Mass spectrometry based hyphenated techniques for microalgal and mammalian metabolomics

Rahul Vijay Kapoore
(B Pharm, MSc)

Thesis submitted for the degree of Doctor of Philosophy
To The University of Sheffield, Sheffield, U.K.
Department of Chemical and Biological Engineering

August 2014

This copy of the thesis has been submitted with the condition that anyone who consults it is understood to recognise that the copyright rests with its author. No quotation and information derived from this thesis may be published without prior written consent from the author or the University (as may be appropriate).

Declaration

I hereby declare that this thesis is my own work and effort which was conducted at The University of Sheffield, U.K. This work has not been submitted anywhere for any other degree of qualification. Where other sources of information have been used, they have been acknowledged.

Signature: ………………………… Date: 29/08/2014

Rahul Vijay Kapoore

Abstract

In metabolomics, the analytical challenge is to capture the chemical diversity of the metabolome. With the current technologies only a portion of the metabolome can be analysed. As a result there is a drive to direct significant analytical efforts towards capturing the metabolome or changes in the metabolome reliably and reproducibly in biological systems. Apart from analytical challenges, the challenges also include development of appropriate methodologies to quench and extract metabolites which is a crucial parameter in sample preparation and is required to achieve an accurate representation of phenotype. This thesis focuses on addressing both the challenges in mammalian and microalgal metabolomics.

Metabolomics in cancer research is gaining momentum as a tool to understand the molecular mechanism of disease progression and for the identification of specific biomarkers which may assist distinguishing between normal, benign and metastatic cancer states. In our first investigation we developed GC-MS based modified direct cell scraping, bead harvesting and LN_2 methods for harvesting three adherently grown mammalian cell lines (two breast cancer cell lines MDA-MB 436, MCF7 and an endothelial cell line HMEC1) which provided rapid and reliable route with three fold improved metabolome coverage and reduced the artifacts due to metabolome leakage compared to conventional methods. Later optimized treatments were employed and the influence of various washing and quenching solvents (buffered/unbuffered) on metabolite leakage was investigated for metastatic cancer cell line MDA-MB-231. This identified one washing step with PBS followed by quenching with 60% methanol (buffered with HEPES) as the best washing and quenching solvents. Further validation and comparison of proposed workflows for metabolomic study of two metastatic TNBC cell lines (MDA-MB-231 and MDA-MB-436) resulted in recovery of 154 unique metabolites and demonstrated the robustness and reliability of these methods in pathway based analysis in cancer.

In case of GC-MS based microalgal metabolomics, with comprehensive evaluation of selected quenching and extraction methods in model microalga *C. reinhardtii*, we have successfully demonstrated that the choice of quenching and extraction solvents have significant impact on recovery of different classes of metabolites. Our results clearly indicate that 60% methanol (buffered with HEPES) and 25 % aqueous methanol are the best suited quenching and extraction solvent respectively for untargeted metabolomic analysis of *C. reinhardtii*, as the highest number of metabolites belonging to various chemical classes were recovered with good

intensities and reproducibilities with this miniaturized proposed method compared to other evaluated methods.

Later impact of various stages involved in biodiesel production workflow from microalga on recovery of biodiesel was assessed in three microalgal species namely *C. reinhardtii, D. salina* and *N. salina*. Within which we have developed an optimized GC-FID method and miniaturized direct TE method for quantification of fatty acids, which can be applied to a small amount of biomass and saves tremendous amounts of time, solvents and reagents required, is less expensive and uses environment friendly solvents making it more suitable for sustainable large scale production.

In our final investigation, we directed our efforts towards preliminary optimization and comparative analysis of HILIC and IP-RP-HPLC based separation for the retention and separation of specific metabolites classes. This identified HILIC as the best available column till date for untargeted metabolomic studies.

The descriptive understanding gained from each of these investigations provides greater insight into biology of mammalian and algal systems by improving the metabolome coverage for various metabolite classes. These insights illustrating the underlying molecular pathways involved in respective biology's, will help scientific communities in identifying as-of-yet-missing reactions in the metabolic network. In addition these insights will surely help in generating many hypothesis based investigations in microalgal and cancer community.

Table of Contents

Declaration.......iii

Abstract.......v

List of Publications.......xv

Oral and Poster presentations.......xvii

List of abbreviations.......xix

List of figures.......xxi

List of tables.......xxxvii

1. Introduction.......3

    1.1 Background on metabolomics.......3

    1.2 Mammalian metabolomics.......4

    1.3 Microalgal systems.......5

    1.4 Aims and objectives.......7

        1.4.1 Aim.......7

        1.4.2 Objectives.......7

2. Literature Review.......11

    2.1 Introduction to metabolomics.......11

        2.1.1 Approaches used in metabolomics.......13

        2.1.2 Classification of metabolites.......14

    2.2 Metabolomics workflow.......16

        2.2.1 Sampling and quenching.......16

        2.2.2 Extraction of metabolites.......19

        2.2.3. Metabolite analysis.......22

        2.2.4. Data export and analysis.......27

    2.3 GC-MS and LC-MS based metabolomics.......29

        2.3.1 GC-MS based metabolomics.......29

        2.3.2 LC-MS based metabolomics.......34

    2.4 Microalgae for biofuel production.......37

        2.4.1 What are microalgae?.......37

        2.4.2 Microalgae for biodiesel production.......39

        2.4.3 Lipids and fatty acids.......39

        2.4.4 Lipid extraction.......40

        2.4.5 In situ lipid extraction and transesterification.......45

        2.4.6 Catalysts used in TE reactions.......46

        2.4.7 Lipid determination.......48

2.5 Microalgal metabolomics 50
2.6 Mammalian metabolomics 51
3. Metabolite leakage in metabolome analyses of adherent mammalian cell lines 57
3.1 Introduction 57
3.2 Materials and Methods 59
3.2.1 Chemicals and reagents 59
3.2.2 Cell lines, cell culture and growth assessment 59
3.2.3 Cell quenching 60
3.2.4 Metabolite Extraction 60
3.2.5 Metabolite Derivatization 62
3.2.6 GC-MS analysis 63
3.2.7 Metabolite identification 63
3.2.8 Data analysis 63
3.3 Results and Discussion 64
3.3.1 Trypsinization versus Cell scraping 65
3.3.2 Modified approaches 68
3.3.3 Modified direct cell scraping (NTcs) 69
3.3.4 Bead harvesting 70
3.3.5 Direct quenching using LN_2 71
3.3.6 Protein precipitation using acetone (AP) 71
3.3.7 Comparison of conventional and modified methods 72
3.4 Conclusions 74
4. Effect of washing and quenching in metabolomics of adherently grown mammalian cells: A case study on the metastatic cancer cell line MDA-MB-231 79
4.1 Introduction 79
4.2 Material and Methods 81
4.2.1 Chemicals and reagents 81
4.2.2 Cell lines, cell culture and growth assessment 81
4.2.3 ATP assay 81
4.2.4 Protein assay as a normalisation to ATP content 84
4.2.5 Scanning electron microscopy (SEM) analysis 85
4.2.6 Cell quenching for metabolome analysis 86
4.2.7 Metabolite Extraction using modified cell scraping method 86
4.2.8 Simultaneous quenching and extraction with liquid nitrogen 88
4.2.9 Metabolite derivatization, GC-MS analysis, metabolite identification and data analysis 88

4.3 Results and Discussion 88
4.3.1 ATP assay.. 89
4.3.2 Leakage of ATP in response to washing steps/solutions 90
4.3.3 Leakage of ATP in response to quenching solutions.. 92
Influence of non-buffered methanol quenching on leakage 93
Influence of buffered methanol quenching on leakage... 94
4.3.4 SEM 94
4.3.5 GC-MS based overall recovery of metabolites 96
4.4 Conclusions 102
5. Comparison of proposed protocols for metabolomics of two metastatic TNBC cell lines: MDA-MB-231 and MDA-MB-436 105
5.1 Background 105
5.2 Comparison of the proposed protocols for adherent mammalian cell metabolomics .. 106
5.3 Are the proposed protocols able to track the changes between the cell lines which share common biological and morphological features? 110
5.4 Impact of these proposed methods in pathway based analysis for deriving meaningful biological interpretations which will aid in understanding the cancer metabolism and developing novel therapeutic regimens 115
5.5 Conclusions 118
6. Optimization of the quenching and extraction workflow for GC-MS based metabolomic investigation of microbial and microalgal species: Protocol validation with hypothesis driven biological study in *C. reinhardtii.* 123
6.1 Introduction 123
6.2 Materials and Methods.. 124
6.2.1 Bacterial/microalgal strain and growth conditions 124
6.2.2 Preparation of standard metabolites mixture 126
6.2.3 Sample extraction and quenching 126
6.2.4 Extraction of metabolites.. 128
6.2.5 Derivatization of metabolites 130
6.2.6 GC-MS for metabolite analysis... 131
6.2.7 Identification of metabolites and data analysis... 131
6.3 Results and Discussion 132
6.3.1 Validation of the use of GC-MS for metabolomics 132
6.3.2 Optimization of the quenching and extraction workflow for GC-MS based metabolomic investigation of one microbial and three microalgal species 133
6.3.3 Comparison of mixotrophic and photoautotrophic growth of *C. reinhardtii* grown under turbidostatic conditions: A metabolic and proteomic investigation. 147

6.4 Conclusions 160

7. Evaluation and optimisation of appropriate sampling and quenching method for microalgal sample: A case study on the unicellular eukaryotic green alga *Chlamydomonas reinhardtii* 165

7.1 Introduction 165

7.2 Material and Methods 168

7.2.1 Chemicals and analytical reagents 168

7.2.2 Microalgal cultivation 169

7.2.3 Sampling and quenching 169

7.2.4 Metabolite extraction 169

7.2.5 Metabolite derivatization 170

7.2.6 GC-MS analysis, metabolite identification and data analysis 170

7.3 Results and Discussion 170

7.3.1 Effect of various quenching solvents, methanol concentration and inclusion of buffer additives on metabolite leakage (Approach 1) 172

7.3.2 Effect of prolonged exposure to quenching solvent on metabolite leakage (Approach 2) 184

7.3.3 Effect of quenching solvent to culture ratio (temperature influence) 193

(Approach 3) 193

7.4 Conclusions 195

8. Influence of extraction solvents on metabolome recovery in the model microalga *Chlamydomonas reinhardtii* 199

8.1 Introduction 199

8.2 Material and Methods 204

8.2.1 Chemicals and analytical reagents 204

8.2.2 Microalgal cultivation 204

8.2.3 Sampling and quenching 204

8.2.4 Metabolite extraction 204

8.2.5 Metabolite derivatization, GC-MS analysis, metabolite identification and data analysis 205

8.3 Results and Discussion 206

8.3.1 Monophasic solvent systems for extraction 207

8.3.2 Biphasic solvent systems for extraction 217

8.3.3 Comparison of extraction efficiencies of monophasic and biphasic solvent systems 227

8.4 Conclusions 239

9. Optimisation and comparative evaluation of conventional and direct trans-esterification methods for reliable quantification of FAMEs from micro-algal biofuel strains 243
9.1 Introduction 243
9.2 Material and Methods 245
9.2.1 Chemicals and analytical reagents 245
9.2.2. Selection of microalgal species, cultivation and harvest 245
9.2.3 Evaluation of cell disruption techniques using experimental design 246
9.2.4 Comprehensive evaluation of conventional two step TE methods 248
9.2.5 Comprehensive evaluation of direct trans-esterification (DT) methods 252
9.2.6 Gravimetric analysis 256
9.2.7 Scanning electron microscopy (SEM) analysis 256
9.2.8 Gas chromatography 256
9.2.9 Statistical analysis of the results 256
9.2.10 Accuracy, precision, LOD and LOQ 257
9.2.11 Biomass water content 257
9.3 Results and Discussion 257
9.3.1 Selection of microalgal strains 257
9.3.2 Selection of disruption techniques 258
9.3.3 Optimization of GC method and parameters 276
9.3.4 Conventional methods of 2 step trans-esterification 280
9.3.5 One step direct trans-esterification (DT) methods 287
9.3.6 Comparison of conventional and DT methods 294
9.4 Conclusions 296
10. Evaluation of HILIC and IP-RP-HPLC for targeted analysis of amino acids, organic acids, water soluble vitamins, nucleotides, nucleosides and nucleobases towards improving metabolome coverage. 303
10.1 Introduction 303
10.2 Material and Methods 306
10.2.1 Equipments and columns 306
10.2.2 Chemicals and reagents 306
10.2.3 Preparation of metabolite standards and cocktails 306
10.2.4 Buffer compositions 307
10.2.5 Hydrophilic interaction liquid chromatography (HILIC) 308
10.2.6 Reversed phase liquid chromatography (RP HPLC) 308
10.2.7 Ion pair reversed phase liquid chromatography (IP RP HPLC) 309
10.2.8 Direct infusion mass spectrometry 309

10.3 Results and Discussion 310
10.3.1 Hydrophilic interaction liquid chromatography (HILIC) 310
10.3.2 Reversed phase HPLC (RP HPLC) 322
10.3.3 Ion pair reversed phase liquid chromatography (IP RP HPLC) 324
10.3.4 Comparison between HILIC and IP RP HPLC 331
10.4 Conclusions 332
11. Conclusions and future work 337
11.1 Thesis conclusions 337
11.2 Future work 342
11.2.1 Analytical challenges 342
11.2.2 Challenges in microalgal metabolomics 344
11.2.3 Challenges in mammalian metabolomics 345
References 347
Appendices 365

Acknowledgements

It is with immense gratitude that I acknowledge the support and help of my supervisor Dr. S. Vaidyanathan, who not only served as my supervisor but also encouraged and challenged me throughout my research work. He has taken pain to go through the project and make necessary correction as and when needed with attention and care. Special thanks goes to Prof. Nicola Brown and Dr Carolyn Staton from Medical School for providing me guidance on biology of adherent mammalian cell cultures. I would also like to thank Prof. Phillip Wright (HoD) for providing me the financial support during my final year of PhD studies and Dr Mark Dickman for his throughout support and guidance in LC-MS techniques.

I would also like to extend my thanks to Dr Caroline Evans, Dr Trong Khoa Pham, Dr Narciso Couto, Dr Esther Karunakaran and Dr Jagroop Pandhal for their considerable guidance with mass spectrometers.

I share the credit of my work with Dr Yimin Chen, Dr Joseph Longworth and Gloria Padmaperuma who provided me the microalgal cultures and Rachael Coyle who provided me the mammalian cells throughout my experimental work. I would also like to thank my housemate Dr Mahendra Raut for his support throughout my PhD studies and my Tech-savvy colleague and housemate Deeptanshu Pandey for his support in computing skills during my writing-up period.

I would also thank The University of Sheffield, ChELSI Institute, Medical School, my faculty members, colleagues and technical team without whom this project would have been a distant reality.

Most of all I would like to thank my wife Rohini Sonawani. Her love and support throughout this endeavor has been never-ending and amazing. Thank you sweetheart. Finally I would like to extend my heartfelt thanks to my parents Vijay and Swati Kapoore, my sweet sister Sumati Deoghare and my brother-in-law Mayur Deoghare for providing me encouragement and financial support throughout my life.

List of Publications

[R1] Kapoore Rahul[1], Coyle Rachael[2], Carolyn A Staton, Brown Nicola[*], Vaidyanathan Seetharaman[*]
A novel optimized method for harvesting, quenching and extraction for GC-MS based metabolomics of adherently growing mammalian cells: MDA 436, MCF 7 and HMEC 1 cell lines. *(Manuscript submitted to Metabolomics. Impact factor: 4.43)*

[R2] Kapoore Rahul[1], Coyle Rachael[2], Carolyn A Staton, Brown Nicola[*], Vaidyanathan Seetharaman[*]
Effective washing and quenching processes for physiologically valid metabolite profiling of adherently grown metastatic cancer cell lines: MDA-MB-231. *(Manuscript submitted to Metabolomics. Impact factor: 4.43)*

[R3] Kapoore Rahul[1], Padmaperuma Gloria[2], Vaidyanathan Seetharaman[*]
Optimisation and comparative evaluation of conventional and direct trans-esterification methods for accurate and reliable quantification of FAMEs from micro-algal biofuel strains. *(Manuscript under preparation for submission to Analytica Chimica Acta. Impact factor: 4.34)*

[R4] Kapoore Rahul, Vaidyanathan Seetharaman[*]
GC-MS based optimised sampling, quenching and extraction workflow for obtaining high quality metabolomic data for microalgal strain: *Chlamydomonas reinhardtii.*
(Manuscript under preparation for submission to Metabolomics. Impact factor: 4.43)

[R5] Pugh John[1], Babakar Mevan[2], Kapoore Rahul[3], Vaidyanathan Seetharaman[*]
In situ Metabolomics via Vapour-mediated Ion Activation Enhanced SIMS
(Manuscript under preparation for submission to Analytica Chimica Acta. Impact factor: 4.34)

[R6] Chen Yimin[1], Kapoore Rahul[2], Vaidyanathan Seetharaman[*]
Influence of nutrient status on the accumulation of algal biomass and lipids.
(Manuscript submitted to Bioresource technology. Impact factor: 5.17)

[R7] Chen Yimin[1], Kapoore Rahul[2], Vaidyanathan Seetharaman[*]
Influence of UVA on growth rate and lipid production in *N.salina* and *D.salina*.
(Manuscript submitted to Bioresource technology. Impact factor: 5.17)

[R8] **Longworth Joseph[1], Kapoore Rahul[2], Wright Phillip[*], Vaidyanathan Seetharaman[*]**
Comparison of mixotrophic and photoautotrophic growth of *Chlamydomonas reinhardtii* grown under turbidostatic conditions: A Proteomic and metabolic investigation.
(Manuscript submitted to Proteomics. Impact factor: 4.13)

Oral and Poster presentations

[**P** – Poster presentation; **O/P** – Both oral and poster presentation; **C** – Conferences attended]

[P1] Babakar Mevan[1], **Kapoore Rahul[2]**, Pugh John[2], Salim Malinda[2], Vaidyanathan Seetharaman[*]
"Biochemical microscopy: ToF SIMS" KTN Industry Day, The University of Sheffield, **Sheffield, U.K., Feb. 2014**

[O1/P2] Kapoore Rahul[1], Coyle Rachael[2], Brown Nicola[*], Vaidyanathan Seetharaman[*]
"Development & Optimization of GC-MS based global metabolite profiling of adherent mammalian cells" Metabolomics 2013, 9th Annual International Conference of the Metabolomics Society, Glasgow, **Scotland, U.K., July 2013**

[C1] Algae Biotechnology Sheffield Network First Annual Conference.
"Research in Algae Biotechnology at the University of Sheffield"
Ideas Space, Chemical and Biological Engineering, **Sheffield, U.K., July 2013**

[O2/P3] Kapoore Rahul[1], Chen Yimin[2], Longworth Joseph[3], Vaidyanathan Seetharaman[*]
"GC-MS based metabolite profiling in microalgae" Metabolomics 2012, Wardman Marriott Park Hotel, **Washington D.C., U.S.A., June 2012** *(Won second price for poster presentation)*

[P4] Pugh John[1], Salim Malinda[2], **Kapoore Rahul[3]**, Vaidyanathan Seetharaman[*]
"*In situ* Metabolomics via Vapour-mediated Ion Activation Enhanced SIMS" Metabolomics 2012, Wardman Marriott Park Hotel, **Washington D.C., U.S.A., June 2012**

[P5] Longworth Joseph[1], **Kapoore Rahul[2]**, Wright Phillip[*], Vaidyanathan Seetharaman[*]
"Carbon source effects on *Chlamydomonas reinhardtii*: A proteomic & metabolic investigation" ChELSI Conference, **Sheffield, U.K., 2012**

[C2] Agilent Applied Metabolomics.
"RECENT ADVANCES AND FUTURE DIRECTIONS IN APPLIED METABOLOMICS"
University of Manchester Centre of Excellence in Biopharmaceuticals, **Manchester, U.K., Feb. 2012**

List of abbreviations

µL	microliter
µM	micrometre
1D GC	One dimensional Gas chromatography
2D GC x GC	Two dimensional gas chromatography
2D LC-MS	Two dimensional Liquid chromatography
ACN	Acetonitrile
ADP	Adenosine diphosphate
AMDIS	Automated Mass Spectral Deconvolution and Identification system
AMP	Adenosine monophosphate
ANOVA	Analysis of variance
APCI	Atmospheric pressure chemical ionization
ATP	Adenosine triphosphate
BSA	Bistrimethylsilylacetamide
BSTFA	N,O-bis-(trimethylsilyl)-trifluoroacetamide
CCAP	Culture collection of algae and protozoa
CCMN	Cross-Contribution Compensating Multiple Standard Normalization
CE-MS	Capillary electrophoresis-Mass spectrometry
CID	Collision induced dissociation
Da	Daltons
DESI	Desorption electrospray ionization
DIMS	Direct Infusion Mass Spectrometry
DOE	Design of experiments
EESI	Extractive Electrospray Ionization
EI	Electron-Impact
ESI	Electrospray ionization
FTICR	Fourier transform ion cyclotron resonance mass spectrometry
FTIR	Fourier transform infrared spectroscopy
GC	Gas chromatography
GC-MS	Gas Chromatography-Mass spectrometry
GC-TOF-MS	Gas chromatography-Time-of-flight mass spectrometry
HCA	Hierarchical cluster analysis
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High performance liquid chromatography
Hz	Hertz
LC-MS	Liquid chromatography-Mass spectrometry
m	meter
m/z	mass/charge
MALDI-MS	Matrix assisted laser desorption ionization-Mass spectrometry
mg	milligram
mL	milliliter

mm/s	millimetre per second
mmol	Millimole
mRNA	Messenger Ribonucleic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MST	Mass spectral tag
MSTFA	N-methyl-N-trimethylsilyltrifluroacetamide
ng	nanogram
NIST	National Institute of Standards and Technology
nm	nanometer
NMR	Nuclear magnetic resonance
NOMIS	Normalization using Optimal selection of Multiple Internal Standards
OD	Optical density
ORF	Open reading frame
PARAFAC	Parallel factor analysis
PCA	Principal component analysis
PFK	Phosphofruktokinase
PK	Pyruvate kinase
pK_a	Dissociation Constant
PLS-DA	Partial Least Squares Discriminant Analysis
pmol	Picomole
ppm	Part(s) per million
QC	Quality control
Q-ToF	Quadrupole Time-of-flight
RI	Retention index
RNA	Ribonucleic acid
RP	Reverse phase
RP-HPLC	Reverse phase High performance liquid chromatography
RT	Retention time
SIM	Selective Ion Monitoring
SIMCA	Soft independent modelling of class analogy
TAP	Tris-Acetate-Phosphate
TBS	t-butyldimethylsilylation
TCA	Tricarboxylic acid cycle
TIC	Total ion chromatogram
TMS	Trimethylsilyl
TOF	Time-of-flight
UPLC	Ultra-high performance liquid chromatography
v/v	Volume by volume
w/v	Weight by volume
XML	Extensible markup language

List of figures

Figure 2.1 Classes of compounds and the analytical techniques with which they are most compatible. 15

Figure 2.2 General flow chart for mass spectrometry (figure modified from Feng et al., 2008). 25

Figure 2.3 A Typical flame ionisation detector design (Figure modified from http://www.chromacademy.com). 49

Figure 3.1 Overview of different metabolomic workflows applied across different adherently growing mammalian cell lines. a) Trypsinization b) Cell scraping c) Modified cell scraping method and d) Bead harvesting e) Direct quenching and extraction with LN_2. 60

Figure 3.2 Effect of trypsinization and cell scraping treatment on leakage of intracellular metabolites during quenching was compared for three different cells lines, MDA-MB-436, MCF-7 and HMEC-1. X-axis represents different sampling protocols; Trp_436= trypsinization on MDA-MB-436; Csr_436 = cell scraping on MDA-MB-436; Trp_MCF7 = trypsinization on MCF-7; Csr_MCF7 = cell scraping on MCF7; Trp_Hmec = trypsinization on HMEC-1 and Csr_Hmec = cell scraping on HMEC-1. Y-axis represents median number of metabolites of each class. After both the treatments the extracted metabolites from cell extracts, cell-free supernatant post quenching and blank samples were analysed by GC-MS. (a) metabolites identified in cell extracts only, (b) metabolites identified in supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both). 66

Figure 3.3 Effect of the sampling protocols on leakage of intracellular metabolites during quenching for MDA-MB-436.After all the treatment the extracted metabolites from cell extract, supernatant following quenching and blank samples were analysed by GC-MS. X-axis represents different sampling protocols; Trp_436 = trypsinization ; Csr_436 = cell scraping; NTcs = modified cell scraping; Ap_436 = trypsinization with acetone precipitation; Bead_436 = bead harvesting and LN_2 = direct quenching and extraction with liquid nitrogen. Y-axis represents median number of metabolites of each class; (a) metabolites identified in cell extracts only, (b) metabolites identified in cell-free supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both). 69

Figure 3.4 Heat maps comparing the response of 11 different classes of metabolite signals identified by GC-MS by each of the applied sampling and harvesting approaches across three different cell lines. A = All metabolites in cell extracts; B = Metabolites present only in cell extracts (unique to cells); C = Metabolites present only in supernatants (unique to supernatants). 73

Figure 4.1 Schematic for determining level of free ATP in response to various washing and quenching processes using luciferase bioluminescence assay in adherently growing metastatic cells MDA-MB-231. 83

Figure 4.2 Schematic displaying workflow for SEM analysis of adherently growing cells. Adherent cells were grown on silicon wafers, which can be removed from the well plates by simply lifting the silicon wafers with forceps and cells along with the silicon wafers can be fixed for SEM analysis. 85

Figure 4.3 Overview of quenching and extraction workflow using modified cell scraping method for metabolome analyses of adherently growing metastatic cancer cell line MDA-MB-231. 1 = Washing step with PBS; 2 = quenching protocol and 3 = extraction protocol. 87

Figure 4.4 Standard curves generated for both luciferase bioluminescence ATP assay (a) and BCA protein assay (b) using mean values of ATP and protein content obtained with set of standard solutions (n=3). 90

Figure 4.5 The effect of washing treatments on intracellular ATP levels of MDA-MB-231: C = No washing step (control); PB1 = washing once with PBS; PB2 = washing twice with PBS; W1 = washing once with water and W2 = washing twice with water. Bars represents the mean of 3 independent experiments ± standard error of the mean (SEM). 91

Figure 4.6 The effect of five quenching reagents on intracellular ATP levels of MDA-MB-231: C = unquenched sample (control); 100%M = quenching with 100% methanol; 60%M = quenching with 60% methanol; AMB = quenching with 60% aqueous methanol supplemented with 0.85% AMBIC; HEP = quenching with 60% aqueous methanol supplemented with 70 mM HEPES and LN_2 = simultaneous quenching and extraction with LN_2 followed by extraction with 100% methanol. Bars represents the mean of 3 independent experiments ± standard error of the mean (SEM). 93

Figure 4.7 The effect of various quenching solutions (buffered/non-unbuffered) on MDA-MB-231 cell membrane integrity was visualised and compared against the non-quenched cells using SEM observations, where: 1 = Non-quenched cells (Control); 2 = quenched with 100% methanol; 3 = quenched with 60% aqueous methanol; 4 = quenched with 60% aqueous methanol supplemented with 0.85% AMBIC and 5 = quenched with 60% aqueous methanol supplemented with 70 mM HEPES. 95

Figure 4.8 The effect of five quenching reagents with (single) or without PBS wash step on leakage of intracellular metabolites during quenching was compared for MDA-MB-231. X-axis represents different sampling protocols; 100M_P = 100% methanol with PBS wash step; 100M_WP = 100% methanol without PBS wash step; 60M_P = 60% methanol (aqueous) with PBS wash step; 60M_WP = 60% methanol (aqueous) without PBS wash step; 60MA_P = 60% methanol (aqueous with 0.85% w/v AMBIC) with PBS wash step; 60MA_WP = 60% methanol (aqueous with 0.85% w/v AMBIC) without PBS wash step; 60MH _P = 60% methanol (aqueous with 0.85% w/v HEPES) with PBS wash step; 60MH _WP = 60% methanol (aqueous with 0.85% w/v HEPES) without PBS wash step; C _P = control (unquenched) with PBS wash step; C _WP = control (unquenched) without PBS wash step; LN_2 _P = liquid nitrogen with PBS wash step and LN_2 _WP = liquid nitrogen without PBS wash step. Y-axis represents median number of metabolites of each class. After both the treatment the extracted metabolites from cell extracts, cell-free supernatant post quenching and blank samples were analysed by GC-MS. (a) metabolites identified in cell extracts only, (b) metabolites identified in supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both). 98

Figure 4.9 GC-MS based analysis of intracellular metabolites extracted from adherent MDA-MB-231 mammalian cells. Cells were washed once with PBS and later quenched with either 60% aqueous methanol alone (black), or 60% aqueous methanol + 0.85% AMBIC (red) or 60% aqueous methanol + 70 mM HEPES (green) prior to extraction of metabolites using cold methanol. For clarity, the Y-axis for 60% aqueous methanol and for 60% aqueous methanol + 0.85% AMBIC has been offset. 101

Figure 5.1 Distribution of metabolite classes detected in metabolome of two metastatic cancer cell lines namely MDA-MB-231 and MDA-MB-436 using two quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN_2). Y-axis represents average number of metabolites identified in the cell pellets. X-axis represents distribution of metabolite classes where 1 = organic acids (non-fatty) and derivatives; 2 = sugar/sugar alcohols and derivatives; 3 = amino acids and derivatives; 4 = nucleotides, nucleosides and nucleobases; 5 = fatty acids/fatty alcohols and derivatives; 6 = biogenic amines/polyamines; 7 = phosphates; 8 = alkanes; 9 = alcohols (others); 10 = ketones and ethers; 11 = others. 107

Figure 5.2 Median intensities of metabolites recovered in individual metabolite class with two metastatic cancer cell lines namely MDA-MB-231 and MDA-MB-436 across two different quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN_2). X-axis represents method and cell line investigated. Y-axis represents median intensities of metabolites recovered in individual metabolite within each metabolite class. In total 157 metabolites (See Appendix 5.1) were identified across all the treatments which were classified into 11 different metabolite classes and plotted separately. 109

Figure 5.3 Heat map comparing the response of 11 different classes of metabolite signals identified by GC-MS in metabolome of two metastatic cancer cell lines namely MDA-MB-231 and MDA-MB-436 using two quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with liquid nitrogen (LN_2). 111

Figure 5.4 PCA analysis of metabolomics data displaying clustering of samples between two metastatic cancer cell lines with both the proposed methods: NTcs (left) and LN_2 (right). PCA is calculated using the feature intensities from all samples. The colours (red/blue) are assigned based on the sample class. Distance to the model (DModX) test was used to verify the presence of outliers and to evaluate whether a submitted sample fell within the model applicability domain. 113

Figure 5.5 Cloud plots generated for four different comparison (C vs D (top left); A vs B (top right); A vs C (bottom left) and B vs D (bottom right)) as listed in Table 5.2, displaying significant features with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value. 114

Figure 5.6 In total 84 identified metabolites with both methods (NTcs and LN2) in two metastatic cancer cell lines namely MDA-MB-231 and MDA-MB-436 were found to be present in cancer pathways (involved in TNBC biology), which were plotted on KEGG's metabolic map and highlighted by red dot. 117

Figure 6.1 Representative chromatogram from *E. coli*, GC-MS metabolomic analysis. The sample was quenched with 60% aqueous methanol and extracted with pure methanol. The extracted metabolites were derivatized by methoximation followed by silylation. The derivatized sample volume of 1 µL was injected in split less mode at 230°C and separated on a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). In total, around 162 metabolites were identified (see table 6.8), where few identified peaks were labelled onto a chromatogram. 134

Figure 6.2 Representative chromatogram from *N. salina*, GC-MS metabolomic analysis. The sample was quenched with 60% aqueous methanol and extracted with methanol-chloroform mixture (2:1 *v/v*). The extracted metabolites were derivatized by methoximation followed by silylation. The derivatized sample volume of 1 µL was injected in split less mode at 230°C and separated on a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). In total, around 116 metabolites were identified (see table 6.8), where few identified peaks were labelled onto a chromatogram. 137

Figure 6.3 Representative chromatogram from *D. salina*, GC-MS metabolomic analysis. The sample was quenched with 60% aqueous methanol and extracted with methanol-chloroform mixture (2.5:2 *v/v*). The extracted metabolites were derivatized by methoximation followed by silylation. The derivatized sample volume of 1 µL was injected in split less mode at 230°C and separated on a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). In total, around 103 metabolites were identified (see table 6.8), where few identified peaks were labelled onto a chromatogram. 139

Figure 6.4 Representative chromatogram from *C. reinhardtii*, GC-MS metabolomic analysis. The sample was quenched with 60% aqueous methanol and extracted with methanol-chloroform-water mixture (5:2:2 *v/v*). The extracted metabolites were derivatized by methoximation followed by silylation. The derivatized sample volume of 1 µL was injected in split less mode at 230°C and separated on a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). In total, around 139 metabolites were identified (see table 6.8), where few identified peaks were labelled onto a chromatogram. 140

Figure 6.5 *C. reinhardtii* central metabolism is visualised using the MapMan tool. Squares represent Chlamydomonas proteins that have been assigned into various pathways. White boxes indicates 159 identified metabolites, red boxes indicated 1069 identified proteins by MS, whereas blue boxes indicates proteins with no peptide support. Figure adopted from (May et al., 2008a). 142

Figure 6.6 Venn diagram displaying the capture of *C. reinhardtii* metabolome in our study in comparison with previously published reports by various researchers (Bölling and Fiehn, 2005, May et al., 2009, May et al., 2008a). Dark blue colour represents 18 metabolites identified by May and co-workers alone (May et al., 2008a), purple colour represents addition of one new metabolite reported by Fiehn and co-workers (Bölling and Fiehn, 2005) to that of 18 metabolites reported by May and co-workers, light blue colour represents 29 new metabolites identified in later studies by both authors in addition to those 18 that were reported by May and co-workers in their previous studies (May et al., 2009), whereas pink colour represents 32 new identified metabolites (see Appendix 6.3) in current study in addition to all of those previously reported. 143

Figure 6.7 Distribution of the metabolite classes detected in metabolome of four different biological species investigated namely *E. coli* (blue)*, C. reinhardtii* (red), *N. salina* (green) *and D. salina* (purple). Y-axis represents average number of metabolites identified. X-axis represents distribution of 13 metabolite classes. 146

Figure 6.8 Summary of conditions studied to investigate the influence of carbon source and availability on *C. reinhardtii* cultures grown under different trophic conditions. 149

Figure 6.9 PCA analysis displaying clustering of *C. reinhardtii* samples for all four culture conditions as summarised in Figure 6.8, where: (a) PCA for metabolomic data and (b) PCA for proteomics data (figure adopted from (Longworth et al, 2013). For metabolomics data, PCA is calculated using the feature intensities from all samples. The colours (red/blue) are assigned based on the sample class. Distance to the model (DModX) test was used to verify the presence of outliers and to evaluate whether a submitted sample fell within the model applicability domain. For proteomics data PCA is based on variance stabilization normalisation (VSN), isotope correction (IC), median corrected reporter ion intensities, limited to ≥ 2 peptide proteins. .. 151

Figure 6.10 Cloud plot generated by comparing condition A against remaining three conditions (see Figure 6.8), displaying 488 features with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value. ... 152

Figure 6.11 PCA analysis for metabolomics data displaying clustering of samples between condition A and condition D. PCA is calculated using the feature intensities from all samples. The colours (red/blue) are assigned based on the sample class. Distance to the model (DModX) test was used to verify the presence of outliers and to evaluate whether a submitted sample fell within the model applicability domain. 154

Figure 6.12 Cloud plot generated by comparing condition A against condition D (see Figure 6.8), displaying 289 features with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value. 154

Figure 6.13 Variations observed in five different metabolite classes across Low Carbon (condition A – blue colour) and High Carbon (condition D – red colour) culture conditions in *C. reinhardtii* (A vs D). X-axis represents unique number of metabolites identified as listed in Appendix 6.3; whereas Y-axis represents median intensities obtained for individual metabolites. 155

Figure 6.14 Cloud plot generated by comparing six two way comparison for each condition against each of the other, displaying significantly different features with p-value ≤ 0.01 and fold change ≥ 1.3. All the compared culture conditions for *C. reinhardtii* are summarised in Figure 6.8. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value. 157

Figure 6.15 Volcano plots displaying the significant changes observed when comparing the various culture conditions of *C. reinhardtii* among them. Details of the culture conditions compared above are summarised in Figure 6.8. Proteins with significant changes ($p < 0.01$) are shown by red star. The more relaxed cut-off of $p < 0.05$ is indicated by a solid line and the more stringent cut-off of p < 0.01 shown by a dotted line (figure adopted from (Longworth, 2013)). 158

Figure 7.1 Experimental design and general workflow adopted for the evaluation and optimisation of quenching protocols for *C. reinhardtii*. 172

Figure 7.2 Schematic displaying the necessary correction required for appropriate calculation of intracellular metabolites after monitoring cell-free supernatant of quenched, non-quenched cells and the blank medium. 173

Figure 7.3 A summary of the unique recovery efficiency of all the nine quenching treatments involved in Approach 1 and 3. X-axis represents different sampling protocols: 33M = 33% methanol; 33A = 33% methanol + AMBIC; 60M = 60% methanol; 60A = 60% methanol + AMBIC; 60H10 = 60% methanol + 10 mM HEPES; 60H70 = 60% methanol + 70 mM HEPES; 70M = 70% methanol; 70A = 70% methanol + AMBIC; 100M = 100% methanol; 60-41 = 60% methanol with solvent to sample ratio 4:1 and 100-21 = 100% methanol with solvent to sample ratio 2:1. Y-axis represents median number of metabolites of each class. After all the treatment the extracted metabolites from cell extracts, cell-free supernatant post quenching and blank samples were analysed by GC-MS. (a) metabolites identified in cell extracts only, (b) metabolites identified in supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both). 174

Figure 7.4 Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis across various quenching treatments (from approach 1 and 3) was compared. X-axis represents different quenching protocols with their respective control sample where: 33M = 33% methanol; 33A = 33% methanol + AMBIC; 60M = 60% methanol; 60A = 60% methanol + AMBIC; 60H10 = 60% methanol + 10 mM HEPES; 60H70 = 60% methanol + 70 mM HEPES; 70M = 70% methanol; 70A = 70% methanol + AMBIC; 100M = 100% methanol; 60-41 = 60% methanol with solvent to sample ratio 4:1; 100-21 = 100% methanol with solvent to sample ratio 2:1 and SN = supernatant. Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 197 metabolites (See Appendix 7.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately. 177

Figure 7.5 A summary of the unique recovery efficiency of all the six quenching treatments involved in Approach 2. X-axis represents different time intervals along with their control samples (C). Y-axis represents median number of metabolites of each class. After all the treatment the extracted metabolites from cell extracts, cell-free supernatant post quenching and blank samples were analysed by GC-MS. (a) metabolites identified in cell extracts only, (b) metabolites identified in supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both). 186

Figure 7.6 Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis across various quenching treatments (from approach 2) was compared. X-axis represents different time intervals in minutes along with their control samples (C) whereas, SN = supernatant. Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 165 metabolites (See Appendix 7.2) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately. 188

Figure 8.1 Experimental design and general workflow adopted for the evaluation and optimisation of extraction protocols for *C. reinhardtii*. 207

Figure 8.2 Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis using various monophasic extraction solvent systems was compared. X-axis represents different extraction protocols where: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol; 75-A = 75% aqueous acetonitrile; 75-M = 75% aqueous methanol; 100-A = pure acetonitrile; 100-E = pure ethanol; 100-M = pure methanol; AMW = acetonitrile/methanol/water (2:2:1) and WiPM = water/isopropanol/methanol (2:2:5). Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 162 metabolites (See Appendix 8.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately. 209

Figure 8.3 Extraction efficiency of various monophasic extraction solvents was compared in recovery of average numbers of metabolites from *C. reinhardtii* cell extracts, based on GC-MS analysis. X-axis represents different extraction protocols where: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol; 75-A = 75% aqueous acetonitrile; 75-M = 75% aqueous methanol; 100-A = pure acetonitrile; 100-E = pure ethanol; 100-M = pure methanol; AMW = acetonitrile/methanol/water (2:2:1) and WiPM = water/isopropanol/methanol (2:2:5). Y-axis represents average number of metabolites recovered. 215

Figure 8.4 Addition of Sudan I granules (A lipid soluble diazo dye used to aid visualisation of phase separation) to Eppendorf tube containing volumes of aqueous and organic components as described in table 8.3, in order to visulise the chloroform phase and estimate the extent of mixing with the methanol/water. Where (a) = without *C. reinhardtii* culture and (b) = with *C. reinhardtii* culture. Colour banding within the aqueous phase (upper) in fig (a) is due to rediffusion of chloroform into the aqueous phase. 221

Figure 8.5 Snapshot of representative samples processed with BI and RK method. (a) = M:C:W – Acetonitrile solvent system (BI) as used by Bi and co-workers (Bi et al., 2013); whereas (b) = M:C:W (2.5:3:1.5) solvent system (RK) with optimised solvent ratios leading to optimal phase separation. 221

Figure 8.6 Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis using various biphasic extraction solvent systems was compared. X-axis represents different extraction protocols where: G = methanol/chloroform/water (5:2:2); BI = methanol/chloroform/water – With inclusion of acetonitrile for protein precipitation; MB = Modified BI protocol, methanol/chloroform/water (2.5:3:1.5) – With inclusion of acetonitrile for protein precipitation; RK = Optimised solvent ratios for MCW mixture, methanol/chloroform/water (2.5:3:1.5). Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 162 metabolites (See Appendix 8.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately. 223

Figure 8.7 Extraction efficiency of various biphasic extraction solvents was compared in recovery of average numbers of metabolites from *C. reinhardtii* cell extracts, based on GC-MS analysis. X-axis represents different extraction protocols where: G = methanol/chloroform/water (5:2:2); BI = methanol/chloroform/water – With inclusion of acetonitrile for protein precipitation; MB = Modified BI protocol, methanol/chloroform/water (2.5:3:1.5) – With inclusion of acetonitrile for protein precipitation; RK = Optimised solvent ratios for MCW mixture, methanol/chloroform/water (2.5:3:1.5). Y-axis represents average number of metabolites recovered. 226

Figure 8.8 Extraction efficiency of various monophasic and biphasic extraction solvents was compared in recovery of median numbers of metabolites within each chemical class from *C. reinhardtii* cell extracts, based on GC-MS analysis. X-axis represents different extraction protocols where: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol; 75-A = 75% aqueous acetonitrile; 75-M = 75% aqueous methanol; 100-A = pure acetonitrile; 100-E = pure ethanol; 100-M = pure methanol; AMW = acetonitrile/methanol/water (2:2:1); WiPM = water/isopropanol/methanol (2:2:5); G = methanol/chloroform/water (5:2:2); BI = methanol/chloroform/water – With inclusion of acetonitrile for protein precipitation; MB = Modified BI protocol, methanol/chloroform/water (2.5:3:1.5) – With inclusion of acetonitrile for protein precipitation; RK = Optimised solvent ratios for MCW mixture, methanol/chloroform/water (2.5:3:1.5). Y-axis represents median number of metabolites. 227

Figure 8.9 Pie charts representing extraction efficiency of various monophasic and biphasic extraction solvents in recovery of average percentage numbers of metabolites within each of the twelve metabolites classes from *C. reinhardtii* cell extracts. 229

Figure 8.10 GC-MS based analysis of intracellular metabolites extracted from *C. reinhardtii* with four different extraction solvents. Cells were quenched with 60% aqueous methanol supplemented with 70 mM HEPES prior to extraction of metabolites with various extraction solvents. X-axis represents retention time in minutes, whereas Y-axis represents relative abundance. The raw chromatograms obtained from GC-MS analysis with various extraction solvents were labelled where: A = 100% methanol (100-M method); B =

methanol/chloroform/water (5:2:2) (G method); C = 25% aqueous acetonitrile (25-A) and D = 25% aqueous methanol (25-M). 233

Figure 8.11 Deviations of retention time of all samples are shown in relation to the retention time of the first sample in the legend (black line) (OBI-Warp method) and in relation to the median retention time of features from all samples (peak groups). 234

Figure 8.12 Cloud plot of extraction methods data set displaying 517 features with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value. 235

Figure 8.13 Non-metric multidimensional scaling (MDS) plot was generated based on the median intensities of all aligned features detected across all the applied extraction methods, in order to explore similarities or dissimilarities between the sample data sets. The colours (red/blue) are assigned based on the extraction solvent system employed where: Red = 100-M (pure methanol) and Blue = other solvents systems which includes: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol and G = methanol/chloroform/water (5:2:2). A, B and C = biological replicates. 236

Figure 9.1 A representative snapshot of the samples processed by DT methods using base catalyst alone and sequential use of both base and acid catalyst. Visual differences on derived biodiesel (top organic layer) across different applied methods (summarised in Table 9.2) can be seen clearly. 254

Figure 9.2 Effect of 10 different bead beating treatments on recovery of FAMEs was compared for *C. reinhardtii*. X-axis represents different applied treatments. Each treatment comprise of 2 factor combination generated by design expert software using 2 factorial design (Factor1: Cycle / Factor2: Time), where A = 1/1; B = 1/2; C = 10/1; D = 10/2; E = 3/1; F = 3/2; G = 5/1; H = 5/2; I = 7/1 and J = 7/2. Y-axis represents amount of FAMEs recovered in mg/L per DCW. After all the treatments the extracted lipids from cell extracts were analysed by GC-FID: (a) = Overall FAME recovery; (b) = recovery of low concentration FAMEs and (c) = recovery of high concentration FAMEs. 260

Figure 9.3 Effect of 10 different bead beating treatments on recovery of FAMEs was compared for *N. salina*. X-axis represents different applied treatments. Each treatment comprise of 2 factor combination generated by design expert software using 2 factorial design (Factor1: Cycle / Factor2: Time); A = 1/1; B = 1/2; C = 10/1; D = 10/2; E = 3/1; F = 3/2; G = 5/1; H = 5/2; I = 7/1 and J = 7/2. Y-axis represents amount of FAMEs recovered in mg/L per DCW. After all the treatments the extracted lipids from cell extracts were analysed by GC-FID: (a) = Overall FAME recovery; (b) = recovery of low concentration FAMEs and (c) = recovery of high concentration FAMEs. 262

Figure 9.4 Effect of 27 different sonication treatments on recovery of FAMEs was compared for *C. reinhardtii*. X-axis represents different applied treatments. Each treatment comprise of 3 factor combination generated by design expert software using 3 factorial design (Factor1: Time / Factor2: Duty cycle / Factor3: Output power). In 1101, first 1 digit stands for time factor (range 1; 2 and 3 minutes); 10 digit stands for duty cycle factor (range 10; 20 and 30%) and last 1 digit stands for output factor (range 1; 3 and 6). Similar trend was followed for all the coding system for further 26 treatments; Y-axis represents amount of FAMEs recovered in mg/L per DCW. After all the treatments the extracted lipids from cell extracts were analysed by GC-FID: (a) = Overall FAME recovery; (b) = recovery of low concentration FAMEs and (c) = recovery of high concentration FAMEs. 264

Figure 9.5 Effect of 27 different sonication treatments on recovery of all FAMEs was compared for *N. salina*. X-axis represents different applied treatments. Each treatment comprise of 3 factor combination generated by design expert software using 3 factorial design (Factor1: Time / Factor2: Duty cycle / Factor3: Output power). In 1101, first 1 digit stands for time factor (range 1; 2 and 3 minutes); 10 digit stands for duty cycle factor (range 10; 20 and 30%) and last 1 digit stands for output factor (range 1; 3 and 6). Similar trend was followed for all the coding system for further 26 treatments; Y-axis represents amount of FAMEs recovered in mg/L per DCW. 266

Figure 9.6 10 Effect of 27 different sonication treatments on recovery of FAMEs was compared for *N. salina*. X-axis represents different applied treatments. Each treatment comprise of 3 factor combination generated by design expert software using 3 factorial design (Factor1: Time / Factor2: Duty cycle / Factor3: Output power). In 1101, first 1 digit stands for time factor (range 1; 2 and 3 minutes); 10 digit stands for duty cycle factor (range 10; 20 and 30%) and last 1 digit stands for output factor (range 1; 3 and 6). Similar trend was followed for all the coding system for further 26 treatments; Y-axis represents amount of FAMEs recovered in mg/L per DCW. After all the treatments the extracted lipids from cell extracts were analysed by GC-FID: (a) = Recovery of low concentration FAMEs; (b) = recovery of medium concentration FAMEs and (c) = recovery of high concentration FAMEs. 267

Figure 9.7 Effect of bead beating and sonication treatments on recovery % total FA content per dry cell weight was compared for *C. reinhardtii* and *N. salina*. X-axis in (a) and (b) represents different applied treatments. Each treatment comprise of 2 factor combination generated by design expert software using 2 factorial design (Factor1: Cycle / Factor2: Time); A = 1/1; B = 1/2; C = 10/1; D = 10/2; E = 3/1; F = 3/2; G = 5/1; H = 5/2; I = 7/1 and J = 7/2, whereas X-axis in (c) and (d) represents different applied treatments. Each treatment comprise of 3 factor combination generated by design expert software using 3 factorial design (Factor1: Time / Factor2: Duty cycle / Factor3: Output power). In 1101, 1 digit stands for time factor (range 1; 2 and 3 minutes); 10 digit stands for duty cycle factor (range 10; 20 and 30%) and last 1 digit stands for output factor (range 1; 3 and 6). Similar trend was followed for all the coding system for further 26 treatments; Y-axis represents % total FA content per DCW. After all the treatment the extracted lipids from cell extracts were analysed by GC-FID: (a) = % total FA content per DCW by bead beating treatments in *C. reinhardtii*; (b) = % total FA content per DCW by bead beating treatments in *N. salina*; (c) = % total FA content per DCW by sonication treatments in *C. reinhardtii* and (d) = % total FA content per DCW by sonication treatments in *N. salina*. 270

Figure 9.8 The cell disruption efficiency of bead beating and sonication treatments on *C. reinhardtii* cell integrity was visualised and compared using SEM observations, where: 1 = untreated wet algal biomass; 2 = bead beating 1_2 treatment with lowest FAME recovery; 3 = bead beating 5_2 treatment with highest FAME recovery; 4 = sonication 3103 treatment with highest FAME recovery and 5 = sonication 1106 with lowest FAME recovery. 273

Figure 9.9 The cell disruption efficiency of bead beating and sonication treatments on *D. salina* cell integrity was visualised and compared using SEM observations, where: 1 = untreated wet algal biomass; 2 = bead beating 1_2 treatment with lowest FAME recovery; 3 = bead beating 5_2 treatment with highest FAME recovery; 4 = sonication 3103 treatment with highest FAME recovery and 5 = sonication 1106 with lowest FAME recovery. 274

Figure 9.10 The cell disruption efficiency of bead beating and sonication treatments on *N. salina* cell integrity was visualised and compared using SEM observations, where: 1 = untreated wet algal biomass; 2 = bead beating 1_2 treatment with lowest FAME recovery; 3 = bead beating 5_2 treatment with highest FAME recovery; 4 = sonication 3103 treatment with highest FAME recovery and 5 = sonication 1106 with lowest FAME recovery. 275

Figure 9.11 Retention time deviations for sequence 1 and 2 are summarised in figure (a) and (b) respectively; whereas calibration curve plots generated for sequence 1 and 2 are displayed in figure (c) and (d)n respectively. Both the plots were generated for C16:0 FAME from *N. salina*, where in sequence 1, 24 samples were ran wheras in sequence 2, 72 samples were ran in total as can be seens in circular froms in figure (a) and (b). The first circle represents calibration standard run; then all samples and the last circle represents the check standard run. The green triangular sign in the calibration curve plots indicate check standard run in both the sequences. 277

Figure 9.12 Representative GC-FID chromatogram of 37 FAME Mix (C8-C24) standard from Supelco (a) and of *N. salina* sample (b). The *N. salina* cells were extracted with toluene and directly transesterified with 0.5M sodium methoxide (alkali catalysts) and later with 5% acetyl chloride in methanol (acid catalyst). The trans-esterified sample volume of 1 µL was injected in split mode at 250°C and separated on a TR-FAME capillary column (25 m x 0.32 mm x 0.25 µm). 279

Figure 9.13 Total lipid contents estimated by conventional methods using gravimetric determination were expressed in mg/DCW of biomass in *C. reinhardtii*, *D. salina* and *N. salina*. The details of codes from 1A to 5 are belonging to conventional methods the details of which are described in Table 9.2. 282

Figure 9.14 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery total FAs content and composition were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total FA content per DCW and individual FA composition expressed in mg/L per DCW. 283

Figure 9.15 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery low concentration FAMEs were compared for.................................... 284

Figure 9.16 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery medium concentration FAMEs were compared.................................. 285

Figure 9.17 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery high concentration FAMEs were compared for 286

Figure 9.18 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery saturated fatty acids (SFAs) content were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total SFA content per DCW and individual FA composition expressed in mg/L per DCW. 291

Figure 9.19 Figure 9.20 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery monounsaturated fatty acid contents (MUFAs) concentration FAMEs were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total MUFA content per DCW and individual FA composition expressed in mg/L per DCW.. 292

Figure 9.20 Figure 9.21 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery polyunsaturated fatty acid contents (PUFAs) concentration FAMEs were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total PUFA content per DCW and individual FA composition expressed in mg/L per DCW.. 293

Figure 10.1 Representative individual chromatograms obtained for phenylalanine, tryptophan and tyrosine (100 µM) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of either 206, 240 or 254 nm. 311

Figure 10.2 Representative chromatogram obtained for mixture of three aliphatic amino acids (100 µM) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of either 206, 240 or 254 nm.. 312

Figure 10.3 Representative chromatogram obtained for mixture of nine (left) and eight (right) amino acids (100 µM) (See Table 10.3) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of either 206, 240 or 254 nm. 314

Figure 10.4 Mass spectra obtained for phenylalanine (top) and methionine (bottom), where purified HILIC fractions were collected from an Agilent 1100 HPLC using fraction collector and infused directly on an ion trap mass spectrometer (HCT Ultra PTM Discovery System, Bruker

Daltonics) at a flow rate of 180 µL min^{-1} and a mass range of 50 – 600 m/z. The capillary voltage was kept at -2500 V to maintain capillary current between 30 – 50 nA. 315

Figure 10.5 Mass spectra obtained for tyrosine (top), tryptophan (middle) and glutamine (bottom), where purified HILIC fractions were collected from an Agilent 1100 HPLC using fraction collector and infused directly on an ion trap mass spectrometer (HCT Ultra PTM Discovery System, Bruker Daltonics) at a flow rate of 180 µL min^{-1} and a mass range of 50 – 600 m/z. The capillary voltage was kept at -2500 V to maintain capillary current between 30 – 50 nA. 315

Figure 10.6 Representative individual chromatograms obtained for individual organic acids (100 µM) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 317

Figure 10.7 Representative chromatograms obtained for mixture of ten (top) and twelve (bottom) nucleotides/nucleosides/nucleobases (100 µM) (See Table 10.6) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 319

Figure 10.8 Representative chromatogram obtained for mixture of fourteen nucleotides/nucleosides/nucleobases (100 µM) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 320

Figure 10.9 Representative chromatogram obtained for mixture of eight water soluble vitamins (100 µM) (See Table 10.7) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 321

Figure 10.10 Representative individual chromatograms obtained for individual nucleotides/nucleosides/nucleobases, amino acids, water soluble vitamins and organic acids (100 µM) by RP-HPLC separation on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). RP-HPLC Buffers used: 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of either 206, 240, 254, 260 or 270 nm. 323

Figure 10.11 Efficiency of IP-RP-HPLC separations in retention and separation of organic acids with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at two different concentrations (7 mM and 1 mM). Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 325

Figure 10.12 Efficiency of IP-RP-HPLC separations in retention and separation of organic acids at varying pH conditions with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at 1 mM concentration. Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 325

Figure 10.13 Efficiency of IP-RP-HPLC separations in retention and separation of individual nucleotides, nucleosides and nucleobases with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at two different concentrations (7 mM (left) and 1 mM (right)). Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 327

Figure 10.14 Efficiency of IP-RP-HPLC separations in retention and separation of nucleotides, nucleosides and nucleobases from their 14 mix cocktail at varying pH conditions with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at 1 mM concentration. Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm 328

Figure 10.15 Efficiency of IP-RP-HPLC separations in retention and separation of individual water soluble vitamins with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at two different concentrations (7 mM and 1 mM). Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 329

Figure 10.16 Efficiency of IP-RP-HPLC separations in retention and separation of water soluble vitamins from their 8 mix cocktail as listed in Table 10.7, at varying pH conditions with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at 1 mM concentration. Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm. 330

List of tables

Table 2.1 Approaches used in metabolomics 13

Table 2.2 Comparison of the advantages and disadvantages between sequential and simultaneous sample processing for metabolomics studies. Table modified from (Mashego et al., 2007) 17

Table 4.1 Assay parameters employed for the estimation of ATP content by luciferase bioluminescence assay in MDA-MB-231 cells. The amount of ATP determined was then normalised to protein content estimated by BCA microplate assay 82

Table 4.2 Eleven different metabolite classes along with their class IDs 97

Table 5.1 Median number of metabolites recovered in MDA-MB-231 and MDA-MB-436 across two different quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN2).. 112

Table 5.2 Number of significantly different features across method and cell line comparison113

Table 6.1 Combination of quenching and extraction solvents employed for various species. (M= methanol; C= chloroform; W= water).. 127

Table 6.2 List of metabolites identified by GC-MS along with their retention time from twelve metabolite mixture 133

Table 6.3 Number of metabolites identified in each replicate in *E. coli* metabolomic study (1, 2, 3 = analytical replicates; Control = Unquenched sample) 135

Table 6.4 Number of metabolites identified in each replicate in *N. salina* metabolomics study: A, B and C = different workflows, where A = no quenching and extraction with methanol-chloroform mixture (2:1 *v/v*); B = quenching with 60% aqueous methanol and extraction with methanol-chloroform mixture (2:1 *v/v*) and C = quenching with 60% aqueous methanol and extraction with pure methanol; whereas 1, 2 and 3 = biological replicates............................ 137

Table 6.5 Number of metabolites identified in each replicate in *D. salina* metabolomic study (A, B = Different quenching and extraction protocols, where A = No quenching, B = quenching with 60% aqueous methanol followed by extraction with methanol/chloroform (2.5:2); whereas 1, 2, 3 = Analytical replicates) 139

Table 6.6 Number of metabolites identified in *C. reinhardtii* metabolomics study (A, B = Different quenching protocols, where A = quenching with 32.5% aqueous methanol, B = quenching with 60% aqueous methanol followed by extraction with a, b or c; where a = extraction with methanol/chloroform/water mixture (10:3:1), b = extraction with methanol/chloroform/water mixture (5:2:2) and c = extraction with 100% methanol) 141

Table 6.7 Number of unique metabolites identified in *C. reinhardtii* under each treatment, where: A, B = Different quenching protocols, where A = quenching with 32.5% aqueous methanol; B = quenching with 60% aqueous methanol followed by extraction with a, b or c; where a = extraction with methanol/chloroform/water mixture (10:3:1); b = extraction with methanol/chloroform/water mixture (5:2:2) and c = extraction with 100% methanol 145

Table 6.8 Summary of the total number and percentage of metabolites detected in each class across all the species investigated in this study 145

Table 6.9 Distribution of all the metabolites identified in this research into different classes (Combined results from all the biological species investigated) 147

Table 6.10 Number of significantly changing features in proteome and metabolome of *C. reinhardtii* against various conditions as summarised in Figure 6.8 150

Table 7.1 Summary of the main different quenching solvents and techniques applied so far for the metabolome analysis of different biological systems. (Sing x stands for no information available or provided in the research papers surveyed) 166

Table 7.2 Variation in quenching treatment tested to investigate the effects of varying methanol concentration, effect of inclusion of various buffer additives to quenching solution, effect of temperature and effect of sample to quenching solvent ratio on metabolite leakage 173

Table 8.1 Summary of the main different extraction solvents and techniques applied so far for the metabolome analysis of various prokaryotic and eukaryotic samples 200

Table 8.2 Summary and IDs used for various extraction solvent systems evaluated for their extraction efficiencies in this study which includes: Monophasic and biphasic solvent systems used in the past literature, and in addition the biphasic solvent systems in which the solvent ratios were optimised prior to their use 205

Table 8.3 Increasing the volumes of chloroform by 100 µL to 800 µL aliquots of 80% aqueous methanol to simulate the extraction step. Subsequent addition of 200 µL of water was done in order to help drive phase separation 220

Table 8.4 Increasing the volumes of chloroform by 50 µL to 800 µL aliquots of 80% aqueous methanol to simulate the extraction step. Subsequent addition of 200 µL of water was done in order to help drive phase separation 220

Table 9.1 Bead beating (a) and sonication (b) treatments generated by two and three factorial design respectively using design of expert software 247

Table 9.2 List of all the methods investigated along with their codes used in this study 249

Table 10.1 List of 20 amino acids prepared individually at a concentration of 100 µM 311

Table 10.2 Summary of retention time and wavelength at which individual amino acids were detected 313

Table 10.3 List of amino acids included in 9 and 8 mix cocktails at a concentration of 100 µM 314

Table 10.4 List of 22 organic acids prepared individually at a concentration of 100 µM......... 316

Table 10.5 List of 18 metabolites belonging to nucleotides/nucleosides/nucleobases class prepared individually at a concentration of 100 µM.. 318

Table 10.6 List of 10 and 12 mix of nucleotides/nucleosides/nucleobases prepared at a concentration of 100 µM 319

Table 10.7 List of eight water soluble vitamins 321

Chapter 1

Introduction

1. Introduction

1.1 Background on metabolomics

The post-genome era has seen the emergence of ‘-omic’ approaches to catalogue biochemical events, especially with respect to the expression of proteins and metabolites. The functions of many open reading frames within the sequenced genomes are still unknown. Hence analysis at the level of functional unit like transcripts, proteins & metabolites has increasingly become relevant. When, mRNA gene expression & proteomic analyses do not tell the whole story of what might be happening in a cell, in such cases metabolic profiling can give us instantaneous snapshot of the physiology of that cell, as metabolites are the end products of cellular regulatory processes, and their levels can be regarded as the ultimate response of biological systems to genetic or environmental changes.

