and 2,3BD which has not been reported before for *C. autoethanogenum*.

5.1.2 Aims and Objectives

The purpose of this study is;

- To demonstrate the use of HILIC-IDMS for targeted metabolomics analysis of intracellular metabolites of *C. autoethanogenum* during a gas shift experiment conducted to optimise ethanol production in *C. autoethanogenum*.
- To use U-¹³C-labelled internal standard generated from *C. pasteurianum* for absolute quantification of intracellular metabolites in *C. autoethanogenum*.
- Use the generated metabolomics data for metabolic pathway analysis for further understanding of the effect of changing gas flow on *C. autoethanogenum*.

5.2 Materials and methods

5.2.1 Reagents and materials

All chemicals and reagents used here are listed in Chapter 3, section 3.2.1.

5.2.2 Culture conditions for *C. autoethanogenum* for gas shift experiment

The gas shift bioreactor experiments were performed by Dr Anne Henstra and Louise Sewell (Synthetic Biology Research Centre, (SBRC), Centre for Biomolecular Science, University of Nottingham).

This reactor was operated with the intent to obtain steady-states over three CO supply conditions (15, 30, and 40 mL/min CO) similar to the pH-shift as detailed in Chapter 4 section 4.2.2, except with pH kept constant (pH 5.0). Three bioreactors (B048, B049, B050) were started from a seed reactor with CO starter cultures. In this experiment the CO supply rate was increased in 5 mL/min steps after an initial step of 15 mL/min.

5.2.3 Sample collection and extraction of *C. autoethanogenum* cell pellets

Cells of *C. autoethanogenum* were collected and extracted as detailed in chapter 4 section 4.2.4. Timeline for sample collection, extraction, LC-MS analysis and data analysis are shown in Figure 5.1 below.

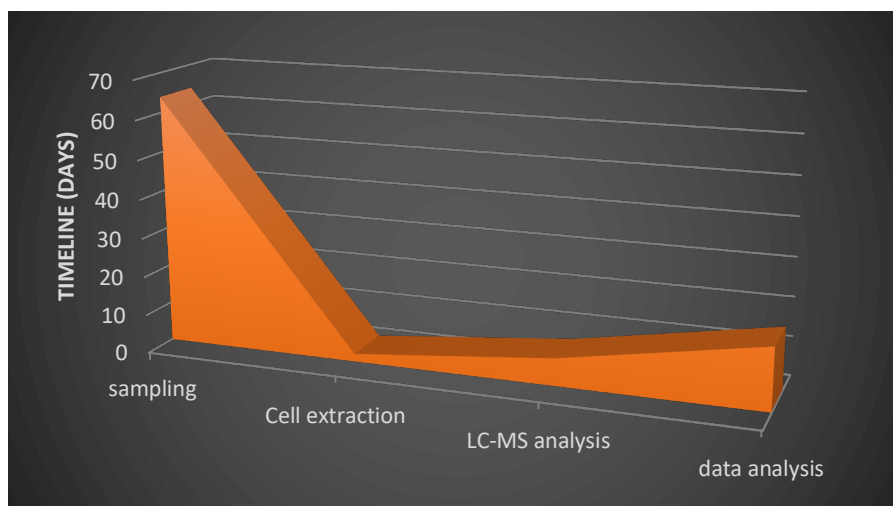


Figure 5.1 The timeline for sampling, extraction, LC-MS analysis and data analysis for gas shift experiment

5.2.4 LC-MS and data analysis

Metabolomics data analysis was done as previously described in Chapter 4 section 4.2.6.

5.2.5 Transcriptomics data analysis

Transcriptomics data analysis was conducted as described in Chapter 4 section 4.2.7.

5.2.6 Measurement of intracellular concentration of acetate and ethanol

Intracellular concentrations of acetate and ethanol phases were averaged over the number of days. Metabolite concentrations for acetate phase were an average of reading from day 1 to day 23 of the acidogenic phase with three replicates per day. Metabolites concentrations for the ethanol phase were an average from day 24 to day 62 with three replicates per day.

5.3 Results and discussion

5.3.1 Metabolomics analysis of cell pellets from gas shift experiment

Out of the 133 metabolites monitored using HILIC-IDMS, 106 metabolites were detected and quantified in all samples at the different time points. The concentrations of metabolites varied depending on the conditions applied and time of sampling. The absence of a particular metabolite may be as a result of the concentration being below the LOQ. Absolute quantification of metabolites was done by using U-¹³C-labelled internal standard generated from *C. pasteurianum*. U-¹³C-labelled internal standard were obtained for 100 metabolites, 2-fluorouridine was used as internal standard for the remaining metabolites. Also, quantified metabolites that were below the limit of quantification (LOD) were excluded.

In Table 5.1. the intracellular concentrations of metabolites present in *C. autoethanogenum* during the acidogenic and solventogenic phases were represented. The acidogenic-solventogenic transition in the gas shift experiments can be compared with pH shift experiment but there will be some differences because the solventogenic phase was induced differently.

Intermediates of TCA cycle such as malate and citrate increased during the acidogenic-solventogenic transition, but the reverse was the case for fumarate as seen in *C. acetobutylicum* during a pH induced acidogenic-solventogenic transition (Amador-Noguez et al., 2011). The concentration of acetyl-CoA increased during the solventogenic phase but levels of coenzyme A (CoA) and dephospho-CoA were lower. CoA and dephospho-CoA are significant to the pantothenate and CoA biosynthesis pathway (Gaddy et al., 2003, Toyota et al., 2016). The level of CoA production in the cell is dependent on levels of L-aspartate, which was found to be lower during solventogenesis hence confirming the lower levels of both dephospho-CoA and CoA. The decrease in glutamate, α -ketoglutarate and most amino acids (Reuse and Skouloubris, 2001) confirm the transition of the cultures from carbon-limitation to nitrogen limitation, which correlates, with the transition from acidogenesis to solventogenesis. The

concentration of most amino acids decreased during solventogenesis such as proline.

Initial t-test performed on the set of metabolites showed the changes observed were significant among 65 metabolites which reduce to 39 when the p-values were corrected using Bonferonni multiple correction test. The standard deviations were also computed for all metabolites to show reproducibility of results.

Table 5.1 Heat map of intracellular concentrations of metabolites in mM extracted from C. autoethanogenum to represent intracellular metabolite concentration during the acetate and ethanol phases. Red colour indicates lower concentration while green represents higher concentration. Acetate phase is an average of concentration from days 1-23 (n=69) and ethanol phase is an average of intracellular metabolites from days 24-62 (n=114). Standard deviation (SD) of both phases are shown. T-test was performed on the dataset and p-values are shown to be significant when p-value<0.05. Bonferroni correction was used to adjust p-values required in multiple comparison test.

Metabolites	Acetate phase (mM)	Ethanol phase(mM)	SD Acetate phase	SD Ethanol phase	P-values	Adjusted p- values
Hydroxyisocaproic acid	0.043	0.160	0.024	0.039	3.140E-21	3.328E-19
Acetoacetic acid	1.360	5.041	0.920	1.259	4.865E-19	5.157E-17
Mesaconic acid	0.270	1.000	0.207	0.188	1.961E-17	2.079E-15
2-Hydroxyglutaric acid	1.717	6.405	1.132	2.027	4.423E-17	4.688E-15
phenyllactic acid	0.012	0.043	0.010	0.008	1.575E-15	1.670E-13
Glyceric acid	0.107	0.291	0.061	0.054	5.092E-15	5.397E-13
Xanthosine	0.001	0.005	0.001	0.002	2.981E-13	3.160E-11
Malic acid	0.021	0.108	0.013	0.053	1.069E-12	1.133E-10
Uridine triphosphate (UTP)	0.024	0.056	0.010	0.019	5.183E-12	5.494E-10
Hydroxyphenyllactic acid	0.009	0.026	0.007	0.008	1.689E-11	1.791E-09
2-Isopropylmalic acid	0.011	0.029	0.007	0.009	2.007E-11	2.127E-09
Succinic acid	5.013	24.874	2.976	14.279	1.274E-10	1.350E-08
Glucosamine-6-phosphate	0.233	0.095	0.068	0.036	5.671E-10	6.011E-08
Phosphoenolpyruvate	0.092	0.187	0.050	0.028	1.730E-09	1.834E-07
Ribose-5-phosphate	0.501	0.953	0.238	0.105	2.794E-09	2.961E-07
3-Phosphoglycerate	0.521	1.070	0.305	0.244	7.615E-09	8.072E-07
Uridine 5'-monophosphate (UMP)	0.635	1.007	0.203	0.195	7.854E-09	8.325E-07
trans-Aconitic acid	0.002	0.010	0.001	0.007	9.943E-09	1.054E-06
Hydroxyphenylpyruvate	0.001	0.003	0.001	0.001	1.046E-08	1.109E-06
Thiamine diphosphate	0.072	0.095	0.013	0.009	1.267E-08	1.343E-06
Adenosine triphosphate (ATP)	0.378	0.960	0.261	0.482	6.679E-08	7.079E-06

Metabolites	Acetate phase (mM)	Ethanol phase(mM)	SD Acetate phase	SD Ethanol phase	P-values	Adjusted p- values
Citric acid	0.082	0.158	0.050	0.058	1.142E-06	1.211E-04
Isocitric acid	0.115	0.217	0.068	0.077	1.287E-06	1.365E-04
2-ketoisovaleric acid	0.003	0.005	0.001	0.001	2.045E-06	2.167E-04
L-Serine	0.048	0.085	0.025	0.029	2.691E-06	2.853E-04
N-Acetylornithine	0.048	0.019	0.023	0.005	3.558E-06	3.771E-04
Uridine diphosphate glucose (UDP-glucose)	0.169	0.240	0.056	0.025	3.724E-06	3.948E-04
NAD+	4.913	4.089	0.646	0.410	4.371E-06	4.633E-04
Adenine	0.016	0.022	0.005	0.005	1.367E-05	1.449E-03
N-acetylglutamic acid	0.582	0.379	0.106	0.229	1.517E-05	1.608E-03
sn-Glycerol-3-phosphate	0.284	0.385	0.089	0.035	2.153E-05	2.282E-03
Coenzyme A (CoA)	0.121	0.063	0.054	0.020	4.355E-05	4.617E-03
Dephospho-CoA	0.039	0.027	0.010	0.012	1.576E-04	1.671E-02
Deoxycytidine monophosphate (dCMP)	0.021	0.029	0.008	0.002	1.695E-04	1.796E-02
Dihydroxyacetone phosphate (DHAP)	0.215	0.121	0.061	0.124	1.783E-04	1.890E-02
Lactic acid	0.478	4.414	0.190	5.973	2.018E-04	2.139E-02
Sedoheptulose-7-phosphate	0.284	0.395	0.118	0.051	2.132E-04	2.260E-02
L-Methionine	0.065	0.049	0.017	0.007	2.217E-04	2.350E-02
Cytidine-diphosphate (CDP)	0.038	0.053	0.012	0.021	3.025E-04	3.206E-02
Deoxyadenosine diphosphate (dADP)	0.012	0.009	0.003	0.001	5.181E-04	5.492E-02
L-Histidine	0.128	0.195	0.068	0.075	6.443E-04	6.829E-02
L-Alanine	2.054	1.035	1.225	0.600	8.410E-04	8.915E-02
Uridine 5'-diphosphate (UDP)	0.733	0.888	0.190	0.108	1.163E-03	1.233E-01
Deoxycytidine diphosphate (dCDP)	0.004	0.002	0.002	0.001	1.299E-03	1.377E-01
L-Lysine	0.290	0.413	0.155	0.116	2.152E-03	2.281E-01
Flavin mononucleotide (FMN)	0.105	0.121	0.019	0.018	2.168E-03	2.298E-01
Acetyllysine	0.007	0.009	0.002	0.001	2.370E-03	2.512E-01

Metabolites	Acetate phase (mM)	Ethanol phase(mM)	SD Acetate phase	SD Ethanol phase	P-values	Adjusted p- values
NADH	0.913	1.483	0.740	0.528	2.616E-03	2.773E-01
Guanine	0.018	0.009	0.013	0.003	4.633E-03	4.911E-01
Cytidine monophosphate (CMP)	0.174	0.143	0.046	0.022	5.937E-03	6.293E-01
Inosine	0.027	0.021	0.009	0.006	6.184E-03	6.555E-01
N-acetyl-glucosamine-6-phosphate	0.376	0.309	0.069	0.119	6.838E-03	7.248E-01
Fructose 1,6-bisphosphate	0.466	0.318	0.219	0.153	7.329E-03	7.769E-01
D-Fructose-6-phosphate	1.066	1.265	0.301	0.214	8.764E-03	9.290E-01
Glycine	0.477	0.572	0.154	0.074	9.892E-03	1.049E+00
L-Ornithine	0.827	0.523	0.513	0.211	1.161E-02	1.230E+00
Adenosine monophosphate (AMP)	4.102	5.438	0.952	2.961	1.222E-02	1.296E+00
L-Tryptophan	0.156	0.091	0.118	0.030	1.658E-02	1.758E+00
Guanosine triphosphate (GTP)	0.080	0.129	0.079	0.068	1.658E-02	1.758E+00
UDP-N-acetylglucosamine	0.486	0.567	0.149	0.060	1.993E-02	2.112E+00
Hypoxanthine	0.003	0.002	0.002	0.001	2.156E-02	2.285E+00
Uracil	0.047	0.061	0.024	0.022	3.013E-02	3.193E+00
O-Acetyl-L-serine	0.011	0.029	0.017	0.044	3.167E-02	3.357E+00
Fumaric acid	0.010	0.086	0.006	0.220	3.641E-02	3.859E+00
Guanosine diphosphate (GDP)	0.603	0.522	0.169	0.110	4.802E-02	5.090E+00
trehalose	0.006	0.071	0.005	0.214	6.240E-02	6.614E+00
Erythrose-4-phosphate	0.032	0.035	0.009	0.005	6.803E-02	7.211E+00
L-Proline	0.541	0.636	0.247	0.066	8.062E-02	8.546E+00
Adenosine	0.048	0.061	0.031	0.024	9.607E-02	1.018E+01
Oxalic acid	0.031	0.150	0.017	0.452	1.084E-01	1.149E+01
Glyoxylic acid	0.019	0.096	0.008	0.295	1.099E-01	1.165E+01
Acetyl-CoA	0.872	1.309	1.027	1.067	1.167E-01	1.237E+01
L-Asparagine	0.291	0.220	0.189	0.143	1.267E-01	1.343E+01
Xanthine	0.580	0.684	0.271	0.235	1.319E-01	1.398E+01

Metabolites	Acetate phase (mM)	Ethanol phase(mM)	SD Acetate phase	SD Ethanol phase	P-values	Adjusted p- values
L-Glutamine	0.301	0.212	0.262	0.133	1.421E-01	1.506E+01
L-Valine	0.736	0.642	0.284	0.158	1.581E-01	1.676E+01
L-Glutamic acid	17.585	19.216	5.265	1.790	1.629E-01	1.727E+01
Thymine	0.008	0.026	0.004	0.079	1.632E-01	1.730E+01
2-Oxoglutaric acid	0.017	0.019	0.006	0.004	1.648E-01	1.747E+01
Guanosine	0.085	0.073	0.039	0.023	1.902E-01	2.017E+01
Glyceraldehyde-3-phosphate	16.792	18.486	4.091	6.938	2.313E-01	2.452E+01
NADPH	0.218	0.179	0.123	0.119	2.317E-01	2.456E+01
NADP	0.568	0.757	0.650	0.555	2.505E-01	2.655E+01
homocysteine	0.016	0.021	0.011	0.025	2.539E-01	2.692E+01
Orotic acid	0.014	0.006	0.037	0.007	3.158E-01	3.347E+01
L-Arginine	0.106	0.118	0.046	0.050	3.374E-01	3.577E+01
L-Citrulline	1.017	1.182	0.649	0.869	4.007E-01	4.247E+01
L-Leucine	0.124	0.140	0.034	0.118	4.454E-01	4.721E+01
L-Phenylalanine	0.340	0.378	0.222	0.205	5.023E-01	5.325E+01
Cytosine	0.031	0.028	0.020	0.015	5.041E-01	5.343E+01
N-Acetylcysteine	0.028	0.030	0.012	0.007	5.182E-01	5.493E+01
Pyroglutamic acid	0.070	0.082	0.026	0.115	5.229E-01	5.543E+01
Dihydroorotic acid	0.006	0.004	0.004	0.018	5.288E-01	5.606E+01
L-Threonine	0.800	0.834	0.250	0.074	5.397E-01	5.721E+01
L-Homoserine	0.089	0.092	0.022	0.008	5.412E-01	5.737E+01
Citraconic acid	0.049	0.022	0.202	0.040	5.419E-01	5.744E+01
Pyruvic acid	0.065	0.059	0.072	0.029	7.161E-01	7.591E+01
N-acetylalanine	0.041	0.042	0.014	0.020	7.587E-01	8.042E+01
Guanosine monophosphate (GMP)	0.042	0.041	0.029	0.014	8.380E-01	8.882E+01
Inosine monophosphate (IMP)	0.327	0.320	0.161	0.069	8.416E-01	8.921E+01
L-Aspartic acid	0.895	1.016	0.907	4.111	8.600E-01	9.116E+01

Metabolites	Acetate phase (mM)	Ethanol phase(mM)	SD Acetate phase	SD Ethanol phase	P-values	Adjusted p- values
L-Tyrosine	0.757	0.745	0.406	0.414	9.145E-01	9.694E+01
S-Adenosylmethionine (SAM)	0.430	0.432	0.094	0.041	9.327E-01	9.887E+01
N-acetylglucosamine	0.019	0.019	0.003	0.002	9.478E-01	1.005E+02
Flavine adenosine diphosphate (FAD)	0.117	0.118	0.031	0.014	9.488E-01	1.006E+02
Acetyl phosphate	1.086	1.086	0.524	0.783	9.991E-01	1.059E+02

5.3.2 Gas shift metabolomics data analysis using MetaboAnalyst 3.0

Further analysis of metabolites using MetaboAnalyst 3.0 was done to explore metabolites of significance. Steady states of acetate and ethanol were averaged into three replicates and submitted to MetaboAnalyst 3.0 for multivariate data analysis (Xia et al., 2015). PCA, PLSDA and OPLS-DA score plots were obtained for the acetate-ethanol transition showing separation between the two phases and clustering between each group (Figure 5.2). As part of the analysis, important features using VIP scores are generated from OPLS-DA plots for each metabolite. Metabolites with VIP scores >1.0 are said have a significant contribution the differences observed (Figure 5.3). The metabolites that showed the most significant change were malate, trans-aconitate and succinate, which are all intermediates of the TCA cycle. Dephospho-CoA and CoA were among the metabolites of significance, which suggests that pantothenate, and CoA biosynthesis are affected. This can be confirmed from metabolic pathway analysis. In Figure 5.3 only 20 metabolites were shown for convenience but a total of 64 metabolites had VIP scored >1.0 . The concentration of intracellular lactic acid increased during solventogenesis and this was also observed in extracellular concentration of lactate. Higher levels of lactate have been associated with cells under stress (Kopke et al., 2011b). A heat map generated from MetaboAnalyst 3.0 for 25 metabolites is shown in Figure 5.4 showing the orderly progression of changes observed during the acidogenic-solventogenic transition.

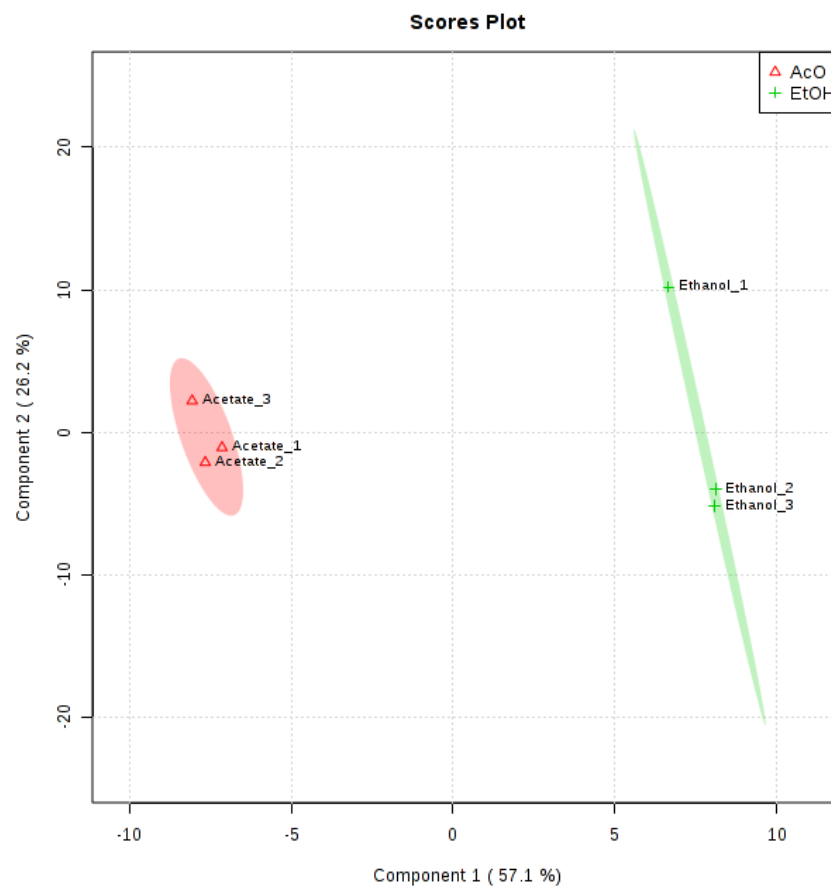


Figure 5.2 OPLSDA plots of acetate and ethanol phases of the gas shift experiment showing separation and clustering of groups. Red triangles represent acetate phase and green pluses represent ethanol phase. Each replicate is an average of samples collected over 24 days with three samples taken daily.

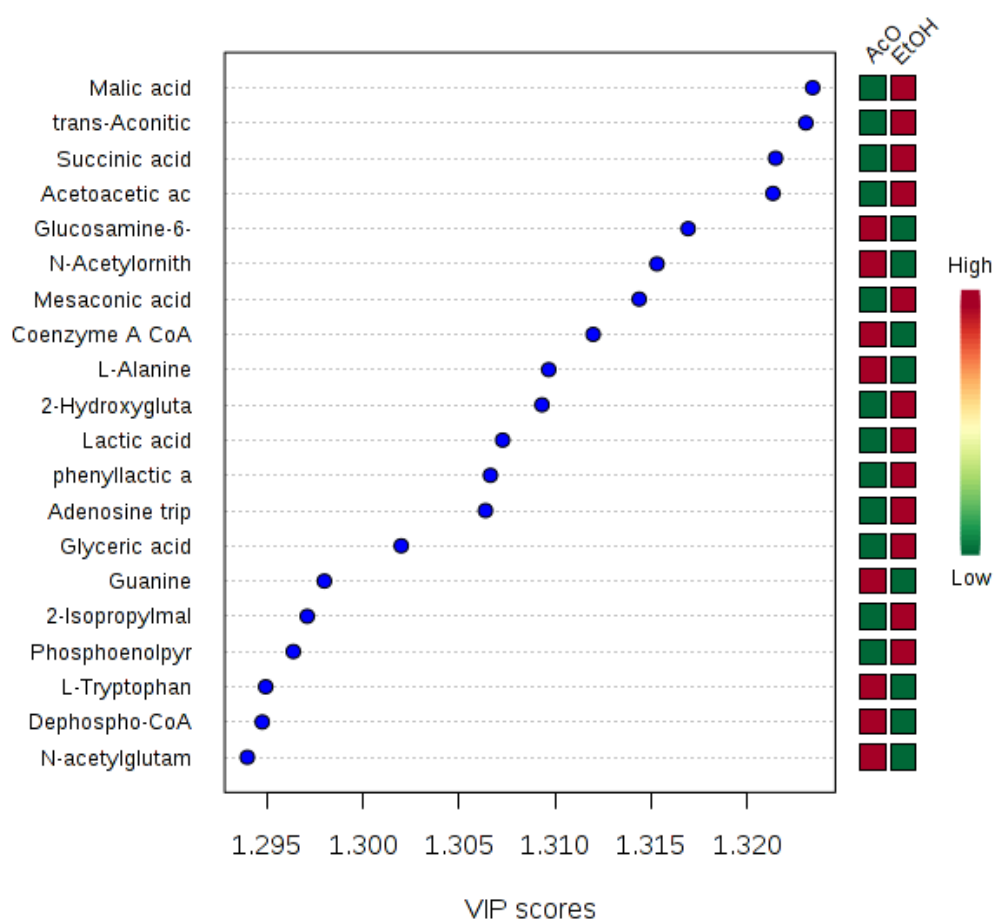


Figure 5.3 Variable Importance in Projection (VIP) scores of 20 metabolites represented in order of significance. On the top right is a mini heat map for all metabolites identified comparing the acetate and ethanol phases. Green colour indicates lower concentrations and high concentrations are coloured red. MetaboAnalyst only reports the first 14 characters of the compound name hence the loss in some of the details.

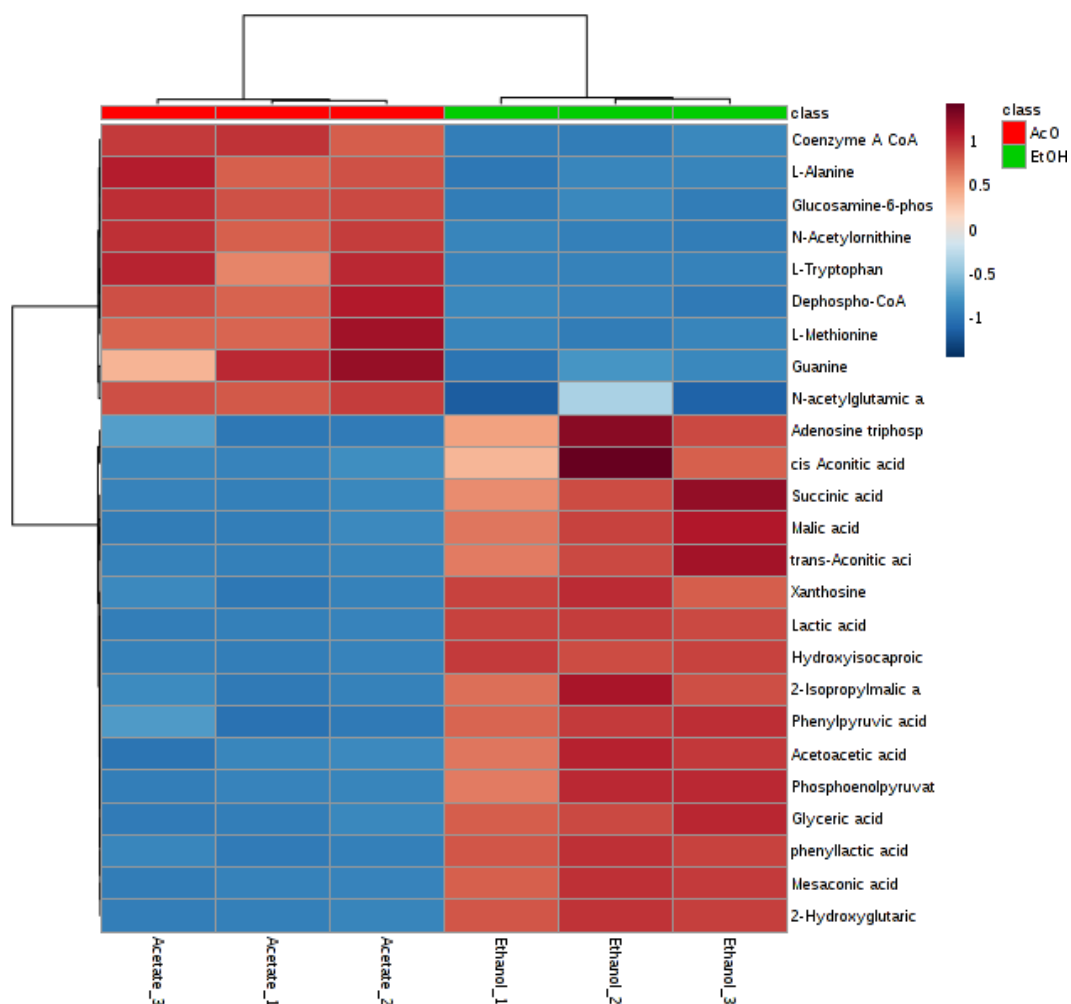


Figure 5.4 Heat map of 25 metabolites showing orderly progression of metabolites during the acidogenic-solventogenic transition in *C. autoethanogenum* obtained from statistical analysis done using MetaboAnalyst 3.0. Acetate phase is associated with the red class while green is for ethanol phase. Higher concentrations of metabolites are highlighted in red while lower concentrations are coloured in blue.

5.3.3 Metabolic pathway analysis of changes in the one carbon metabolic pathways during the gas shift experiment

Intracellular metabolites were mapped onto the one carbon pathway (Figure 5.5) and the levels of pyruvate lower in the solventogenic phase compared to that of the pH shift experiment. There is a higher accumulation of acetylphosphate compared to the pH shift experiment. The conversion of pyruvate to 2,3BD was up-regulated during the gas shift experiment, which was not the case for the pH induced solventogenic phase. Transcriptomics data analysis also confirmed the observation from metabolomics (Figure 5.6). The enzyme responsible for conversion of acetyl CoA to pyruvate (PFOR) was up-regulated as observed in the pH shift experiment but the pathways responsible for the conversion of pyruvate to 23BD were active which was not the case in the pH shift. This could account for the lower levels of pyruvate due to its conversion to 23BD. The pathway for conversion of acetate to acetaldehyde is active but conversion of acetyl CoA to acetaldehyde is inactive, hence production of ethanol is through one pathway. The higher levels of acetylphosphate observed can be attributed to the fact that there is no conversion of acetylphosphate to acetyl CoA to make acetaldehyde.

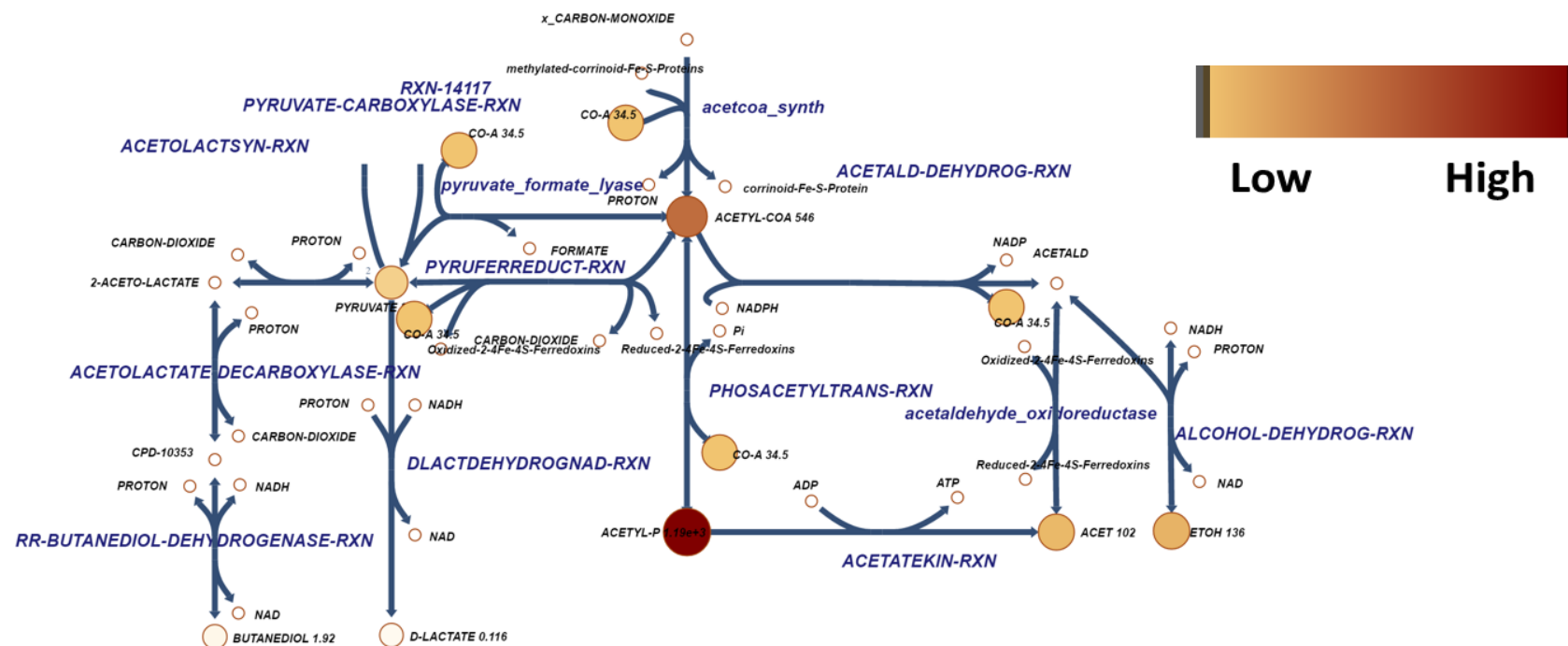


Figure 5.5 Mapping of intracellular metabolites on the one carbon metabolic pathway during the gas shift experiment for the solventogenic phase using Escher online platform for pathway analysis. Colour code goes from light brown for lower concentration to darker red for higher concentrations. Circles represent metabolites and the concentration of metabolites is indicated by size and colour. Small plain circles are unmapped metabolites.

5.3.4 Metabolic pathway analysis and mapping

The 64 metabolites with VIP scores >1.0 were submitted for metabolic pathway analysis. As would be expected TCA cycle was the most affected pathway as shown in Figure 5.7 and Table 5. 2. All metabolic pathways are of significance having p-values <0.05 . Only the three metabolic pathways stay significant after FDR correction. The metabolic pathway visualization shows a plot of the scores from enrichment analysis (y axis) and from topology analysis (x axis).

In comparison to the pH shift experiment reported in chapter 4, the TCA cycle was the most significantly affected metabolic pathway, which was also the case in *C. acetobutylicum* (Amador-Noguez et al., 2010, Crown et al., 2011). Another pathway that is directly dependent on the TCA cycle is the alanine, aspartate and glutamate (AAG) metabolism. The AAG metabolism pathway produces the aspartic acid required for the pantothenate CoA biosynthesis pathway. This is confirmed with metabolomics data presented above with decrease in concentration of L-Aspartate and CoA.

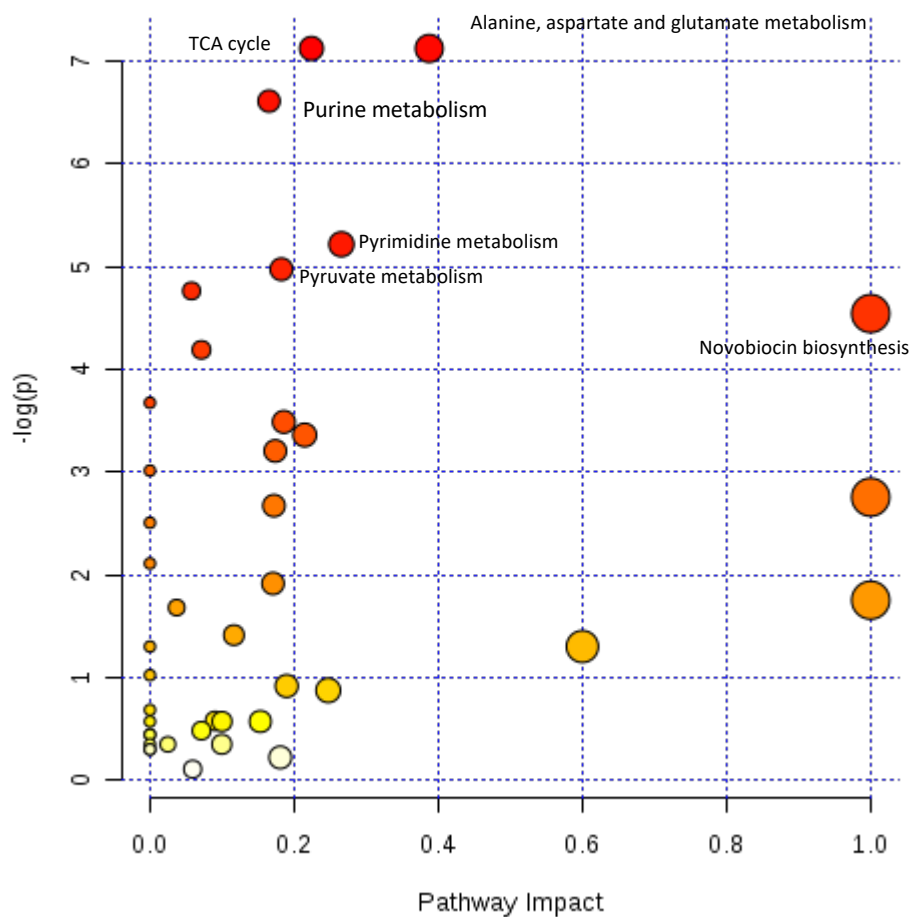


Figure 5.7 Metabolic pathway analysis and visualisation using MetaboAnalyst 3.0. This is the metabolome view showing all metabolic pathways arranged according to the scores from enrichment analysis (y axis) and from topology analysis (x axis).

Table 5.2 List of metabolic pathways in order of most affected during the acidogenic –solventogenic transition in the gas shift experiment conducted using *C. autoethanogenum*.

Metabolic pathways	p-values	FDR	Impact
Citrate cycle (TCA cycle)	0.001	0.032	0.224
Alanine, aspartate and glutamate metabolism	0.001	0.032	0.388
Purine metabolism	0.001	0.036	0.165

Pyrimidine metabolism	0.005	0.109	0.266
Pyruvate metabolism	0.007	0.111	0.182
Phenylalanine, tyrosine and tryptophan biosynthesis	0.009	0.114	0.058
Novobiocin biosynthesis	0.011	0.122	1.000
Aminoacyl-tRNA biosynthesis	0.015	0.152	0.071
Tyrosine metabolism	0.025	0.225	0.000
Arginine and proline metabolism	0.031	0.244	0.186
Pantothenate and CoA biosynthesis	0.035	0.252	0.215
Glyoxylate and dicarboxylate metabolism	0.041	0.270	0.174
Nitrogen metabolism	0.049	0.302	0.000

Metabolic pathways of significance determined by using p values <0.05 together with their node of importance score. P values calculated by t -test. FDR-false discovery rate. Holm adjust- Holm-Bonferroni adjusted p -values (used to prevent errors associated with multiple comparisons).

5.3.5 Using metabolomics to understand the effect of increasing gas flow on ethanol production

To elucidate the effect of the CO supply on culture growth and product formation, three independent gas-shift experiments were conducted in which the rate of CO inflow was changed in a stepwise manner. The CO flow rate was started at 15 mL/min, then increased to 30 mL/min followed by a series of increments of approximately 5 mL/min. Later analysis of the gas inflow revealed the following averaged supply rate for the three independent experiments, 15.4 mL/min, 30.5 mL/min, 34.2 mL/min, 39.4 mL, 44.2 mL/min, 50 mL/min and finally 54.3 mL/min. The external pH was kept constant at pH of 5 throughout the experiments.

Analysis of the gases in the headspace revealed that approximately 70 % of the CO supplied was utilised by the culture.

In Figure 5.8 it appeared that the cells shift from a carbon-limited growth phase to a non-carbon-limited growth phase around an external CO supply rate of 1.2 mol/(d·L). Nitrogen became growth limiting when gas flow rate was increased to 30 mL/min. There is a sharp increase in biomass production for the initial gas shift but the OD₆₀₀ remained nearly constant with gradual increment of gas flow rate.

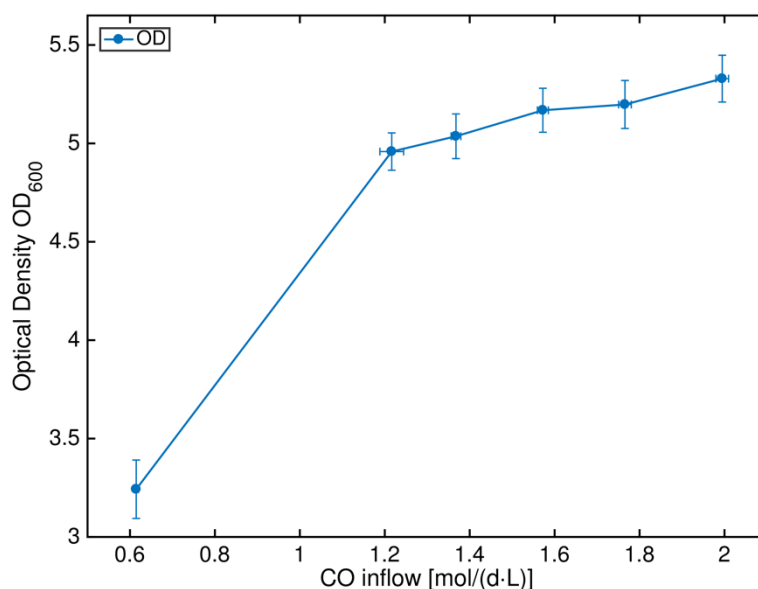


Figure 5.8 Optical density at steady state as a function of carbon monoxide supply rate. A shift from carbon-limited growth to non-carbon-limited growth occurred around a CO inflow of 1.2 mol/(d·L).

The experimentally observed product formation, see Figure 5.9, revealed that the transition between the two growth phases is accompanied by a metabolic shift from acidogenesis to solventogenesis. With the conditions investigated, acetate and ethanol were the major liquid products. During carbon-limited growth, their formation increased with increasing carbon monoxide supply, but their yields were little sensitive to the supply. The latter finding agrees with the results of a previous study reported by (Cotter et al., 2009b). This changed dramatically, when the culture entered the non-carbon-limited growth phase. Whereas ethanol formation rises further, the production of acetate drops. As a consequence, dominant liquid product and the ratio of ethanol and acetate completely reverses from one third to three which is in contrast to the conclusion

drawn in that *C. autoethanogenum* cannot exceed an equal flux distribution with respect to these two products.

In addition, 2,3-BD and lactate were formed in minor but increasing amounts. Whereas lactate could not be detected in the medium during the carbon-limited growth phase, its rising concentration during the second growth phase might indicate increased cellular stress due to the increased carbon monoxide levels.

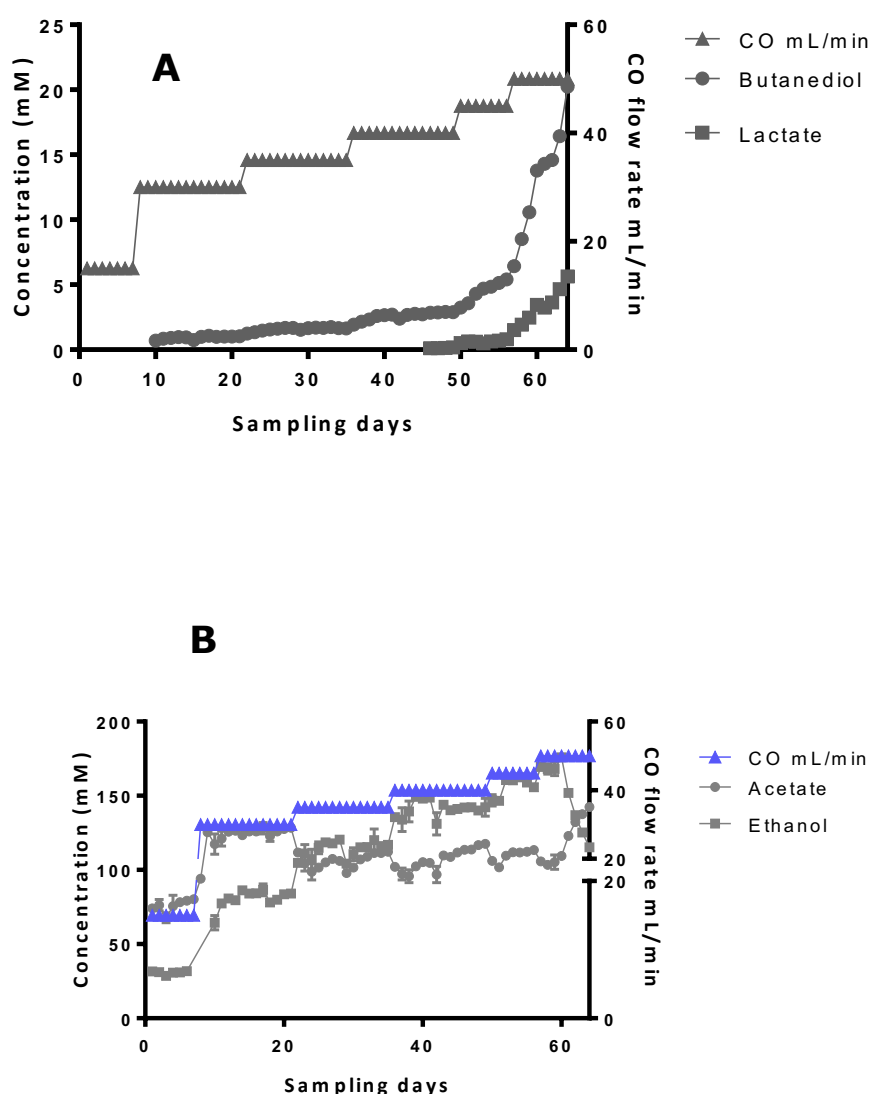


Figure 5.9 A- shows production of ethanol and acetate ad function of carbon monoxide (CO) uptake. B- shows the production of 2,3-butanediol (2,3BD) and lactate as a function of the CO uptake. Whereas acetate and ethanol production is increasing, with acetate as the major product, under carbon-limited conditions, formation of ethanol further rises as that of acetate drops during

non-carbon limited conditions. Furthermore, the concentration of the minor products 2,3-butanediol and lactate increase with increasing carbon monoxide supply, however lactate could be detected under non-carbon-limited conditions only.

Metabolomics data indicated possible limitation of pantothenate in the gas-shift study during the nitrogen limitation phase. Pantothenate is a direct precursor of coenzyme A (CoA), thus of key importance for production of acetyl-CoA through the WLP. Metabolomics data shown in Figure 5.10 confirms decrease in CoA concentration during the solventogenic phase where nitrogen was limiting. Pathway analysis also confirms the pantothenate and CoA metabolism (Table 5.2) pathway as one of the pathways significantly affected during the acidogenic-solventogenic transition. A limited pantothenate supply may affect the cell's ability to recruit CoA effectively in reactions that use this cofactor, thus flux towards acetyl-CoA and subsequently acetate is possibly affected.

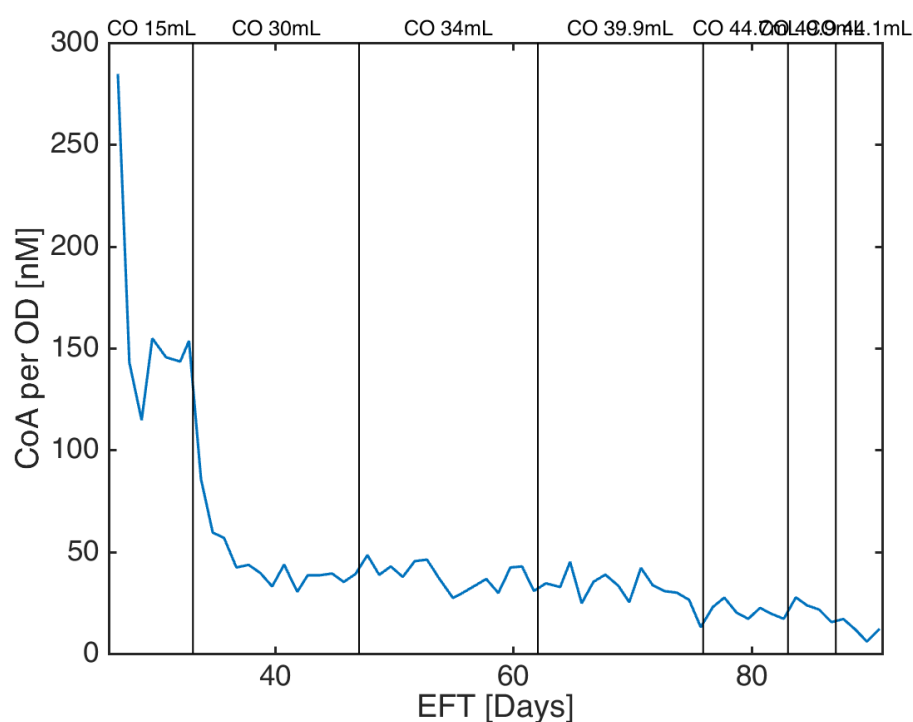


*Figure 5.10 Intracellular concentration of CoA normalised to OD during the gas shift experiment using *C. autoethanogenum*. There is a decrease in CoA concentration during the non-carbon limiting phase.*

The use of metabolomics data can improve the knowledge of the transitions that occur in the organism. As shown in Figure 5.11, intracellular concentrations of metabolites were mapped onto the central metabolic network to show the concentrations of metabolites during the different limitation phases. When carbon was no longer the limiting nutrient the concentrations of most metabolites were higher than the carbon limiting phase except for acetyl-phosphate which has already been discussed in previous sections.

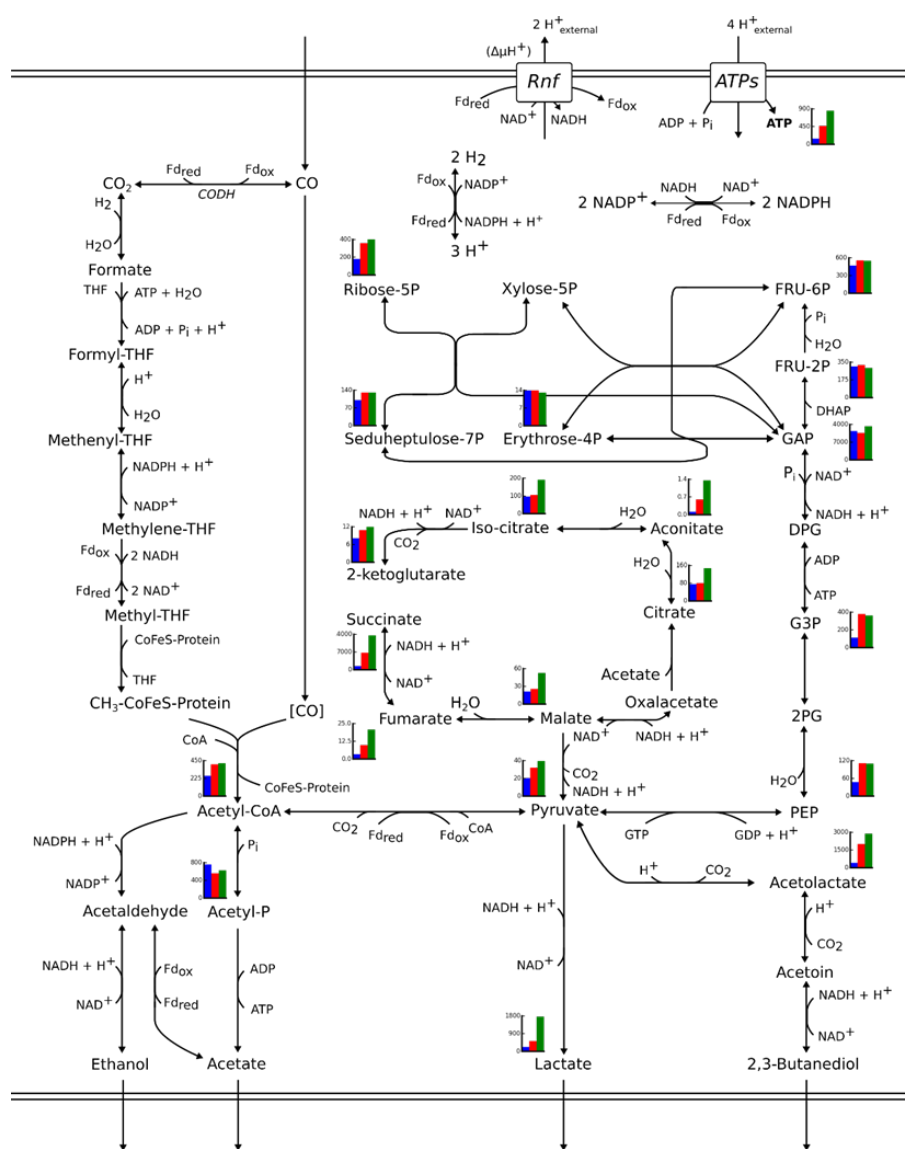


Figure 5.11 The central metabolic network of gas fermentation in *C. autoethanogenum*. The bar graphs represent the metabolite concentrations at days 15, 34 and 44 (carbon limitation, transition, non-carbon limitation).

(Designed by Dr Thomas Millat and Rupert Norman with both intracellular and extracellular concentrations of metabolites)

5.3.6 Effects of gas shift in *C. autoethanogenum* on folates (Wood-Ljungdahl pathway intermediates)

As discussed above the Wood-Ljungdahl pathway (WLP) is an important pathway to consider for gas fermentation in acetogens. The WLP has shown in Figure 1.1 comprises of four folate intermediates that are crucial for the one carbon metabolism. Cells assimilate CO₂ which is reduced to form formic acid. The formic acid then reacts with tetrahydrofolate (H₄folate) to produce formyltetrahydrofolate. Formyl tetrahydrofolate is then reduced to 5,10-Methenyltetrahydrofolate. 5, 10-Methenyltetrahydrofolate undergoes further reduction to form 5,10-Methylenetetrahydrofolate which is further reduced to produce 5-methyltetrahydrofolate. The methyl group on 5-Methyltetrahydrofolate is cleaved off to form the methyl part of acetyl-CoA and the h₄folate starts the cycle again binding to the formic acid formed. There are two forms of folates present in the cells, folates with one glutamic acid residue (monoglutamated) and folates with more than one glutamic acid residues (polyglutamated).

Analysis of the polyglutamated chains can be complicated especially with increasing glutamic acid residues. They also difficult to quantify because of the lack of standard mixture used to obtain calibration curves. The only way around this is to measure the polyglutamated folates as total folates by removing the glutamic acid residues and measuring it as monoglutamated folates. Gamma-glutamyltransferase is an enzyme that catalyses the transfer of gamma glutamic acid residues in the compound. Rat serum is known to possess this enzyme and there was used in our method to remove the glutamic acid residues on the polyglutamated folates.

In Figure 5.12 shows intracellular folate concentrations during the gas shift experiment represented for acetate, ethanol and 2,3-butanediol phases. Concentration of the total folates were higher than the monoglutamated folates. The polyglutamated folates are broken down into the monoglutamated forms for absorption into the gastro-intestinal tract in mammal (Garratt et al., 2005). The polyglutamated folates are known to be the storage forms of the folates that are subsequently converted to other forms of folates depending on the requirements of the cell. Therefore, there is the expectation of having higher levels of the polyglutamated folates in the cell in comparison to the monoglutamated forms.

Quantification of folates in microorganisms is less reported due to the difficulty and sensitivity problems that exist of this class of compounds. Isotope ratio-based profiling of folates were reported for *Escherichia coli* and *Saccharomyces cerevisiae* by (Lu et al., 2007) but the method quantified 5-Methyltetrahydrofolate and 5/10-Formyltetrahydrofolate suggesting the 5-Methyltetrahydrofolate with three glutamic acid residues was the most abundant in *E. coli*. The lack of authentic standards make them semi-quantitative and concentrations normalised with cell dry weight (CDW) which has its disadvantages. The process of de-glutamation provides the monoglutamated form of the folate that can be referenced against an authentic standard and calibration given confidence on the reported concentrations.

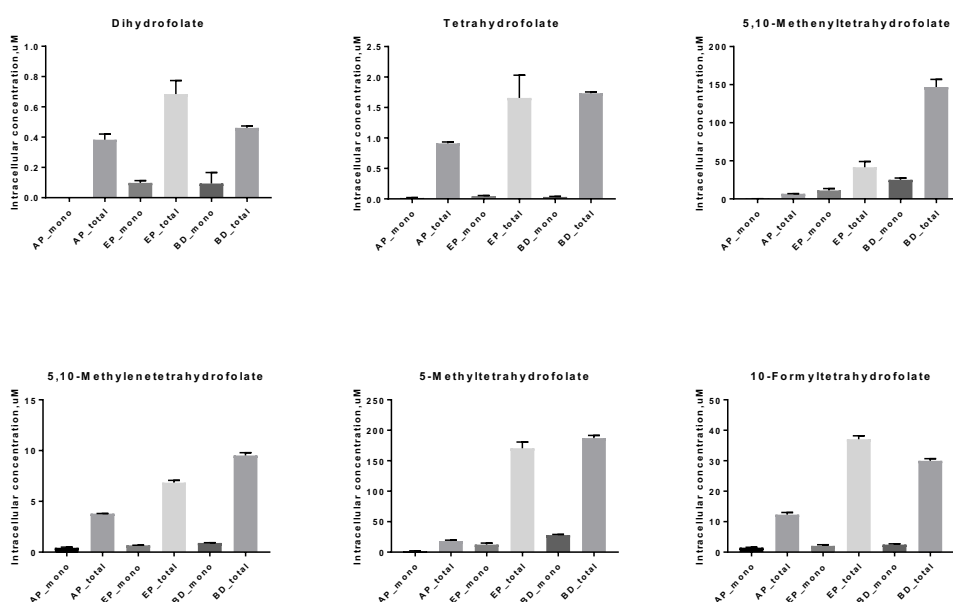


Figure 5.12 Plots of intracellular concentrations of folates during the three steady states of the gas shift experiment. Intracellular concentration of monoglutamated folates and total folates were evaluated for the acetate, ethanol and 2,3-butanediol phases. (AP_mono- Acetate phase monoglutamated folate, AP_total - Acetate phase total folates, EP_mono- Ethanol phase monoglutamated, EP_total- Ethanol phase total folates, BD_mono- 2,3-butanediol phase monoglutamated, BD_total- 2,3-butanediol phase total folates).

5.3.7 Statistical and pathway analysis for intracellular folate concentration in *C. autoethanogenum* during gas shift experiment

Statistical analysis conducted using MetaboAnalyst to compare the acetate phase and ethanol phase during the gas shift experiment showed the intensities of the folates at the different phases. Metabolites of significance were indicated by their VIP scores as shown in Figure 5.13. Four out of the seven folates had VIP scores >1 to indicate their significant contribution to the changes observed.

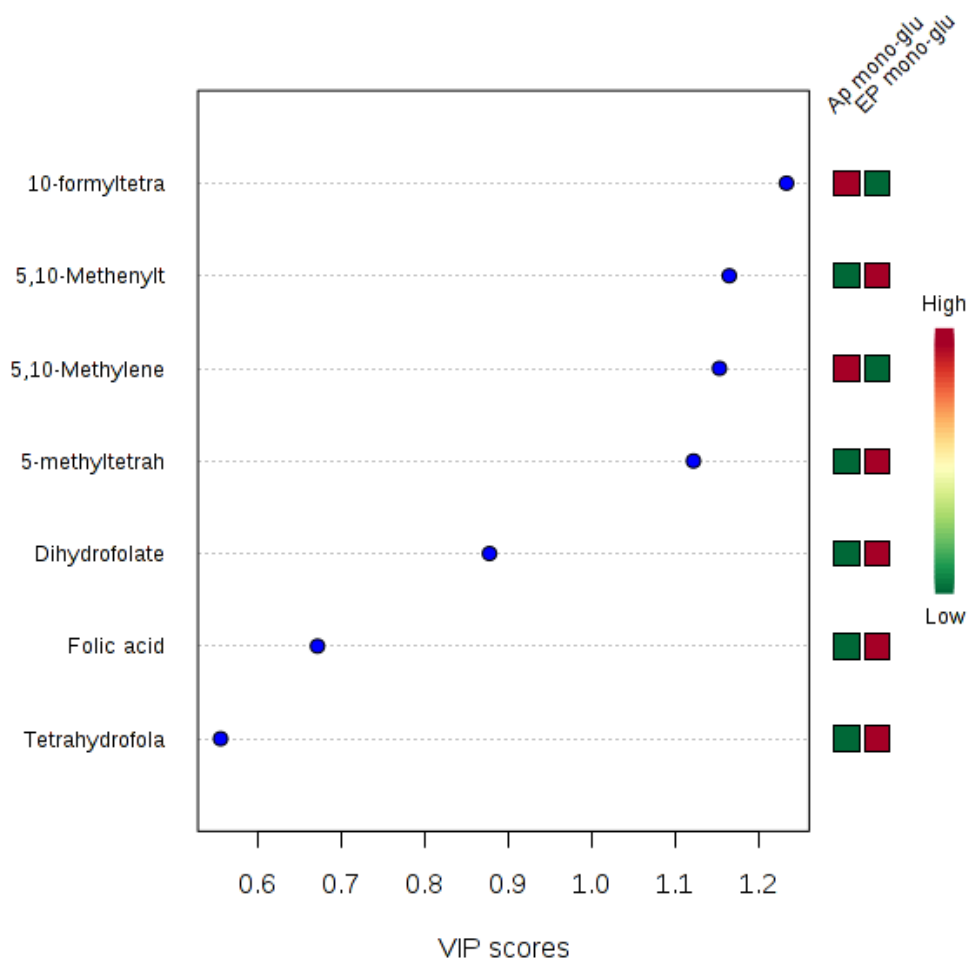


Figure 5.13 List of metabolites that contributed significantly to the changes observed in the folate analysis during the gas shift experiment. Metabolites with VIP scores > 1 were chosen to determine metabolites of significant contribution. The metabolites are in descending order of importance. The coloured boxes on the right indicate absolute concentrations of the corresponding metabolites for each group, red means high and green means low. (AP_mono-glu- Acetate phase monoglutamated folate, EP_mono-glu - Ethanol phase monoglutamated)

The four most significant metabolites identified using VIP scores > 1 were submitted for pathway analysis using MetaboAnalyst 3.0. The list of significant pathways are listed in Table 5.3 along with the p-values and adjusted p-values. Two pathways were significant after adjusting the p-values. The first is the one carbon pool by folates which is central to folate biosynthesis. The second pathway was Glyoxylate and dicarboxylate metabolism (Figure 5.14) which is

directly linked to Alanine, aspartate and glutamate metabolism, TCA cycle and pyruvate metabolism. These pathways were also confirmed to be affected during the acidogenic-solventogenic transition reported in 5.3.3 with the first being significant. This confirms the role of the folates in the acidogenic-solventogenic transition through the Wood-ljungdahl pathway which is observed in the lower levels of the total folates during acetate and ethanol phases. The lower levels of total folates in the acetate and ethanol phases confirm the conversion of the polyglutamates to generate monoglutamated folates required for acetate and ethanol production. During the 2-,3-butanediol phase where cells crashed, higher levels of the total forms of folates were observed indicating slower conversion of polyglutamated folates to monoglutamated folates.

Table 5.3 List of metabolic pathways obtained from MetaboAnalyst 3.0 using metabolite list obtained from folate analysis of the gas shift experiment.