In the case of genomics and proteomics the analytical methodologies have stabilized into several well characterized approaches, whereas analytical procedures for metabolomics analysis still remain in a considerable flux due to the diversity and dynamic nature of metabolome. With the current technologies only a portion of the metabolome can be analysed. As a result there is a drive to direct significant analytical efforts towards capturing metabolome or changes in the metabolome reliably and reproducibly in biological systems with high enough precision, resolution and sensitivity. Challenges also exist in increasing analyte throughput and reliably detecting quantitative changes. Most of the metabolites involved in the corresponding pathways are negatively charged and many are phosphorylated. As a result, to date, there is no single analytical method that is suitable for detection of all the metabolite classes. Apart from this, there are several metabolite classes that cannot be analysed with the currently available analytical techniques. Thus, for both targeted and un-targeted metabolomic investigations it is essential to define a uniform, comprehensive, reproducible and user-friendly system of classifying metabolites based on their physicochemical properties, which will serve as a basis for selection of analytical methods/protocols for the different metabolite classes in any biological system. Chromatographic separations coupled to mass spectrometry offer the potential to address these challenges and generate reliable metabolomic data. Hence we propose to investigate and develop the optimized method for metabolomics by examining application of both optimized GC-MS and LC-MS (offline MS) based techniques that has the potential to comprehensively profile the metabolome in a given biological species.

Apart from analytical challenges, the challenges also include development of appropriate methodologies to quench the metabolism before analysis, techniques for reliable and comprehensive extraction of metabolites from the biological matrix, data interpretation using statistical approaches, metabolite identification, optimal detections and quality control and validation of all the steps involved in metabolomics pipeline. Analysis of the metabolome is sample and cell dependent. Differences in cellular characteristics including membrane, wall structure and, size and sampling techniques employed, can influence the efficiency of quenching, recovery of different metabolite classes and the rate of metabolite leakage. To achieve an accurate representation of phenotype, development of rapid quenching and metabolite extraction techniques is a crucial parameter in sample preparation workflows for metabolomic investigations.

1.2 Mammalian metabolomics

Metabolome analyses have been increasingly applied in human studies on body fluids, for various clinical applications to define biomarkers related to prognosis or diagnosis of a disease or drug toxicity/efficacy to provide improved understanding of the pathophysiology of disease or therapeutic pharmacokinetics (Mamas et al., 2011). Metabolomics is being recognised as a valuable tool to evaluate cellular state and the deterioration of structure –pathway-activity-relationships (SPARs) considered suitable for drug discovery (Khoo and Al-Rubeai, 2007). However, as reported previously there have been insufficient reports on the application of metabolomics in the analysis of cultured mammalian cell lines (Čuperlović-Culf et al., 2010, Khoo and Al-Rubeai, 2007). Thus the potential of metabolomics in traditional and emerging biological areas have yet to be fully realized.

The application of metabolomics to mammalian cell culture provides a controlled reproductive but relevant model without the requirement of ethical consideration for *in vivo* and human studies. Data are easier to interpret as focusing on a specific cell line reduces variability and provide a more constant background against which subtle metabolic changes can be identified (Čuperlović-Culf et al., 2010). Consequently, metabolome analyses on immortalized cancer cell lines in cancer research is gaining momentum as a tool to understand the molecular mechanism of disease progression and, response and resistance to therapeutics, leading to the identification of specific biomarkers which may assist distinguishing between normal, benign and metastatic cancer states (Sheikh et al., 2011).

Several studies have been published on optimization of sample pre-processing protocols, for example, on suspension cell culture (Dietmair et al., 2012, Dietmair et al., 2010, Kronthaler et al., 2012a, Sellick et al., 2011, Volmer et al., 2011, Wiendahl et al., 2007), but there are only a few reports on the handling of adherent mammalian cells for GC-MS based metabolomic analysis (Danielsson et al., 2010, Dettmer et al., 2011, Hutschenreuther et al., 2012b). The most commonly used pre-treatments for handling of adherent mammalian cells are associated with the higher metabolome leakage during sampling and quenching process. Therefore, the development of optimised methods/workflows which will improve the metabolome coverage without altering the internal metabolite signatures and with minimal metabolite leakage is required. It is also important to establish the level of metabolome leakage for the specific cell line investigated. Rapid approaches that combine quenching and extraction steps may be more effective in retaining valid metabolome data, with minimal metabolome leakage occurring.

For valid metabolomic data, it is also important to assess and compare the impact of newly developed and proposed methods/workflows with the previously reported methods with respect to metabolome leakage and coverage. Moreover, the proposed protocols should able to track the changes between the cell lines, which share common biological and morphological features, such as two metastatic triple negative breast cancer (TNBC) cell lines, namely MDA-MB-231 and MDA-MB-436. Due to their aggressive phenotypes, TNBC only partially respond to chemotherapy and present lack of clinically established targeted therapies. Therefore, further validation of optimised workflows for handling adherent mammalian culture which can track the changes between TNBC cell lines is essential. To our knowledge we are not aware of such kind of analysis in adherent mammalian cell metabolomics where emphasis was given to improve the metabolome coverage with a comprehensive approach. Such a comprehensive analysis with the pathway-based metabolome coverage specific to cancer will surely help in generating many hypothesis based investigations in cancer community such as, for setting up a platform for classifying cancer sub-type, in defining the relative contribution of major metabolic pathways in TNBC and illustrating the underlying molecular mechanism involved in TNBC biology and finally, for generating novel therapeutic regimes by identifying novel targets for anticancer therapies.

1.3 Microalgal systems

Algae are one of the most important bio-factories on earth based on their photosynthesis/CO_2 fixation capacity and supposed to be biofuels of the third generation (Wienkoop et al., 2010a). Not only do these organisms fix carbon dioxide, but they also have the potential to be used for the production of inexpensive bulk chemicals because the major inputs into the system (light and CO_2) are essentially free. However, to harness this potential through metabolic engineering, a deeper understanding of photosynthetic metabolism is required (Boyle and Morgan, 2009, Field et al., 1998, May et al., 2008b). Bottlenecks in algal biodiesel production within the cell can be identified by metabolomics approaches in combination with proteomics and transcriptomics, as the identification of differentially expressed metabolites gives clues to rate-limiting processes in cell, which can be backed up by the determination of metabolite flux. Finally, the integration of such a reliable metabolomics data will help us in identifying the as-of-yet missing reactions in the metabolic network which later will allow the fine tuning of algal properties by genetic or metabolite engineering in combination with other system biology approaches.

The metabolic response of microalgae to environmental stimuli or stressors have to date only been carried out in *Chlamydomonas, Dunaliella* and *Nannochloropsis* species to obtain as much information as possible about key metabolites, which will provide first insight into the metabolic pathways employed. For example, the metabolic network of primary metabolism for *C. reinhardtii,* was constructed recently using genomic and biochemical information, which include 484 metabolic reactions and 458 intracellular metabolites (Boyle and Morgan, 2009). However, the metabolome that can be covered with the currently available techniques is only one fifth of the *C. reinhardtii* metabolome and there is considerable scope for improvement in metabolome coverage. The major bottlenecks associated with the sample preparation workflow are efficient sampling, quenching and extraction of metabolites without altering the internal metabolites signatures.

Past literature, demonstrated the potential of metabolomics in understanding the metabolic pathways involved, and their underlying regulation. However, very limited numbers of studies were reported on optimised non-targeted metabolic profiling approach that can potentially detect and quantify hundreds of metabolites, in order to obtain as much information as possible about key metabolites in micro algal species such as *C. reinhardtii, N. salina* and *D. salina*. Optimised quenching and extraction protocols with parallel application of optimized MS-based hyphenated techniques will provide first insight into the metabolic pathways employed, which requires further attention.

C. reinhardtii is a unicellular green alga and was recently sequenced. In the era of global climate change and biofuels it is a primary biological model system for basic studies on photosynthesis and CO_2 –neutral biomass. *N. salina* is well appreciated in aquaculture due to its nutritional value and the ability to produce valuable chemical compounds, such as pigments chlorophyll a, zeaxanthin, astaxanthin and polyunsaturated fatty acids (EPA). Apart from this *D. salina* is also known to have higher lipid content thus can be considered suitable for biodiesel production.

Biodiesel can be produced from the microalgal lipids through trans-esterification process, however overall workflow involves different stages which includes selection of suitable species with higher lipid contents, optimising the parameters for cultivation and harvesting of microalgae, selection of appropriate lipid extraction technique and solvent, trans-esterification of extracted lipids into FAMEs and accurate and reliable quantification of lipids and obtained FAMEs at various stages. For sustainable industrial scale production of biodiesel from microalgae, the impact of above stages on quality of biodiesel yield and on environment should be carefully considered. Moreover, during quantification of fatty acids, it is essential to develop an optimised miniaturised lab based protocol that can save tremendous amount of time, solvents and reagents required, is less expensive and uses environment friendly solvents thus making it more suitable for sustainable large scale production.

1.4 Aims and objectives

1.4.1 Aim

The aim of this project is to capture an accurate and reliable metabolome data by improving the metabolome coverage for all the major metabolite classes by developing an optimised GC-MS and LC-MS based sample preparation workflows that will aid in understanding of underlying molecular mechanisms involved and deeper insight into biology of both the adherently growing mammalian and microalgal systems.

1.4.2 Objectives

1. To assess the extent of metabolite leakage with conventional approaches used for harvesting adherently growing mammalian cells. Furthermore, characterisation of combined quenching and extraction protocols for rapid and reliable metabolome analysis towards improved metabolome coverage and minimal loss of intracellular metabolites during harvesting in three adherently grown mammalian cell lines (two breast cancer cell lines MDA-MB 436, MCF7 and an endothelial cell line HMEC1). (Chapter 3)
2. To examine the influence of various washing and quenching solvents (buffered/non-buffered) in metabolomics of adherently grown mammalian cells: A case study on the metastatic cancer cell line MDA-MB-231. (Chapter 4)
3. Further validation and comparison of proposed protocols for metabolomics of two metastatic TNBC cell lines: MDA-MB-231 and MDA-MB-436. Analysis with respect to pathway specific metabolome coverage. (Chapter 5)
4. To examine GC-MS based optimised sampling, quenching and extraction workflow for obtaining high quality metabolomic data for microalgal strain: *Chlamydomonas reinhardtii.* (Chapter 6, 7 and 8)

5. To optimise and make a comparative evaluation of conventional and direct trans-esterification methods for reliable quantification of FAMEs from micro-algal biofuel strains. Three microalgal species were investigated namely *C. reinhardtii, N. salina* and *D. salina*. (Chapter 9)
6. To evaluate the use of HILIC and IP-RP-HPLC for targeted analysis of amino acids, organic acids, water soluble vitamins, nucleotides, nucleosides and nucleobases towards improving metabolome coverage. (Chapter 10)

Chapter 2

Literature review

2. Literature Review

2.1 Introduction to metabolomics

In the pre-genomic era, the main focus was to seek function first and then to seek the corresponding genes that were involved in providing that function. Whereas, the post-genomic era tends to elucidate the function of all the genes, for many of which there is no corresponding biochemical activity or function known. However, these methods do not provide direct information about how a change in mRNA or protein is coupled to a change in biological function, which led to the evolution of systems biology (Fiehn et al., 2000, Kell and Oliver, 2004, Kitano, 2002). Systems biology is the study of an organism, viewed as an integrated and interacting network of genes, proteins and biochemical reactions. The post-genomic sciences basically include proteomics, transcriptomics and metabolomics (Kussmann et al., 2006). The term metabolomics was first coined in 1998 (Oliver et al., 1998) to describe the "change in the relative concentration of metabolites as the result of deletion or over-expression of gene". At the same time, the term metabolome analysis was coined, referred to the analysis of metabolites in phenotypic profile of *E. coli* (Tweeddale et al., 1998). In brief, the metabolome includes all the small molecular weight intermediates of metabolism (typically of mass <1500 Da). In addition, xenobiotics and their biotransformation products can also constitute the metabolome (Roux et al., 2011). The term metabonomics was coined by (Nicholson et al., 2002, Nicholson et al., 1999) and was defined as, "The quantitative measurement of the dynamic multi-parametric response of a living system to pathophysiological stimuli or genetic modification". The metabonomics studies generate global metabolite profiles that enable diagnostic changes of metabolites in a given biological sample, which help in generating metabolic phenotypes. Further, metabolic phenotype studies are used to obtain new and unexpected insights into biological processes like toxicity, onset and progression of disease as well as ageing and therapy (Roux et al., 2011, Theodoridis et al., 2008). The main difference between metabolomics and metabonomics is that metabonomics is concerned with dynamic changes within a system in response to some sort of stimulus, whereas metabolomics looks at a, "snap-shot of the cellular metabolome" (Tang and Wang, 2006).

Exhaustive work has been done on genomics, proteomics and transcriptomics, which has made significant contribution to the field of functional genomics (Garcia et al., 2008). Now the question arises as to what the metabolomics can add? In a biological system, the metabolome represents the final -omic level and its component can be viewed as an end product of gene expression, which defines the biochemical phenotype of a cell or tissue (Ryan and Robards, 2006). In addition, metabolites are highly connected, and changes in gene expression and protein levels are amplified on the metabolome level. Also, metabolomics can be easily extended to different organisms, across all species, as a given metabolite has the same chemical structure whilst proteins can have species specific isoforms (Wu et al., 2005). Quantitative and qualitative analysis of large number of cellular metabolites represents a broad view of biochemical status of an organism which is useful in monitoring and assessing the gene functions (Sumner et al., 2003). In contrast, proteomics and transcriptomics technologies have some limitations. Any changes in the transcriptome and proteome due to increase in RNA do not always corresponds to alteration in the biochemical phenotype. This may be because the translated protein may require post-translational modifications to be enzymatically active (Bedair and Sumner, 2008a). Furthermore, mRNA and proteins are identified by modern techniques through sequence similarity or database matching. Identification is based on quality of the match and is therefore indirect. In the absence of existing database of information, profiling of the transcripts or proteins yield very limited information (Sumner et al., 2003).

Based on the above limitations, metabolomics is advantageous and complements other -omics technologies in many ways. It also provides the most functional information of the -omics technologies and facilitates metabolic engineering and system biology efforts towards designing superior biocatalysts and cell factories (Bedair and Sumner, 2008b, Garcia et al., 2008, Morris and Watkins, 2005, Ryan and Robards, 2006, Sumner et al., 2003, Wu et al., 2005). The metabolomics field is currently developing rapidly for the study of several biological systems including microbial, plant and mammalian systems. Proteomics studies generate generic protein identification, however metabolomics studies has the ability to reveal that, the accumulated enzyme is more specifically related to a specific biochemical pathway (Bedair and Sumner, 2008a). Furthermore, knowledge of the metabolite level has led to the identification of bottlenecks in the metabolic reaction network. For example, reactions far from equilibrium, such as phosphofructokinase (PFK) and pyruvate kinase (PK), have been confirmed as key regulatory points of glycolytic flux through metabolomics. Metabolite concentration plays an important direct regulatory role through which rapid response to metabolic flux changes via allosteric or feedback inhibition mechanism of enzymes as a mode of pathway flux control is effected (Mashego et al., 2007, Theobald et al., 1993). *Metabolomics* (Springer) journal dedicated to publishing research related results to metabolomics was launched in 2005 and as of today has published 732 articles.

2.1.1 Approaches used in metabolomics

Due to the wide diversity of metabolites in terms of their physicochemical properties, a major challenge in metabolomics is the comprehensive profiling of all the metabolites in a biological system. The chemical nature of metabolites ranges from ionic, inorganic species to hydrophilic carbohydrates, hydrophobic lipids and complex natural products, extending over 7-9 orders of magnitude in concentration (pM-mM) (Monton and Soga, 2007, Ryan and Robards, 2006). However, there is no single analytical platform that is currently available, which is capable of identifying and quantifying the entire metabolome in a given biological system, in an unbiased, reproducible way (Monton and Soga, 2007, Sumner et al., 2003). The approaches used in metabolomics today which are concerned with investigating subsets of the metabolome are listed in table 2.1 (Dettmer et al., 2007, Hall, 2006).

Table 2.1 Approaches used in metabolomics

Term	Definition
Metabolomics	Identification and quantification of all metabolites in a biological system in a given physiological state at a given point of time.
Metabonomics	The quantitative measurement of the dynamic, multi-parametric metabolic response of living system to pathophysiological stimuli or genetic modification.
Targeted analysis	Mainly used for screening purpose – study of primary effects of any alteration. Biochemical profiling can be performed in greater detail on selected groups of metabolites by using optimized extraction and dedicated separation/detection technique.
Metabolite profiling	Identification and quantification of the metabolites present in an organism. For practical reasons this is generally only feasible for a limited number of components, which are generally chosen on the basis of discriminate analysis or on molecular relationships based upon molecular pathways/networks.
Metabolic fingerprinting	High throughput qualitative screening of the metabolic composition of an organism or tissue with the primary aim of sample comparison and discrimination analysis. Generally no attempt is initially made to identify the metabolites present. All steps from sample preparation, separation and detection should be rapid and simple as is feasible. Often used as a forerunner to metabolic profiling. Used for functional genomics, plant breeding and disease diagnosis.
Metabolic footprinting	Analysis of extracellular metabolites in cell medium as a reflection of metabolite excretion or uptake by cells.

Absolute quantification (to make data comparable) would be possible with metabolomics analysis, if standard reference materials are available. However, in most cases standard reference materials for metabolomics analysis are not available. Therefore metabolites are mostly quantified using relative quantification by determining response ratio between the metabolite and an internal standard or other metabolite (Koek et al., 2010).

2.1.2 Classification of metabolites

Metabolites are the end products of cellular regulatory processes and their levels can be regarded as the ultimate response of biological systems to genetic or environmental changes. On the basis of their origin, metabolites are classified as endogenous and exogenous metabolites. Endogenous metabolites are further classified as primary and secondary metabolites. Primary metabolites are those which are directly involved in the essential life processes like growth, development and reproduction e.g. amino acids and glycolysis intermediates. Whereas, secondary metabolites are species-specific and include those which are not directly involved in those processes but usually has important ecological functions, as they are synthesized for a particular biological function e.g. antibiotics, alkaloids and pigments. On the other hand, exogenous metabolites are those that are secreted by an organism or transformed in its immediate vicinity. The concept of xenometabolome was proposed by (Holmes et al., 2008), which is a description of xenobiotic metabolic profile of an individual exposed to environmental pollutants, drugs and phytochemicals.

In metabolomics, the analytical challenge is to capture the chemical diversity of the metabolome. Estimated number of metabolites present only in the plant and human metabolome ranges from 2,766 to well over 200,000 (Duarte et al., 2007, Fiehn, 2002). These metabolites vary in their polarity from non-polar lipids to extremely polar metabolites. Functional groups include alcohols, amines, organic acids, aromatic alcohols, esters, aldehydes, ethers, phosphates and many others. Lipidome and glycome are some of the metabolic subgroups which are so large and important that they have their own classification. Numerous isomers exist in the metabolome, as it comprises a relatively narrow band of molecular weights (typically < 1,500 Da). Compound chirality, chemical complexity of the metabolome and the physiological concentrations of metabolites that range from mM to pM is also a concern and a challenge to analysis (Dwivedi et al., 2008).

It is clear that, there is no single analytical technique that can detect all metabolites present in a metabolome with high enough precision, resolution and sensitivity due to the large diversity and complexity of the metabolome (Van Der Werf et al., 2005). To overcome this problem it is essential to have an array of comprehensive analytical methods, as different analytical techniques are capable of analysing different classes of metabolites. While developing the analytical platforms that can allow the detection of all metabolites present in microbial, algal, plant or mammalian metabolome, it is essential to have access to information about the class and number of metabolites potentially present in the metabolome. Hence, it is essential to classify all the metabolites depending upon their physicochemical properties, which will serve as a base, for setting up the analytical methods for different metabolite classes (Werf et al., 2007). The classes of compounds and the analytical techniques with which they are most compatible are described in figure 2.1.

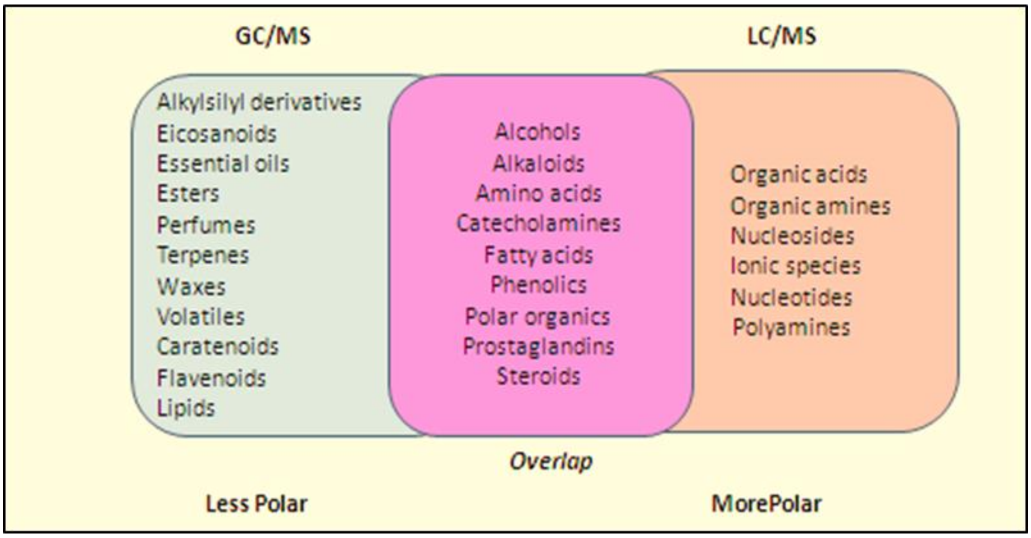

GC/MS
LC/MS
Alkylsilyl derivatives
Eicosanoids
Essential oils
Esters
Perfumes
Terpenes
Waxes
Volatiles
Caratenoids
Flavenoids
Lipids
Alcohols
Alkaloids
Amino acids
Catecholamines
Fatty acids
Phenolics
Polar organics
Prostaglandins
Steroids
Organic acids
Organic amines
Nucleosides
Ionic species
Nucleotides
Polyamines
Overlap
Less Polar
MorePolar

Figure 2.1 Classes of compounds and the analytical techniques with which they are most compatible.

Recently, a validated and comprehensive analytical platform was set up (Coulier et al., 2006) in which both GC-MS and LC-MS methods were developed to quantitatively analyze as many metabolites in microorganisms as possible. GC-MS platform was used to analyze several classes of small polar metabolites, i.e. sugars, amino acids and sugar phosphates. While the LC-MS platform was developed to analyze different classes of important metabolites that could not be analyzed by GC-MS such as nucleotides, coenzyme A esters, and sugar nucleotides. After analysis by different analytical platforms in different biological systems, metabolites were classified into different classes such as amino acid and its derivatives, polyamines, nucleotide pathways intermediates, TCA cycle intermediates, photorespiration intermediates, sugars/sugar alcohols, isoprenoids, organic acids, phosphates etc. Another approach was to classify them as cationic and anionic metabolites (Matthew et al., 2009, Saito et al., 2006, Vijayendran et al., 2008).

To date, different authors have tried to classify metabolites based on their research perspective rather than using a generalised approach. It is therefore essential to have a comprehensive, reproducible and user-friendly system for the classification of metabolites based on their physicochemical properties, which will serve as the basis for selecting analytical methods.

2.2 Metabolomics workflow

The steps involved in comprehensive metabolome analysis in a given biological sample are discussed in this section which includes,

1) Sampling and quenching
2) Extraction of metabolites
3) Metabolite analysis
   a) Direct detection methods such as NMR, FTICR-MS, DIMS and MALDI-MS
   b) Separation based techniques such as GC-MS, LC-MS and CE-MS
4) Data export and analysis

2.2.1 Sampling and quenching

Rapid sampling and instantaneous quenching of enzyme activities are essential because of the high turnover rates of the intracellular metabolites. It is essential to inactivate the enzymes under mild conditions and as fast as possible during cell harvesting in order to retain a valid snapshot of the metabolic processes and to minimize changes to the internal metabolites signature (Lee and Fiehn, 2008b). Thus the ideal quenching solution should instantly arrest cellular metabolic activity without causing any significant cell membrane damage thus inhibiting leakage of intracellular metabolites from the cells (Dietmair et al., 2010).

The two general strategies that are commonly used for quenching are: a) simultaneous sample processing, in which quenching and extraction are combined by rapid sampling of the broth directly into cold quenching solution. However, this method may lead to the risk of overestimating intracellular pools, as it doesn’t involve removal of metabolites present in the extracellular medium b) sequential sample processing involves quenching followed by a separation of biomass from the extracellular medium then extraction of intracellular metabolites (Link et al., 2008, Villas-Bôas et al., 2005b). A comparison of the different strategies is shown in Table 2.2.

Table 2.2 Comparison of the advantages and disadvantages between sequential and simultaneous sample processing for metabolomics studies. Table modified from (Mashego et al., 2007)

	Advantage	Disadvantage
Sequential sample processing		
Cold methanol or liquid nitrogen quenching	Separation of biomass from supernatant	Possible metabolite leakage
	Sample matrix is cleaner	Multiple extraction procedure for specific metabolites
	Separate extraction step (Targeted metabolites)	Laborious
Simultaneous sample processing		
Quenching & direct extraction with either alkali, acid, boiling ethanol or water	Procedure is simpler to perform	Difficult to interpret data
	No separation of biomass from supernatant required	Sample matrix complex
	Total quantification of both intra and extracellular metabolites	High salt content in samples
	Non-specific	Metabolites are too dilute leading to poor detection and quantification

The cold methanol and liquid nitrogen quenching are most widely used. However quenching with the liquid nitrogen does not allow separation between intra- and extracellular metabolites. Quenching with 60% v/v cold methanol at -40 or -50°C was originally pioneered in *Saccharomyces cerevisiae* and has been used widely for microbial metabolomics. The advantages of using methanol include: a) it is fully miscible with water b) it has a very low freezing point (-97.6°C) c) it is less harmful to the cells than other organic solvents and d) methanol-water solutions are not very viscous thus allowing easy centrifugation and cell separation (Canelas et al., 2008). However, quenching with cold methanol in prokaryotic cells, results in leakage of intracellular metabolites caused by cold shock.

Alternative methods to cold methanol quenching has been developed by Rabinowitz and co-workers (Rabinowitz, 2007). A filter culture methodology was used, in which cells were grown on filter on top of agarose media support, which enables quick quenching of metabolism simply by transferring a filter from an agarose plate into cold quenching solution. Bolten and co-workers used the fast filtration method which is only applicable if metabolites with high time constants are quantified (Bolten et al., 2007). Similarly, Kim and co-workers used the fast filtration method with a water washing step (Kim et al., 2013). Although filtration can be quick, there is a risk of the filter getting blocked if the number of cells required for analysis is high (Dietmair et al., 2010). Glycerol-saline quenching fluid was also used (Villas-Bôas and Bruheim, 2007) which showed reduction in leakage of intracellular metabolites compared to methanol/water, however the application of glycerol leads to additional problems as it is difficult to remove from the samples. Similarly, Link and co-workers used ATP as indicator for cell leakage in case of microbial metabolomics and based the analysis on recovery of adenylates both in cell extract and supernatant following quenching (Link et al., 2008). The authors suggested that the use of methanol/glycerol quenching fluid could reduce leakage of ATP as compared to the commonly applied methanol/water solution. However, the effect of glycerol on other metabolites has not yet been assessed, requiring further investigations in future. Moreover, analysis of ATP is just a representative of a small subset of metabolome, and results for other high-turnover metabolites might be different. All suggested alternative methods to cold methanol quenching has its own advantages and disadvantages and more importantly cannot be directly applied/extended to given organism.

In order to limit metabolite leakage when cold methanol quenching is used, Faijes et al have suggested the inclusion of additives, which will act as a buffer or will restrict osmotic shock (Faijes et al., 2007). Later several researchers explored the inclusion of various buffer additives to the methanol quenching method to minimise metabolite leakage and reported contradicting results on the influence of buffers and salt concentration in the quenching solutions (Bolten et al., 2007, Canelas et al., 2008, Kim et al., 2013, Link et al., 2008, Spura et al., 2009). The commonly employed buffer additives at different concentration involves methanol supplemented with either HEPES, AMBIC, tricine or NaCl. Influence of these buffer additives in preserving the membrane integrity and therefore in minimising metabolite leakage is well studied for bacteria (Bolten et al., 2007, Faijes et al., 2007, Link et al., 2008, Marcinowska et al., 2011, Taymaz-Nikerel et al., 2009), fungi (Hajjaj et al., 1998) and yeast (Canelas et al., 2008, Castrillo et al., 2003, Kim et al., 2013, Spura et al., 2009). However, Canelas and co-workers investigated use of additives such as HEPES and AMBIC, and reported less recoveries of intracellular metabolites for buffered sample compared to use of non-buffered methanol in yeast, and suggested no additives are required for metabolome quenching (Canelas et al., 2008). Faijes and co-workers reported higher ATP leakage in *L. plantarum* when using methanol that was supplemented with NaCl, compared to that of 60% aqueous methanol (Faijes et al., 2007). On the other hand, some researchers (Canelas et al., 2008, Taymaz-Nikerel et al., 2009, Tredwell et al., 2011) evaluated addition of tricine to methanol in order to reduce the osmotic shock during quenching, however all reports demonstrated lower recovery of metabolites in terms of number and concentration with use of methanol/tricine compared to that of 60% aqueous methanol alone. Moreover, Tredwell and co-workers reported strong interference of tricine with the derivatization reactions which affected GC-MS analysis and suggested use of methanol supplemented with AMBIC (Tredwell et al., 2011). Contrary to this, Castrillo and co-workers employed ESI-LC-MS/MS technique and reported higher recoveries with tricine in yeast samples compared to use of HEPES and non-buffered methanol (Castrillo et al., 2003). In *E. coli*, Taymaz and co-workers evaluated influence of various additives HEPES, NaCl and tricine and reported no significant improvement in the recovery of intracellular metabolites compared to non-buffered aqueous methanol (Taymaz-Nikerel et al., 2009). Mashego and co-workers suggested use of different methods, where leakage of metabolites should be corrected by measuring the metabolites in both total broth and culture supernatants (Mashego et al., 2007). However quantification of such leaked metabolites is a daunting task and will require highly sensitive analytical method, as the concentration of leaked metabolites in culture supernatants is expected to be below 1 µM (Link et al., 2008). In addition this approach also has drawbacks because of the high salt content in the growth medium is likely to interfere with the MS analysis. To date it is not completely clear, if this leakage is a constant loss or if it is linked to biological parameters such as the state of the cells and length of time spent in the cold. Therefore, there is a need for novel optimized metabolic quenching steps in the experimental procedure in metabolomics for a given biological sample that can overcome all of the challenges in the field of metabolic quenching (Wellerdiek et al., 2009a).

2.2.2 Extraction of metabolites

Due to the diverse physicochemical properties of metabolites, the identification of an optimal extraction solvent to quantitatively extract all intracellular metabolites represents a major challenge in metabolomics. Different procedures and a range of extraction solvents have been used till date which includes acids, bases, organics, alcohols, water and the usage of hot or cold temperatures. However, each extraction condition and solution may favour a set of metabolites and certain cell types (Dietmair et al., 2010).

In order to achieve efficient extraction, it is essential that the sample is homogenised. This can be achieved by various techniques such as liquid N_2 grinding with a pestle and mortar (Erban et al., 2007), milling with vibrational mill with chilled holder (Weckwerth et al., 2004a), sonification and freeze-thaw cycle with the use of liquid N_2 and dry ice (Roessner et al., 2001). Various techniques have been used in order to increase the efficiency of extraction and these include: supercritical fluid extraction (Chen et al., 2008, Huie, 2002), microwave assisted extraction (Brachet et al., 2002, Kaufmann and Christen, 2002, Proestos and Komaitis, 2008), and pressurized liquid extraction (Barbero et al., 2006, Benthin et al., 1999, Ong, 2002).

Extraction with boiling ethanol is capable of extracting a wide set of metabolites from glycolysis, TCA cycle and pentose phosphate pathway. However Mashego et al suggested that, it is still essential to classify the extraction protocols, in order to target classes of like metabolites (Mashego et al., 2007). For example, polar solvents for the extraction of polar metabolites, non-polar solvents for extraction of non-polar metabolites, acids for extraction of acid stable metabolites, alkali’s for extraction alkali stable metabolites, low temperature for thermo labile metabolites high temperature for thermo stable metabolites.

In the case of microbial metabolomics**,** it has been (Prasad Maharjan and Ferenci, 2003a) suggested that cold methanol/water extracted more metabolites from *E. coli,* whereas it was suggested later (Rabinowitz, 2007) that extraction with this solvent mixture causes marked decomposition of nucleotide triphosphates, Additionally, they used the extraction with acidic acetonitrile for microbial samples which minimized loss of high energy metabolites and their conversion into low-energy derivatives. Hiller et al demonstrated that extraction with buffered hot water (30 mM triethanolamine, pH 7.5, 95°C) results in more reliable metabolite extraction in comparison to buffered ethanol, unbuffered hot water or perchloric acid for *E. coli* metabolomics (Hiller et al., 2007a). Yanes et al explored the influence of solvent polarity, temperature and pH, while extracting both polar and non-polar metabolites, and concluded that hot polar solvents are most efficient for this purpose (Yanes et al., 2011). Recently, Duportet et al compared different chemical extraction methods coupled to sonication for different microbial cell types including gram positive bacterium, gram negative bacterium, yeast and filamentous fungus, and concluded that different chemical extraction methods yield significantly different metabolic profiles (mainly due to polarity of solvents used), therefore it is essential to identify chemical extraction method using one or more organic solvents as an extraction agent for a given biological sample (Duportet et al., 2011). Canelas et al used an approach of addition of ^{13}C-labeled internal standards at different stages of the sample treatment process, which is useful in determination of metabolite recoveries (Canelas et al., 2008).

In the case of mammalian cells optimization of extraction protocols was performed by fractional factorial design, based on the evaluation of different conditions on different response variables (Ritter et al., 2008). Recently, Dietmair et al compared 12 different extraction methods with respect to mammalian cells and concluded that cold extraction using 50% acetonitrile was the best method (Dietmair et al., 2010).

In the case of micro-algal metabolomics, the combination methanol:chloroform:water not only yields superior recoveries of amino acids, organic acids, and sugar alcohol but also adequately recovers phosphorylated compounds in comparison with methanol extraction methods (Jensen et al., 1999, Smits et al., 1998, Villas-Bôas et al., 2005a). Lee and Fiehn tested five different extraction solvent system with respect to *C. reinhardtii* which include methanol:chloroform:water (5:2:2) as published for plant organs (Gullberg et al., 2004a, Weckwerth et al., 2004a), methanol: isopropanol: water (5:2:2), 100% methanol (Villas-Bôas et al., 2005a), acetonitrile: isopropanol: water (5:2:2) as published for blood plasma (Fiehn and Kind, 2007), and methanol: chloroform: water (10:3:1) which previously found to be most suitable for *C. reinhardtii* (Bölling and Fiehn, 2005). In addition to this two alternative methods for cell disruption were also tested for efficiency of extraction which include use of mixer mill with glass beads and use of single 5 mm i.d. steel ball widely used in plant tissues. Authors concluded use of methanol:chloroform:water (5:2:2) coupled to steel ball milling for cell homogenization was the most efficient extraction method. Recently, Shin et al carried out evaluation of extraction methods for global metabolic profiling of gram negative marine bacterium *Saccharophagus degradans* which is capable of degrading wide range of polysaccharides, which enables its application to the saccharification of lignocellulose and marine microalgae for biofuel production (Shin et al., 2010). The evaluation criteria used was, the number of peaks detected for each extraction solvent, the peak intensity of structurally identified compound, and the reproducibility of metabolite quantification. Authors demonstrated that acetonitrile:methanol:water (2:2:1) and water: isopropanol: methanol (2:2:5) were found to be superior extraction solvents.

Hence for the unbiased analysis of metabolites it is essential that all metabolites need to be completely, non-selectively and reproducibly extracted by avoiding their degradation (e.g., chemical, thermal) and conversion to other metabolites (e.g., due to enzymatic activity). For global metabolite profiling, the ideal extraction solvent should cover a broad range of chemical properties of metabolites, and should be capable of extraction of all metabolites in high yield regardless of their chemical class and in addition the resulting sample matrix should be compatible or amenable to the analytical method of choice (Dietmair et al., 2010, Mashego et al., 2007, Prasad Maharjan and Ferenci, 2003a, Shin et al., 2010, Villas-Bôas et al., 2005a). However it is impossible to generate such an extraction solvent, hence it is essential to acquire an optimal solvent system to develop a method with high extraction efficiency and reproducibility for a given biological sample (Canelas et al., 2008, Gullberg et al., 2004a, Shin et al., 2010). Completeness of extraction cannot be determined theoretically, as no one knows initially the number of metabolites present in the cells, hence determining the extent of metabolite degradation and efficacy of method should be tested to validate the optimal method.

Efficacy can be tested by comparing the different methods for the identical biological sample, whereas extent of metabolite degradation and the absence of enzyme activity can be tested by metabolite recoveries, by means of spiking or standard additions, which will give information on metabolite degradation (metabolite-specific recoveries much below 100%) and inter-conversion (metabolite-mixed recoveries above and below 100%).

However several authors reported that it is essential to identify the optimal extraction condition with respect to the biological system considered for research study (Hiller et al., 2007a, Mashego et al., 2007, Oldiges et al., 2004). Different groups of metabolites have different extraction efficiencies, hence method optimization should be based not only on a few compounds, as seen in the previous studies, but rather on a broad range of metabolites (Dietmair et al., 2010). It is also essential to validate the performance of different extraction techniques available by utilizing methods described above with respect to microalgal species which has not been seen in the past literature.

2.2.3. Metabolite analysis

Historically, spectrophotometric assays capable of detecting single metabolites or by simple chromatographic separation of mixtures of low complexity were used for the measurement of metabolites. However, over the past decade, several analytical methods have been developed offering both high sensitivity and accuracy for metabolome analysis (Lisec et al., 2006).

2.2.3.1. Direct detection methods

2.2.3.1.1 NMR

In the case of more concentrated metabolites, NMR has been widely used in the past and has significantly contributed to the metabolite profiling studies (Gowda et al., 2008, Lindon et al., 2004, Serkova and Niemann, 2006). NMR offers many advantages including high information content of resulting spectra, non-destructiveness, the relative stability of the resulting chemical shifts, the ease of quantification of multiple classes of metabolites and the lack of any need to pre-select the conditions employed for the analysis. One dimensional NMR laid the foundation of metabolomics studies, which offers an array of detection schemes that can be tailored to the nature of the sample and the metabolic problem that needs to be addressed. However complexity of the NMR spectra limits the identification of metabolites. Several attempts were made to improve peak assignments with the help of a metabolite database. However, it is still inadequate for a complete and unambiguous assignment, particularly for metabolites at lower concentrations. In addition, NMR has a low sensitivity compared to MS and suffers from overlapping signals, leading to smaller number of absolute identifications. The hyphenation of NMR with HPLC can increase its efficiency by reducing the co-resonant peaks and improving the dynamic range, but their use still remains limited by low throughput and difficulties associated with metabolite identification (Bacher et al., 1998, Martin et al., 2007, Scalbert et al., 2009).

2.2.3.1.2 FTICR MS

FTICR has great potential to unravel metabolomes. This technique is capable of identifying many metabolite classes simultaneously without the need for chromatographic separations. It has been described for metabolome analysis of plant extracts and characterization of lipo-oligosaccharides (Leavell et al., 2002).

2.2.3.1.3 DIMS

DIMS provide the high-throughput metabolic fingerprinting and foot printing analyses in which sample are directly injected in a flow injection mode into mass spectrometers without any prior separation (Vaidyanathan et al., 2002). Direct infusion is often employed with ESI-MS analysis for rapid characterization of organisms, cells or tissues for foot printing purposes (Vaidyanathan et al., 2001). This techniques has also been described for profiling of secondary metabolites in plants and fungi (Villas-Bôas et al., 2005b).

2.2.3.1.4 MALDI-MS

MALDI-MS has not been widely used for metabolomics, due to interference from matrix ions in the low molecular weight range. However, introduction of 9-aminoacridine as a matrix has been shown successful application of this technique in quantitative detection of metabolites in negative ion mode (Edwards and Kennedy, 2005). Later to assess the performance of this method, a metabolite cocktail was spiked with different concentrations of single metabolites, which lead to conclusion that changes in the concentration of one analyte is often influenced the quantification of other metabolites (Vaidyanathan and Goodacre, 2007).

However, FTICR-MS and DIMS are preferred techniques for metabolic fingerprinting, as they require minimal sample preparation and provide high throughput analysis. However the spectra obtained is very complex and elucidation of such spectra is very complicated task (Koek et al., 2010).

2.2.3.2. Separation based techniques

2.2.3.2.1 Separation by chromatographic techniques

In the case of metabolites having lower concentration, sequential approaches of concentration, separation and detection methods are commonly employed. Separation can be achieved by the use of chromatographic or electrophoresis methods, such as GC, LC or CE. Separation by GC can be applied for those metabolites which have high vapour pressures or can be derivatized to form volatile products. A major disadvantage of GC for metabolic profiling is that, non-volatile compounds cannot be determined and requires time consuming derivatization procedures to render them volatile. LC can be used for the separation of non-volatile metabolites, which makes up most of the metabolome. RP-HPLC has been widely applied for separation of non-volatile metabolites at lower concentration and varying polarities, however as compared to GC separation; this method is time consuming, has low resolution and does not adequately address the hydrophilic component of the metabolome (Asbury and Hill, 2000). Recently, UPLC has been introduced to overcome these limitations, with small diameter (≤2 µm) packing's and the novel silicoethyl bonded phase, thus linear mobile phase velocities of up to 10 mm/s can be achieved without sacrifice in resolution . Apart from these chromatographic techniques, CE can also be used for the efficient separation of cations and anions and/or hydrophilic metabolites in the metabolome, but it is too slow and can take up to 16 hours for some separations. Furthermore, the buffers used are not readily compatible with the ion source of MS.

After chromatographic or electrophoretic separation of metabolites, identification of metabolites is achieved by MS (Dwivedi et al., 2008). MS is generally employed in combination with pre-fractionation methods, such as GC, LC and CE to reduce the complexity of target sample before MS or tandem MS analysis.

2.2.3.2.2 Detection by Mass spectrometry (MS)

Mass spectrometry is a well-established technique for the analysis of diverse chemicals and bio-molecules on the basis of their mass to charge ratio (*m/z*). Over the last decade, MS has secured a pinnacle position and holds additional promise for the advancement of metabolomics based upon sensitivity, selectivity, relative cost and depth of coverage. The MS based hyphenated techniques combine chromatographic techniques for separation of metabolites based on their physicochemical properties coupled to mass detection with mass spectrometry.

MS instrumentation

Typical mass spectrometer consists of an ionization source, mass analyzer and detector as shown in figure 2.2;

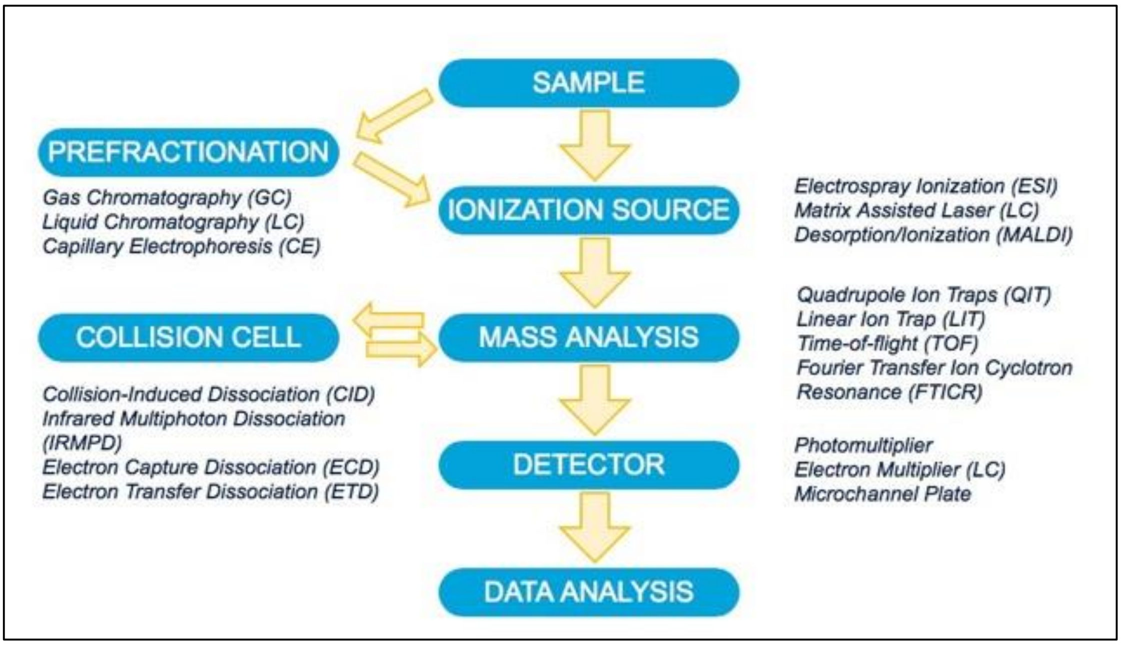

SAMPLE
PREFRACTIONATION
Gas Chromatography (GC)
Liquid Chromatography (LC)
Capillary Electrophoresis (CE)
IONIZATION SOURCE
Electrospray Ionization (ESI)
Matrix Assisted Laser (LC)
Desorption/Ionization (MALDI)
COLLISION CELL
MASS ANALYSIS
Quadrupole Ion Traps (QIT)
Linear Ion Trap (LIT)
Time-of-flight (TOF)
Fourier Transfer Ion Cyclotron
Resonance (FTICR)
Collision-Induced Dissociation (CID)
Infrared Multiphoton Dissociation
(IRMPD)
Electron Capture Dissociation (ECD)
Electron Transfer Dissociation (ETD)
DETECTOR
Photomultiplier
Electron Multiplier (LC)
Microchannel Plate
DATA ANALYSIS

Figure 2.2 General flow chart for mass spectrometry (figure modified from Feng et al., 2008).

Mass spectrometers operate in a four step process. Following the sample introduction, ions are generated in an ion source, at vacuum pressures, in order to create charged species, a requirement for the separation of ions according to mass-to-charge *(m/z)* ratio. Subsequently separation, in space or time, of ions based on their *m/z* ratio occurs in a mass analyzer, followed by detection of ions either physically at a detector as an ion current or by the detection of orbital frequencies as image currents.

Ionisation sources

In MS, there are two types of ionization techniques, hard ionization techniques which result in a large amount of fragmentation of the analyte and soft ionization techniques which result in little ion source fragmentation and molecular ion peaks. Ionization methods can be further classified on the basis of the source of the ions, either in the gas-phase, by desorption or in a spray. In gas-phase sources, the sample is directly converted to gaseous ions and this mode is applicable to compounds with molecular masses of less than 10^3 Da. The electron impact ionization and chemical ionization methods employ gas-phase sources and can be easily coupled with GC, but not with LC. In desorption methods, the sample in either gas or liquid state is converted to gaseous ions and is applicable to analysis of much higher masses e.g. MALDI. Lastly, the spray sources involves ionization of an aerosolized spray, which are commonly employed in techniques such as atmospheric pressure chemical ionization (APCI) and electrospray ionization (ESI).

Analyzers

In time of flight (TOF) mass analyzer, all ions having the same charge are imparted with the same kinetic energy by acceleration through an electrical field and are then allowed to drift freely through a vacuum tube. As all ions have the same kinetic energy, the lighter particles will transverse the flight tube faster than heavier particles, thus allowing separation of ions based on *m/z.* The TOF mass analyzers have several advantages over quadrupole and 3D ion traps because they have a wide mass coverage. It provides greater sensitivity by detecting all ions simultaneously rather than scanning mass ranges as is the case with many quadrupole instruments, thus prior knowledge of the metabolites to be detected is not required, as would be required for Q-TOF and triple quadrupole instruments in tandem MS. It has high acquisition rates greater than 100 Hz, thus it is suitable for sampling of data points across chromatographic peaks. It also provides accurate mass measurement of the molecular ion, with typical mass accuracies of less than 5 ppm. The hybrid Q-TOF instrument provides two routes to identify metabolites. First route involves accurate mass measurements of the molecular ion while the other provides accurate mass measurements of the precursor (molecular) and product ions using tandem mass spectrometry (MS/MS). FTICR mass spectrometers are well known to provide higher mass accuracy and mass resolution, but they are very expensive. However recent developments in Q-TOF instruments serves as a cheaper alternative to that of FTICR MS and are capable of providing higher mass accuracy, sensitivity and dynamic range (Dunn, 2008).

Detectors

The most commonly used detectors in MS are the electron multiplier and Faraday cup. Faraday cups are highly accurate, as the current produced is directly proportional to the number of ions, whereas it has less sensitivity as compared to electron multiplier because no amplification of signal takes place in Faraday cups (Dunn, 2008).

In conclusion, coupling of high-resolution separations to MS can also substantially increase the depth-of-metabolome coverage, add an additional dimension for metabolite identification, and enhances the biological context through more rigorous identification of a greater number of metabolites (Bedair and Sumner, 2008a). Hence nowadays, hyphenation of GC, LC and CE based separation techniques to MS are preferred over NMR and other analytical techniques, as they are more robust and has the potential to be applied for the analysis of more than one metabolite class (Dettmer et al., 2007, Scalbert et al., 2009, Van Der Werf et al., 2005).

2.2.4. Data export and analysis

The resulting data burden arising from the complexity and richness of the metabolome is regarded as one of the major bottlenecks. The raw instrumental data requires processing in order to extract information for further analysis and interpretation. A standard extensible markup language (XML) format for metabolomics is, a yet to be developed. Such a development can provide not only standardization of quantitative data output, but also a thorough sample annotation, which is required for data interpretation (Dettmer et al., 2007). All analytical instruments generally have software for quantitative data processing but unfortunately not all are capable of processing very rich comprehensive metabolomics data generated by GC-MS and/or LC-MS. Nowadays most of the instrument companies such as Agilent, Bruker and Thermo have become active in producing automated data extraction tools for metabolomics, however these tools work only for specific type of data and data formats (Scalbert et al., 2009). Hence, there is need for improvements in the area of data extraction.

2.2.4.1 Data processing

For data processing, essentially three types of methods are available: target analysis, peak picking and deconvolution. The main challenges in data processing are: the amount of data, unbiased data processing, alignment of peaks shifted along the retention time axis and lastly obtaining only one entry for each metabolite (Koek et al., 2010). The GC-MS and/or LC-MS experiments can generate 2 general types of data or mass spectral tags (MSTs): 1) parent mass + chromatographic retention time or 2) parent mass + fragment mass + chromatographic retention time. The identification of both known and unknown compounds is possible if these properties are properly documented. The processing of obtained raw sets of chromatogram involves noise reduction, spectrum deconvolution, peak detection and integration, chromatogram alignment, compound identification and quantification (Lu et al., 2011).

The construction of mass spectral libraries and metabolite databases is essential so as to define the metabolome more comprehensively. The most commonly used databases for compound identification are NIST 08, GMD, MassBank, METLIN, MMCD, and HMDB (Go, 2010). However these libraries are not specific from the metabolomics perspective. In GC-MS, a number of research groups are developing their own metabolite mass spectral/retention index libraries, an example currently available is the GOLM metabolome database. The variation in the retention times across multiple samples can be normalized by addition of retention indices like *n*-alkanes to the sample which increases the reproducibility of chromatographic alignment. A match factor greater than 70% is generally employed for a mass spectral match. Recently, NIST libraries have been carefully curated and available as NIST 14 with 242,477 spectra and GC retention indices for 82,337 compounds. However there are no mass spectral/retention index libraries available for 2D GC-MS (Dunn and Ellis, 2005). In LC-MS applications, these libraries are much less developed. Recent development of the MZmine and MZmine 2 provides important platform independent software for the processing of the LC-MS data in metabolomics and proteomics (Lu et al., 2008a). Similarly, a web based platform known as XCMS online has been recently developed for analysis of LC-MS based metabolomics data which offers feature detection, RT correction, alignment, annotation, statistical analysis and data visualization followed by putative identification of metabolites against the METLIN database (Patti et al., 2012, Smith et al., 2006, Tautenhahn et al., 2007, Tautenhahn et al., 2008, Tautenhahn et al., 2012).

Deconvolution makes the use of the differences in mass-spectra information between different metabolites to separate overlapping peaks. One of the tools, performing spectra deconvolution and application of retention indices is the AMDIS, which is freely available from NIST. It is most promising method for GC-MS, as it can handle huge datasets, automated processing, automatic peak alignment and provides just one quantitative value per metabolite per sample. However, AMDIS is not compatible with LC-MS or CE-MS, but ESI-LC-MS data can be processed using component detection algorithm. Recently freely available software tool have been developed by (Behrends et al., 2011) for backfilling missing values obtained from AMDIS-processed GC-MS spectra, producing a data matrix more suitable for subsequent chemometric analysis.

Further development of LC-MS and GC-MS mass spectral libraries with respect to mammalian and microalgal samples, and their applicability and availability across many research laboratories is essential, but not yet achieved.

2.2.4.2 Data interpretation and visualisation

Once raw data has been converted to a quantitative description, one can, in principle apply statistical tools which are commonly employed for normalization, alignment, corrections, noise removal and transformation of large data sets. The selection of statistical tool, in metabolomics, highly depends on aim of the study. For sample classification; unsupervised methods such as PCA, HCA or ICA are used, while for biomarker discovery, supervised methods, such as PLS, ANOVA or SIMCA can be used. Data normalization and data transformations are used to minimize the impact of variability of high intensity peaks. Each sample is usually normalized to the sum of its intensities, while transformation is intended to reduce high intensity values but keep low intensity values (Dettmer et al., 2007). The widely used pattern recognition tools are

PCA and HCA, which enable the ease of comparison and visualization of similarities and differences between data sets. PCA is a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components, which are linear combinations of the original variables (Hotelling, 1933, Lu et al., 2011). To gain further insight into data characteristics such as *p*-value, fold change, retention time, *m/z* ratio and signal intensity of all detected features, cloud plot visualization tool have been recently developed by the XCMS online platform (Patti et al., 2012).

2.2.4.3 Metabolite identification

In order to extract the biological context from the data, it is essential to identify the metabolites. Metabolites should be identified or classified as either, unknown, belonging to a certain chemical class, putatively identified by a match to a database MST or confirmed with an authentic standard. Recently the Metabolomics Standard Initiative (MSI), proposed a guideline for the identification of non-novel metabolites, in which a minimum of two independent and orthogonal data relative to an authentic standard compound analyzed under identical experimental conditions are proposed as necessary for metabolite identification (Fiehn et al., 2007, Sansone et al., 2007). Metabolites can also be identified by a library search. The largest spectra libraries are available from Wiley and NIST 14.

Recently the first draft of minimum standard reporting requirements for biological samples in metabolomics experiment was proposed and this has allowed the scientific community to evaluate, understand, repeat, compare and re-investigate metabolomics studies (Fiehn et al., 2006, Sumner et al., 2007, van der Werf et al., 2007). However this approach has not been widely used in mammalian and microalgal metabolomics and requires further attention.

2.3 GC-MS and LC-MS based metabolomics

2.3.1 GC-MS based metabolomics

GC-MS is one of the most widely used analytical techniques in metabolomics. It combines the high separation efficiency and resolution of capillary GC that is essential for complex metabolic profiling with the high sensitivity of mass selective detection. A wide range of volatile and/or derivatized non-volatile metabolites can be analyzed qualitatively and quantitatively with high analytical reproducibility and at lower costs compared to LC-MS and CE-MS. Coupling of GC to quadrupole MS, provides high sensitivity and a wide dynamic range but it has nominal mass accuracy and slow scan speed. These limitations can be overcome by GC-TOF-MS, which is widely used in metabolic profiling as it offers higher mass accuracy and mass resolution. Furthermore, it offers high scan speed essential for adequate sampling of high-resolution chromatographic peak widths in the range of 0.5 to 1 s which also facilitates the implementation of fast GC methods, thus reducing the analysis time and increases the productivity. However a major disadvantage associated with GC-MS is the requirement of analyte volatility and thermal stability, which incorporates derivatization steps in analysis, which are time consuming. Derivatization introduces an additional source of variance to the experimental procedures and increases complexity of data, as a single metabolite can produce multiple derivatized metabolite peaks (Dunn, 2008).

The derivatized samples are then introduced into the GC capillary column which provides high resolution and peak capacity, through split or split less injection mode. GC columns are silica capillaries that have lengths between 10-60 m with internal diameter between 100-500 µm, coated externally with an imide layer to reduce column fragility and internally with a liquid stationary phase of siloxane of thickness 10-50 µm. In metabolomics DB5 or DB17 columns are generally employed which contains a 95/5 mix of methyl/phenyl groups and a 50/50 mix of methyl/phenyl groups respectively. Separation can be achieved by non-polar stationary phase that separates analytes according to their boiling point and polar stationary phase that separates analytes based on their polarity (Bedair and Sumner, 2008a, Dunn, 2008). However, for complex biological samples, peak capacities, resolving power and depth of metabolome coverage can be further increased by the use of 2-D GC (GC x GC), that utilizes two columns having different stationary phase selectivity's and are connected serially. Small portions of the effluent from a first-dimension column, usually non-polar, are continuously trapped and released via a modulator onto the second column, usually more polar, for further separation. Therefore, two metabolites of similar volatility but different polarity can be separated. This mode of separation provides higher sensitivity and wider dynamic range as compared to 1-D GC. In order to acquire sufficient data points across the sharp narrow peaks, 2-D GC, is often coupled with TOF-MS. However, the data generated by GC x GC-TOF-MS is large and complex requiring the use of chemometric methods in order to gather the information from these profiles, using multivariate tools like PCA, HCA, PLS-DA, Fisher ratio and PARAFAC. PARAFAC provides a fully deconvoluted mass spectrum as an aid to convert raw data to biological significance (Dunn, 2008, Zhang et al., 2008).

2.3.1.1 Sensitivity of GC-MS

Intracellular metabolites are present at low concentrations and therefore the detector has to be highly sensitive. Analytical sensitivity is dependent on choice of the mass analyzer, the data acquisition mode (full scan/SIM), and the chromatographic setup. In SIM mode, TOF-MS and ion traps provide higher sensitivity, however the dynamic range is lower than for quadrupoles operated in SIM mode. The full scanning mode requires sufficient number of data points, in order to describe the chromatographic elution profile adequately. SIM mode is more sensitive for quantitative analysis of isotopes; however it is not the method of choice for metabolite profiling. The sensitivity also depends on the injection techniques used. Split-less injection is most commonly used for quantitative analysis of low concentration [0.1-50 ppm (ng/µL)] and very dilutes samples, whereas split injection methods are non-quantitative and generally employed for target analysis (Chen et al., 1998, Kopka et al., 1995, Villas-Bôas et al., 2005b).