Metabolic pathways	p-values	Holm adjust	FDR	Impact
One carbon pool by folate	1.04E-06	8.35E-05	8.35E-05	0.65625
Glyoxylate and dicarboxylate metabolism	4.54E-05	0.003586	0.001816	0.28261
Methane metabolism	0.054839	1	1	0
Glycine, serine and threonine metabolism	0.11528	1	1	0.03165
Aminoacyl-tRNA biosynthesis	0.25542	1	1	0.00893

5.4 Conclusion

Application of the HILIC-IDMS method to complex bioreactor samples during the gas shift experiment using *C. autoethanogenum* was shown to be successful. The intracellular concentration obtained showed TCA intermediates were most significantly affected. The reported concentrations of some metabolites such as amino acids and CoA were very significantly changed and indicated other effects that could be confirmed through pathway analysis. Metabolic pathway analysis done using MetaboAnalyst 3.0 showed the TCA cycle was the most significant pathway affected followed by associated pathways such as alanine, aspartate and glutamate metabolism.

A gradual increment of the gas flow enabled a metabolic shift from acetate to ethanol, which is characterised as the acidogenic-solventogenic transition. The TCA cycle has been reported to play a significant role in the acidogenic-solventogenic transition regardless of the conditions applied which was confirmed by metabolomics studies.

Metabolomics data generated here complements transcriptomics data to provide an in-depth understanding of intracellular metabolites and their effects on specific pathways.

CHAPTER SIX

**Application of HILIC-IDMS method to study the
effects of pantothenate limitation on the
metabolism of *Clostridium autoethanogenum***

6. Application of HILIC-IDMS method to study the effects of pantothenate limitation on the metabolism of *Clostridium autoethanogenum*

6.1 Introduction

6.1.1 Pantothenate and Phosphate limitation in *C. autoethanogenum*

Nutrient limitation in *Clostridia* *sp* is a common approach to induce solvent product instead of acid production. Phosphorus is an essential nutrient for bacteria but is usually scarce. Introduction of phosphate into the external media is the only source of phosphorus to the bacteria (Romano et al., 2017). Phosphate limitation has been reported to improve acetone, butanol and ethanol (solvents) production in *C. acetobutylicum* (Bahl et al., 1982, Millat et al., 2013). The fermentation of Clostridia in continuous cultures is characterised into two distinct metabolic phases, which is dependent on the substrate used. The two main metabolic states are acidogenesis where higher levels of acid production are observed and solventogenesis where higher levels of solvents are observed.

Pantothenic acid is an essential vitamin that is required for the synthesis of Coenzyme A (CoA). It is also responsible for the synthesis and metabolism of proteins, fats and carbohydrates. The pantothenate and CoA biosynthesis pathway is the pathway for production of CoA from aspartate. This pathway is directly associated with glycolysis which produces the pyruvate and alanine, aspartate and glutamate metabolism pathway which, in turn, produces the aspartate for pantothenate synthesis and eventually CoA.

The addition of pantothenate as a vitamin in the growth medium is essential for the growth and viability of *C. autoethanogenum* cells. Slower growth rates or fewer colonies on plates have been associated with pantothenate concentrations below 10% of the standard addition (50 mg/L) and dilution to 1% showed limited growth. As previously, stated acetogens tend to favour the production of ethanol over acetate in reduced growth conditions. Limitation of pantothenate may well reduce the organism's ability to perform reactions requiring CoA, affecting the flux towards acetyl-CoA and hence having a direct effect on the production of

acetate. Other studies have confirmed this effect on acetate and demonstrated a shift to higher ethanol production (Valgepea et al.). It seems reasonable in this case that a decrease in available pantothenate may force this limitation of acetyl-CoA to occur earlier.

Attempts have been made to change the natural product ratios by omitting yeast extract and B-vitamin within the media and, although these changes reduced cell growth, there was no observed increase in ethanol production (Phillips et al., 1994). The acetate-ethanol ratios in the bioreactor have been reported to increase when the concentration of yeast in the media is higher and this could possibly be due the higher levels of pantothenate which, in turn, increased the levels of acetyl-CoA within the cell (Klasson et al., 1993). This also led to a diversion of reducing potentials away from ethanol production and towards biomass or other growth-related products (Richter et al., 2001, Abubackar et al., 2012a, Abubackar et al., 2015b). Although measurement of the end product (ethanol) can give information of the effects of changing pantothenate concentration in the media, an understanding of the specific changes to individual intracellular metabolite levels in the relevant pathways would give a better understanding of the mechanism the organism uses to produce the end products.

A method reporting effects of limiting calcium pantothenate in *Clostridium ljungdahlii* has been described (Gaddy et al., 2003). Acetyl CoA is derived from calcium pantothenate and limiting concentration levels of pantothenate in the media were reported to reduce acetyl-CoA formation in relation to the rate of oxidation of CO in the Wood-Ljungdahl pathway. This results in build-up of reduced ferredoxin which leads to NADPH regeneration (Kundiyana et al., 2011).

6.1.2 The use of metabolomics approaches in understanding the role of pantothenate in ethanol production

The mechanism by which biological systems such as bacteria respond to physical and chemical stimuli is of importance. The advantages of metabolomics to determine a link between changes in environmental stimuli and intracellular metabolites is relevant to nutrient limitation studies, such as the studies proposed here on pantothenate. Two approaches using metabolomics analysis are suitable, a targeted metabolomics analysis where the metabolites of interest are known prior to analysis and an untargeted metabolomics approach where there is no prior knowledge of the metabolites to be analysed (Begou et al., 2017). A combination of both methods on the same sample will give us an opportunity to compare the different types of metabolic information on both specific metabolic pathways and global metabolism as a result of changes in pantothenate concentrations.

An example of metabolomics application in pantothenate measurements in water flea has been reported (Toyota et al., 2016) and will be followed closely to compare results

6.1.3 Aims and Objectives

The overall aim was to study the effect of pantothenate and phosphate limitation on ethanol production using the previously developed targeted LC-IDMS method (Chapter 3) combined with an untargeted metabolomics approaches as a means of comparison.

The objectives of this chapter are:

- To investigate the effects of pantothenate and phosphate limitation on intracellular metabolism of *C. autoethanogenum* during fermentation.
- To compare the results obtained from targeted and untargeted metabolomics for each experiment.

- To use the results from metabolomics analysis to investigate metabolic pathways that were directly or indirectly affected by the pantothenate and phosphate limitation.

6.2 Materials and Methods

6.2.1 Bioreactor operation parameters

All bioreactor experiments and preparation of growth media in this chapter were performed by Dr Anne Henstra and Louise Sewell of the Synthetic Biology Research Centre (SBRC), Centre for Biomolecular Science, University of Nottingham. A gas composition of approximately 15 mL/min of CO and 80 mL/min of nitrogen was added throughout the experiment. The bioreactor used in this project was a New Brunswick Scientific BioFlo 115 bioreactor system (Eppendorf) and BioCommand Batch Control software (Eppendorf). The total capacity of the bioreactor was 3 L, and the *C. autoethanogenum* was grown in 1.5 L of medium. A diagram of the reactor set up is shown in Figure 6.2. A constant temperature of 37°C was maintained in the bioreactor by a water jacket. The continuously stirred tank reactor was fitted with Bioflo115 baffles, Rushton impellers and open-pipe sparger. An open pipe sparger was used and adapted from a standard BioFlo115 micro-sparger. Gases were supplied through thermal mass flow controllers (SmartTrak C100, Sierra Instruments). A standard BioFlo115 head plate with 13 ports to allow the addition of media, gases, antifoam, Sodium hydroxide (NaOH), sodium sulfide (Na₂S), and to allow the removal of waste gases and media, as well as access for sampling media, pH and redox probes and for the addition of supplements through a sterile port using needles and syringes. A peristaltic pump (Masterflex L/S precision, Cole-Parmer) fitted with two pump-heads and Marprene pump tubing (Watson-Marlow) was used to add growth medium to the culture at a constant rate and remove excess liquid (waste) from the culture. NaOH, Na₂S and antifoam were added using program-controlled pumps. Exhaust gas was cooled through a condenser that was cooled to 10°C and an 80 cm water column served to generate back-pressure to direct exhaust gases to a gas chromatograph for analysis.

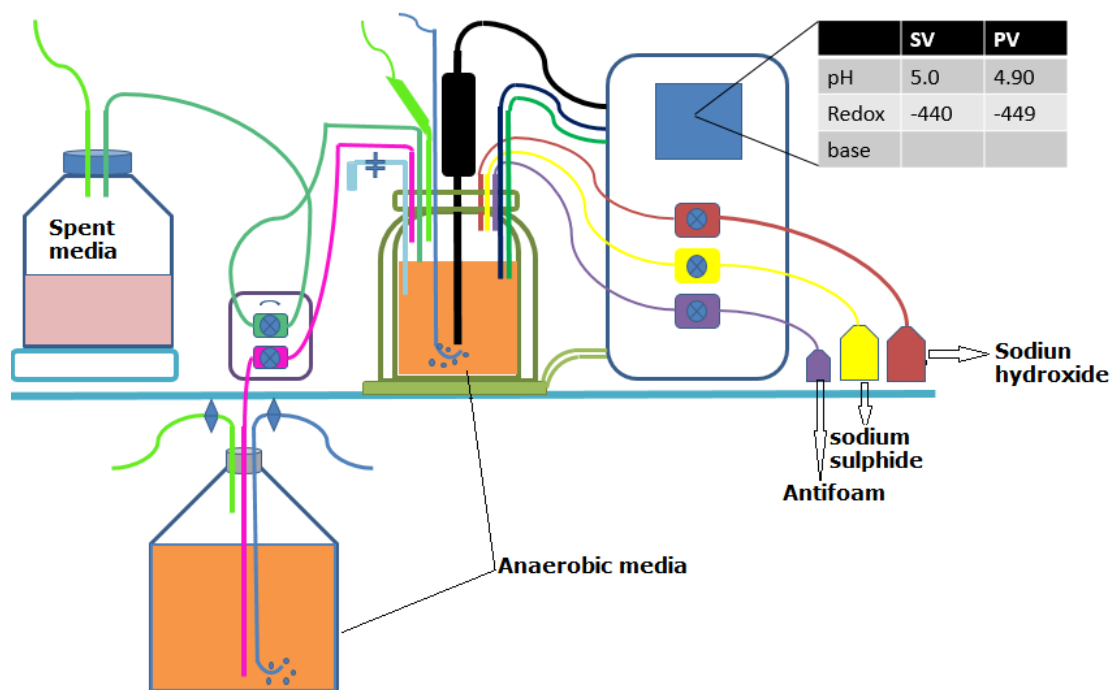


Figure 6.2 A typical example of bioreactor setup. Orange indicates anaerobic media, pink indicated spent media. Blue indicated gases in, neon green represents waste gases. Black indicates agitator. Light green indicates fermentation vessel with water jacket. Red tubing indicates NaOH, yellow indicates Na₂S and purple indicates antifoam. Hot pink indicates media addition and the blue green indicates waste media tubing. Purple and green lines indicate the pH and redox probes. Light blue is the sampling port for the culture. The system is controlled by a BioFlo computer and data recorded on a computer. (Modified from diagram drawn by Florence Annan).

6.2.2 Composition of B vitamin solution

The B vitamin solution used in the bioreactor was listed in Table 6.1. The solution was balanced to pH 7.4 in order to prevent precipitation and then sterilised by autoclaving in a 2 L Duran flask.

Table 6.1A list of B-vitamins used in media for continuous gas fed bioreactor

<i>STOCK COMPONENT</i>	<i>1.0 L</i>
<i>Thiamine HCl</i>	<i>50 mg</i>
<i>Riboflavin</i>	<i>50 mg</i>
<i>Nicotinic Acid</i>	<i>50 mg</i>
<i>Pantothenic Acid</i>	<i>50 mg</i>
<i>Pyridoxine HCl</i>	<i>10 mg</i>
<i>Biotin</i>	<i>20 mg</i>
<i>Folic Acid</i>	<i>20 mg</i>
<i>4-Aminobenzoic Acid</i>	<i>50 mg</i>
<i>Cyanocobalamin</i>	<i>50 mg</i>
<i>Lipoic Acid</i>	<i>50 mg</i>
<i>Distilled water</i>	<i>Up to total volume of 1 L</i>

6.2.3 Metabolomics sampling

Metabolomics sampling was done as reported previously in chapter 4, section 4.2.4 with the exception of the volume of samples taken was 1.5 mL instead of 1 mL. Samples for metabolomics analysis were taken when steady states were reached and when media condition is changed.

6.2.4 Extraction of intracellular metabolites from *C. autoethanogenum*

The extraction protocol was the same as reported in chapter 3, section 3.2.4.2 using freeze thaw cycles.

6.2.5 LC-MS and data analysis

LC-MS data analysis were conducted as previously described in chapter 4 section 4.2.6.

6.2.5.1 Targeted metabolomics data analysis

Authentic standards at eight calibration levels (0.5 μ M – 200 μ M) were analysed prior to the samples. Targeted metabolomics data were processed using Tracefinder (ThermoFisher) and metabolites were identified by their accurate mass ($\pm$ 5 ppm) and retention times. Metaboanalyst 3.0 was used for metabolic pathway analysis by uploading metabolites of significance to obtain pathways that were significantly affected.

6.2.5.2 Untargeted metabolomics data analysis

Samples analysed for untargeted metabolomics were randomised with analysis of pooled QC after every 10 injections. Raw data files obtained analysis of high pantothenate (HP) and low pantothenate (LP) conditions were processed using Progenesis QI software (Waters, USA). This software starts the data analysis process by picking peaks followed by peak alignments. QC is used as a reference sample to identify adducts and isotopes deconvolution, normalisation and transformation results for the dataset. The normalised data from Progenesis QI containing the abundance of extracted metabolites peaks were submitted to Simca P+14 (Umetrics AB, Sweden) for multivariate analysis (MVA). The data was also scaled to unit variance (UV) which gives the weight of each variable by its standard deviation, which shrinks the weight of intense features and stretches the weight of smaller ones so that all features have equal weights (Eriksson et al., 2006b). Principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) were used to investigate metabolic changes between HP and LP conditions. To maximise the classification between HP and LP groups, successive OPLS-DA models were generated comprising only two classes.

UV scaling was found to generate multivariate models with higher fitness of model (R^2Y) and predictive ability (Q^2) values (Eriksson et al., 2006a) compared to other forms of scaling such as pareto scaling. Cross-validation was done to

evaluate the robustness of the models by monitoring the fitness of model (R^2Y) and predictive ability (Q^2) values. Permutation test was used to evaluate the validity of the predictive ability (Q^2 values) of the models.

6.2.5.3 Metaboanalyst 3.0 data analysis

Data for pantothenate limitation experiment was done by submitting sample in triplicates for the two conditions. The sample normalisation was done using the median, no data transformation was used, and auto scaling was used for data scaling. Auto scaling uses mean-centred and divided by the standard deviation (SD) of each variable. Figure 6.3 shows before and after normalisation of the data, which shows great improvement to the peaks. After this step the results were generated, and they included chemometrics analysis, feature analysis, cluster analysis and classification and feature selection.

The identification of important compounds is achieved by three approaches; 1) a differential expression analysis that the user can use to identify compounds that significantly different between groups, 2) a coexpression analysis helps the user to identify compounds that share similar or opposite trends in concentration changes in different conditions and 3) a pattern search that allows users to search for specific trends of interest from their method by matching with specific templates (Xia et al., 2012).

The same approach stated above was used for data analysis of phosphate limitation experiment. Statistical analysis and metabolic pathway analysis were conducted using dataset from high phosphate concentration (HP) and phosphate limitation phase (P_lim) in three replicates. For HP and P_lim phases the concentrations were averaged for the 7 days and 9 days respectively.

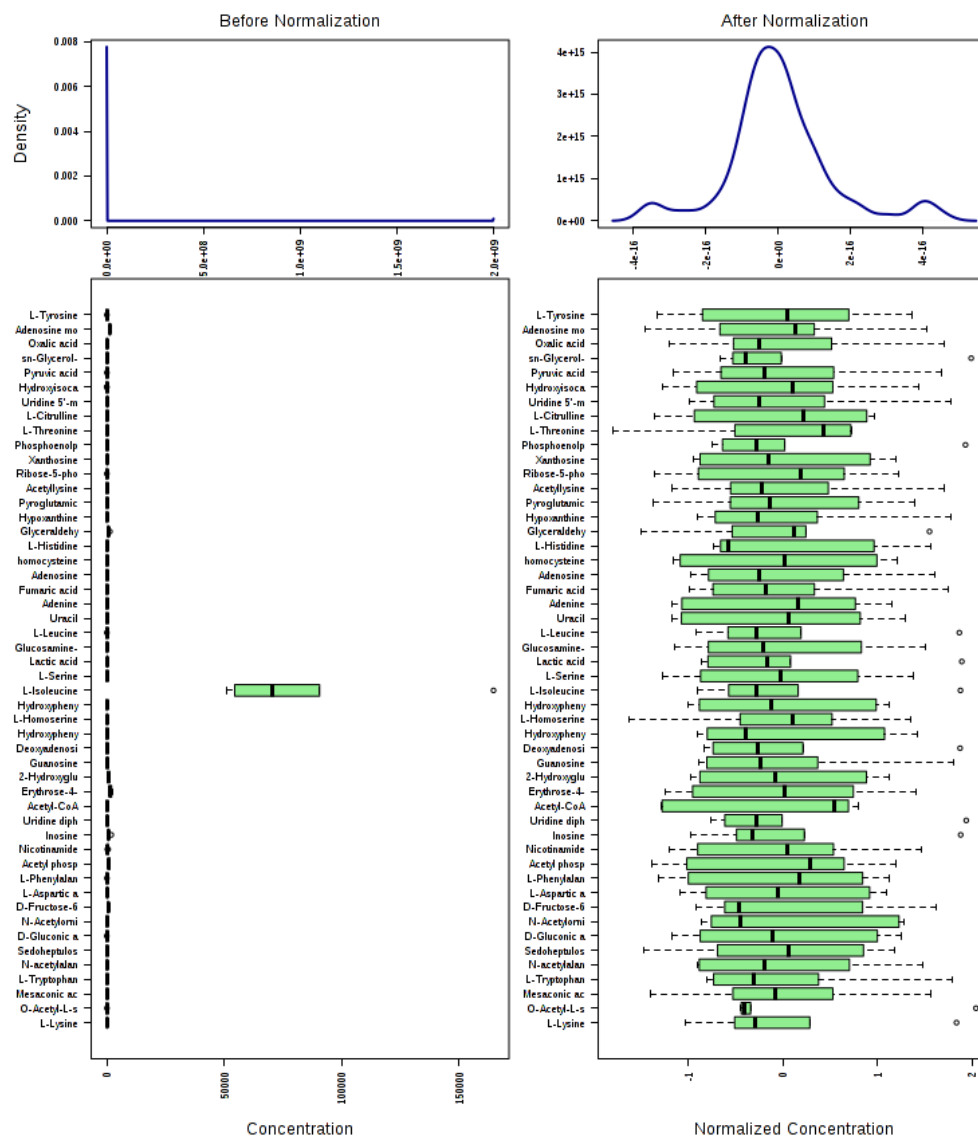


Figure 6.3 Representation of data before and after normalisation in Metaboanalyst 3.0. Data was normalised by the median of the data.

6.3 Results and discussion

6.3.1 Targeted metabolomics analysis of pantothenate limitation in *C. autoethanogenum*

Targeted metabolomics data were generated for set of samples and normalised using the U-¹³C-labelled internal standard generated from *C. pasteurianum*. Out of the 133 metabolites of interest, 88 metabolites were quantified from *C. autoethanogenum* cell pellets. Interpreting this type of complex data can be difficult and, in this section, several different approaches are used to analyse and visualise the data. In Table 6.2 is a heat map to show the intracellular concentrations of metabolites during pantothenate limiting conditions. The two pantothenate states were compared to confirm metabolites that were significantly affected when concentration of pantothenate in the media was reduced. As it would be expected, the concentration of acetyl-CoA during optimal concentration of pantothenate was reduced from 0.05 mM to 0.04 mM when pantothenate concentration was decreased due to reduced CoA production (Das et al., 1966, Rock et al., 2000, Spry et al., 2008, Toennies et al., 1966). The decline in acetyl-CoA concentration during low pantothenate (LP) conditions confirms the results observed with reduction in acetate production and increase in ethanol production which was also been observed in *C. ragsdalei* (Kundiyana et al., 2011). This reduction in acetyl-CoA concentration also accounts for the fact that during LP conditions little CoA was produced from the pathway using L-aspartic acid to produce CoA that links to Wood–Ljungdahl pathway to produce acetyl-CoA. Another confirmation of the redundancy of the biosynthesis of pantothenate and CoA pathway can be found in the reduction of the concentration of L-aspartic acid from 400 µM to 138 µM from HP to LP showing less favourability for production of L-aspartic acid for use in this metabolic pathway.

Due to the number of metabolites detected, it can be difficult to see the wider effects of pantothenate limitations when examining individual metabolite data. Hence, a further analysis using MetaboAnalyst 3.0 was used to give a more meaningful interpretation of the data and the potential advantage of pathway analysis. Although MetaboAnalyst is usually used for untargeted metabolomics

data, we can draw from its advantages by using the dataset generated from targeted analysis where the metabolites are validated and fully characterised.

Table 6.2 Heatmap showing concentration of intracellular metabolites comparing high pantothenate (HP) and low pantothenate (LP) phases. Red shows lower concentration while green shows higher concentration. Concentrations are in mM and averaged over 7 days for HP and 6 days for LP. Standard deviation (SD) for each condition was determined with p values <0.05 showing significance.

Metabolites	HP	LP	SD HP	SD LP	P values	Adjusted p values
2-Hydroxyglutaric acid	0.319	1.230	0.769	2.458	4.16E-05	3.66E-03
2-ketoisovaleric acid	0.002	0.003	0.003	0.011	0.055	4.88E+00
Adenine	0.011	0.020	0.092	0.095	0.022	1.91E+00
Citraconic acid	0.001	0.003	0.002	0.019	0.013	1.13E+00
Glycine	0.076	0.166	0.942	1.577	0.091	8.05E+00
D-Fructose-6-phosphate	0.392	0.870	2.991	6.844	0.045	3.92E+00
Dihydroxyacetone phosphate (DHAP)	0.006	0.006	0.039	0.030	0.974	8.57E+01
Hydroxyphenyllactic acid	0.001	0.000	0.005	0.001	0.043	3.78E+00
L-Asparagine	0.014	0.023	0.138	0.265	0.2788	2.45E+01
Inosine monophosphate (IMP)	0.365	0.269	5.463	4.932	0.611	5.37E+01
L-Asparagine	0.014	0.023	0.138	0.265	0.279	2.45E+01
L-Citrulline	0.035	0.024	0.265	0.257	0.236	2.07E+01
L-Lysine	0.046	0.186	0.661	2.224	0.061	5.39E+00
L-Lysine	0.046	0.186	0.661	2.224	0.061	5.39E+00
N-acetylalanine	0.032	0.096	0.144	0.657	0.015	1.30E+00
Phosphoenolpyruvate	0.015	0.131	0.268	1.768	0.056	4.94E+00
Orotic acid	0.004	0.012	0.018	0.065	0.008	7.02E-01
Phosphoenolpyruvate	0.015	0.131	0.268	1.768	0.056	4.94E+00
Succinic acid	2.181	6.943	4.791	18.560	0.000	2.52E-02
trans-Aconitic acid	0.002	0.007	0.006	0.075	0.052	4.58E+00

Metabolites	HP	LP	SD HP	SD LP	P values	Adjusted p values
Thiamine	0.458	0.168	8.419	6.788	0.310	2.73E+01
trans-Aconitic acid	0.002	0.007	0.006	0.075	0.052	4.58E+00
Uridine diphosphate-N-acetylglucosamine (UDP-N-acetylglucosamine)	0.148	0.260	0.964	2.170	0.109	9.61E+00
Uracil	0.014	0.025	0.150	0.159	0.065	5.75E+00
Acetyl phosphate	0.788	1.497	14.194	23.864	0.338	2.97E+01
Cytidine	0.000	0.000	0.002	0.002	0.293	2.58E+01
Cytidine monophosphate (CMP)	0.042	0.086	0.186	0.698	0.062	5.47E+00
Dihydroorotic acid	0.001	0.002	0.009	0.017	0.172	1.51E+01
Deoxycytidine monophosphate (dCMP)	0.005	0.007	0.008	0.018	0.007	6.04E-01
Fumaric acid	0.005	0.016	0.027	0.062	0.001	8.81E-02
Glyceraldehyde-3-phosphate	0.616	0.639	3.936	3.509	0.867	7.63E+01
Glyoxylic acid	0.054	0.104	0.990	1.850	0.373	3.28E+01
Glyceraldehyde-3-phosphate	0.616	0.639	3.936	3.509	0.867	7.63E+01
Hypoxanthine	0.001	0.001	0.012	0.016	0.831	7.32E+01
Hydroxyphenylpyruvate	0.001	0.001	0.005	0.005	0.508	4.47E+01
Hydroxyphenylpyruvate	0.001	0.001	0.005	0.005	0.508	4.47E+01
Hydroxyphenylpyruvate	0.001	0.001	0.005	0.005	0.508	4.47E+01
Inosine	1.213	0.920	13.457	10.854	0.514	4.52E+01
Inosine	1.213	0.920	13.457	10.854	0.514	4.52E+01
L-Glutamic acid	1.516	1.737	8.825	17.554	0.665	5.85E+01
L-Proline	0.138	0.241	0.842	2.314	0.1498	1.32E+01
L-Glutamic acid	1.516	1.737	8.825	17.554	0.665	5.85E+01
L-Leucine	0.019	0.032	0.129	0.457	0.309	2.72E+01

Metabolites	HP	LP	SD HP	SD LP	P values	Adjusted p values
L-Valine	0.197	0.135	2.324	1.082	0.371	3.27E+01
Mesaconic acid	0.095	0.337	1.346	5.085	0.128	1.13E+01
Malic acid	0.005	0.016	0.001	0.060	0.002	1.52E-01
Mesaconic acid	0.095	0.337	1.346	5.085	0.128	1.13E+01
N-acetyl-glucosamine-6-phosphate	0.112	0.084	0.885	0.776	0.355	3.12E+01
Nicotinamide adenine dinucleotide phosphate (NADP)	0.580	0.536	3.661	1.355	0.666	5.86E+01
trehalose	0.021	0.020	0.111	0.258	0.876	7.71E+01
Sedoheptulose-7-phosphate	0.027	0.028	0.199	0.356	0.953	8.38E+01
Uridine 5'-monophosphate (UMP)	0.229	0.501	0.994	3.117	0.021	1.86E+00
Adenosine monophosphate (AMP)	1.228	1.725	6.603	9.405	0.122	1.07E+01
D-Gluconic acid	0.001	0.002	0.009	0.017	0.233	2.05E+01
Cytosine	2.771	2.963	34.958	36.586	0.882	7.76E+01
Glucosamine-6-phosphate	0.059	0.101	0.349	0.760	0.088	7.78E+00
Guanosine	0.162	0.392	1.364	3.453	0.052	4.56E+00
Isocitric acid	0.001	0.001	0.023	0.032	0.881	7.75E+01
L-Histidine	0.012	0.036	0.086	0.393	0.063	5.52E+00
L-Methionine	0.035	0.012	0.387	0.193	0.075	6.63E+00
L-Phenylalanine	0.046	0.036	0.327	0.135	0.341	3.00E+01
L-Threonine	0.175	0.191	1.113	1.952	0.777	6.84E+01
O-Acetyl-L-serine	0.001	0.003	0.003	0.038	0.0619	5.45E+00
N-Acetylcysteine	0.010	0.022	0.039	0.086	0.002	1.98E-01
N-Acetylmethionine	0.020	0.036	0.139	0.372	0.167	1.47E+01
phenyllactic acid	0.000	0.000	0.001	0.000	0.243	2.14E+01

Metabolites	HP	LP	SD HP	SD LP	P values	Adjusted p values
Phenyllactic acid	0.000	0.000	0.001	0.000	0.243	2.14E+01
Acetyl CoA	0.0489	0.0358	0.299	0.454	0.3609	3.18E+01
Xanthine	0.002	0.003	0.026	0.034	0.400	3.52E+01

6.3.1.1 Targeted metabolomics analysis using MetaboAnalyst 3.0

Intracellular concentration metabolites obtained for both HP and LP (three replicates each) were submitted for further analysis using MetaboAnalyst 3.0. Multivariate analysis (MVA) of the data showed separation between the two states. PCA and OPLS-DA score plots shown in Figure 6.4 A and B respectively were used to further look for significant features such as VIP scores.

Identification of compounds of significance was obtained using VIP scores. The VIP scores >1 was selected for compounds that contributed significantly to the changes observed as shown in Figure 6.5. A total of 40 metabolites were identified. The metabolites that contributed significantly to the changes observed included 2-hydroxuglutarate, L-aspartic acid, hydroxyphenyllactic acid and some TCA intermediates such as succinic acid and malic acid. As explained before, it is not surprising to see L-aspartic acid and Acetyl- CoA in the list and the data from MetaboAnalyst 3.0 confirms observations made from the targeted dataset.

With the list of metabolites obtained from VIP scoring, a further analysis can be conducted to further interpret the data. The pathway analysis of the metabolites can be achieved using MetaboAnalyst 3.0, which was the next step for the data analysis.

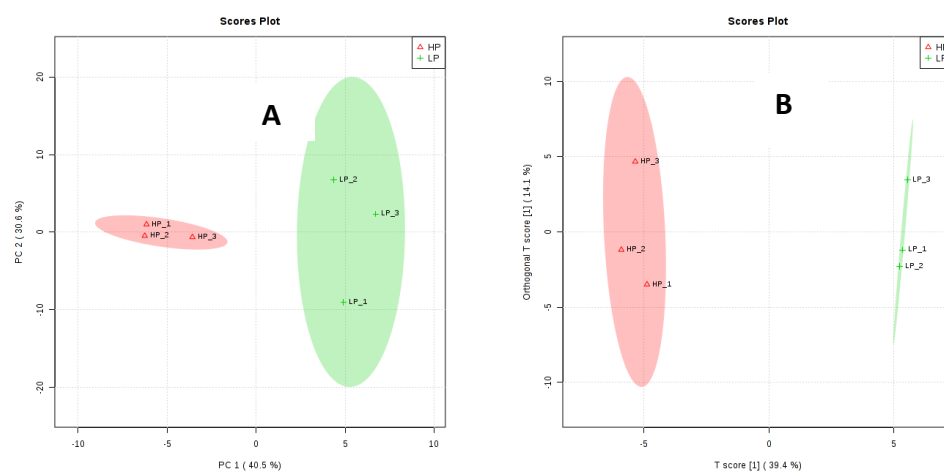


Figure 6.4 PCA (A) and PLS-DA (B) score plots of high pantothenate (HP) and low pantothenate (LP) showing separation and clustering of the two pantothenate conditions. Pink-HP and green-LP.

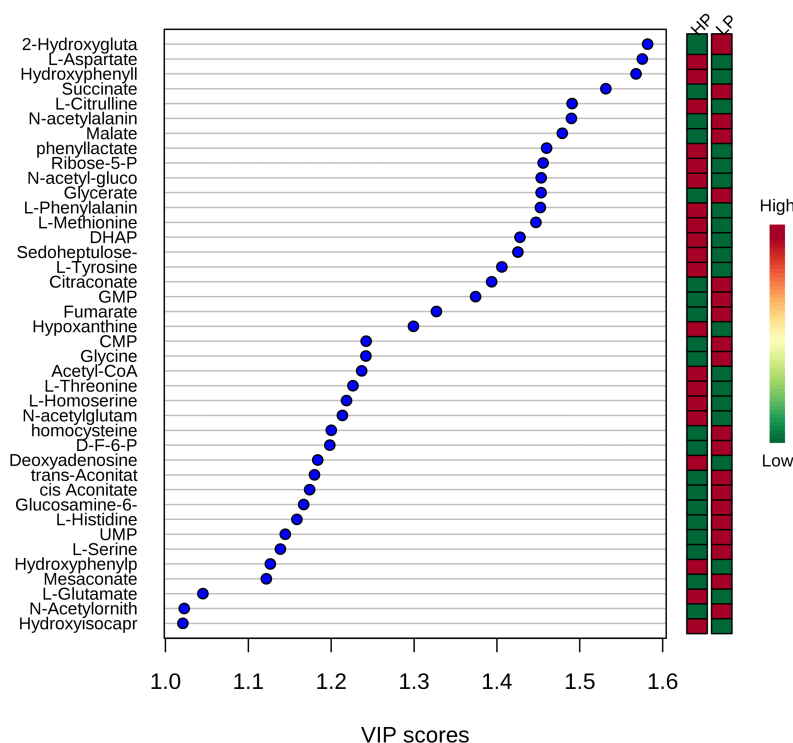


Figure 6.5 List of 40 metabolites that contributed significantly to the changes observed when pantothenate concentration was decreased. Metabolites with variable importance in projection (VIP) scores > 1 were chosen to determine

metabolites of significant contribution. Vip scores are calculated as the weighted sum of squared correlations that exist between the PLSDA components and the original variable. The metabolites are in descending order of importance. The coloured boxes on the right indicate absolute concentrations of the corresponding metabolites for each group.

6.3.1.2 Metabolic pathway analysis for targeted metabolomics data using MetaboAnalyst 3.0

Metabolic pathway analysis was done by submitting the 40 metabolites with VIP scores >1 to MetaboAnalyst 3.0 considering p values obtained from univariate analysis. The metabolic pathways found in MetaboAnalyst 3.0 are obtained from KEGG database. The pathways are presented as networks of compounds using metabolites as nodes and the reactions as the edges.

As shown in Table 6.3 are the *p* values and false discovery rate (FDR) for all metabolic pathways in order of importance. The pathways outlined were investigated to see how they contribute or affect CoA production within the cells. A few of the pathways were selected to further confirm the effects these pathways have on pantothenate and CoA biosynthesis.

Table 6.3 List of metabolic pathways from Metaboanalyst 3.0 obtained after submitting 40 metabolites with VIP>1 for metabolic pathway analysis. The results shown are the pathways of significance using p values<0.05.

Metabolic pathways	p value	Holm adjust	FDR	Impact
Glycine, serine and threonine metabolism	0.0004	0.036	0.018	0.50216
Aminoacyl-tRNA biosynthesis	0.0005	0.042	0.018	0.17857
Alanine, aspartate and glutamate metabolism	0.0006	0.052	0.018	0.6
Nitrogen metabolism	0.0014	0.111	0.029	0
Cyanoamino acid metabolism	0.0026	0.198	0.042	0
Arginine and proline metabolism	0.0032	0.242	0.043	0.19453
Novobiocin biosynthesis	0.0042	0.316	0.049	1

Benzoate degradation via CoA ligation	0.0053	0.387	0.052	0
Citrate cycle (TCA cycle)	0.0058	0.424	0.052	0.15142
Tyrosine metabolism	0.0071	0.504	0.057	0
Methane metabolism	0.0116	0.817	0.085	0.1
Pyruvate metabolism	0.0438	1	0.250	0.23019
Glyoxylate and dicarboxylate metabolism	0.0494	1	0.250	0
Phenylalanine, tyrosine and tryptophan biosynthesis	0.0494	1	0.250	0.0398
Butanoate metabolism	0.0494	1	0.250	0.04878
Cysteine and methionine metabolism	0.0501	1	0.250	0.17867

Metabolic pathways of significance determined by using p values <0.05 together with their node of importance score. P values calculated by t-test. FDR-false discovery rate. Holm adjust- Holm-Bonferroni adjusted p-values (used to prevent errors associated with multiple comparisons).

Glycine, serine and threonine metabolism is a central pathway connected other pathways that are linked to pantothenate and CoA biosynthesis. Pathways such as alanine, aspartate and glutamate metabolism, TCA cycle and pyruvate metabolism are directly linked to Glycine, serine and threonine metabolism pathway.

Pyruvate metabolism (Figure 6.6) is an important pathway for acetyl-CoA production. Pyruvate produced through glycolysis is a precursor for acetyl-CoA production which then leads to acetate production, hence when this pathway is affected; acetyl-CoA production reduces leading to less acetate production. The pyruvate metabolism pathway is also connected to citrate acid cycle (TCA cycle) (Figure 6.7). The conversion of pyruvate into oxaloacetate which is then converted to malate and fumarate to fuel the TCA cycle. The concentration of TCA intermediates such as malate, fumarate and succinate increase when pantothenate concentration is reduced. A key intermediate 2-ketoglutarate (2-oxoglutarate) stays constant from HP to LP and is directly linked to alanine, aspartate and glutamate (AAG) metabolism (Figure 6.8). It has been reported in

E. coli that reduction in pantothenate concentration led to the inability to generate enough quantities of precursors through the TCA cycle which leads to lack of amino acid biosynthesis (Jackowski and Rock, 1986).

The alanine, aspartate and glutamate (AAG) metabolism pathway has L-aspartate as the main product from associated pathways but it serves as the precursor metabolite for the pantothenate and CoA biosynthesis pathway. The reduction of pantothenate concentration showed a reduction in the concentration of key metabolites such as acetyl-CoA and L-aspartate. An overview of the metabolic pathways show that targeted metabolomics of the samples gave a good understanding of the metabolism of *C. autoethanogenum* during this growth limitation. The results below show pathways that are directly or indirectly linked to pantothenate and CoA biosynthesis with intracellular concentrations of metabolites confirming their effects on acetyl-CoA production.

6.3.2 Untargeted metabolomics data analysis

6.3.2.1 Metabolomics data analysis for pantothenate limitation study in *C. autoethanogenum*

For a complete modelling of the significant differences between steady states, data generated by LC-MS analysis were submitted for multivariate analysis. PCA score plots showed little clustering between HP and LP conditions (Figure 6.9) and the stability of the analytical run was confirmed by the clustering of the pooled QC samples to the centre of the PCA score plots (Beckmann et al., 2008). A supervised multivariate analysis model, orthogonal partial least square-discriminant analysis (OPLS-DA) was created to find differences between HP and LP conditions. There was visual evidence of the clustering and separation between the two conditions as shown in Figure 6.10. Cross validation of the OPLS-DA model of HP versus LP were done by monitoring fitness of the model, R^2Y and the predictive ability, Q^2 values that were 0.993 and 0.941 respectively, demonstrating a valid model. Further permutation test was conducted for the OPLS-DA score plots to confirm that the predictive capability of the model was

not due to over fitting of the data (Eriksson et al., 2006b). The obtained intercept values for Q^2 value from the regression line of Y permuted Q^2 value was less the Q^2 value of the OPLS-DA model (Figure 6.11). The intercept was -0.378 which indicates that the model has a reliable predictive power and not due to over fitting of the data.

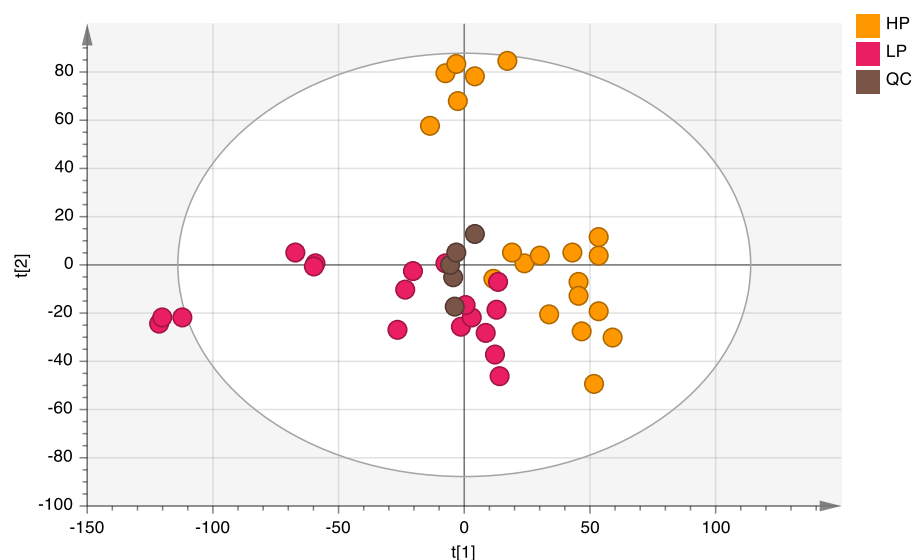


Figure 6.6 PCA score plots for all samples obtained from the pantothenate limiting experiment. Pink represents low pantothenate concentration (LP), orange for high pantothenate concentration (HP) and brown represents pooled QC.

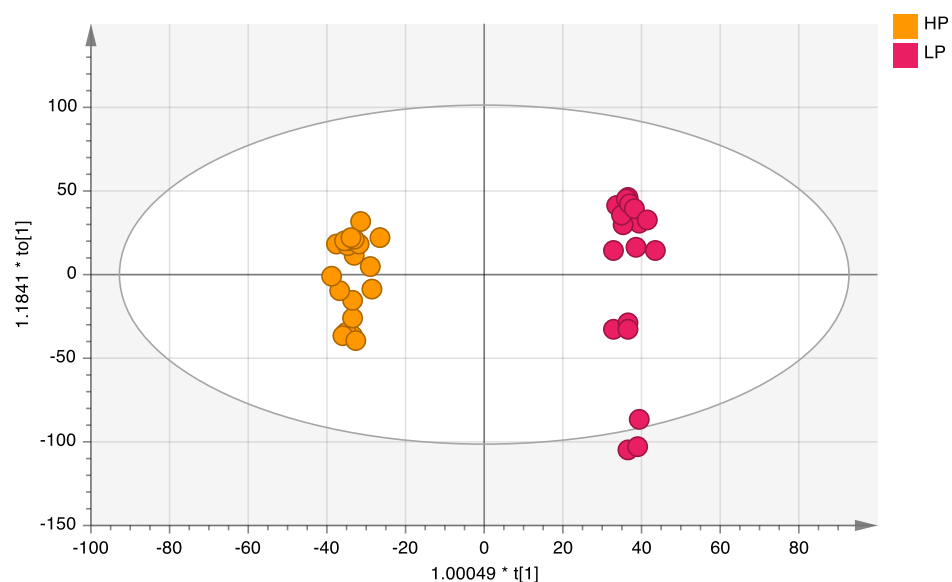


Figure 6.7 OPLS-DA score plots for all samples obtained from the pantothenate limiting experiment. Pink represents low pantothenate concentration (LP) and orange for high pantothenate concentration (HP)

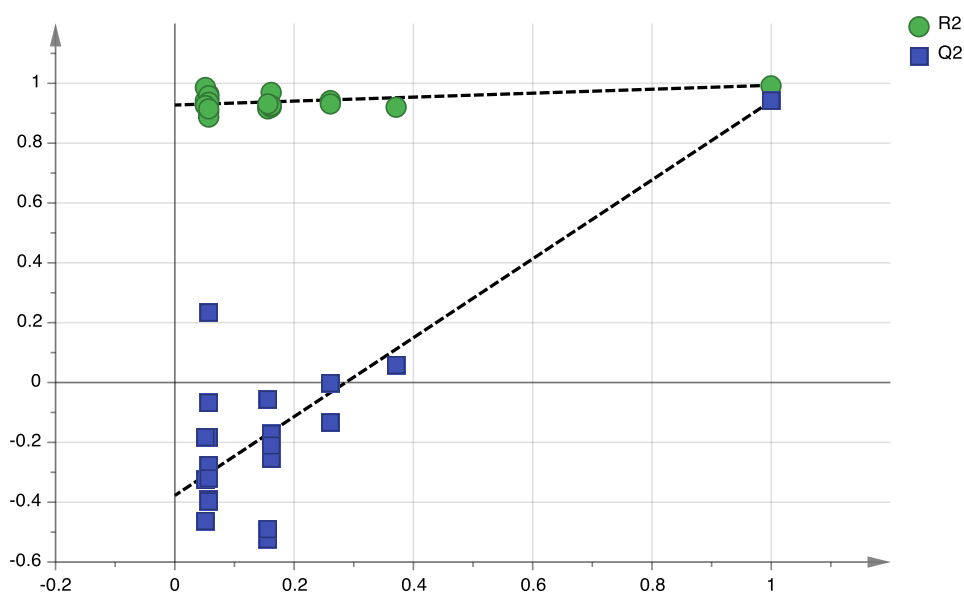


Figure 6.8 Validation of OPLS-DA models of higher pantothenate and lower pantothenate conditions using permutation test ($n=38$). R^2 is the model fit and Q^2 is the predictive ability of the model. The Q^2 intercept values of the regression line of randomly permuted Q^2 value was -0.378, indicating that the

model is statically sound, and the high prediction ability of the model is not due to over fitting.

6.3.2.2 Tentative identification of compound using Compound Discoverer 2.1

There were limitations with identification of compounds using VIP scores alone from SIMCA-P from the database used for the Progenesis Q1. Identification of metabolites was difficult due to differences with metabolites identified and those in the database used. Compound discoverer incorporates different databases such as KEGG, biocyc and chemspider making it possible to have identification from at least one of these databases. For this reason, raw data files were submitted to Compound Discoverer (Hartmann et al.) 2.1 (Thermo Scientific) for multivariate analysis as before but this has the added advantage of using different databases (BIOCYC, KEGG etc.) to putatively identify compounds (Vrobel et al., 2017) and a further step for metabolic pathway analysis for the specific organism. CD has been used over the years for metabolite profiling and identification for different applications (Najdekr et al., 2016, Scheibner et al., 2016, Vrobel et al., 2017). The cited journals used the previous version of CD 2.0 but in this work the current version CD2.1 was used. A total of 3217 ions were detected in both positive and negative mode with 688 identified from Chemspider, Biocyc and KEGG databases. The lists of compounds identified from these databases are detailed with chemical structure and formula, retention times (RT) and molecular weight (Figure 6.9).

To find metabolites that contributed significantly when pantothenate concentration was reduced, metabolites with *p values* <0.05 obtained from multivariate analysis were filtered and 27 metabolites were found fully annotated (Table 6.4). Metabolic pathway analysis was conducted with the list of metabolites for further pathway interpretation of the results.

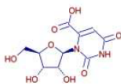
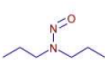
Structure	Name	RT [min]	Formula	Molecular Weight	Group Areas			
	Orotidine	8.04	C10 H12 N2 O8	288.0593	1.60e5	7.14e4	1.83e5	3.29e5
	4-Amino-4-deoxypentopyranose	9.80	C5 H11 N O4	149.0687	1.58e6	5.84e5	1.85e6	3.00e6
	D-Alanyl-D-alanine	7.11	C6 H12 N2 O3	160.0846	7.50e6	3.60e6	6.15e6	8.04e6
	Oryzalin metabolite	12.33	C6 H14 N2 O	130.1106	4.79e5	9.84e4	4.29e4	1.59e5

Figure 6.9 A representative compound identification output obtained from compound discoverer 2.1 showing compound structure, name, retention time, chemical formula and molecular weight

Table 6.4 List of metabolites from compound discoverer 2.1 in order of significant importance using P values < 0.05

Compounds	Formula	Molecular Weight	RT [min]	P-value	Adj. P-value
Noradrenaline	C ₈ H ₁₁ N O ₃	169.0737	7.189	1E-15	1.4E-13
Uric Acid	C ₅ H ₄ N ₄ O ₃	168.0281	8.231	1E-15	1.4E-13
L-(+)-Aspartic acid	C ₄ H ₇ N O ₄	133.0374	8.72	1.56E-10	7.92E-09
DL-Tyrosine	C ₉ H ₁₁ N O ₃	181.0737	8.744	2.29E-07	3.77E-06
8-hydroxy-deoxyguanosine	C ₁₀ H ₁₃ N ₅ O ₅	283.091	8.346	4.52E-07	6.82E-06
L-(-)-Methionine	C ₅ H ₁₁ N O ₂ S	149.0509	8.165	6.72E-07	9.86E-06
Formyl-L-methionine	C ₆ H ₁₁ N O ₃ S	177.0457	5.051	1.36E-06	1.8E-05
DL-Glutamic acid	C ₅ H ₉ N O ₄	147.053	7.665	5.35E-06	6.49E-05
Ethyl lactate	C ₅ H ₁₀ O ₃	118.0628	5.046	1.03E-05	0.000115
Phenylacetic acid	C ₈ H ₈ O ₂	136.0522	4.995	2.89E-05	0.000303
Adenine	C ₅ H ₅ N ₅	135.0545	9.343	3.68E-05	0.000367
Itaconic acid	C ₅ H ₆ O ₄	130.0264	7.025	6.45E-05	0.00059
Niacin	C ₆ H ₅ N O ₂	123.0319	7.023	0.000119	0.00101

Compounds	Formula	Molecular Weight	RT [min]	P-value	Adj. P-value
L-(-)-Threonine	C ₄ H ₉ N O ₃	119.0581	8.915	0.000449	0.003293
L-(-)-Serine	C ₃ H ₇ N O ₃	105.0424	9.34	0.00069	0.004815
L-Pyroglutamic acid	C ₅ H ₇ N O ₃	129.0425	7.12	0.001097	0.006583
Ethyl 3-methylthiopropionate	C ₆ H ₁₂ O ₂ S	148.0565	8.498	0.001335	0.007762
S-Allyl-L-cysteine	C ₆ H ₁₁ N O ₂ S	161.051	5.054	0.001379	0.007931
L-(+)-Leucine	C ₆ H ₁₃ N O ₂	131.0945	8.147	0.001467	0.008354
Pterin	C ₆ H ₅ N ₅ O	163.0492	8.084	0.001967	0.010762
DL-Citrulline	C ₆ H ₁₃ N ₃ O ₃	175.0956	9.284	0.002131	0.011322
oxiracetam	C ₆ H ₁₀ N ₂ O ₃	158.069	9.278	0.002129	0.011322
L-alpha-Aspartyl-L-lysine	C ₁₀ H ₁₉ N ₃ O ₅	261.1319	8.617	0.002625	0.013317
L-(+)-Valine	C ₅ H ₁₁ N O ₂	117.0788	8.421	0.003277	0.016181
Malonic acid	C ₃ H ₄ O ₄	104.011	8.924	0.003422	0.016749
N-Acetyl-L-glutamic acid	C ₇ H ₁₁ N O ₅	189.0635	8.212	0.00483	0.02246
α-aminoadipic acid	C ₆ H ₁₁ N O ₄	161.0687	5.388	0.009046	0.038531

Compounds	Formula	Molecular Weight	RT [min]	P-value	Adj. P-value
DL-Arginine	C ₆ H ₁₄ N ₄ O ₂	174.1115	12.599	0.010391	0.042322

6.3.3 Pathway analysis using metabolite identification from Compound Discoverer 2.1

Metabolic pathway analysis was done as before but this time with the list of putative metabolites in Table 6.4. The lists of metabolites were narrowed down to metabolites that were present in the organism in question, *C. autoethanogenum*. For example, phaseic acid was removed from the list because is a plant hormone linked photosynthesis arrest and not particularly related to the organism and false identifications were removed. In Table 6.5 are list of metabolic pathways in order of importance obtained from the untargeted metabolomics analysis. Interestingly there are similar pathways that were also found with the targeted metabolomics analysis. Aminoacyl-tRNA biosynthesis pathway was common between targeted metabolomics and untargeted metabolomics data analysis. Aminoacyl-tRNA synthetases (aatRS) play a pivotal function in the biosynthesis of protein by catalysing the attachment of amino acids to the 3' end of its associated tRNA (Putzer and Laalami, 2000-2013). Aminoacyl-tRNAs act as substrates for translation and they are central in determining how genetic codes are interpreted as amino acids (Ibba and Soll, 2000). It is not surprising that the aminoacyl-tRNA biosynthesis (Figure 6.10) pathway is upregulated because the pantothenate and CoA biosynthesis is very dependent on amino acid intermediates and other amino acid biosynthesis pathways, such as alanine, aspartate and glutamate metabolism, which is directly linked to pantothenate and CoA biosynthesis pathway. The metabolic pathways obtained from untargeted metabolomics and targeted metabolomics have a lot of similar pathways between them. There were more metabolic hits from the targeted analysis compared to the untargeted analysis, but the same conclusions could be deduced from both with targeted metabolomics been able to provide in-depth interpretation.

Table 6.5 List of metabolic pathways obtained from MetaboAnalyst 3.0 using metabolite list obtained from untargeted metabolomics analysis using Compound discoverer 2.1

Metabolic pathways	Raw p values	Holm adjust	FDR	Impact
Aminoacyl-tRNA biosynthesis	0.00756	0.60481	0.60481	0.07143
Cyanoamino acid metabolism	0.018985	1	0.75941	0
Glycine, serine and threonine metabolism	0.039963	1	0.97064	0.2528
Nicotinate and nicotinamide metabolism	0.048532	1	0.97064	0.08108
Cysteine and methionine metabolism	0.080258	1	1	0.116
Arginine and proline metabolism	0.096571	1	1	0.19073
Valine, leucine and isoleucine biosynthesis	0.16225	1	1	0.00085
C5-Branched dibasic acid metabolism	0.18056	1	1	0
D-Glutamine and D-glutamate metabolism	0.18056	1	1	0.37931
beta-Alanine metabolism	0.2261	1	1	0
Benzoate degradation via CoA ligation	0.24795	1	1	0
Sulphur metabolism	0.30999	1	1	0
Methane metabolism	0.30999	1	1	0.1
Nitrogen metabolism	0.32956	1	1	0
Lysine biosynthesis	0.34859	1	1	0
Alanine, aspartate and glutamate metabolism	0.43621	1	1	0.3125
Pantothenate and CoA biosynthesis	0.45231	1	1	0
Purine metabolism	0.56287	1	1	0.04659
Valine, leucine and isoleucine degradation	0.6571	1	1	0

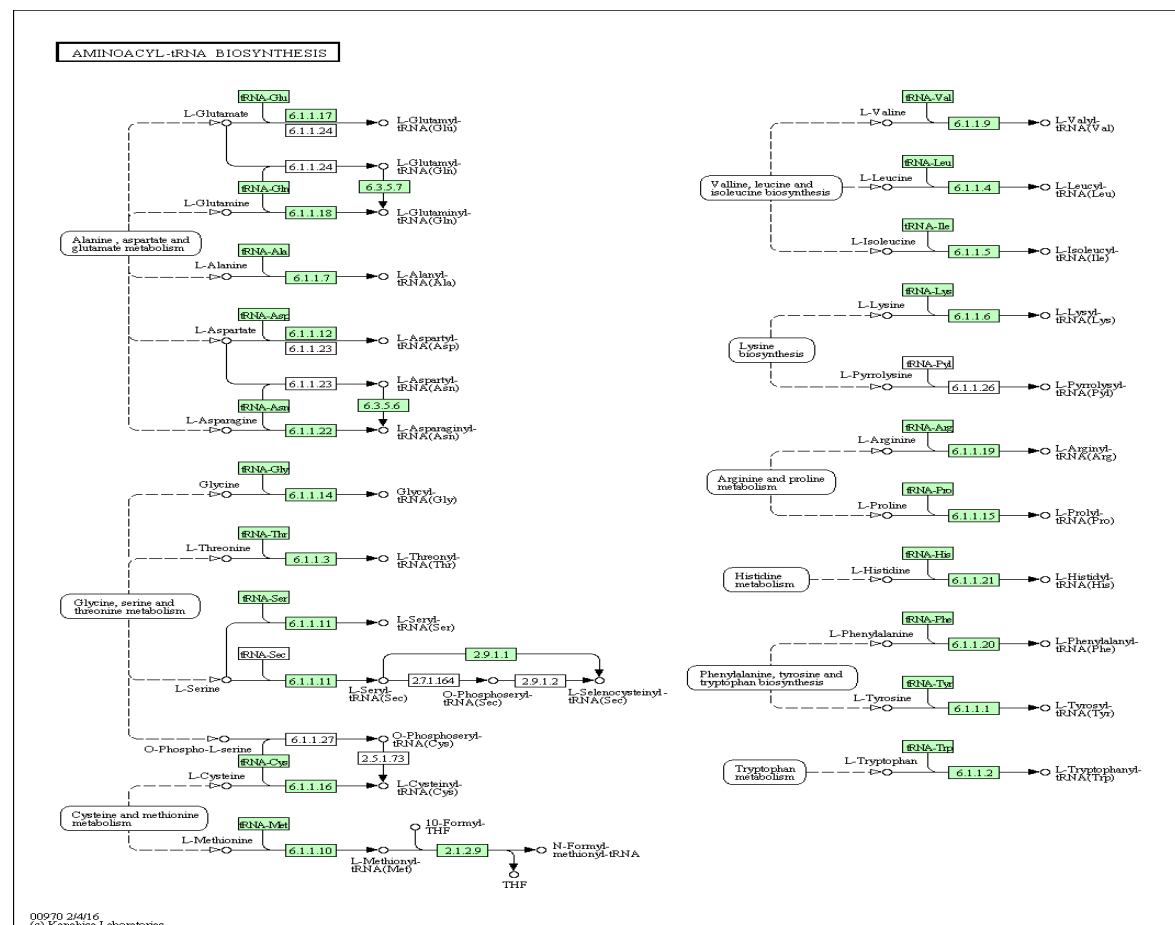


Figure 6.10 Aminoacyl-tRNA biosynthesis pathway of *C. autoethanogenum* (KEGG Pathway database)

6.3.4 Using metabolomics to understand the effect of pantothenate limitation on ethanol production

Changing physical and chemical environment of the organism can optimise the products produced during syngas fermentation. In this study, the effects of limiting pantothenate concentration on the Wood-Ljungdahl pathway was tested in continuous gas fed bioreactors. In Figure 6.11, acetate was the major product at optimal concentrations of pantothenate in steady state (SS) 3 but at SS4 when the concentration of pantothenate was reduced to 10 % ethanol was the major product. Explanation for the results seen in the steady states with standard pantothenate production and with pantothenate limitation lie in the methods the cells use to produce the economic products. As stated, the major differences between the two different conditions are a switch between high acetate production and low ethanol production in high pantothenate concentrations and low acetate, high ethanol production in low pantothenate concentrations with only a very slight change in the average optical density. Acetogens have been shown to produce increased amounts of ethanol in growth limiting conditions caused by lowering B-vitamin concentrations or reduction of yeast extract in the media (Phillips et al., 1994). The effects of limiting pantothenate in this experiment resulted in similar outcome. *C. autoethanogenum* uses the Wood-Ljungdahl pathway to reduce CO and CO₂ to form acetyl-CoA, which it uses to produce a range of downstream chemicals, notably, 2,3-butanediol, ethanol and acetate as well as utilising it for growth. The production of acetyl-CoA is impossible in *C. autoethanogenum* without the presence of pantothenate.

The results obtained here correlates with the results from metabolomics studies. The observation of decrease in acetyl CoA formation during lower pantothenate concentration confirms the increase in ethanol production.

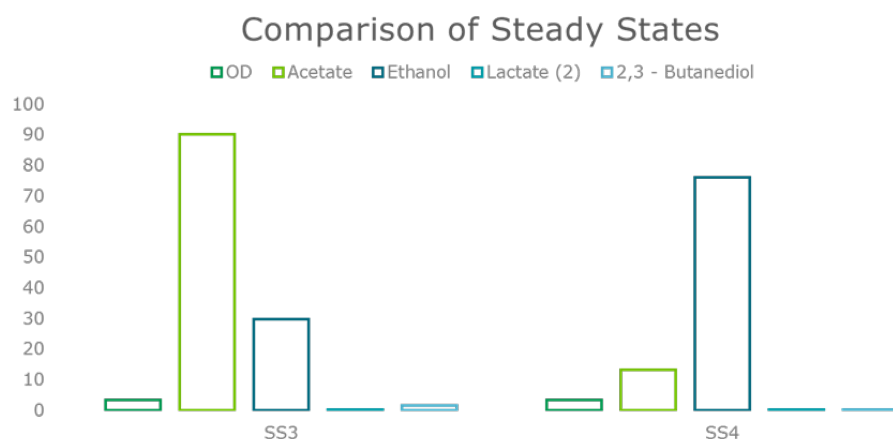


Figure 6.11 Comparison of steady states for solvent production during pantothenate limitation. In steady state 3(SS), optimal concentrations of pantothenate was present in the media and in SS4 the pantothenate concentration was reduced to 10 % of the standard concentration. There is an increase in ethanol levels during this reduction in pantothenate concentration. (diagram form Florence Annan).

It has been suggested that *C. ljungdahlii* shifts from acetate to ethanol production at high biomass concentrations in order to maintain ATP homeostasis, but this leads to the depletion of CoA (Mohammadi et al., 2012b). *C. autoethanogenum*, is so closely related to *C. ljungdahlii* that it is likely that these conclusions could be drawn for both organisms. However, in the limitation phase, it is not high biomass depleting the CoA availability but the low pantothenate availability, producing the same results and changing the flux of products accordingly.

6.3.5 The use of HILIC-IDMS method to investigate phosphate limitation in *C. autoethanogenum*

Analysis of intracellular metabolites of *C. autoethanogenum* to investigate the effect of phosphate limitation on intracellular compounds was done using developed and validated HILIC-IDMS method reported in chapter 3. A total of 80 metabolites were quantified for this experiment. A heat map showing the

orderly progression of metabolites before and after the phosphate-limiting phase is shown in Table 6.6. The concentrations are reported in mM with red colour indicating lower concentrations and green represents higher concentrations.

Intermediates of the TCA cycle such as succinate, fumarate and aconitic acid were observed to increase during the phosphate-limiting phase. The TCA cycle has been reported in *C. acetobutylicum* to be active during the acidogenic-solventogenic transition (Amador-Noguez et al., 2011) and this was observed in pH induced and gas induced metabolic shift experiments reported in chapters 4 and 5. Since the phosphate limitation was intended to improve ethanol production hence not surprising to see TCA intermediates increasing during this phase and this hopefully will be confirmed with the pathway analysis. Another important observation was the concentration of 2-Hydroxyglutaric acid increased nearly 8-fold during the phosphate limiting phase. 2-Hydroxyglutaric acid can be converted to α -ketoglutaric acid which is a key intermediate in the TCA cycle.

The concentration of L-Aspartate was also observed to increase during the phosphate limitation, but the concentration of acetyl-CoA decreased. L-Aspartate is the precursor for the production of coenzyme A through the pantothenate and CoA biosynthesis pathway (Toennies et al., 1966). The accumulation of L-aspartate during phosphate limitation phase is due to little or no production of CoA hence the decrease in acetyl-CoA concentration.

Univariate analysis was conducted as before to confirm if the changes observed were significant and changes associated with 33 metabolites were significant but upon correction using Bonferroni test only changes associated with 5 metabolites were significant.