2.3.1.2 Reproducibility of GC-MS

The major sources of variability in GC-MS analysis are caused by the analytical methodology (chromatography, detection and stability of derivatized samples); and the sample preparation (sampling, quenching, extraction, and concentration). Reproducibility of retention indices between GC columns from different manufacturers allows transfer of libraries between GC columns, however production of reproducible fragmentation mass spectra on different instruments is difficult and is a major limitation to the development of these libraries (Dunn, 2008). The application of CID is more variable across different instruments when compared to the higher reproducibility of fragmentation patterns with EI ionization acquired on different instruments. Hence use of multiple collision energies are recommended for the construction of MS/MS libraries, as the precursor ion of one metabolite may extensively fragment at a given collision energy whereas a separate metabolite may not produce significant fragmentation of the molecular ion at the same collision energy. Large standard deviations may be observed if the derivatized samples are not sufficiently stable (within 24 hrs). Use of cold tray at the GC auto sampler could reduce the variations caused by loss of sample due to evaporation. The sample to sample variation can be minimized by use of internal standards, by adding it to sample as early as possible in the process. Typical internal standards used are synthetic compounds such as N-methylanilline and 5α-cholestane, which should not exhibit significantly higher or lower boiling points than the analytes, and should not interfere with the metabolites present in the sample (Villas-Bôas et al., 2005b).

2.3.1.3 Derivatization

The primary requirement for GC-MS analysis is that analyte must be volatile and thermally stable. Short-chain alcohols, esters, and low molecular weight hydrocarbons and lipids meet these requirements; however majority of the metabolites which includes, amino acids, organic acids and sugars are highly polar and non-volatile, thus requires chemical derivatization at the functional group in order to reduce their polarity, render them volatile and increase their thermal stability prior to GC-MS analysis (Dettmer et al., 2007, Wittmann, 2007). To date, the commonly used derivatization methods for GC-MS analysis of metabolites involves silylation, alkylation or acylation reactions which are simple and lead to high yield with efficient conversion of non-volatile analyte to volatile derivative. In alkylation, silylation and acylation reactions the functional groups such as –COOH, -OH, -NH and –SH can be derivatized (Wittmann, 2007).
Silylation reactions involve replacement of the active hydrogen by an alkylsilyl group [trimethylsilyl (TMS)] with formation of TMS ethers, TMS esters, TMS sulfides or TMS amines. The formed silyl derivatives shows better thermal stability, higher volatility and can produce more distinct MS spectra as compared to their underivatized precursors. In metabolomics, the predominantly used reagents for trimethylsilylation are BSTFA (N, O-bis-(trimethylsilyl)-trifluoroacetamide) and MSTFA (N-methyl-tri-methylsilyltrifluoroacetamide). The problem with BSTFA is, the leaving group is trifluoroacetamide, which is responsible for causing interference with early eluting peaks, whereas, comparable in silyl donor strength, MSTFA is the most volatile of the TMS acetamides and therefore preferred over BSTFA for GC-MS analysis. In certain cases, 1% trimethylchlorosilane is also added as a catalyst. The disadvantage with silylation is its sensitivity to water, which requires that reaction must be carried out under anhydrous conditions and that derivatized samples must be stored in a dry environment, in order to prevent degradation. In addition to this, silylation reactions are also responsible for conversion reactions, as arginine is being converted into ornithine by reaction with MSTFA or BSTFA (Bedair and Sumner, 2008a, Dettmer et al., 2007). Alternatively *t*-butyldimethylsilylation (TBS) derivatives, can be used in addition to trimethylsilylation, which are less sensitive to moisture than TMS derivatives, however TBS derivatives have higher molecular mass than TMS derivatives, which often leads to higher degree of partial derivatization in compounds with multiple functional groups due to steric hindrance. Alkylation and esterification have been used but they can derivatize a narrower range of metabolites than silylation (Bedair and Sumner, 2008a). As a result silylation is preferred in metabolomics over other available alternatives.

In addition to this, hydroxylamine or alkoxyamines are also used to transform carbonyl groups to corresponding oximes in order to stabilize the reducing sugars in the open-chain conformation and also to prevent the decarboxylation of α-ketoacids. Depending upon the orientation at the carbon-nitrogen double bond, oximes derivatives can be formed as *syn* and *anti*-isomers that can be chromatographically resolved, leading to two GC peaks for each compound (Bedair and Sumner, 2008a, Dettmer et al., 2007). For identification purposes, alkoxyamines are preferred over hydroxylamines (Fiehn et al., 2000).

The two-step derivatization procedure is most commonly utilized for GC-MS metabolite profiling, in order to obtain broad coverage of the metabolome, which include methoximation followed by silylation. So in the first step alkoxyamines convert carbonyl groups to oximes as described above, while in the second step, the active hydrogen in polar functional groups is replaced with trimethylsilyl group by silylation with MSTFA.

In the past, attempts have been made to study the influence of oximation temperature and time followed by silylation with MSTFA (with 1% trimethylchlorosilane) and variation of silylation temperature, amount of reagent, and addition of co-solvent such as hexane, acetonitrile. The derivatization efficiency study with respect to plant metabolomics was carried out by (Gullberg et al., 2004a) by using reference compounds and concluded that methoximation with methoxyamine in pyridine for 17 hours at room temperature, followed by silylation with MSTFA for 1 hour at room temperature to be the optimum derivatization procedure.

More recently, (Koek et al., 2006) conducted an experiment with respect to microbial metabolomics, in order to calculate the derivatization efficiency, by using a set of *n*-alkanes as reference compounds. Several parameters were tested including the choice of derivatization solvents (i.e., acetonitrile, dimethylforamide, dimethyl sulfoxide, pyridine, tetrahydrofuran), oximation reagents (hydroxylamine, ethoxyamine), silylation reagents (BSA, MSTFA, BSTFA and a mixture of trimethylsilylimidazole/BSA/trimethylchlorosilane 3:3:2 v/v), derivatization times (15-90 min), and temperature (30-70°C).

Moreover, in view of analytical performance of different metabolites which is mainly governed by the stability of the silylation product after derivatization, Koek and co-workers, classified metabolites based on their derivatization efficiencies, where authors demonstrated higher derivatization efficiencies for metabolites containing hydroxylic and carboxylic functional groups such as organic acids and sugars. Results for these metabolites showed higher derivatization efficiencies (60 – 115%), analysis of repeatability (RSDs < 5% in cell extracts) and reproducibility (RSDs 8 – 14%). Metabolites containing amine or phosphoric functional groups showed satisfactory derivatization efficiencies (30 – 110%), repeatability (1 – 7%) and reproducibility (10%), whereas metabolites with amide, thiol, or sulfonic functional groups were found to be difficult to analyse with the silylation reagents (Koek et al., 2006). Similar conclusions on these metabolite classes with the use of silylation reagents were reported by Gullberg and co-workers

(Gullberg et al., 2004a) as that of Koek and co-workers (Koek et al., 2006) with respect to derivatization efficiencies and application range.

Recently, the source of bias in GC-MS based metabolomics was discussed briefly by (Mishur and Rea, 2011). The authors suggested that bias in metabolomics occurs in two forms. Type A bias which are universal and affect all the metabolites equally and can be corrected by the addition of an internal standard, whereas type B bias, are those which affect individual metabolites differently. They concluded that the primary source of bias in GC-MS is the sample derivatization step, which introduces both Type A and to a greater extent Type B bias. Therefore it is essential to ensure that enough derivatizing agent and optimum experimental conditions exist for the efficient derivatization of hundreds to thousands of organic compounds with a variety of functional groups all competing for the same reagent. In addition the correct duration and temperature also needs to be identified in order to ensure complete derivatization step, as incomplete derivatization of compound with multiple functional groups may result in eluting multiple peaks for the same metabolite.

Few derivatization efficiency studies have been conducted with respect to plant and microbial samples, and methods have been optimized with respect to them, however there is no evidence of such studies in the literature with respect to microalgal metabolomics.

2.3.2 LC-MS based metabolomics

The LC-MS platform in metabolomics offers several advantages over GC-MS, such as operation at lower analysis temperature, thus enabling analysis of thermo labile metabolites which often degraded during GC analysis. More importantly LC-MS analysis does not require chemical derivatization, thus simplifying the sample preparation steps and identification of the metabolites. Ionization in the positive ion mode (amines and amino acids) and negative ion mode (sugars and lipids) leads to the detection of two sets of analytes that can differ significantly. Detection in both the positive and negative ion mode simultaneously is possible with LC-MS, thus reducing the time required for analysis and reduce bias due to injection errors. However, the major disadvantage of LC-MS is ion suppression, which can be overcome to some extent by miniaturization of ESI to nanospray ionization and by better separation of the metabolites. Other major drawbacks with the use of LC-MS metabolomics is contamination of the MS source, and adduct formation. Both have significant consequences for the robustness of the method and result in the lack of transferable LC-MS libraries for metabolite identification (Coulier et al., 2006).

LC operates by the passage of a liquid mobile phase through a stainless steel column packed with the stationary phase, silica or related small diameter particles to which the stationary phase is chemically bonded. The liquid mobile phase is pumped through the high resistance system, a mixing system for combining multiple mobile phases, the column and the detector. The most commonly used columns in LC separation are reverse-phase columns, such as C18 or C8, for efficient separation of non-polar/hydrophobic metabolites (such as aromatics, lipids, phospholipids and bile acids). However retention of polar metabolites on reverse-phase columns is not sufficient and/or often co-elute close to the column void volume, thus making their detection difficult with MS. This problem can be overcome by use of HILIC, which uses water miscible polar organic solvents, such as acetonitrile and water and offer a complementary metabolic profile providing greater retention of hydrophilic metabolites. In HILIC a stagnant water layer is established within the stationary phase and the separation is achieved by partitioning the analyte between that polar layer and the mobile phase. However, greater metabolome coverage can be obtained by employing both reverse and HILIC phase columns (Dunn and Ellis, 2005). An alternative option for the analysis of hydrophilic metabolites with reverse phase columns is use of ion pair reagents containing non-polar moieties and as a result hydrophilic metabolites are retained on hydrophobic stationary phase. This approach has been widely applied in microbial metabolomics to study glycolysis, pentose phosphate and TCA cycle metabolites (Coulier et al., 2006, Luo et al., 2007a).

Recently, UHPLC system which uses a polar mobile phase, typically consisting of water and acetonitrile has been introduced. Separation is achieved at relatively higher flow rate and high pressure, around 15,000 psi, and involve the use of column packed with sub-2-µm stationary phase. Reduction in the particle size, greatly enhances the chromatographic resolution and efficiency, thus can be applied in resolving complex biological mixtures in non-targeted metabolic profiling (Zhang et al., 2012). Miniaturization, such as reduced column dimension and reduced flow rates, increases the sensitivity in ESI-MS. Several applications of the latter in metabolomics and proteomics as demonstrated with *Shewanella oneidensis (Bedair and Sumner, 2008a)* have been reported. Capillary columns packed with nine different stationary phases were evaluated by Ding and co-workers (Ding et al., 2007) for efficient separation of small polar metabolites in Cyanobacterium sp. Good separation was achieved with a Synergi-fusion-RP column with packing material having a pore size less than 0.01 µm and a surface area larger than 400m^2/g. The separation of polar analytes can be further improved with the use of stationary phases which have polar functional groups imbedded in the C18 chain for better interaction with polar metabolites. Alternatively, long capillary monolithic columns can also be used as seen with metabolomics study of *Arabidopsis thaliana,* the use of which resulted in the detection of several hundred peaks (Tolstikov et al., 2003).

Due to complicated experimental set-up, the implementation of 2D LC-MS for metabolomics has lagged behind than that of 2D-GC-MS. In practice, effluents from the first column are transferred to the second at constant interval via automated valves equipped with multiple sample loops. The first column is used for the primary separation mode, usually with long analysis time and gradient elution, while the second column is short with a larger inner diameter, used for the orthogonal separation of the modulated fractions from the first dimensional separation. However, in order to minimize the loss in the first dimensional resolution, fast modulation is required, while to decrease the extra band broadening it requires a very fast second-dimensional separation. A recent study showed that, the loss in the first dimensional resolution can be minimized to 20- 30 % by employing sampling rate of less than 1.5 times the standard deviation of the first dimensional peaks (Schoenmakers et al., 2006). Alternatively, the use of high temperature > 100°C RP gradient elution chromatography, using thermally stable stationary phase results in a second dimensional analysis time of 20 s, thus permitting a high modulation rate. Recently 2D LC-ESI-TOF-MS was successfully applied for the analysis and quantification of *Stevia rebaudiana* glycosides (Pól et al., 2007).

2.3.2.1 Issues and challenges in LC-MS

The off-line identification of metabolites at multiple collision energies with CID is possible; however it is not suitable for the online identification of metabolites in LC applications. There are two types of artefacts introduced by the analytical system in LC-MS analysis: 1) Chemical noise, i.e. real peaks that are generated in the instrument or apparent changes that are instrument related, such as ion suppression 2) Apparent sample differences introduced by the experimental protocol or the data processing (Burton et al., 2008). However, it is essential to detect these artefacts, prior to normalization of the data by using visualization, univariate and multivariate tools, so that they can be identified (Burton et al., 2008). The consistency of the analytical results are assessed by injecting a mixture of reference compounds at regular intervals to assess the performance of both the chromatographic column and the mass spectrometer during the measurements. Alternatively, QC samples can also be used that are representative of the biological samples, but this approach is not sufficient to normalize peak intensities, hence other normalization approaches have been developed to facilitate comparison of datasets, such as NOMIS or CCMN in order to overcome analytical variability. Other instrumental artefacts which are mainly linked to the sample acquisition order can be determined and minimized by randomizing the order of samples throughout the sequence of injections prior to data acquisition. With this approach, the only systematic differences between these samples will be due to time related artefacts and, if all real classes are roughly equally represented, class effect should cancel out (Burton et al., 2008, Roux et al., 2011).

However no single analytical method is suitable for the detection of all the metabolites. As the physicochemical properties such as pK_a, polarity and size of metabolites cover a wide range, hence it is not possible to separate all metabolites with one separation method (GC, LC or CE). In addition, it is not possible to measure all metabolites with one detector. MS cannot detect all metabolites, as some metabolites do not ionize well under certain conditions, or their concentrations are too low for detection. The dynamic range of most MS is also still only 3-4 orders of magnitude, whereas the range of concentration of metabolites in biological samples is often much larger. In addition to being complex, a metabolome is transient, as biochemical pathways are especially sensitive to stresses such as minor changes in pH, temperature, concentration, reaction rates, composition and location. Due to this, metabolome which is measured at one set of conditions will not necessarily be the same as that measured at another set of conditions. Hence, to fully characterize a biological system, the evolving nature of its metabolome must be monitored through flux analysis (Dunn, 2008, Dwivedi et al., 2008).

Thus for the acquisition of a comprehensive metabolite profile, it is essential to set up the generic metabolomics platforms that can analyze the maximum number of metabolites, by parallel application of several techniques, using a combination of GC-MS, LC-MS and/or CE-MS methods.

2.4 Microalgae for biofuel production

2.4.1 What are microalgae?

Algae are photosynthetic aquatic organisms. The term is often used to refer specifically to eukaryotic organisms, thus excluding photosynthetic bacteria (such as cyanobacteria, which are also referred to as ‘blue-green algae’). They may be unicellular (microalgae) or multicellular (macroalgae). The latter category includes seaweeds. The algae are very diverse in evolutionary terms. The red algae and green algae belong to one group, while the others, including diatoms, brown algae, heterokonts, and dinoflagellates have all evolved separately. Whilst some aquatic photosynthetic organisms, such as sea-grasses, are more closely related to flowering plants (Scott et al., 2010).

The biodiversity of microalgae is enormous and they represent an almost untapped resource. It has been estimated that about 200,000-800,000 species exist of which about 40,000 species have been identified. Algae are classified in multiple major groupings as follows: cyanobacteria

(Cyanophyceae), green algae (Chlorophyceae), diatoms (Bacillariophyceae), yellow-green algae (Xanthophyceae), golden algae Chrysophyceae), red algae (Rhodophyceae), brown algae (Phaeophyceae), dinoflagellates (Dinophyceae) and 'pico-plankton' Prasinophyceae and Eustigmatophyceae).

2.4.1.1 Microalgae for non-biofuel products

Over 15,000 novel compounds originating from algal biomass have been chemically determined. Most of these microalgae species produce unique products such as carotenoids, antioxidants, polyunsaturated fatty acids (PUFAs), enzymes, polymers, peptides, toxins, stable isotopes, polysaccharides and sterols (Lamers et al., 2010a, Pulz and Gross, 2004, Wijesekara et al., 2011). The chemical composition of microalgae is not an intrinsic constant factor but varies over a wide range, depending on species and on cultivation conditions. It is possible to accumulate the desired products in microalgae to a large extent by changing environmental factors such as temperature, illumination, pH, CO_2 supply, salt and nutrients.

2.4.1.2 Microalgae for other biofuel products

The first generation bioethanol suffered from food vs fuel dilemma whereas the second generation bioethanol (from lignocellulose feedstock) is associated with higher investment costs and still lacks economic viability for large scale production. The conversion of lignocellulose to biofuel is costly and resistance to degradations adds to the cost of overall biomass saccharification. In contrast, bioethanol can be easily produced from algal biomass through fermentation, as unlike plant biomass, algal biomass do not contain lignin and hence it is less resistant to conversions into simple sugars (Daroch et al., 2013). Furthermore, in addition to cellulose and hemicellulose *C. reinhardtii* (Choi et al., 2010) and *C. vulgaris* (Hirano et al., 1997) have high starch contents as reserve material. Therefore bioethanol from microalgal biomass has the potential to replace conventional gasoline.

During photosynthesis microalgae converts water molecule into hydrogen ions (H^+) and oxygen, and under anaerobic conditions expression of a hydrogenase enzyme which recombines H^+ and e^- (resulting from the photosynthetic electron transfer chain) resulting in H_2 production. *C. reinhardtii* is considered as one of the most efficient eukaryotic producers of H_2 (Amaro et al., 2013). Thermochemical conversion of algal biomass via the gasification process have been shown to convert carbonaceous materials into synthetic gas (syngas) such as methane and nitrogen by means of partial oxidation, in the presence of air, at around 850-1000°C (Amin, 2009). In addition, hydrothermal liquefaction under sub-critical water conditions can be employed to convert wet algal biomass to crude bio-oil (Amin, 2009). Anaerobic digestion of microalgal biomass is also known to yield methane gas. Recently, methane recovery from post trans-esterified microalgal residues of *Chlorella sp.* coupled with optimised biomass and lipid production have been reported by (Ehimen et al., 2011).

2.4.2 Microalgae for biodiesel production

Several microalgal strains have been reported to be significantly rich in oil which can be converted into biodiesel. Biodiesel is a mixture of fatty acid alkyl esters which differ in chain length and unsaturation, obtained conventionally by trans-esterification (alcoholysis) of triglycerides from vegetable oils and animal fats. It is a complex mixture consisting of 90-98% of triglycerides by weight, small amounts of free fatty acids (1-5%), mono and diglycerides, residual amounts of phospholipid, phosphatides, carotenes, tocopherols, sulphur compounds and traces of water (Mata et al., 2010). The overall biodiesel properties are determined by properties of individual fatty esters. The properties of the fatty esters are determined by structural features which include the chain length, degree of unsaturation and branching of the chain. The latter have greater impact on fuel properties such as cetane number, exhaust emission, heat of combustion, cold flow, oxidative stability, viscosity and lubricity (Knothe, 2005).

2.4.3 Lipids and fatty acids

There is no such widely accepted definition for the lipids. However according to (Fahy et al., 2005), lipids are defined as “Hydrophobic or amphipathic small molecules that may originate entirely or in part by carbanion-based condensation of thioesters such as fatty acids, polyketides etc. and/or by carbanion based condensation of isoprene units such as prenols, sterols etc.” On the other hand, lipids are loosely defined by (Manirakiza et al., 2001), as “A diverse group of biological substances made up primarily of non-polar compounds (triglycerides, diglycerides, monoglycerides and sterols) and more polar compounds (free fatty acids, sphingolipids and phospholipids)”. Lipids are covalently bound to carbohydrates and proteins to form glycolipids and lipoproteins respectively. Lipids are broadly classified into simple and complex lipids where simple lipids are those which yield two types of products on hydrolysis (fatty acids, sterols and acylglycerols) whereas complex lipids are those which yield three or more products on hydrolysis like glycerophospholipids and glycosphingolipids.

Fatty acids are the major building blocks of phospholipid and glycolipids in biological membranes and are widely known as fuel molecules for biofuel production. In general, fatty acids contain a long hydrocarbon chain and a terminal carboxylic group. On the basis of their chemical properties, fatty acids are classified into two categories, saturated and unsaturated. Saturated fatty acids (SFAs), are solid at room temperature, in contrast unsaturated fatty acids, are liquid at room temperature. A fatty acid is said to be saturated when each of the carbon atoms except the two at the ends of the chain is bonded to two hydrogen atoms i.e. all the bonding capacity of the carbon is saturated with hydrogen. Examples of SFAs includes: caprylic acid, lauric acid, myristic acid and palmitic acid etc. The fatty acid is said to be unsaturated when two adjacent carbon atoms in the chain except the two at the ends of the chain is bonded to only one hydrogen, and a double bond exist between the pair of carbons. The unsaturated fatty acids are further classified into two categories: monounsaturated fatty acids (MUFAs) and polyunsaturated fatty acids (PUFAs). In case of MUFAs, the chain contains only one double bond e.g. oleic acid whereas in PUFAs the chain contains two or more double bonds e.g. linoleic acid. On the basis of the location of their double bonds PUFAs are further classified into omega 3 and omega 6 fatty acids. An omega or n notation indicates the number of carbon atoms from the methyl end of the acyl chain to the first double bond. Each family is derived from a specific essential fatty acid that cannot be made by humans, and the two families cannot be interconverted. The n-6 family is derived from linoleic acid (LA) whereas n-3 family is derived from alpha-linoleic acid (ALA).

Microalgal lipids

All marine and freshwater microalgae are primarily made up of lipids, proteins, carbohydrates and nucleic acid in different proportions. The microalgal lipids mainly comprise of neutral lipids such as triglycerides, pigments and trace amounts of hydrocarbons and polar lipids such as phospholipids, phosphotidylcholine, sterols, and prenyl derivatives such as tocopherols, terpenes and quinines. Lipid content of algae varies from 1 to 70 % of dry weight and can be escalated to around 90% of dry weight under controlled growth conditions. Lipid content is species dependent and also on the growth conditions and induced stress factors. Recently, Mata et al reviewed, summarised and compared the lipid content and productivities of different algal species among which *Dunaliella, Nannochloropsis* and *Chlorella* species were shown to have oil yields between 20 to 50%, which could be further increased, suitable for biodiesel production (Mata et al., 2010).

2.4.4 Lipid extraction

The sustainability of microalgal biofuel production relies on efficient extraction of lipids. For microalgal biodiesel production, an ideal extraction method should be more selective towards extraction of specific lipid classes and simultaneously minimise the co-extraction of non-lipid contaminants. The bottlenecks in microalgal lipid extraction include:

1) **Selection of suitable microalgal strain for biofuel production:** For biodiesel production, the desired lipid fraction includes neutral lipids containing mono-, di- and trienoic fatty acyl chains (Halim et al., 2011). In brief, biodiesel produced from long chain FAs possess high cloud point, representing a bad fuel characteristic whereas biodiesel derived from short chain FAs have low cloud point, improved flow characteristics and are anti-oxidative (Stournas et al., 1995). Furthermore, the overall biodiesel fuel characteristics such as cetane number, oxidative stability and cold flow are largely dependent on fatty acid composition of the microalgal species, hence selection of appropriate strain is crucial (Slocombe et al., 2013).
2) **Selection of appropriate disruption technique:** The chemical complexity and the variety of structural robustness of algal cell wall can influence the choice of extraction technique. For efficient extraction, it is essential to weaken cell wall by liberation of cell wall mono- or polysaccharides followed by complete disruption of the cell wall prior to lipid extraction. Currently, there is no such disruption technique that can be widely applied to all microalgal species belonging to different taxonomic classes.
3) **Selection of appropriate solvent and solvent ratios:** The commonly applied conventional protocols use solvent mixtures which include chloroform and methanol because of their ability to penetrate the cell wall as they have high polarity index values suggesting higher solubility for all polar lipid compounds. However, both the solvents are toxic, non-environment-friendly, pose health risk to humans and are expensive. As a result many researchers have focused on the use of non-chlorinated solvents which are cheaper, less toxic, and have lesser impact on the environment compared to chlorinated solvents and more importantly they have nearly the same lipid extraction efficiency. However, in the past literature, the extraction efficiencies achieved with the use of chlorinated solvents have been shown to be much higher than achieved with the use of non-chlorinated solvents. In addition, the use of different solvent ratios with the same solvent system can also affect extraction efficiency. Therefore, it is essential to carefully select the optimum extraction solvent system and their ratios prior to use. The selected solvent system should also be evaluated for their selectivity towards desired lipid classes of interest with minimum extraction of non-lipid contaminants.
4) **Effect of pre-treatments:** Pre-treatments for lipid extraction from microalgae include lyophilisation, inactivation of lipases and addition of an antioxidant. All the pre-treatments are time consuming and will incur costs when used on an industrial scale, hence it is essential to assess the requirement for these and the effect of them on lipid extraction efficiency.

2.4.4.1 Lipid extraction techniques and methods

Conventional methods

The original Bligh and Dyer method (Bligh and Dyer, 1959) was optimised for marine tissues such as cod muscle and with the recommendation that the protocol should only be extrapolated to those samples which contain at least 80 ± 1% of water and about 1% lipid. Microalgae often have lipid contents greater than 10%. In spite of this, the method has been treated as a gold standard and is widely used for algal lipid estimation. However, it is essential to keep the actual volumes of chloroform, methanol and water in the proportion of 1:2:0.8, *v/v/v* (before homogenisation) and 2:2:1.8, *v/v/v* respectively (ratios considering the amount of water present in the wet sample biomass).

The original Folch method (Folch et al., 1957) was regarded as a classical method, for lipid analysis in animal tissue, which was based on the assumption that the tissue/sample has specific gravity of water, so 1 g of tissue/sample is computed as 1 ml. Moreover, Folch method relies on initial homogenization of the sample with chloroform-methanol mixture (2:1, *v/v*) to a final dilution, 20 fold the volume of wet sample and later addition of water to remove non-lipid materials, while keeping the final ratios of chloroform-methanol-water to 8:4:3, *v/v/v*.

Solvent extraction

The use of toxic solvents such as chloroform and methanol on an industrial scale will have a great impact on the environment and will pose a threat to health. Considering toxicity of methanol and chloroform, attempts were made in which lipids were extracted from plaice, mussel and herring samples using non-chlorinated solvents (Modified Bligh and Dyer method) as suggested previously (Smedes, 1999a), where methanol was replaced by propan-2-ol or ethanol and chloroform was replaced by cyclohexane or DIPE or MTBE in hexane. Among them propan-2-ol and cyclohexane mixture was found to be the best alternative to the chloroform-methanol mixture. Moreover, cyclohexane has a lower density than chloroform, hence separates on top of the extraction mixture, making isolation of organic phase easier and minimises dripping losses. Another study investigated various combinations of solvents, which are less toxic such as MTBE, ethanol, hexane, IPA, butanol and acetic acid esters for lipid extraction from wild-type *Synechosystis* PCC 6803 (Sheng et al., 2011). The authors concluded that, methanol-MTBE (1.5:5 *v/v*) mixture was the most suitable alternative to methanol-chloroform mixture, although it resulted in slightly lower yield than the conventional solvents, however it is less toxic and less harmful to the environmental as the proportion of methanol was reduced because of the higher proportion of MTBE used. Other solvent mixtures were ineffective in extracting lipids from wild-type *Synechosystis* PCC 6803. In another investigation

(Matyash et al., 2008), the use of methanol-MTBE (1.5:5 *v/v*) for shotgun lipidomics was demonstrated as an alternative with equal or better recoveries for all major lipid classes from mouse brain, human blood plasma, *E. coli* and *Caenorhabditis elegans* samples. Like cyclohexane, MTBE has a lower density than chloroform, hence separates on top of the extraction mixture, facilitating better recovery and minimising dripping losses. The use of dichloromethane (DCM) has been recommended (Cequier-Sánchez et al., 2008) as an alternative to chloroform for lipid extraction for animal and plant samples. DCM is much less hazardous, less toxic and cheaper than chloroform, and equal recovery of FAs to that of chloroform was obtained. Lee and co-workers evaluated different solvent systems for extraction of lipids from green alga *Botryococcus braunii* which included acetone-DCM 1:1, hexane-IPA 3:2, chloroform-methanol 2:1 and DCM-ethanol 1:1 among which the use of chloroform-methanol 2:1 mixture resulted in the highest recovery of lipids (Lee et al., 2010). Recently, Ryckebosch and co-workers evaluated seven different solvent systems for lipid extraction across four different microalgal species and concluded that chloroform-methanol (1:1) resulted in highest lipid yields (Ryckebosch et al., 2012). A recent study (Kumari et al., 2011) compared conventional methods with the method suggested by Cequier-Sanchez et al. (Cequier-Sánchez et al., 2008), and concluded that methanol-chloroform mixture has more potential in extracting lipids compared to that of the DCM-methanol mixture.

Bead beating

Bead beating is the most commonly used and recommended cell disruption technique in the current literature. Recently, it was reported (Halim et al., 2012) that the most effective disruption methods includes high-pressure homogenization, sulphuric acid treatment and bead beating, however the authors did not account for the energy expenses while comparing those methods. Similarly, Serive and co-workers evaluated nine disruption techniques on two microalgal species for metabolomics studies and concluded that the use of glass beads with the use of mixer mill was the most effective technique (Serive et al., 2012). However, in the same study, the author's abandoned the method of using bead beaters because of excessive rise in temperature of the samples. The possible reasons for the higher temperatures might be due to the use of 2 mm i.d. glass beads, which have an upper limit of milling time and number of cycles as 4 minute and 4 cycles respectively, while keeping the relaxation time constant as 1 minute. In contrast to the above report, Ryckebosch and co-workers suggested bead beating was the most effective technique that yielded higher lipid contents compared to sonication and freeze-thaw cycles with liquid nitrogen (LN_2) (Ryckebosch et al., 2012). This finding was in agreement with that of Lee and co-workers (Lee et al., 2010). Some report suggests combined use of bead beating with microwave (Lee et al., 2010) or sonication bath (Breuer et al., 2013), which might not be a practical solution in terms of time and cost incurred. It is clear that there is an important missing link while comparing these disruption techniques. There is a lack of agreement on the selection of the experimental parameters.

Sonication

In contrast to bead beating, the use of sonication in conjunction with deep freezing in LN_2 and mechanical grinding has been reported (Jaki et al., 2006) as a better disruption technique for metabolomics studies of mycobacterial cells. Similarly the combined use of grinding, microwave treatment and sonication has been demonstrated (Šoštarič et al., 2012) to achieve better yields of lipids from microalga *C. vulgaris*. Recently, Araujo and co-workers evaluated and suggested that sonication improves the efficiency and recovery of lipids from microalga *C. vulgaris* while investigating the efficiency of five lipid extraction methods (Araujo et al., 2013). In addition, Adam and co-workers reported a solvent free ultra-sound assisted extraction of fresh microalgae as an optimum method to extract lipids from *N. oculata* (Adam et al., 2012). Other studies (Kumari et al., 2011, Paik et al., 2009) have also incorporated sonication as an efficient disruption techniques for fatty acid profiling analysis of different microalgal species by GC-FID.

Other lipid extraction techniques

The other techniques widely used for lipid extraction includes: conventional Soxhlet apparatus, supercritical fluid extractor (SFE), thermal liquefaction, solid phase extraction (SPE), accelerated solvent extraction (ASE) and microwave assisted extraction (MAE). However, overall use of ASE, Soxhlet, and thermal liquefaction requires longer extraction times, larger solvent requirements and high energy inputs making them inappropriate candidates for large scale biofuel production from microalgae.

MAE seems to be a more environmentally friendly option compared to ASE, Soxhlet, thermal liquefaction and SPE because the high heating rates effectively reduces the extraction time and volume of extraction solvent required. The working principle of this method has been described by Balasubramanian and co-workers (Balasubramanian et al., 2011). Briefly, when an oscillating electric field at frequencies of MHz – GHz was applied to wet microalgal biomass containing water and other polar compounds, oscillations of polar molecules occurs resulting in inter- and intra-molecular friction coupled with the movement and collision of very large number of charged ions resulting in rapid heating of the sample within seconds. This subsequent intracellular heating causes pressurised effects resulting in rupturing of microalgal cell wall along with electroporation effects. MAE has several advantages over conventional extraction methods. In addition to the excellent recovery of interested components other key advantages include: ease of use, fast and shorter reaction times, single step conversion process, removal of water (expensive step) is not required instead water serves as an excellent solvent for extraction and is non-toxic (Patil et al., 2012). Recently, (Balasubramanian et al., 2011) developed and optimised continuous microwave system for efficient lipid extraction from wet green alga *Scenedesmus obliquus*, where higher percentages of unsaturated and essential fatty acids were recovered with superior quality compared to conventional Soxhlet extraction method. Despite all these advantages offered by MAE, the method usually involves, the use of organic solvent such as hexane for microalgal lipid extraction, which is not considered to be environmentally friendly for industrial scale application. Moreover closed vessel MAE system poses high risk of injuries to analysts due to use of higher pressures, whereas polytetrafluoroethylene (PTFE) material used in the construction of vessels does not allow the use of high temperature and therefore additional cooling steps are required in order to avoid the loss of extracted volatile components. Open vessel systems are safer than closed vessel systems but the extraction conditions are less reproducible. In addition it cannot process many samples simultaneously and requires longer extraction times to achieve efficiency similar to that of closed vessel system (Tatke and Jaiswal, 2011).

MAE seems to be a suitable candidate as it offers higher extraction efficiencies, however problems associated with the use of environment friendly solvent are still under investigation and in addition the system poses high risk to analyst dues to use of higher pressures in closed vessel system.

2.4.5 In situ lipid extraction and transesterification

Fatty acids can occur in nature in the free unesterified form, however most often fatty acids are found as esters, linked to glycerol, cholesterol or long chain aliphatic alcohols, and as amides in sphingolipids. The physical state of the lipids varies greatly, and they can be isolated as pure lipid classes or else remain as a mixed lipid extract (Christie, 1993). The esterification or trans-esterification (TE) reactions are carried out either after extraction of the lipids or else they can be effected while the lipids are present inside the sample matrix. It is important to note that there is a difference between the terms esterification and transesterification. Derivative reactions are usually catalysed by acidic or basic media or in the presence of acid or base catalyst. The reactions usually occur with the following: a) free fatty acids in sample in presence of catalyst. The reaction is called esterification. b) Fatty acids included in complex lipids by *O*-acyl bonds (ester) (such as phospholipids, triacylglycerols (TAGs)) in acid or basic media. The reaction is called TE or trans-methylation. TE reactions are also known as alcoholysis or methanolysis. As most of the fatty acids in the biological samples are included in TAGs and phospholipids, thereby derivatization reactions of fatty acids from such biological samples are often referred to as transesterification or FAME synthesis from fatty acid mixtures (Carrapiso and García, 2000).

Direct transesterification (DT) methods offer several advantages over two step TE methods (extraction followed by TE). In DT methods, a separate extraction step is avoided and FAs are simultaneously extracted and transesterified in presence of a catalyst.

2.4.6 Catalysts used in TE reactions

Homogeneous acid catalysed TE

The main advantage of acid catalysed reactions are their ability to catalyse the esterification from samples containing higher content of FFAs (free or linked). They also have an ability to simultaneously catalyse both the fatty acid esterification and oil TE reactions. However, the main drawbacks associated with the use acid catalyst compared to base catalyst include: they are 4000 times slower reaction rate than the base catalysed reactions, high molar ratios of alcohol to oil are required, and higher temperatures (60°C to 90°C) and longer reaction times. The commonly used acid catalysts in methanol are HCl, H_2SO_4 and BF_3.

In 1964, Morrison and Smith (Morrison and Smith, 1964) recommended the use of Lewis acid BF_3/methanol as a powerful acidic catalyst for the preparation of FAMEs and dimethylacetals. It has potential to transesterify most lipid classes but usually longer reaction times are required. Subsequently, the TE efficiency of acetyl chloride/methanol and BF_3/methanol was investigated and the results demonstrated that acetyl chloride was the most appropriate reagent (Lepage and Roy, 1984). In addition, severe peak tailing of solvent peak was another major disadvantage and this finding was in agreement with the findings of Rogiers and co-workers (Rogiers, 1977). Other problems associated with the use of BF_3/methanol are well summarised in reviews (Carrapiso and García, 2000, Christie, 1993). Briefly, the BF_3/methanol can be quickly destroyed in the presence of water, it is more toxic, expensive and has a limited shelf life even when refrigerated. In addition to this, it suffers from the formation of major artefacts, and the use of old or highly concentrated (~50%) solutions may result in loss of PUFAs. If the extract contain antioxidants such as BHT, BF_3/methanol is known to react with BHT to produce methoxy derivatives which co-elute with methyl pentadecanoate or hexadecanoate in GC analysis. Recently, Xu and co-workers suggested the use of acetyl chloride/methanol as the best option compared to other reagents (Xu et al., 2010). This finding is in agreement with the findings of Lepage and Roy (Lepage and Roy, 1984), however the authors suggested addition of the reagent should be done in a dry ice bath, as addition at room temperature leads to exothermic reaction which can lead to the loss of PUFAs due to their lower stability and this can result in overestimation of C16:0 and C18:1 FAs. On the other hand, Harmanescu et al evaluated TE efficiency of BF_3/methanol and HCl/methanol and concluded that BF_3/methanol is better, as it resulted in less interfering compounds and better statistical parameters compared to that of HCl/methanol (Harmanescu, 2012). It is important to note that the findings are in contrast to that of Lepage and Roy (Lepage and Roy, 1984). However, the authors did not consider the use of acetyl chloride/methanol for evaluation purposes.

Homogeneous base catalysed TE

The main advantage of base catalysed reactions over acid catalysed reactions are, they are widely preferred for industrial scale application because of their higher efficiency, lower cost and less corrosiveness. In addition, they have faster reaction rate and require lower temperature and pressure conditions (Ejikeme et al., 2010). In contrast, the main drawbacks are the formation of soap between FFAs (content > 3%) and base catalyst. Algae usually contains as high as 35.1% of FFAs which make them highly unsuitable for biodiesel production from microalgae (Daroch et al., 2013). In addition, the post-reaction treatments to remove catalyst from biodiesel and treatment of the alkaline waste water is more expensive on industrial scale. The most commonly used alkali catalysts are NaOH, KOH, TMG and sodium methoxide (CH_3ONa). CH_3ONa is the most popular among them, because at much lower concentration about 0.5M, 98% yield could be easily obtained with shorter reaction time (30 minutes), but it requires complete absence of water, thereby making it inappropriate for industrial scale application (Ejikeme et al., 2010). On the other hand, a special type of non-ionic base catalyst such as TMG has been reported in a comprehensive review by Schuchardt et al (Schuchardt et al., 1998) to be the most suitable alkaline catalyst for *in situ* derivatization as it has the ability to react with FFAs as well.

Combined: acid and base catalysed TE

The long reaction times required with the use of homogeneous acid or base catalyst can be reduced drastically by sequential use of base catalyst followed by an acid catalyst. As FAs from TAG are saponified or trans-esterified at a faster rate using base catalyst and FFAs are then esterified using acid catalyst owing to their high esterifying power. Use of CH_3ONa followed by BF_3/methanol has been widely applied. Recently, Griffiths and co-workers investigated the sequential use of CH_3ONa followed by BF_3/methanol in presence of water for two microalgal species and concluded that 100% TE efficiency was obtained with sequential use of both catalysts compared to each separately (Griffiths et al., 2010). Similarly, Laurens and co-workers applied the standard AOAC method 991.39 to four microalgal species involving the use of

CH_3ONa and BF_3/methanol, however in this case both the acid and base catalyst were added simultaneously and incubated together which was in contrast to that of (Griffiths et al., 2010) method (Laurens et al., 2012b). However, so far sequential use of CH_3ONa with acetyl chloride in methanol or methanolic HCl has not been applied to microalgae for evaluating FAME recoveries.

2.4.7 Lipid determination

The most popular and commonly used conventional methods include extraction of lipids by using single or mixed solvent followed by lipid quantification by gravimetry (weighing) or by gas chromatography (GC). The conventional gravimetric estimation of total lipid content has been in practice for many years and is still routinely used. This method usually involves lipid extraction by organic solvent followed by estimation of total lipid content by gravimetric determination. Gravimetric determinations are less time consuming and do not require specialised instruments such as GC or HPLC. However, the quantitative determination of lipids by gravimetry is highly inaccurate and could be overestimated due to the presence of non-fatty acid compounds such as proteins, pigments and steroids. In addition, the accuracy and precision of this method is dependent on the accuracy and precision of the weight and volume measurements which cannot be corrected with the use of an internal standard (Phillips et al., 1997). Therefore, for biodiesel production, where only fatty acid content of the microalgal extract is of interest, gravimetric determinations may not provide required accuracy.

On the other hand, GC along with flame ionisation detector (FID) is the most popular method for the fatty acid profiling of lipids as compared to that of gravimetric determinations. GC is capable of quantifying individual fatty acids as well as the total fatty acid content in a very short time. However the sample needs derivatization to volatile samples prior to GC-FID analysis, which is usually achieved by TE of lipid-bound FAs (TAGs and PLs) and esterification of FFAs to FAMEs in the presence of excess methanol and a catalyst. These procedures eliminate the probability of overestimation of total FA content as unlike gravimetric estimations, non-fatty-acid substances are not detected by GC.

The identification of fatty acids in a sample is relatively simple and is based on retention times of commercially available authentic FAMEs standards. However, in certain application where identification of complex organic compounds is essential, GC-MS is the preferred choice because it combines the retention time parameters determined by GC with the structural information provided by the mass spectrometry (MS). GC-FID with fused-silica capillary column comprising of polar cyanopropyl polysilphenylenesiloxane phase (70%) is the most preferred and frequently used analytical methods for FAME analysis due to its high sensitivity and linear range response for carbon-containing compounds.

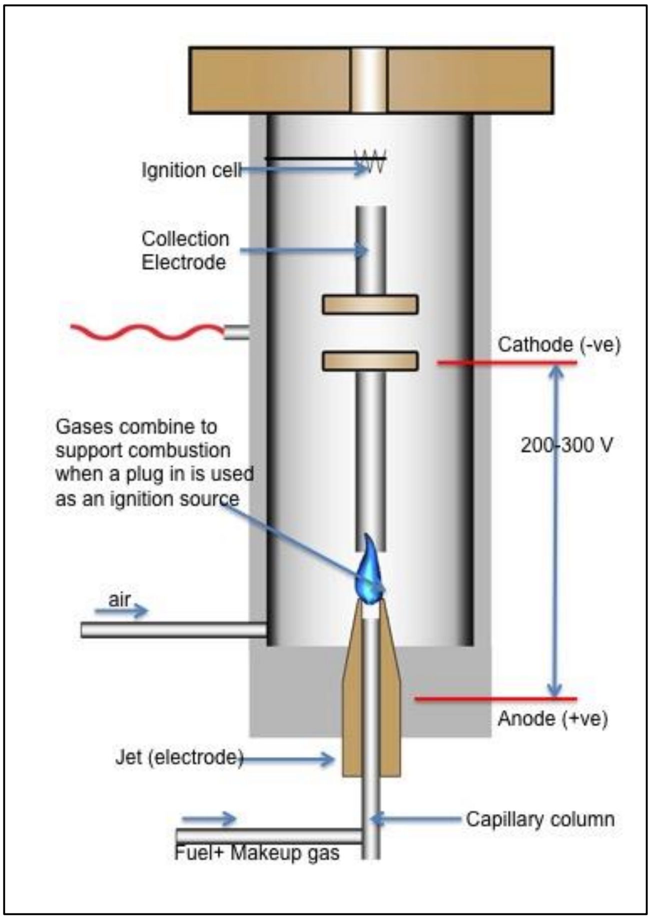

Ignition cell
Collection Electrode
Cathode (-ve)
Gases combine to support combustion when a plug in is used as an ignition source
200-300 V
air
Anode (+ve)
Jet (electrode)
Capillary column
Fuel+ Makeup gas

Figure 2.3 A Typical flame ionisation detector design (Figure modified from http://www.chromacademy.com).

A typical FID design is shown in figure 2.3. Briefly, a small voltage (200-300V) is applied between anode and cathode resulting in ion formation, as the column effluent is burned in the flame produced from a mixture of hydrogen, make up gas and air. The exact mechanism of ion production is not well characterised, however it is suggested that any hydrocarbons in the sample will produce carbon ions via pyrolysis and small organic fragment ions via high energy combustion. FID produces a constant proportional response to the number of carbon atoms in a molecule. Sensitivity of FID detector can be severely reduced when heteroatoms are present. Therefore, for quantitative determinations one should always generate calibration curves for each analyte in order to account for the response variations due to the nature of the compounds.

2.5 Microalgal metabolomics

"Microalgae not only play an important ecological role, but are also of commercial importance and therefore call for an in depth knowledge of basic biological functions through metabolomics" (Jamers et al., 2009a). In algae, most metabolic analyses have so far been focused on the quantification and identification of secondary metabolites with economic value in food science, the pharmaceutical industry and public health, among others. Fatty acids, steroids, carotenoids, polysaccharides, lectins, polyketides and algal toxins are among the algal products that have been studied (Siew Moi, 2004).

Algae are one of the most important bio-factories on earth based on their photosynthesis/CO_2 fixation capacity, and rich sources of biofuels of the third generation (Wienkoop et al., 2010a). Not only do these organisms fix carbon dioxide, but they also have the potential to be used for the production of inexpensive bulk chemicals because the major inputs (light and CO_2) are essentially free. However, to harness this potential through metabolic engineering, a deeper understanding of photosynthetic metabolism is required (Boyle and Morgan, 2009, Field et al., 1998, May et al., 2008b). Bottlenecks in algal biodiesel production within the cell can be identified by metabolomics approaches in combination with proteomics and transcriptomics, as the identification of differentially expressed metabolites gives clues to rate-limiting processes in cell. Finally the integration of such a reliable metabolomics data will help us in identifying the as-of-yet missing reactions in the metabolic network which may help to fine tune algal properties by genetic or metabolite engineering in combination with other system biology approaches.

Metabolomics also has a good potential to support environmental risk assessment of chemicals in future, even if the genome of a test species is not known. Kluender et al in their study showed that, the parallel analysis of algal growth parameters (cell volume) and metabolite composition indicates that metabolomics can support interpretation of well-established toxicological assessment parameters. The study also revealed multiple metabolic markers, responding to exposure, providing additional observation parameters to traditional endpoints in phytotoxicity assessment (Kluender et al., 2009a). Understanding toxicant induced deviations from developmental metabolic trajectories may help to discriminate between primary and secondary toxic effects.

As metabolites are the first to react to stressors, (Jamers et al., 2009b, Kluender et al., 2009a, Viant, 2007b), it will be advantageous to carry out the evaluation of stress induced effects in microalgae in combination with multivariate data analysis. Examples of such studies, as applied by various researchers in past includes: evaluation of phytotoxic effects on green alga

Scenedesmus vacuolatus during the light phase of the cell cycle and profiling altered responses of gene expression or metabolite level in response to availability of macronutrients such as phosphate, sulphur, carbon and nitrogen in *C. reinhardtii.* (Kluender et al., 2009a, Lee and Fiehn, 2008b). Recently, the metabolomics study enabled researchers to propose a working model for the co-ordination regulation of cellular metabolism during the induction of the carbon concentrating mechanism in *C. reinhardtii* (Kempa et al., 2009a, Renberg et al., 2010).

These observations indicate the importance of understanding metabolic pathways and their underlying regulation. In particular, the increased energy burden of any new pathway, and more importantly, the interdependence of pathways (Scott et al., 2010). The integration of metabolomics data in the draft metabolic network of microalgae will help in the identification of missing reactions in the network. Considering the fraction of metabolites whose presence cannot yet be explained by the draft network, it can be estimated how incomplete the network actually still is.

However very limited numbers of studies were reported on optimised non-targeted metabolic profiling approach that can potentially detect and quantify hundreds of metabolites, in order to obtain as much information as possible about key metabolites in micro algal species such as *C. reinhardtii, N. salina and D. salina*. Optimised quenching and extraction protocols with parallel application of optimized MS-based hyphenated techniques, will provide initial insight into the metabolic pathways, which will form the basis for a detailed investigation.

2.6 Mammalian metabolomics

Mammalian metabolomics has received special attention in recent years mainly because of its potential in both *in vitro* and *in vivo* oncology studies, specifically in the prognosis and diagnosis of cancer and for assessing the treatment efficacy by analysing cells, fluids or tissues for specific biomarkers. Metabolomic biomarkers are generally explored pre-clinically using animal models or human subjects followed by their quantification and validation in biofluids (urine, serum samples) or directly using tumour tissues (Mamas et al., 2011). The metabolomics workflow for biofluid analysis is well developed and several reports have been published for exploring tumour biochemistry with metabolome analysis of urine, serum, plasma, saliva, bronchial washes, prostatic secretions and/or faecal water (Spratlin et al., 2009). The application of metabolomics directly to tumour tissues is more complex and expensive as it requires microdissection techniques, without which the protocols could be extremely tedious, as contamination from surrounding stromal and epithelial cells could severely alter the metabolomics profile (Spratlin et al., 2009). In contrast to this, application of metabolomics directly to mammalian cell cultures instead of tumour tissues, animal models or human subjects offers several advantages. The application of metabolomics to mammalian cell culture provides a controlled reproductive but relevant model without the requirement of ethical consideration for *in vivo* and human studies. Data are easier to interpret as focusing on a specific cell line reduces variability and provide a more constant background against which subtle metabolic changes can be identified (Čuperlović-Culf et al., 2010). Consequently, metabolome analyses on immortalized cancer cell lines in cancer research is gaining momentum as a tool to understand the molecular mechanism of disease progression, response and resistance to therapeutics, leading to the identification of specific biomarkers which may assist distinguishing between normal, benign and metastatic cancer states (Sheikh et al., 2011). However, as reported previously there have been insufficient reports on the application of metabolomics in the analysis of cultured mammalian cell lines (Čuperlović-Culf et al., 2010, Khoo and Al-Rubeai, 2007). As a result the potential of metabolomics in traditional and emerging biological areas have yet to be fully realized.

Sampling and quenching methods have extensively studied in case of prokaryotes as discussed in section 2.2.1, but the methods cannot be simply adopted for mammalian cells due to basic differences in cell structure. Briefly, mammalian cells lack cell wall, and instead have a cell membrane which makes it more prone to metabolite leakage due to the delicate and fragile nature of the cell envelope compared to that of prokaryotes (Dietmair et al., 2010). As a result, in case of CHO cells, numerous authors (Dietmair et al., 2010, Sellick et al., 2008, Volmer et al., 2011, Wurm and Zeng) have used various washing solutions and additives for quenching solutions which aid in preserving the membrane integrity by maintaining the ionic strength and preventing osmotic shock. However, rigorous studies demonstrating effective quenching methods for adherent cell cultures have been few and far between. The sampling and quenching of adherently grown cells is much more complicated than suspension cultures, as the former involves detachment of adhered cells from the bottom of the cultured flask. This might be a possible obstacle and deterrent for many researchers while development of sampling and quenching protocols for adherent cells. The detachment of adherent cells from bottom of culture flask is popularly achieved by either trypsinization (Beloueche-Babari et al., 2006, Lane and Fan, 2007) or cell scraping (Metallo et al., 2009, Yuan et al., 2012). In trypsinization, which is performed before quenching, the enzyme trypsin is used to cleave membrane-based integrin attachments and release the cells in the medium to get a suspension from which the metabolites can be extracted. EDTA, a calcium chelator is often added to mop up divalent cations which would otherwise inhibit trypsin. In cell scraping, which is performed after quenching, the cells are detached by mechanically lifting the cells or scrapping them off gently by means of a cell lifter or a scraper. In case of simultaneous quenching methods, Teng and co-workers suggested direct quenching and extraction of human breast cancer cells using methanol (Teng et al., 2009).

Similar approach with use of LN_2 was recently reported (Lorenz et al., 2011), where the authors recommended washing with water, direct LN_2 quenching and rapid single step extraction for LC-MS based metabolomics of adherent cultures. Various uses of LN_2 have been reported in the literature including for preserving mammalian cell lines, extraction of metabolites using freeze-thaw cycles and so on; where the procedures are simple and various safety measures are available to conduct these procedures safely. In the case of suspension cultures, quenching with LN_2 does not allow separation between intra- and extracellular metabolites. However in the case of adherent cultures, direct quenching with LN_2 can be successfully applied (Lorenz et al., 2011), where LN_2 is directly poured into culture flask or well plates. As previously reported (Teng et al., 2009) trypsinized cells exhibit a chemical deviation from cells in their attached state and the effective time of incubation employed for sufficient detachment of cells by trypsin treatment can lead to significant cell leakage. Given differences in the cell types, metabolite leakage during trypsinization or cell scrapping in different cell types could be affected to different degrees.

Apart from simultaneous quenching methods, influence of various washing steps and inclusion of various buffer additives to methanol quenching on extent of metabolite leakage in sequential quenching protocols with adherent cultures is not well documented. To date, we are not aware of any report except that of a very recent report by Purwaha and co-workers (Purwaha et al., 2014), where the authors evaluated the use of buffer additives such as AMBIC, HEPES and NaCl and concluded that the use of methanol/AMBIC is an effective quenching solution for OVCAR-8 cells (ovarian cancer cells). Despite potential advantages of metabolomics, there are several bottlenecks in sample preparation workflow for metabolome analyses of mammalian cell cultures, among which the most important is the development of accurate and reliable sampling and quenching methods.

Chapter 3

Metabolite leakage in metabolome analyses of adherent mammalian cell lines

3. Metabolite leakage in metabolome analyses of adherent mammalian cell lines

3.1 Introduction

Metabolome analyses has been increasingly applied to the human body fluids, for various clinical applications in order to define biomarkers related to the prognosis or diagnosis of a disease or to test drug toxicity/efficacy. These investigations provide improved understanding of the pathophysiology of disease or information on the therapeutic pharmacokinetics of drugs (Mamas et al., 2011). Metabolomics is recognised as a valuable tool for the evaluation of the cellular state and the determination of structure-pathway-activity-relationships (SPARs) considered essential for drug discovery (Khoo and Al-Rubeai, 2007). However, as reported previously there have been insufficient reports on the application of metabolomics in the analysis of cultured mammalian cell lines (Čuperlović-Culf et al., 2010, Khoo and Al-Rubeai, 2007). Thus the potential of metabolomics in traditional and emerging biological areas have yet to be fully realized. The application of metabolomics to mammalian cell culture provides a controlled reproducible model which can be used without the requirement for ethical approval for *in vivo* and human studies. Data are easier to interpret as the focus is on for a specific cell line thus reducing variability and providing a more constant background against which subtle metabolic changes can be identified (Čuperlović-Culf et al., 2010). Consequently, metabolome analyses on immortalized cell lines in cancer research is gaining momentum as a tool that can be used to understand the molecular mechanism of disease progression, response and drug resistance resulting in the identification of specific biomarkers, which may assist in distinguishing between normal, benign and metastatic cancer states (Sheikh et al., 2011).

Recent advances and developments in analytical technologies have enabled profiling of the mammalian cell metabolome. GC-MS is one of the most widely used analytical techniques in metabolomics, which combines the high separation efficiency and resolution of capillary GC that is essential for complex metabolic profiling with the high sensitivity of mass selective detection. A variety of volatile and/or derivatized non-volatile metabolites can be analysed qualitatively and quantitatively with high analytical reproducibility and lower costs as compared to LC-MS, CE-MS and NMR. However improvements in the analytical techniques are not sufficiently supported by similar improvements in methods for sample preparation. The entire metabolomics experimental workflow should be planned, and requires careful consideration, because the instrumental data are only as good as the experimental design and sample pre-treatments, which are invariably linked to the quality and reliability of the metabolomic data.

For the development of an analytical protocol to quantify metabolomes, it is essential to arrest the metabolic activities, by inactivating enzymes as rapidly and with minimal disruption to metabolism as possible, during the period of harvesting, in order to obtain a valid snapshot of metabolism without altering the internal metabolite signature. The ideal quenching solution should instantly arrest cellular metabolic activity without causing any significant cell membrane damage thus preventing leakage of intracellular metabolites from the cells. The quenching methods commonly employed are highly sample and cell dependent, and differences in cell membrane, cell wall structure and cell size influences the efficiency of quenching and the rate of metabolite leakage (Sellick et al., 2008). Often the methods employed for quenching and other sample pre-treatment approaches in metabolomics are not sufficiently characterised with respect to the effect they have on metabolite leakage from the cells.

Several studies have been published on the optimization of sample pre-processing protocols, for example, on suspension cell cultures (Dietmair et al., 2012, Dietmair et al., 2010, Kronthaler et al., 2012a, Sellick et al., 2011, Volmer et al., 2011, Wiendahl et al., 2007), but there are only a few reports on the handling of adherent mammalian cells for GC-MS based metabolomic analysis (Danielsson et al., 2010, Dettmer et al., 2011, Hutschenreuther et al., 2012b). The sampling and quenching of adherently grown cells is more complicated than for suspension cultures, because of the need to detach adherent cells from the culture flask. Trypsinized cells exhibit a chemical deviation from cells in their attached state and the effective incubation time to achieve sufficient cell detachment can lead to significant cell leakage. Given differences in the cell types, metabolite leakage during trypsinization or cell scrapping in different cell types is likely to be affected by varying degrees. Although this has been suggested in published literature, there is minimal experimental evidence to support these claims (Dettmer et al., 2011, Hutschenreuther et al., 2012b, Lorenz et al., 2011, Teng et al., 2009). It is not clear whether, and to what degree, trypsinization and cell scrapping affect different cell types. This is a crucial step for devising appropriate reproducible protocols for representative metabolome analyses of adherently grown mammalian cells.

An alternative to the conventional approaches involves simultaneous quenching and extraction of adherently grown cells using a single organic solvent such as cold methanol (Martineau et al., 2011, Teng et al., 2009) or a combination of methanol and water (Dettmer et al., 2011). The use of acetone for deproteination of mammalian cells prior to metabolome profiling has also been suggested for both suspension cell cultures (Aranibar et al., 2011, Bruce et al., 2009, Tiziani et al., 2008) & adherently grown cells (Danielsson et al., 2010). Recently, the use of liquid nitrogen for direct quenching, followed later by extraction using a suitable extraction solvent, was suggested for LC-MS based metabolomics of adherently grown clonal *β*-cell line INS-1 (Lorenz et al., 2011) and human pancreatic cancer cells Panc-1 (Bi et al., 2013). The breast cancer MCF-7 cell line has been analysed using GC-MS based metabolomics after a similar quenching and extraction approach (Hutschenreuther et al., 2012b).

The objective of this study is to assess whether and to what extent the conventional methods of trypsinization and cell scrapping affect metabolite leakage in different adherently grown mammalian cell lines. Two breast cancer cell lines, the primary MCF and the metastatic MDA-MB 436, in addition to the human microvascular endothelial cell line HMEC-1, were studied. Furthermore, we characterised the effect of acetone precipitation and combined quenching and extraction protocols, for rapid and reliable metabolome analysis of adherently grown mammalian cells, with the aim of improving metabolome coverage and reducing loss of intracellular metabolites during harvesting and quenching. We also investigated metabolome leakage with respect to recoveries of different metabolite classes during the quenching and extraction steps.

3.2 Materials and Methods

3.2.1 Chemicals and reagents

The RPMI-1640 medium was obtained from Lonza (Gibco-BRL, Paisley, UK). All other reagents and consumables were obtained from Sigma-Aldrich (Dorset, U.K.), unless stated otherwise.

3.2.2 Cell lines, cell culture and growth assessment

The MDA-MB-436, MCF-7 breast cancer cell lines and the HMEC-1 cell line (Human dermal micro vascular endothelial cells) were obtained from American Type Culture Collection (ATCC), (http://www.atcc.org/). MDA-MB-436, HMEC-1 and MCF-7 cell lines were cultured at 37˚C with 5% CO_2 in T75 flasks with approximately $1x10^6$ cells per flask containing 10 ml RPMI media, 10% fetal bovine serum, 1% penicillin/streptomycin and 1% glutamine. Growth curves were produced by seeding cells in 24 well plates with 1 mL media and a seeding density of $0.1x10^6$ cells per plate. Viable cell counts are obtained at 24, 48 and 72 hours using a Beckman Coulter Vi-Cell XR cell viability analyser (Beckman Coulter, Germany) after trypsinization. Control flasks were prepared without seeding cells, to determine any unwanted background signal.