Table 6.6 Intracellular concentration (mM) of intracellular metabolites during the phosphate limitation experiment conducted in *C. autoethanogenum*

Metabolites	HP	P_lim	SD_HP	SD P_lim	P value	Adjusted p value
UDP-N-acetylglucosamine	0.074	0.021	0.014	0.008	0.000	0.006
Deoxycytidine monophosphate (dCMP)	0.003	0.001	0.001	0.001	0.000	0.007
Citraconic acid	0.000	0.004	0.000	0.002	0.000	0.016
Uridine 5'-monophosphate (UMP)	0.117	0.060	0.020	0.018	0.000	0.019
Pyruvic acid	0.023	0.052	0.006	0.014	0.000	0.020
Acetyl CoA	0.021	0.010	0.005	0.006	0.001	0.118
L-Proline	0.100	0.022	0.034	0.008	0.002	0.178
Lactic acid	0.485	0.893	0.162	0.252	0.002	0.198
Adenosine monophosphate (AMP)	1.159	0.352	0.387	0.128	0.003	0.271
L-Aspartic acid	0.143	0.628	0.087	0.363	0.004	0.336
2-Hydroxyglutaric acid	0.260	1.070	0.138	0.655	0.006	0.527
Orotic acid	0.003	0.005	0.001	0.002	0.007	0.594
L-Serine	0.040	0.072	0.012	0.025	0.007	0.637
L-Citrulline	0.048	0.014	0.021	0.010	0.009	0.854
N-Acetylornithine	0.017	0.004	0.008	0.005	0.010	0.942
L-Valine	0.081	0.140	0.037	0.035	0.011	1.014
N-acetylglutamic acid	0.165	0.078	0.057	0.025	0.011	1.031
Fumaric acid	0.003	0.012	0.002	0.008	0.012	1.101
L-Glutamine	0.076	0.014	0.041	0.013	0.013	1.218
L-Glutamic acid	1.770	0.926	0.572	0.411	0.013	1.221
Succinic acid	0.072	0.040	0.036	0.011	0.083	7.533

Metabolites	HP	P_lim	SD_HP	SD P_lim	P value	Adjusted p value
Cytidine monophosphate (CMP)	0.013	0.006	0.005	0.003	0.016	1.452
Ribose-5-phosphate	0.032	0.007	0.017	0.004	0.018	1.641
N-acetylalanine	0.013	0.025	0.005	0.012	0.020	1.785
Malic acid	0.004	0.015	0.001	0.012	0.024	2.194
Glucosamine-6-phosphate	0.050	0.019	0.025	0.007	0.025	2.287
D-Gluconic acid	0.001	0.002	0.000	0.001	0.025	2.306
L-Tryptophan	0.025	0.008	0.014	0.002	0.031	2.840
trans-Aconitic acid	0.005	0.008	0.002	0.002	0.033	2.963
Inosine monophosphate (IMP)	0.467	0.043	0.366	0.048	0.036	3.273
N-Acetylcysteine	0.019	0.010	0.008	0.003	0.045	4.080
sn-Glycerol-3-phosphate	0.035	0.019	0.015	0.010	0.047	4.313
Acetyl phosphate	1.479	0.000	1.388	0.000	0.048	4.343
Erythrose-4-phosphate	0.936	0.071	0.847	0.145	0.054	4.931
Inosine	3.332	9.598	1.730	8.999	0.073	6.647
Guanine	0.039	0.025	0.011	0.016	0.075	6.786
N-acetyl-glucosamine-6-phosphate	0.139	0.019	0.133	0.025	0.078	7.088
L-Tyrosine	0.072	0.040	0.036	0.011	0.083	7.533
Mesaconic acid	0.062	0.011	0.057	0.033	0.086	7.858
Adenosine	0.019	0.007	0.015	0.003	0.104	9.451
S-Adenosylmethionine (SAM)	1.835	4.048	2.260	2.831	0.118	10.776
Sedoheptulose-7-phosphate	0.030	0.017	0.017	0.008	0.129	11.716
Uracil	0.017	0.009	0.010	0.004	0.135	12.324
Acetyllysine	0.002	0.001	0.001	0.000	0.137	12.451
L-Arginine	0.011	0.006	0.007	0.002	0.156	14.151

Metabolites	HP	P_lim	SD_HP	SD P_lim	P value	Adjusted p value
Glyceric acid	0.026	0.037	0.008	0.020	0.159	14.468
Pyroglutamic acid	0.013	0.108	0.006	0.185	0.162	14.721
2-Isopropylmalic acid	0.001	0.002	0.000	0.002	0.187	17.060
D-Fructose-6-phosphate	0.217	0.130	0.138	0.076	0.199	18.153
Adenine	0.021	0.014	0.012	0.005	0.221	20.069
2-ketoisovaleric acid	0.002	0.002	0.000	0.001	0.239	21.723
trehalose	0.015	0.024	0.003	0.022	0.239	21.733
L-Histidine	0.009	0.013	0.005	0.010	0.253	23.036
Oxaloacetic acid	0.009	0.002	0.013	0.004	0.258	23.496
L-Leucine	0.012	0.016	0.005	0.008	0.301	27.421
Dihydroorotic acid	0.001	0.045	0.001	0.119	0.302	27.439
homocysteine	0.107	0.157	0.062	0.122	0.323	29.402
Phosphoenolpyruvate	0.027	0.005	0.049	0.004	0.325	29.539
Xanthine	0.054	0.026	0.068	0.028	0.365	33.250
L-Threonine	0.085	0.067	0.042	0.030	0.386	35.121
L-Phenylalanine	0.067	0.054	0.033	0.017	0.410	37.350
L-Lysine	0.008	0.007	0.003	0.002	0.420	38.217
Oxalic acid	0.166	0.136	0.089	0.058	0.488	44.386
L-Homoserine	0.043	0.035	0.024	0.018	0.500	45.545
Glycine	0.022	0.028	0.015	0.018	0.505	45.918
L-Methionine	0.013	0.017	0.011	0.012	0.570	51.860
Xanthosine	0.677	0.884	0.571	0.876	0.590	53.651
L-Asparagine	0.004	0.005	0.001	0.002	0.626	56.921
Glyceraldehyde-3-phosphate	0.738	0.614	0.501	0.548	0.659	59.946

Metabolites	HP	P_lim	SD_HP	SD P_lim	P value	Adjusted p value
Thiamine	0.672	0.418	1.342	0.345	0.668	60.771
Dihydroxyacetone phosphate (DHAP)	0.007	0.006	0.005	0.006	0.672	61.172
Guanosine	0.127	0.137	0.062	0.045	0.733	66.730
Glyoxylic acid	0.101	0.089	0.075	0.045	0.735	66.845
Ketoglutaric acid (2-Oxoglutaric acid)	0.005	0.005	0.001	0.001	0.875	79.637
Hydroxyisocaproic acid	0.016	0.015	0.005	0.016	0.923	83.984
Hypoxanthine	0.010	0.011	0.013	0.009	0.951	86.569
Hydroxyphenylpyruvate	0.002	0.002	0.000	0.001	0.954	86.821
Cytosine	5.512	5.291	9.846	4.871	0.961	87.431
Acetoacetic acid	0.832	0.824	0.362	0.358	0.971	88.324
L-Alanine	0.865	0.870	0.341	0.643	0.984	89.507

6.3.6 Data analysis and pathway analysis using MetaboAnalyst 3.0

Statistical analysis comparing metabolite concentration in the two phases showed clear separation and clustering of the groups in the multivariate analysis observed in PCA, and OPLS-DA plots. Multivariate data analysis will be validated using another software (SIMPCA-P14). MetaboAnalyst 3.0 also gives VIP scores for metabolites showing metabolites that contributed significantly to the separation between groups. Metabolites with VIP scores >1 are said to be of the most significance as shown in Figure 6.12. In Figure 6.13 the list of metabolites are listed in order of decreasing VIP scores with a legend on the top right indicating the concentrations. Green colour indicates lower concentration while red indicates high concentration. As discussed above 2-Hydroxyglutaric acid is the most significant metabolite with some TCA intermediates following the trend. A total of 51 metabolites were found to be significant with VIP scores >1.0 .

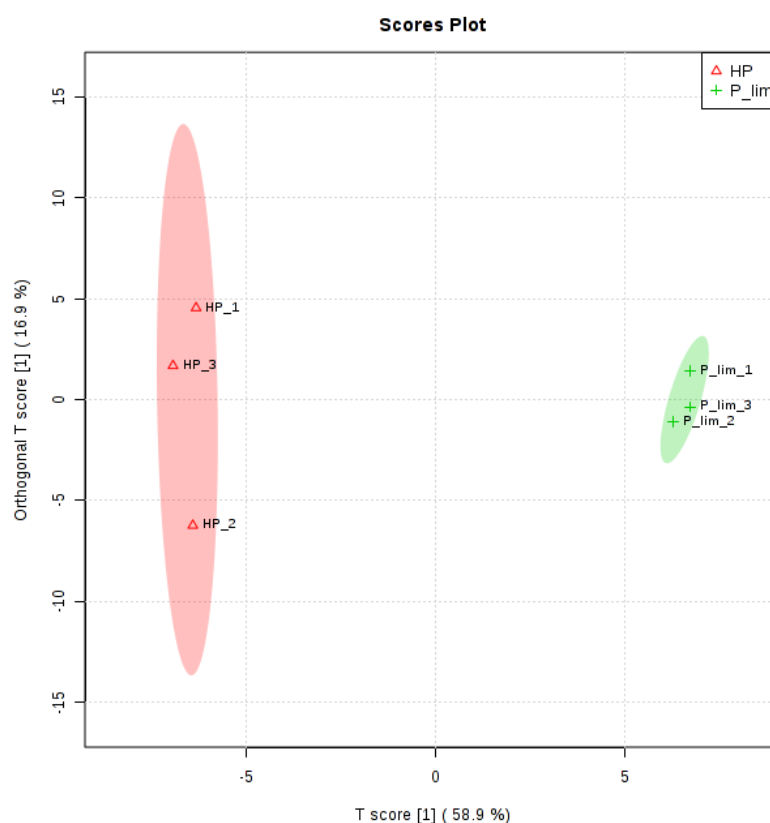


Figure 6.12 OPLS-DA of high phosphate concentration (HP) and phosphate limitation (P_lim) showing separation and clustering of the groups. Red clustering represents HP and green indicates P_lim.

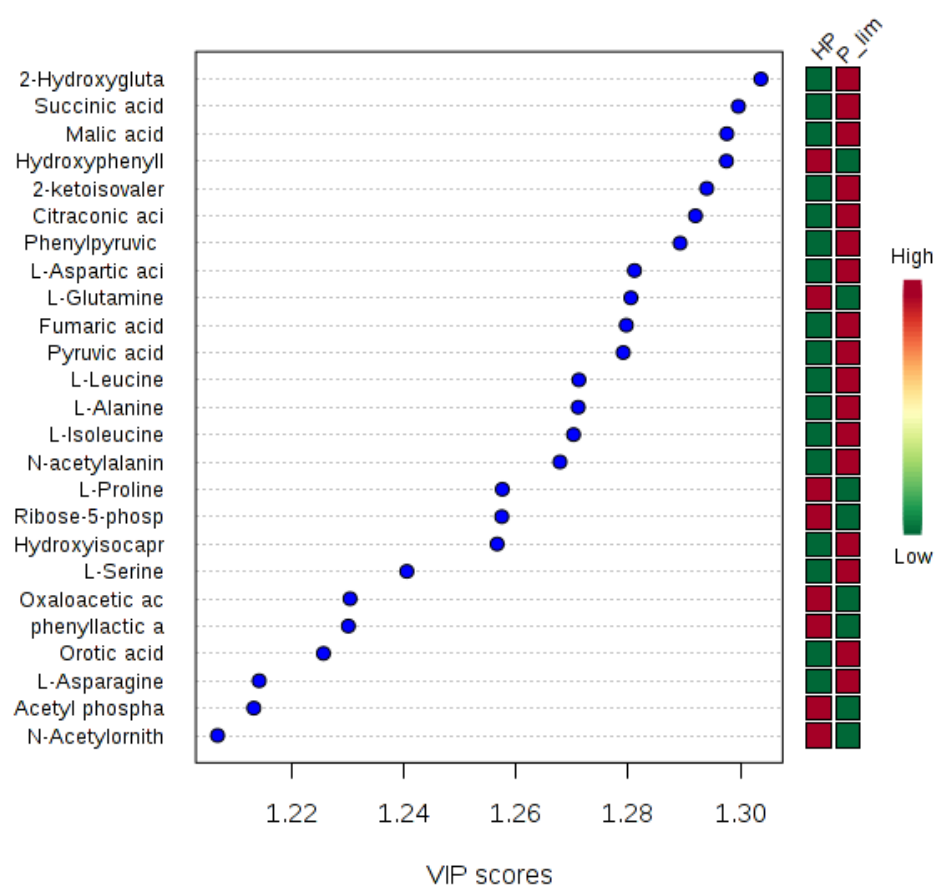


Figure 6.13 Metabolites of significance using VIP scores >1 obtained from MetaboAnalyst 3.0

Metabolic pathway analysis was conducted using the 51 metabolites showing most significance with VIP scores >1. The pathway analysis was done as already reported in previous chapters using MetaboAnalyst 3.0. An overview of the pathways is shown in Figure 6.14, which was generated by pathway impact and p values. The key metabolites shown are alanine, aspartate and glutamate metabolism, aminoacyl-tRNA biosynthesis and TCA cycle. The remaining pathways are shown in Table 6.5 with the corresponding p values <0.05 and false discovery rate (FDR). The main observation with the pathways are that most of the pathways are linked to other pathways and this makes interpretation of the results in depth. Alanine, aspartate and glutamate metabolism pathway produces the L-Aspartate, which produces oxaloacetate that feeds into the TCA cycle. Another pathway of importance was Taurine and hypotaurine metabolism

(Figure 6.15) that leads to acetyl CoA production through acetylphosphate. The acetyl-CoA produced here then links directly to pyruvate metabolism pathway, which was also highlighted as of significance.

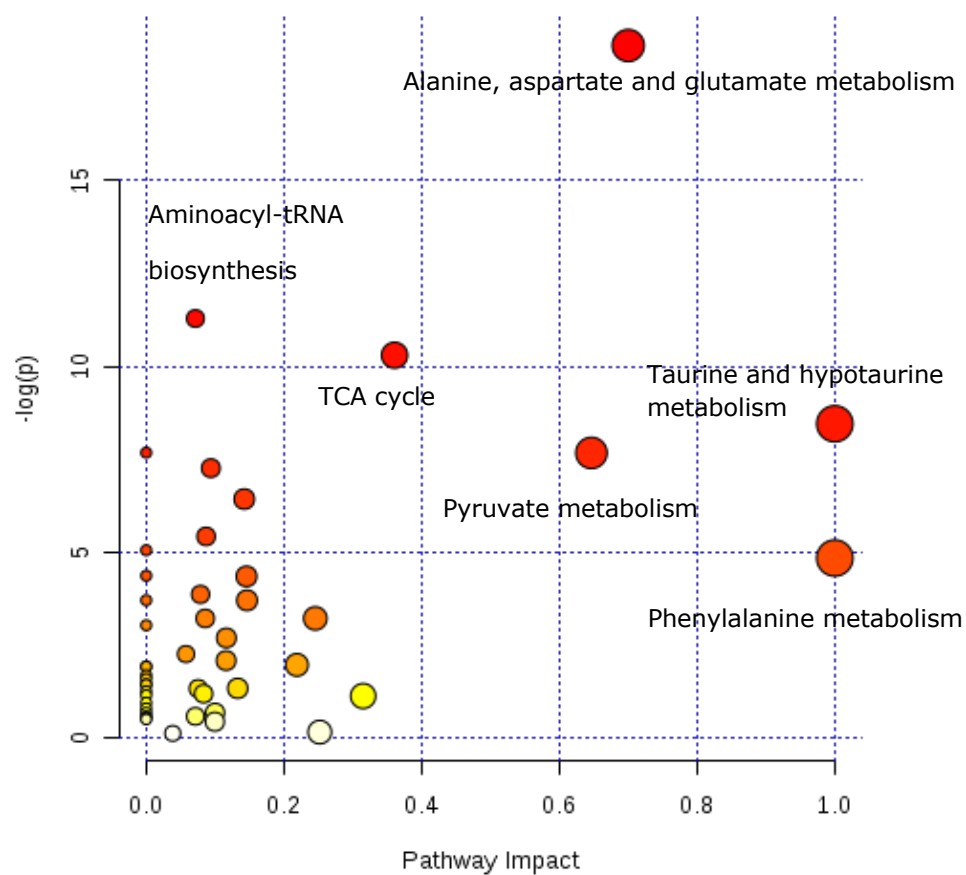


Figure 6.14 Overview of pathway analysis using MetaboAnalyst 3.0

Table 6.7 List of significant metabolic pathways affected during the phosphate limiting experiment in C. autoethanogenum using significant metabolites with VIP scores >1. This analysis was done using the pathway analysis algorithm on MeataboAnalyst 3.0.

Pathway Name	p value	FDR
Alanine, aspartate and glutamate metabolism	0	0
Aminoacyl-tRNA biosynthesis	0	0.001
Citrate cycle (TCA cycle)	0	0.001
Taurine and hypotaurine metabolism	0	0.004
Nitrogen metabolism	0	0.006
Pyruvate metabolism	0	0.006
Arginine and proline metabolism	0.001	0.008
Valine, leucine and isoleucine biosynthesis	0.002	0.016
Glyoxylate and dicarboxylate metabolism	0.004	0.039
Cyanoamino acid metabolism	0.006	0.051
Phenylalanine metabolism	0.008	0.057
Benzoate degradation via CoA ligation	0.013	0.08
Glycine, serine and threonine metabolism	0.013	0.08
Pantothenate and CoA biosynthesis	0.021	0.12
D-Alanine metabolism	0.025	0.123
Butanoate metabolism	0.025	0.123
Pyrimidine metabolism	0.04	0.178
Valine, leucine and isoleucine degradation	0.04	0.178
C5-Branched dibasic acid metabolism	0.048	0.203

6.3.7 Multivariate data analysis using SIMCA-P-14 for phosphate limitation experiment

Multivariate data analysis was conducted by generating normalised data using progenesis QI as described earlier. The normalised data was then imported into SIMCA-P14. PCA plots in Figure 6.16 shows clustering and separation between high phosphate concentration (HP) and phosphate limitation (P_lim) conditions. Stability of the analysis was evaluated by injecting pooled QC samples after every 10 injections. The PCA score plots shows QC centred between the separation of the two groups, which shows a good and stable analytical run. OPLS-DA shown in Figure 6.17 gives a good separation and clustering between the two groups. Due to the tendency to over fitting in OPLS-DA models, a further validation is required to confirm the separation is not due to over fitting. The validation was done by conducting a permutation test as shown in Figure 6.18. The Q2 value is the determinant of whether the separation of groups in the model was not due to overfitting. The Q2 value obtained was -0.303 which is an indication that the separation in the model was not due to overfitting.

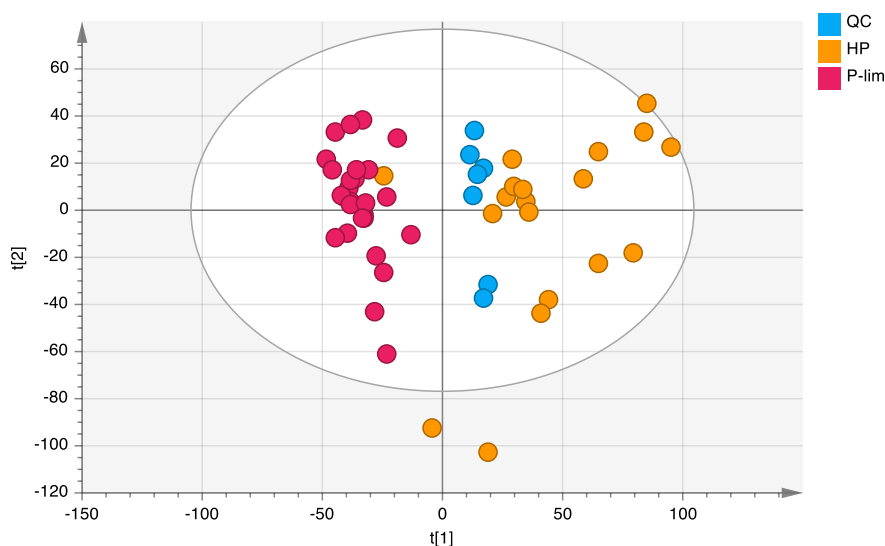


Figure 6.16 PCA score plots for all samples obtained from the phosphate limiting experiment. Pink represents phosphate limiting phase (P_lim), orange for high phosphate concentration (HP) and blue represents pooled QC. ($R^2X=0.591$ $Q^2=0.31$)

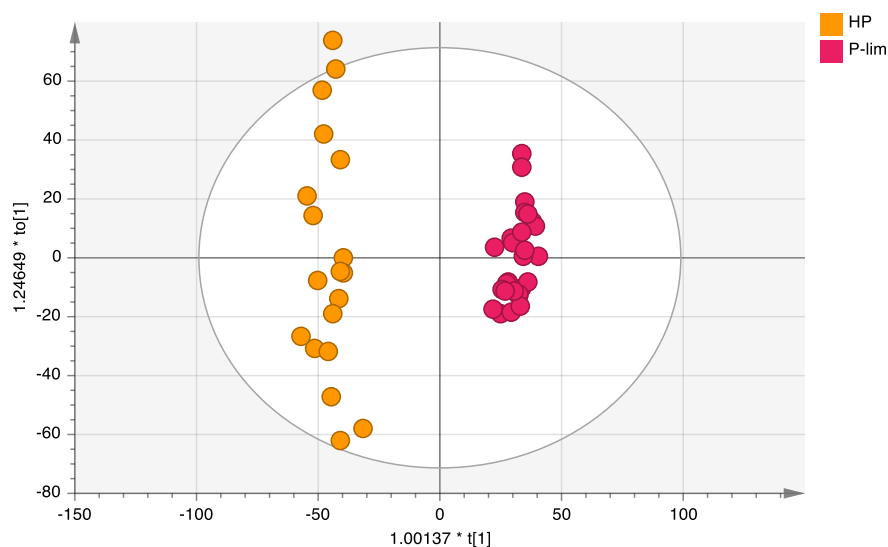


Figure 6.17 OPLS-DA score plots for all samples obtained from the phosphate limiting experiment. Pink represents phosphate limiting phase (P_lim), orange for high phosphate concentration (HP). (R^2X 0.326, R^2Y 0.981 Q^2 = 0.837)

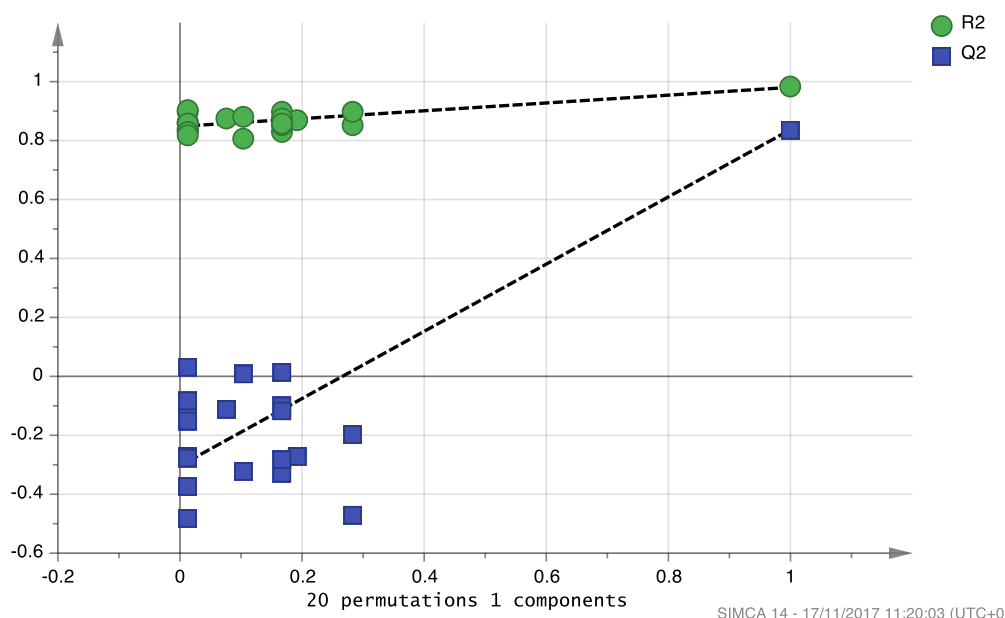


Figure 6.18 Validation of OPLS-DA models of higher phosphate and lower phosphate concentration using permutation test ($n=54$). R^2 is the model fit and Q^2 is the predictive ability of the model. The Q^2 intercept values of the regression line of randomly permuted Q^2 value was -0.303, indicating that the

model is statically sound, and the high prediction ability of the model is not due to over fitting.

6.4 Conclusion

The use of targeted metabolomics approach monitoring 133 metabolites resulted in the quantification of 88 metabolites for pantothenate limitation and 80 metabolites in phosphate limitation experiment carried out in *C. autoethanogenum*. Targeted metabolomics data confirmed reduction in acetyl-CoA concentration when pantothenate concentration was reduced. This confirms that there is no or limited production of CoA from the pantothenate and CoA biosynthesis pathway to link with the Wood-Ljungdahl pathway for the formation of acetyl-CoA. L-Aspartate intracellular levels were reduced in lower pantothenate concentration. L-aspartate is the precursor amino acid for pantothenate and CoA production confirming the effects of pantothenate concentration on L-aspartate levels.

PCA and OPLS-DA score plots from untargeted metabolomics analysis showed clear separation between the two-pantothenate conditions. The use of Compound discoverer 2.1 identified putative metabolites that contributed significantly to the changes observed. Putative metabolites were grouped in order of significance using *P* values <0.05. L-aspartate was identified as one of the metabolites of significance as observed in targeted metabolomics.

Metabolic pathway analysis using both dataset from targeted and untargeted metabolomics separately, gave some pathways in common. Both methods showed the pathways such as TCA cycle and alanine, aspartate and glutamate metabolism pathways linked to pantothenate and CoA biosynthesis were affected when pantothenate concentration was decreased in the growth limiting experiment. Untargeted metabolomics gave a good indication of metabolites of significance, but the targeted metabolomics dataset provided in depth

explanation and there is a higher degree of confidence since the metabolites have been characterised and validated.

The reduction of pantothenate concentration in the media resulted in preferred ethanol production over acetate production, which was confirmed, from metabolomics data using intracellular metabolites.

CHAPTER SEVEN

Conclusions and Future work

7. Conclusions and Future work

7.1 Method development and validation for microbial metabolomics application

In this thesis, two analytical methods were compared to show their compatibility with the analysis of the range of polar compounds of interest. The two methods were the ion pair method and HILIC and the latter was found to be more suited for the analysis of the range of polar compounds selected. The use of volatile ion pairing reagents on C₁₈ column to analyse polar compounds is well reported (Coulier et al., 2006, McGinnis et al., 2013). The results obtained shows that peak separation and peak shapes were enhanced when the ion pair reagents were used compared to when it was omitted. Ion pair reagents are less compatible with MS detection due to increased ion suppression ion positive mode ionisation. Due to the hydrophilic stationary phase of the HILIC column, it is best suited for the analysis of polar compounds compared to a standard C₁₈ column. The HILIC-MS method developed showed good separation and retention of the 133 metabolites of interest.

A HILIC-IDMS method was developed using high resolution mass spectrometry and validated for microbial targeted and untargeted metabolomics. The main challenge associated with method development is validation of the method which is a requirement for quantitative analysis and evaluating if the method is to be fit for purpose. The developed HILIC-IDMS method was validated in accordance with FDA guidelines. Precision and accuracy were evaluated for all metabolites and were within acceptable limits set out in the FDA guidelines. LOD and LOQ were also determined for each metabolite by using their signal to noise ratio of 3:1 and 10:1 respectively. Initial validation of the HILIC method involved the selection of 133 metabolites present in relevant metabolic pathways in *C. autoethanogenum*. Most reported methods do not cover this wide range of compounds (Bajad et al., 2006, Cui et al., 2013) or combine HILIC with ion pair chromatography to achieve a decent metabolites coverage but less than reported with this HILIC-IDMS method. HILIC provides an advanced separation method suitable for retention and separation of polar compounds

which does not require the use of ion pair reagents that interfere with mass spectrometry detection. The developed method was initially applied to cell extracts of *C. autoethanogenum* and peak retention and separation matched those observed in the standard mixture confirming the method was fit for purpose.

The above method can be widely applied to for other analytical purposes and it is superior to current methods published due the number of metabolites that can be detected within the 15 minutes run time which has never been investigated. This method also has the advantage of replacing methods that involve the use of ion pair reagents because separation and detection of various classes of compounds are excellent.

7.2 Considerations for quantitative metabolomics analysis

Mass spectrometry is a more sensitive technique compared to its counterparts such as nuclear magnetic resonance (NMR). The increased sensitivity and selectivity of MS comes with limitations such as ion suppression/enhancement associated with biological matrices. In order to overcome the matrix effects associated with MS detection the use of isotope labelled multiple internal standard was studied. The use of stable isotope labelled internal standard in targeted metabolomics studies is a well-known approach and has been generated from different sources using microorganisms (Kim et al., 2015). Due to the high cost of isotope labelled internal standards, it can be very costly to cover a wide range of metabolites, as was the case with this project. Also, the availability of commercially available isotope labelled internal standard can be limited and it can be difficult to find internal standard for all metabolites of interest. The use of microbes to generate isotope labelled internal standard is a well-known approach to generate internal standards from one microorganism for application in both mammalian and microbial cells. Uniformly (U) ^{13}C -labelled internal standards are the most reported to be generated from microbes such as *E. coli* (Kim et al., 2015, Lu et al., 2007) due to the ease in finding U- ^{13}C -labelled substrates such as glucose. Metabolites generated from microorganisms such *E. coli* are usually similar to metabolites from other organisms and can be used to generate U- ^{13}C -labelled internal standards although limited in the coverage of metabolites for the specific organism. The ideal scenario will be to generate the

U-¹³C-labelled internal standards from the same organism of interest, but this can be challenging due to a number of factors including cost and availability of U-¹³C-labelled substrate.

Commercially available U-¹³C-labelled cells of *Arthrospira platensis* was used initially for generation of multiple U-¹³C-labelled internal standard (Schatschneider et al., 2017). There were a good number of metabolites in *A. platensis* that were similar to metabolites in *C. autoethanogenum*, but became limited in coverage of metabolites in when the number of metabolites of interest were increased. As the metabolites of interest increased in order to have a better understanding of the mode of operation of *C. autoethanogenum*, *A. platensis* was no longer enough to be used as a source of multiple U-¹³C-labelled internal standard. Due to this limitation, closely related organisms were investigated for generation of multiple U-¹³C-labelled internal standard, because *C. autoethanogenum* grows very slowly on gas or fructose. U-¹³C-labelled CO or fructose is comparatively more expensive to U-¹³C-labelled glucose. Therefore, an organism closely related to *C. autoethanogenum* that can grow well on glucose will be a suitable alternative. *C. pasteurianum* was investigated as the closest organism to *C. autoethanogenum* that utilises glucose for growth.

Sometimes the coverage of internal standard for the set of metabolites monitored can be small because the internal standard is generated in a completely different organism compared to organism of interest (Cui et al., 2013, Kim et al., 2015). The approach taken on this project to overcome this challenge was to generate the internal standard from a closely related organism *C. pasteurianum* in order to improve the coverage for the 133 set of metabolites found in *C. autoethanogenum*. *C. pasteurianum* provided a good coverage as a U-¹³C-labelled internal standard for application in *C. autoethanogenum*. The generation of internal standard using *C. pasteurianum* can be applied to other metabolomics studies especially those on other *Clostridium* species.