3.2.3 Cell quenching

The cells were grown to mid-“exponential” phase (36 hours) and then rapidly quenched before extraction. Initially, the culture medium was removed from the T75 flask and cells were quickly washed twice with 5 mL ice-cold phosphate-buffered saline (PBS, pH 7.4). The residual PBS was removed by vacuum. The cells were rapidly quenched by addition of fivefold volumes of precooled 60% aqueous methanol supplemented with 0.85% (w/v) ammonium bicarbonate (AMBIC, pH 7.4) at -50°C to 1 x 10^7 cells, unless stated otherwise (Sellick at al. 2008). Addition of cells to the quenching solution increased the temperature by no more than 15°C.

3.2.4 Metabolite Extraction

Quenching and extraction were performed using the methods described below and illustrated schematically in Figure 3.1.

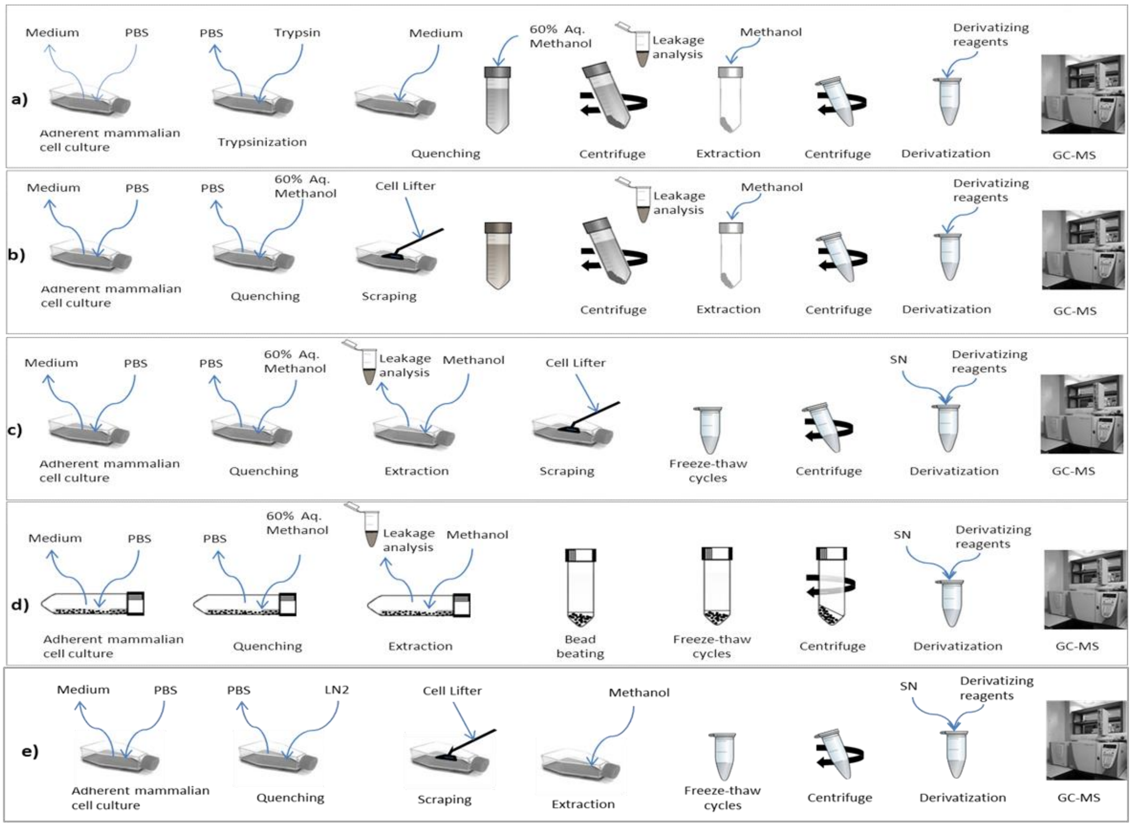

a)
Medium
PBS
Adherent mammalian cell culture
PBS
Trypsin
Trypsinization
Medium
Quenching
60% Aq. Methanol
Leakage analysis
Centrifuge
Methanol
Extraction
Centrifuge
Derivatizing reagents
Derivatization
GC-MS
b)
Medium
PBS
Adherent mammalian cell culture
PBS
60% Aq. Methanol
Quenching
Cell Lifter
Scraping
Leakage analysis
Centrifuge
Methanol
Extraction
Centrifuge
Derivatizing reagents
Derivatization
GC-MS
c)
Medium
PBS
Adherent mammalian cell culture
PBS
60% Aq. Methanol
Quenching
Leakage analysis
Methanol
Extraction
Cell Lifter
Scraping
Freeze-thaw cycles
Centrifuge
SN
Derivatizing reagents
Derivatization
GC-MS
d)
Medium
PBS
Adherent mammalian cell culture
PBS
60% Aq. Methanol
Quenching
Leakage analysis
Methanol
Extraction
Bead beating
Freeze-thaw cycles
Centrifuge
SN
Derivatizing reagents
Derivatization
GC-MS
e)
Medium
PBS
Adherent mammalian cell culture
PBS
LN2
Quenching
Cell Lifter
Scraping
Methanol
Extraction
Freeze-thaw cycles
Centrifuge
SN
Derivatizing reagents
Derivatization
GC-MS

Figure 3.1 Overview of different metabolomic workflows applied across different adherently growing mammalian cell lines. a) Trypsinization b) Cell scraping c) Modified cell scraping method and d) Bead harvesting e) Direct quenching and extraction with LN_2.

3.2.4.1 Trypsinization

Trypsinization is the conventional sample preparation method for adherently grown mammalian cells, described in Figure 3.1a (Beloueche-Babari et al., 2006, Lane and Fan, 2007, Teng et al., 2009). After removal of culture medium and two washings with PBS, 3 mL of trypsin was added to the T75 culture flask with 5 minutes incubation time to detach the cells, which were then re-suspended in 3 mL of culture medium and rapidly quenched. The quenched biomass was then centrifuged for 5 minutes at 2500*g* at -9°C, with the rotor pre-chilled at -24°C to minimize any increase in the temperature during centrifugation. The supernatant was removed rapidly, and an aliquot (1 mL) was transferred to a 2 mL pre-chilled Eppendorf to assess the leakage of internal metabolites. The cell pellets and supernatants were rapidly snap frozen in liquid nitrogen and stored at -80°C for further analysis. The procedure was repeated simultaneously for the RPMI 1640 media which was treated as a control sample.

3.2.4.2 Cell scraping

Cell scraping (Figure 3.1b), is another conventional sample preparation method used previously for adherent cells (Metallo et al., 2009, Yuan et al., 2012). In contrast to trypsinization, cells are quenched after removal of culture medium and two PBS washings. The cells are then suspended in the quenching solution by cell lifting with a pre-chilled cell lifter, all these steps were performed on dry ice. The quenched biomass is then centrifuged; with cell pellets and supernatants collected and stored at -80°C.

3.2.4.3 Modified cell scraping

The modified sample preparation method, in which cells were rapidly quenched after removal of culture medium and PBS washings, in a similar manner to the cell scraping approach, but in contrast the quenching solution was removed rapidly at this stage, and an aliquot (1 mL) was transferred to a 2 mL pre-chilled Eppendorf to assess the leakage of internal metabolites (Figure 3.1c). This step was performed on dry ice. The adherent cells in the T75 flask were then subjected to extraction by addition of 750 µL of pre-chilled 100% methanol, followed by cell lifting with a pre-chilled cell lifter, all these steps were performed on dry ice. The cell suspension was then transferred to 2 mL pre-chilled Eppendorf and stored on dry ice. A further 750 µL aliquot of 100% methanol was added to culture flask and the procedure was repeated. The first and second aliquot was mixed together and metabolites were extracted by performing freeze-thaw cycles as suggested elsewhere (Winder et al., 2008). Briefly the aliquot was snap frozen in liquid nitrogen for 3 minutes, followed by thawing on dry ice and vortexed. The freeze-thaw cycle was repeated three times for complete cell disruption and followed by centrifugation at 16000*g* at -

9°C for 5 minutes. The pellet was re-extracted with 500 µL of 100% methanol. The first and second aliquots were then combined together and the extract was lyophilized overnight prior to metabolite derivatization.

3.2.4.4 Bead harvesting

In this novel experimental approach, the cells were cultured on beads in 50 mL Falcon tube (in a fixed horizontal position) instead of the conventional T75 flask, as shown in Figure 3.1d. The quenching and extraction was performed as described for the modified cell scraping method except that the cells were bead beaten prior to freeze-thaw cycles.

3.2.4.5 Acetone precipitation

The quenching and extraction was performed as described for the trypsinization method except that the extract was subjected to protein precipitation by acetone prior to lyophilisation. (Tiziani et al., 2008). Briefly, after performing the freeze-thaw cycles for extraction of intracellular metabolites, the four times extract volume of pre-chilled acetone (-20°C) was added to the extract. The tube was vortexed and incubated for 60 minutes at -20°C, followed by centrifugation for 10 minutes at 15,000 *g*. The supernatant was carefully transferred to a new Eppendorf and lyophilized overnight prior to metabolite derivatization.

3.2.4.6 Direct quenching with liquid nitrogen

A recently described method was used (Bi et al., 2013, Hutschenreuther et al., 2012b, Lorenz et al., 2011) with little modification as shown in figure 3.1e. Briefly, after removal of culture medium and two washings with PBS, cells were rapidly quenched by directly adding ~15 mL liquid nitrogen (LN_2) to the T75 flask. The adherent cells in the T75 flask were then immediately subjected to extraction by addition of 750 µL of pre-chilled 100% methanol (-40°C) followed by lifting with a pre-chilled cell lifter, all these steps were performed on dry ice. The cells were re-extracted and further steps were performed as described in the modified cell scraping method.

3.2.5 Metabolite Derivatization

Metabolite derivatization was performed as described elsewhere (Winder et al., 2008). Briefly, to the lyophilized extract, 40 µL of 20mg/mL methoxyamine hydrochloride in pyridine was added and samples were shaken for 80 min at 37°C. The samples were then derivatized by trimethylsilylation of acidic protons by addition of 40 µL MSTFA (N-methyl-N-trimethylsilyltrifluoroacetamide) with further incubations in shaking conditions at 40°C for 80 min. A retention index solution was added for the chromatographic alignment prior to analysis by GC-MS.

3.2.6 GC-MS analysis

Metabolite data was acquired on a Thermo Finnigan TRACE DSQ GC-MS System (Thermo Scientific, Hertfordshire, UK) operating in EI mode onto a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). The derivatized sample volume of 1 µL was injected in split less mode at 230°C, with the transfer line temperature was maintained at 250°C. The GC was operated at a constant flow of 1 mL/min helium. The temperature program was started at 80°C for 6 min, followed by temperature ramping at 6°C/min to final temp of 290°C and held constant at 310°C for 5 min. Data acquisition was performed on a DSQ MS system with a mass range of 50 to 650.

3.2.7 Metabolite identification

The metabolites were identified as TMSi derivatives by comparing their mass spectral and RI index with online databases (The GOLM Metabolome database: http://csbdb.mpimp-golm.mpg.de/ and NIST 05 database). The acquired spectra were deconvolulated by AMDIS (Automated Mass Spectral Deconvolution and Identification System), before comparing with the database. Spectra of individual components were further transferred to the NIST mass spectral search system and matched with NIST main library, RI index library and the GMD (GOLM metabolome database).

3.2.8 Data analysis

GC-MS analysis yield complex data sets (time x mass x intensity) which require deconvolution, as fragment ions may be shared between two co-eluting compounds. All GC-MS chromatograms were processed using freely available AMDIS 2.70 software. The peaks were deconvoluted and the retention indices (RIs) were automatically calculated according to the retention time of the alkane mixture by exporting the RI calibration file into AMDIS. AMDIS deconvolution parameters used are as follows: resolution was set to high, sensitivity was high, shape requirement was medium, and component width was at 12 (Validated with 12 metabolite standard mixture). For identification, the minimum match factor was kept at 60, resolution: high;

sensitivity: high; shape requirement: medium. Finally, a report was generated in *.xls format and the first hit considered. Compounds found at least in two out of three biological replicates were considered true hits. Data for retention time, S/N ration, peak tailing, m/z value and peak area was collected manually by exporting to MS Excel 2013. MATLAB 7.0 (MathWorks, Natick, MA, USA) with in house routines employed for analysing the data containing metabolite identities. For intracellular metabolites the GC-MS data was normalised to viable cell count and sum of peak areas as suggested elsewhere (Hutschenreuther et al., 2012b).

3.3 Results and Discussion

A wide range of sampling and quenching protocols are reported in the literature for profiling metabolomes in mammalian cell cultures. We directed our efforts towards evaluating and minimising the leakage of intracellular metabolites during different sampling and quenching protocols. We adopted previously published quenching and extraction solvents for sampling of adherently grown mammalian cells with a few modifications. For all the sampling methods, except for direct quenching with LN_2, cells were quenched by 60% aqueous methanol supplemented with 0.85% AMBIC at - 40^0 C to minimise the leakage of intracellular metabolites (Sellick et al., 2008), followed by an extraction with pure methanol (Winder et al., 2008). However, when the study objective is to maximise metabolome coverage, characterisation of leakage with respect to metabolite class is required. For this, it is essential to first identify and classify metabolites detected based on their physicochemical properties, thus generating a metabolite matrix that can be utilized for interpretations. Hence, all the identified metabolites with the different sample treatment approaches were initially classified based on their physicochemical properties. Although, there are issues with metabolite identifications in GC-MS, changes in metabolite classes within a sample set, provide a reasonable measure to compare and assess.

The first objective of the investigation was to assess if the cell lines show differences in leakage of intracellular metabolites when conventional methods of trypsinization and cell scraping are employed, during quenching of adherently grown mammalian cells. This was carried out on three different cell lines, namely, two breast cancer cell lines MDA-MB-436 and MCF-7, and the human microvascular endothelial cell line HMEC-1. Subsequently, recently suggested optimized methods that involve direct scraping (Bi et al., 2013, Dettmer et al., 2011, Hutschenreuther et al., 2012b, Lorenz et al., 2011, Martineau et al., 2011, Teng et al., 2009) were assessed as alternative approaches after appropriate modifications, and compared with the two conventional approaches, with respect to evidence for metabolite leakage. In addition, we examined the influence of acetone precipitation on leakage and changes in metabolite classes captured by the GC-MS approach adopted. Further, the use of glass beads to cultivate and rapidly harvest metabolomes as a method of minimising metabolite leakage, was also examined. These latter investigations were performed only on the breast cancer cell line, MDA-MB-436.

3.3.1 Trypsinization versus Cell scraping

In order to compare the degree to which different cell lines are influenced by the conventional cell detachment methods, MDA-MB-436, MCF-7 and HMEC-1 cells were either detached by trypsinization or by scraping. Adherent cells attach to the surface of tissue culture flasks or dishes using secreted proteins to form a tight bridge between the cell and the surface. This needs to be disrupted to lift the cells off the surface and extract the metabolites within. This is traditionally achieved biochemically with the help of a protease, typically trypsin, to cleave the protein bridges and release the cells for harvest and analysis. Alternatively, the cells can be mechanically lifted off the surface by scrapping. Both methods have the potential (have been shown) to introduce artefacts (Dettmer et al., 2011, Hutschenreuther et al., 2012b, Lorenz et al., 2011, Teng et al., 2009).

For an overall comparison, the harvested cell extracts, cell-free supernatant of quenched cells and the blank sample (culture medium) were analysed to determine the extent of leakage of intracellular metabolites during quenching. After monitoring cell-free supernatant of quenched cells and the blank medium, the necessary correction was done for appropriate calculation of intracellular metabolites. Only features that were present in at least two biological replicates out of three were considered for further analysis. A summary of the unique recovery efficiency of both methods for three different cell lines is shown in figure 3.2.

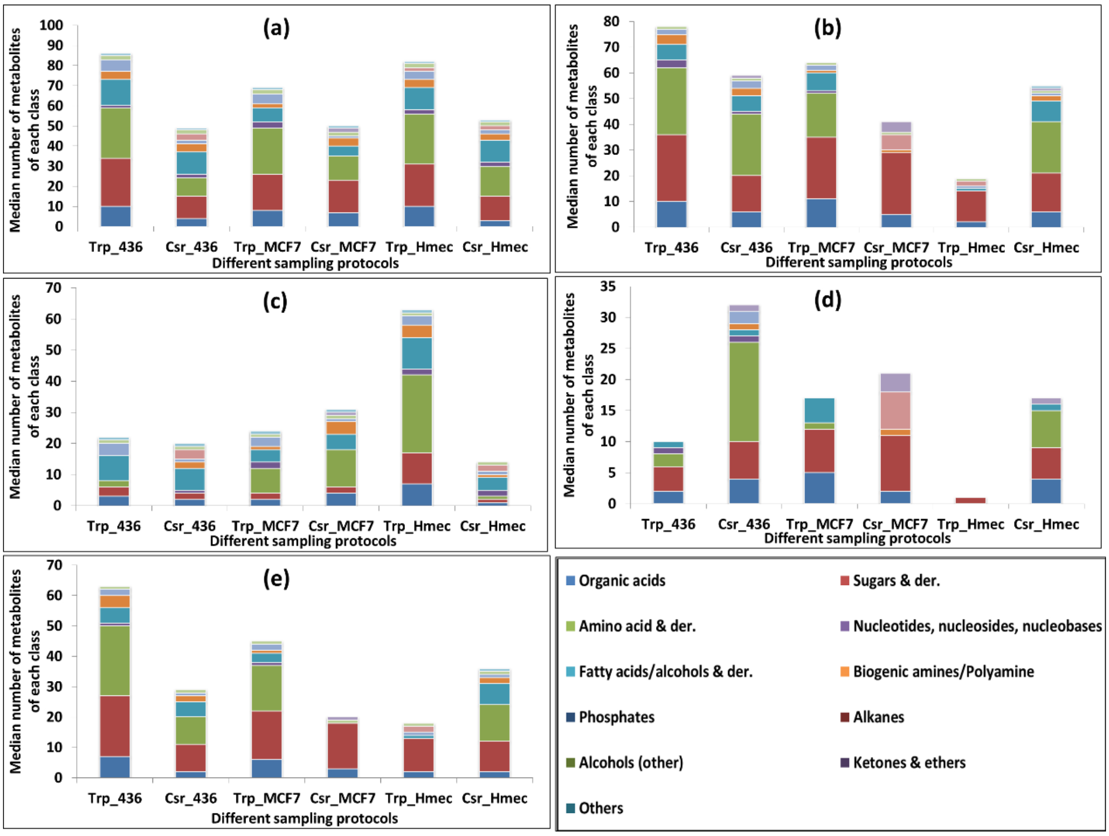


Figure 3.2 Effect of trypsinization and cell scraping treatment on leakage of intracellular metabolites during quenching was compared for three different cells lines, MDA-MB-436, MCF-7 and HMEC-1. X-axis represents different sampling protocols; Trp_436= trypsinization on MDA-MB-436; Csr_436 = cell scraping on MDA-MB-436; Trp_MCF7 = trypsinization on MCF-7; Csr_MCF7 = cell scraping on MCF7; Trp_Hmec = trypsinization on HMEC-1 and Csr_Hmec = cell scraping on HMEC-1. Y-axis represents median number of metabolites of each class. After both the treatments the extracted metabolites from cell extracts, cell-free supernatant post quenching and blank samples were analysed by GC-MS. (a) metabolites identified in cell extracts only, (b) metabolites identified in supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both).

In total, after correction of leaked metabolites, 129, 116 and 119 unique metabolites were identified with the trypsinization treatment in MDA-MB-436, MCF-7 and HMEC-1 respectively, whereas with the cell scraping treatment, 115,112 and 103 unique metabolites were identified in MDA-MB-436, MCF-7 and HMEC-1 respectively. Metabolite identification and data analysis was performed as detailed in section 3.2.7 and 3.2.8. In contrast to a previously published report (Dettmer et al., 2011), for all the three cell lines we have found higher TIC for trypsinization treatment as compared to cell scraping, which is in agreement with the recently published report (Hutschenreuther et al., 2012b). The leakage of metabolites during sampling and quenching method is suggested to be highly dependent on the cell membrane structure (Dettmer et al., 2011, Hutschenreuther et al., 2012b, Winder et al., 2008). There is, however, a lack of experimental data to show the degree to which the specific treatments affect metabolite leakage, therefore this was assessed in three different cell lines.

The results of the investigation are summarised in Figure 3.2, where the metabolite class and numbers detected for each of the cell lines and treatment are plotted. Figure 3.2a summarises the metabolites detected in the cell pellets, Figure 3.2b those detected in the cell-free supernatant, Figure 3.2c indicates the metabolites present in the cells but not in the supernatants, Figure 3.2d indicates metabolites present in the supernatants and not in the cells, and Figure 3.2e metabolites present in both the cell pellet and the supernatant. Metabolites detected in the supernatant are from: a) the culture medium component, b) intracellular metabolites actively secreted into the medium through overflow metabolism, and/or c) intracellular metabolites leaked into the medium as sample treatment artefact. High metabolite numbers in the supernatant (Figure 3.2b), relatively high numbers detected in both the cells and the supernatants (Figure 3.2e), and corresponding low numbers unique to the cell pellets (Figure 3.2c) indicates high metabolite leakage. Higher proportion of metabolites detected in Figure 3.2e, compared to that detected in Figure 3.2c or Figure 3.2d indicates that there is an increased chance of metabolite leakage.

As can be seen from Figure 3.2, for the trypsinization treatment of HMEC-1 cells (Trp_Hmec), metabolites unique to the cells are the highest (77% of cell extract) (Figure 3.2c), and correspondingly, metabolites common to both the cells and supernatant (Figure 3.2e) are the lowest (22% of cell extract). The number of metabolites detected in the supernatant is also the least for all cell lines and treatments. This suggests minimum leakage from cells into the supernatant following trypsinization, whereas cell scraping for this cell line results in 26% and 68% leakage respectively, suggesting significantly greater metabolite leakage. Compared to the breast cell lines, trypsinization of HMEC shows minimum metabolite leakage. For the other two cell lines, both trypsinization and cell-scraping protocols adopted result in metabolite leakage, but this is greater (by 2 fold) for trypsinization than for cell-scraping, as indicated by the proportion of metabolites detected that are unique to the cells (Figure 3.2c as a proportion of Figure 3.2a). This suggests that cell lines are differentially influenced by trypsinization and cell-scraping, and that, dependent on the specific cell line, the ideal approach requires optimisation, to minimise metabolite leakage. A PCA analysis revealed that, the reproducibility of metabolome recovery for biological replicates was better with the cell scraping treatment compared to the trypsinization protocol adopted, for all three cell lines.

Analysis of these “method dependent” groups revealed major differences in five major classification groups as shown in figure 3.2. In the case of MDA-MB-436 & HMEC-1, trypsinization yielded a larger proportion of organic acids, amino acids, sugars and phosphates, whereas in contrast to this, for MCF-7 cells there was negligible variance between the two treatments in the recovery of organic acids & sugars [Figure 3.2 (a)], whilst the scraped samples comprised greater amino acids and amines than trypsinized samples [Figure 3.2 (c)]. However, the extent of leakage of organic acids, amino acids and sugars was higher following scraping compared to trypsinization for both MDA-MB-436 & HMEC-1. In contrast MCF-7 cells, showed severe leakage of alkanes, ketones & ethers on scrapping whereas trypsinization resulted in loss of organic acids & fatty acids [Figure 3.2 (d)].

In conclusion, the three cell lines investigated appear to be differently influenced by the two conventional methods, with respect to metabolite leakage, both in terms of numbers and the metabolite class detected. Trypsinized cells exhibit a chemical deviation from cells in their attached state and the effective time for sufficient detachment of cells by trypsin treatment may lead to significant metabolite leakage (Teng et al. 2009). The cell-scraping protocol is subjective, in the sense that it might be difficult to reproduce between experiments, although there was good reproducibility in our hands. However as shown in this investigation, it is important to note that the leakage of metabolites during trypsinization and scraping treatments is highly dependent on the cell membrane structure which varies greatly across the different cell lines and can significantly affect recovery of different metabolite classes.

3.3.2 Modified approaches

From our data we have shown that the conventional methods of sample preparation result in leakage of intracellular metabolites, hence there is a need to address this problem in the experimental protocol and preparation for the metabolome analyses of adherently grown mammalian cells. Therefore, we have investigated a further 3 modified approaches as comparisons to the conventional approaches of trypsinization and cell scraping in a single cell line MDA-MB 436 cell line. In addition, the effect of protein precipitation using acetone was also investigated. The data was analysed as described previously, and is presented Figure 3.3.

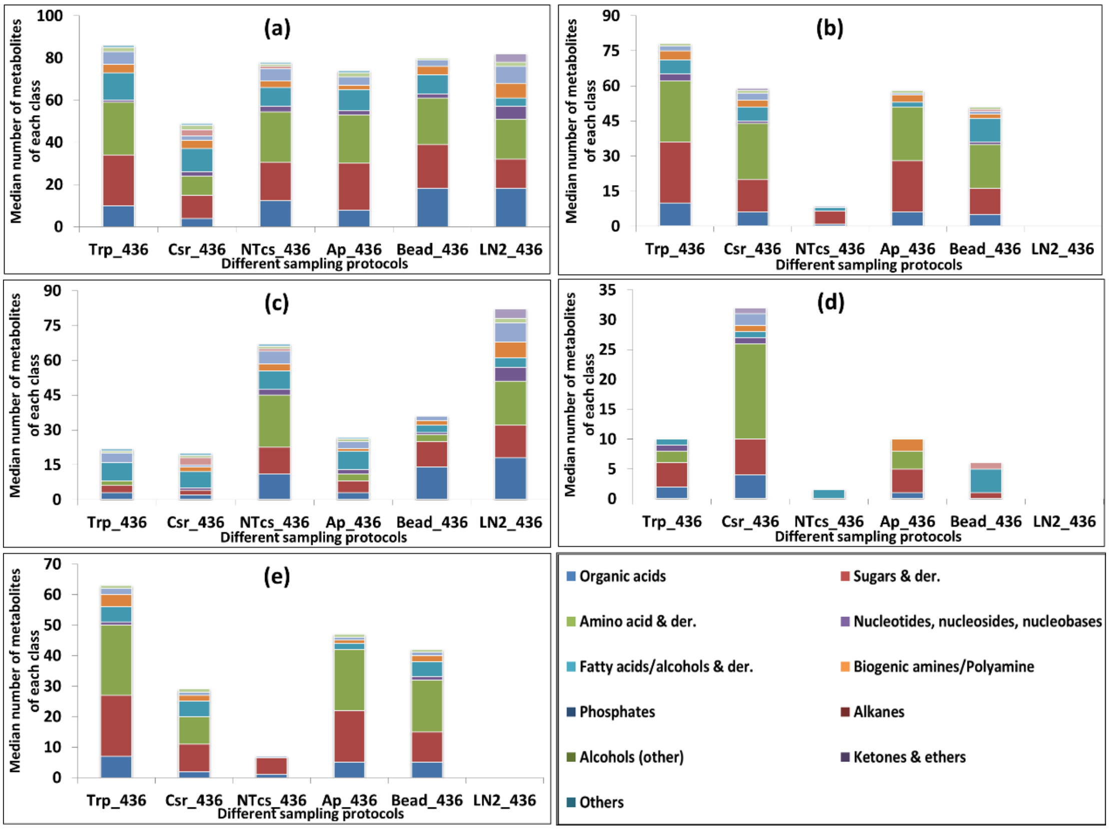


Figure 3.3 Effect of the sampling protocols on leakage of intracellular metabolites during quenching for MDA-MB-436.After all the treatment the extracted metabolites from cell extract, supernatant following quenching and blank samples were analysed by GC-MS. X-axis represents different sampling protocols; Trp_436 = trypsinization ; Csr_436 = cell scraping; NTcs = modified cell scraping; Ap_436 = trypsinization with acetone precipitation; Bead_436 = bead harvesting and LN_2 = direct quenching and extraction with liquid nitrogen. Y-axis represents median number of metabolites of each class; (a) metabolites identified in cell extracts only, (b) metabolites identified in cell-free supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both).

3.3.3 Modified direct cell scraping (NTcs)

We investigated the recently suggested direct scraping method (Dettmer et al., 2011, Teng et al., 2009), with a modification. Instead of combined cell harvesting and extraction, MDA-MB-436 cells were quenched and extracted separately for assessing the leakage of intracellular metabolites and for comparison with the other applied treatments (Figure 3.3c).

As can be seen in figure 3.3 (c), and (e), the highest recoveries of metabolites unique to the cell extracts and minimal proportion common to the cells and supernatants was determined for this modified protocol, suggesting minimal leakage over all the applied protocols. The exception to this is with LN_2 treatment where the supernatants could not be collected and analysed, as the LN_2 evaporates.

As previously reported (Hutschenreuther et al., 2012b), the combined cell harvesting and extraction technique does not allow determination of cell number, as cell detachment is a prerequisite to cell counting. The protein content of the cell pellet can be determined for normalisation (Dettmer et al. 2011), however, a recently published report (Hutschenreuther et al., 2012b) suggested normalization to the sum of peak areas, as a superior approach, since normalisation to cellular protein content does not allow for adjustment of extract concentration prior to metabolite extraction. In the method employed in the current studies, conventional cell count for normalisation of data can be achieved by parallel cultivation of flasks, as quenching and extraction are performed separately. In addition to this, the modified method offers an advantage for selection of appropriate quenching and extraction solvents, depending upon the cell type.

3.3.4 Bead harvesting

Here, we investigated a novel approach of cell culture in beads for quenching and harvesting with minimal time lag. The cells were cultured on beads in a 50 mL Falcon tube (kept in a fixed horizontal position), and harvested as described in the modified direct cell scraping method. The motivation behind culturing the cells on beads and in a flacon tube originated in a quest to find a suitable means to culture and harvest the cells from a platform that aids rapid quenching and extraction of metabolites from cells. We have adopted two extraction techniques together, bead beating followed by freeze-thaw cycle to achieve higher extraction efficiency, as it was much easier to apply in this approach, as cells were cultured in a Falcon tube and on beads, which was not possible to achieve with the conventional T75 flasks.

The quantitative data is shown in figure 3.3. The overall recovery of intracellular metabolites was improved compared to cell scraping, and similar to other sampling and quenching techniques applied, as can be seen in figure 3.3 (a). With respect to the unique recoveries of metabolites in cell extracts (Figure 3.3 (c)), overall this approach proved superior to deproteination (Ap_436) and the conventional methods for recovery of organic acids and sugars. A similar trend was observed with leakage analysis, with this approach superior to deproteination and conventional methods, with fewer metabolites in supernatants as shown in figure 3.3 (d). However yields of fatty acids were much reduced compared to all other treatments applied as can be seen in figure 3.3 (c). Further investigations on the composition of the cells grown on beads will be required to interpret this observation. However, this is an appropriate method to use if the focus is on non-lipid metabolites.

3.3.5 Direct quenching using LN_2

The recently described use of LN_2 (Bi et al., 2013, Hutschenreuther et al., 2012b, Lorenz et al., 2011) for direct quenching and storage of cells in -80 °C for later extraction was also investigated in this study, and compared with the other sampling and quenching protocols. The overall recovery of metabolites with this approach (LN_2_436 in Figure 3.3 (a)) was similar to that of bead harvesting and modified direct cell scraping treatment. However, compared to the other modified approaches nucleotides/nucleosides/nucleobases, biogenic amines/polyamines and phosphates demonstrated increased recovery with the LN_2 approach, whereas recovery of fatty acids/alcohols appeared compromised. Recoveries of organic acids was reasonable, but that of amino acids and derivatives were less than that of the modified direct cell scraping (NTcs_436) method.

3.3.6 Protein precipitation using acetone (AP)

The presence of proteins in metabolome extracts of adherently grown mammalian cell is likely to interfere with the composition of the metabolite pool, with a proportion of these are enzymes, potentially affecting the metabolite pool post-extraction, changing composition. In addition, they could interfere with the chromatographic separation of the metabolites. As demonstrated in the past with LC-MS data (Bruce et al., 2009), efficient removal of proteins before injection onto an analytical chromatographic column is important to help in ensuring high-quality data (efficient extraction of metabolites, reproducibility) and also to preserve the lifetime of the column. In the case of serum samples, the use of organic solvents such as acetone for deproteination of metabolome extracts has been regarded as an effective method in reports for the removal of detergents, lipids, small molecular-mass nucleic acids, and other contaminant species, while providing high metabolite recoveries, most reliable data and also minimal post-preparation problems such as ion source contamination(Aranibar et al., 2011, Bruce et al., 2009, Duan et al., 2009, Tiziani et al., 2008). We were therefore interested in examining if a protein precipitation step prior to metabolome analysis enabled improved recoveries or perceivable changes in composition. Accordingly, cells were harvested by trypsinization and deproteinized using cold acetone, prior to derivatization.

As shown in figure 3.3 (a), surprisingly the overall recovery of metabolites appears to be slightly reduced with the deproteination treatment as compared to the trypsinization protocol. With respect to the unique recoveries, deproteination treatment shows a slightly higher unique cell recoveries for sugars and nucleotides (figure 3.3 (c)), as compared to the trypsinization protocol.

However, we did not see any significant improvements following deproteination treatment, with the cell line investigated. It is possible that co-precipitation of some metabolites along with the precipitation of the macromolecules occurred, suggesting careful optimization is required to prevent artificially introduced metabolome compositional changes. Overall, acetone precipitation of proteins does not dramatically alter the pattern of different classes of metabolites detected in MDA-MB 436, as compared with the other treatments.

3.3.7 Comparison of conventional and modified methods

In order to understand better the coverage of different classes of metabolites in cell extracts and the extent of leakage during quenching by each of the applied approaches, we assessed and compared all applied sampling protocols for the recovery of eleven different classes of metabolites identified by GC-MS in the cell extracts and the cell free supernatants post quenching. Metabolites from each of the applied sampling protocols were identified using the NIST 05 and GOLM metabolome database, which were then classified into eleven different classes of metabolites, based on their physicochemical properties. Further, the unique metabolite recoveries obtained from the cell extract and the extent of metabolite leakage in supernatants following quenching for each of the applied sampling protocol was assessed and then used to generate a heat map, as displayed in figure 3.4. Each band represents the median metabolite levels from three independent determinations, normalised for each condition and expressed on a logarithmic scale to capture the variation that was over orders of magnitude.

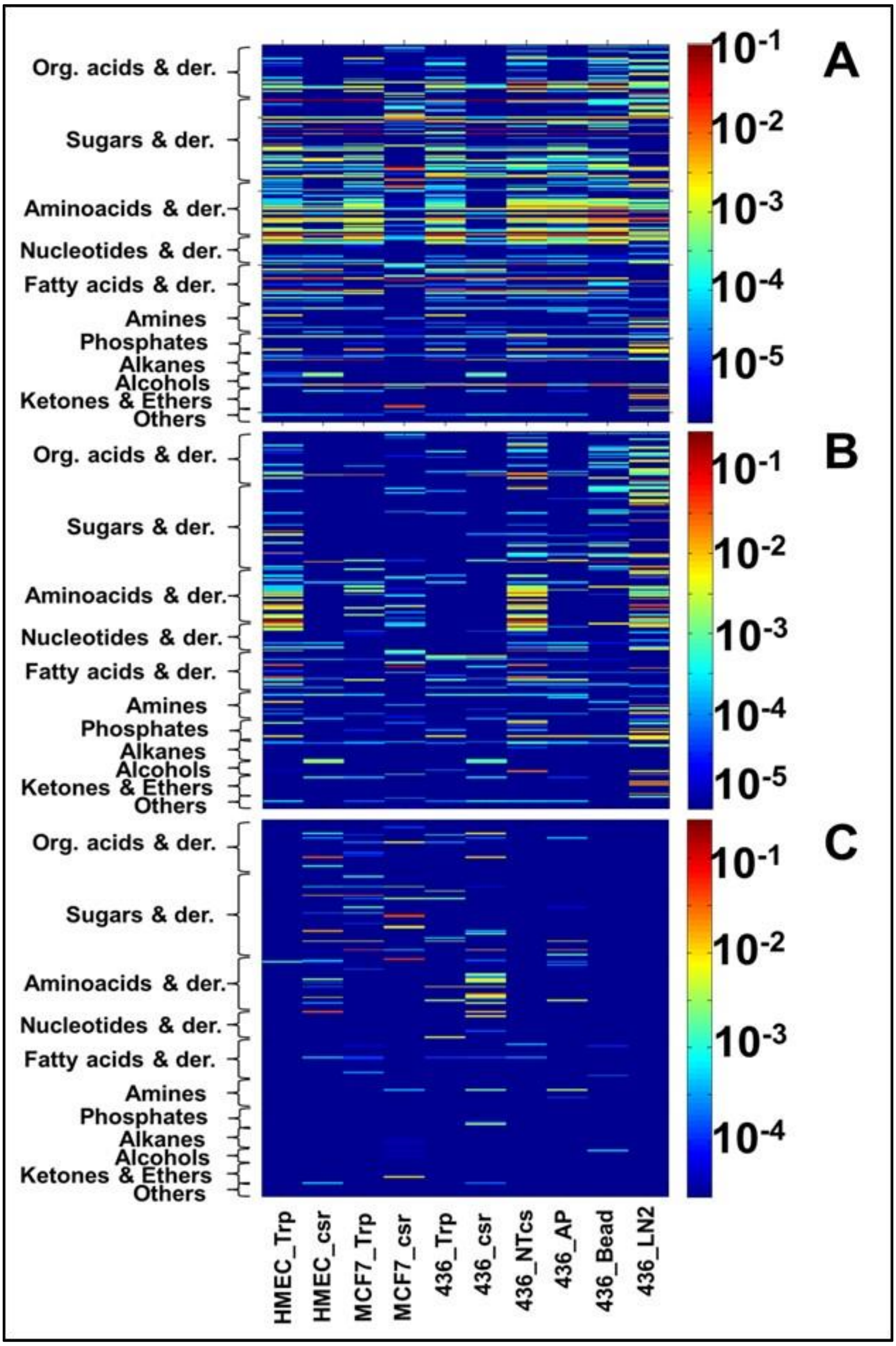


Figure 3.4 Heat maps comparing the response of 11 different classes of metabolite signals identified by GC-MS by each of the applied sampling and harvesting approaches across three different cell lines. A = All metabolites in cell extracts; B = Metabolites present only in cell extracts (unique to cells); C = Metabolites present only in supernatants (unique to supernatants).

As displayed in figure 3.4A, the recovery of all classes of metabolites was significantly improved with the modified direct cell scraping, direct LN_2 quenching and bead harvesting methods, compared to the conventional methods. The recovery of amino acids and its derivatives with the modified direct scraping treatment was found to be superior over all the compared treatments, as no leakage was observed in the supernatant with the modified direct scraping treatment. Whereas recovery of sugars/sugar alcohols and its derivatives was found to be superior with the bead harvesting and deproteination treatment. The direct quenching with LN_2 resulted in superior recovery of nucleotides/nucleosides/nucleobases, biogenic amines/polyamines and phosphates whereas recovery of fatty acids/alcohols and its derivatives was severely impeded as compared to modified direct scraping, bead harvesting and deproteination treatments. With respect to the recoveries of all metabolites classes unique to the cells (figure 3.4B), there was an approximate threefold increase in the recoveries with the modified cell scraping method compared to the conventional method, whereas a two fold increase was noted with the bead harvest method when compared to the conventional cell scraping. However, LN_2 treatment could not be analysed due to nature of the protocol, henceforth displaying highest recoveries in figure 3.4B and no intracellular leakage in figure 3.4C.

The heat map in figure 3.4C shows that the recovery of metabolites unique to the supernatant for all three cell lines was higher for cells harvested by conventional cell scraping method indicating higher metabolite leakage during treatment particularly with amino acids. This finding was in contrast to the previously published report, where the author reported higher leakage of amino acids with the trypsinization treatment (Dettmer et al., 2011). However, with the modified cell scraping and bead harvest method, the extent of metabolite leakage was significantly reduced, except for leakage of few fatty acids/alcohols.

3.4 Conclusions

In this study, we have initially compared the conventional methods of handling adherent mammalian cells across three different cell lines and successfully demonstrated that the leakage of intracellular metabolites is highly dependent on cell type and the sampling and quenching protocol used and it significantly affects recovery of different classes of metabolites. Moreover, trypsinized cells exhibit a chemical deviation from cells in their attached state and that the effective time for sufficient detachment of adherent cells by trypsin treatment. In conclusion, it is important to assess the treatment methodology for each cell type and optimise it to minimise metabolome leakage. Inclusion of a step to precipitate proteins using acetone in the trypsinization protocol, prior to metabolite profiling in MDA-MB-436 showed some changes in the metabolite levels detected, suggesting that inclusion of this step in the protocol needs to be assessed for the specific cell type before being included.

The direct quenching and extraction techniques yielded higher recovery of all metabolite classes, approximately three fold higher than the conventional sampling techniques. However comparison within them, showed little variations with respect to number of metabolites recovered, although recovery of different classes of metabolites appears to be altered significantly across different simultaneous quenching and extraction treatments. In this investigation we have introduced bead harvesting, which is rapid, effective and an efficient approach for recovery of different classes of metabolites, and does compare favourably with the conventional approaches. The modified direct cell scraping & LN_2 treatments provide a rapid and reliable route to quenching and extraction, and could be combined with bead harvesting to minimise artefacts due to metabolome leakage. Whichever approach is chosen, it is imperative that the option is first characterised for metabolome leakage with the particular cell type being investigated before it is adopted. Moreover, when particular class of metabolites are of interest, it is necessary to carefully consider the impact of different treatments on recovery of different classes of metabolites.

Chapter 4

Effect of washing and quenching in metabolomics of adherently grown mammalian cells: A case study on the metastatic cancer cell line MDA-MB-231

4. Effect of washing and quenching in metabolomics of adherently grown mammalian cells: A case study on the metastatic cancer cell line MDA-MB-231

4.1 Introduction

In Chapter 3, we demonstrated that selection of the sampling methods largely depend on the type of cells, culture format (adherent or non-adherent) and analytes of interest (exo-metabolome or endo-metabolome), similarly different washing steps/solution and various quenching additives might influence the analysis. In the case of suspension cultures, cells are separated from the media by means of filtration or centrifugation followed by quenching and washing steps. Whereas in the case of adherent cultures, the media can be simply removed by suction followed by a washing step. However, cells need to be detached from cell surface prior to quenching. In metabolomics, potential problems connected to sampling have not been considered properly, which often leads to enormous errors, as they are normally adopted from the literature without due consideration and critical evaluation for the given case and the investigated organism as noted by several investigators (Canelas et al., 2008, Sellick et al., 2008, Wellerdiek et al., 2009a).

As discussed in Chapter 2, despite contradicting reports, sampling methods are rigorously studied in case of prokaryotes which cannot be simply adopted for mammalian cells due to basic differences in the cell structure. In case of suspension mammalian cultures (CHO cells), methanol/water (buffered/un-buffered) was compared with cold isotonic saline (0.9% w/v NaCl) (Dietmair et al., 2010), and it was shown that quenching with isotonic saline did not damage cells and resulted in proper metabolic arrest, as it effectively halted the conversion of ATP to ADP and AMP. However in contrast, no improvement in metabolite recoveries was reported with the use of 0.9% NaCl compared to that of 60% aqueous methanol (Sellick et al., 2008) and various authors have suggested the use of methanol supplemented AMBIC. With this solution the membrane integrity preserved compared to other additives, thereby minimising the leakage of intracellular ATP. A fluorescent marker such as green fluorescence protein (GFP) has been used as a visualisation marker by Wurm and co-workers to estimate the rate of metabolite efflux calculated by numerical modelling from CHO cells upon membrane damage (Wurm and Zeng). Authors reported 90% of a small metabolites would be lost within ≤ 1 s when 5% of membrane damage is caused by washing, quenching solutions or harsh sampling techniques. This clearly underlines the requirement of rapid sampling methods with suitable quenching or washing solutions in order to avoid metabolite leakage.

Rigorous studies demonstrating effective washing and quenching methods for adherent cell cultures have been few and far between. Moreover composition of media used for mammalian cultures are very rich in nutrients such as amino acids, vitamins, sugars etc. which are also present inside the cells as an intracellular metabolites. Hence subsequent extraction of cells followed by quenching, without any washing step will result in false higher recoveries for specific intracellular metabolites that are found to be present in the medium as well. Moreover these media carryovers might also interfere in the post extraction analytical protocols. Inclusion of a washing step for adherent cultures would be advantageous as it can be performed rapidly prior to quenching, and does not prolong the quenching time frame and thus might improve the validity of the intracellular metabolite measurements. However, this needs to be validated to see if the use of washing solution and a number of washing steps employed influences the intracellular metabolite leakage from adherent cultures? To our knowledge, no quantitative data on effectiveness of various washing step/solutions and quenching solvents (buffered or non-buffered) on adherent cultures is available.

Hence, one objective of this research is to determine quantitatively, whether leakage occurs in adherent cultures during the quenching step with the variety of quenching solutions employed. In case of intracellular leakage, the research will be directed towards determination of the metabolite levels in the fractions in order to establish mass balances and to trace the fate of metabolites. In addition, the effect of other factors which includes: the properties of the quenching solution with various buffer additives, the number of washing steps required, interferences with the analytical techniques (GC-MS) and validity check of data will be investigated. To achieve our objectives, we have employed an approach which involves initial quantification of intracellular ATP levels from adherent metastatic MDA-MB-231 cell line using bioluminescence assay. The concentration of ATP was determined after the various washing steps/solutions and normalised to total cellular protein content determined by BCA assay. In addition, the effect of various quenching solutions (buffered or non-buffered) on MDA-MB-231 cell membrane integrity was visualised using scanning electron microscopy (SEM). Furthermore, we investigated the overall recovery of eleven different metabolite classes using GC-MS technique and compared the results with those obtained from the ATP assay and SEM analysis.

4.2 Material and Methods

4.2.1 Chemicals and reagents

The RPMI-1640 medium was obtained from Lonza (Gibco-BRL, Paisley, UK). All other reagents and consumables were obtained from Sigma-Aldrich (Dorset, U.K.), unless stated otherwise.

4.2.2 Cell lines, cell culture and growth assessment

The cell lines where maintained, cultured and provided by the Microcirculation research group led by Prof. Nicola Brown at the Medical School. The MDA-MB-231 epithelial breast cancer cell line was obtained from American Type Culture Collection (ATCC), (http://www.atcc.org/). Cell line was cultured at 37˚C with 5% CO_2 in 100mm Nucleon dish (Nucleon®, Thermo Scientific) with approximately $2x10^5$ cells per dish containing 10 ml RPMI media which contains 10% fetal bovine serum (FBS), 1% penicillin/streptomycin (GIBCO®) and 1% glutamine. Growth curves were produced by seeding cells in a 24 well plates with 1 mL media and seeding density of $0.1x10^6$ cells per plate. Viable cell counts are obtained at 24, 36, 48 and 72 hours using a Beckman coulter Vi-Cell XR cell viability analyser (Beckman Coulter, Germany) after trypsinization. Control flasks were performed but omitted seeding cells to determine any unwanted background signal.

4.2.3 ATP assay

For all ATP assay, all assay vials, glassware and pipette tips were soaked in 1N HCl overnight followed by washing 3 times with ultrapure water and then dried in an oven at 40°C for 1 hour. MDA-MB-231 cells were seeded in triplicate (5×10^4 per well) in 24 well plate (Corstar®). Cells were incubated in 500 µL of RPMI 1640 media containing 1% fungazone, 10% fetal bovine serum (FBS), 1% penicillin/streptomycin (GIBCO®) and 1% glutamine at 37˚C with ±5% CO_2. After 36 hours of incubation, culture medium was removed from each well and cells were treated differently based on experimental design which include various washings and buffered quenching reagents including direct quenching with LN_2. The level of free ATP in response to various treatments indicating leakage, was determined using ATP bioluminescent somatic cell assay kit (FLASC) purchased from Sigma-Aldrich (Dorset, U.K.). The assay parameters are shown in Table 4.1. Briefly, 100 µL of ATP assay mix was added to the assay vials and incubated for at least 3minutes at room temperature in order to hydrolyse any endogenous ATP. Cells in each well were suspended in 150 µL of ATP releasing agent (ARR) which includes: *p*-tertiary-Octylphenoxy polyethyl alcohol and edetic acid. Finally 75 µL of sterile dH_2O was added to vial containing luciferase assay mix followed by addition of 75 µL of sample as shown in figure 4.1. The amount of light emitted from each reaction was measured immediately in luminescence units (RLU) using microplate luminometer (Centro LB 960, Berthold, Germany). The concentration of ATP in nM in each sample was determined from the log-log plot of eight or seven ATP standards (0 to 1.5 nM) against relative luminescence units (RLU). To normalise for differences in cell number between various treatments, the concentration levels of free ATP were corrected for the levels of protein (in micrograms) present in the same cell extracts prepared for the ATP assay. Determination of total protein content was described in detail in section 4.2.4. The nM amounts of ATP per mg of protein produced by MDA-MB-231 cells, after each treatment (n=3) was determined. Finally, the mean, standard deviation and standard error of the mean (SEM) values were determined for each treatment and levels of free ATP determined were reported in nM ATP per mg of protein.

Table 4.1 Assay parameters employed for the estimation of ATP content by luciferase bioluminescence assay in MDA-MB-231 cells. The amount of ATP determined was then normalised to protein content estimated by BCA microplate assay

Assay Parameters	
Scale	*24 well plate (Corstar®)*
Seeding density	*50000 cells per well (cpw)*
Incubation	*36hr - 500µl full * RPMI 1640 (37°C + 5% CO_2)*
Luciferase assay mix dilution factor	*1:25 (FLASC Kit dilution buffer)*
Amount of ATP releasing reagent (ARR)	*150 µl ATP Releasing Reagent, 100µl Luciferase*
ATP standard range (n=3) for preliminary experiments	*0, 0.2, 0.3, 0.44, 0.67, 1.0, 1.25, 1.5 (nM)*
ATP standard range (n=3) for other experiments.	*0, 0.44, 0.67, 1.0, 1.25, 1.5, 2 (nM)*
Base line ATP MDA-231	*1.2-1.3 nM*
Additional test std. dilution (0.87nM) (n=3)	*6.5µl (1.5nM ATP) + 1493µl dH_2O*
Protein quantification	*BCA Microplate assay (ab. 570nm) (Pierce)*
**Full = supplemented: 1% fungazone, 1% strep/P, 1% glutamine and 10% fetal bovine serum (FBS)*	

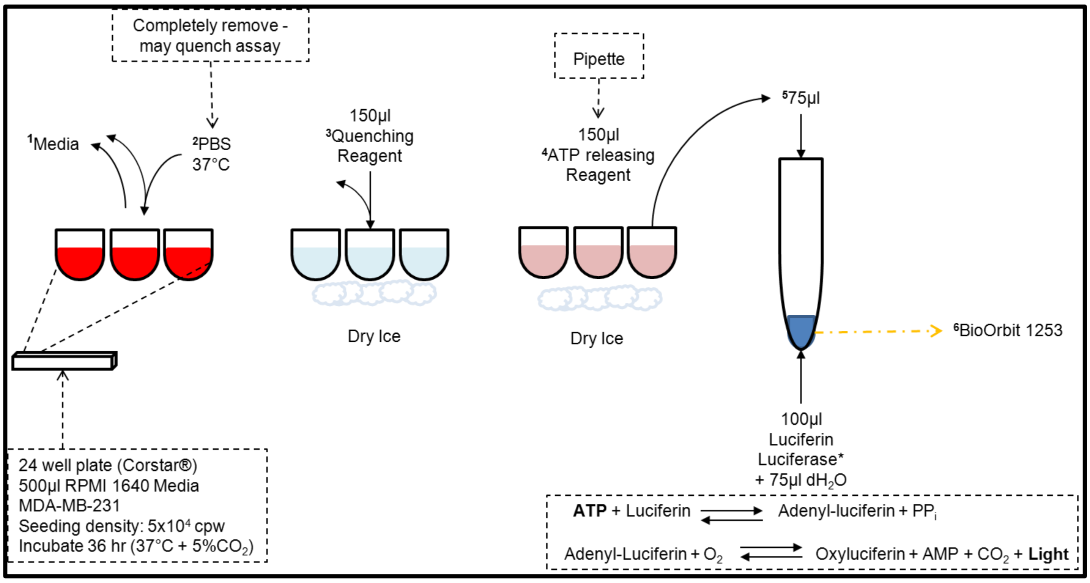

Completely remove - may quench assay
1Media
2PBS 37°C
150µl 3Quenching Reagent
Pipette
150µl 4ATP releasing Reagent
575µl
Dry Ice
Dry Ice
6BioOrbit 1253
100µl Luciferin Luciferase* + 75µl dH2O
24 well plate (Corstar®)
500µl RPMI 1640 Media
MDA-MB-231
Seeding density: 5x104 cpw
Incubate 36 hr (37°C + 5%CO2)
ATP + Luciferin ⇌ Adenyl-luciferin + PPi
Adenyl-Luciferin + O2 ⇌ Oxyluciferin + AMP + CO2 + Light

Figure 4.1 Schematic for determining level of free ATP in response to various washing and quenching processes using luciferase bioluminescence assay in adherently growing metastatic cells MDA-MB-231.

4.2.3.1 Preliminary ATP assay optimisation

As a part of optimisation and to evaluate successful application of ATP assays to determine the extent of leakage from adherently growing mammalian cell lines, preliminary ATP assay was performed on eight ATP standard solutions, additional test ATP standard solution and MDA-MB-231 cell extract (n=3). The eight ATP standard solutions were prepared with ultrapure water in different nM concentrations (0, 0.2, 0.3, 0.44, 0.67, 1, 1.25 and 1.5 nM). The standard solutions were prepared by making serial dilutions of the ATP standard stock solution (20 µM) supplied with the ATP bioluminescent somatic cell assay kit (FLASC). The level of free ATP in response to various ATP standard solutions, test solution and MDA-MB-231 cell extract sample was determined using ATP assay as described above.

4.2.3.2 ATP assay in response to various washing solutions and steps

After 36 hours of incubation, the culture medium was removed from each well and the cells were subjected to various ice-cold washing steps, which includes no washing (control), washing once with PBS (PB1), washing twice with PBS (PB2), washing once with water (W1) and washing twice with water (W2). The level of free ATP in response to various washing solvents and steps was determined using ATP assay as described above.

4.2.3.3 ATP assay in response to various quenching solvents

After 36 hours of incubation, the culture medium was removed from each well and the cells were washed with appropriate selected washing solution based on results of the above experiment (section 2.2.3.2). After washing step, the cells were rapidly quenched separately with five different quenching solvents (either buffered or non-buffered) along with the control sample. The various treatments includes no quenching (control), quenching with 100% methanol, 60% (*v/v*) aqueous methanol, 60% (*v/v*) aqueous methanol buffered with 0.85% ammonium bicarbonate (AMBIC), 60% (*v/v*) aqueous methanol buffered with 70 mM of 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethane sulfonic acid (HEPES) and direct quenching using LN_2. The level of free ATP in response to various quenching solvents was determined using the ATP assay described above.

4.2.4 Protein assay as a normalisation to ATP content

The nM amounts of free ATP produced in response to various washing solutions/steps and quenching reagents were normalised to the protein levels within each whole-cell lysate. The total protein content was determined by detergent-compatible bicinchoninic acid (BCA) protein assay using the MicroBCA Protein Assay Kit purchased from Thermo Scientific. Briefly, seven diluted albumin (BSA) standards in mg/mL (1, 0.8, 0.6, 0.4, 0.2, 0.1 and 0 mg/mL) were prepared initially from a BSA stock solution (2 mg/mL) using ARR as a diluent. Total volume of Micro BCA working reagent (WR) required was determined and WR was prepared according to manufacturer's instruction (Micro BCA Protein Assay Kit, Pierce/Thermo Scientific, Rockford, IL, USA) by mixing 25 parts of Micro BCA reagent MA and 24 parts of reagent MB with 1 part of reagent MC (25:24:1, reagent MA:MB:MC). Finally, 75 µL of cell extract (MDA-MB-231) was added to the 96 well plate (Corstar®) followed by addition of 75 µL of BCA WR. The components were gently mixed for 30 sec and incubated at 37°C for 2 hours wrapped in an aluminium foil. The plates were cooled to room temperature for 5 minutes after incubation and absorbance was measured at 570 nm on a plate reader (Centro LB 960, Berthold, Germany). The average absorbance reading of blank standard was subtracted from the absorbance reading of all individual standards. Average values of all replicates samples (blank-corrected) were determined and used to generate a BCA standard curve by plotting the mean standard absorbance values vs the respective concentrations. A best-fit polynomial equation was used for the standard curve and to estimate the amount of total protein content (mg).

4.2.5 Scanning electron microscopy (SEM) analysis

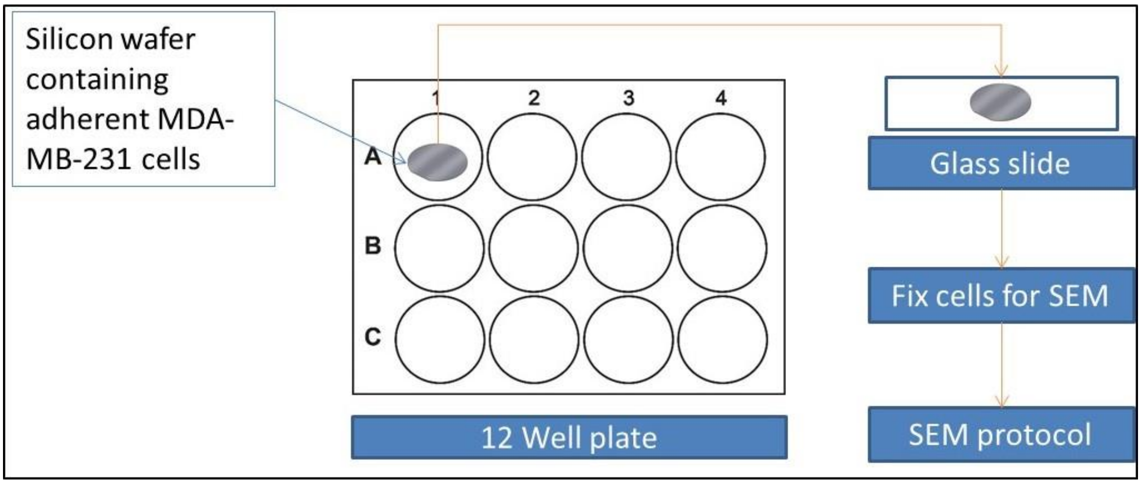


Figure 4.2 Schematic displaying workflow for SEM analysis of adherently growing cells. Adherent cells were grown on silicon wafers, which can be removed from the well plates by simply lifting the silicon wafers with forceps and cells along with the silicon wafers can be fixed for SEM analysis.