Intracellular metabolite concentrations are an essential requirement for quantitative metabolomics data required for pathway modelling. In order to report intracellular concentration for pathway modelling, the cell volume of the organism or cell is required but this approach is less reported. Also, the cell

volume of *C. autoethanogenum* was never reported hence had to be determined using microscopy and flow cytometer. The cell volume of *C. autoethanogenum* was determined to be between 2-2.1 fL, which is the first time the cell volume of this organism, has been reported. With the cell volume of *C. autoethanogenum*, intracellular concentrations were obtained from different applications of *C. autoethanogenum*. Cordell also used a similar approach to report intracellular concentrations of Chinese hamster ovaries (CHO) by measuring the cell volume using confocal microscopy (Cordell et al., 2008) but less reported for microbial cells (Amador-Noguez et al., 2011, Kim et al., 2015). This is the only work where the cell volume of *C. autoethanogenum* has been measured and further used to report intracellular concentrations for metabolomics studies and potential pathway modelling.

The HILIC-IDMS method developed in this thesis has the following strengths:

- 1 The runtime of the method was optimised to 15 minutes, which is very difficult to achieve for HILIC column, which requires longer equilibrating times.
- 2 U-¹³C-labelled internal standard generated from *C. pasteurianum* provided coverage for more metabolites than have been reported previously.
- 3 The cell volume reported for *C. autoethanogenum* in this thesis can be applied for other studies involving *C. autoethanogenum*.
- 4 The validation methods used gives confidence in identified metabolites reducing the risk of false positives.
- 5 This method can resolve isobaric compounds by using the accurate mass of the compounds.

The main limitations of the HILIC-IDMS method reported in this chapter is the inability to differentiate between isomers which can be resolved by using tandem mass spectrometry approach. Isomers such as citric acid and isocitric acid can easily be differentiated according to their fragmentation patterns in MS/MS but in the full scan mode of the Orbital trap they are unresolved as the retention times and accurate masses are the same. Due to the large set of samples that were analysed in this thesis it was necessary to monitor the stability of the instrument when the run times exceeded 48 hours or plan the experiment to ensure all samples were analysed within 48 hours.

7.3 Applications of the developed HILIC-IDMS method to metabolomics analysis in *C. autoethanogenum*.

With the world's energy needs growing rapidly and with limited geological resources, there is the need to investigate sustainable and reliable ways of producing fuel and energy. The use of acetogenic bacteria has proven a viable and sustainable way of producing high value chemicals for the manufacture of fuel and energy (Abubackar et al., 2011). As these acetogens are non-pathogenic, they are easy to handle, and the safety issues associated with pathogenic bacteria are eliminated. There are some products such as acetate that can be problematic on a large-scale production and will need to be converted into useful compounds of interest. The use of metabolomics can help improve the understanding of the mechanism the organism uses to produce the end products and enable genetic manipulation of the organism to produce only the products of interest.

The validated HILIC-IDMS method was applied to different experimental approaches to induce solvent production on *C. autoethanogenum*. The method provided detailed metabolic changes observed during the acidogenic/solventogenic transitions of pH and gas shift experiments carried out using *C. autoethanogenum*. Intracellular metabolite concentrations for each phase could be reported using the determined cell volume of the organism. Metabolomics data generated were also used for pathway analysis which in both cases confirmed the TCA cycle as a major contributory pathway during the acidogenic-solventogenic transition (Amador-Noguez et al., 2011). The method was further applied to nutrient limitation in *C. autoethanogenum* as an approach to improve solvent production. The method showed changes in metabolites such as CoA and L-aspartate in the experiment that where pantothenate was limited. Metabolic pathway analysis conducted using the data also confirmed the pantothenate and CoA biosynthesis pathway as the most significant pathway affected.

The method was also applied for untargeted metabolomics analysis and there were clear separation and clustering of the different conditions using OPLS-DA score plots. PCA score plots confirmed the stability of the analysis with the pooled QC samples centred between the plots. The method proved to be versatile

for both targeted and untargeted metabolomics analysis.

In this thesis, it has been demonstrated that the HILIC -IDMS method is suitable for the analysis of a complex bioreactor samples and suitable for application to other samples. The fully validated method using IDMS used can provide the high-quality data needed to do pathway mapping. A fully quantitative method with intracellular concentrations is the best way forward for metabolomics analysis in targeted metabolomics studies. Also, the use of *Clostridium*-generated ^{13}C multiple IS metabolites is a major step forward in generating multiple internal standards for the range of metabolites investigate.

The above manipulations of conditions of *C. autoethanogenum* all induce ethanol production either by conversion of acetate to ethanol or driving carbon flux more in favour of ethanol production. On large scale industrial applications, the use of higher volumes of carbon monoxide to induce ethanol production in the gas shift approach will be very expensive. Also, *C. autoethanogenum* has less tolerance for higher concentrations of CO leading to cell death. Using the gas shift approach on an industrial scale will be more problematic than anticipated in the small-scale experiment conducted in this experiment. The approach of limiting nutrients to induce ethanol production can be a much better approach but for essential vitamins such as pantothenate, limiting these can affect cell viability over a long period of time. The reduction in pH to induce ethanol production is a more favourable approach when industrial scale application is concerned. The logistics involved in this approach is feasible and the organism will not be subject to high concentrations of CO, making the cells viable for longer since growth in *C. autoethanogenum* happens slowly.

7.4 Evaluation of transcriptomics vs. metabolomics

The integration of metabolomics data and transcriptomics data to interpret observations in cells or organisms is a well-known process. Both metabolomics and transcriptomics on their own have advantages and limitations but combining the two improves the confidence in the data generated. Transcriptomics data shows the genes that are upregulated or down regulated when conditions in the media are altered whiles metabolomics reports changes to metabolites. These

two are interlinked, hence, interpretation of data should give the same conclusion.

Data from metabolomics and transcriptomics were compared for pH shift and gas shift experiments for the one carbon metabolic pathway during the solventogenic phases. Comparison of metabolomics data and transcriptomics data confirmed changes in concentrations of metabolites corresponded to gene regulations in the transcriptomics data. Interpretation of transcriptomics data generated for these experiments were straight forward using Escher maps and easy colour coding to distinguish between upregulated and down regulated genes. Comparison of transcriptomics and metabolomics data generated for the experiments conducted improved the understanding of the mode of operation of the *C. autoethanogenum* when conditions in the media are altered.

7.5 Future Work

Further work, which was beyond the scope of this project but would be beneficial, includes quenching of the metabolic activity of the bacteria before extraction. Although this was not crucial to this organism because it grows very slowly on gas, it is important to quench metabolic activity in the cell to obtain the concentration of metabolites at that particular time point. The use of the full scan mode on the Orbital trap MS did not help distinguish between isomers; hence LC-MS/MS analysis to distinguish between isomers would have been ideal.

For untargeted metabolomics analysis, metabolites were tentatively identified that could lead to false negatives and false positives. A detailed analysis to confirm the identifications would be ideal. The approach will be to use targeted MS/MS method to improve compound identification and give more confidence in the results.

An approach to use metabolic flux analysis would be ideal to measure the rate of production and consumption of metabolites in *C. autoethanogenum*. Isotope labelled can be used but in the case of *C. autoethanogenum*, this would have

been challenging but a good thing to explore to improve the understanding already gained from how the organism uses substrate to produce useful products.

Due to time and funds constraints, optimisation of the cell volume measurements could not be achieved because further tests to confirm if the cell volume changes under different growth conditions as was done for *E. coli* in this journal (Volkmer and Heinemann, 2011).

CHAPTER EIGHT

REFERENCES

8 References

- ABRINI, J., NAVEAU, H. & NYNS, E. J. 1994. *Clostridium autoethanogenum*, sp. nov., an anaerobic bacterium that produces ethanol from carbon-monoxide. *Archives of Microbiology*, 161, 345-351.
- ABUBACKAR, H. N., BENGLSDORF, F. R., DUERRE, P., VEIGA, M. C. & KENNES, C. 2016. Improved operating strategy for continuous fermentation of carbon monoxide to fuel-ethanol by clostridia. *Applied Energy*, 169, 210-217.
- ABUBACKAR, H. N., VEIGA, M. C. & KENNES, C. 2011. Biological conversion of carbon monoxide: rich syngas or waste gases to bioethanol. *Biofuels Bioproducts & Biorefining-Biofpr*, 5, 93-114.
- ABUBACKAR, H. N., VEIGA, M. C. & KENNES, C. 2012a. Biological conversion of carbon monoxide to ethanol: effect of pH, gas pressure, reducing agent and yeast extract. *Bioresour Technol*, 114, 518-22.
- ABUBACKAR, H. N., VEIGA, M. C. & KENNES, C. 2012b. Biological conversion of carbon monoxide to ethanol: Effect of pH, gas pressure, reducing agent and yeast extract. *Bioresource Technology*, 114, 518-522.
- ABUBACKAR, H. N., VEIGA, M. C. & KENNES, C. 2015a. Carbon monoxide fermentation to ethanol by *Clostridium autoethanogenum* in a bioreactor with no accumulation of acetic acid. *Bioresource Technology*, 186, 122-127.
- ABUBACKAR, H. N., VEIGA, M. C. & KENNES, C. 2015b. Ethanol and acetic acid production from carbon monoxide in a *Clostridium* strain in batch and continuous gas-fed bioreactors. *Int J Environ Res Public Health*, 12, 1029-43.
- ABUBACKAR, H. N., VEIGA, M. C. & KENNES, C. 2015c. Ethanol and Acetic Acid Production from Carbon Monoxide in a *Clostridium* Strain in Batch and Continuous Gas-Fed Bioreactors. *International Journal of Environmental Research and Public Health*, 12, 1029-1043.
- AKERLUND, T., NORDSTROM, K. & BERNANDER, R. 1995. Analysis of cell-size and dna content in exponentially growing and stationary-phase batch cultures of *Escherichia coli*. *Journal of Bacteriology*, 177, 6791-6797.
- ALPERT, A. J. 1990. Hydrophilic-interaction chromatography for the separation of peptides, nucleic-acids and other polar compounds. *Journal of Chromatography*, 499, 177-196.
- ALSAKER, K. V. & PAPOUTSAKIS, E. T. 2005. Transcriptional program of early sporulation and stationary-phase events in *Clostridium acetobutylicum*. *Journal of Bacteriology*, 187, 7103-7118.
- AMADOR-NOGUEZ, D., BRASG, I. A., FENG, X. J., ROQUET, N. & RABINOWITZ, J. D. 2011. Metabolome Remodeling during the Acidogenic-Solventogenic Transition in *Clostridium acetobutylicum*. *Applied and Environmental Microbiology*, 77, 7984-7997.
- AMADOR-NOGUEZ, D., FENG, X. J., FAN, J., ROQUET, N., RABITZ, H. & RABINOWITZ, J. D. 2010. Systems-Level Metabolic Flux Profiling Elucidates a Complete, Bifurcated Tricarboxylic Acid Cycle in *Clostridium acetobutylicum*. *Journal of Bacteriology*, 192, 4452-4461.
- BAHL, H., ANDERSCH, W. & GOTTSCHALK, G. 1982. CONTINUOUS PRODUCTION OF ACETONE AND BUTANOL BY *CLOSTRIDIUM-ACETOBUTYLICUM* IN A 2-STAGE PHOSPHATE LIMITED CHEMOSTAT. *European Journal of Applied Microbiology and Biotechnology*, 15, 201-205.
- BAJAD, S. & SHULAEV, V. 2011. LC-MS-Based Metabolomics. In: METZ, T. O. (ed.) *Metabolic Profiling: Methods and Protocols*.

- 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, 76-88.
- BALAT, M. & BALAT, H. 2009. Recent trends in global production and utilization of bio-ethanol fuel. *Applied Energy*, 86, 2273-2282.
- BAUMAN, J. N. & DING, C. 2015. METABOLITE PROFILING AND IDENTIFICATION IN CRYOPRESERVED HUMAN HEPATOCYTES USING THERMO SCIENTIFIC COMPOUND DISCOVERER SOFTWARE. *Drug Metabolism Reviews*, 47, 200-200.
- BECKMANN, M., PARKER, D., ENOT, D. P., DUVAL, E. & DRAPER, J. 2008. High-throughput, nontargeted metabolite fingerprinting using nominal mass flow injection electrospray mass spectrometry. *Nature Protocols*, 3, 486-504.
- BEGOU, O., GIKA, H. G., WILSON, I. D. & THEODORIDIS, G. 2017. Hyphenated MS-based targeted approaches in metabolomics. *Analyst*, 142, 3079-3100.
- BEJERANO-SAGIE, M., OPPENHEIMER-SHAANAN, Y., BERLATZKY, I., ROUVINSKI, A., MEYEROVICH, M. & BEN-YEHUDA, S. 2006. A checkpoint protein that scans the chromosome for damage at the start of sporulation in *Bacillus subtilis*. *Cell*, 125, 679-690.
- BENGELSDORF, F. R., STRAUB, M. & DUERRE, P. 2013. Bacterial synthesis gas (syngas) fermentation. *Environmental Technology*, 34, 1639-1651.
- BERG, T. & STRAND, D. H. 2011. C-13 labelled internal standards-A solution to minimize ion suppression effects in liquid chromatography-tandem mass spectrometry analyses of drugs in biological samples? *Journal of Chromatography A*, 1218, 9366-9374.
- BERLATZKY, I. A., ROUVINSKI, A. & BEN-YEHUDA, S. 2008. Spatial organization of a replicating bacterial chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 14136-14140.
- BERTINI, I., HU, X. Y. & LUCHINAT, C. 2014. Global metabolomics characterization of bacteria: pre-analytical treatments and profiling. *Metabolomics*, 10, 241-249.
- BIAGI, S., GHIMENTI, S., ONOR, M. & BRAMANTI, E. 2012. Simultaneous determination of lactate and pyruvate in human sweat using reversed-phase high-performance liquid chromatography: a noninvasive approach. *Biomedical Chromatography*, 26, 1408-1415.
- BIEBL, H. 2001. Fermentation of glycerol by *Clostridium pasteurianum* - batch and continuous culture studies. *Journal of Industrial Microbiology & Biotechnology*, 27, 18-26.
- BLOEM, J., VANMULLEM, D. K. & BOLHUIS, P. R. 1992. Microscopic counting and calculation of species abundances and statistics in real-time with an ms-dos personal-computer, applied to bacteria in soil smears. *Journal of Microbiological Methods*, 16, 203-213.
- BLOEM, J., VENINGA, M. & SHEPHERD, J. 1995. Fully-automatic determination of soil bacterium numbers, cell volumes, and frequencies of dividing cells by confocal laser-scanning microscopy and image-analysis. *Applied and Environmental Microbiology*, 61, 926-936.
- BOLTEN, C. J., KIEFER, P., LETISSE, F., PORTAIS, J.-C. & WITTMANN, C. 2007. Sampling for metabolome analysis of microorganisms. *Analytical Chemistry*, 79, 3843-3849.
- BOUVIER, T., TROUSSELLIER, M., ANZIL, A., COURTIES, C. & SERVAIS, P. 2001. Using light scatter signal to estimate bacterial biovolume by flow cytometry. *Cytometry*, 44, 188-194.

- BOYE, E., STEEN, H. B. & SKARSTAD, K. 1983. Flow-cytometry of bacteria - a promising tool in experimental and clinical microbiology. *Journal of General Microbiology*, 129, 973-980.
- BRAMONO, S. E., LAM, Y. S., ONG, S. L. & HE, J. 2011. A mesophilic *Clostridium* species that produces butanol from monosaccharides and hydrogen from polysaccharides. *Bioresource Technology*, 102, 9558-9563.
- BRINDLE, J. T., ANTTI, H., HOLMES, E., TRANTER, G., NICHOLSON, J. K., BETHELL, H. W. L., CLARKE, S., SCHOFIELD, P. M., MCKILLIGIN, E., MOSEDALE, D. E. & GRAINGER, D. J. 2002. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using H-1-NMR-based metabolomics. *Nature Medicine*, 8, 1439-1444.
- BRZEZINKA, H., DALLAKIAN, P. & BUDZIKIEWICZ, H. 1999. Thin-layer chromatography and mass spectrometry for screening of biological samples for drugs and metabolites. *Jpc-Journal of Planar Chromatography-Modern Tlc*, 12, 96-108.
- BUCHHOLZ, A., HURLEBAUS, J., WANDREY, C. & TAKORS, R. 2002. Metabolomics: quantification of intracellular metabolite dynamics. *Biomolecular Engineering*, 19, 5-15.
- BUESCHER, J. M., CZERNIK, D., EWALD, J. C., SAUER, U. & ZAMBONI, N. 2009. Cross-Platform Comparison of Methods for Quantitative Metabolomics of Primary Metabolism. *Analytical Chemistry*, 81, 2135-2143.
- BUESCHER, J. M., MOCO, S., SAUER, U. & ZAMBONI, N. 2010. Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectrometry Method for Fast and Robust Quantification of Anionic and Aromatic Metabolites. *Analytical Chemistry*, 82, 4403-4412.
- BYLESJO, M., RANTALAINEN, M., CLOAREC, O., NICHOLSON, J. K., HOLMES, E. & TRYGG, J. 2006. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics*, 20, 341-351.
- BYLUND, D., NORSTROM, S. H., ESSEN, S. A. & LUNDSTROM, U. S. 2007. Analysis of low molecular mass organic acids in natural waters by ion exclusion chromatography tandem mass spectrometry. *Journal of Chromatography A*, 1176, 89-93.
- CANELAS, A. B., RAS, C., TEN PIERICK, A., VAN DAM, J. C., HEIJNEN, J. J. & VAN GULIK, W. M. 2008. Leakage-free rapid quenching technique for yeast metabolomics. *Metabolomics*, 4, 226-239.
- CARNICER, M., CANELAS, A. B., TEN PIERICK, A., ZENG, Z., VAN DAM, J., ALBIOL, J., FERRER, P., HEIJNEN, J. J. & VAN GULIK, W. 2012. Development of quantitative metabolomics for *Pichia pastoris*. *Metabolomics*, 8, 284-298.
- CARNICER, M., VIEIRA, G., BRAUTASET, T., PORTAIS, J. C. & HEUX, S. 2016. Quantitative metabolomics of the thermophilic methylotroph *Bacillus methanolicus*. *Microbial Cell Factories*, 15.
- CASCANTE, M. & MARIN, S. 2008. Metabolomics and fluxomics approaches. *Essays in Biochemistry: Systems Biology, Vol 45*, 45, 67-81.
- CHALLAMALLA, P., GHOSH, S., PARTHIBAN, N., RAO, K. N. V. & BANJI, D. 2012. Negative ion mode mass spectrometry- an overview. *Journal of Chemical, Biological and Physical Sciences*, 2, 1462-1471.
- CHAMBERS, E., WAGROWSKI-DIEHL, D. M., LU, Z. & MAZZEO, J. R. 2007. Systematic and comprehensive strategy for reducing matrix effects in LC/MS/MS analyses. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 852, 22-34.
- CHEN, M. J., RAO, R. S. P., ZHANG, Y. M., ZHONG, C. X. & THELEN, J. J. 2014. A modified data normalization method for GC-MS-based metabolomics to minimize batch variation. *Springerplus*, 3.

- CHUANG, C. K., WANG, T. J., YEUNG, C. Y., LIN, D. S., LIN, H. Y., LIU, H. L., HO, H. T., HSIEH, W. S. & LIN, S. P. 2009. A method for lactate and pyruvate determination in filter-paper dried blood spots. *Journal of Chromatography A*, 1216, 8947-8952.
- CLOAREC, O., DUMAS, M. E., CRAIG, A., BARTON, R. H., TRYGG, J., HUDSON, J., BLANCHER, C., GAUGUIER, D., LINDON, J. C., HOLMES, E. & NICHOLSON, J. 2005. Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic H-1 NMR data sets. *Analytical Chemistry*, 77, 1282-1289.
- COEN, M. & KUCHEL, P. W. 2004. Metabonomics based on NMR spectroscopy. *Chem Aust.*, 71, 13-17.
- CORDELL, R. L., HILL, S. J., ORTORI, C. A. & BARRETT, D. A. 2008. Quantitative profiling of nucleotides and related phosphate-containing metabolites in cultured mammalian cells by liquid chromatography tandem electrospray mass spectrometry. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 871, 115-124.
- COTTER, J. L., CHINN, M. S. & GRUNDEN, A. M. 2009a. Ethanol and acetate production by *Clostridium ljungdahlii* and *Clostridium autoethanogenum* using resting cells. *Bioprocess and Biosystems Engineering*, 32, 369-380.
- COTTER, J. L., CHINN, M. S. & GRUNDEN, A. M. 2009b. Influence of process parameters on growth of *Clostridium ljungdahlii* and *Clostridium autoethanogenum* on synthesis gas. *Enzyme and Microbial Technology*, 44, 281-288.
- COUCHMAN, L. & MORGAN, P. E. 2011. LC-MS in analytical toxicology: some practical considerations. *Biomedical Chromatography*, 25, 100-123.
- COULIER, L., BAS, R., JESPERSEN, S., VERHEIJ, E., VAN DER WERF, M. J. & HANKEMEIER, T. 2006. Simultaneous quantitative analysis of metabolites using ion-pair liquid chromatography - Electrospray ionization mass spectrometry. *Analytical Chemistry*, 78, 6573-6582.
- CREEK, D. J., JANKEVICS, A., BREITLING, R., WATSON, D. G., BARRETT, M. P. & BURGESS, K. E. V. 2011. Toward Global Metabolomics Analysis with Hydrophilic Interaction Liquid Chromatography-Mass Spectrometry: Improved Metabolite Identification by Retention Time Prediction. *Analytical Chemistry*, 83, 8703-8710.
- CROWN, S. B., INDURTHI, D. C., AHN, W. S., CHOI, J., PAPOUTSAKIS, E. T. & ANTONIEWICZ, M. R. 2011. Resolving the TCA cycle and pentose-phosphate pathway of *Clostridium acetobutylicum* ATCC 824: Isotopomer analysis, in vitro activities and expression analysis. *Biotechnology Journal*, 6, 300-305.
- CUBBON, S., BRADBURY, T., WILSON, J. & THOMAS-OATES, J. 2007. Hydrophilic interaction chromatography for mass spectrometric metabonomic studies of urine. *Analytical Chemistry*, 79, 8911-8918.
- CUI, J. T., ZHANG, J. T., ZHU, X. S., BAI, F. L., FENG, Y. G., GUAN, W. N. & CUI, Q. 2013. Separation and quantification of water-soluble cellular metabolites in *Clostridium thermocellum* using liquid chromatography-isotope dilution tandem mass spectrometry. *Analytical Letters*, 46, 2767-2786.
- DAMS, R., HUESTIS, M. A., LAMBERT, W. E. & MURPHY, C. M. 2003. Matrix effect in bio-analysis of illicit drugs with LC-MS/MS: Influence of ionization type, sample preparation, and biofluid. *Journal of the American Society for Mass Spectrometry*, 14, 1290-1294.
- DANIELL, J., KOEPKE, M. & SIMPSON, S. D. 2012. Commercial Biomass Syngas Fermentation. *Energies*, 5, 5372-5417.
- DAS, D. N., FENG, F. W. & TOENNIES, G. 1966. COENZYME A AND PANTOTHENATE METABOLISM IN BACTERIAL GROWTH. *Federation Proceedings*, 25, 217-&.

- DELCHER, A. L., PHILLIPPY, A., CARLTON, J. & SALZBERG, S. L. 2002. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Research*, 30, 2478-2483.
- DERVIEUX, T., LEIN, D. O., MARCELLETTI, J., PISCHEL, K., SMITH, K., WALSH, M. & RICHERSON, R. 2003. HPLC determination of erythrocyte methotrexate polyglutamates after low-dose methotrexate therapy in patients with rheumatoid arthritis. *Clinical Chemistry*, 49, 1632-1641.
- DETTMER, K., ARONOV, P. A. & HAMMOCK, B. D. 2007. Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews*, 26, 51-78.
- DOLE, M., MACK, L. L. & HINES, R. L. 1968. MOLECULAR BEAMS OF MACROIONS. *Journal of Chemical Physics*, 49, 2240-&.
- DONG-HYUN KIM, F. A., RAINER BREITLING, KARL E. BURGESS, MICHAEL P. BARRETT 2015. LC-MS-based absolute metabolite quantification: Application to metabolic flux measurement in trypanosomes. *Metabolomics*.
- DOUMA, R. D., VERHEIJEN, P. J. T., DE LAAT, W., HEIJNEN, J. J. & VAN GULIK, W. M. 2010. Dynamic Gene Expression Regulation Model for Growth and Penicillin Production in *Penicillium chrysogenum*. *Biotechnology and Bioengineering*, 106, 608-618.
- DRAKE, H. L. & KUSEL, K. 2005. *Acetogenic clostridia*, Boca Raton, FL, CRC Press.
- DRAKE, H. L., KUSEL, K. & MATHIES, A. C. 2006. *Acetogenic Prokaryotes*, New York, (ecophysiology and biochemistry) Springer.
- DUNN, W. B. & WINDER, C. L. 2011. Sample preparation related to the intracellular metabolome of yeast: Methods for quenching, extraction, and metabolite quantitation. *Methods in Enzymology, Vol 500: Methods in Systems Biology*, 500, 277-297.
- DURRE, P. 2017. Gas fermentation - a biotechnological solution for today's challenges. *Microbial Biotechnology*, 10, 14-16.
- EDWARDS, J. L., EDWARDS, R. L., REID, K. R. & KENNEDY, R. T. 2007. Effect of decreasing column inner diameter and use of off-line two-dimensional chromatography on metabolite detection in complex mixtures. *Journal of Chromatography A*, 1172, 127-134.
- ERIKSSON, L., JOHANSSON, E., KETTANEH-WOLD, N. J., TRYGG, C. W. & WOLD, S. 2006a. Multi- and megavariable data analysis: advanced applications and method extensions. *Umea, Umetrics AB*.
- ERIKSSON, L., JOHANSSON, E., KETTANEH-WOLD, N. J., TRYGG, C. W. & WOLD, S. 2006b. Multi- and megavariable data analysis: Basic principles and applications. *Umea, Umetrics AB*.
- FAIRCHILD, J. N., HORVATH, K., GOODING, J. R., CAMPAGNA, S. R. & GUIOCHON, G. 2010. Two-dimensional liquid chromatography/mass spectrometry/mass spectrometry separation of water-soluble metabolites. *Journal of Chromatography A*, 1217, 8161-8166.
- FENN, J. B., MANN, M., MENG, C. K., WONG, S. F. & WHITEHOUSE, C. M. 1989. ELECTROSPRAY IONIZATION FOR MASS-SPECTROMETRY OF LARGE BIOMOLECULES. *Science*, 246, 64-71.
- FERREIRA, V., CULLERE, L., LOSCOS, N. & CACHO, J. 2006. Critical aspects of the determination of pentafluorobenzyl derivatives of aldehydes by gas chromatography with electron-capture or mass spectrometric detection - Validation of an optimized strategy for the determination of oxygen-related odor-active aldehydes in wine. *Journal of Chromatography A*, 1122, 255-265.
- FLACH, B., BENDZ, K., KRAUTGARTNER, R. & LIEBERZ, S. 2013. EU Biofuels Annual 2013. EU-27.
- FLORES, P., HELLIN, P. & FENOLL, J. 2012. Determination of organic acids in fruits and vegetables by liquid chromatography with tandem-mass spectrometry. *Food Chemistry*, 132, 1049-1054.