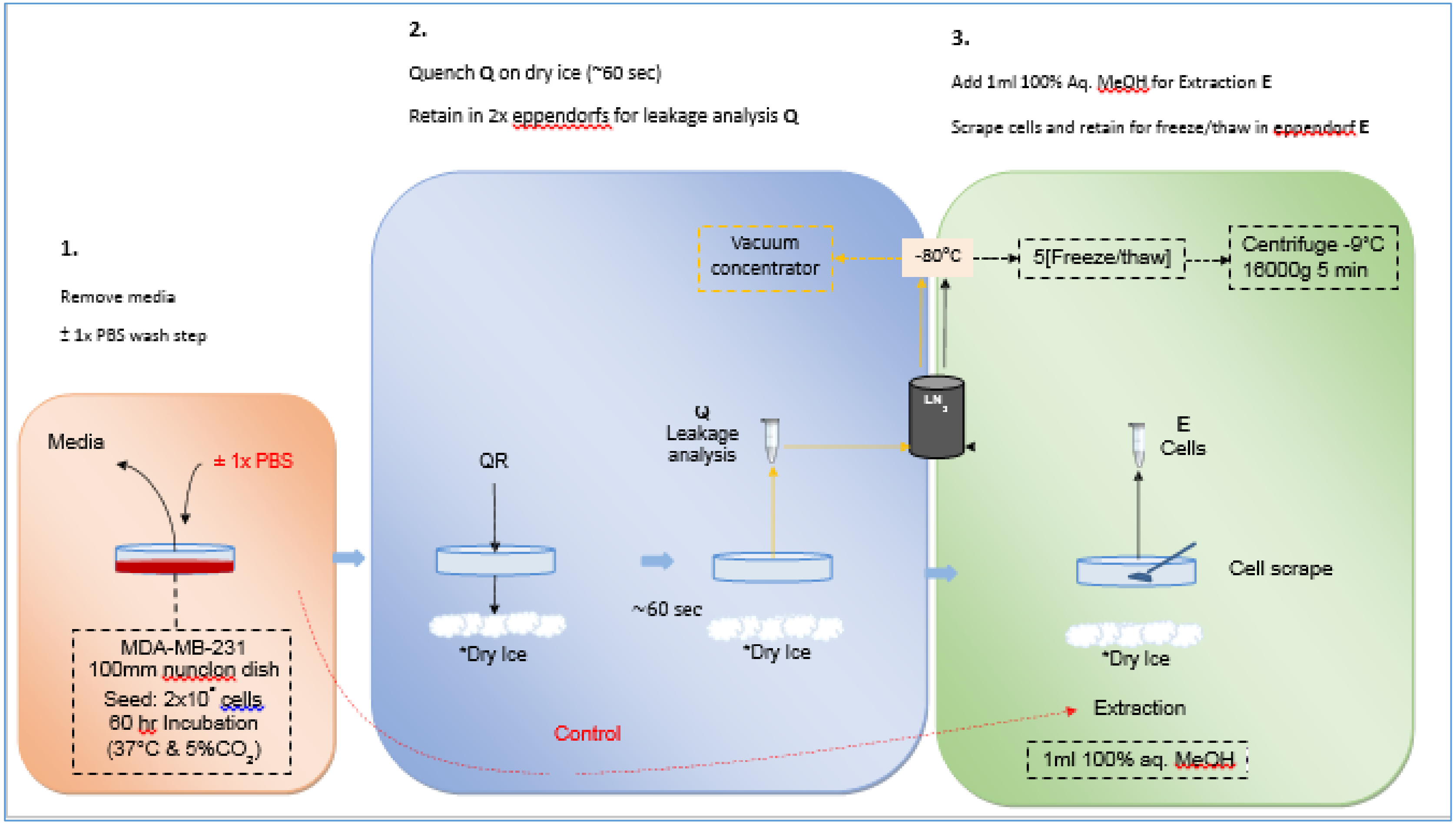
SEM analysis was performed on samples of control (non-quenched cells) and on MDA-MB-231 cells quenched with four different quenching reagents. The cells were grown in 12 well plate under similar conditions as described in section 4.2.2. However for ease of SEM analysis we used silicon wafers as shown in figure 4.2, for growing adherent mammalian cells. Briefly, the specimens were fixed in 3% glutaraldehyde in 0.1M phosphate buffer at 4°C. The specimens were then washed thrice in 0.1M phosphate buffer with 30 mins intervals at 4°C. Secondary fixation was carried out in 1% osmium tetroxide aqueous for 1 hour at room temperature. Dehydration was carried out through a graded series of ethanol concentrations in the following order: 75% ethanol for 15 mins, 95% ethanol for 15 mins, 100% ethanol for 15 mins, 100% ethanol for 15 mins, 100% ethanol dried over anhydrous copper sulphate for 15 mins. All the above dehydration steps were carried out at room temperature. Later the specimens were placed in a 50/50 mixture of 100% ethanol and 100% hexamethyldisilazane for 30 min followed by 30 minutes in 100% hexamethyldisilazane. The specimens were then allowed to air dry overnight at room temperature before mounting on aluminium stubs. Upon completion of drying, the specimens were mounted on 12.5 mm diameter stubs and attached with Carbon-Sticky Tabs and then coated in an Edwards S150B sputter coater with approximately 25nm of gold. The specimens were examined in a Philips XL-20 Scanning Electron Microscope at an accelerating voltage of 20Kv.

4.2.6 Cell quenching for metabolome analysis

The cells were grown to mid-“exponential” phase (48 hours) and then rapidly quenched before extraction. Initially, the culture medium was removed from the Nucleon dish and cells were quickly washed once with about 3 mL ice-cold phosphate-buffered saline (PBS, pH 7.4). The residual PBS was removed by suction. The cells were rapidly quenched with 3 mL of pre-chilled (-50°C) quenching solutions (buffered or non-buffered): 100% methanol alone, 60% aqueous methanol, 60% aqueous methanol supplemented with 0.85% (w/v) ammonium bicarbonate (AMBIC, pH 7.4) and 60% aqueous methanol supplemented with 70 mM HEPES. Addition of the cells to the quenching solution increased the temperature by no more than 15˚C. After 60 sec, the quenching solution was removed, and an aliquot (2 mL) was transferred to a 2 mL pre-chilled Eppendorf to assess the leakage of the internal metabolites. This step was performed on dry ice. An aliquot was then snap frozen in LN_2 and stored at -80°C freezer for further extraction step.

4.2.7 Metabolite Extraction using modified cell scraping method

Quenching and extractions were performed using the methods illustrated schematically in figure 4.3.



2.
Quench Q on dry ice (~60 sec)
Retain in 2x eppendorfs for leakage analysis Q
3.
Add 1ml 100% Aq. MeOH for Extraction E
Scrape cells and retain for freeze/thaw in eppendorf E
1.
Remove media
± 1x PBS wash step
Vacuum concentrator
-80°C
5[Freeze/thaw]
Centrifuge -9°C
16000g 5 min
LN$_2$
Media
± 1x PBS
QR
Q
Leakage analysis
E
Cells
Cell scrape
~60 sec
*Dry Ice
*Dry Ice
*Dry Ice
MDA-MB-231
100mm nunclon dish
Seed: 2x10 cells
60 hr Incubation
(37°C & 5%CO$_2$)
Control
Extraction
1ml 100% aq. MeOH

Figure 4.3 Overview of quenching and extraction workflow using modified cell scraping method for metabolome analyses of adherently growing metastatic cancer cell line MDA-MB-231. 1 = Washing step with PBS; 2 = quenching protocol and 3 = extraction protocol.

The adherent cells in the Nucleon dish were then subjected to extraction by addition of 1 mL of pre-chilled 100% methanol (-40°C). The cells were then suspended in extraction solvent on dry ice by cell lifting with a pre-chilled rubber tipped cell lifter. The cell suspension was then transferred to a 2 mL pre-chilled Eppendorf and stored on dry ice. A further 1 mL aliquot of 100% methanol was added to the culture flask and the same procedure was repeated. The first and second aliquot were mixed together and the metabolites were extracted by performing freeze-thaw cycles as suggested elsewhere (Winder et al., 2008). Briefly the aliquot was snap frozen in liquid nitrogen for 3 minutes, followed by thawing on dry ice & vortexed. The freeze-thaw cycle was repeated five times for complete cell disruption and followed by centrifugation at 16000*g* at -9°C for 5 minutes. The pellet was re-extracted with 500 µL of 100% methanol. The first and second aliquots were then combined and the extract was lyophilized overnight prior to derivatization.

4.2.8 Simultaneous quenching and extraction with liquid nitrogen

Recent protocols (Bi et al., 2013, Hutschenreuther et al., 2012b, Lorenz et al., 2011) were adopted with a few modifications. Briefly, after removal of the culture medium and washings with PBS, the cells were rapidly quenched by directly adding about 15 mL liquid nitrogen (LN_2) to the Nucleon dish. The adherent cells in the Nucleon dish were then immediately subjected to extraction on dry ice by addition of 1 mL of pre-chilled 100% methanol followed by lifting the cell with a pre-chilled rubber tipped cell lifter. The cells were re-extracted and further steps were performed as described in section 4.2.7.

4.2.9 Metabolite derivatization, GC-MS analysis, metabolite identification and data analysis

Metabolite derivatization, GC-MS analysis, metabolite identification and data analysis was performed as described in Chapter 3.

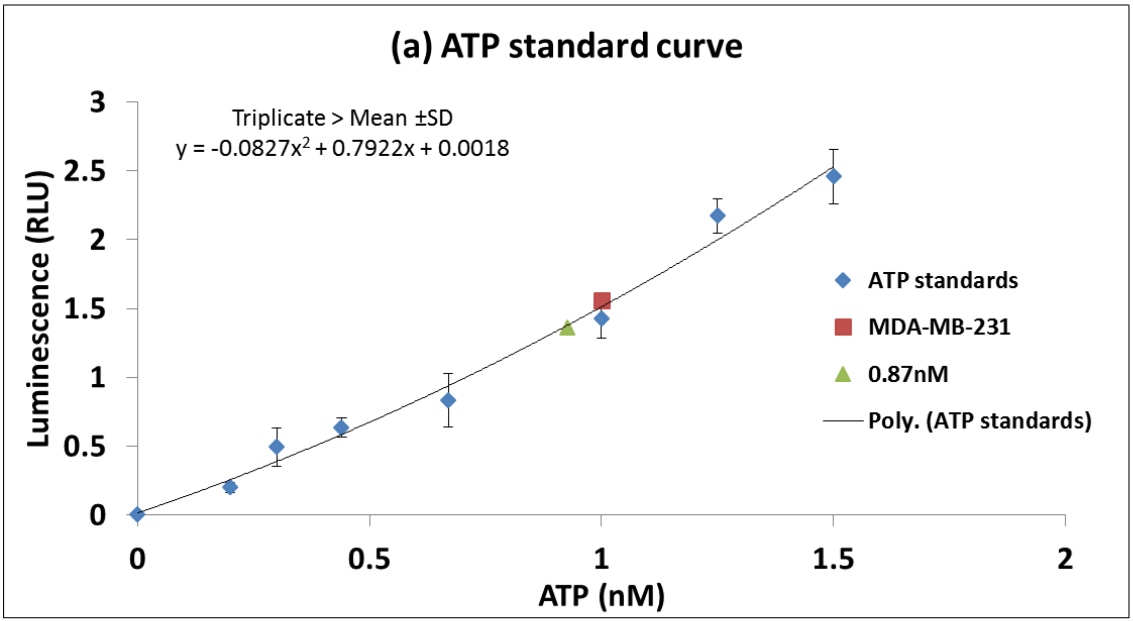
4.3 Results and Discussion

To obtain a snapshot of the intracellular metabolome, it is essential that the employed washing solutions are effective in removing the extracellular media components and similarly the quenching solution is efficient in halting the metabolic activities within the cell without altering the cell wall/membrane integrity. To achieve these primary objectives, required for any metabolomics research, it is important to evaluate the efficiency of the washing and quenching solutions. In this context, we have directed our approach for evaluating metabolite leakage in adherently grown metastatic cancer cell line MDA-MB-231 using a variety of washing and quenching solutions (buffered/non-buffered). Results from the different treatments are evaluated and compared using ATP assay, SEM analysis (for visual observation of membrane integrity) and GC-MS based untargeted metabolomics.

4.3.1 ATP assay

ATP is a key central metabolite to all live cells with high turnover rate (1.5 mM/s) (Dietmair et al., 2010) and with constant concentration across each cell. Unlike other metabolites, ATP is never secreted into the extracellular environment by cells under normal conditions. However, when cells are subjected to environmental stress or membrane damage, rapid changes in ATP concentration occurs because of the high turnover rate and subsequent leakage into the extracellular medium can occur (Lee and Fiehn, 2008a, Schwiebert and Zsembery, 2003). Quenching protocols in metabolomics studies are well known to cause severe leakage of metabolites from cells due to cold shock. Therefore, estimating changes in the total intracellular ATP content in response to various washing and quenching solution will provide an indication of the extent of metabolite leakage and quenching efficiency of the solutions.

Based on the above rationale, the first objective of this investigation was to evaluate the extent of metabolite leakage in response to different washing and quenching solutions (buffered/non-buffered) using the ATP assay. The ATP content was determined using a luciferase bioluminescence assay and normalised to protein content estimated by the BCA assay. ATP and protein assays have been successfully applied to the bacterial and mammalian cell suspension cultures, in contrast application to the adherent cell cultures are few. Hence prior to application of these assays to actual experiments on adherent cultures, standard curves were generated for both ATP and protein assay using standard solutions, an additional test standard solution and the MDA-MB-231 cell extract (n=3). The standard curves for both assays are shown in figure 4.4.



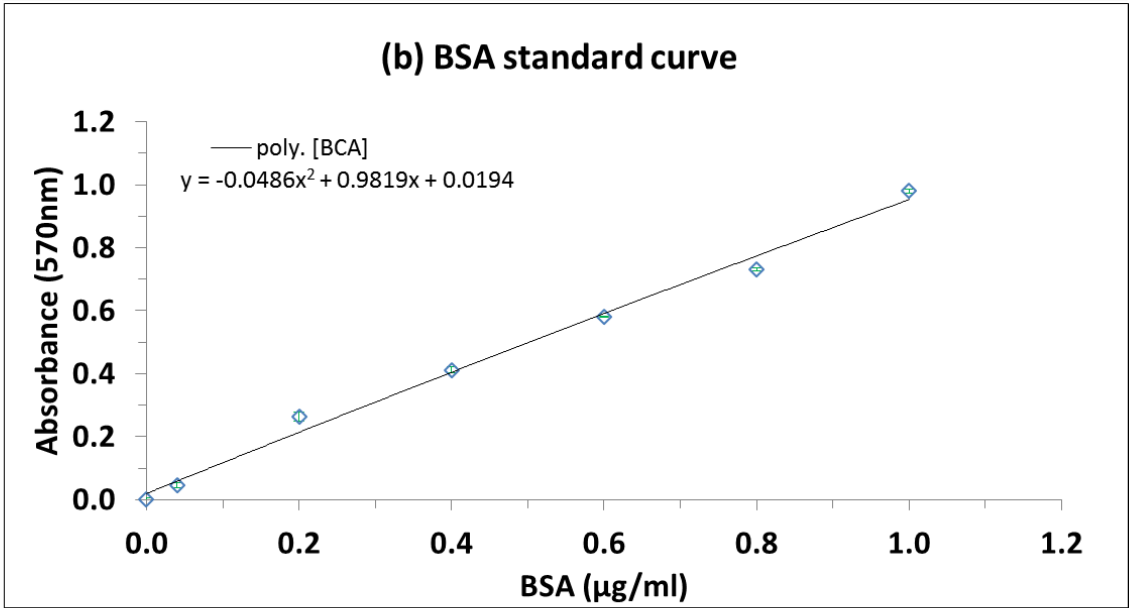


Figure 4.4 Standard curves generated for both luciferase bioluminescence ATP assay (a) and BCA protein assay (b) using mean values of ATP and protein content obtained with set of standard solutions (n=3).

4.3.2 Leakage of ATP in response to washing steps/solutions

The composition of the media used for mammalian cultures is rich in nutrients such as amino acids, vitamins, sugars etc., which are also present inside the cells as intracellular metabolites. Hence subsequent extraction of the cell content followed by quenching, without any washing step, will result in high recoveries for specific intracellular metabolites that are present in the medium as well. In the suspension culture, the washing step is performed after quenching and centrifugation of the sample. As addition of fivefold volume of quenching reagent to that of suspension culture will dilute the media components, some reports have suggested exclusion of washing step for rapid quenching of the culture and have demonstrated minimum leakage by doing so (Dietmair et al., 2010, Kronthaler et al., 2012a, Sellick et al., 2008). Moreover, inclusion of a washing step extends the duration of quenching resulting in more leakage.

In adherent cultures, similar workflow as that of suspension cultures cannot be applied as lifting of cells from their attached state by trypsinization or cell scraping will result in more mechanical damage causing severe leakage than would be caused by quenching or washing steps alone. Hence, it is essential to quench the adherent cultures in the adhered state only after removal of the medium. It is known that usage of large volumes of quenching solvents virtually removes all traces of most media components and extracellular metabolites. However, in cases where the media components are relatively more concentrated than the intracellular metabolites, the inclusion of a washing step to remove media components or extracellular metabolites could be advantageous. However, a key concern is whether the use of washing solutions and number of washing steps employed influence the metabolite leakage from adherent cultures?

In order to evaluate the effects of washing, the levels of free ATP in the washing solutions were determined after the washing step. (See Appendix 4.1 for calculation part).

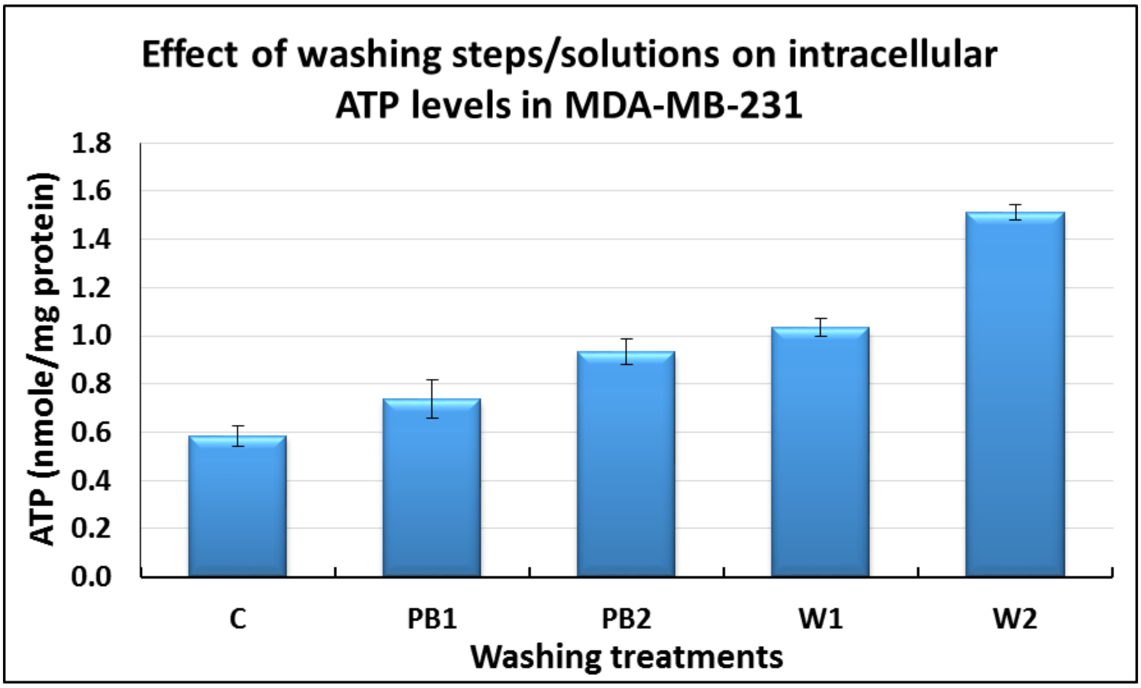


Figure 4.5 The effect of washing treatments on intracellular ATP levels of MDA-MB-231: C = No washing step (control); PB1 = washing once with PBS; PB2 = washing twice with PBS; W1 = washing once with water and W2 = washing twice with water. Bars represents the mean of 3 independent experiments ± standard error of the mean (SEM).

As can be seen in figure 4.5, washing with de-ionised water (W1 & W2) resulted in higher leakage of ATP (by 2 fold) compared to washing with PBS (PB1 & PB2). As the number of water washes increased to two, more leakage of ATP can be noticed (W1 vs W2; PB1 vs PB2). On the other hand, washing once with PBS (PB1) appears to show similar results to the control sample (C) suggesting that one PBS wash has a lower detrimental effect than a two-step PBS wash. As mentioned earlier in this report, mammalian cells are more delicate and to preserve their membrane integrity, it is essential to keep the ionic strength of their medium highly compatible. The extremely low ionic strength of water might be responsible for more leakage of intracellular ATP into extracellular environment caused by osmotic shock. The effect of more than 2 washes with PBS was not investigated as it will increase the processing time and will result in the decrease in ATP level due to the conversion of ATP to ADP and AMP and the estimation of metabolite leakage will not be reliable. For subsequent experiments, washing wash with PBS prior to quenching was selected as the optimal washing step.

4.3.3 Leakage of ATP in response to quenching solutions

As quenching is highly cell and sample dependent, differences in the use of optimal quenching additives with respect to specific sample can be explained. However, within the same biological system contradictory reports have been published with respect to extent of metabolite leakage (Dietmair et al., 2010, Kronthaler et al., 2012a, Sellick et al., 2008). On the other hand, in the case of adherent cell cultures, except for the recent report by Purwaha and co-workers (Purwaha et al., 2014), where the authors evaluated the use of buffer additives for OVCAR-8 cells (ovarian cancer cells), there is little information available. Even here, it is important to note that the evaluation was based on a very narrow approach, as the conclusions were drawn based on only amino acid analysis by HPLC. Moreover, there was no separate washing step included in the protocol, instead the quenching solvent was employed in two washings, which in fact might have caused more leakage than a single step rapid quenching. Furthermore, no analysis of supernatant following quenching was performed. However, to our knowledge, no quantitative data on the effectiveness of quenching solvents on adherent cultures have been presented with a broader approach. The broader approach in this context refers to the conclusions which are based highly correlated findings and obtained using two or more analytical platforms.

We have evaluated levels of free ATP in supernatant following quenching with various quenching solvents, the results of which are shown in 4.6.

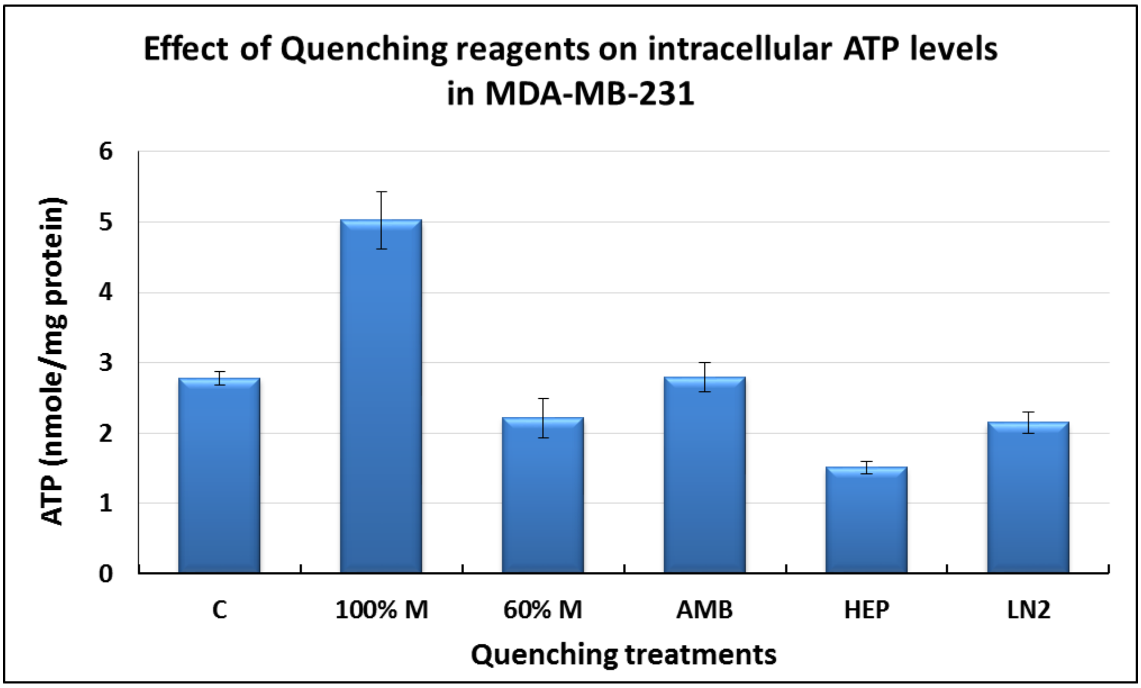


Figure 4.6 The effect of five quenching reagents on intracellular ATP levels of MDA-MB-231: C = unquenched sample (control); 100%M = quenching with 100% methanol; 60%M = quenching with 60% methanol; AMB = quenching with 60% aqueous methanol supplemented with 0.85% AMBIC; HEP = quenching with 60% aqueous methanol supplemented with 70 mM HEPES and LN_2 = simultaneous quenching and extraction with LN_2 followed by extraction with 100% methanol. Bars represents the mean of 3 independent experiments ± standard error of the mean (SEM).

Influence of non-buffered methanol quenching on leakage

As can be seen in figure 4.6, in case of non-buffered methanol quenching, quenching with 100% methanol compared to 60% aqueous methanol resulted in higher leakage of intracellular ATP, because of its detrimental effect on cell membrane integrity. The levels of free ATP in nmole/mg of proteins obtained with 100% methanol and 60% aqueous methanol are 2.21 nmole/mg and 5.02 nmole/mg respectively. However Koning and co-workers suggested use of 60% aqueous methanol as the most appropriate choice for effective quenching step because of its lower freezing point (Koning and Dam, 1992). Whilst evaluating the influence of increasing methanol concentration above 60% our results showed severe leakage of metabolite (figure 4.6). Our findings are in agreement with those of Canelas and co-workers where higher leakage of intracellular metabolites in *S. cerevisiae* was reported with the increase in methanol concentration (Canelas et al., 2008). In GC-MS analysis, we expect higher recoveries of intracellular metabolites with 60% aqueous methanol compared to that of 100% methanol which will be discussed later in detail.

Influence of buffered methanol quenching on leakage

In the case of buffered methanol quenching, methanol supplemented with HEPES showed a lower leakage of intracellular ATP (1.51 nmole/mg) compared to that supplemented with AMBIC (2.79 nmole/mg). In addition, an overall comparison of the non-quenched and methanol quenched (buffered/non-buffered) samples clearly suggests that methanol supplemented with HEPES causes minimum leakage of intracellular ATP and could be said to have the least detrimental effect on the membrane integrity of metastatic MDA-MB-231 cells. Our results of ATP assay are in partial agreement with that of Faijes and co-workers, where authors reported less leakage with buffered methanol compared to that of methanol alone (Faijes et al., 2007). However, in contrast to our results the study reported equal quenching efficiencies for both HEPES and AMBIC in case of *Lactobacillus plantarum*. In addition, contradicting data was published by Sellick and co-workers for CHO cells, where the authors reported AMBIC buffered methanol to be better than HEPES buffered methanol in preserving membrane integrity (Sellick et al., 2008). As stated elsewhere in this report, this might be due to considerable differences between the sample and cell size and/or cell wall/membrane structure. Methanol supplemented with 70 mM HEPES has shown less leakage of intracellular ATP and thereby can be said to preserve the membrane integrity of MDA-MB-231 cells compared to other quenching solvents.

4.3.4 SEM

In addition to the ATP assay, the cell membrane integrity of MDA-MB-231 was further studied with SEM analysis after the application of the quenching solutions. The adherent mammalian cell integrity was studied on samples of unquenched cells with single PBS wash (as a control) and cells quenched with four different quenching solutions such as with 100% methanol, 60% aqueous methanol, 60% aqueous methanol supplemented with 0.85% AMBIC and with 60% methanol supplemented with 70 mM HEPES. The results of SEM observation on MDA-MB-231 in response to various quenching solutions are summarised in figure 4.7.

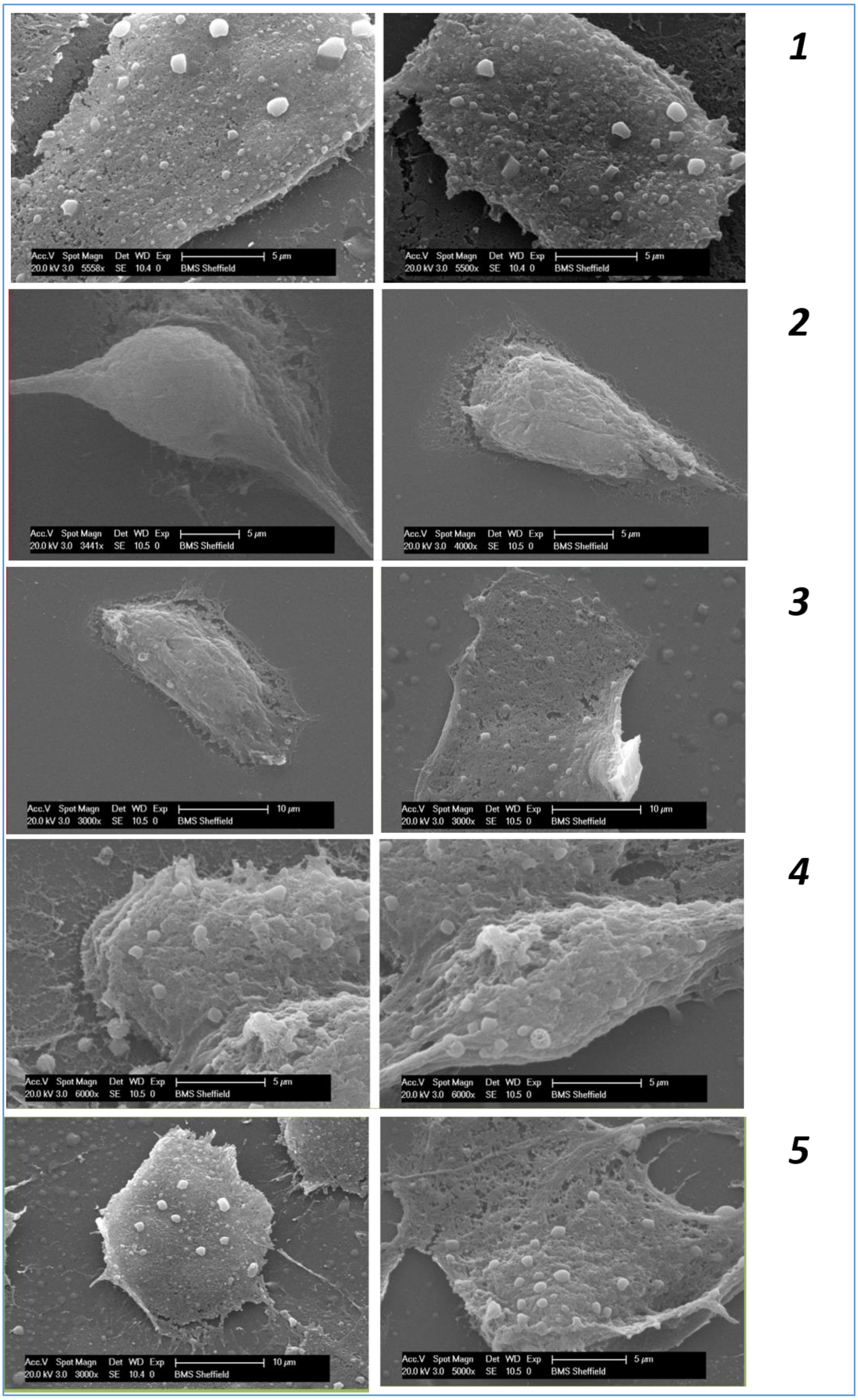

1
Acc.V Spot Magn Det WD Exp 5 µm
20.0 kV 3.0 5558x SE 10.4 0 BMS Sheffield
Acc.V Spot Magn Det WD Exp 5 µm
20.0 kV 3.0 5500x SE 10.4 0 BMS Sheffield
2
Acc.V Spot Magn Det WD Exp 5 µm
20.0 kV 3.0 3441x SE 10.5 0 BMS Sheffield
Acc.V Spot Magn Det WD Exp 5 µm
20.0 kV 3.0 4000x SE 10.5 0 BMS Sheffield
3
Acc.V Spot Magn Det WD Exp 10 µm
20.0 kV 3.0 3000x SE 10.5 0 BMS Sheffield
Acc.V Spot Magn Det WD Exp 10 µm
20.0 kV 3.0 3000x SE 10.5 0 BMS Sheffield
4
Acc.V Spot Magn Det WD Exp 5 µm
20.0 kV 3.0 6000x SE 10.5 0 BMS Sheffield
Acc.V Spot Magn Det WD Exp 5 µm
20.0 kV 3.0 6000x SE 10.5 0 BMS Sheffield
5
Acc.V Spot Magn Det WD Exp 10 µm
20.0 kV 3.0 3000x SE 10.4 0 BMS Sheffield
Acc.V Spot Magn Det WD Exp 5 µm
20.0 kV 3.0 5000x SE 10.5 0 BMS Sheffield

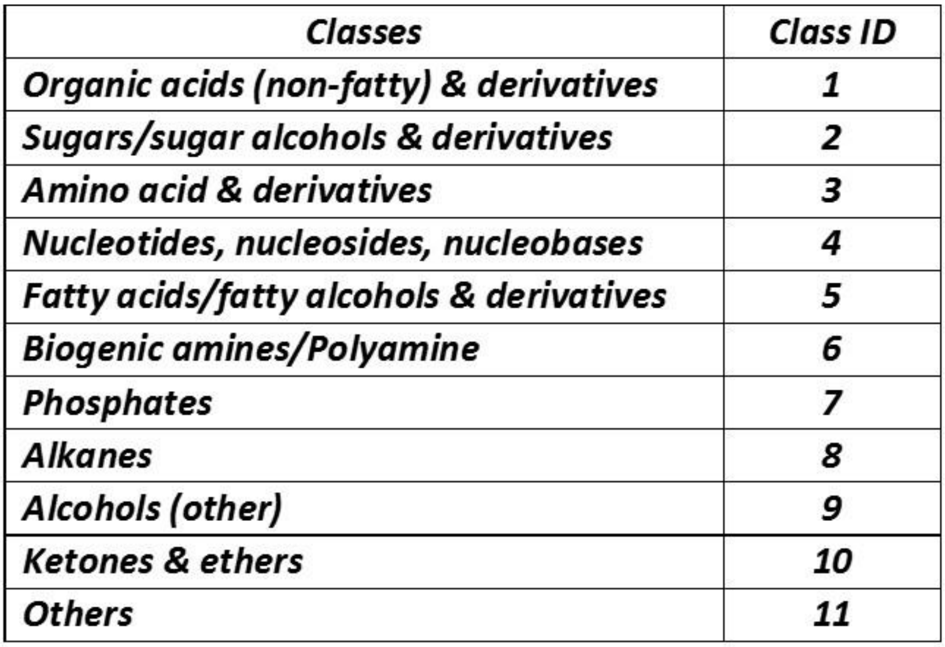
Figure 4.7 The effect of various quenching solutions (buffered/non-unbuffered) on MDA-MB-231 cell membrane integrity was visualised and compared against the non-quenched cells using SEM observations, where: 1 = Non-quenched cells (Control); 2 = quenched with 100% methanol; 3 = quenched with 60% aqueous methanol; 4 = quenched with 60% aqueous methanol supplemented with 0.85% AMBIC and 5 = quenched with 60% aqueous methanol supplemented with 70 mM HEPES.

In the case of non-quenched cells (control) (Fig 4.7-1), it can be clearly seen that MDA-MB-231 cells adopted a polygonal and a flat morphology in the adhered/fixed state, where the patterned network and vascular channels are clearly visible. However, when the cells are quenched with 100% methanol (Fig 4.7-2), detrimental effects on cellular morphology (completely shrinked) can be clearly seen where patterned network, vascular channels and pores have completely vanished. The use of 60% aqueous methanol (Fig 4.7-3) also shown to have a similar effect, however managed to preserve the membrane integrity to some extent compared to that of 100% methanol. The use of additives such as HEPES and AMBIC managed to preserve the membrane integrity to a greater extent, where slight damage to cellular network can be seen with use of AMBIC (Fig 4.7-4) compared to control. Overall, quenching with methanol supplemented with 70 mM HEPES (Fig 4.7-5) seems to be a better choice compared to other quenching solutions and correlates very well with our finding based on ATP assay.

4.3.5 GC-MS based overall recovery of metabolites

In order to obtain confirmation of the findings from the ATP assay and SEM, GC-MS analysis of the samples was undertaken after reaction with the quenching solutions (buffered or non-buffered). In total 140 unique metabolites were identified (Appendix 4.2) across all treatments in adherently growing metastatic cancer cell line MDA-MB-231 However, when the study objective is to maximise metabolome coverage, characterisation of the leakage with respect to metabolite class is required. For this, it is essential to first identify and classify the metabolites detected based on their physicochemical properties, thus generating a metabolite matrix that can be utilized for interpretations. Hence, all the identified metabolites with the different sample treatment approaches were initially classified into eleven different metabolite classes based on their physicochemical properties (Table 4.2).

Table 4.2 Eleven different metabolite classes along with their class IDs



Classes	***Class ID***
Organic acids (non-fatty) & derivatives	***1***
Sugars/sugar alcohols & derivatives	***2***
Amino acid & derivatives	***3***
Nucleotides, nucleosides, nucleobases	***4***
Fatty acids/fatty alcohols & derivatives	***5***
Biogenic amines/Polyamine	***6***
Phosphates	***7***
Alkanes	***8***
Alcohols (other)	***9***
Ketones & ethers	***10***
Others	***11***

4.3.5.1 Metabolite leakage in response to washing steps

After monitoring cell-free supernatant of quenched cells and the blank medium, the necessary correction was applied for appropriate calculation of intracellular metabolites. Only features that were present in at least two biological replicates out of three were considered for further analysis. A summary of the recovery efficiency for all the applied sampling protocols are shown in figure 4.8, where the metabolite class and numbers detected for various treatment are plotted. Figure 4.8a summarises the metabolites detected in the cell pellets, Figure 4.8b those detected in the cell-free supernatant, Figure 4.8c indicates the metabolites present in the cells but not in the supernatants, Figure 4.8d indicates metabolites present in the supernatants and not in the cells, and Figure 4.8e shows metabolites present in both the cell pellet and the supernatant. Higher metabolite numbers in the supernatant (Fig 4.8b), relatively high numbers detected in both the cells and the supernatants (Fig 4.8e), and corresponding low numbers unique to the cell pellets (Fig 4.8c) indicates high metabolite leakage. Higher proportion of metabolites detected in Figure 4.8e, compared to that detected in Figure 4.8c or Figure 4.8d indicates that there is an increased chance of metabolite leakage.

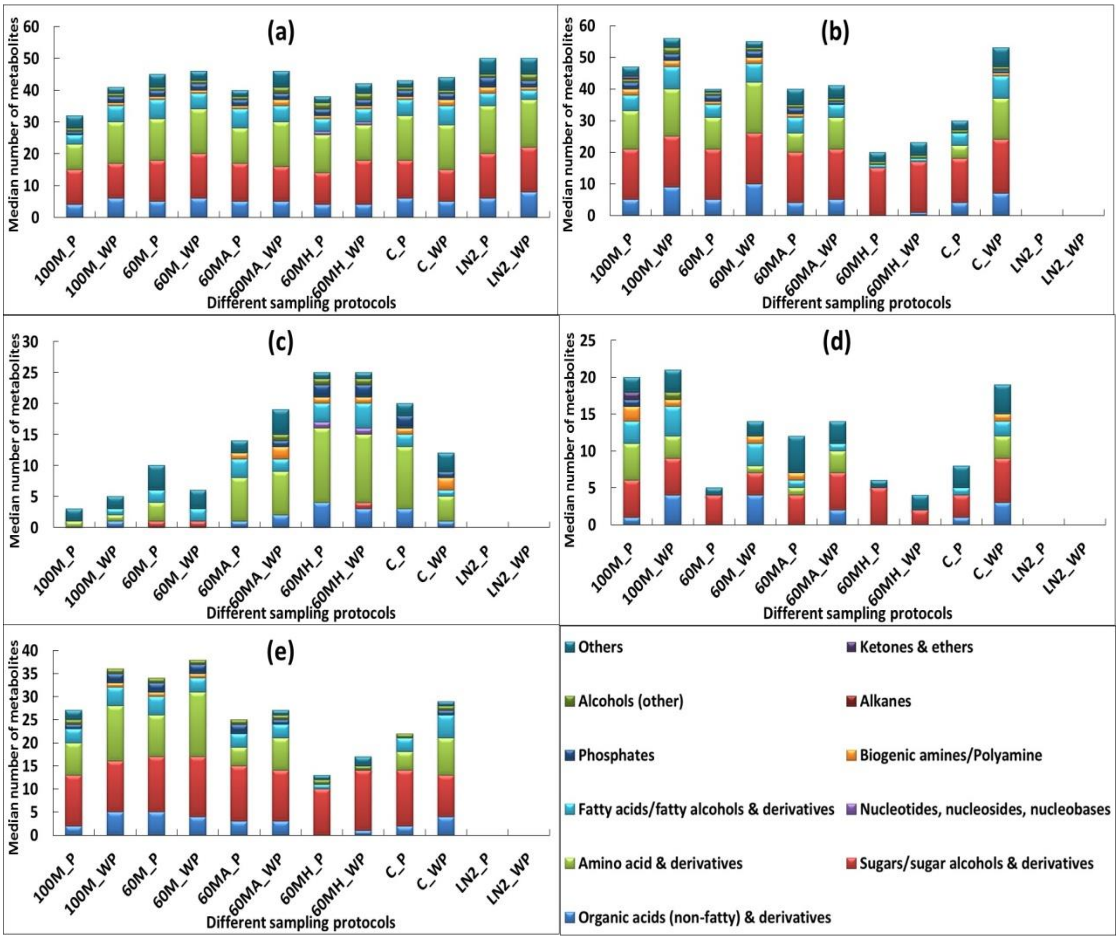


Figure 4.8 The effect of five quenching reagents with (single) or without PBS wash step on leakage of intracellular metabolites during quenching was compared for MDA-MB-231. X-axis represents different sampling protocols; 100M_P = 100% methanol with PBS wash step; 100M_WP = 100% methanol without PBS wash step; 60M_P = 60% methanol (aqueous) with PBS wash step; 60M_WP = 60% methanol (aqueous) without PBS wash step; 60MA_P = 60% methanol (aqueous with 0.85% w/v AMBIC) with PBS wash step; 60MA_WP = 60% methanol (aqueous with 0.85% w/v AMBIC) without PBS wash step; 60MH _P = 60% methanol (aqueous with 0.85% w/v HEPES) with PBS wash step; 60MH _WP = 60% methanol (aqueous with 0.85% w/v HEPES) without PBS wash step; C _P = control (unquenched) with PBS wash step; C _WP = control (unquenched) without PBS wash step; LN_2 _P = liquid nitrogen with PBS wash step and LN_2 _WP = liquid nitrogen without PBS wash step. Y-axis represents median number of metabolites of each class. After both the treatment the extracted metabolites from cell extracts, cell-free supernatant post quenching and blank samples were analysed by GC-MS. (a) metabolites identified in cell extracts only, (b) metabolites identified in supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both).

As can be seen from figure 4.8c, 60% aqueous methanol supplemented with HEPES yielded similar recoveries of metabolites unique to the cells with both inclusion and exclusion of washing step and highest compared to other treatments (except that of LN_2). Correspondingly metabolites common to both the cells and supernatant (Fig 4.8e) are the lowest (with slightly higher recovery with exclusion of washing step (60MH_WP)). The number of metabolite detected in the supernatant are also the least compared to other treatments (Fig 4.8d). Among other treatments, exclusion of washing step with 100M_WP and 60MA_WP treatments resulted in overall higher recovery of metabolites unique to the cells compared to that of with inclusion of PBS wash (Fig 4.8c). In contrast, higher recoveries were obtained with 60M_P and C_P treatments involving washing step compared to that of exclusion of washing step. However, recoveries of metabolites common to both the cells and supernatant (Fig 4.8e) exclusion of washing step in all the treatments resulted in higher recoveries compared to that of inclusion of washing step. Possible reason might be carry-over of media components along with the intracellular metabolites contributing to the higher recoveries with all the treatments without washing step. Only use of HEPES, direct quenching with LN_2 (with or without PBS wash) showed negligible variations in recovery of metabolites in cell extracts (Fig 4.8a).

Our findings are in disagreement to the previously published report (Sellick et al., 2008), where the authors suggested no washing step is required in the case of suspension CHO cultures, based on their ATP assay results. The authors reported loss of ATP with all the evaluated treatments which includes methanol alone and that supplemented with 70 mM HEPES, 0.85% AMBIC and 0.9% NaCl. However it is important to note that, the authors did the washing step with quenching reagents which may further aggravate the metabolite leakage due to cold shock phenomenon and extended time limits for quenching step. In another report (Kronthaler et al., 2012a), similar strategy was adopted for CHO cells, where the authors investigated the effect of additional washing steps (one, two and three) against no washing, where PBS was used for both washing and quenching steps. The authors reported severe leakage with the all the washing steps. The possible reason might be the longer duration of the protocol which might have accelerated the conversion of ATP to ADP and AMP. We have demonstrated minor variation in ATP leakage between non-washed (control) and cells washed once with PBS, which clearly suggests that the washing step with PBS does not result in major leakage of metabolites in adherent mammalian cultures. Moreover GC-MS analysis confirmed this conclusion, where minor variations in recovery of intracellular metabolites were observed (Fig 4.8c) across all the applied treatments.

4.3.5.2 Metabolite leakage in response to various quenching solutions

With respect to analysis of metabolite classes recovered, non-buffered methanol quenching treatments (both 60% aqueous and 100%) resulted in severe leakage of nearly all metabolite classes (Fig 4.8c) compared to that of cells quenched with buffered methanol and direct quenching with LN_2. Data is not shown for LN_2, as there was no supernatants analysed with this treatment due to the nature of the protocol. Among buffered quenching solutions, methanol supplemented with HEPES resulted in the highest recoveries of organic acids, amino acids, nucleotides and phosphates compared to that of AMBIC supplemented methanol. With all the applied treatments no recovery for alkanes and ketones was observed in MDA-MB-231 cells. This finding is in contrast to a previously published report (Purwaha et al., 2014), where the authors reported similar recoveries obtained for amino acids in OVCAR-8 ovarian cancer cell line with both treatments. However as discussed in section 2.3.3, the possible reasons for this contradictory result might be the use of different analytical techniques employed for the evaluation of the treatment method, and in this case conclusions were drawn based on only amino acid analysis by HPLC.

In contrast to both buffered and non-buffered methanol, direct quenching with LN_2 yielded highest recoveries for all the metabolite classes (Fig 4.8a). However, there were problems in using LN_2 for simultaneous quenching and extraction of metabolites. The procedure involves pouring of LN_2 directly into T75 culture flasks (small neck diameter) or well plates, and this could result in frostbite or cryogenic burns, and possible asphyxiation. In addition the compatibility of culture flasks or well plates used for culturing adherently growing mammalian cells should be carefully tested for direct use with LN_2, in order to avoid the loss of samples.

4.3.5.3 Evaluating the interference of quenching additives with derivatization reactions and GC-MS analysis

In past, Sellick and co-workers successfully demonstrated the use of AMBIC with no apparent interference with the derivatization reactions and GC-MS analysis (Sellick et al., 2008). However, no such evaluation has been reported for HEPES. Kronthaler and co-workers suggested the use of simultaneous quenching and washing step with PBS for suspension cultures (Kronthaler et al., 2012a), although addition of fivefold volume of simultaneous quenching and washing solvent will result in extreme concentration of phosphate with the subsequent potential of interference with the GC-MS analysis.

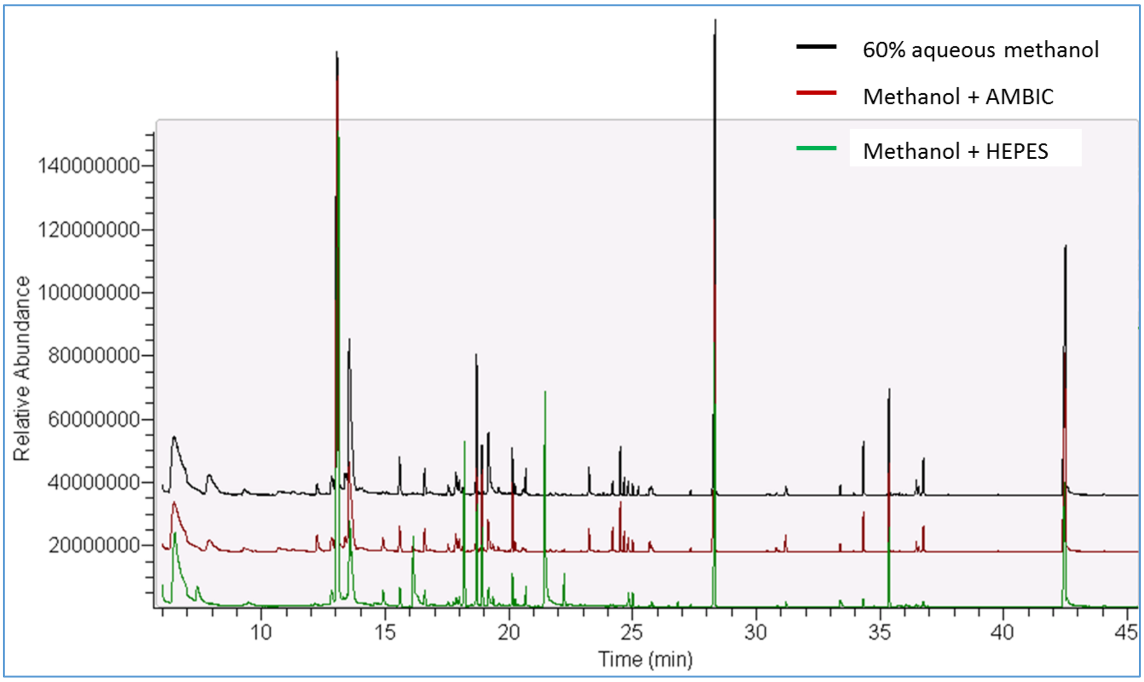


Figure 4.9 GC-MS based analysis of intracellular metabolites extracted from adherent MDA-MB-231 mammalian cells. Cells were washed once with PBS and later quenched with either 60% aqueous methanol alone (black), or 60% aqueous methanol + 0.85% AMBIC (red) or 60% aqueous methanol + 70 mM HEPES (green) prior to extraction of metabolites using cold methanol. For clarity, the Y-axis for 60% aqueous methanol and for 60% aqueous methanol + 0.85% AMBIC has been offset.

To evaluate the effects of PBS washing step and use of quenching additives (HEPES or AMBIC) on two step derivatization protocols (methoximation followed by silylation) and GC-MS based analysis, we have profiled the metabolome and compared the chromatograms of methanol supplemented with HEPES to that of methanol supplemented with AMBIC and methanol alone. Fig 4.9 represents GC-MS based analysis of intracellular metabolites extracted from adherent MDA-MB-231 mammalian cells. Cells were washed once with PBS and later quenched with either 60% aqueous methanol alone (black), or 60% aqueous methanol + 0.85% AMBIC (red) or 60% aqueous methanol + 70 mM HEPES (green) prior to extraction of metabolites using cold methanol. For clarity, the Y-axis for 60% aqueous methanol and 60% aqueous methanol + 0.85% AMBIC has been offset. Comparison between three different sample types with respect to total numbers and amounts of metabolites recoved showed no major diffrences between them. In total 41 ± 3 unique metabolites were identified with each of the three different sample types. However, slightly higher relative peak areas for all the metabolites were obtained with methanol + HEPES treatment compared to other two treatments. This analysis clearly reveals that the presence of HEPES and AMBIC does not interefere with the derivatization reactions and GC-MS analysis. Our results for inclusion of AMBIC were in agreement with that of Sellick and co-workers (Sellick et al., 2008), however in addition to that we note no apparent interference of HEPES with the derivatization rections and the GC-MS analysis.

4.4 Conclusions

Estimation of the intracellular ATP levels in response to various washing steps/solutions resulted in minor variations in ATP leakage between non-quenched (control) cells and cells washed once with PBS, indicating a single washing step with PBS does not alter the membrane integrity of adherent mammalian cultures. Moreover, GC-MS analysis confirmed this conclusion, where similar numbers of intracellular metabolites were recovered (Fig 4.8c) with all the applied treatments. On the other hand, estimation of intracellular ATP levels in response to various quenching solvents has shown less leakage of intracellular ATP with the use of methanol supplemented with 70 mM HEPES, and appears to preserve the membrane integrity of MDA-MB-231 cells compared to other quenching solvents. Similar conclusions were drawn from GC-MS based metabolomic analyses, where analysis based on recoveries of eleven different metabolite classes clearly demonstrated severe leakage of nearly all metabolite classes with the use of non-buffered methanol quenching treatments compared to that of cells quenched with buffered methanol and direct quenching with LN_2. Among buffered quenching solutions containing 60% aqueous methanol, HEPES supplemented methanol resulted in highest recoveries of organic acids, amino acids, nucleotides and phosphates compared to that of AMBIC supplemented methanol. Furthermore, SEM analysis in response to various quenching solvents yielded similar conclusion to that of ATP assay and GC-MS based analyses. In addition, we have demonstrated no interferences on the derivatization reactions and GC-MS based analysis when PBS and HEPES are employed in the protocol. All these findings from different analytical platforms correlate very well and clearly indicate that a single washing step with PBS and quenching with methanol supplemented with 70 mM HEPES resulted in minimum leakage of intracellular metabolites.

Chapter 5

Comparison of proposed protocols for metabolomics of two metastatic TNBC cell lines: MDA-MB-231 and MDA-MB-436

5. Comparison of proposed protocols for metabolomics of two metastatic TNBC cell lines: MDA-MB-231 and MDA-MB-436

5.1 Background

Breast cancer (BC) based on their genomic profiles and tumour characteristics is a highly heterogeneous type of cancer for which different subgroups of cell lines that are morphologically and clinically distinct have been established (Hutschenreuther et al., 2013). In human mammary gland, there are two distinct types of epithelial cells namely basal and luminal cells which are immuno-histochemically distinct from each other (Perou et al., 2000). Steroid hormone receptors such as estrogen receptor (ER) and progesterone receptor (PR) in concert with the oncogene ErbB-2 (HER-2) are critical determinants of BC subtypes. The triple negative breast cancer (TNBC) is characterised by absence of ER, PR and lack of overexpression of HER-2. TNBC represents approximately 20% of all BC and is typically associated with poor prognosis. Moreover, due to its aggressive phenotype TNBC only partially responds to chemotherapy and present lack of clinically established targeted therapies thus increases the probability of fatality in patients (Criscitiello et al., 2012, Gluz et al., 2009, Podo et al., 2010).

In order to sustain growth and proliferation of tumour cells, they constantly require supplements of macromolecular precursors thereby exhibiting altered metabolism compared to quiescent cells. As a result, several researchers have employed the metabolomics approach to catalogue these and have focused on the classification approach where healthy cells were compared against tumour cells (Podo et al., 2010). Nevertheless, while doing so it is important to obtain reliable metabolomics data using an optimised workflow which will provide maximum coverage for all classes of metabolites with minimum leakage.

In Chapter 3, conventional techniques were compared against the simultaneous quenching and extraction techniques using three different adherently growing mammalian cell lines and successfully demonstrated that the direct quenching and extraction techniques yielded three fold higher recovery of all metabolites classes compared to the conventional sampling techniques. Among the simultaneous quenching and extraction techniques the modified direct cell scraping (NTcs) and LN_2 treatments provided a rapid and reliable route which can be combined with the newly suggested bead harvesting technique in order to minimise artefacts due to metabolome leakage. In Chapter 4, we employed the optimised NTcs and LN_2 treatments, and investigated the effect of washing steps/solvents and quenching solvent additives on metabolite leakage from adherently growing metastatic cancer line MDA-MB-231. This identified one washing step with PBS followed by quenching with 60% aqueous methanol supplemented with the 70mM HEPES as the best washing and quenching solvents for NTcs treatment which resulted in minimal leakage of metabolites compared to 60% aqueous methanol supplemented with 0.85% AMBIC.

It is important to note that, in both Chapters it was demonstrated that whichever approach is chosen, it is imperative that the option is first characterised for the metabolome leakage with the particular cell type being investigated before being adopted, moreover it is essential to consider the impact of different treatments on recovery of different classes of metabolites. In comparing the previously characterised metabolite profiles of two metastatic cancer lines (MDA-MB-231 and MDA-MB-436) after optimised NTcs and LN_2 treatments will provide preliminary insight into the impact of these two treatments on the recovery of different classes of metabolites. Data obtained from these measurements were analysed with the following objectives:

1) To examine the proposed protocols for adherent mammalian cell metabolomics to identify significant variations between them?
2) To examine whether the proposed protocols able to track the changes between the cell lines, which share common biological and morphological features?
3) To examine the impact of these proposed methods in pathway based analysis for deriving meaningful biological interpretations, which will aid in understanding the cancer metabolism and developing novel therapeutic regimens.

In order to meet these objectives, data obtained from the proposed protocols was selected from Chapter 3 and 4 (NTcs and LN_2) and compared with the results obtained with them for two metastatic cancer cell line (MDA-MB-231 and MDA-MB-436).

5.2 Comparison of the proposed protocols for adherent mammalian cell metabolomics

To investigate the influence of two proposed treatments (NTcs and LN_2) on recovery of different classes of metabolites and thereby numbers of metabolites recovered within each class were further investigated for two different cell lines (MDA-MB-231 and MDA-MB-436), thereby resulting in four different conditions, as detailed below:

Condition A: Metabolome coverage with NTcs method for MDA-MB-231

Condition B: Metabolome coverage with LN_2 method for MDA-MB-231

Condition C: Metabolome coverage with NTcs method for MDA-MB-436

Condition D: Metabolome coverage with LN_2 method for MDA-MB-436

The results of the treatments are summarised in figure 5.1, where the distribution of metabolite classes detected in metabolome of the two metastatic cancer cell lines are plotted.

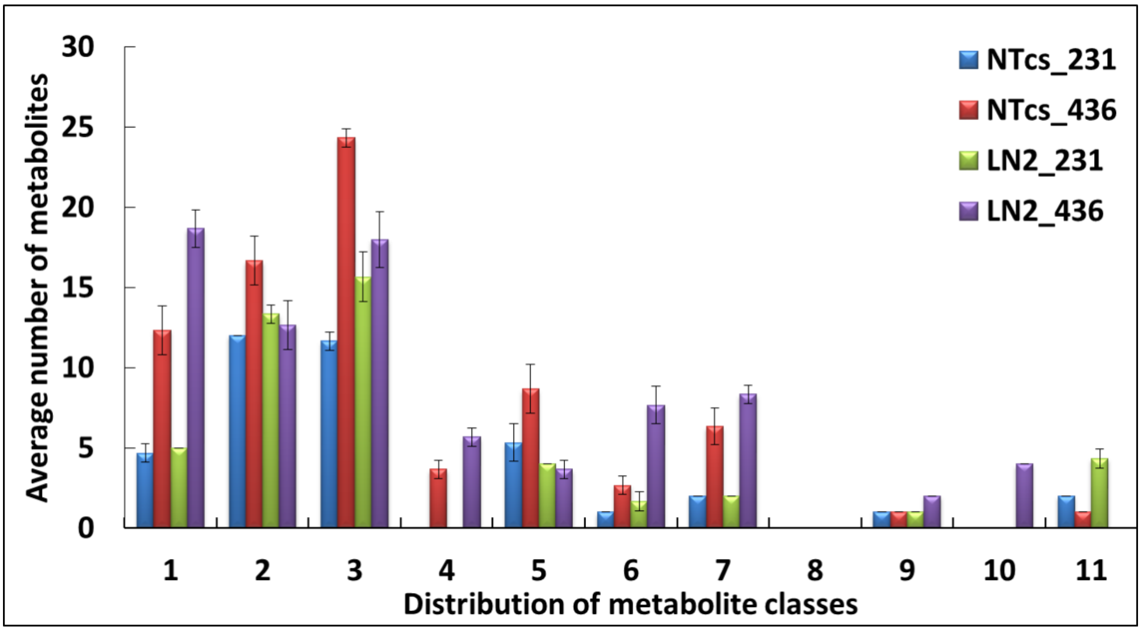


Figure 5.1 Distribution of metabolite classes detected in metabolome of two metastatic cancer cell lines namely MDA-MB-231 and MDA-MB-436 using two quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN_2). Y-axis represents average number of metabolites identified in the cell pellets. X-axis represents distribution of metabolite classes where 1 = organic acids (non-fatty) and derivatives; 2 = sugar/sugar alcohols and derivatives; 3 = amino acids and derivatives; 4 = nucleotides, nucleosides and nucleobases; 5 = fatty acids/fatty alcohols and derivatives; 6 = biogenic amines/polyamines; 7 = phosphates; 8 = alkanes; 9 = alcohols (others); 10 = ketones and ethers; 11 = others.

As can be seen from figure 5.1, for MDA-MB-231 cell line similar recoveries for five organic acids were observed with both the applied methods, however with MDA-MB-436 higher recoveries were obtained with LN_2 method (around 18 organic acids). For sugar/sugar alcohols and amino acids and its derivatives, higher recoveries were obtained with NTcs method (recovered 16 sugars and 24 amino acids) than with LN_2 method (recovered 12 sugars and 18 amino acids) for MDA-MB-436, whereas in contrast to this higher recoveries were obtained with LN_2 method (recovered 13 sugars and 28 amino acids) than with NTcs method (recovered 12 sugars and 11 amino acids) in MDA-MB-231. With both methods, no recoveries for nucleotides/nucleosides/nucleobases was seen in MDA-MB-231, whereas in MDA-MB-436 slightly higher recoveries were obtained with LN_2 method than with NTcs method, for this class. With both the cell lines, NTcs method showed higher recoveries for fatty acids/fatty alcohols compared to LN_2 method, whereas in contrast to this, LN_2 method showed higher recoveries for biogenic amines/polyamines class compared to NTcs method.

With both methods and in both the cell lines no recoveries were observed for alkanes (8), whereas minor variations in recoveries of alcohols were observed. Only LN_2 method recovered few metabolites in MDA-MB-436 belonging to ketone and ether class, whereas except LN_2 in MDA-MB-436, all other methods recovered few other classes of metabolites in both the cell lines.

To further investigate, influence of these methods in recovering different metabolites within the same class, we have plotted the median intensities of metabolites recovered in each class, the result of which are summarised in figure 5.2. In total 157 unique metabolites, were identified with both the cell line and methods, using NIST 05 and GOLM metabolome databases, which were then further classified into eleven different classes of metabolites, based on their physicochemical properties. The list of 157 unique metabolites along with their class ID as mentioned in figure 5.2, are summarised in Appendix 5.1.

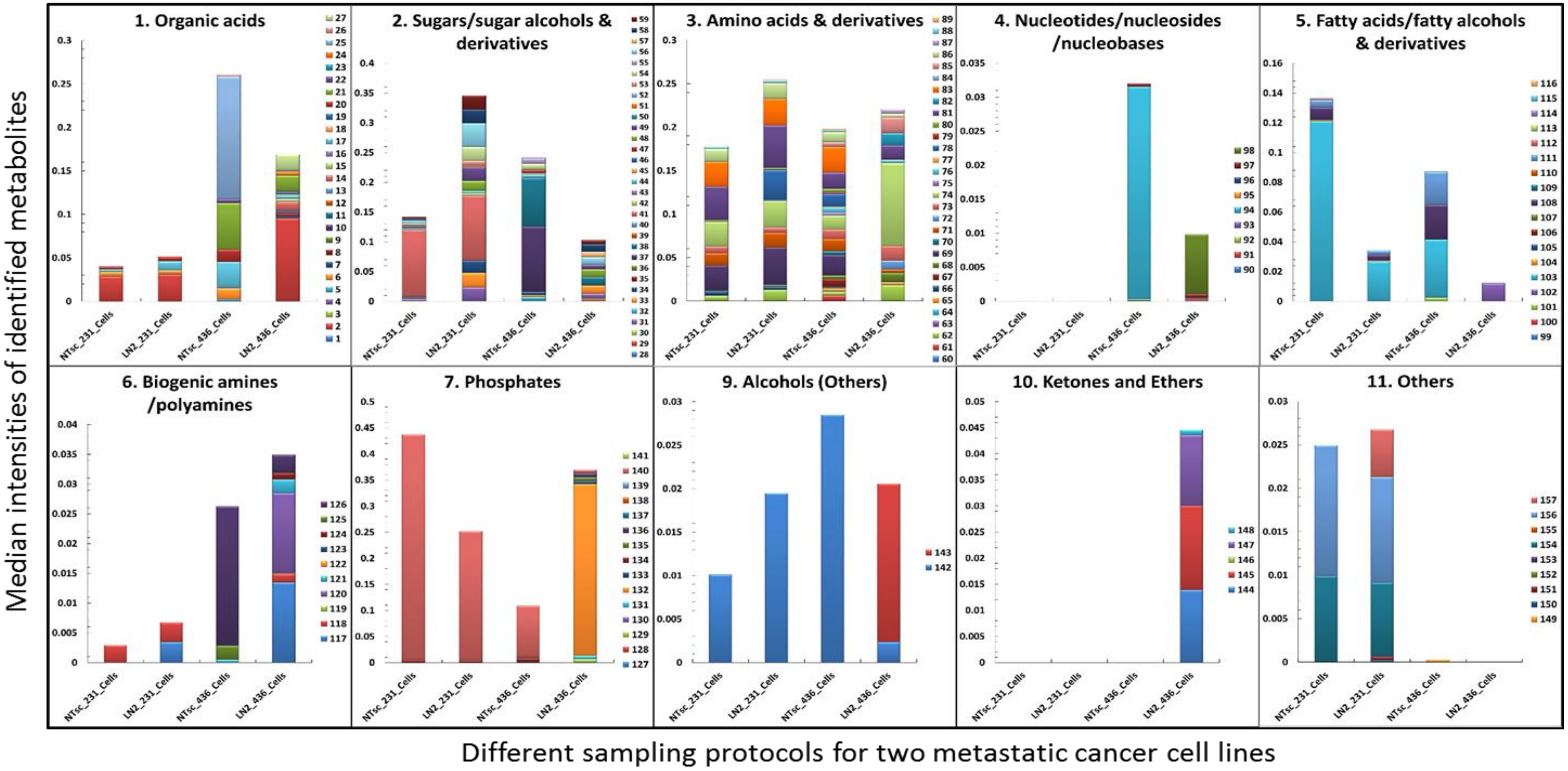


Figure 5.2 Median intensities of metabolites recovered in individual metabolite class with two metastatic cancer cell lines namely MDA-MB-231 and MDA-MB-436 across two different quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN_2). X-axis represents method and cell line investigated. Y-axis represents median intensities of metabolites recovered in individual metabolite within each metabolite class. In total 157 metabolites (See Appendix 5.1) were identified across all the treatments which were classified into 11 different metabolite classes and plotted separately.

As can be seen from figure 5.2, in case of organic acids, the recovery obtained with both methods were similar in the case of MDA-MB-231. In MDA-MB-436, higher numbers were recovered with LN_2 method (20), whereas higher intensities were observed with NTcs method. It is important to note that both methods showed variation in recovery of different organic acids. In case of sugars/sugar alcohols and amino acids and its derivatives, both higher number and intensities were recovered with LN_2 method than with NTcs method in case of MDA-MB-231, whereas contrast results were observed with MDA-MB-436, where NTcs method seems to be superior to LN_2. With both methods, no recoveries for nucleotides/nucleosides/nucleobases was seen in MDA-MB-231, whereas in case of MDA-MB-436, higher recoveries in numbers with lower intensities were obtained with LN_2 method. In both the cell lines, NTcs method seems to be superior than the LN_2 method in recovery of fatty acids/fatty alcohols and derivatives. In contrast, for remaining metabolites classes in both the cell lines, LN_2 method seems to be superior to that of NTcs method.

Overall, it can be seen that, in both the cell lines, both the LN_2 method and NTcs method are superior with respect to recoveries of different metabolite classes and intensities recovered.

5.3 Are the proposed protocols able to track the changes between the cell lines which share common biological and morphological features?

In order to understand better the coverage of different classes of metabolites in cell extracts of both the cell lines with both the applied methods and to evaluate whether both the proposed methods are able to track the variation in metabolome of these two metastatic cancer cell lines, we assessed and compared both the sampling protocols in both the cell lines for the recovery of eleven different metabolites classes and then used to generate heat map, as displayed in figure 5.3. Each band represents the median metabolite levels from three independent determinations, normalised for each condition and expressed on a logarithmic scale to capture variation that was over orders of magnitude. Data for retention time, S/N ration, peak tailing, m/z value and peak area was collected manually by exporting to MS Excel 2013. MATLAB 7.0 (MathWorks, Natick, MA, USA) with in house routines employed for analysing the data containing metabolite identities to obtain median values for each identified metabolite. For intracellular metabolites the GC-MS data was normalised to viable cell count and sum of peak areas as suggested elsewhere (Hutschenreuther et al., 2012b).

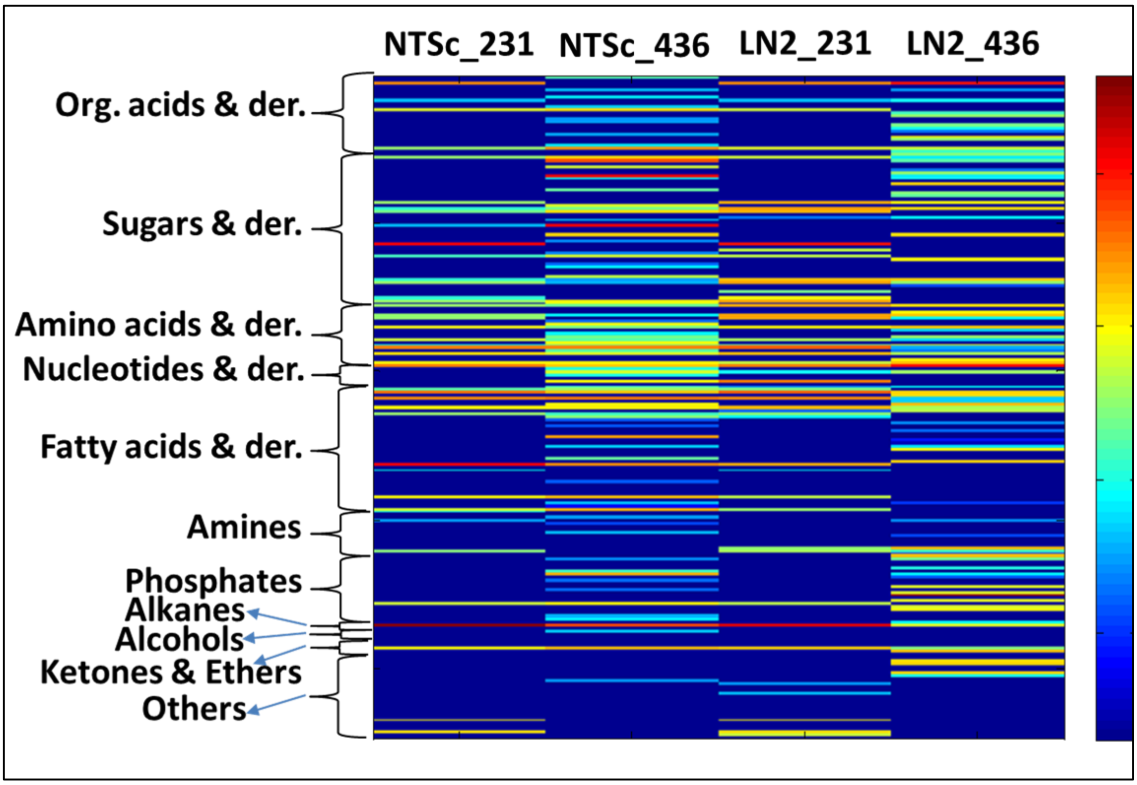
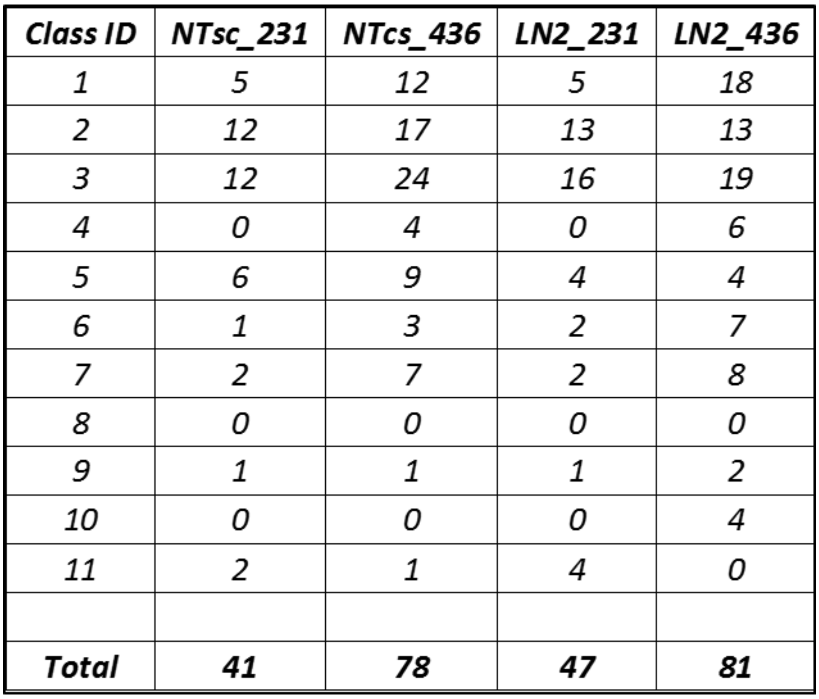


Figure 5.3 Heat map comparing the response of 11 different classes of metabolite signals identified by GC-MS in metabolome of two metastatic cancer cell lines namely MDA-MB-231 and MDA-MB-436 using two quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with liquid nitrogen (LN_2).

As displayed in figure 5.3, minor variations in the recovery of organic acids was observed in MDA-MB-231 with both methods, whereas higher recoveries were observed in MDA-MB-436 with the LN_2 method. In MDA-MB-231, the recovery of amino acids and its derivatives and sugars/sugar alcohols and its derivatives with the LN_2 method was found to be superior over NTcs method, however contradictory results were observed in MDA-MB-436, where NTcs method show higher recoveries than the LN_2 method. In both the cell lines, the LN_2 method resulted in superior recovery of nucleotides/nucleosides/nucleobases, biogenic amines/polyamines and phosphates whereas recovery of fatty acids/alcohols and its derivatives was severely impeded as compared to NTcs method. Overall the recovery with respect to metabolome coverage with both methods in both cell lines showed minor variations as summarised in table 5.1, although slightly higher variations were observed with respect to recovery of different metabolite classes as displayed in figure 5.3.

Table 5.1 Median number of metabolites recovered in MDA-MB-231 and MDA-MB-436 across two different quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN2)



Class ID	*NTsc_231*	*NTcs_436*	*LN2_231*	*LN2_436*
1	*5*	*12*	*5*	*18*
2	*12*	*17*	*13*	*13*
3	*12*	*24*	*16*	*19*
4	*0*	*4*	*0*	*6*
5	*6*	*9*	*4*	*4*
6	*1*	*3*	*2*	*7*
7	*2*	*7*	*2*	*8*
8	*0*	*0*	*0*	*0*
9	*1*	*1*	*1*	*2*
10	*0*	*0*	*0*	*4*
11	*2*	*1*	*4*	*0*
Total	***41***	***78***	***47***	***81***

Despite these minor variations between the two proposed methods, we selected both methods for further investigation, where we directed our efforts in tracking the variations in the metabolome with both methods in two metastatic cancer cell lines which share common biological and morphological similarities. To assess the sample arrangements, PCA is calculated (figure 5.4) using the feature intensities from all the analysed samples using XCMS online. Comparison within the two cell lines with both methods (436 vs 231 -NTcs) and (436 vs 231 -NTcs) displayed around 5 fold more significant changes along the X-axis than those on the Y-axis. To further gain access to different data characteristics such as *p*-value, fold change, retention time, *m/z* ratio and signal intensity of all detected features, XCMS online, a web based platform was employed for further validation of GC-MS results using visualisation of the results by PCA analysis and cloud plot. The results of these investigations are summarised in figure 5.4 and 5.5, and table 5.2.

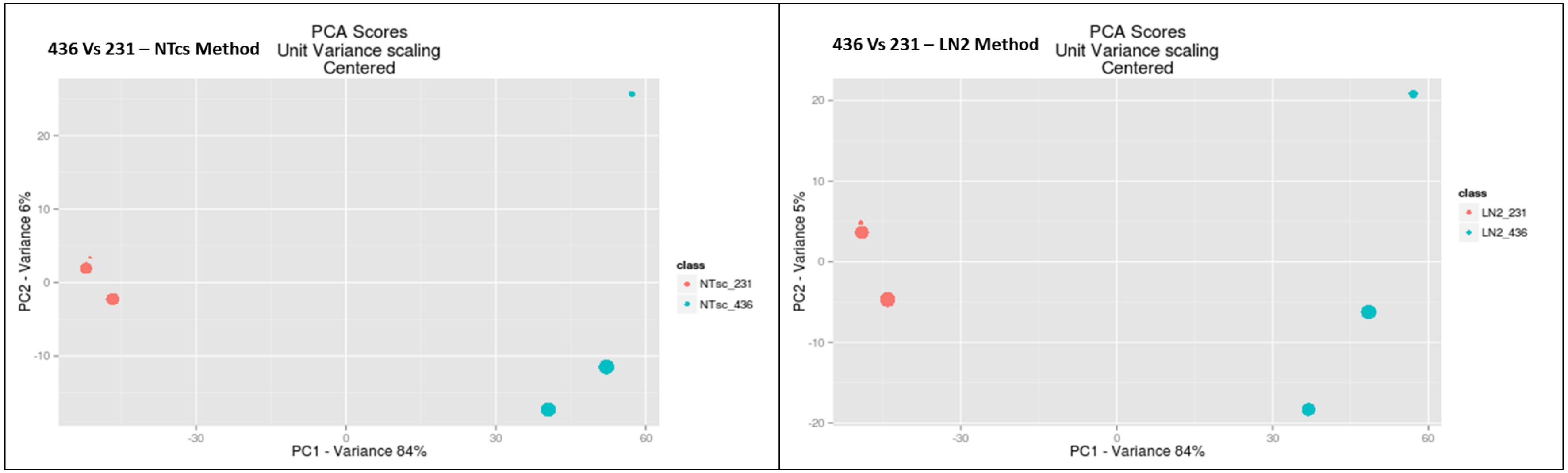


Figure 5.4 PCA analysis of metabolomics data displaying clustering of samples between two metastatic cancer cell lines with both the proposed methods: NTcs (left) and LN_2 (right). PCA is calculated using the feature intensities from all samples. The colours (red/blue) are assigned based on the sample class. Distance to the model (DModX) test was used to verify the presence of outliers and to evaluate whether a submitted sample fell within the model applicability domain.

Table 5.2 Number of significantly different features across method and cell line comparison

Comparison type	**Comparison**	**Significant features**
Method comparison	A Vs B (NTsc Vs LN2 for 231)	83
	C Vs D (NTsc Vs LN2 for 436)	308
Cell line comparison	A Vs C (231 Vs 436 By NTsc)	1204
	B Vs D (231 Vs 436 By LN2)	1149

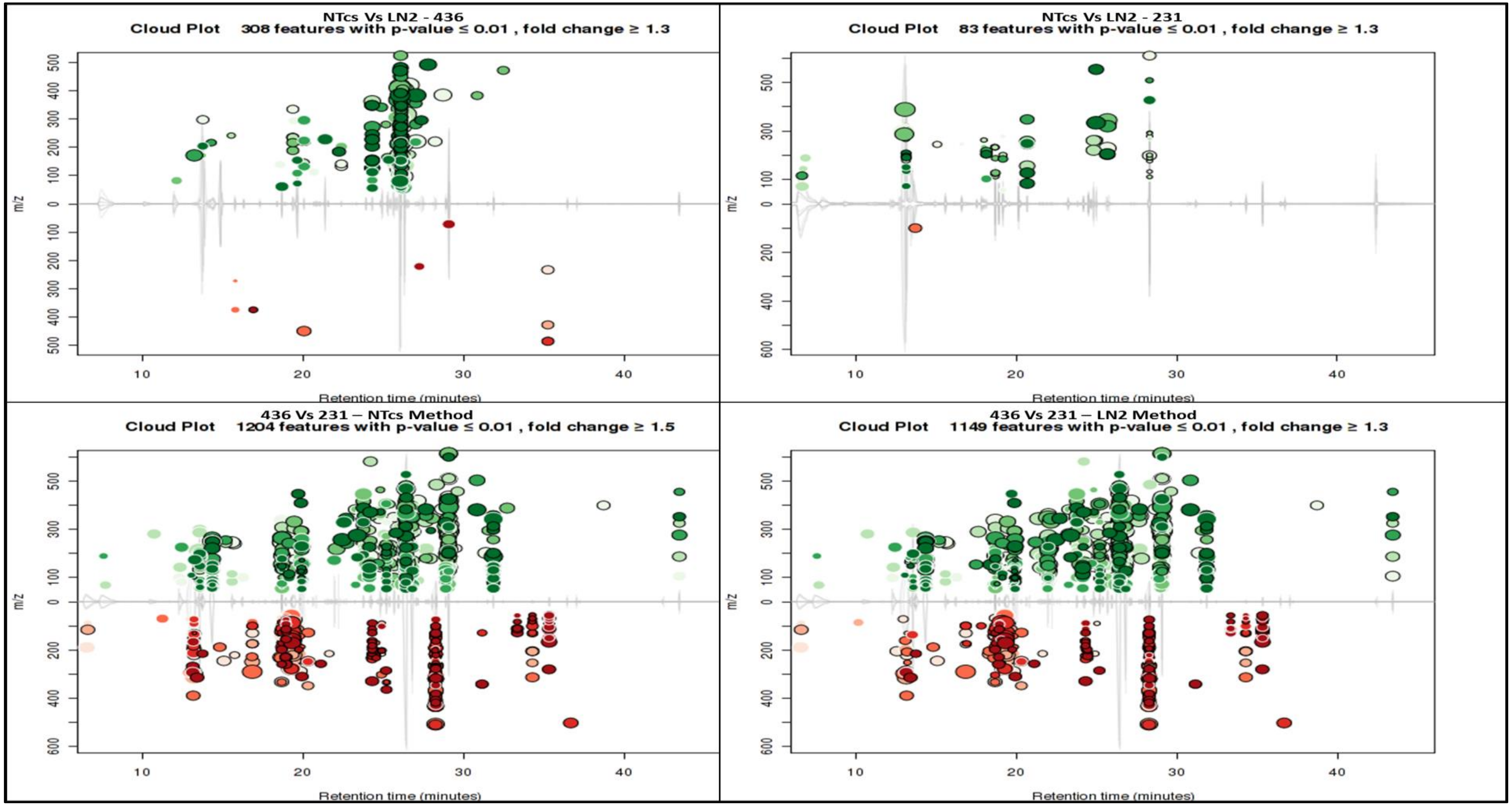


Figure 5.5 Cloud plots generated for four different comparison (C vs D (top left); A vs B (top right); A vs C (bottom left) and B vs D (bottom right)) as listed in Table 5.2, displaying significant features with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value.

As can be seen from figure 5.5 and table 5.2, fewer dysregulated features were detected when both the cell lines (83 with A vs B and 308 with C vs D) are compared, although very high number of dysregulated features were detected across the cell line comparison with both methods (1204 with A vs C and 1149 with B vs D). "Significant features" in this case, refers to those intensities were altered between sample groups with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubble corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value, darker the colour, the smaller the p-value. The statistical significance of the fold change, as calculated by a Welch *t* test with unequal variances, is represented by the intensity of the features colour, where features with low *p*-values are brighter compared to features with high *p*-values, and therefore more significant. The Y co-ordinate of each feature corresponds to the *m/z* ratio of the compound as determined by the MS. The features that were identified with the METLIN database are being colour coded with the black outline surrounding the bubble, whereas those not identified in the database are displayed without a black outline.

These observations clearly demonstrate the ability of both the proposed methods to track the variations in the metabolome of two biologically and morphologically similar metastatic cancer cell lines.

5.4 Impact of these proposed methods in pathway based analysis for deriving meaningful biological interpretations which will aid in understanding the cancer metabolism and developing novel therapeutic regimens

The biology and pathways involved in TNBC is still very poorly understood, therefore there is demand for a comprehensive approach that will aid in defining the relative contribution of major metabolic pathways in TNBC (Ossovskaya et al., 2011). Gross and co-workers recently demonstrated that for the growth and survival of TNBC cells, they are largely dependent on glutamine and key glutamate derived intermediates, which are responsible for supporting macromolecule synthesis, cellular redox balance and ATP production to meet the constant metabolic demands of proliferative TNBC cells (Gross et al., 2013). This highlights the significance of amino acid biosynthesis pathways with respect to metabolome coverage. Analysis of differential gene expression in metabolic pathways revealed significant variations in purine, pyrimidine and folate metabolism in TNBC cells (Bobrovnikova-Marjon and Hurov, 2014, Li et al., 2012, Ossovskaya et al., 2011). Few reports have suggested significant variations in arginine and proline metabolism (Traina et al., 2013). Further research with TNBC cells, revealed importance of understanding alterations in adenosine triphosphate-binding cassette (ABC) in order to improve our ability in rationally designing appropriate chemotherapy regimens (Ablett et al., 2012, Alison et al., 2012, Marquette and Nabell, 2012) and in mitochondrial phosphorylation (Walsh et al., 2012). Recent research while exploring the novel therapeutic targets for TNBC suggested elevated levels of serine and glycine in TNBC due to shift in metabolic flux from phosphotidylcholine to glycine, which might serve as a precursor for macromolecules and antioxidant defence required for rapid proliferation of TNBC, therefore anticancer therapy should aim at targeting serine and glycine biosynthesis (Amelio et al., 2014, Moestue et al., 2010). Some reports also suggested alanine, aspartate and glutamate metabolism as TNBC specific pathways and thus employed for functional characterisation of breast cancer (Tian et al., 2014). To our knowledge only a single report implicated the importance of glyoxylate and dicarboxylate metabolism in TNBC (JOHNSON et al., 2014). The pentose/glucuronate inter-conversions pathways in TNBC was recently reported while uncovering the therapeutic vulnerabilities using genome and transcriptome sequencing (Craig et al., 2013).

From the above discussion, it is clear that various attempts have been made in past three years to illustrate the metabolic pathways involved in TNBC, however the approach has been restricted to coverage of specific metabolic pathways. As mentioned earlier, and the metabolic pathways involved in TNBC are still very poorly understood, and demands for a comprehensive approach that will aid in defining the relative contribution of major metabolic pathways in TNBC has increased. Given this, we focused our analysis on evaluating the coverage of major breast cancer pathways in the analysed metabolome for the two metastatic cancer cell lines with our optimised NTcs and LN_2 methods.

For comprehensive analysis, we have classified all the identified metabolites with both methods in both the metastatic cancer cell lines based on their pathway involvement. Some of the metabolites were found to be involved in more than one pathway. The list of 126 metabolites with their corresponding pathway involvement are summarised in Appendix 5.2.

Initially C number for all the identified metabolites were searched using DBGET search tool in KEGG compound category (Kanehisa and Goto, 2000). All the identified metabolites were then mapped on cancer pathways using KEGG mapper tool. In total, 84 metabolites were successfully mapped which were highlighted by red dot in figure 5.6,

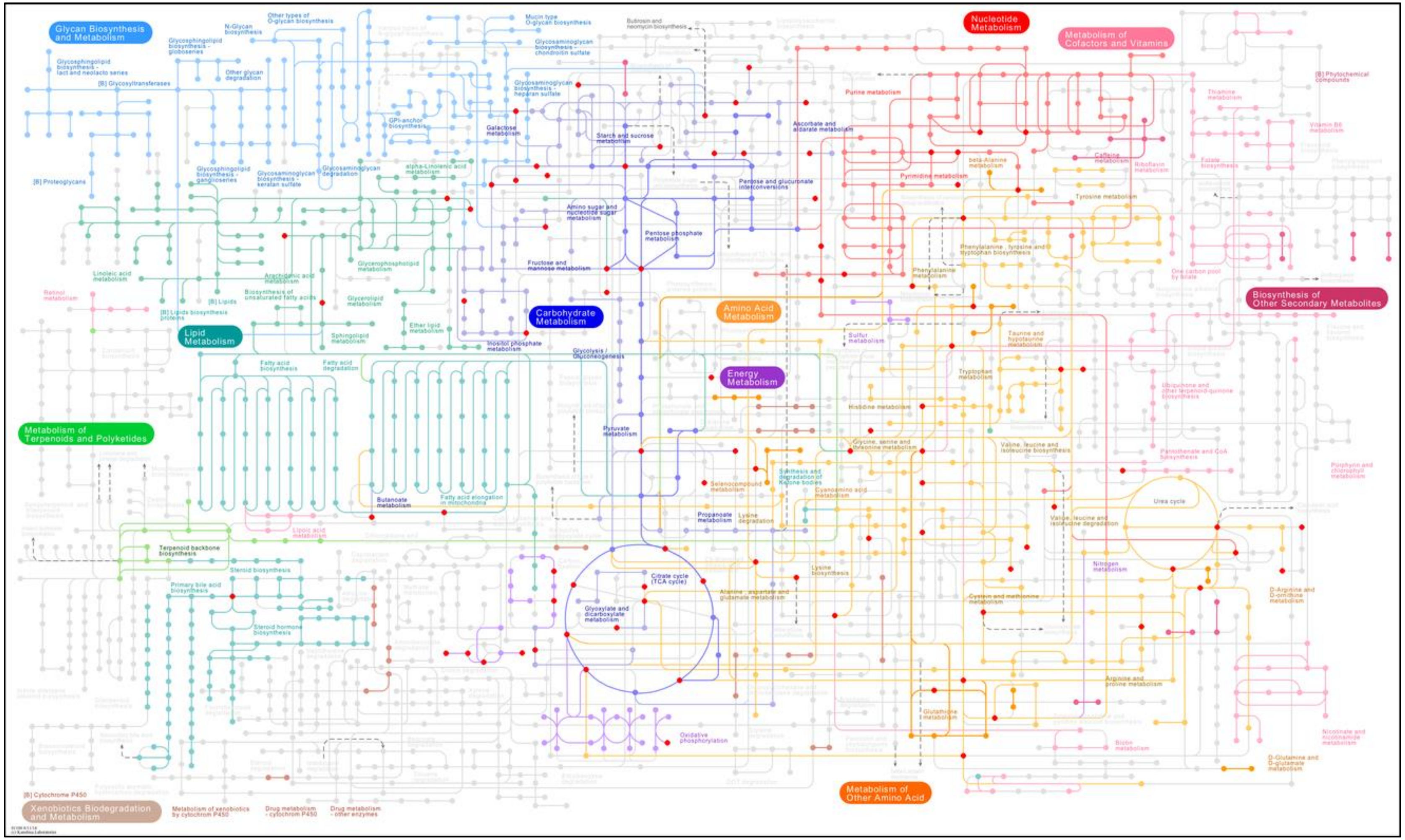

Glycan Biosynthesis and Metabolism
N-Glycan biosynthesis
Other types of O-glycan biosynthesis
Mucin type O-glycan biosynthesis
Glycosaminoglycan biosynthesis - chondroitin sulfate
Butirosin and neomycin biosynthesis
Nucleotide Metabolism
Metabolism of Cofactors and Vitamins
Glycosphingolipid biosynthesis - globoseries
Glycosphingolipid biosynthesis - lact and neolacto series
Other glycan degradation
[B] Glycosyltransferases
Glycosaminoglycan biosynthesis - heparan sulfate
Purine metabolism
[B] Phytochemical compounds
Thiamine metabolism
GPI-anchor biosynthesis
Galactose metabolism
Starch and sucrose metabolism
Ascorbate and aldarate metabolism
Vitamin B6 metabolism
[B] Proteoglycans
Glycosphingolipid biosynthesis - ganglioseries
Glycosaminoglycan biosynthesis - keratan sulfate
Glycosaminoglycan degradation
alpha-Linolenic acid metabolism
Pentose and glucuronate interconversions
beta-Alanine metabolism
Pyrimidine metabolism
Caffeine metabolism
Riboflavin metabolism
Folate biosynthesis
Tyrosine metabolism
Amino sugar and nucleotide sugar metabolism
Pentose phosphate metabolism
Phenylalanine, tyrosine and tryptophan biosynthesis
Phenylalanine metabolism
One carbon pool by folate
Linoleic acid metabolism
Arachidonic acid metabolism
Glycerophospholipid metabolism
Fructose and mannose metabolism
Retinol metabolism
[B] Lipids
Biosynthesis of unsaturated fatty acids
Glycerolipid metabolism
Carbohydrate Metabolism
Amino Acid Metabolism
Biosynthesis of Other Secondary Metabolites
[B] Lipids biosynthesis proteins
Lipid Metabolism
Ether lipid metabolism
Sphingolipid metabolism
Inositol phosphate metabolism
Glycolysis / Gluconeogenesis
Sulfur metabolism
Taurine and hypotaurine metabolism
Fatty acid biosynthesis
Fatty acid degradation
Energy Metabolism
Tryptophan metabolism
Ubiquinone and other terpenoid-quinone biosynthesis
Histidine metabolism
Metabolism of Terpenoids and Polyketides
Pyruvate metabolism
Glycine, serine and threonine metabolism
Valine, leucine and isoleucine biosynthesis
Pantothenate and CoA biosynthesis
Porphyrin and chlorophyll metabolism
Selenocompound metabolism
Synthesis and degradation of ketone bodies
Cyanoamino acid metabolism
Butanoate metabolism
Fatty acid elongation in mitochondria
Propanoate metabolism
Lysine degradation
Urea cycle
Valine, leucine and isoleucine degradation
Lipoic acid metabolism
Terpenoid backbone biosynthesis
Nitrogen metabolism
Steroid biosynthesis
Lysine biosynthesis
Primary bile acid biosynthesis
Citrate cycle (TCA cycle)
Alanine, aspartate and glutamate metabolism
Cysteine and methionine metabolism
D-Arginine and D-ornithine metabolism
Glyoxylate and dicarboxylate metabolism
Steroid hormone biosynthesis
Arginine and proline metabolism
Glutathione metabolism
Oxidative phosphorylation
Biotin metabolism
Nicotinate and nicotinamide metabolism
D-Glutamine and D-glutamate metabolism
[B] Cytochrome P450
Metabolism of Other Amino Acid
Xenobiotics Biodegradation and Metabolism
Metabolism of xenobiotics by cytochrome P450
Drug metabolism - cytochrome P450
Drug metabolism - other enzymes

Figure 5.6 In total 84 identified metabolites with both methods (NTcs and LN2) in two metastatic cancer cell lines namely MDA-MB-231 and MDA-MB-436 were found to be present in cancer pathways (involved in TNBC biology), which were plotted on KEGG's metabolic map and highlighted by red dot.

As highlighted above, there were numerous major pathways likely to be involved in TNBC biology and therefore comprehensive coverage of these pathways with the aid of metabolomics will surely help in illustrating the underlying biology and mechanisms involved in TNBC cells and for finding novel therapeutic targets. Given this, we have specifically mapped all the identified metabolites with both methods in two metastatic cancer cell line to get a first glimpse at metabolome coverage with respect to specific important pathways involved in TNBC biology. Selected pathways that were mapped with the identified metabolites are listed below:

- ABC transporters: 35 metabolites
- Biosynthesis of amino acids: 27 metabolites
- Pyrimidine metabolism: 10 metabolites
- Glycine, serine and threonine metabolism: 11 metabolites
- Alanine, aspartate and glutamate metabolism: 11 metabolites
- Arginine and proline metabolism: 15 metabolites
- Glyoxylate and dicarboxylate metabolism: 14 metabolites
- Pentose and glucuronate interconversions: 10 metabolites
- Galactose metabolism: 14 metabolites
- Oxidative phosphorylation: Fumarate
- Protein digestion and absorption: 23 metabolites

From above information it is clear that the proposed two methods for quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN_2) show superior metabolome coverage for the two metastatic cancer cell lines. Furthermore, most of the identified metabolites with these methods were successfully mapped on significant pathways involved in TNBC cells, thereby demonstrating the robustness and reliability of these methods in pathways based analysis of TNBC cell lines.

5.5 Conclusions

The previously optimised sampling protocols in Chapter 3 and 4, namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN_2) showed superior metabolome coverage in the two metastatic cancer cell lines investigated. With preliminary investigations in both the cell lines, we have concluded that both methods are superior to each other with respect to the recoveries of different metabolite classes and intensities recovered for them. However, the use of visualisation tools resulted in fewer number of significantly changing features, thereby demonstrating the minimum variation between both methods. The importance of both methods in recovering specific metabolite classes with specific cell line under investigation have been demonstrated. Therefore, when particular class of metabolites are of interest, it is essential to carefully consider the impact of these two treatments on the recovery of different classes of metabolites. Comparison within the metabolome recovered in two different metastatic cancer cell lines with both methods revealed variations in significantly changing features, where number of metabolites identified with MDA-MB-436 were around two fold higher than that of identified in MDA-MB-231. These observations clearly demonstrate the ability of both the proposed methods to track the variations in the metabolome of two biologically and morphologically similar metastatic cancer cell lines.

In total 154 metabolites were identified with both methods in both the metastatic cancer cell lines which were classified based on their pathway involvement. Most of the identified metabolites with these methods were successfully mapped on significant pathways involved in TNBC cells, thereby demonstrating the robustness and reliability of these methods in pathways based analysis in TNBC cell lines. No work has yet been published on adherent mammalian cell metabolomics where the emphasis was on improving the metabolome coverage. The observations made with the pathway-based coverage specific to cancer will help in generating many hypothesis based investigations. For example, in setting up a platform for classifying cancer sub-types, in defining the relative contribution of major metabolic pathways in TNBC and illustrating the underlying molecular mechanism involved in TNBC biology and finally for generating novel therapeutic regimes by identifying novel targets for anticancer therapies.

Chapter 6

Optimization of the quenching and extraction workflow for GC-MS based metabolomic investigation of microbial and microalgal species: Protocol validation with hypothesis driven biological study in *C. reinhardtii.*

6. Optimization of the quenching and extraction workflow for GC-MS based metabolomic investigation of microbial and microalgal species: Protocol validation with hypothesis driven biological study in *C. reinhardtii.*

6.1 Introduction

Microalgae not only play an important ecological role, but are also of commercial importance and therefore there is a need for an in depth study of the metabolites through metabolomics (Jamers et al., 2009b). *Nannochloropsis* is well appreciated in aquaculture because of their nutritional value and ability to produce valuable chemical compounds such as pigments, chlorophyll a, zeaxanthin, astaxanthin and polyunsaturated fatty acids (PUFAs). The metabolism of such pigments is very dynamic, as exposure of cells to changes in light environment causes rapid changes in pigments. Therefore, in order to elucidate the underlying mechanisms involved, it is essential to carry out a metabolomic study first with the optimised quenching and extraction protocols (Rocha et al., 2003). *Dunaliella salina* is well known for the commercial production of β-carotene as it accumulates carotenoids to high concentrations. However, little is known about the underlying mechanisms of carotenoid accumulation. Therefore, better understanding of the regulatory mechanisms involved in the β-carotene overproduction using metabolomics will facilitate the production of specific valuable carotenoids and other products in *D. salina* and in related organisms (Lamers et al., 2008).

The metabolic response of an organism to environmental stimuli or stressors have, to date, only rarely been carried out in algae. *C. reinhardtii* is a unicellular green algae and its the genome was recently sequenced. Recent report on the metabolite levels of *C. reinhardtii* under nutrient deprivation described a procedure for cell preparation and metabolite extraction (Bölling and Fiehn, 2005). When cells were grown under nutrient deficient conditions (depletion of nitrogen, phosphorus, sulphur or iron), highly distinct metabolic phenotypes were observed, as changes in the metabolites are first to be observed in relation to changes in the nutrients (Jamers et al., 2009a, Kluender et al., 2009b, Viant, 2007a). Recently, a working model for the coordinated regulation of cellular metabolism during the induction of the carbon concentrating mechanism in *C. reinhardtii* has been reported (Kempa et al., 2009b, Renberg et al., 2010). These observations underline the importance of understanding the metabolic pathways, and their underlying regulation. Metabolomics approaches in combination with proteomics and transcriptomics, can be used to identify differentially expressed metabolites and provide clues to rate-limiting processes in the cell. Thus systems biology approach will enable the fine tuning of algal properties by genetic or metabolite engineering (Schenk et al., 2008).

In systems biology, proteomics studies generates generic protein identification, however metabolomics studies has the ability to reveal that, the accumulated enzyme is more specifically related to a specific biochemical pathway. For an in-depth understanding of biology within given organism and at a given condition, it is essential to adopt metaproteogenomic strategy, which involves integration of both proteomics and metabolomics data. In Chapter 3, we have successfully demonstrated that, quenching methods are highly sample and cell dependent in case of the adherent mammalian cultures, and the optimised quenching protocols cannot be simply adopted for suspension cultures without validating them for a given case. Therefore, it is essential to develop optimised quenching protocols for a given biological sample, in this case bacterial and microalgal cultures. Followed by quenching, the major issue with metabolomics studies is to find the suitable extraction solvent that can quantitatively extract all the intracellular metabolites. For an unbiased analysis of the metabolome it is essential to extract all the metabolites completely, non-selectively and reproducibly by avoiding their degradation (e.g., chemical, thermal) and conversion to other metabolites (e.g., due to enzymatic activity).

The objective of this investigation is to develop an un-targeted GC-MS based metabolic profiling approach that can potentially be used to detect and quantify hundreds of metabolites with highest possible recoveries to obtain as much information as possible about key metabolites in microbial model *E. coli* and microalgal species such as *C. reinhardtii, N. salina and D. salina.* We have adopted few previously published quenching and extraction protocols in order to test the efficiency of a variety of workflows for GC-MS based un-targeted metabolomic studies and to identify major bottlenecks in the sample preparation steps. In addition, the optimised protocol for *C. reinhardtii* was used to study the influence of carbon source and availability on the metabolome and proteome profiles in *C. reinhardtii* cultures grown under different trophic conditions.

6.2 Materials and Methods

6.2.1 Bacterial/microalgal strain and growth conditions

a) Escherichia coli

The *E. coli* K12 strain MG1655 was cultured at 37°C in 500 mL shake flasks containing 100 mL of Luria-Bertani (LB) medium. After inoculation with a pre-culture in the same medium to an OD of

0.1 (wavelength = 600 nm using an Ultraspec 2100 Pro from Pharmica Biotech, Uppsala Sweden), cells were harvested in the stationary phase. Cell dry weight was determined simultaneously by centrifugation of 10 mL of medium containing cells at different optical densities. The pellets were re-suspended in 1 mL of sterile water and transferred to a new Eppendorf tube, and dried in vacuum concentrator. In order to obtain the dry weight the Eppendorf tubes were first weighed empty and after drying the cells.

b) *Chlamydomonas reinhardtii*

For optimisation of the quenching and extraction workflows

The *C. reinhardtii* strain CCAP 11/32C was obtained from the Culture Collection of Algae and Protozoa, UK. The cells were grown in a Sanyo incubator (MLR-351H, Sanyo Versatile Environmental Test Chamber, from Japan) at 25° C in 1 L shake flasks, containing 1 litre of TAP medium, under constant illumination at 85 µmol photons $m^{-2}s^{-1}$ and 25˚C. The medium (per litre) is composed of 2.42 g Tris; 25 ml TAP salts (1g $MgSO_4.7H_2O$, 0.5g $CaCl_2.2H_2O$ in 250 mL dH_2O); 25 ml Ammonia chloride solution (containing 0.75g NH_4CL in 50 mL dH_2O); 0.375 ml phosphate Solution (28.8g K_2HPO_4, 14.4g KH_2PO_4 in 100 mL dH_2O); 1.0 ml solution Hunter media (containing 2.2g $ZnSO_4.7H_2O$, 1.14g H_3BO_3, 0.506g $MnCl_2.4H_2O$, 0.161g $CoCl_2.6H_2O$, 0.157g $CuSO_4.5H_2O$, 0.11g $(NH_4)_6Mo_7O_{24}$ and 0.499g $FeSO_4.7H_2O$ in 100 mL dH_2O); and 1.0 ml glacial acetic acid. Cells were harvested in stationary phase (OD at 750 nm = 0.700). Cell dry weight measurement was performed simultaneously.

For biology based experiments

The *C. reinhardtii* strain CCAP 11/32CW15+ was obtained from the Culture Collection of Algae and Protozoa, UK. The culture was maintained, cultured and provided by Joseph Longworth for metabolomic investigations whereas the proteomic and biochemical analysis were simultaneously carried out by the Longworth and co-workers (Longworth, 2013). Cultures were grown in either TAP media as described above or TAP media without acetate and pH adjusted to 7 using HCl. Culture were grown in a Sartorius Biostat B plus bioreactor. Culture were maintained at OD of 0.450 ± 0.05 at 750nm for at least 12 h prior to any sampling event.

c) *Nannochloropsis salina*

The microalga strain used in this study was CCAP 849/1, which was obtained from the Culture Collection of Algae and Protozoa, UK. The cells were grown in Sanyo incubator (MLR-351H, Sanyo versatile environmental test chamber, from Japan) at 25°C in 1 L shake flasks containing 500 mL of f/2 medium with 33.6 g artificial sea salt per litre (Ultra Marine Synthetica Sea Salt, Waterlife). The medium (per litre) is composed of 33.6 g artificial seawater salts, 75 mg $NaNO_3$, 5.65 mg $NaH_2PO_4 \cdot 2H_2O$,1ml trace elements stock and 1 ml vitamin mix stock. The trace elemental solution (per litre) includes 4.16 g Na_2EDTA, 3.15 g $FeCl_3 \cdot 6H_2O$, 0.18 g $MnCl_2 \cdot 4H_2O$, 10 mg $CoCl_2 \cdot 6H_2O$, 10 mg $CuSO_4 \cdot 5H_2O$, 22 mg $ZnSO_4 \cdot 7H_2O$, 6 mg $Na_2MoO_4 \cdot 2H_2O$. The vitamin mix solution (per litre) includes 100 mg vitamin B1, 0.5 mg vitamin B12 and 0.5 mg biotin. Cells were harvested in stationary phase (OD at 680 nm = 0.729). Cell dry weight measurement was performed simultaneously.

d) Dunaliella salina

The microalga strain used in this study was CCAP 19/30 obtained from the Culture Collection of Algae and Protozoa, UK. The cells were grown at the similar conditions used for *N. salina*. Cells were harvested in stationary phase (OD at 680 nm = 0.519). Cell dry weight measurement was performed simultaneously.

6.2.2 Preparation of standard metabolites mixture

The amino acid kit (purchased from Sigma), was used to prepare a standard metabolite mixture, containing 50 µM of each metabolite, for quality control purposes and/or used as a pure reference compound for method validation and to assess analytical performance. Stepwise dilution of stock solution was done and was used as a reference compounds in the current research. The metabolite mixture contains L-arginine, L-aspartate, L-glutamate, L-glutamine, L-histidine, L-lysine, L-methionine, L-phenylalanine, L-threonine, Malic acid, Citric acid and Glucose (All obtained from Sigma). The mixture was derivatized prior to GC-MS analysis as detailed in section 6.2.5 c).

6.2.3 Sample extraction and quenching

Combinations of various quenching and extraction solvent previously reported for the species used in this investigation are listed in table 6.1.

Table 6.1 Combination of quenching and extraction solvents employed for various species. (M= methanol; C= chloroform; W= water)

Species	Quenching with	Extraction with
E. coli	1) 60 % aq. methanol	1)100% M
C. reinhardtii	1) 32.5% aq. methanol 2) 60% aq. methanol	1) M:C:W (10:3:1) 2) M:C:W (5:2:2) 3) 100% methanol
N. salina	1) 60 % aq. methanol	1) M:C (2:1) 2) 100% M
D. salina	1) 60 % aq. methanol	1) M:C (2.5:2)

a) E. coli

10 mL of culture samples in triplicate were rapidly plunged into 50 mL pre-chilled Falcon tube containing an equal volume of pre-chilled 60% aqueous methanol solution (-48°C) (Winder et al., 2008). The centrifuge was set at -9°C and the rotor was pre-chilled at -20°C. The quenched biomass was then centrifuged for 8 min at 2500*g* at -9°C. The supernatant was removed rapidly, and an aliquot (1 mL) was transferred to a 2 mL Eppendorf to assess the leakage of internal metabolites. The pellets and supernatant were snap frozen in liquid nitrogen and stored at -80°C for further analysis. The procedure was repeated simultaneously for the control sample, except for the quenching step.

b) C. reinhardtii

For optimisation of the quenching workflows

In the case of *C. reinhardtii* different protocols were applied for intracellular metabolite quenching and extraction. The two quenching solutions were tested in triplicate, quenching with 32.5% aqueous methanol (Renberg et al., 2010) and with 60% aqueous methanol (Lee and Fiehn, 2008b). At the incubation site 3 mL of cell suspensions were rapidly plunged into 50 mL pre-chilled Falcon tube containing 12 mL of pre-chilled quenching solution (-48°C). The centrifuge was set at -9°C and the rotor was pre-chilled at -20°C. The quenched biomass was then centrifuged for 5 min at 2500*g* at -9°C. The supernatant was removed rapidly, and an aliquot (1 mL) was transferred to a 2 mL Eppendorf to assess the leakage of internal metabolites. The pellets and supernatant were snap frozen in liquid nitrogen and stored at -80°C for further analysis.

For biology based experiments

For biology based investigations, at the incubation site 5 mL of cell suspensions were rapidly plunged into 15 mL pre-chilled tube containing 5 mL of pre-chilled quenching solution (-48°C). The remaining steps were carried out similarly as described above.

c) N. salina

In the case of *N. salina,* 60% aqueous methanol solution was used as a quenching solution, and another set was also tested simultaneous without any quenching step, to determine the efficiency of quenching in metabolomics. At the incubation site 5 mL of cell suspensions were rapidly plunged into 15 mL pre-chilled tube containing 5 mL of pre-chilled quenching solution (-48°C). The centrifuge was set at -9°C and the rotor was pre-chilled at -20°C. The quenched biomass was then centrifuged for 5 min at 2500*g* at -9°C. The supernatant was removed rapidly, and an aliquot (1 mL) was transferred to a 2 mL Eppendorf to assess the leakage of internal metabolites. The pellets and supernatant were snap frozen in liquid nitrogen and stored at -80°C for further analysis.

d) D. salina

In the current work with *D. salina,* 60% aqueous methanol solution was used as a quenching solution, whereas another set was also tested simultaneously without any quenching step, to determine the efficiency of quenching in metabolomics. For further treatment, similar protocol, as followed for *N. salina* was adopted.

6.2.4 Extraction of metabolites

a) E. coli

The extraction method was based on a previously published report (Winder et al., 2008). Briefly, to the snap frozen biomass, 500 µL of 100% methanol precooled at -80°C was added. The Falcon tube was then transferred to liquid nitrogen for 3 min, followed by thawing on dry ice for 10 min, and briefly vortexed. The freeze-thaw cycle was repeated three times for complete cell disruption. The sample was then centrifuged at 16000*g*, at -9°C for 5 min. The supernatant was transferred to a new tube, and the pellet was re-extracted twice with 500 µL of 100% methanol precooled at -80°C. The first and second aliquots were then combined and stored on dry ice. The 700 µL of extract was spiked with 100 µL of internal standard solution containing 0.19 mg m L^{-1} succinic d_4 acid, 0.27 mg m L^{-1} malonic-d_2 acid in HPLC grade water. The samples were then lyophilized overnight prior to derivatization.

b) C. reinhardtii

For optimisation of the extraction workflows

The samples were lyophilized at -50°C overnight prior to extraction. Three different extraction solvent systems were tested against two quenching solutions. The different extraction solvent systems employed were: methanol: chloroform: water (10:3:1), previously found to be most suitable for *C. reinhardtii* (Bölling and Fiehn, 2005), methanol: chloroform : water (5:2:2) as published for plant organs (Gullberg et al., 2004a, Weckwerth et al., 2004a) and 100 % methanol as published for microbial metabolomics (Winder et al., 2008). Solvent ratios are given as volumetric measures. Briefly, to lyophilized cells, 500 µL of extraction solvent was added, pre-chilled at -48°C, along with an equal volume of glass beads (425-600 µm i.d., acid washed, from Sigma). To the reaction mixture, 100 µL of the internal standard solution containing 0.19 mg m L^{-1} succinic d_4 acid, 0.27 mg m L^{-1} malonic-d_2 acid in HPLC grade water was added. Cells were then disrupted using (Cell disruptor, from Genie), 13 cycles of disruption was performed with 1 min disruption, with an interval of 1 min on ice. The sample was then centrifuged at 14,000 rpm, at -9°C for 10 min to remove any cell debris. The supernatant was transferred to new pre-chilled Eppendorf tube (-20°C) and evaporated to dryness using vacuum concentrator (5301 Vacuum Concentrator, from Eppendorf). The dried extract was then stored at -80°C for further analysis.

For biology based experiments

Extraction was carried out with methanol: chloroform: water (5:2:2) mixture. All further steps were carried out similarly as described above.

c) N. salina

Two different extraction solvent systems: methanol: chloroform (2:1) as published previously for Daphnia magna (Vandenbrouck et al., 2010b) and 100 % methanol as published for microbial metabolomics (Winder et al., 2008) were tested. Briefly, to the snap frozen biomass, 1.2 mL of extraction solvent was added, pre-chilled at -48°C. Cells were then disrupted using sonicator (450 Sonifier, from Branson), 15 cycles of sonication (Duty cycle 50%, noise 40 times) was performed with 1 min sonication, with an interval of 1 min on ice. After this time, 400 µL of chloroform and 400 µL of ultrapure water was added to the reaction mixture. The sample was then centrifuged at 14,000 rpm, at -9°C for 15 min to separate polar and non-polar phase. The resultant aqueous and organic fractions as well as pellet were transferred to a new pre-chilled Eppendorf tube (-20°C).150 µL of each extract is then evaporated to dryness using vacuum concentrator (5301 vacuum concentrator, from Eppendorf). The dried extract was then stored at -80°C for further analysis.

d) D. salina

The extraction method was based on a previously published report (Lamers et al., 2010b). Briefly, to the snap frozen biomass, 4 mL of chloroform: methanol (2:2.5) pre-chilled at -80°C was added. To the reaction mixture, 66 µg ml^{-1} of glyceryl trinonadecanoate was added as an internal standard and vortexed for 1 min. Cells were then disrupted using sonicator (450 Sonifier, from Branson), 15 cycles of sonication (Duty cycle 50%, noise 40 times with output control 6) was performed with 1 min sonication, with an interval of 1 min on ice. After sonication, 2.5 mL of a 50 mM Tris-buffer (pH 7.5) containing 1M NaCl was added to the suspension. The samples were vortexed and sonicated once more and subsequently centrifuged at 14,000 rpm, at -9°C for 15 min to separate the polar and non-polar phases. The non-polar phase was transferred to a new Eppendorf tube and the polar phase and cell debris were re-extracted with 1 mL chloroform. The first and second chloroform fractions were combined and 150 µL of the pooled chloroform fraction and polar phase were evaporated to dryness using vacuum concentrator (5301 vacuum concentrator, from Eppendorf). The dried extract was then stored at -80°C for further analysis.

6.2.5 Derivatization of metabolites

a) E. coli

Metabolite derivatization was performed as suggested elsewhere (Winder et al., 2008). Briefly, to the lyophilized sample, 40 µL of 20mg/mL methoxyamine hydrochloride in pyridine was added and samples were shaken for 80 min at 40°C. The samples were then derivatized by trimethylsilylation by the addition of 40 µL MSTFA (N-methyl-N-trimethylsilyltrifluoroacetamide) and incubating them further in shaking condition at 37°C for 80 min. A retention index solution was added for the chromatographic alignment prior to analysis by GC-MS. Chromatographic alignment approach usually account for the drifts in retention times caused by the alteration in the column selectivity which is directly dependent on various factors which includes: temperature fluctuation, irreversible adsorption, column degradation and machine drifts. The alignment is performed in a *z*-score transformed retention time domain to standardize chromatographic peak distribution across samples using a retention index solution such as mixture of n-alkanes. A mixture score is developed to assess the similarity between two peaks by simultaneously evaluating the mass spectral similarity and the closeness of retention time.

b) C. reinhardtii

Metabolite derivatization was performed as suggested elsewhere (Lee and Fiehn, 2008b). Briefly, to the dried extract, 30 µL of 20mg/mL methoxyamine hydrochloride in pyridine was added and samples were shaken for 90 min at 30°C to protect aldehyde and ketone groups. The samples were then derivatized by trimethylsilylation by the addition of 45 µL MSTFA (N-methyl-N-trimethylsilyltrifluoroacetamide) and incubating them further in shaking condition at 37°C for 30 min. A retention index solution was added for the chromatographic alignment prior to analysis by GC-MS.

c) N. salina

Metabolite derivatization was performed as suggested elsewhere (Vandenbrouck et al., 2010b), with little modification. Briefly, to the dried extract, 30 µL of 20mg/mL methoxyamine hydrochloride in pyridine was added and samples were shaken for 90 min at 30°C to protect aldehyde and ketone groups. The samples were then derivatized by trimethylsilylation by the addition of 30µL MSTFA (N-methyl-N-trimethylsilyltrifluoroacetamide) and incubating them further in shaking condition at 37°C for 30 min. A retention index solution was added for the chromatographic alignment prior to analysis by GC-MS.

d) D. salina

Similar protocol, as followed for *N. salina* was adopted, See section 6.2.5 (c)

6.2.6 GC-MS for metabolite analysis

GC-MS analysis was performed as described in Chapter 3, section 3.2.6

6.2.7 Identification of metabolites and data analysis

Identification of metabolite and further data analysis was performed as described in Chapter 3, section 3.2.7 and 3.2.8.

6.3 Results and Discussion

6.3.1 Validation of the use of GC-MS for metabolomics

The objective of this study was to validate a GC-MS method that can be used to study metabolites found in algae. Known metabolite mixtures provided a means of assessing the performance and reliability of algorithms, databases and protocols used in MS-based metabolite identification and quantification. Indeed, the use of standardized metabolite mixtures could be used to provide much-needed validation of existing methods and deconvolution softwares.

The AMDIS software for deconvolution of metabolomics data sets based on GC-MS was evaluated. Attention was paid to the extent of detection, identification and agreement of qualitative results. In addition flexibility and the productivity of this deconvolution software program in their application was also monitored. To achieve this objective, we performed the analysis of test-mixture solution containing 50 µM of each of the 12 metabolites. The resulting mass spectra, after deconvolution using the AMDIS search program, were compared with the NIST mass spectral libraries containing both mass spectra and the retention indices of derivatives. Depending upon the various parameters used the results showed a large number of false positives (data not shown). It was found that the component width, match factor, resolution and sensitivity are the most important parameters for predicting the accuracy of the deconvolution results.

AMDIS deconvolutes 87 components and 28 target peaks as displayed in Appendix 6.2, out of 12 metabolite derivatives with the set of parameters mentioned above. It was noted that the number of peaks detected does not equal the number of metabolites before derivatization, as the method of derivatization produces more than one derivative of a single metabolite. In addition one single peak was deconvoluted as multiple components. So taking this into consideration, the parameters were optimised. The number of components detected by AMDIS were closer to the true number of metabolites derivatives in solution, which made it easier to define the true metabolites in real samples analyzed later in this current research. However only 11 metabolites out of 12 were detected, and futher analysis revealed that arginine was not detectable (table 6.2). Previous studies have shown that, arginine is not detectable by GC-MS analysis, as silylation reactions converts arginine to ornithine or its degradation products, which were difficult to analyze by GC-MS (Bedair and Sumner, 2008a, Dettmer et al., 2007). In addition, histidine was found to be degraded to putresine. Still there is a not a good balance between avoiding false positive and negatives. Therefore the use of AMDIS for deconvoluting spectra requires further validation and optimisation.

Table 6.2 List of metabolites identified by GC-MS along with their retention time from twelve metabolite mixture

Number	Metabolites	R.T.
1	Citric acid (4TMS)	25.027
2	Glucose methoxyamine (5TMS)	26.118
3	L-Aspartic acid (3TMS)	19.475
4	L-Glutamic acid (3TMS)	21.496
5	L-Glutamine (3TMS)	18.683
6	L-Lysine (4TMS)	26.751
7	L-Methionine (2TMS)	19.547
8	L-Phenylalanine (2TMS)	21.7
9	L-Threonine (3TMS)	16.428
10	Malic acid (3TMS)	18.805
11	Putrescine (4TMS)	17.35

Analysis of repeatability: The analysis was performed three times to test the repeatability of the software, which resulted in repeatable result, thus showed the reliability of experimental method was very high.

6.3.1.1 Amino acid analysis for quality control of GC-MS based metabolomics

It is essential to check the consistency of analytical results before making any biological interpretation. The purpose of this study was to detect and/or correct for detector drift and to control the inertness of the GC-MS technique. The amino acid mixture was treated as a QC sample (normalization step) and used as an early marker for the decline of the performance of the analytical system, as metabolites are more prone to adsorb or degrade on the surface of the analytical column in the absence of sample matrix. The QC sample was injected periodically and used to define intra-batch response curves that can then be used to adjust the response of the analytical sample by assessing the performance of chromatographic column and the MS detector during the experiments.

6.3.2 Optimization of the quenching and extraction workflow for GC-MS based metabolomic investigation of one microbial and three microalgal species

In this section, we have evaluated the efficient applications of metabolomic workflow (as it was not established before in our laboratory) for handling suspension cultures. The idea behind these investigations was to assess the bottlenecks in metabolomics sample preparation workflow for handling suspension cultures, and address them in the subsequent Chapters.

6.3.2.1 Overview of E. coli metabolomics data

The metabolome of microbe, while not as complex as that of a plant and microalgae, still contains at least several hundred metabolites. In this research work, *E. coli* was selected for preliminary metabolomics study, due to its importance as both model organism, human pathogen and ease of growth. The optimized analytical protocol suggested elsewhere (Winder et al., 2008) was applied for the comprehensive metabolite profiling of *E. coli* K12 strain MG1655. Two biological replicates were analysed, with three analytical replicates taken from each culture. For sampling and quenching 60% aqueous methanol was used, while for extraction 100% methanol was used as suggested elsewhere (Winder et al., 2008), followed by GC-MS analysis. The procedure was repeated simultaneously for the control sample, except the quenching step.