- FRY, J. C. 1990. Direct methods and biomass estimation. *Methods in Microbiology*, 22, 41-85.
- GADDY, J. L., ARORA, D. K., KO, C. W., PHILLIPS, R. J., BASU, R., WIKSTRIM, C. V. & CLAUSEN, E. C. 2003. Methods for Increasing the Production of Ethanol from Microbial Fermentation. *US20030211585*.
- GANGL, E. T., ANNAN, M., SPOONER, N. & VOUIROS, P. 2001. Reduction of signal suppression effects in ESI-MS using a nanosplitting device. *Analytical Chemistry*, 73, 5635-5644.
- GARRATT, L. C., A., O. C., A., T. G., F., S., J., B. M. & A., B. D. 2005. Comprehensive metabolic profiling of mono- an polyglutamated folates and their precursors in plant and animal tissue using liquid chromatography/negative ion electrospray ionisation tandem mass spectrometry. *Rapid communications in mass spectrometry*, 19, 2390-2398.
- GAVAGHAN, C. L., WILSON, I. D. & NICHOLSON, J. K. 2002. Physiological variation in metabolic phenotyping and functional genomic studies: use of orthogonal signal correction and PLS-DA. *Febs Letters*, 530, 191-196.
- GIKA, H. G., THEODORIDIS, G. A., VRHOVSEK, U. & MATTIVI, F. 2012. Quantitative profiling of polar primary metabolites using hydrophilic interaction ultrahigh performance liquid chromatography-tandem mass spectrometry. *Journal of Chromatography A*, 1259, 121-127.
- GIKA, H. G., THEODORIDIS, G. A. & WILSON, I. D. 2008. Hydrophilic interaction and reversed-phase ultra-performance liquid chromatography TOF-MS for metabonomic analysis of Zucker rat urine. *Journal of Separation Science*, 31, 1598-1608.
- GOSETTI, F., MAZZUCCO, E., ZAMPIERI, D. & GENNARO, M. C. 2010. Signal suppression/enhancement in high-performance liquid chromatography tandem mass spectrometry. *Journal of Chromatography A*, 1217, 3929-3937.
- GREEN, E. M., BOYNTON, Z. L., HARRIS, L. M., RUDOLPH, F. B., PAPOUTSAKIS, E. T. & BENNETT, G. N. 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. *Microbiology-Sgm*, 142, 2079-2086.
- GRIFFIN, D. W. & SCHULTZ, M. A. 2012. Fuel and Chemical Products from Biomass Syngas: A Comparison of Gas Fermentation to Thermochemical Conversion Routes. *Environmental Progress & Sustainable Energy*, 31, 219-224.
- GRIFFITHS, W. J., KOAL, T., WANG, Y. Q., KOHL, M., ENOT, D. P. & DEIGNER, H. P. 2010. Targeted Metabolomics for Biomarker Discovery. *Angewandte Chemie-International Edition*, 49, 5426-5445.
- GUO, K., JI, C. & LI, L. 2007. Stable-isotope dimethylation labeling combined with LC-ESI MS for quantification of amine-containing metabolites in biological samples. *Analytical Chemistry*, 79, 8631-8638.
- GUO, K., PENG, J., ZHOU, R. K. & LI, L. 2011. Ion-pairing reversed-phase liquid chromatography fractionation in combination with isotope labeling reversed-phase liquid chromatography-mass spectrometry for comprehensive metabolome profiling. *Journal of Chromatography A*, 1218, 3689-3694.
- GUO, Y. & GAIKI, S. 2005. Retention behavior of small polar compounds on polar stationary phases in hydrophilic interaction chromatography. *Journal of Chromatography A*, 1074, 71-80.
- GUO, Y., XU, J., ZHANG, Y., XU, H., YUAN, Z. & LI, D. 2010. Medium optimization for ethanol production with *Clostridium autoethanogenum* with carbon monoxide as sole carbon source. *Bioresource Technology*, 101, 8784-8789.
- GUPTA, R. S. & GAO, B. 2009. Phylogenomic analyses of clostridia and identification of novel protein signatures that are specific to the genus

- Clostridium sensu stricto* (cluster I). *International Journal of Systematic and Evolutionary Microbiology*, 59, 285-294.
- HAANDEL, L. V., BECKER, M. L., LEEDER, J. S., AND, T. D. W. & STOBAUGH, J. F. 2009. A novel high-performance liquid chromatography/mass spectrometry method for improved selective and sensitive measurement of methotrexate polyglutamation status in human red blood cells. *Rapid communications in mass spectrometry*, 23, 3693-3702.
- HAGGARTY, J. & BURGESS, K. E. V. 2017. Recent advances in liquid and gas chromatography methodology for extending coverage of the metabolome. *Current Opinion in Biotechnology*, 43, 77-85.
- HALLIDAY, N. M., HARDIE, K. R., WILLIAMS, P., WINZER, K. & BARRETT, D. A. 2010. Quantitative liquid chromatography-tandem mass spectrometry profiling of activated methyl cycle metabolites involved in LuxS-dependent quorum sensing in *Escherichia coli*. *Analytical Biochemistry*, 403, 20-29.
- HAN, J., DATLA, R., CHAN, S. & BORCHERS, C. H. 2009. Mass spectrometry-based technologies for high-throughput metabolomics. *Bioanalysis*, 1, 1665-1684.
- HANSON, A. D. & ROJE, S. 2001. One-carbon metabolism in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 52, 119-137.
- HARTMANN, C., SMEYERS-VERBEKE, J., MASSART, D. L. & MCDOWALL, R. D. 1998. Validation of bioanalytical chromatographic methods. *Journal of Pharmaceutical and Biomedical Analysis*, 17, 193-218.
- HELDAL, M., NORLAND, S., BRATBAK, G. & RIEMANN, B. 1994. Determination of bacterial-cell number and cell-volume by means of flow-cytometry, transmission electron-microscopy, and epifluorescence microscopy. *Journal of Microbiological Methods*, 20, 255-263.
- HELDAL, M., NORLAND, S. & TUMYR, O. 1985. X-ray microanalytic method for measurement of dry-matter and elemental content of individual bacteria. *Applied and Environmental Microbiology*, 50, 1251-1257.
- HELLER, D. N. 2007. Ruggedness testing of quantitative atmospheric pressure ionization mass spectrometry methods: the effect of co-injected matrix on matrix effects. *Rapid Communications in Mass Spectrometry*, 21, 644-652.
- HEUBERGER, A. L., ROBISON, F. M., LYONS, S. M. A., BROECKLING, C. D. & PRENNI, J. E. 2014. Evaluating plant immunity using mass spectrometry-based metabolomics workflows. *Frontiers in Plant Science*, 5.
- HEUX, S., MEYNIAL-SALLES, I., O'DONOHUE, M. J. & DUMON, C. 2015. White biotechnology: State of the art strategies for the development of biocatalysts for biorefining. *Biotechnology Advances*, 33, 1653-1670.
- HEWAVITHARANA, A. K. 2011. Matrix matching in liquid chromatography-mass spectrometry with stable isotope labelled internal standards-Is it necessary? *Journal of Chromatography A*, 1218, 359-361.
- HIDALGO, F. J., NAVARRO, J. L., DELGADO, R. M. & ZAMORA, R. 2013. Determination of alpha-keto acids in pork meat and Iberian ham via tandem mass spectrometry. *Food Chemistry*, 140, 183-188.
- HU, S., WANG, J. H., JI, E. H., CHRISTISON, T., LOPEZ, L. & HUANG, Y. Y. 2015. Targeted Metabolomic Analysis of Head and Neck Cancer Cells Using High Performance Ion Chromatography Coupled with a Q Exactive HF Mass Spectrometer. *Analytical Chemistry*, 87, 6371-6379.
- HUANG, Y., TIAN, Y., ZHANG, Z. & PENG, C. 2012. A HILIC-MS/MS method for the simultaneous determination of seven organic acids in rat urine as biomarkers of exposure to realgar. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 905, 37-42.
- HUEGE, J., KRALL, L., STEINHAUSER, M. C., GIAVALISCO, P., RIPPKA, R., DE MARSAC, N. T. & STEINHAUSER, D. 2011. Sample amount

- alternatives for data adjustment in comparative cyanobacterial metabolomics. *Analytical and Bioanalytical Chemistry*, 399, 3503-3517.
- HUMPHREYS, C. M., MCLEAN, S., SCHATTSCHNEIDER, S., MILLAT, T., HENSTRA, A. M., ANNAN, F. J., BREITKOPF, R., PANDER, B., PIATEK, P., ROWE, P., WICHLACZ, A. T., WOODS, C., NORMAN, R., BLOM, J., GOESMAN, A., HODGMAN, C., BARRETT, D., THOMAS, N. R., WINZER, K. & MINTON, N. P. 2015. Whole genome sequence and manual annotation of *Clostridium autoethanogenum*, an industrially relevant bacterium. *Bmc Genomics*, 16.
- IBBA, M. & SOLL, D. 2000. Aminoacyl-tRNA synthesis. *Annual Review of Biochemistry*, 69, 617-650.
- ISOM, C. E., NANNY, M. A. & TANNER, R. S. 2015. Improved conversion efficiencies for n-fatty acid reduction to primary alcohols by the solventogenic acetogen "*Clostridium ragsdalei*". *Journal of Industrial Microbiology & Biotechnology*, 42, 29-38.
- JACKOWSKI, S. & ROCK, C. O. 1986. Consequences of reduced intracellular coenzyme A content in *Escherichia coli*. *Journal of Bacteriology*, 166, 866-871.
- JAITZ, L., MUELLER, B., KOELLENSPERGER, G., HUBER, D., OBURGER, E., PUSCHENREITER, M. & HANN, S. 2011. LC-MS analysis of low molecular weight organic acids derived from root exudation. *Analytical and Bioanalytical Chemistry*, 400, 2587-2596.
- JANSSEN, H., DORING, C., EHRENREICH, A., VOIGT, B., HECKER, M., BAHL, H. & FISCHER, R. J. 2010. A proteomic and transcriptional view of acidogenic and solventogenic steady-state cells of *Clostridium acetobutylicum* in a chemostat culture. *Applied Microbiology and Biotechnology*, 87, 2209-2226.
- JEMAL, M., OUYANG, Z. & XIA, Y.-Q. 2010. Systematic LC-MS/MS bioanalytical method development that incorporates plasma phospholipids risk avoidance, usage of incurred sample and well thought-out chromatography. *Biomedical Chromatography*, 24, 2-19.
- JESSOME, L. L. & VOLMER, D. A. 2006. Ion suppression: A major concern in mass spectrometry. *Lc Gc North America*, 24, 498-+.
- JONES, S. W., PAREDES, C. J., TRACY, B., CHENG, N., SILLERS, R., SENGHER, R. S. & PAPOUTSAKIS, E. T. 2008. The transcriptional program underlying the physiology of clostridial sporulation. *Genome Biology*, 9.
- KHOURY, S., EL BANNA, N., TFAILI, S. & CHAMINADE, P. 2016. A study of inter-species ion suppression in electrospray ionization-mass spectrometry of some phospholipid classes. *Analytical and Bioanalytical Chemistry*, 408, 1453-1465.
- KIEFER, P., PORTAIS, J.-C. & VORHOLT, J. A. 2008. Quantitative metabolome analysis using liquid chromatography-high-resolution mass spectrometry. *Analytical Biochemistry*, 382, 94-100.
- KIM, D. H., ACHCAR, F., BREITLING, R., BURGESS, K. E. & BARRETT, M. P. 2015. LC-MS-based absolute metabolite quantification: application to metabolic flux measurement in trypanosomes. *Metabolomics*, 11, 1721-1732.
- KINGDON, K. H. 1923. A method for the neutralization of electron space charge by positive ionization at very low gas pressures. *Physical Review*, 21, 408-418.
- KLASSON, K. T., ACKERSON, M. D., CLAUSEN, E. C. & GADDY, J. L. 1993. Biological conversion of coal and coal-derived synthesis gas. *Fuel*, 72, 1673-1678.
- KLEPACKI, J., KLAWITTER, J., THURMAN, J. M. & CHRISTIANS, U. 2015. A high-performance liquid chromatography - tandem mass spectrometry - based targeted metabolomics kidney dysfunction marker panel in human urine. *Clinica Chimica Acta*, 446, 43-53.

- KNIGHT, R. D. 1981. STORAGE OF IONS FROM LASER-PRODUCED PLASMAS. *Applied Physics Letters*, 38, 221-223.
- KOEPKE, M., GERTH, M. L., MADDOCK, D. J., MUELLER, A. P., LIEW, F., SIMPSON, S. D. & PATRICK, W. M. 2014. Reconstruction of an Acetogenic 2,3-Butanediol Pathway Involving a Novel NADPH-Dependent Primary-Secondary Alcohol Dehydrogenase. *Applied and Environmental Microbiology*, 80, 3394-3403.
- KOEPKE, M., HELD, C., HUJER, S., LIESEGANG, H., WIEZER, A., WOLLHERR, A., EHRENREICH, A., LIEBL, W., GOTTSCHALK, G. & DUERRE, P. 2010. Clostridium ljungdahlii represents a microbial production platform based on syngas. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 13087-13092.
- KOPKE, M., MIHALCEA, C., BROMLEY, J. C. & SIMPSON, S. D. 2011a. Fermentative production of ethanol from carbon monoxide. *Current Opinion in Biotechnology*, 22, 320-325.
- KOPKE, M., MIHALCEA, C., LIEW, F. M., TIZARD, J. H., ALI, M. S., CONOLLY, J. J., AL-SINAWI, B. & SIMPSON, S. D. 2011b. 2,3-Butanediol Production by Acetogenic Bacteria, an Alternative Route to Chemical Synthesis, Using Industrial Waste Gas. *Applied and Environmental Microbiology*, 77, 5467-5475.
- KUNDIYANA, D. K., HUHNKE, R. L. & WILKINS, M. R. 2011. Effect of nutrient limitation and two-stage continuous fermentor design on productivities during "Clostridium ragsdalei" syngas fermentation. *Bioresource Technology*, 102, 6058-6064.
- LATIF, H., ZEIDAN, A. A., NIELSEN, A. T. & ZENGLER, K. 2014. Trash to treasure: production of biofuels and commodity chemicals via syngas fermenting microorganisms. *Current Opinion in Biotechnology*, 27, 79-87.
- LEHMANN, D. & LUTKE-EVERSLOH, T. 2011. Switching Clostridium acetobutylicum to an ethanol producer by disruption of the butyrate/butanol fermentative pathway. *Metabolic Engineering*, 13, 464-473.
- LI, J., VON PFOESTL, V., ZALDIVAR, D., ZHANG, X., LOGOTHETIS, N. & RAUCH, A. 2012. Measuring multiple neurochemicals and related metabolites in blood and brain of the rhesus monkey by using dual microdialysis sampling and capillary hydrophilic interaction chromatography-mass spectrometry. *Analytical and Bioanalytical Chemistry*, 402, 2545-2554.
- LI, X., LONG, D., JI, J., YANG, W., ZENG, Z., GUO, S., JI, Z., QI, G. & CHEN, S. 2013. Sample preparation for the metabolomics investigation of poly-gamma-glutamate-producing Bacillus licheniformis by GC-MS. *Journal of Microbiological Methods*, 94, 61-67.
- LINK, H., BUESCHER, J. M. & SAUER, U. 2012. Targeted and quantitative metabolomics in bacteria. In: HARWOOD, C. & WIPAT, A. (eds.) *Systems Biology of Bacteria*.
- LIU, K., ATIYEH, H. K., TANNER, R. S., WILKINS, M. R. & HUHNKE, R. L. 2012. Fermentative production of ethanol from syngas using novel moderately alkaliphilic strains of Alkalibaculum bacchi. *Bioresource Technology*, 104, 336-341.
- LJUNGDAHL, L., IRION, E. & WOOD, H. G. 1966. ROLE OF CORRINOIDS IN TOTAL SYNTHESIS OF ACETATE FROM CO₂ BY CLOSTRIDIUM THERMOACETICUM. *Federation Proceedings*, 25, 1642-+.
- LOFERER-KROSSBACHER, M., KLIMA, J. & PSENNER, R. 1998. Determination of bacterial cell dry mass by transmission electron microscopy and densitometric image analysis. *Applied and Environmental Microbiology*, 64, 688-694.

- LOKHOV, P. G. & ARCHAKOV, A. I. 2009. Mass spectrometry methods in metabolomics. *Biochemistry (Moscow) Supplement Series B: Biomedical Chemistry*, 3, 1-9.
- LU, W., BENNETT, B. D. & RABINOWITZ, J. D. 2008. Analytical strategies for LC-MS-based targeted metabolomics. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 871, 236-242.
- LU, W., KWON, Y. K. & RABINOWITZ, J. D. 2007. Isotope ratio-based profiling of microbial folates. *Journal of the American Society for Mass Spectrometry*, 18, 898-909.
- LU, W. Y., KIMBALL, E. & RABINOWITZ, J. D. 2006. A high-performance liquid chromatography-tandem mass spectrometry method for quantitation of nitrogen-containing intracellular metabolites. *Journal of the American Society for Mass Spectrometry*, 17, 37-50.
- 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, 153-164.
- MAGDENOSKA, O., KNUDSEN, P. B., SVENSSEN, D. K. & NIELSEN, K. F. 2015. Quantifying intracellular metabolites in yeast using a matrix with minimal interference from naturally occurring analytes. *Analytical Biochemistry*, 487, 17-26.
- MAHER, A. D., ZIRAH, S. F. M., HOLMES, E. & NICHOLSON, J. K. 2007. Experimental and analytical variation in human urine in H-1 NMR spectroscopy-based metabolic phenotyping studies. *Analytical Chemistry*, 79, 5204-5211.
- MAKAROV, A. 2000. Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. *Analytical Chemistry*, 72, 1156-1162.
- MAKAROV, A., DENISOV, E., KHOLOMEEV, A., BAISCHUN, W., LANGE, O., STRUPAT, K. & HORNING, S. 2006. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Analytical Chemistry*, 78, 2113-2120.
- MAO, H. L., WANG, H. M., WANG, B., LIU, X., GAO, H. C., XU, M., ZHAO, H. S., DENG, X. M. & LIN, D. H. 2009. Systemic Metabolic Changes of Traumatic Critically Ill Patients Revealed by an NMR-Based Metabonomic Approach. *Journal of Proteome Research*, 8, 5423-5430.
- MARCELLIN, E., BEHRENDORFF, J., NAGARAJU, S., DETISSERA, S., SEGOVIA, S., PALFREYMAN, R., DANIELL, J., LICONA-CASSANI, C., QUEK, L.-E., SPEIGHT, R., HODSON, M., SIMPSON, S. D., MITCHELL, W., KOPKE, M. & NIELSEN, L. K. 2016a. Low carbon fuels and commodity chemicals from waste gases-systematic approach to understand energy metabolism in a model acetogen. *Green Chemistry*.
- MARCELLIN, E., BEHRENDORFF, J. B., NAGARAJU, S., DETISSERA, S., SEGOVIA, S., PALFREYMAN, R. W., DANIELL, J., LICONA-CASSANI, C., QUEK, L. E., SPEIGHT, R., HODSON, M. P., SIMPSON, S. D., MITCHELL, W. P., KOPKE, M. & NIELSEN, L. K. 2016b. Low carbon fuels and commodity chemicals from waste gases - systematic approach to understand energy metabolism in a model acetogen. *Green Chemistry*, 18, 3020-3028.
- MASHEGO, M. R., WU, L., VAN DAM, J. C., RAS, C., VINKE, J. L., VAN WINDEN, W. A., VAN GULIK, W. M. & HEIJNEN, J. J. 2004. MIRACLE: mass isotopomer ratio analysis of U-C-13-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. *Biotechnology and Bioengineering*, 85, 620-628.

- MATUSZEWSKI, B. K. 2006. Standard line slopes as a measure of a relative matrix effect in quantitative HPLC-MS bioanalysis. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 830, 293-300.
- MATUSZEWSKI, B. K., CONSTANZER, M. L. & CHAVEZ-ENG, C. M. 2003. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. *Analytical Chemistry*, 75, 3019-3030.
- MAURER, H. H. 2005. Multi-analyte procedures for screening for and quantification of drugs in blood, plasma, or serum by liquid chromatography-single stage or tandem mass spectrometry (LC-MS or LGMS/MS) relevant to clinical and forensic toxicology. *Clinical Biochemistry*, 38, 310-318.
- MCGINNIS, A. C., GRUBB, E. C. & BARTLETT, M. G. 2013. Systematic optimization of ion-pairing agents and hexafluoroisopropanol for enhanced electrospray ionization mass spectrometry of oligonucleotides. *Rapid Communications in Mass Spectrometry*, 27, 2655-2664.
- MILLAT, T., JANSSEN, H., THORN, G. J., KING, J. R., BAHL, H., FISCHER, R. J. & WOLKENHAUER, O. 2013. A shift in the dominant phenotype governs the pH-induced metabolic switch of *Clostridium acetobutylicum* phosphate-limited continuous cultures. *Applied Microbiology and Biotechnology*, 97, 6451-6466.
- MILROY, S. P. 2016. Introduction to Univariate Analysis. *Field Methods in Marine Science: from Measurements to Models*, 157-179.
- MOHAMMADI, M., NAJAFPOUR, G. D., YOUNESI, H., LAHIJANI, P., UZIR, M. H. & MOHAMED, A. R. 2011. Bioconversion of synthesis gas to second generation biofuels: A review. *Renewable & Sustainable Energy Reviews*, 15, 4255-4273.
- MOHAMMADI, M., YOUNESI, H., NAJAFPOUR, G. & MOHAMED, A. R. 2012a. Sustainable ethanol fermentation from synthesis gas by *Clostridium ljungdahlii* in a continuous stirred tank bioreactor. *Journal of Chemical Technology and Biotechnology*, 87, 837-843.
- MOHAMMADI, M., YOUNESI, H., NAJAFPOUR, G. & MOHAMED, A. R. 2012b. Sustainable ethanol fermentation from synthesis gas by *Clostridium ljungdahlii* in a continuous stirred tank bioreactor. *Journal of Chemical Technology & Biotechnology*, 87, 837-843.
- MOLITOR, B., RICHTER, H., MARTIN, M. E., JENSEN, R. O., JUMINAGA, A., MIHALCEA, C. & ANGENENT, L. T. 2016. Carbon recovery by fermentation of CO-rich off gases - Turning steel mills into biorefineries. *Bioresource Technology*, 215, 386-396.
- NAJDEKR, L., FRIEDECKY, D., TAUTENHAHN, R., PLUSKAL, T., WANG, J., HUANG, Y. & ADAM, T. 2016. Influence of Mass Resolving Power in Orbital Ion-Trap Mass Spectrometry-Based Metabolomics. *Analytical Chemistry*, 88, 11429-11435.
- NALAKATH ABUBACKAR, H., VEIGA, M. C. & KENNES, C. 2011. Biological conversion of carbon monoxide: rich syngas or waste gases to bioethanol. *Biofuels Bioproducts & Biorefining-Biofpr*, 5, 93-114.
- NICHOLSON, J. K., CONNELLY, J., LINDON, J. C. & HOLMES, E. 2002. Metabonomics: a platform for studying drug toxicity and gene function. *Nature Reviews Drug Discovery*, 1, 153-161.
- NICHOLSON, J. K. & LINDON, J. C. 2008. Systems biology - Metabonomics. *Nature*, 455, 1054-1056.
- NICHOLSON, J. K., LINDON, J. C. & HOLMES, E. 1999. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, 29, 1181-1189.

- NUNEZ, O., NAKANISHI, K. & TANAKA, N. 2008. Preparation of monolithic silica columns for high-performance liquid chromatography. *Journal of Chromatography A*, 1191, 231-252.
- OLDIGES, M., LUTZ, S., PFLUG, S., SCHROER, K., STEIN, N. & WIENDAHL, C. 2007. Metabolomics: current state and evolving methodologies and tools. *Applied Microbiology and Biotechnology*, 76, 495-511.
- PACZIA, N., ZHU, C. C. & LINSTER, C. 2015. Dynamic quantitative metabolomics for comparative studies in yeast. *Yeast*, 32, S254-S254.
- PENG, J., GUO, K., XIA, J. G., ZHOU, J. J., YANG, J., WESTAWAY, D., WISHART, D. S. & LI, L. 2014. Development of Isotope Labeling Liquid Chromatography Mass Spectrometry for Mouse Urine Metabolomics: Quantitative Metabolomic Study of Transgenic Mice Related to Alzheimer's Disease. *Journal of Proteome Research*, 13, 4457-4469.
- PFEIFFER, C. M., FAZILI, Z., MCCOY, L., ZHANG, M. & GUNTER, E. W. 2004. Determination of folate vitamers in human serum by stable-isotope-dilution tandem mass spectrometry and comparison with radioassay and microbiologic assay. *Clinical Chemistry*, 50, 423-432.
- PHILLIPS, J. R., CLAUSEN, E. C. & GADDY, J. L. 1994. Synthesis gas as substrate for the biological production of fuels and chemicals. *Applied Biochemistry and Biotechnology*, 45-46, 145-157.
- PISITKUN, T., HOFFERT, J. D., YU, M.-J. & KNEPPER, M. A. 2007. Tandem mass spectrometry in physiology. *Physiology*, 22, 390-400.
- POGLIANO, J., OSBORNE, N., SHARP, M. D., ABANES-DE MELLO, A., PEREZ, A., SUN, Y. L. & POGLIANO, K. 1999. A vital stain for studying membrane dynamics in bacteria: a novel mechanism controlling septation during *Bacillus subtilis* sporulation. *Molecular Microbiology*, 31, 1149-1159.
- PUTZER, H. & LAALAMI, S. 2000-2013. Regulation of expression of Aminoacyl-tRNA synthetases and translation factors. *Landes Bioscience*.
- QIAN, T. X., CAI, Z. W. & YANG, M. S. 2004. Determination of adenosine nucleotides in cultured cells by ion-pairing liquid chromatography-electrospray ionization mass spectrometry. *Analytical Biochemistry*, 325, 77-84.
- RABINOWITZ, J. D. & KIMBALL, E. 2007. Acidic acetonitrile for cellular metabolome extraction from *Escherichia coli*. *Analytical Chemistry*, 79, 6167-6173.
- RAGAUSKAS, A. J., WILLIAMS, C. K., DAVISON, B. H., BRITOVSEK, G., CAIRNEY, J., ECKERT, C. A., FREDERICK, W. J., HALLETT, J. P., LEAK, D. J., LIOTTA, C. L., MIELENZ, J. R., MURPHY, R., TEMPLER, R. & TSCHAPLINSKI, T. 2006. The path forward for biofuels and biomaterials. *Science*, 311, 484-489.
- RAGSDALE, S. W. 1997. The Eastern and Western branches of the Wood/Ljungdahl pathway: how the East and West were won. *Biofactors*, 6, 3-11.
- RAGSDALE, S. W. & PIERCE, E. 2008. Acetogenesis and the Wood-Ljungdahl pathway of CO₂ fixation. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 1784, 1873-1898.
- RAHMAN, A. M. A., PAWLING, J., RYCZKO, M., CAUDY, A. A. & DENNIS, J. W. 2014. Targeted metabolomics in cultured cells and tissues by mass spectrometry: Method development and validation. *Analytica Chimica Acta*, 845, 53-61.
- REBEILLE, F., RAVANEL, S., JABRIN, S., DOUCE, R., STOROZHENKO, S. & VAN DER STRAETEN, D. 2006. Foliates in plants: biosynthesis, distribution, and enhancement. *Physiologia Plantarum*, 126, 330-342.
- REUSE, H. & SKOULOUBRIS, S. 2001. Nitrogen metabolism. In: MOBLEY, H. M., GL. HAZELL, SL. (ed.) *Helicobacter pylori: Physiology and Genetics*. Washington DC: ASM Press.

- RICHTER, H., VLAD, D. & UNDEN, G. 2001. Significance of pantothenate for glucose fermentation by *Oenococcus oeni* and for suppression of the erythritol and acetate production. *Archives of Microbiology*, 175, 26-31.
- ROBB, D. B., COVEY, T. R. & BRUINS, A. P. 2000. Atmospheric pressure photoionisation: An ionization method for liquid chromatography-mass spectrometry. *Analytical Chemistry*, 72, 3653-3659.
- ROCK, C. O., CALDER, R. B., KARIM, M. A. & JACKOWSKI, S. 2000. Pantothenate kinase regulation of the intracellular concentration of coenzyme A. *Journal of Biological Chemistry*, 275, 1377-1383.
- ROMANO, S., BONDAREV, V., KOLLING, M., DITTMAR, T. & SCHULZ-VOGT, H. N. 2017. Phosphate Limitation Triggers the Dissolution of Precipitated Iron by the Marine Bacterium *Pseudovibrio* sp FO-BEG1. *Frontiers in Microbiology*, 8.
- SATO, Y., NAKAMURA, T., AOSHIMA, K. & ODA, Y. 2010. Quantitative and Wide-Ranging Profiling of Phospholipids in Human Plasma by Two-dimensional Liquid Chromatography/Mass Spectrometry. *Analytical Chemistry*, 82, 9858-9864.
- SAWADA, Y., AKIYAMA, K., SAKATA, A., KUWAHARA, A., OTSUKI, H., SAKURAI, T., SAITO, K. & HIRAI, M. Y. 2009. Widely Targeted Metabolomics Based on Large-Scale MS/MS Data for Elucidating Metabolite Accumulation Patterns in Plants. *Plant and Cell Physiology*, 50, 37-47.
- SCHATSCHNEIDER, S., ABDELRAZIG, S., SAFO, L., HENSTRA, A. M., MILLAT, T., KIM, D.-H., WINZER, K., MINTON, N. & BARRETT, D. 2017. Quantitative isotope dilution high-resolution mass spectrometry analysis of multiple intracellular metabolites in *Clostridium autoethanogenum* using uniformly ¹³C-labelled standards derived from *Spirulina*. *Analytical Chemistry*, SUBMITTED.
- SCHEIBNER, O., SQUIBB, A., GRECO, G. & STEINER, F. 2016. Target, Suspected-Target, and Non-Target LC-MS(/MS) Screening: New (Practical) Strategies for CECs in Water Bodies. *Assessing Transformation Products of Chemicals by Non-Target and Suspect Screening - Strategies and Workflows, Vol 2*.
- SCHIEL-BENGELSDORF, B. & DURRE, P. 2012. Pathway engineering and synthetic biology using acetogens. *Febs Letters*, 586, 2191-2198.
- SCHIESEL, S., LAEMMERHOFER, M. & LINDNER, W. 2010. Multitarget quantitative metabolic profiling of hydrophilic metabolites in fermentation broths of beta-lactam antibiotics production by HILIC-ESI-MS/MS. *Analytical and Bioanalytical Chemistry*, 396, 1655-1679.
- SCHUCHMANN, K. & MULLER, V. 2016. Energetics and Application of Heterotrophy in Acetogenic Bacteria. *Applied and Environmental Microbiology*, 82, 4056-4069.
- SCHWARTZ, A., SUGG, H., RITTER, T. W. & FERNANDEZREPOLLET, E. 1983. DIRECT DETERMINATION OF CELL DIAMETER, SURFACE-AREA, AND VOLUME WITH AN ELECTRONIC VOLUME SENSING FLOW CYTOMETER. *Cytometry*, 3, 456-458.
- SEIFAR, R. M., RAS, C., VAN DAM, J. C., VAN GULIK, W. M., HEIJNEN, J. J. & VAN WINDEN, W. A. 2009. Simultaneous quantification of free nucleotides in complex biological samples using ion pair reversed phase liquid chromatography isotope dilution tandem mass spectrometry. *Analytical Biochemistry*, 388, 213-219.
- SHEN, Y., FATEMEH, T., TANG, L. H. & CAI, Z. W. 2016. Quantitative metabolic network profiling of *Escherichia coli*: An overview of analytical methods for measurement of intracellular metabolites. *Trac-Trends in Analytical Chemistry*, 75, 141-150.

- SHI, G., WU, J. T., LI, Y., GELEZIUNAS, R., GALLAGHER, K., EMM, T., OLAH, T. & UNGER, S. 2002. Novel direct detection method for quantitative determination of intracellular nucleoside triphosphates using weak anion exchange liquid chromatography/tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 16, 1092-1099.
- SHIN, M. H., LEE, D. Y., LIU, K.-H., FIEHN, O. & KIM, K. H. 2010. Evaluation of Sampling and Extraction Methodologies for the Global Metabolic Profiling of *Saccharophagus degradans*. *Analytical Chemistry*, 82, 6660-6666.
- SIEGEL, D., MEINEMA, A. C., PERMENTIER, H., HOPFGARTNER, G. & BISCHOFF, R. 2014. Integrated Quantification and Identification of Aldehydes and Ketones in Biological Samples. *Analytical Chemistry*, 86, 5089-5100.
- SOUVERAIN, S., RUDAZ, S. & VEUTHEY, J. L. 2004. Matrix effect in LC-ESI-MS and LC-APCI-MS with off-line and on-line extraction procedures. *Journal of Chromatography A*, 1058, 61-66.
- SPRY, C., KIRK, K. & SALIBA, K. J. 2008. Coenzyme A biosynthesis: an antimicrobial drug target. *Fems Microbiology Reviews*, 32, 56-106.
- STOKVIS, E., ROSING, H. & BEIJNEN, J. H. 2005. Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? *Rapid Communications in Mass Spectrometry*, 19, 401-407.
- TANG, L. & KEBARLE, P. 1993. Dependence of ion intensity in electrospray mass-spectrometry on the concentration of the analytes in the electrosprayed solution. *Analytical Chemistry*, 65, 3654-3668.
- TANG, Y. J., MARTIN, H. G., MYERS, S., RODRIGUEZ, S., BAIDOO, E. E. K. & KEASLING, J. D. 2009. Advances in analysis of microbial metabolic fluxes via C-13 isotopic labeling. *Mass Spectrometry Reviews*, 28, 362-375.
- TOENNIES, G., DAS, D. N. & FENG, F. 1966. PANTOTHENATE AND COENZYME A IN BACTERIAL GROWTH. *Journal of Bacteriology*, 92, 707-&.
- TOYOTA, K., GAVIN, A., MIYAGAWA, S., VIAN, M. R. & IGUCHI, T. 2016. Metabolomics reveals an involvement of pantothenate for male production responding to the short-day stimulus in the water flea, *Daphnia pulex*. *Scientific Reports*, 6.
- TYCKO, D. H., METZ, M. H., EPSTEIN, E. A. & GRINBAUM, A. 1985. Flow-cytometric light-scattering measurement of red blood-cell volume and hemoglobin concentration. *Applied Optics*, 24, 1355-1365.
- TZUR, A., MOORE, J. K., JORGENSEN, P., SHAPIRO, H. M. & KIRSCHNER, M. W. 2011. Optimizing Optical Flow Cytometry for Cell Volume-Based Sorting and Analysis. *Plos One*, 6.
- URAN, S., LANDMARK, K. E., HJELLUM, G. & SKODAND, T. 2007. Quantification of C-13 pyruvate and C-13 lactate in dog blood by reversed-phase liquid chromatography-electrospray ionization mass spectrometry after derivatization with 3-nitrophenylhydrazine. *Journal of Pharmaceutical and Biomedical Analysis*, 44, 947-954.
- VALGEPEA, K., DE SOUZA PINTO LEMGRUBER, R., MEAGHAN, K., PALFREYMAN, R. W., ABDALLA, T., HEIJSTRA, B. D., BEHRENDORFF, J. B., TAPPEL, R., KÖPKE, M., SIMPSON, S. D., NIELSEN, L. K. & MARCELLIN, E. Maintenance of ATP Homeostasis Triggers Metabolic Shifts in Gas-Fermenting Acetogens. *Cell Systems*.
- VAN DAM, J. C., EMAN, M. R., FRANK, J., LANGE, H. C., VAN DEDEM, G. W. K. & HEIJNEN, S. J. 2002. Analysis of glycolytic intermediates in *Saccharomyces cerevisiae* using anion exchange chromatography and electrospray ionization with tandem mass spectrometric detection. *Analytica Chimica Acta*, 460, 209-218.

- VAN EECKHAUT, A., LANCKMANS, K., SARRE, S., SMOLDERS, I. & MICHOTTE, Y. 2009. Validation of bioanalytical LC-MS/MS assays: Evaluation of matrix effects. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 877, 2198-2207.
- VAN GROENESTIJN, J. W., ABUBACKAR, H. N., VEIGA, M. C. & KENNES, C. 2013. Bioethanol. *Air Pollution Prevention and Control: Bioreactors and Bioenergy*, 431-463.
- VAN GULIK, W. M. 2010. Fast sampling for quantitative microbial metabolomics. *Current Opinion in Biotechnology*, 21, 27-34.
- VAN HOOFF, N., COURTHEYN, D., ANTIGNAC, J. P., VAN DE WIELE, M., POELMANS, S., NOPPE, H. & DE BRABANDER, H. 2005. Multi-residue liquid chromatography/tandem mass spectrometric analysis of beta-agonists in urine using molecular imprinted polymers. *Rapid Communications in Mass Spectrometry*, 19, 2801-2808.
- VOLKMER, B. & HEINEMANN, M. 2011. Condition-Dependent Cell Volume and Concentration of Escherichia coli to Facilitate Data Conversion for Systems Biology Modeling. *Plos One*, 6.
- VROBEL, I., FRIEDECKY, D., FABER, E., NAJDEKR, L., MICOVA, K., KARLIKOVA, R. & ADAM, T. 2017. Novel sulphur-containing imatinib metabolites found by untargeted LC-HRMS analysis. *European Journal of Pharmaceutical Sciences*, 104, 335-343.
- WANG, S., CYRONAK, M. & YANG, E. 2007. Does a stable isotopically labeled internal standard always correct analyte response? A matrix effect study on a LC/MS/MS method for the determination of carvedilol enantiomers in human plasma. *Journal of Pharmaceutical and Biomedical Analysis*, 43, 701-707.
- WANG, Y., LEHMANN, R., LU, X., ZHAO, X. J. & XU, G. W. 2008. Novel, fully automatic hydrophilic interaction/reversed-phase column-switching high-performance liquid chromatographic system for the complementary analysis of polar and apolar compounds in complex samples. *Journal of Chromatography A*, 1204, 28-34.
- WARD, J. L., BAKER, J. M. & BEALE, M. H. 2007. Recent applications of NMR spectroscopy in plant metabolomics. *Febs Journal*, 274, 1126-1131.
- WIELING, J. 2002. LC-MS-MS experiences with internal standards. *Chromatographia*, 55, S107-S113.
- WINDER, C. L. & DUNN, W. B. 2011. Fit-for-Purpose Quenching and Extraction Protocols for Metabolic Profiling of Yeast Using Chromatography-Mass Spectrometry Platforms. *Yeast Systems Biology: Methods and Protocols*, 759, 225-238.
- WINDER, C. L., DUNN, W. B., SCHULER, S., BROADHURST, D., JARVIS, R., STEPHENS, G. M. & GOODACRE, R. 2008. Global metabolic profiling of Escherichia coli cultures: An evaluation of methods for quenching and extraction of intracellular metabolites. *Analytical Chemistry*, 80, 2939-2948.
- WISHART, D. S., JEWISON, T., GUO, A. C., WILSON, M., KNOX, C., LIU, Y. F., DJOUMBOU, Y., MANDAL, R., AZIAT, F., DONG, E., BOUATRA, S., SINELNIKOV, I., ARNDT, D., XIA, J. G., LIU, P., YALLOU, F., BJORNDAHL, T., PEREZ-PINEIRO, R., EISNER, R., ALLEN, F., NEVEU, V., GREINER, R. & SCALBERT, A. 2013. HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Research*, 41, D801-D807.
- WOOD, H. G. 1991. Life with co or co2 and h2 as a source of carbon and energy. *Faseb Journal*, 5, 156-163.
- WU, L., MASHEGO, M. R., VAN DAM, J. C., PROELL, A. M., VINKE, J. L., RAS, C., VAN WINDEN, W. A., VAN GULIK, W. M. & HEIJNEN, J. J. 2005. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly C-13-labeled cell extracts as internal standards. *Analytical Biochemistry*, 336, 164-171.

- XIA, J. G., MANDAL, R., SINELNIKOV, I. V., BROADHURST, D. & WISHART, D. S. 2012. MetaboAnalyst 2.0-a comprehensive server for metabolomic data analysis. *Nucleic Acids Research*, 40, W127-W133.
- XIA, J. G., SINELNIKOV, I. V., HAN, B. & WISHART, D. S. 2015. MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Research*, 43, W251-W257.
- XIAO, J. F., ZHOU, B. & RESSOM, H. W. 2012. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *Trac-Trends in Analytical Chemistry*, 32, 1-14.
- YAMASHITA, M. & FENN, J. B. 1984. ELECTROSPRAY ION-SOURCE - ANOTHER VARIATION ON THE FREE-JET THEME. *Journal of Physical Chemistry*, 88, 4451-4459.
- ZAKHARTSEV, M., VIELHAUER, O., HORN, T., YANG, X. L. & REUSS, M. 2015. Fast sampling for quantitative microbial metabolomics: new aspects on cold methanol quenching: metabolite co-precipitation. *Metabolomics*, 11, 286-301.
- ZHAO, X. H., CONDRUZ, S., CHEN, J. K. & JOLICOEUR, M. 2016. A quantitative metabolomics study of high sodium response in *Clostridium acetobutylicum* ATCC 824 acetone-butanol-ethanol (ABE) fermentation. *Scientific Reports*, 6.
- ZUBAREV, R. A. & MAKAROV, A. 2013. Orbitrap Mass Spectrometry. *Analytical Chemistry*, 85, 5288-5296.

APPENDICES

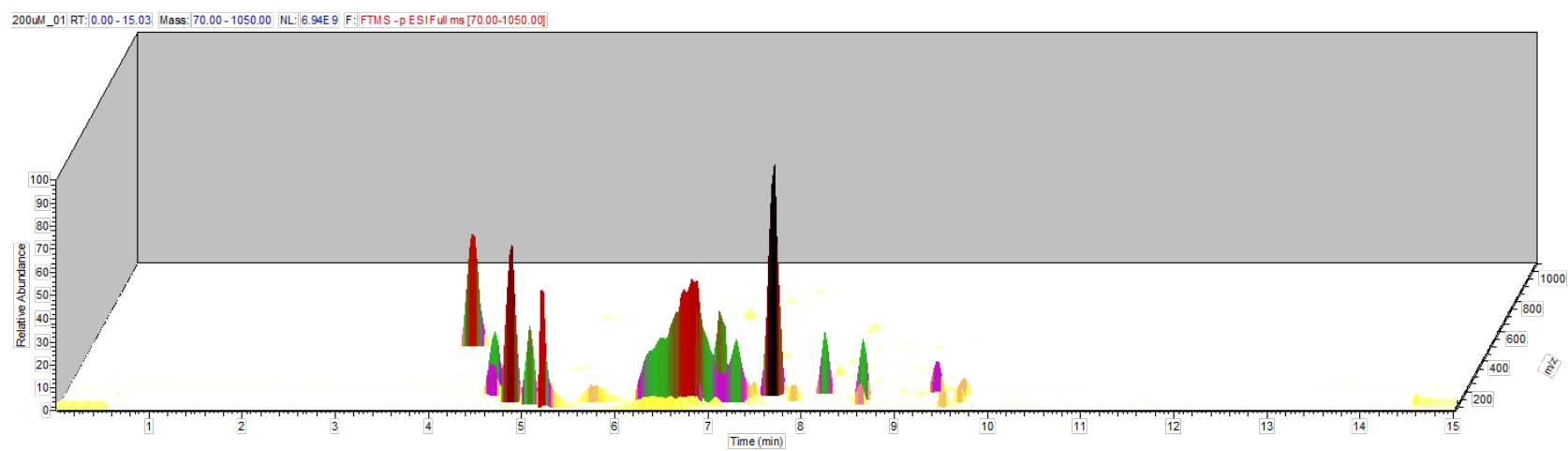
Appendix 1 Parent and product ions for monoglutamated and poly-glutamated folates in positive ionisation mode confirmed using relevant publication (Lu et al., 2007).

<u>Compounds</u>	Parent formula	¹²C- parent <i>m/z</i>	Collision energy (eV)	¹²C- product <i>m/z</i>	Product formula
PteGlu	C ₁₉ H ₂₀ N ₇ O ₆ ⁺	442.15	27	295.09	C ₁₄ H ₁₁ N ₆ O ₂ ⁺
H ₂ PteGlu	C ₁₉ H ₂₂ N ₇ O ₆ ⁺	444.16	27	178.07	C ₇ H ₈ N ₅ O ₁ ⁺
H ₄ PteGlu	C ₁₉ H ₂₄ N ₇ O ₆ ⁺	446.18	20	299.13	C ₁₄ H ₁₅ N ₆ O ₂ ⁺
5,10- CH=H ₄ PteGlu	C ₂₀ H ₂₂ N ₇ O ₆ ⁺	456.16	31	412.17	C ₁₉ H ₂₂ N ₇ O ₄ ⁺
5,10-CH ₂ - H ₄ PteGlu	C ₂₀ H ₂₄ N ₇ O ₆ ⁺	458.18	23	311.13	C ₁₅ H ₁₅ N ₆ O ₂ ⁺
5-CH ₃ - H ₄ PteGlu	C ₂₀ H ₂₆ N ₇ O ₆ ⁺	460.19	25	313.14	C ₁₅ H ₁₇ N ₆ O ₂ ⁺
5/10-CHO- H ₄ PteGlu	C ₂₀ H ₂₄ N ₇ O ₇ ⁺	474.17	20	327.12	C ₁₅ H ₁₅ N ₆ O ₃ ⁺
PteGlu ₂	C ₂₄ H ₂₇ N ₈ O ₉ ⁺	571.19	30	295.09	C ₁₄ H ₁₁ N ₆ O ₂ ⁺
H ₂ PteGlu ₂	C ₂₄ H ₂₉ N ₈ O ₉ ⁺	573.21	30	178.07	C ₇ H ₈ N ₅ O ₁ ⁺
H ₄ PteGlu ₂	C ₂₄ H ₃₁ N ₈ O ₉ ⁺	575.22	30	299.13	C ₁₄ H ₁₅ N ₆ O ₂ ⁺
5,10- CH=H ₄ PteGlu ₂	C ₂₅ H ₂₉ N ₈ O ₉ ⁺	585.22	30	412.17	C ₁₉ H ₂₂ N ₇ O ₄ ⁺
5,10-CH ₂ - H ₄ PteGlu ₂	C ₂₅ H ₃₁ N ₈ O ₉ ⁺	587.22	30	311.13	C ₁₅ H ₁₅ N ₆ O ₂ ⁺
5-CH ₃ - H ₄ PteGlu ₂	C ₂₅ H ₃₃ N ₈ O ₉ ⁺	589.24	25	313.14	C ₁₅ H ₁₇ N ₆ O ₂ ⁺
5/10-CHO- H ₄ PteGlu ₂	C ₂₅ H ₃₁ N ₈ O ₁₀ ⁺	603.22	24	327.12	C ₁₅ H ₁₅ N ₆ O ₃ ⁺
PteGlu ₃	C ₂₉ H ₃₄ N ₉ O ₁₂ ⁺	700.23	32	295.09	C ₁₄ H ₁₁ N ₆ O ₂ ⁺
H ₂ PteGlu ₃	C ₂₉ H ₃₆ N ₉ O ₁₂ ⁺	702.25	35	178.07	C ₇ H ₈ N ₅ O ₁ ⁺
H ₄ PteGlu ₃	C ₂₉ H ₃₈ N ₉ O ₁₂ ⁺	704.26	35	299.13	C ₁₄ H ₁₅ N ₆ O ₂ ⁺
5,10- CH=H ₄ PteGlu ₃	C ₃₀ H ₃₆ N ₉ O ₁₂ ⁺	714.25	35	412.17	C ₁₉ H ₂₂ N ₇ O ₄ ⁺
5,10-CH ₂ - H ₄ PteGlu ₃	C ₃₀ H ₃₈ N ₉ O ₁₂ ⁺	716.26	35	311.13	C ₁₅ H ₁₅ N ₆ O ₂ ⁺

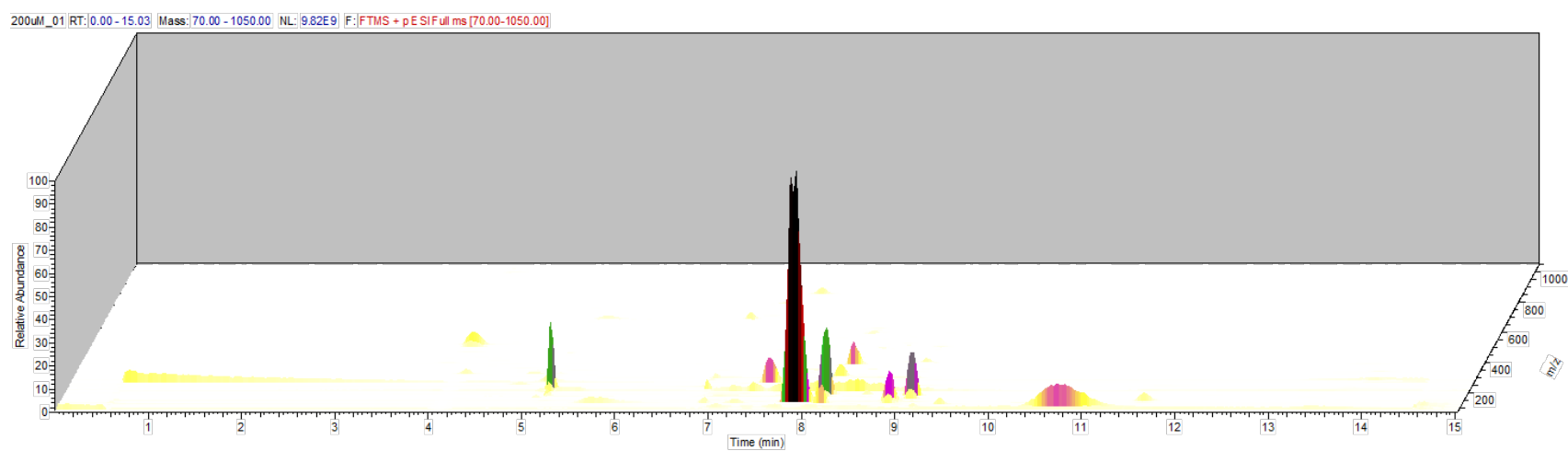
5-CH ₃ - H ₄ PteGlu ₃	C ₃₀ H ₄₀ N ₉ O ₁₂ ⁺	718.28	32	313.14	C ₁₅ H ₁₇ N ₆ O ₂ ⁺
5/10-CHO- H ₄ PteGlu ₃	C ₃₀ H ₃₈ N ₉ O ₁₃ ⁺	732.26	36	327.12	C ₁₅ H ₁₅ N ₆ O ₃ ⁺
PteGlu ₄	C ₃₄ H ₄₁ N ₁₀ O ₁₅ ⁺	829.28	34	295.09	C ₁₄ H ₁₁ N ₆ O ₂ ⁺
H ₂ PteGlu ₄	C ₃₄ H ₄₃ N ₁₀ O ₁₅ ⁺	831.29	40	178.07	C ₇ H ₈ N ₅ O ₁ ⁺
H ₄ PteGlu ₄	C ₃₄ H ₄₅ N ₁₀ O ₁₅ ⁺	833.31	40	299.13	C ₁₄ H ₁₅ N ₆ O ₂ ⁺
5,10- CH=H ₄ PteGl u ₄	C ₃₅ H ₄₃ N ₁₀ O ₁₅ ⁺	843.29	40	412.17	C ₁₉ H ₂₂ N ₇ O ₄ ⁺
5,10-CH ₂ - H ₄ PteGlu ₄	C ₃₅ H ₄₅ N ₁₀ O ₁₅ ⁺	845.31	40	311.13	C ₁₅ H ₁₅ N ₆ O ₂ ⁺
5-CH ₃ - H ₄ PteGlu ₄	C ₃₅ H ₄₇ N ₁₀ O ₁₅ ⁺	847.32	38	313.14	C ₁₅ H ₁₇ N ₆ O ₂ ⁺
5/10-CHO- H ₄ PteGlu ₄	C ₃₅ H ₄₅ N ₁₀ O ₁₆ ⁺	861.3	45	327.12	C ₁₅ H ₁₅ N ₆ O ₃ ⁺
PteGlu ₅	C ₃₉ H ₄₈ N ₁₁ O ₁₈ ⁺	958.32	20	295.09	C ₁₄ H ₁₁ N ₆ O ₂ ⁺
H ₂ PteGlu ₅	C ₃₉ H ₅₀ N ₁₁ O ₁₈ ⁺	960.33	20	178.07	C ₇ H ₈ N ₅ O ₁ ⁺
H ₄ PteGlu ₅	C ₃₉ H ₅₂ N ₁₁ O ₁₈ ⁺	962.35	20	299.13	C ₁₄ H ₁₅ N ₆ O ₂ ⁺
5,10- CH=H ₄ PteGl u ₅	C ₄₀ H ₅₀ N ₁₁ O ₁₈ ⁺	972.33	20	412.17	C ₁₉ H ₂₂ N ₇ O ₄ ⁺
5,10-CH ₂ - H ₄ PteGlu ₅	C ₄₀ H ₅₂ N ₁₁ O ₁₈ ⁺	974.35	20	311.13	C ₁₅ H ₁₅ N ₆ O ₂ ⁺
5-CH ₃ - H ₄ PteGlu ₅	C ₄₀ H ₅₄ N ₁₁ O ₁₈ ⁺	976.36	43	313.14	C ₁₅ H ₁₇ N ₆ O ₂ ⁺
5/10-CHO- H ₄ PteGlu ₅	C ₄₀ H ₅₂ N ₁₁ O ₁₉ ⁺	990.34	53	327.12	C ₁₅ H ₁₅ N ₆ O ₃ ⁺
PteGlu ₆	C ₄₄ H ₅₅ N ₁₂ O ₂₁ ⁺	1087.36	21	295.09	C ₁₄ H ₁₁ N ₆ O ₂ ⁺
H ₂ PteGlu ₆	C ₄₄ H ₅₇ N ₁₂ O ₂₁ ⁺	1089.38	21	178.07	C ₇ H ₈ N ₅ O ₁ ⁺
H ₄ PteGlu ₆	C ₄₄ H ₅₉ N ₁₂ O ₂₁ ⁺	1091.39	21	299.13	C ₁₄ H ₁₅ N ₆ O ₂ ⁺
5,10- CH=H ₄ PteGl u ₆	C ₄₅ H ₅₇ N ₁₂ O ₂₁ ⁺	1101.38	21	412.17	C ₁₉ H ₂₂ N ₇ O ₄ ⁺
5,10-CH ₂ - H ₄ PteGlu ₆	C ₄₅ H ₅₉ N ₁₂ O ₂₁ ⁺	1103.39	21	311.13	C ₁₅ H ₁₅ N ₆ O ₂ ⁺
5-CH ₃ - H ₄ PteGlu ₆	C ₄₅ H ₆₁ N ₁₂ O ₂₁ ⁺	1105.41	21	313.14	C ₁₅ H ₁₇ N ₆ O ₂ ⁺
5/10-CHO- H ₄ PteGlu ₆	C ₄₅ H ₅₉ N ₁₂ O ₂₂ ⁺	1119.39	25	327.12	C ₁₅ H ₁₅ N ₆ O ₃ ⁺

PteGlu ₇	C ₄₉ H ₆₃ N ₁₃ O ₂₄ ² ₊	608.71	22	295.09	C ₁₄ H ₁₁ N ₆ O ₂ ₊
H ₂ PteGlu ₇	C ₄₉ H ₆₅ N ₁₃ O ₂₄ ² ₊	609.71	22	178.07	C ₇ H ₈ N ₅ O ₁ ⁺
H ₄ PteGlu ₇	C ₄₉ H ₆₇ N ₁₃ O ₂₄ ² ₊	610.71	22	299.13	C ₁₄ H ₁₅ N ₆ O ₂ ₊
5,10- CH=H ₄ PteGl u ₇	C ₅₀ H ₆₅ N ₁₃ O ₂₄ ² ₊	615.71	22	412.17	C ₁₉ H ₂₂ N ₇ O ₄ ₊
5,10-CH ₂ - H ₄ PteGlu ₇	C ₅₀ H ₆₇ N ₁₃ O ₂₄ ² ₊	616.71	22	311.13	C ₁₅ H ₁₅ N ₆ O ₂ ₊
5-CH ₃ - H ₄ PteGlu ₇	C ₅₀ H ₆₉ N ₁₃ O ₂₄ ² ₊	617.72	22	313.14	C ₁₅ H ₁₇ N ₆ O ₂ ₊
5/10-CHO- H ₄ PteGlu ₇	C ₅₀ H ₆₇ N ₁₃ O ₂₅ ² ₊	624.71	26	327.12	C ₁₅ H ₁₅ N ₆ O ₃ ₊
PteGlu ₈	C ₅₄ H ₇₀ N ₁₄ O ₂₇ ² ₊	673.22	23	295.09	C ₁₄ H ₁₁ N ₆ O ₂ ₊
H ₂ PteGlu ₈	C ₅₄ H ₇₂ N ₁₄ O ₂₇ ² ₊	674.23	23	178.07	C ₇ H ₈ N ₅ O ₁ ⁺
H ₄ PteGlu ₈	C ₅₄ H ₇₄ N ₁₄ O ₂₇ ² ₊	675.23	23	299.13	C ₁₄ H ₁₅ N ₆ O ₂ ₊
5,10- CH=H ₄ PteGl u ₈	C ₅₅ H ₇₂ N ₁₄ O ₂₇ ² ₊	680.23	23	412.17	C ₁₉ H ₂₂ N ₇ O ₄ ₊
5,10-CH ₂ - H ₄ PteGlu ₈	C ₅₅ H ₇₄ N ₁₄ O ₂₇ ² ₊	681.23	23	311.13	C ₁₅ H ₁₅ N ₆ O ₂ ₊
5-CH ₃ - H ₄ PteGlu ₈	C ₅₅ H ₇₆ N ₁₄ O ₂₇ ² ₊	682.24	23	313.14	C ₁₅ H ₁₇ N ₆ O ₂ ₊
5/10-CHO- H ₄ PteGlu ₈	C ₅₅ H ₇₄ N ₁₄ O ₂₈ ² ₊	689.23	27	327.12	C ₁₅ H ₁₅ N ₆ O ₃ ₊

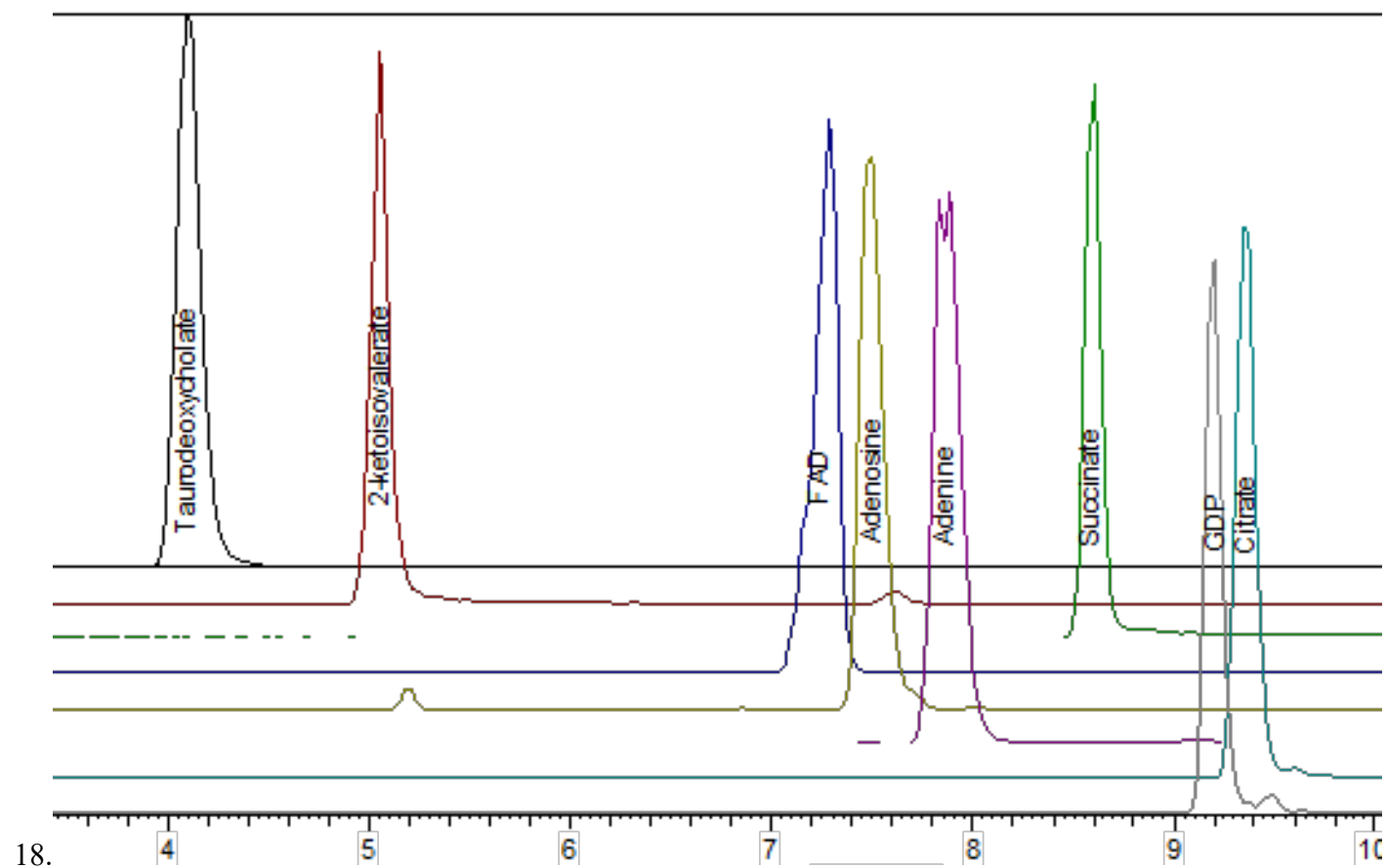
Appendix 2 3D ion map view of negative mode total ion chromatogram of metabolites using full scan mode



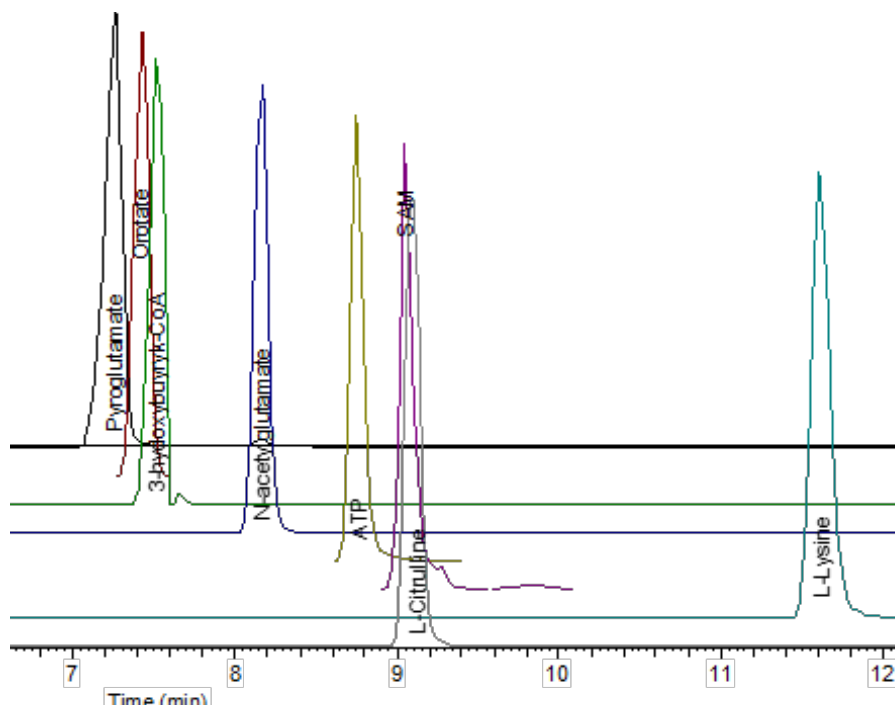
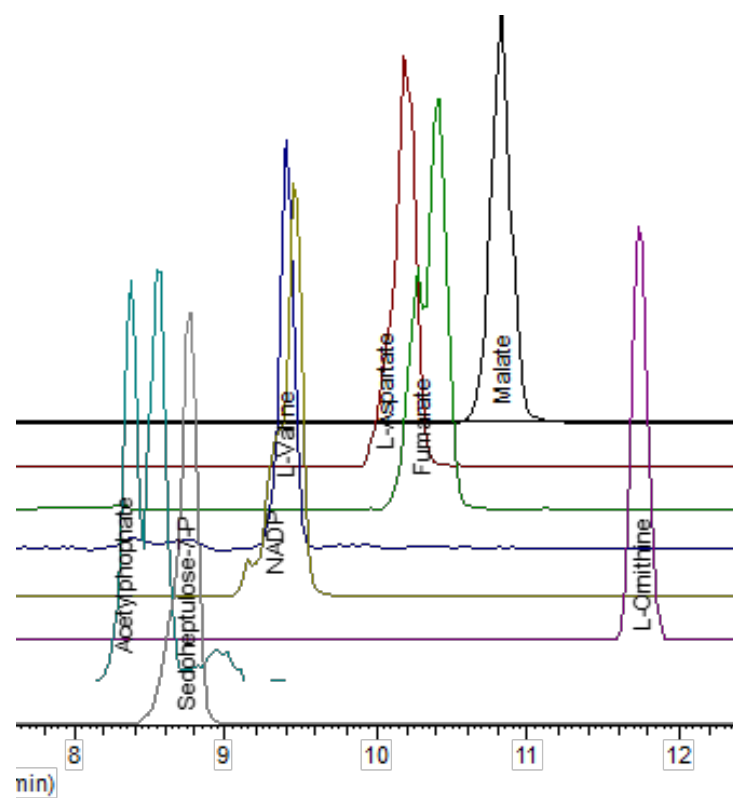
Appendix 3 3D ion map view of positive mode total ion chromatogram of metabolites using full scan mode



Appendix 4 Chromatograms of compounds in order of retention after switching polarity full scan mode on Q-Exactive (Taurodeoxycholate, 2-ketoisovalerate, FAD, Adenosine, Adenine, Succinate, Guanosine diphosphate (GDP), citrate)



Appendix 5 Chromatograms of compounds in order of retention after switching polarity full scan mode on Q-Exactive (Acetylphosphate, Sedoheptulose-7-phosphate, NADP, L-Valine, L-Aspartate, Fumarate, Malate, L-Ornithine, pyroglutamate, Orotate, 3-hydroxybutyryl-CoA, N-acetylglutamate, ATP, SAM, L-citrulline , L-Lysine)



Appendix 6 Chromatograms of compounds in order of retention after switching polarity full scan mode on Q-Exactve (Phenylpyruvate, N-acetylcysteine, Pyruvate, L-phenylalanine, l-Methionine, N-Acetyllysine, α -ketoglutarate)