A representative chromatogram from *E. coli,* GC-MS metabolomic analysis is shown in figure 6.1. The number of metabolites identified by GC-MS in each replicates can be seen in table 6.3;

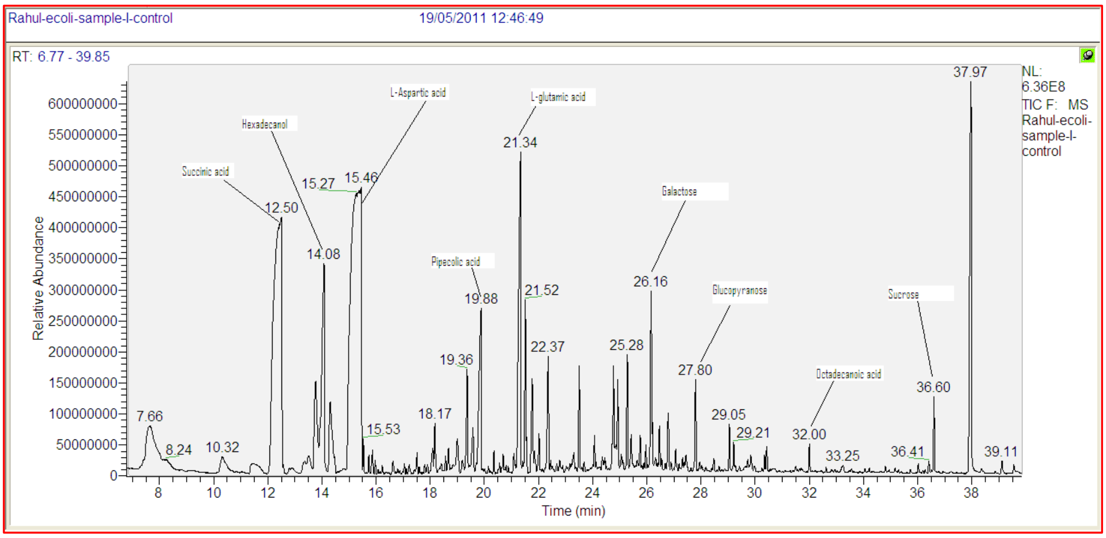
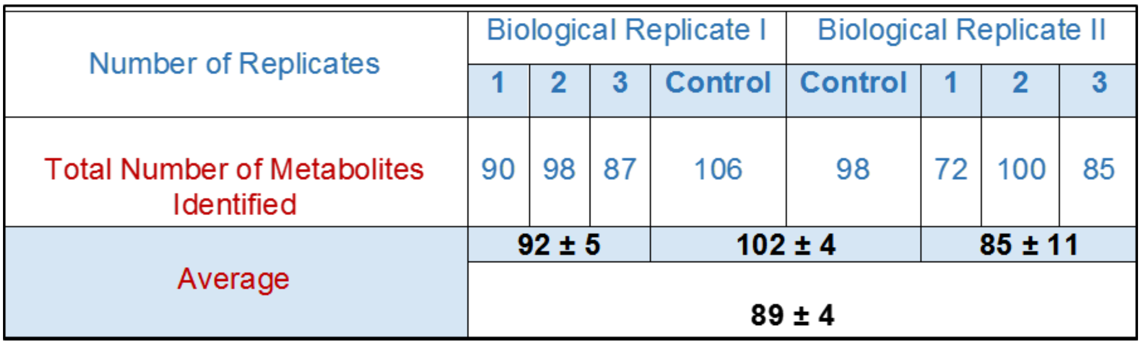


Figure 6.1 Representative chromatogram from *E. coli*, GC-MS metabolomic analysis. The sample was quenched with 60% aqueous methanol and extracted with pure methanol. The extracted metabolites were derivatized by methoximation followed by silylation. The derivatized sample volume of 1 µL was injected in split less mode at 230°C and separated on a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). In total, around 162 metabolites were identified (see table 6.8), where few identified peaks were labelled onto a chromatogram.

Table 6.3 Number of metabolites identified in each replicate in *E. coli* metabolomic study (1, 2, 3 = analytical replicates; Control = Unquenched sample)



Number of Replicates	Biological Replicate I				Biological Replicate II			
	1	2	3	Control	Control	1	2	3
Total Number of Metabolites Identified	90	98	87	106	98	72	100	85
Average	**92 ± 5**			**102 ± 4**		**85 ± 11**		
	89 ± 4							

We compared our *E. coli* metabolite data with a previously published optimized method for global metabolite profiling of *E. coli* cultures (this will be referred to henceforth as previous study) (Winder et al., 2008). In the previous study, the number of peaks detected with the methanol quenching solution was reported as 152. The number of metabolites identified in our research was 89 ± 4. The possible reasons for the lower number identified by us as compared to previously published method are listed below;

1) The analysis was carried out with GC-ToF-MS, whereas our results were based on conventional quardupole GC-MS. As mentioned in section 6.1, GC-ToF-MS offers increased productivity with higher mass accuracy and mass resolution as compared to conventional GC-MS method

2) The metabolites were detected in both positive and negative mode, whereas the number of reported metabolites in this research was based only on analysis in the positive ion mode.

3) The leakage of internal metabolites was assessed, and accurately balanced to calculate final number of metabolites, whereas in this study, the leakage of internal metabolites was not assessed and corrected.

However, in our study we concentrated more on the choice of the GC column, which does not appear to be considered in the previous studies. The GC-column used in this research study was TR-5-MS, which was selected purposely for metabolomics studies because of the diversity of metabolome. The column characteristics such as column phase, internal diameter, film thickness and length were properly considered while developing the method in order to determine the best column for separation. The column selection guide was referred prior to selection of appropriate GC column for metabolomics (http://www.polygen.com.pl/thermo/download/chromatography-catalog-column-selection.pdf). Very few studies (Al Awam et al., 2014, Jones and Hügel, 2013, Liu et al., 2014) employed this column for GC-MS based metabolomic investigations, whereas other protocols seems to be simply adopted from the literature without critically validating them for the given case. In addition, the extraction methods considered in the previous study were not suitable for extraction of highly lipophilic metabolites. Furthermore, GC-MS is not ideally suited to all classes of metabolites such as nucleotides and sugar phosphates, which are extremely labile. Therefore, identifying and addressing the bottlenecks in quenching and extracting workflow with parallel application of GC-ToF-MS (in both positive and negative mode with TR-5-MS column for GC separation) and available LC-MS techniques might improve the metabolome coverage of *E. coli* further.

Qualitative analysis of *E. coli* metabolome coverage

The total number of unique metabolites identified in this research work with *E. coli* was 162 including quenched and non-quenched sample. Summary of the number of metabolites detected in each class are presented in table 6.8. Out of these, the total number of metabolites identified with quenching protocol was 142; whereas numbers of metabolites identified were reduced to 118 in case of non-quenched sample which was processed without quenching step. We also attempted to find out metabolites which are unique and those that are common to the different treatments. The results showed that, out of the 162 metabolites identified, 66 metabolites were found to be identified under only one particular treatment, whereas 96 metabolites were found to present in both the treatments.

6.3.2.2 Overview of N. salina metabolomics data

For this study, *N. salina* which is a small yellow-green heterokont alga was selected because of its high fatty acid content and its potential for use as a biodiesel producer. Investigation of its metabolism through metabolomic study will help us in the development of strategies for increasing the lipid content and in understanding CO_2 fixation mechanism, which can be then enhanced by controlling the environmental conditions. To date, there is no evidence of metabolomic study with *N. salina.* So the protocols for quenching and extraction were adopted from previously published studies on other microalgae as mentioned in section 6.2.3 (c) and 6.2.4 (c) were tested. In addition to this, different combinations of quenching and extraction solvents were also tested. Two biological replicates were analysed, with three replicates taken from each culture.

A representative chromatogram from *N. salina,* GC-MS metabolomic analysis is shown in figure 6.2. The number of metabolites identified by GC-MS in each replicates can be seen in table 6.4.

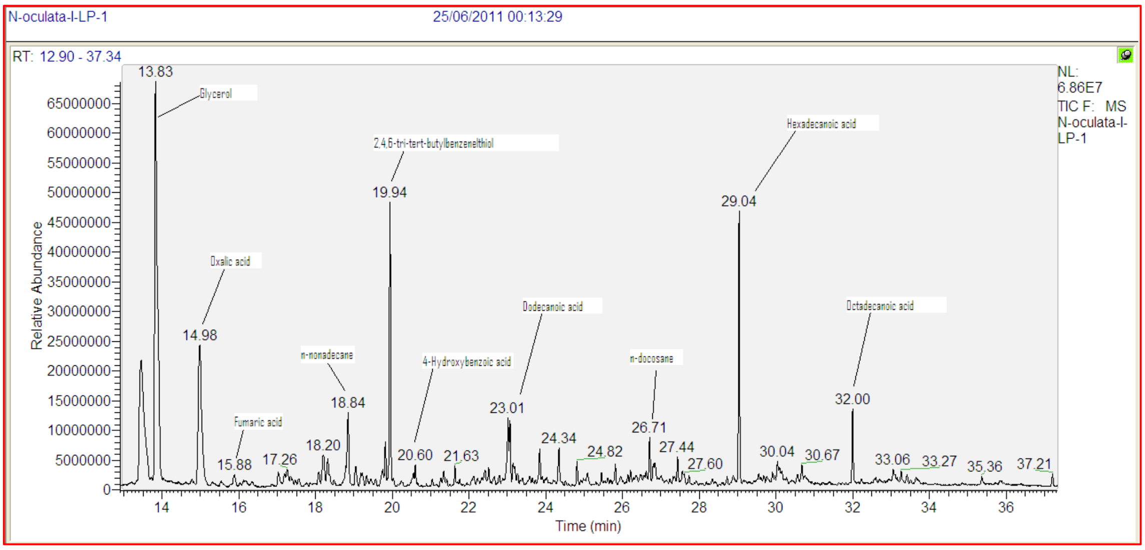


Figure 6.2 Representative chromatogram from *N. salina*, GC-MS metabolomic analysis. The sample was quenched with 60% aqueous methanol and extracted with methanol-chloroform mixture (2:1 *v/v*). The extracted metabolites were derivatized by methoximation followed by silylation. The derivatized sample volume of 1 µL was injected in split less mode at 230°C and separated on a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). In total, around 116 metabolites were identified (see table 6.8), where few identified peaks were labelled onto a chromatogram.

Table 6.4 Number of metabolites identified in each replicate in *N. salina* metabolomics study: A, B and C = different workflows, where A = no quenching and extraction with methanol-chloroform mixture (2:1 *v/v*); B = quenching with 60% aqueous methanol and extraction with methanol-chloroform mixture (2:1 *v/v*) and C = quenching with 60% aqueous methanol and extraction with pure methanol; whereas 1, 2 and 3 = biological replicates

Number of Replicates	SET A			SET B		SET C		
	1	**2**	**3**	**1**	**2**	**1**	**2**	**3**
Lower Phase (Non-polar)	36	32	31	40	37	42	39	41
Upper Phase (Polar)	36	43	50	56	56	32	62	19
Total Number of Metabolites Identified (Excluding common)	54 (18)	50 (17)	61 (19)	74 (19)	71 (19)	65 (9)	73 (9)	49 (11)
Average	**55 ± 4**			**73 ± 2**		**62 ± 10**		

Among different quenching and extraction methods employed in this study, we proposed that quenching with 60% aqueous methanol and extraction with methanol: chloroform (2:1) (Set B) is the most suitable for metabolomic study of *N. salina.* The metabolites were identified in both the polar and non-polar phases. The total numbers of metabolites identified in both polar and non-polar phases by GC-MS are shown in table 6.4. The omission of the quenching step (Set A) resulted in a lower number of metabolites detected and extraction with 100% methanol (Set C) resulted in detection of lesser number of metabolite as compared to that of extraction with methanol: chloroform (2:1). As there are no previously published studies with *N. salina* in the literature, we cannot compare our data with other studies, however the adoption of similar strategy such as, parallel application of 2D GC-ToF-MS (in both positive and negative mode with TR-5-MS column for GC separation) and available LC-MS techniques in future along with the appropriate correction for leaked intracellular metabolites will surely increase the number of metabolites identified as compared to recently adopted method in this study.

Qualitative analysis of *N. salina* metabolome coverage

The total number of unique metabolites identified in this research work with *N. salina* was 116 (including all samples from three various treatments). Out of these, total number of metabolites identified with quenching protocol (Set B and C) was much higher (around 90) than that of identified with no quenching (Set A) (only 72 identified). The individual treatment results showed that treatments given in Set B are superior to others, with which 90 unique metabolites were identified. Further data analysis showed that 50 metabolites were found to be identified in all the given treatments, 28 found to be identified in only two treatments given, whereas 41 metabolites were found to be identified only under one treatment. Summary of the number of metabolites detected in each class could be found in table 6.8.

6.3.2.3 Overview of D. salina metabolomics data

For this study, *D. salina,* a unicellular green halophilic alga, was selected because it is used for production of β-carotenes, chlorophylls and xanthophylls such as lutein, zeaxanthin and neoxanthin. Further investigation of its metabolism through metabolomic study will help us increase carotenoid accumulation which are powerful antioxidants used in human diet supplements and colorants used as fish food pigments (Denery et al., 2004). The protocol for quenching and extraction was adopted from previously published studies as mentioned in section 6.2.3 (d) and 6.2.4 (d). Two biological replicates were analysed, with three replicates taken from each culture. A representative chromatogram from *D. salina,* GC-MS metabolomic analysis is shown in figure 6.3. The number of metabolites identified by GC-MS in each replicates can be seen in table 6.5;

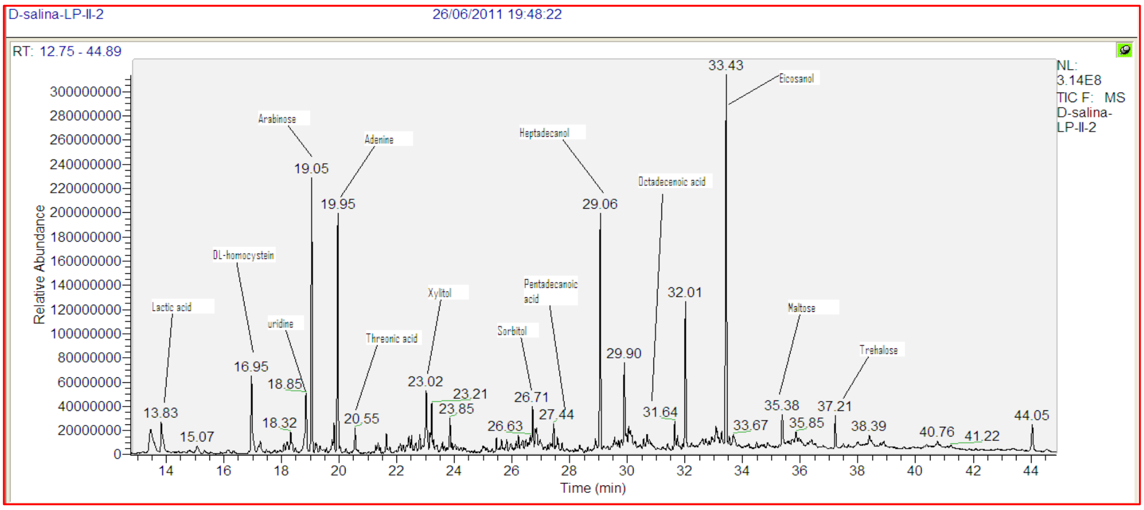


Figure 6.3 Representative chromatogram from *D. salina*, GC-MS metabolomic analysis. The sample was quenched with 60% aqueous methanol and extracted with methanol-chloroform mixture (2.5:2 *v/v*). The extracted metabolites were derivatized by methoximation followed by silylation. The derivatized sample volume of 1 µL was injected in split less mode at 230°C and separated on a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). In total, around 103 metabolites were identified (see table 6.8), where few identified peaks were labelled onto a chromatogram.

Table 6.5 Number of metabolites identified in each replicate in *D. salina* metabolomic study (A, B = Different quenching and extraction protocols, where A = No quenching, B = quenching with 60% aqueous methanol followed by extraction with methanol/chloroform (2.5:2); whereas 1, 2, 3 = Analytical replicates)

Number of Replicates	SET A			SET B	
	1	**2**	**3**	**1**	**2**
Lower Phase (Non-polar)	41	37	35	33	48
Upper Phase (Polar)	60	53	55	70	68
Total Number of Metabolites Identified (Excluding common)	75	64	66	76	88
Average	**68 ± 5**			**82 ± 6**	

From results, it can be seen that the number of metabolites identified with incorporation of quenching step is higher as compared to method without quenching. The total number of metabolites identified in this study was 82 ± 6. Very few published reports are there in the literature on metabolomics study of *D. salina*. However no attempts were made in this study to improve the coverage of metabolome and protocol was simply adopted from previously published study. The objective in this case was to construct metabolite specific library for *D. salina*.

Qualitative analysis of *D. salina* metabolome coverage

The total number of unique metabolites identified in this research work with *D. salina* was 103 (including all samples from two various treatments). Overall comparison, negligible amount of variation was found in results as analyses showed that nearly equal numbers of metabolites were identified in both treatments 91 in Set A (without quenching) and 92 in Set B (with quenching step). However it should be noted that the leakage of internal metabolites during quenching is yet to be quantified for Set B, which may lead to results having greater variation in numbers. Further analysis of data showed that 56 unique metabolites were found to be identified under both the treatments whereas 47 metabolites were identified under only one treatment. Summary of the total number of metabolites detected in each class could be found in table 6.8.

6.3.2.4 Overview of C. reinhardtii metabolomics data

Among photosynthetic eukaryotes, *C. reinhardtii* was selected for this research work as it is a sequenced organism and offers unique opportunities for the generation of high-quality metabolomics data. The optimized analytical protocols suggested elsewhere (Lee and Fiehn, 2008b, May et al., 2008b, Renberg et al., 2010) were adopted in this study as mentioned in section 6.2.3 (b) and 6.2.4 (b). Two biological replicates were analysed, with three technical replicates taken from each culture. A representative chromatogram from *C. reinhardtii,* GC-MS metabolomic analysis is shown in figure 6.4. The average number of metabolites identified by GC-MS with two different quenching and extraction protocols are summarized in table 6.6;

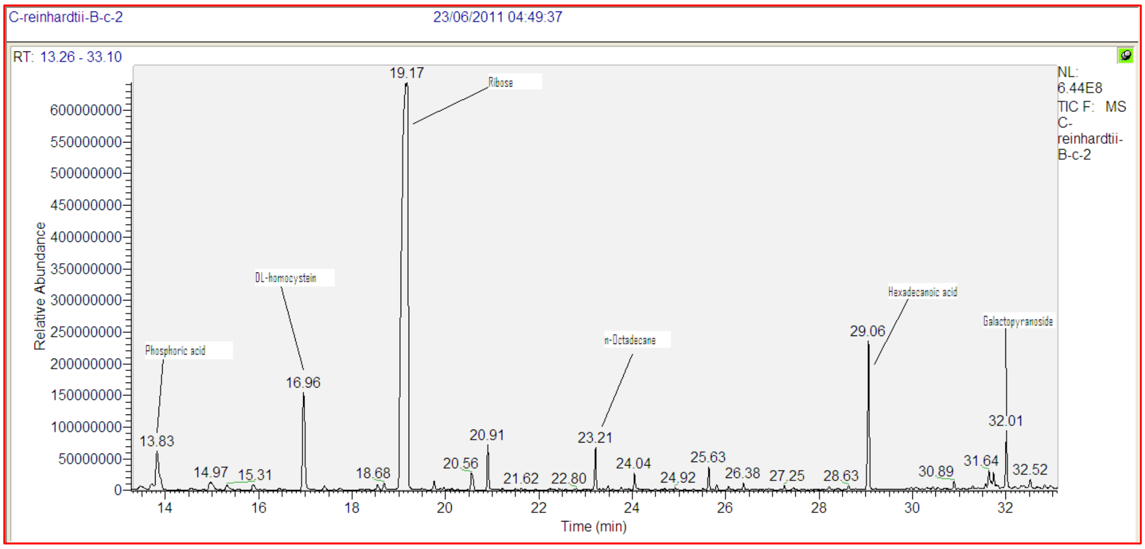


Figure 6.4 Representative chromatogram from *C. reinhardtii*, GC-MS metabolomic analysis. The sample was quenched with 60% aqueous methanol and extracted with methanol-chloroform-water mixture (5:2:2 *v/v*). The extracted metabolites were derivatized by methoximation followed by silylation. The derivatized sample volume of 1 µL was injected in split less mode at 230°C and separated on a TRACE TR-5MS capillary column (30 m x 0.25 mm x 0.25 µm). In total, around 139 metabolites were identified (see table 6.8), where few identified peaks were labelled onto a chromatogram.

Table 6.6 Number of metabolites identified in *C. reinhardtii* metabolomics study (A, B = Different quenching protocols, where A = quenching with 32.5% aqueous methanol, B = quenching with 60% aqueous methanol followed by extraction with a, b or c; where a = extraction with methanol/chloroform/water mixture (10:3:1), b = extraction with methanol/chloroform/water mixture (5:2:2) and c = extraction with 100% methanol)

Quenching treatment	**A**			**B**		
Extraction treatment	**A-a**	**A-b**	**A-C**	**B-a**	**B-b**	**B-c**
Average recoveries in numbers	57 ± 8	67 ± 2	74 ± 6	74 ± 4	77 ± 2	63 ± 16

From the results, it is clear that quenching with 60% aqueous methanol and extraction with methanol: chloroform: water (5:2:2) (B-b), is most suitable for metabolome analysis of *C. reinhardtii* as it resulted in highest recoveries in numbers (77±2) with good reproducibility's compared to other applied treatments. The total number of unique metabolites identified in this research work with *C. reinhardtii* was 139 (including all the samples with various treatments)
In past, May and co-workers reported identification 159 known metabolites using GCxGC-ToF-MS analysis, among which 119 metabolites were actually identified from *C. reinhardtii* culture, whereas the rest metabolites were spiked into the samples (May et al., 2009, May et al., 2008a). Authors reported two fold increase in number of metabolites identified compared to that reported by Bölling and Fiehn (Bölling and Fiehn, 2005), as they employed 2D GC-ToF-MS with a cold injection system (in order to protect thermo labile metabolites from undergoing degradation or inter-conversions) instead of GC-ToF-MS. Out of 119 identified metabolites, 24 metabolites were identified only by Bölling and Fiehn (Bölling and Fiehn, 2005), 41 metabolites were identified by both research groups (Bölling and Fiehn, 2005, May et al., 2008a), whereas 57 newly identified metabolites were reported only by May and co-workers (May et al., 2008a). All 159 identified metabolites from both studies (Bölling and Fiehn, 2005, May et al., 2009, May et al., 2008a) were then integrated along with the identified proteins from the proteomic investigations and were successfully mapped using the MapMan visualization platform as displayed in figure 6.5. MapMan is a visualisation tool that has been developed for the display of metabolite and transcript data onto metabolic pathways of Arabidopsis and other plant genomes and thus features a special emphasis on plant-specific pathways (May et al., 2008a).

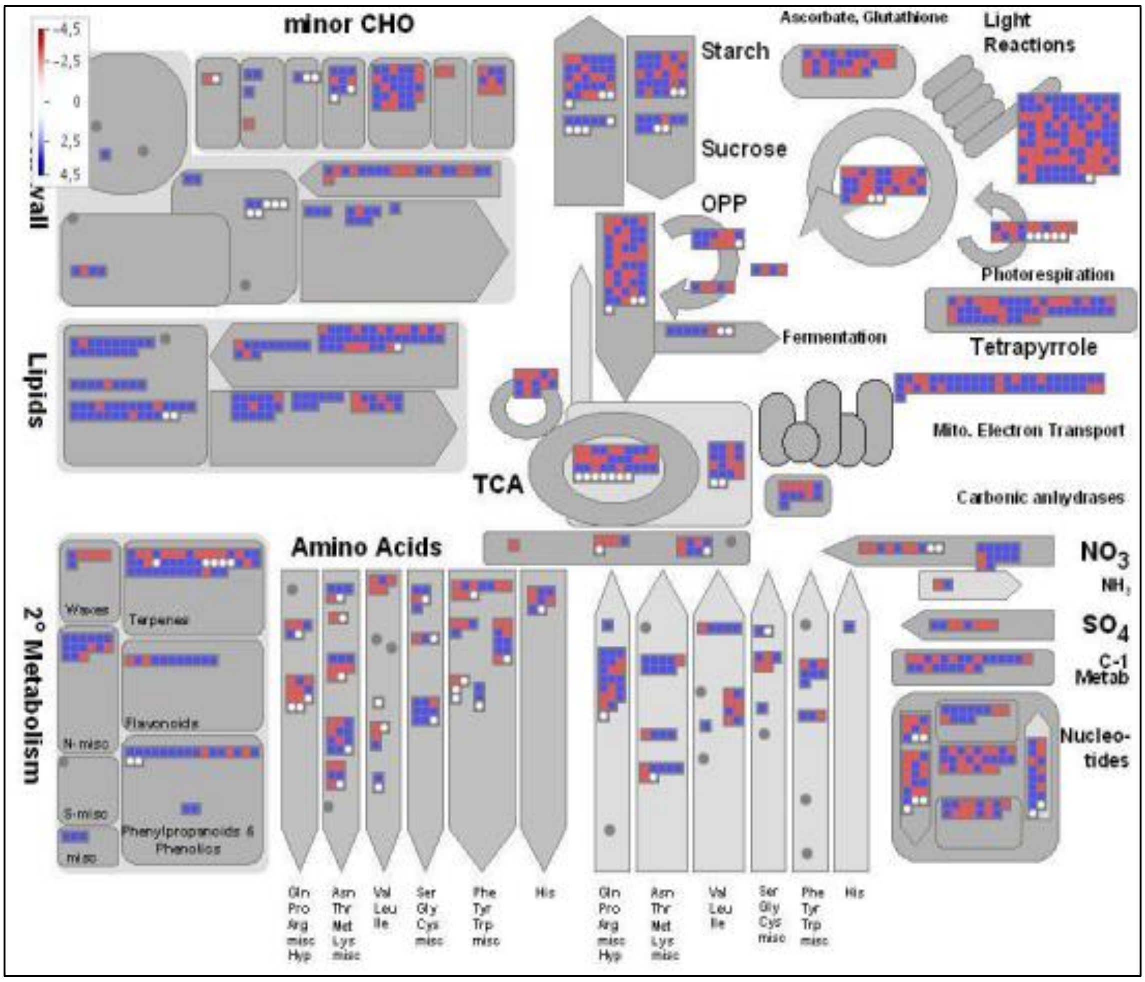

-4,5
-2,5
0
2,5
4,5
minor CHO
Cell wall
Starch
Sucrose
Ascorbate, Glutathione
Light Reactions
OPP
Photorespiration
Fermentation
Tetrapyrrole
Lipids
TCA
Mito. Electron Transport
Carbonic anhydrases
Amino Acids
NO3
NH3
SO4
C-1 Metab
Nucleo-tides
2° Metabolism
Waxes
Terpenes
Flavonoids
N-misc
S-misc
Phenylpropanoids & Phenolics
misc
Gln Pro Arg misc Hyp
Asn Thr Met Lys misc
Val Leu Ile
Ser Gly Cys misc
Phe Tyr Trp misc
His

Figure 6.5 *C. reinhardtii* central metabolism is visualised using the MapMan tool. Squares represent Chlamydomonas proteins that have been assigned into various pathways. White boxes indicates 159 identified metabolites, red boxes indicated 1069 identified proteins by MS, whereas blue boxes indicates proteins with no peptide support. Figure adopted from (May et al., 2008a).

We have compared the overall recovery of unique metabolites identified in our study against the metabolite database generated by May and co-workers (May et al., 2009, May et al., 2008a), who compared his method with that of Bölling and Fiehn (Bölling and Fiehn, 2005). The results of this investigations are summarized in figure 6.6, where we have generated Venn diagram for only those metabolites that were found to be commonly identified with the previously published reports. In addition, the newly identified metabolites that were not identified by both the previous studies and therefore not included in the MapMan visualization platform were represented by the pink colour.

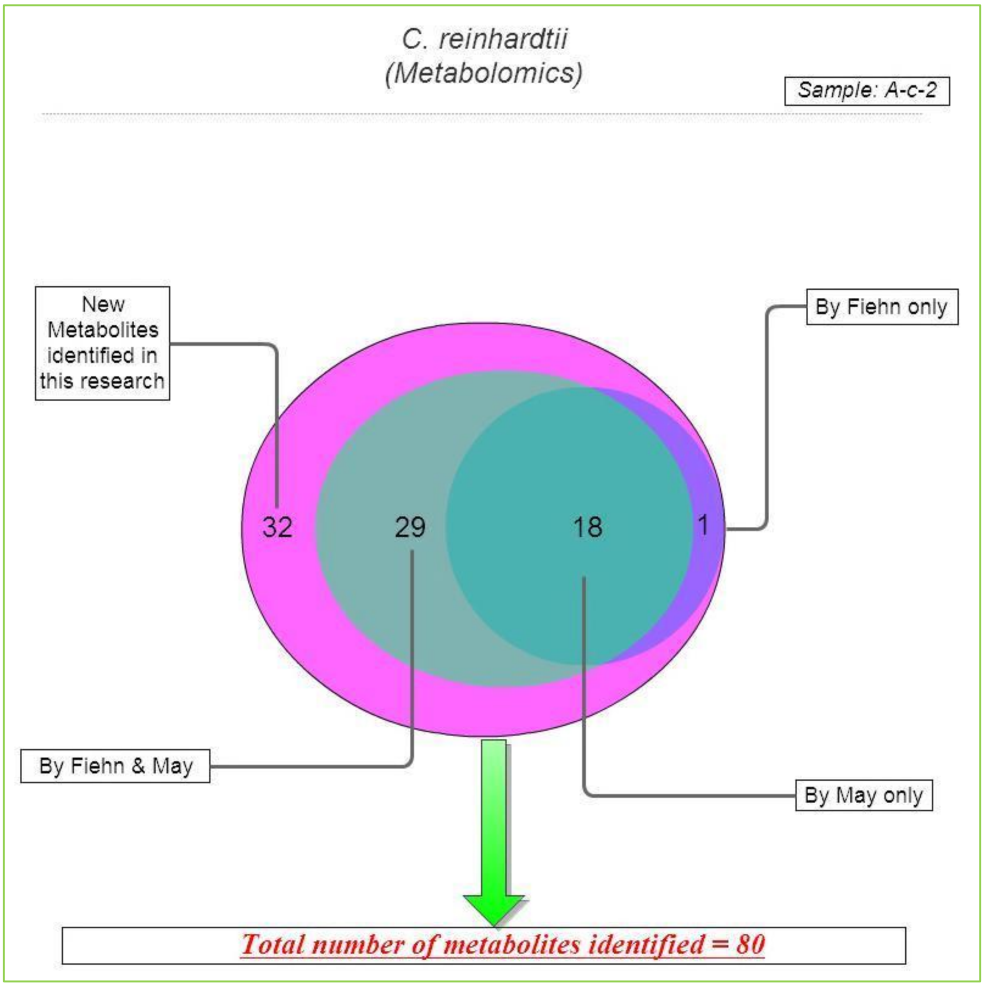

C. reinhardtii
(Metabolomics)
Sample: A-c-2
New Metabolites identified in this research
By Fiehn only
32
29
18
1
By Fiehn & May
By May only
Total number of metabolites identified = 80

Figure 6.6 Venn diagram displaying the capture of *C. reinhardtii* metabolome in our study in comparison with previously published reports by various researchers (Bölling and Fiehn, 2005, May et al., 2009, May et al., 2008a). Dark blue colour represents 18 metabolites identified by May and co-workers alone (May et al., 2008a), purple colour represents addition of one new metabolite reported by Fiehn and co-workers (Bölling and Fiehn, 2005) to that of 18 metabolites reported by May and co-workers, light blue colour represents 29 new metabolites identified in later studies by both authors in addition to those 18 that were reported by May and co-workers in their previous studies (May et al., 2009), whereas pink colour represents 32 new identified metabolites (see Appendix 6.3) in current study in addition to all of those previously reported.

From figure 6.6, it can be seen that, in total 19 metabolites were identified by Bölling and co-workers (Bölling and Fiehn, 2005) along with all the 18 metabolites that were reported by May and co-workers (May et al., 2009, May et al., 2008a) in their previously published report. Later both researchers conducted collaborative investigation on capturing the *C. reinhardtii* metabolome and reported recoveries of in total 47 metabolites, which were then incorporated in the metabolic network of primary metabolism for *C. reinhardtii*. Our GC-MS based analysis covered all those metabolites identified in previously published reports, in addition we have captured 32 new metabolites (See Appendix 6.3). Despite capturing 32 new metabolites, the metabolome coverage demonstrated in our study was not comparable to that of previously published reports, as authors identified 119 (May et al., 2009, May et al., 2008a) and 100 (Bölling and Fiehn, 2005) metabolites, whereas we managed to cover only 80. The possible reasons for a lower number of metabolites detected in our research work as compared to previously published methods are as discussed in section 6.3.2.1. In contrast, the number of new metabolites identified in this study was quite high, which were not reported with the previously published benchmark methods and further validation of this data would lead to publishable results.

The metabolic network of primary metabolism for *C. reinhardtii,* was constructed recently using genomic and biochemical information, which include 484 metabolic reactions and 458 intracellular metabolites (Boyle and Morgan, 2009). This suggests that we currently have identified only one fifth of the *C. reinhardtii* metabolome (with 139 identified metabolites) and there is considerable scope for improvement.

Qualitative analysis of *C. reinhardtii* metabolome coverage

The total number of unique metabolites identified in *C. reinhardtii* was 139 (including all samples with various treatments). In overall comparison with respect to quenching solvent employed, negligible amount of variation was found in results, as analysis showed that quenching with 60% aqueous methanol (Set A) leads to identification of 118 metabolites, whereas quenching with 32.5% aqueous methanol resulted in identification of 116 metabolites. However it should be noted that the leakage of internal metabolites during quenching is yet to be quantified for both the sets, which may lead to results having greater variation in numbers. If we consider individual treatments, the results clearly suggest that quenching with 60 % aqueous methanol and extraction with M: C: W (5:2:2) (Set B-b) is superior to other treatments and with which 96 unique metabolites were identified (table 6.7). In order to provide further qualitative information, we grouped all the identified metabolites in a way which will help us in sorting those which are identified only under one treatment, two treatments and so on up to six different treatments. Analysis showed that 26 metabolites were found to be present only under one treatment whereas 48 metabolites were found to be identified with both treatments as detailed in table 6.6. Summary of the number of metabolites detected in each class could be found in table 6.8.

Table 6.7 Number of unique metabolites identified in *C. reinhardtii* under each treatment, where: A, B = Different quenching protocols, where A = quenching with 32.5% aqueous methanol; B = quenching with 60% aqueous methanol followed by extraction with a, b or c; where a = extraction with methanol/chloroform/water mixture (10:3:1); b = extraction with methanol/chloroform/water mixture (5:2:2) and c = extraction with 100% methanol

SET	**Unique metabolites identified**
A-a	83
A-b	84
A-c	90
B-a	92
B-b	96
B-c	94
Set A	**116**
Set B	**118**
In Total	**139**

Table 6.8 Summary of the total number and percentage of metabolites detected in each class across all the species investigated in this study

Metabolite class	***E. coli***		***C. reinhardtii***		***N. salina***		***D. salina***	
	Number	%	Number	%	Number	%	Number	%
Organic acids	43	26.5	31	22.3	25	21.5	20	19.4
Sugars or Sugar-alcohols or Polyols	27	16.6	22	15.8	18	15.5	20	19.4
Amino acid and its derivatives	27	16.6	19	13.6	16	13.7	13	12.6
Others	24	14.8	23	16.5	21	18	15	14.5
Alcohols or Fatty alcohols	10	6.1	7	5	8	7	6	5.8
Biogenic amines	8	5	5	3.5	4	3.4	1	0.9
Fatty acids	7	4.3	6	4.3	7	6	8	7.7
Phosphates	5	3	3	2.1	0	0	0	0
Isopropenoids	4	2.5	6	4.3	3	2.5	3	2.9
Photorespiration intermediates	2	1.2	4	2.8	3	2.5	1	0.9
Antioxidants	2	1.2	1	0.7	1	0.8	1	0.9
Nucleotides	2	1.2	2	1.4	2	1.7	1	0.9
n-alkane	1	0.6	11	8	11	9.5	11	10.6
Total metabolites identified	162		139		116		103	

6.3.2.5 Comparing metabolome of all the species investigated

In this section, we combined the identified metabolites under various treatments across all the four species. The objective of this type of data analysis is to examine the occurrence of different metabolites under given treatment. In total, 216 unique metabolites were identified with GC-MS based analysis under various treatments across all four species. The analysis showed that 63 unique metabolites were found to be identified under only one treatment, whereas 24 metabolites were found to be identified in all the various 12 treatments employed in this research work across all four species. Further all the metabolites identified in this research work were classified into different classes and compared across all the species. The comparison of different classes of unique metabolites identified across all the four species investigated can be seen in figure 6.7 and listed in table 6.9.

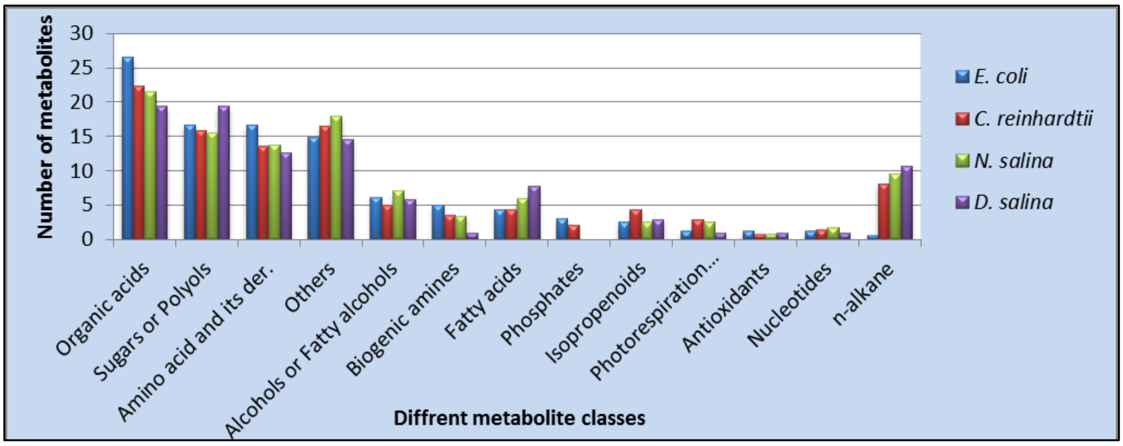


Figure 6.7 Distribution of the metabolite classes detected in metabolome of four different biological species investigated namely *E. coli* (blue)*, C. reinhardtii* (red), *N. salina* (green) *and D. salina* (purple). Y-axis represents average number of metabolites identified. X-axis represents distribution of 13 metabolite classes.

Table 6.9 Distribution of all the metabolites identified in this research into different classes (Combined results from all the biological species investigated)

Metabolite Class	Number	%
Organic acids	45	21
Others	41	19
Sugar or Sugar-alcohol and Polyols	30	14
Amino acids and its derivatives	29	13
n-alkane	14	6.5
Fatty acids	13	6
Biogenic amines or polyamines	11	5
Alcohols or fatty-alcohols	11	5
Isopropenoids	7	3.2
Phosphates	6	2.7
Photo-respiration intermediates	3	1.3
Antioxidants	3	1.3
Nucleotides	3	1.3
Total metabolites identified	216	

As can be seen from figure 6.8, the coverage of GC-MS was excellent for most of the metabolite classes which includes amino acids, organic acids, sugars/sugar alcohols, however recovery of nucleotides/nucleosides/nucleotides and sugar phosphates was severely compromised, thereby requiring the development of an optimized protocol with alternative analytical platform such LC-MS for coverage of these classes of metabolites. This is addressed in Chapter 10. The above preliminarily optimized quenching and extraction protocols were based on previously published benchmark protocols that had been adopted and modified slightly without critically evaluating them for the given condition and organism. In addition, with the above quenching and extraction protocols we currently covered only one fifth of the *C. reinhardtii* metabolome and there is considerable scope for improvement by developing reliable protocols for sampling and quenching (with appropriate correction for the metabolite leakage) followed by extraction which we propose to address in subsequent chapters.

6.3.3 Comparison of mixotrophic and photoautotrophic growth of *C. reinhardtii* grown under turbidostatic conditions: A metabolic and proteomic investigation.

6.3.3.1 Background

The major precursor for the biofuel production from feedstock organism are the carbon molecules in the form of carbohydrates and lipids, which constitute majority of the dry biomass. Therefore the method by which the organisms derives its carbon compounds serve as a key factor for any feedstock organism. Lab based *C. reinhardtii* cultures were generally grown under mixotrophic conditions (with acetate as a carbon source) in order to increase specific growth rates. However for the industrial scale biodiesel production, these cultures needs to be grown under photoautotrophic conditions, where for the carbon source, cultures have to rely on atmospheric or CO_2 supplemented growth. However prior to extrapolating the lab based investigations to industrial scale, it is essential to understand the influence of carbon source and availability on *C. reinhardtii* cultures.

As discussed in section 6.1, bottlenecks in algal biodiesel production within the cell can be identified by metabolomics approaches in combination with proteomics and transcriptomics, as the identification of differentially expressed metabolites gives clues to rate-limiting processes in cell, which can be backed up by the determination of metabolite flux. Thus, system biology approach will allow fine tuning of algal properties by genetic or metabolite engineering (Schenk et al., 2008).

To our knowledge only Wienkoop and co-workers have investigated alterations in the metabolome and proteome while comparing the mixotrophic and photoautotrophic batch cultures (Wienkoop et al., 2010b). However it is important to note that the approach was kept targeted for both the -omics approaches, focusing more on low CO_2 induced proteins. Similarly Renberg and co-workers reported significant differences in 128 metabolites while comparing the mixotrophic and autotrophic cultures of *C. reinhardtii* (Renberg et al., 2010). No proteomic investigation was carried out in this study. In both previous studies, the analyses was carried out by coupling the two different factors, carbon source and availability, which ideally needs to be evaluated separately under turbidostatic conditions.

Therefore, to assess the influence of these two factors, carbon source and availability, on *C. reinhardtii* cultures, the cultures were grown mixotrophically and photoautotrophically, each with and without CO_2 supplementation, using turbidostatic conditions in a stirred tank photo bioreactor, followed by metabolomic and proteomic analyses. Turbidostatic cultures (fixed cell density) were preferred over chemostatic cultures (fixed nutrient supply), as use of chemostatic culture will lead to dramatic impact on cell culture density, not an ideal condition for the desired comparison of carbon source and availability. Given this, we employed turbidostatic culture to gain control on the carbon availability at a per cell basis, where optical density (OD) was being used as a measure of cell numbers.

The influence of carbon source and availability on *C. reinhardtii* metabolome and proteome under the four conditions listed in figure 6.8 were further investigated;

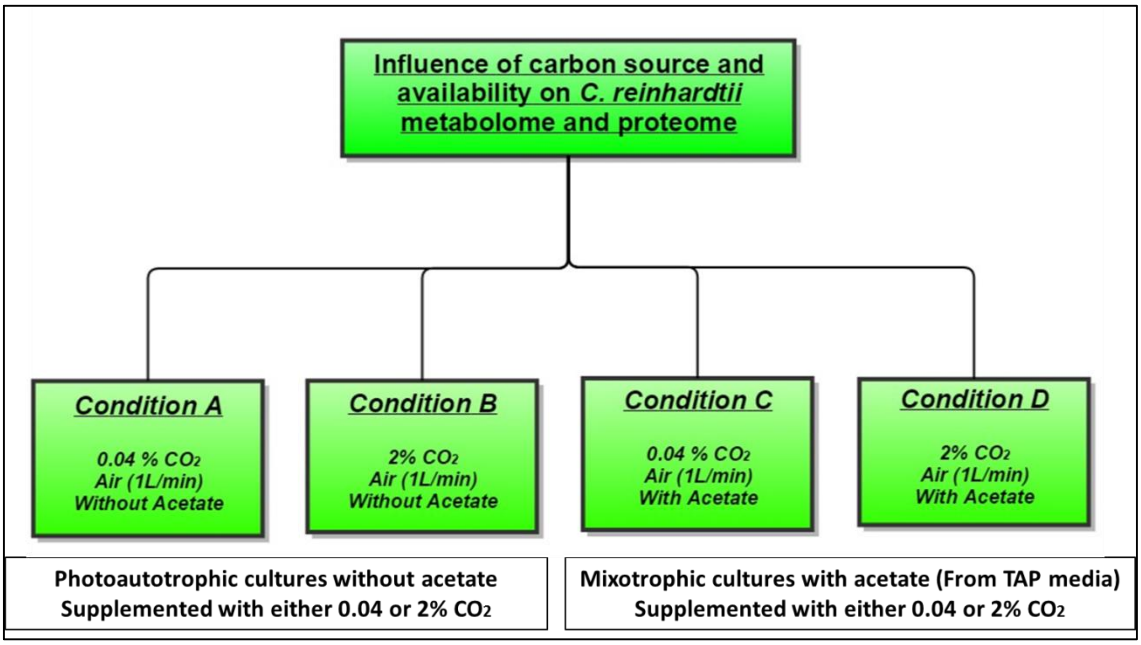

Influence of carbon source and availability on C. reinhardtii metabolome and proteome
Condition A
0.04 % CO2
Air (1L/min)
Without Acetate
Condition B
2% CO2
Air (1L/min)
Without Acetate
Condition C
0.04 % CO2
Air (1L/min)
With Acetate
Condition D
2% CO2
Air (1L/min)
With Acetate
Photoautotrophic cultures without acetate
Supplemented with either 0.04 or 2% CO2
Mixotrophic cultures with acetate (From TAP media)
Supplemented with either 0.04 or 2% CO2

Figure 6.8 Summary of conditions studied to investigate the influence of carbon source and availability on *C. reinhardtii* cultures grown under different trophic conditions.

Condition [A]: Un-supplemented Photoautotrophic culture: without acetate and atmospheric air supplemented with 0.04% CO_2

Condition [B]: Supplemented Photoautotrophic culture: without acetate and atmospheric air supplemented with 2% CO_2

Condition [C]: Un-supplemented Mixotrophic culture with acetate from TAP media and atmospheric air supplemented with 0.04% CO_2

Condition [D]: Supplemented Mixotrophic culture with acetate from TAP media and atmospheric air supplemented with 2% CO_2

6.3.3.2 Overview of metabolomics and proteomics results

To investigate the changes occurring under different trophic conditions an 8-plex iTRAQ proteomic investigation along with biochemical analysis (carbohydrate, protein and pigment assays) was carried out by Longworth and co-workers (Longworth, 2013), whilst we simultaneously focused on metabolomic analysis. It is important to note that our primary aim was to assess, if the optimized metabolomics protocol can be applied successfully to track the changes in the metabolome of *C. reinhardtii* cultures grown under different trophic conditions. Moreover, validation of our results by other omics approaches, in this case with proteomics, was also investigated to evaluate and compare the linearity between both the -omics studies.

Multiple comparisons can be made from the datasets obtained with both the -omics approaches. The results of these multiple comparisons, where only significant differences observed in both the proteins and metabolites were summarized in table 6.10. For GC-MS based metabolomic investigations, two biological replicates with three technical replicates were used, resulting in total six replicates for each condition. Only features that were found to be identified in four out of six replicates were considered as true hits for further data analysis.

Table 6.10 Number of significantly changing features in proteome and metabolome of *C. reinhardtii* against various conditions as summarised in Figure 6.8

Comparative analysis		***Conditions***	***Number of Significantly Changing Features In***	
			Proteome	***Metabolome***
Comparison 1	***1***	***A vs B***	***155***	***81***
	2	***A vs C***	***191***	***278***
	3	***A vs D***	***165***	***285***
	4	***B vs C***	***60***	***55***
	5	***B vs D***	***92***	***91***
	6	***C vs D***	***36***	***17***
Comparison 2		***A vs B, C & D***	***x***	***488***

In case of proteomic investigations, significantly different features refers to $p < 0.05$, whereas in case of metabolomics investigations significantly different features refers to $p \leq 0.01$ with fold change ≥ 1.3.

6.3.3.3 Condition A vs B, C and D

To assess the sample arrangement, PCA is calculated (figure 6.9) using the feature intensities from all the analysed samples using XCMS online (web based platform). With PC1 38% variance was observed compared to 11% variance of PC2, indicating 4.1 fold more significant changes along the X-axis than those on the Y-axis. From figure 6.9 (a), it is clear that condition A is well separated from other conditions B, C and D. Similar observations were made with proteomic investigations (figure 6.9(b)), where PC1 accounts for 74% variations compared to PC2 which accounted for 10% variations, indicating 7.4 fold more significant changes along the X-axis than those on the Y-axis. Both the approaches suggesting that, the primary effect of differentiation is due to the carbon availability.

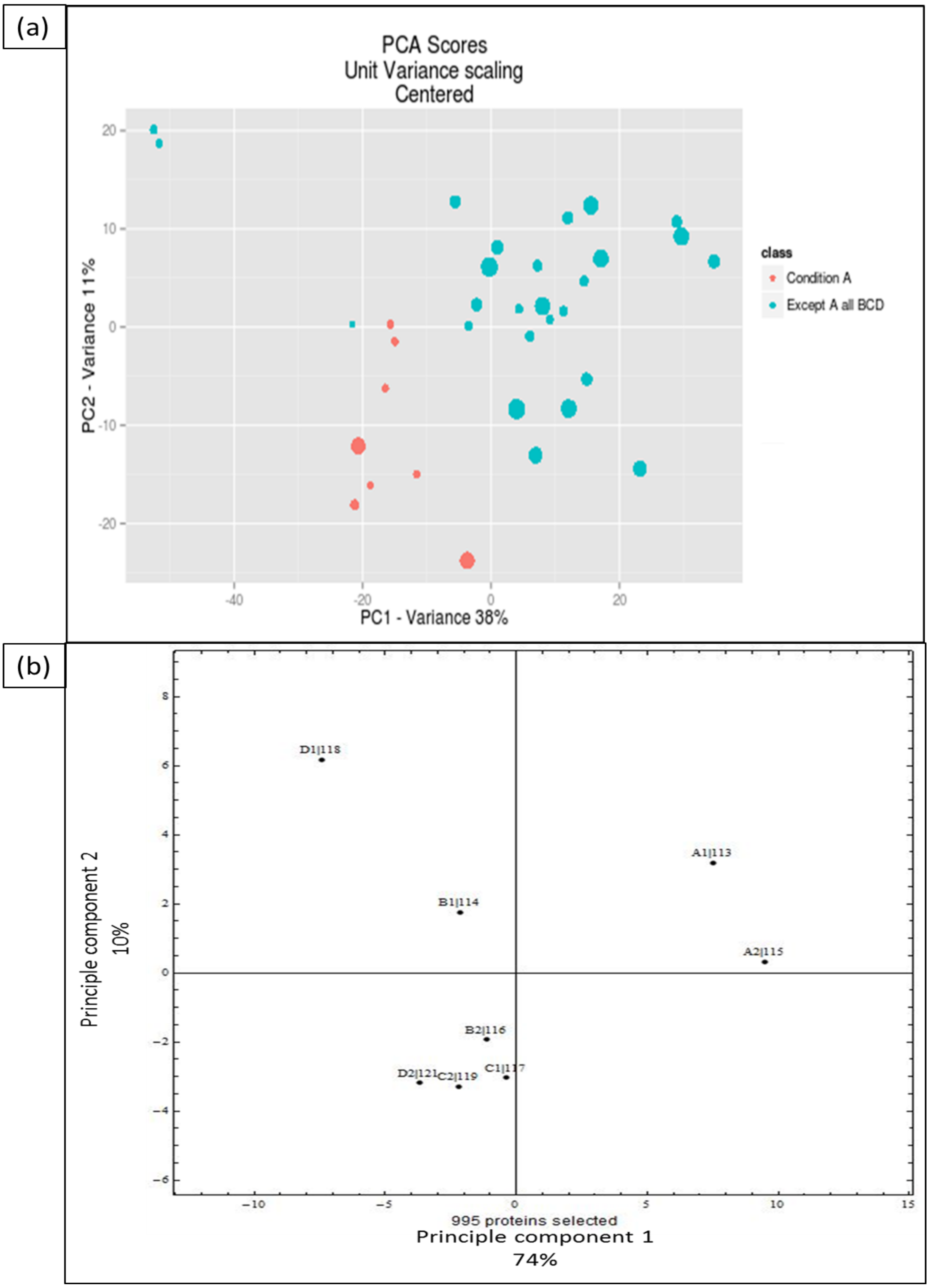


Figure 6.9 PCA analysis displaying clustering of *C. reinhardtii* samples for all four culture conditions as summarised in Figure 6.8, where: (a) PCA for metabolomic data and (b) PCA for proteomics data (figure adopted from (Longworth et al, 2013). For metabolomics data, PCA is calculated using the feature intensities from all samples. The colours (red/blue) are assigned based on the sample class. Distance to the model (DModX) test was used to verify the presence of outliers and to evaluate whether a submitted sample fell within the model applicability domain. For proteomics data PCA is based on variance stabilization normalisation (VSN), isotope correction (IC), median corrected reporter ion intensities, limited to ≥ 2 peptide proteins.

To gain further insight into data characteristics such as *p*-value, fold change, retention time, *m/z* ratio and signal intensity of all detected features, cloud plot visualization tool was employed using the XCMS online platform. The results of these investigations are summarized in figure 6.10.

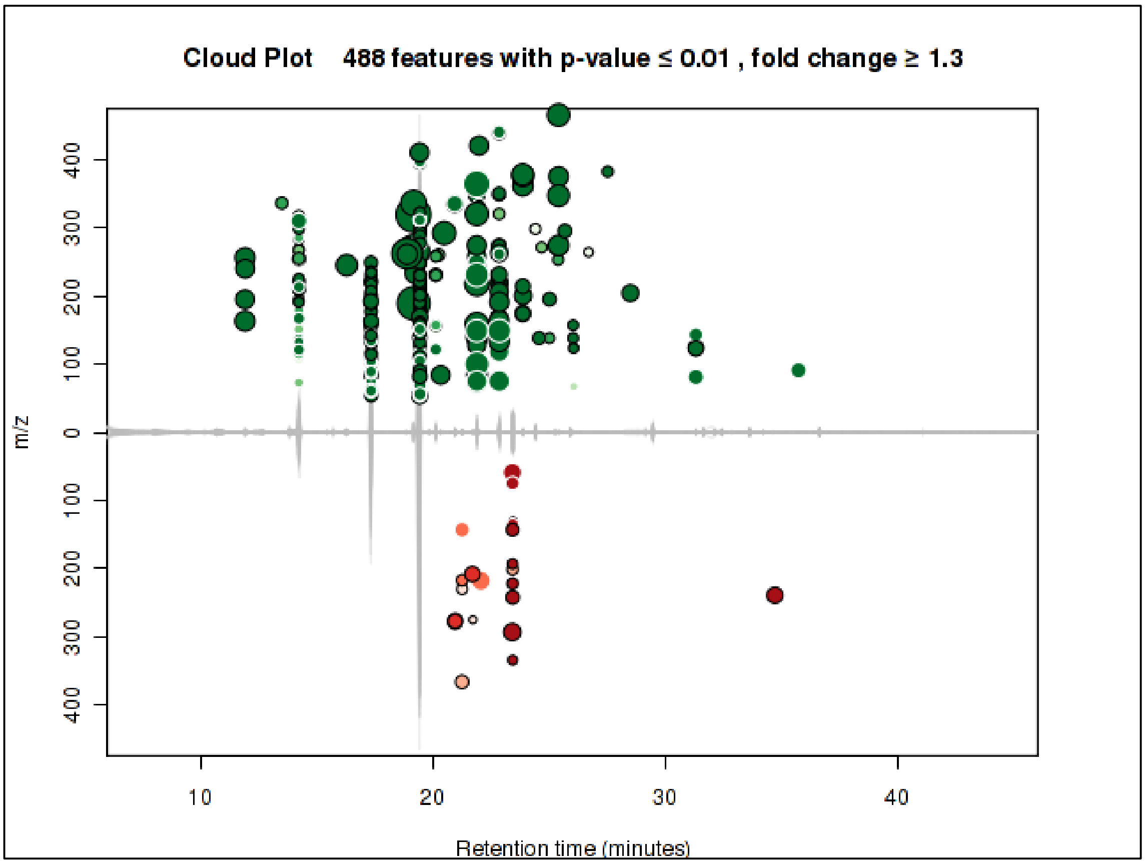


Figure 6.10 Cloud plot generated by comparing condition A against remaining three conditions (see Figure 6.8), displaying 488 features with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value.

As can be seen from figure 6.10, in total 488 dysregulated features were detected whose intensities were altered between sample groups with *p*-value ≤ 0.01 and fold change ≥ 1.3, when condition A (Dataset 1 as control) was compared against other conditions (Dataset 2), which includes condition B, C and D. Up-regulated features were shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the *p*-value, darker the colour, the smaller the *p*-value. The statistical significance of the fold change, as calculated by a Welch *t* test with unequal variances, is represented by the intensity of the features colour, where features with low *p*-values are brighter compared to features with high *p*-values. The Y- co-ordinate of each feature corresponds to the *m/z* ratio of the compound as determined by the mass spectrometry. The features that were identified with the METLIN database were colour coded with the black outline surrounding the bubble, whereas those not identified in the database are displayed without a black outline.

For qualitative and quantitative analysis, metabolite identification and data analysis was performed as described in section 6.2.7. Out of 488 significant features, 138 unique metabolites were identified across all the conditions.

6.3.3.4 Condition A (Low carbon) vs Condition D (High carbon)

PCA analysis from both the omics approaches clearly suggests that there is majority of variance between condition A and other three conditions. The carbon availability was extremely low with the condition A, whereas in contrast it was extremely high with condition D. Considering this, we have further selectively compared the condition A against the condition D.

To assess the sample arrangement between low vs high carbon conditions, PCA is calculated (figure 6.11) using the feature intensities from samples belonging to condition A and D. With PC1 13% variance was observed compared to 44% variance of PC2, indicating 5.7 fold more significant changes along the X-axis than those on the Y-axis. Further analysis with cloud plot visualization (figure 6.12) reported 289 features that were significantly different from each other among both the conditions, out of which 94 unique metabolites were identified (Appendix 6.4). Among 94 unique metabolites, 63 metabolites were found to be present in both the condition A and D, whereas 13 and 18 metabolites were identified only in condition A and D respectively.

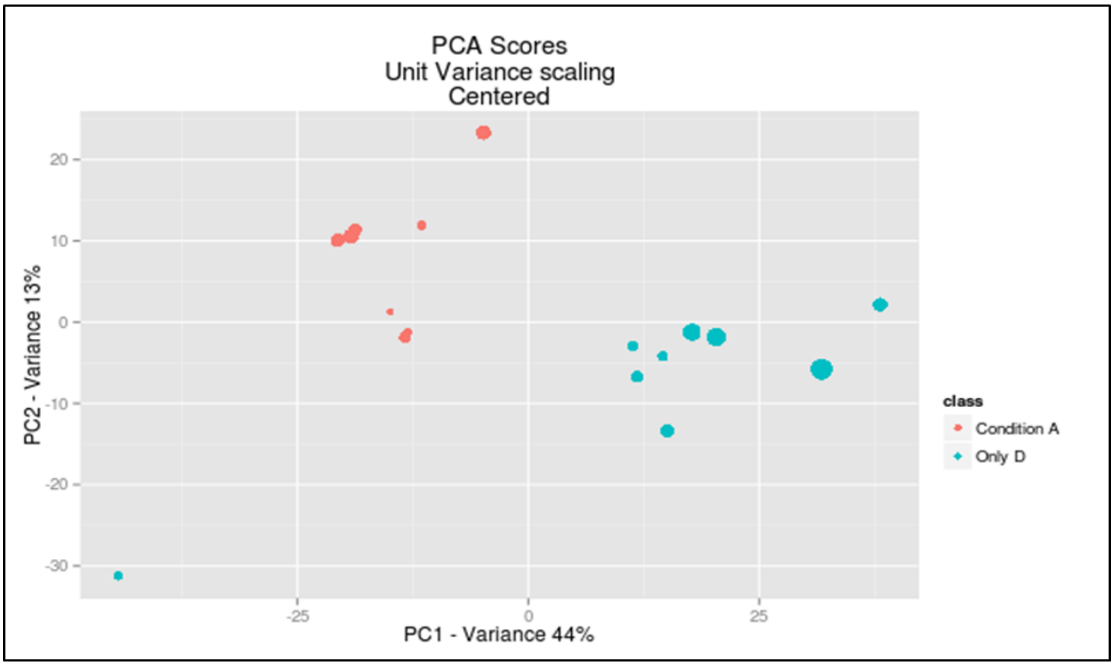


Figure 6.11 PCA analysis for metabolomics data displaying clustering of samples between condition A and condition D. PCA is calculated using the feature intensities from all samples. The colours (red/blue) are assigned based on the sample class. Distance to the model (DModX) test was used to verify the presence of outliers and to evaluate whether a submitted sample fell within the model applicability domain.

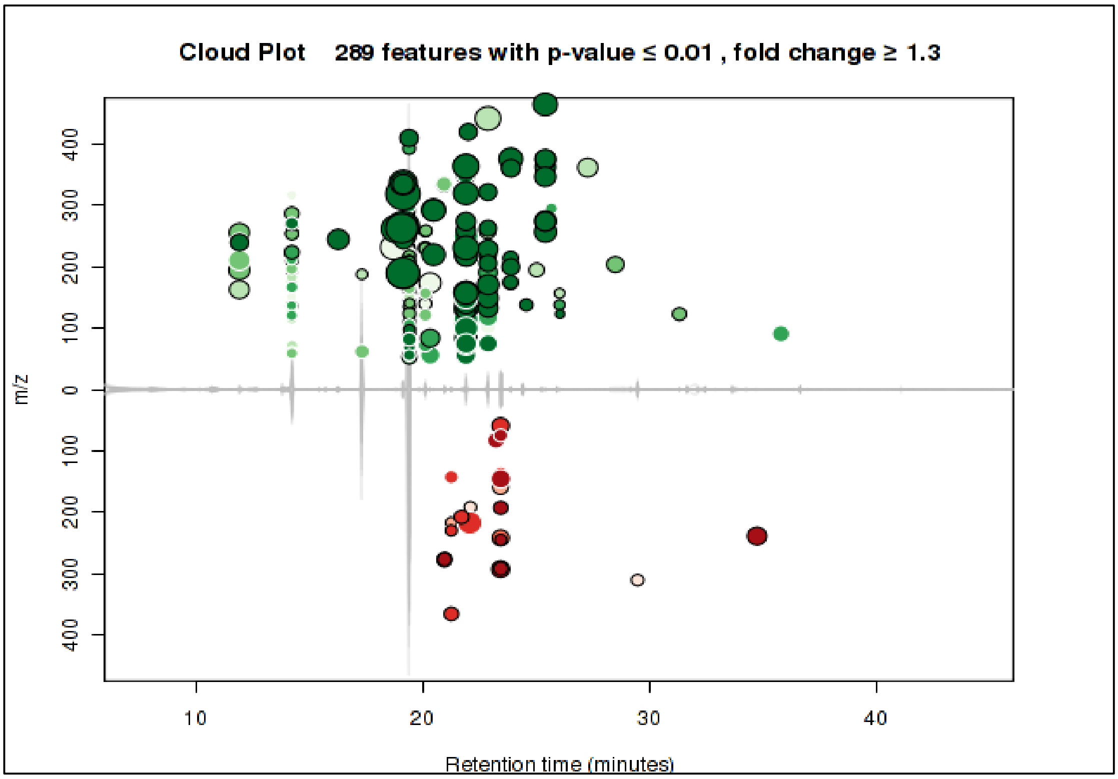


Figure 6.12 Cloud plot generated by comparing condition A against condition D (see Figure 6.8), displaying 289 features with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value.

In the proteomic analysis led by Longworth and co-workers (Longworth, 2013), the authors grouped the significant proteins based on Gene Oncology (GO), the results of which revealed an increase in the overall biosynthetic processes and therefore increased cellular production with the increased carbon availability (condition D) compared to lower carbon availability (condition D). In case of metabolomics analysis we have classified the significant metabolites based on their physicochemical characteristics, the results of which were in agreement with that of proteomic investigation, as we have observed an increase in intracellular metabolite concentration across various classes of metabolites. The results of these investigation are summarized in figure 6.13.

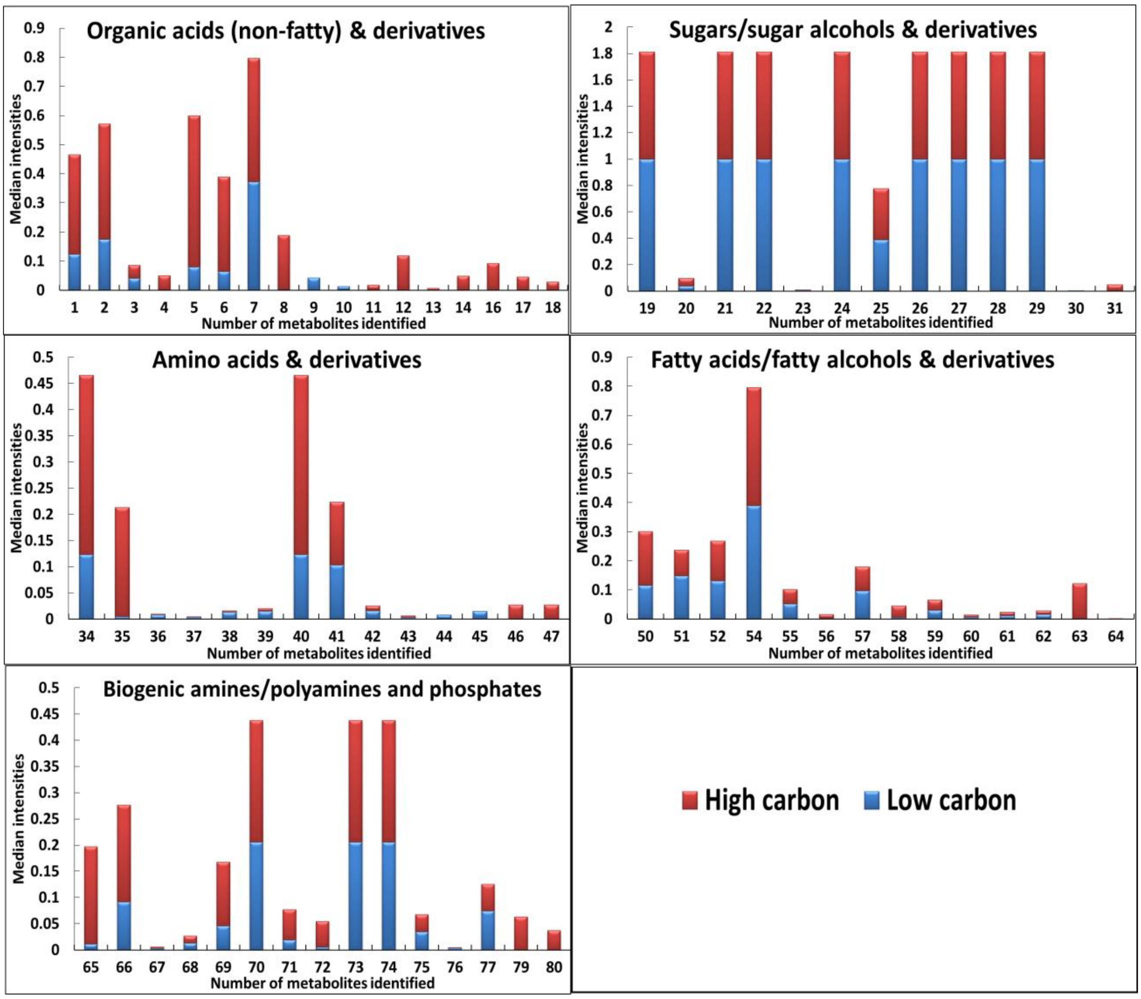


Figure 6.13 Variations observed in five different metabolite classes across Low Carbon (condition A – blue colour) and High Carbon (condition D – red colour) culture conditions in *C. reinhardtii* (A vs D). X-axis represents unique number of metabolites identified as listed in Appendix 6.3; whereas Y-axis represents median intensities obtained for individual metabolites.

As can be seen from figure 6.13, compared to low carbon availability (condition A), higher recoveries were observed for organic acids, amino acids, fatty acids, biogenic amines and phosphates with the high carbon availability condition (condition D). In case of sugars/sugar alcohols and derivatives, minor variations were observed with both the conditions, which was in agreement with the previously published report (Renberg et al., 2010). Similar observations were reported by Longworth and co-workers (Longworth, 2013) with carbohydrate assay, where no significant changes in total carbohydrates was observed. All these observations are in agreement with the results obtained from proteomic and gross biochemical/physiological analysis (Longworth, 2013) conducted on the same sample set.

6.3.3.5 Six two way comparison for each condition against other

In order to assess the influence of these two factors (carbon source and availability) independently and dependently, six two way comparisons were made for each condition against other. With all the six comparisons number of significant differences observed are listed in table 6.10. The analysis was performed using cloud plot visualization tool, results of which are summarized figure 6.14.

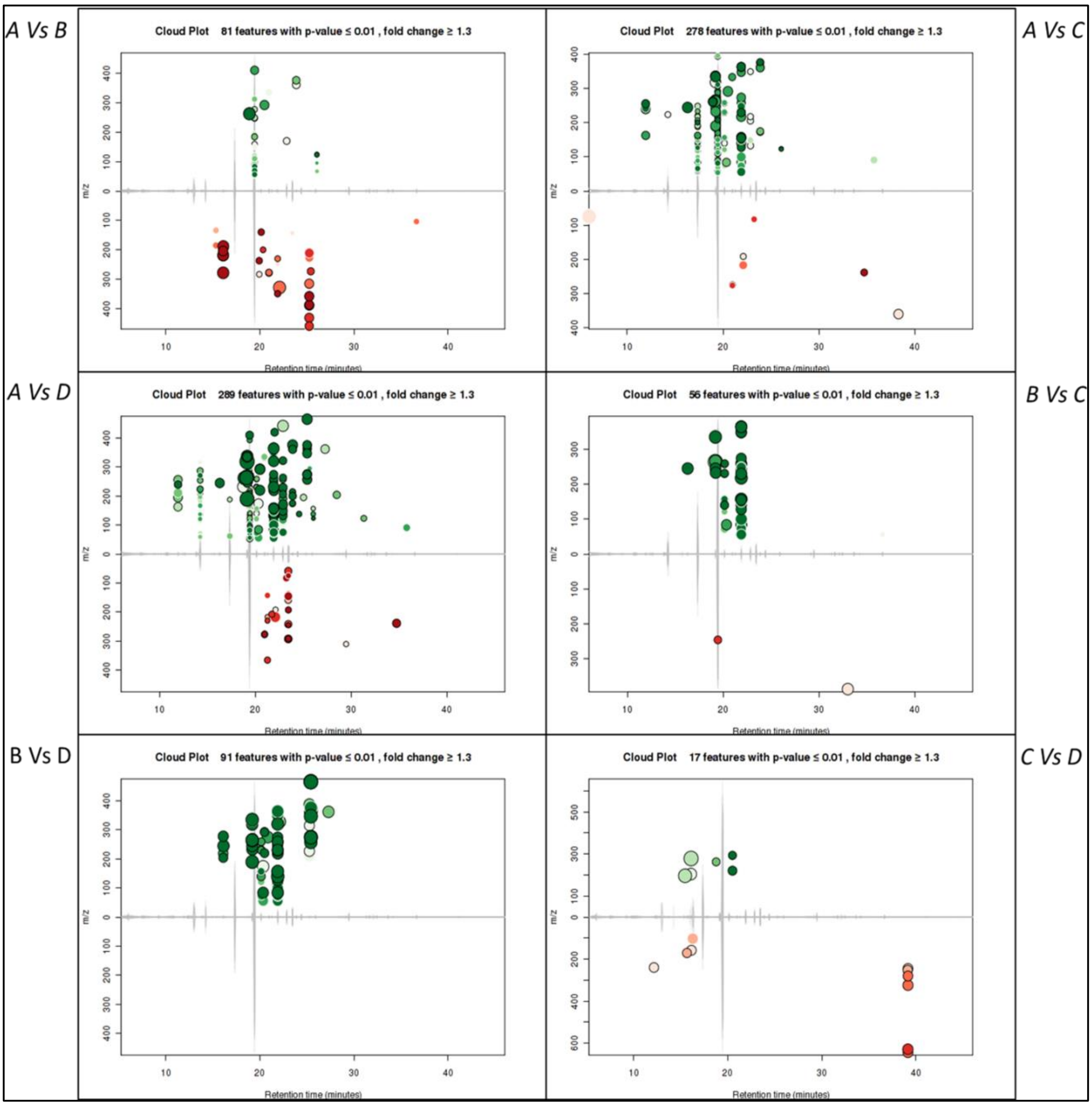


Figure 6.14 Cloud plot generated by comparing six two way comparison for each condition against each of the other, displaying significantly different features with p-value ≤ 0.01 and fold change ≥ 1.3. All the compared culture conditions for *C. reinhardtii* are summarised in Figure 6.8. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value.

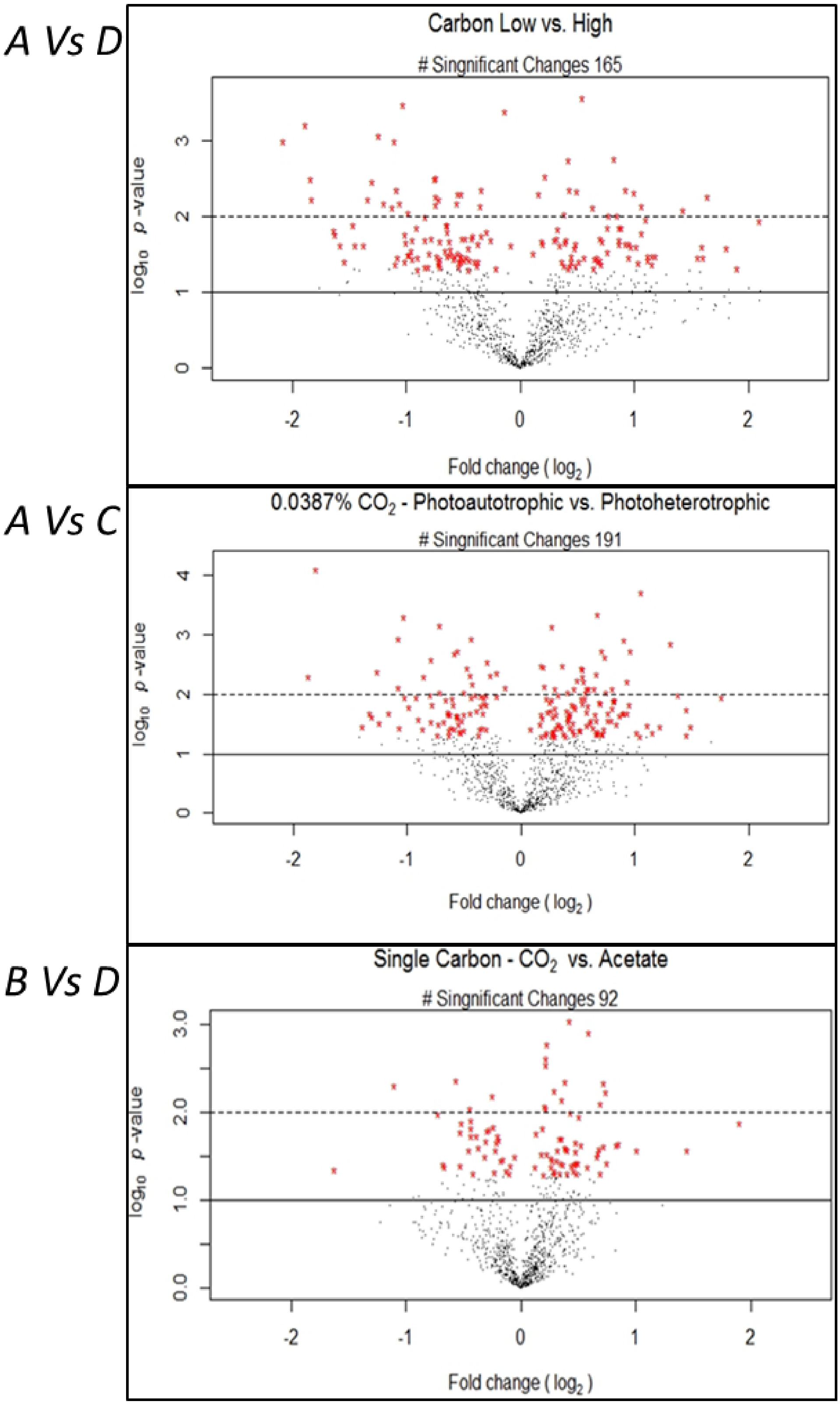


Figure 6.15 Volcano plots displaying the significant changes observed when comparing the various culture conditions of *C. reinhardtii* among them. Details of the culture conditions compared above are summarised in Figure 6.8. Proteins with significant changes ($p < 0.01$) are shown by red star. The more relaxed cut-off of $p < 0.05$ is indicated by a solid line and the more stringent cut-off of $p < 0.01$ shown by a dotted line (figure adopted from (Longworth, 2013)).

A vs D comparison

As demonstrated above, major variation in the metabolome of condition A and D (figure 6.14 and 6.15) was mainly attributed to the fact of variations in both the carbon source and availability.

A vs C comparison

Higher number of significant changes (278 features) were observed (figure 6.14), when condition A was compared against condition C. Similar observation were made with the proteomic analysis where 191 significant protein changes were observed (figure 6.15), suggesting there is a considerable reorganization of the metabolome and proteome with these two culture conditions.

B vs D comparison

Reduction in number of significant differences, when condition B was compared with that of D (91 features) (figure 6.14), clearly suggests that supplemented mixotrophic growth (Condition D) is similar to that of supplemented photoautotrophic culture. Similar observation were made with the proteomic analysis where only 92 significant protein changes were observed (figure 6.15).

B vs C; A vs B and C vs D comparison

Major reduction in number of significant features were also observed when following conditions were compared with each other (figure 6.14);
B vs C (55 & 60 features in metabolome and proteome respectively),
A vs B (81 & 155 features in metabolome and proteome respectively) and
C vs D (17 & 36 features in metabolome and proteome respectively).

This investigation identified significant changes based on the trophism. It identifies the need to ensure consistent carbon availability between compared conditions during investigations. Lowest number of significant features (17 features) (figure 6.14) within mixotrophic cultures (C vs D), suggests that, the culture maintains its true mixotrophic growth without any reduction in the photosynthetic apparatus, as also demonstrated with the proteomics results.

The interpretations made from proteomics and biochemical analysis (Longworth, 2013) in this case suggests that, the most predominant factor influencing the *C. reinhardtii* cultures is the carbon availability rather than the source of carbon, where metabolomic analysis seems to support this conclusion based on multivariate analysis performed on un-targeted metabolomics data, which further requires in-depth pathway based analysis to strengthen and support the conclusions derived from both the -omics approaches.

However as mentioned earlier, the objective of this investigation was to demonstrate efficient application of preliminarily optimized metabolomics workflow to hypothesis driven biological study, along with the other -omics approaches such as proteomics. The GC-MS based metabolomic results clearly showed the ability to track the variations in the metabolome of *C. reinhardtii* cultures grown under different trophic conditions. In addition, we directed our efforts towards demonstrating and validating our results with that of proteomic results, which showed highly linear relationship, when outcomes of both the -omics approaches were compared using visualization tools as shown in table 6.10. However, it would be advantageous to assess the variations in differentially expressed metabolites (and classes) with respect to their involvement and role in numerous pathway, in order to fully explore the potential of metabolomics. Only such comprehensive investigation will help in understanding the biology behind rate-limiting processes in cell, which can be then backed up by the determination of metabolite flux. However, we restricted the scope of this investigation to assessing the metabolome coverage and refrained from further pathways based metabolomic analysis due to the adopted sampling, quenching and extraction protocols, which require prior comprehensive evaluation and optimization in order to make accurate and reliable biological interpretations. As discussed in Chapter 3 and 4, different solvent based quenching and extraction methods can yield significantly different metabolite profiles, which has major influence on the biological interpretations made from such metabolomics data. In addition, as mentioned in section 6.3.2.4, with the above quenching and extraction protocols we currently covered only one fifth of the *C. reinhardtii* metabolome and there is considerable scope for improvement which we propose to address in subsequent chapters.

6.4 Conclusions

Results from optimized quenching and extraction protocols with microbial and microalgal species have been used to demonstrate that GC-MS is a valuable method for metabolomics investigations. Superior coverage for most of the metabolite classes was observed with GC-MS based analysis. However, recovery of nucleotides/nucleosides/nucleotides, pigments and sugar phosphates was severely compromised thus requiring development of an optimized protocol with alternative analytical platform such LC-MS for coverage of these classes of metabolites. This aspect is addressed in Chapter 10. Further to improve the metabolome coverage with GC-

MS, we proposed employing 2D GC-ToF-MS or GC-ToF-MS, but since we did not have access to such facilities, we directed our approach towards establishing the strategies for evaluation and optimization of various steps involved in the metabolomics workflow that is needed irrespective of the platform subsequently employed.

Biology based metabolomics investigation identified significant changes in the metabolome of *C. reinhardtii* cultures grown under different trophic conditions and showed high co-relation and agreements with the conclusions drawn from proteomic and biochemical analysis. We evaluated the influence of two factors, carbon source and availability, on mixotrophically and photoautotrophically growing cultures of *C. reinhardtii*, and proposed that, the most predominant factor influencing the *C. reinhardtii* cultures is the carbon availability rather than the source. Results of these investigations clearly demonstrated suitability of GC-MS based metabolomics protocol for making a biological interpretations from the metabolomics data, however we restricted ourselves from further pathway based in depth analysis, as the applied sampling, quenching and extraction protocols were optimized based on very narrow approach, where previously published benchmark protocols were simply adopted and modified slightly without critically evaluating them for the given condition and organism. As demonstrated in Chapters 3 and 4, different solvent based quenching and extraction methods can yield significantly different metabolite profiles, which can influence the biological interpretations of the metabolomics data. In addition, with the above quenching and extraction protocols we currently covered only one fifth of the *C. reinhardtii* metabolome and there is considerable scope for improvement by developing reliable and appropriate methods for sampling and quenching with appropriate correction for the metabolite leakage, which we propose to address in Chapter 7. Following quenching, the major issue in metabolomic studies is to find a suitable extraction solvent that can quantitatively extract all the intracellular metabolites, which we propose to address in Chapter 8. Bottlenecks in algal biodiesel production within the cell can be identified by metabolomics approaches, however prior to identifying and addressing these bottlenecks it is essential to optimize the FAME production workflow for biodiesel application as discussed in in Chapter 9.

Chapter 7

Evaluation and optimisation of appropriate sampling and quenching method for microalgal sample: A case study on the unicellular eukaryotic green alga *Chlamydomonas reinhardtii*

7. Evaluation and optimisation of appropriate sampling and quenching method for microalgal sample: A case study on the unicellular eukaryotic green alga *Chlamydomonas reinhardtii*

7.1 Introduction

Quenching with 60% *v/v* cold methanol at -40°C or -50°C has been used widely in the past for microbial, fungi, yeast and plant metabolomics. However, potential problems connected to leakage of intracellular metabolites with cold methanol quenching was reported later for yeast (Gonzalez et al., 1997) and bacterial cells (Bolten et al., 2007, Villas-Bôas and Bruheim, 2007, Villas-Bôas et al., 2005, Wittmann et al., 2004). As discussed in the Chapter 2 – section 2.2.1, various attempts were made in previously published reports, in order to overcome and/or minimise the metabolite leakage during cold methanol quenching. However, no systematic solution to this problem has been reported as yet. It is important to understand that, the leakage occurs not only due to the contact of cold methanol solution with cells, but also because of the other causes. In prokaryotes, a sudden change in temperature alone might cause drastic loss of metabolites induced by the cold shock, whereas in eukaryotes leakage is not only governed by the cold shock, but also by various other factors as detailed below, which has not been investigated thoroughly:

1. Inclusion of various buffer additives (such as HEPES, AMBIC, tricine, NaCl) to the quenching solvent at different concentration.
2. Quenching solvents other than conventional cold methanol/water solution.
3. Various concentration of methanol in quenching solvent.
4. Starting and final temperature of quenching solvent.
5. Contact time of sample with the quenching solvent.
6. Sample to quenching solvent ratio.

The inclusion of additives to methanol, will act as a buffer or will restrict osmotic shock, leading to a decrease in leakage, as has been reported with bacterial cells (Faijes et al., 2007) yeast (Canelas et al., 2008, Spura et al., 2009) and mammalian cells (Kronthaler et al., 2012b, Sellick et al., 2008). Inclusion of tricine buffer in 60% aqueous methanol have been shown to result in lower recoveries for the intracellular metabolites because of the interference of tricine with derivatization reactions (Tredwell et al., 2011). However, with other additives within the same biological system contradictory conclusions have been reported (Canelas et al., 2008) as can be seen from Table 7.1.

Table 7.1 Summary of the main different quenching solvents and techniques applied so far for the metabolome analysis of different biological systems. (Sing x stands for no information available or provided in the research papers surveyed)

Sample	***Culture volume***	***Quenching solvent (QS)***	***Quenching buffer***	***Temp***	***Ratio (Sample to QS)***	***Washing solutions***	***Analytical Technique***	***Recommended QS***	***Reference***
Penicillium chrysogenum	5 mL	100% Methanol	x	(-40°C)	x	x	GC-MS & LC-ESI-MS/MS	40% Methanol	(de Jonge et al. 2012)
		60% Methanol		(-40°C)	1:5				
	10 mL	40% Methanol		(-25°C)	1:1				
Pichia pastoris	2 mL	60% Methanol	x	(-50°C)	1:7	Not required	GC-MS and NMR	Minor differences with all 4 QS, 60% Methanol + 0.11M AMBIC	(Tredwell et al. 2011)
		86% Methanol							
		60% Methanol +10 mM Tricine	10mM Tricine						
		60% Methanol + 0.11M AMBIC	0.11 M AMBIC						
S. cerevisiae	1 mL	100% Methanol	x		1:5	5 mL water	GC-ToF-MS	Fast filtration with water washing	(Kim et al. 2013)
S. cerevisiae	x	60% Methanol + 10 &/or100 mM HEPES	10 &/or100 mM HEPES	*(-40°C)*	1:5	*Not required*	*GC-MS & LC-ESI-MS/MS*	*100% Methanol*	*(Canelas et al. 2008)*
		60% Methanol + 10 &/or100 mM AMBIC	10 &/or100 mM AMBIC						
		60% Methanol + 10 &/or100 mM Tricine	10 &/or100 mM Tricine						
		40% Methanol	x						
		50% Methanol							
		60% Methanol							
		70% Methanol							
		80% Methanol							
Yeast (*Yarrowia lipolytica*)	7 mL	60% Methanol	x	(-40°C)	1:4	50% Methanol	GC-MS	DMSO-saline	(Zhao et al. 2014)
		Glycerol /NaCl (3:2)	NaCl	(-20°C)		Glycerol /NaCl (1:1)			
		20, 40 or 60% DMSO/NaCl	NaCl	(-4, -20 & -40°C)					
Yeast and Bacteria	x	40% Ethanol/NaCl	NaCl	(-20°C)	x	Not required	GC-MS	40% Methanol/0.8% NaCl	(Spura et al. 2009)
		60% Methanol	x	(-50°C)					
		Glycerol /NaCl (3:2)	NaCl	(-20°C)					
Yeast & Bacteria	5 mL	Glycerol-water (3:2)	x	(-23°C)	1:5	x	GC-ToF-MS	Cold glycerol-saline	(Villas-Bôas and Bruheim 2007)
		Glycerol-saline (3:2)	NaCl						
Bacillus subtilis, Corynebacterium glutamicum, E. coli, Gluconobacter oxydans, P. putida, Zymononas mobilis	x	60% Methanol + 10 mM HEPES	10 mM HEPES	(-58°C)	1:2	x	HPLC & IC-MS/MS	60% Methanol + 10 mM HEPES	(Bolten et al. 2007)
Corneybacterium glutamicum	5 mL	60% Methanol + 10 mM HEPES	HEPES	(-58°C)	1:2	With wash	HPLC	Buffered methanol	(Wittmann et al. 2004)
		60% Methanol + 10 mM HEPES	HEPES			Without wash			
		60% Methanol	x			x			
		0.9% NaCl	x	(0.5°C)					
Corneybacterium glutamicum		60% Methanol	x	(-50°C)	x	x	LC-MS		(Wellerdiek et al. 2009)
E. coli	5 mL	LN2	x	(-40°C)	1:1	PBS	NMR	LN2	(Bertini et al.)
		60% methanol							

The use of alternative quenching solvents to cold methanol/water which includes: glycerol-saline (Villas-Bôas and Bruheim, 2007), methanol/glycerol (Link et al., 2008) and methanol/NaCl (Faijes et al., 2007) have showed minimum leakage compared to use of 60% aqueous methanol. However, these alternatives have also been shown to add difficulties in the overall metabolomics workflow. For example, the higher viscosity and hygroscopicity of glycerol has been shown to result in prolonged processing time (during separation of glycerol from cells) (Schädel et al., 2011) and interference with the commonly employed silyalation derivatization reaction (required for GC-MS analysis) (Spura et al., 2009, Zhao et al., 2014). Faijes and co-workers reported higher leakage of ATP from *S. cerevisiae* with the use of methanol/NaCl indicating more damage to the cell integrity compared to 60% aqueous methanol (Faijes et al., 2007). To date, there are no effective alternative quenching solvents to that of 60% aqueous methanol, which usually results in leakage of the intracellular metabolites, thereby requires accurate balancing of the quenching supernatant and sample (Wellerdiek et al., 2009b).

Few reports have investigated the influence of various concentration of methanol on metabolite leakage (Canelas et al., 2008, de Jonge et al., 2012), however the approach was kept limited for the quantification of specific metabolites. Another important factor that needs to be evaluated include, the influence of initial (before quenching) and final temperature of methanol (sample-quenching solvent mixture). Lower the temperature, slower will be the turnover rate of all the enzymes within the cell and hence more efficient is the quenching process (Sellick et al., 2008). However, only Canelas and co-workers evaluated this factor for yeast (Canelas et al., 2008), where authors concluded that “leakage free quenching” can be achieved with the use of pure methanol rather than 60% methanol with quenching solvent to sample ratio of 5:1, and at -40°C or lower (Canelas et al., 2008). However with the similar “leakage free quenching” method, Kim and co-workers have demonstrated severe leakage in *S. cerevisiae*, suggesting reduced membrane integrity caused by the use extreme cold quenching conditions (Kim et al., 2013).

The influence of contact time of sample with the quenching solvent has not been investigated yet, where higher leakage of specific classes of metabolites is likely to occur upon increasing the exposure time of sample to that of quenching solvent (Bolten et al., 2007). Finally, the influence of sample to quenching solvent ratio should also be carefully considered, as it directly alters the temperature of the quenching process and might help in minimising the intracellular metabolite leakage. Schädel and co-workers evaluated this factor for *E. coli* metabolomics, and observed no change in the impact of methanol with the temperature variation (Schädel et al., 2011).

As discussed above, despite contradicting reports, quenching protocols have been rigorously studied and optimised for yeast, mammalian and bacterial model. Lee and Fiehn suggested, yeast can be regarded as good proxy for *C. reinhardtii* as both are eukaryotes and have sturdy cell walls compared to that of bacterial models which are easily prone to metabolite leakage caused by harsh quenching treatments (Lee and Fiehn, 2008a). However, as discussed and demonstrated in Chapter 3, minor differences in cellular characteristics including membrane, wall structure, size and sampling techniques employed, can influence the efficiency of quenching, recovery of different metabolite classes and the rate of metabolite leakage. Therefore, optimised quenching methods for yeast, mammalian and bacterial models cannot be simply adopted for microalga without critically evaluating them.

In case of microalgal samples we are not aware of any report to date, except that of Lee and Fiehn (Lee and Fiehn, 2008a), where the authors reported quenching cultures of *C. reinhardtii* with 70% aqueous methanol (pre-cooled to -70°C) with 1:1 ratio to sample, resulted in minimum leakage of intracellular metabolites. In this case, the resultant final concentration of methanol was 35% and final temperature recorded was -20°C. This finding was in contrast to their previously published report, where no leakage was reported and results concluded that *C. reinhardtii* cultures are resistant to quenching with cold methanol (Bölling and Fiehn, 2005).

The objective of this investigation is to examine quantitatively, the influence of the above mentioned factors on the extent of metabolite leakage in *C. reinhardtii* cultures. To achieve our objective, we have designed and categorised the experiments into three different approaches. Approach 1 involves evaluation of selected quenching solvents (nine in total) with varying methanol concentration and with various buffer additives, on the extent of metabolite leakage. Approach 2 investigates the effect of prolonged quenching duration on metabolite leakage and approach 3 investigates the effect of the ratio of quenching solvent to culture on metabolite leakage. Furthermore, we investigated recovery of twelve different metabolite classes using GC-MS technique across different quenching treatments. The evaluation was based on recovery of median number of metabolites within each class in all possible sample fractions and recovery of normalised median peak intensities of metabolites within each class.

7.2 Material and Methods

7.2.1 Chemicals and analytical reagents

All chemicals and reagents were obtained from Sigma-Aldrich (Dorset, U.K.), unless stated otherwise.

7.2.2 Microalgal cultivation

The *C. reinhardtii* strain (CC4323) was grown under constant illumination in a Sanyo incubator (MLR-351H, Sanyo versatile environmental test chamber, from Japan) at 25°C in 1L shake flasks, containing 1L of TAP medium, under constant illumination at 85 µmol/m^2/s. The medium (per L) is composed of 2.42 g Tris; 25mL of TAP salts (15g NH_4Cl, 4g $MgSO_4 \bullet 7H_2O$, 2g $CaCl_2 \bullet 2H_2O$ in 1L dH_2O); 0.375mL of phosphate solution (28.8g K_2HPO_4, 14.4g KH_2PO_4 in 100mL dH_2O); 1mL solution Hunter's trace elements purchase from the Chlamydomonas Resource Centre (Minnesota, US) and 1mL of glacial acetic acid. Cells were harvested at an OD of 1.2 at 680nm wavelength.

7.2.3 Sampling and quenching

To evaluate and minimise the leakage of metabolites during quenching treatments the overall design of quenching experiments within this Chapter has been categorised into three different approaches as summarised in Figure 7.1. At the incubation site, 1 mL of cell suspensions were rapidly plunged into a 2 mL of pre-chilled Eppendorf containing 1 mL of pre-chilled quenching solvent (-50°C) unless stated otherwise. Addition of the cells to the quenching solvent increased the temperature by no more than 15°C. The centrifuge was set at -9°C and the rotor was pre-chilled at -24°C. The quenched biomass was then centrifuged for 2 min at 2500*g* at -9°C. The supernatant was removed rapidly, and transferred to a 2 mL pre-chilled Eppendorf to assess the leakage of internal metabolites. The pellets and supernatant were snap frozen in liquid nitrogen and stored at -80°C for 3 weeks for further analysis.

7.2.4 Metabolite extraction

After 3 weeks of storage at -80°C, the samples were lyophilized at -50°C for overnight prior to extraction. Briefly, 500 µL of extraction solvent (M:C:W (5:2:2)) was added to lyophilized cells, as suggested elsewhere (Gullberg et al., 2004b, Lee and Fiehn, 2008a, Weckwerth et al., 2004b), pre-chilled at -48°C, along with an equal volume of glass beads (425-600 µm i.d., acid washed, from Sigma). Cells were then disrupted using a Cell disruptor, (Genie, VWR, U.K.), 11 cycles of disruption was performed with 2 min disruption, with an interval of 1 min on ice. The sample was then centrifuged at 13,000 rpm, at -9°C for 15 min to remove any cell debris. The supernatant was transferred to a new pre-chilled Eppendorf tube (-20°C) and the remaining cell debris was subjected to re-extraction (5 Cycles) with 500 µL of extraction solvent & centrifuged at 13,000 rpm, at -9°C for 15 min. The supernatant was then evaporated to dryness using vacuum concentrator (5301 vacuum concentrator, from Eppendorf). The dried extract was then stored at -80°C for further analysis.

7.2.5 Metabolite derivatization

Metabolite derivatization was performed immediately after the drying step, as suggested elsewhere (Lee and Fiehn, 2008a). Briefly, 30 µL of 20mg/mL methoxyamine hydrochloride in pyridine was added to the dried extract and samples were shaken for 80 min at 37°C to protect the aldehyde and ketone groups. The samples were then derivatized by trimethylsilylation of acidic protons by addition of 45 µL MSTFA (N-methyl-N-trimethylsilyltrifluoroacetamide) and incubating them further in shaking condition at 40°C for 80 min.

7.2.6 GC-MS analysis, metabolite identification and data analysis

GC-MS analysis, metabolite identification and data analysis was conducted as described in Chapter 3.

7.3 Results and Discussion

In the investigation discussed in this Chapter, we have directed our approach towards evaluating and minimising the leakage of intracellular metabolites from microalga *C. reinhardtii* during quenching treatments. To achieve our objective in a broader sense, we have designed the experiments and categorised them into three different approaches as illustrated in Figure 7.1. In all, the applied quenching treatments, temperature during the quenching processes were maintained below -20°C. Although metabolites with high turnover rates such as ATP and ADP might still remain active at -20°C, they cannot be detected with GC-MS based analysis. In most of the studies that evaluated the quenching methods for bacterial models (Faijes et al., 2007, Schädel et al., 2011) conclusions were drawn primarily based on ATP assay and adenylate energy charge. However such analysis confirms the disruption of in vivo metabolite ratios, however comment on quenching efficiency protocols cannot be made only based on these assay as it does not take into consideration the possible alterations in the rest of the metabolome. As with the Chapter 4, we have demonstrated high correlation between the results obtained with three different types of analyses which included ATP assay, SEM and GC-MS based analysis. Therefore, in the present study, we focussed primarily on GC-MS based analysis and the optimisation of quenching protocols were done by taking into consideration quantification of detectable compounds with GC-MS analysis.

In all the three approaches, evaluation and comparison within different treatments were based on two response variables where only features that were present in at least three biological replicates out of five were considered for further analysis:

- Response variable 1: Median number of metabolites recovered in;

a) Cell extracts
b) In quenching supernatant
c) Only in cell extract and not in supernatant (Unique to cells)
d) Only in supernatant and not in cell extract (Unique to supernatant) and
e) In both the cell extracts and supernatants (Common to both)

- Response variable 2: Recoveries of metabolites within twelve different classes of metabolites with respect to their normalised median peak intensities.

We analysed the unquenched samples, simply by centrifugation of the culture sample without any quenching step in order to get an estimate of extracellular metabolites accurately, as unquenched sample serves as a reliable standard for comparison of the extracellular metabolome data from the quenched sample. As we have discarded the washing step in all the applied quenching protocols as suggested in the literature (section 7.1), we have also analysed the culture media as a control for appropriate quantification of intracellular pool, to account for the leftover media components if any.

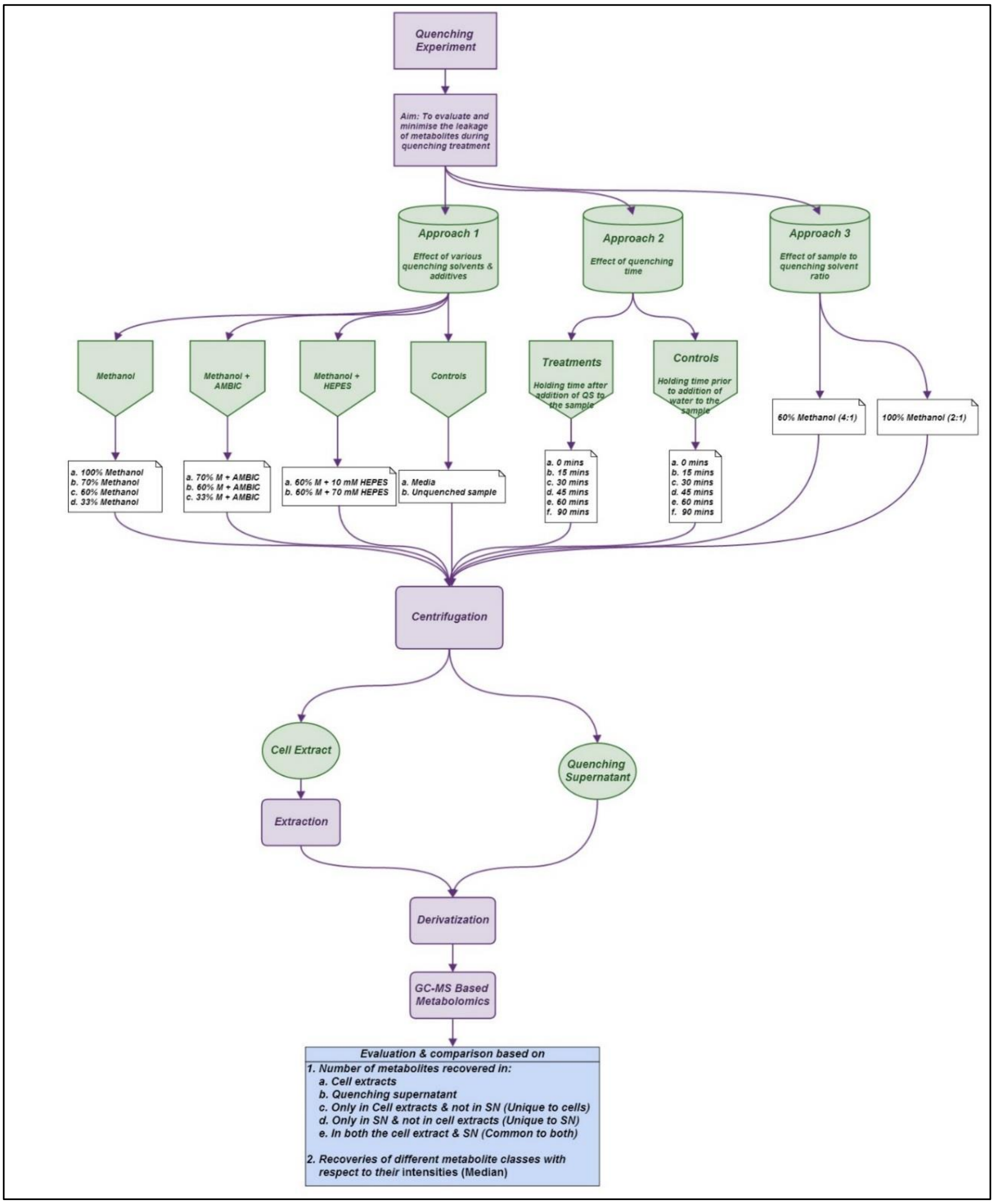

Quenching Experiment
Aim: To evaluate and minimise the leakage of metabolites during quenching treatment
Approach 1
Effect of various quenching solvents & additives
Approach 2
Effect of quenching time
Approach 3
Effect of sample to quenching solvent ratio
Methanol
Methanol + AMBIC
Methanol + HEPES
Controls
Treatments
Holding time after addition of QS to the sample
Controls
Holding time prior to addition of water to the sample
60% Methanol (4:1)
100% Methanol (2:1)
a. 100% Methanol
b. 70% Methanol
c. 60% Methanol
d. 33% Methanol
a. 70% M + AMBIC
b. 60% M + AMBIC
c. 33% M + AMBIC
a. 60% M + 10 mM HEPES
b. 60% M + 70 mM HEPES
a. Media
b. Unquenched sample
a. 0 mins
b. 15 mins
c. 30 mins
d. 45 mins
e. 60 mins
f. 90 mins
a. 0 mins
b. 15 mins
c. 30 mins
d. 45 mins
e. 60 mins
f. 90 mins
Centrifugation
Cell Extract
Quenching Supernatant
Extraction
Derivatization
GC-MS Based Metabolomics
Evaluation & comparison based on
1. Number of metabolites recovered in:
a. Cell extracts
b. Quenching supernatant
c. Only in Cell extracts & not in SN (Unique to cells)
d. Only in SN & not in cell extracts (Unique to SN)
e. In both the cell extract & SN (Common to both)
2. Recoveries of different metabolite classes with respect to their intensities (Median)

Figure 7.1 Experimental design and general workflow adopted for the evaluation and optimisation of quenching protocols for *C. reinhardtii*.

7.3.1 Effect of various quenching solvents, methanol concentration and inclusion of buffer additives on metabolite leakage (Approach 1)

In order to compare the degree to which extent of metabolite leakage and recoveries of intracellular metabolites were influenced by various quenching solvents and additives, cells of *C. reinhardtii* were harvested and quenched as described in section 7.2.3 with nine different quenching solvents and additives. Variation in quenching treatments were tested to investigate the different effects and the results are summarised in table 7.2. For an overall comparison, the harvested cell extracts, cell-free supernatant of both quenched and non-quenched cells along with the blank sample (culture medium) were analysed to determine the extent of leakage of intracellular metabolites during quenching. After monitoring cell-free supernatant of quenched, non-quenched cells and the blank medium, the necessary correction was done for appropriate calculation of intracellular metabolites as shown in figure 7.2.

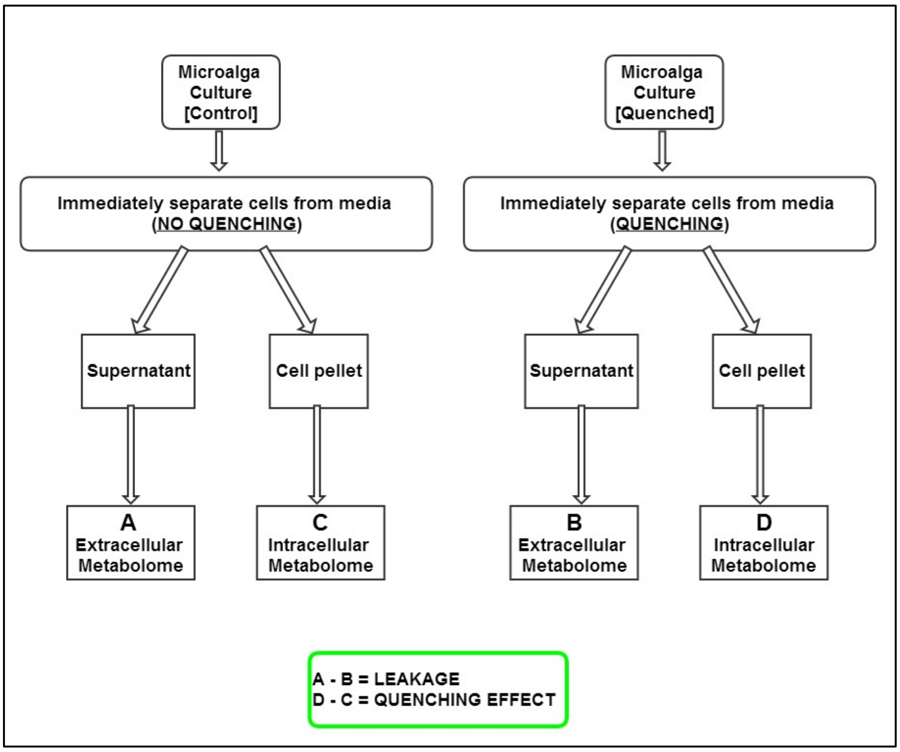

Microalga Culture [Control]
Microalga Culture [Quenched]
Immediately separate cells from media (NO QUENCHING)
Immediately separate cells from media (QUENCHING)
Supernatant
Cell pellet
Supernatant
Cell pellet
A Extracellular Metabolome
C Intracellular Metabolome
B Extracellular Metabolome
D Intracellular Metabolome
A - B = LEAKAGE
D - C = QUENCHING EFFECT

Figure 7.2 Schematic displaying the necessary correction required for appropriate calculation of intracellular metabolites after monitoring cell-free supernatant of quenched, non-quenched cells and the blank medium.

Table 7.2 Variation in quenching treatment tested to investigate the effects of varying methanol concentration, effect of inclusion of various buffer additives to quenching solution, effect of temperature and effect of sample to quenching solvent ratio on metabolite leakage

	Method code	Addition of buffer	Initial quenching solvent Temperature (°C)	Initial concentration of methanol solutions (v/v) (%)	Sample/quenching solvent ratio	Final concentration of methanol after quenching (v/v) (%)	Final temperature of resultant mixture after quenching (°C)
Approach 1	33M	x	-50	33	1:1	17	-20
	33A	0.85% AMBIC	-50	33	1:1	17	-20
	60M	x	-50	60	1:1	30	-30
	60A	0.85% AMBIC	-50	60	1:1	30	-30
	60H10	10 mM HEPES	-50	60	1:1	30	-30
	60H70	70 mM HEPES	-50	60	1:1	30	-30
	70M	x	-50	70	1:1	35	-35
	100M	x	-50	100	1:1	50	-40
Approach 3	60-41	x	-50	60	1:4	48	-40
	100-21	x	-50	100	1:2	77	-45
Controls	Control (Unquenched)	x	x	x	x	x	x
	Media	x	x	x	x	x	x

A summary of the unique recovery efficiency of all the nine quenching treatments is shown in Fig. 7.3, where the metabolite class and numbers detected for various treatments are plotted. Figure 7.3a summarises the metabolites detected in the cell pellets, Figure 7.3b those detected in the cell-free supernatant, Figure 7.3c indicates the metabolites present in the cells but not in the supernatants, Figure 7.3d indicates metabolites present in the supernatants and not in the cells, and Figure 7.3e metabolites present in both the cell pellet and the supernatant. Higher metabolite numbers in the supernatant (Fig 7.3b), relatively high numbers detected in both the cells and the supernatants (Fig 7.3e), and corresponding low numbers unique to the cell pellets (Fig 7.3c) indicates high metabolite leakage. Higher proportion of metabolites detected in Figure 7.3e, compared to that detected in Figure 7.3c or Figure 7.3d indicates that there is an increased chance of metabolite leakage.

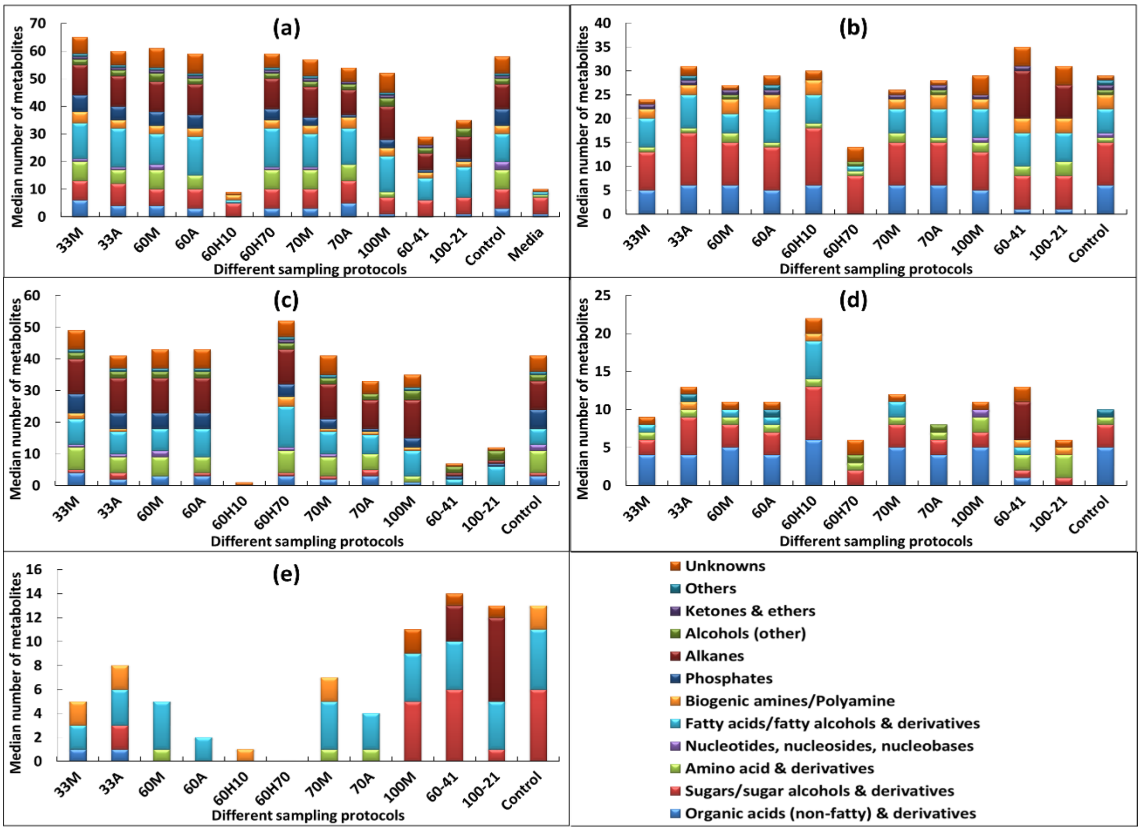
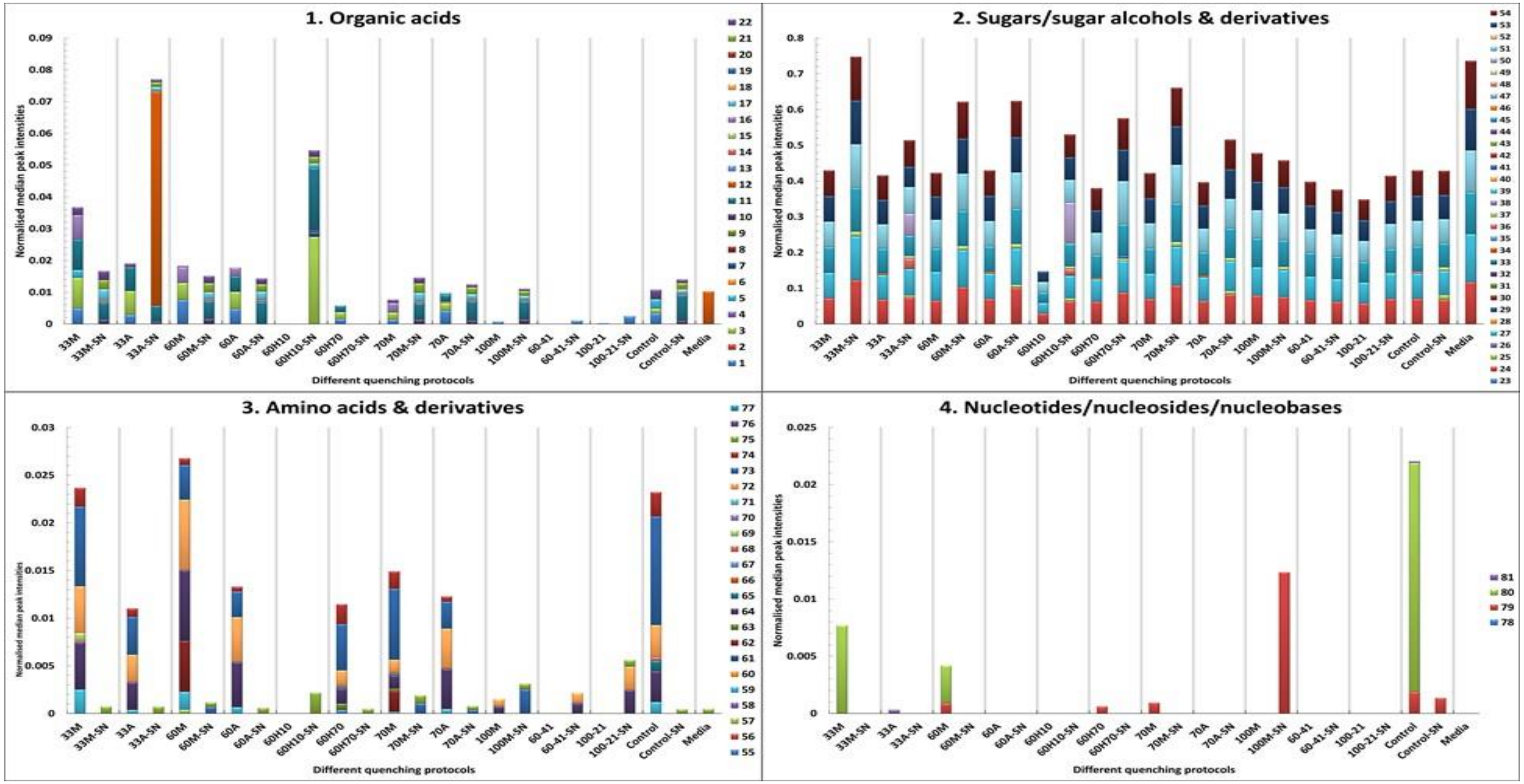


Figure 7.3 A summary of the unique recovery efficiency of all the nine quenching treatments involved in Approach 1 and 3. X-axis represents different sampling protocols: 33M = 33% methanol; 33A = 33% methanol + AMBIC; 60M = 60% methanol; 60A = 60% methanol + AMBIC; 60H10 = 60% methanol + 10 mM HEPES; 60H70 = 60% methanol + 70 mM HEPES; 70M = 70% methanol; 70A = 70% methanol + AMBIC; 100M = 100% methanol; 60-41 = 60% methanol with solvent to sample ratio 4:1 and 100-21 = 100% methanol with solvent to sample ratio 2:1. Y-axis represents median number of metabolites of each class. After all the treatment the extracted metabolites from cell extracts, cell-free supernatant post quenching and blank samples were analysed by GC-MS. (a) metabolites identified in cell extracts only, (b) metabolites identified in supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both).

As can be seen from figure 7.3c, 60% aqueous methanol supplemented with 70 mM HEPES (60H70) yielded higher recoveries of metabolites unique to cells as compared to other treatments. Correspondingly, metabolites common to both the cells and supernatant (Fig 7.3e) are the lowest. The number of metabolites detected in the supernatant is also the least compared to other treatments (Fig 7.3d). Therefore, among all the treatments 60% aqueous methanol supplemented with 70 mM HEPES (60H70) seems to preserve the integrity of microalgal cells resulting in higher recoveries of intracellular metabolites with minimal leakage into the extracellular environment.

Among other treatments, we did not observe any variations between quenching with 60% aqueous methanol and 60% aqueous methanol supplemented with 0.85% AMBIC. On the other hand, recovery of intracellular metabolites decreases with corresponding increase in extracellular levels as the methanol concentration increases. As can be seen from figure 7.3c, higher recovery of intracellular metabolites were observed with 33% aqueous methanol compared to 60%, 70% and 100% methanol whereas metabolites common to both the cells and supernatants (Fig. 7.3e) are lower with 33% aqueous methanol compared to other treatments. Among treatments where methanol was supplemented with 0.85% AMBIC, 60% methanol yielded higher recoveries of metabolites unique to cells (Fig.7.3c) compared to 33% and 70% aqueous methanol supplemented with AMBIC. Correspondingly, metabolites common to both the cells and supernatants (Fig. 7.3e) are the lowest. However as mentioned above we did not observe any major change in recoveries between 60% aqueous methanol alone compared to that of 60% aqueous methanol supplemented with 0.85% AMBIC, therefore use of 60% aqueous methanol alone would be advantageous over the use of methanol supplemented with AMBIC. Moreover, comparison within treatments where non-buffered methanol was used for quenching, quenching with 33% aqueous methanol seems to be a better option among non-buffered methanol and buffered with AMBIC. Therefore, in studies involving use of MS based hyphenated techniques (especially LC-MS), where use of salts/buffers as quenching additives introduces an additional source of variance to the experimental procedure, increases complexity of data and causes ion suppression in LC-MS, we strongly recommend use of 33% aqueous methanol as a second alternative to 60% aqueous methanol supplemented with 70 mM HEPES. However as demonstrated earlier in Chapter 4, addition of HEPES has no apparent interference with the derivatization reactions and GC-MS analysis. Our results are in contrast to that of Lee and Fiehn (Lee and Fiehn, 2008a), where authors suggested quenching with 70% aqueous methanol (-70°C) with sample to quenching solvent ratio of 1:1 resulted in minimal leakage of intracellular metabolites in *C. reinhardtii.*

With respect to analysis of metabolite classes recovered and median intensities of individual metabolites identified within each class, all the identified metabolites with the nine different quenching treatments were initially classified into twelve different classes (Table 4.2 from Chapter 4) based on their physicochemical properties. The results of the investigation are summarised in Fig. 7.4, where within individual metabolite class, median intensities of individual metabolites were plotted against various quenching treatments. Within each class the median intensities of individual metabolites identified in cell extracts and cell free quenching supernatant with various quenching treatments were plotted separately adjacent to each other. In total 197 unique metabolites were identified in approach 1 and 3 among all the applied quenching treatments in both the cell extract and the supernatant which are listed in Appendix 7.1.


1. Organic acids

2. Sugars/sugar alcohols & derivatives

3. Amino acids & derivatives

4. Nucleotides/nucleosides/nucleobases

Figure 7.4 Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis across various quenching treatments (from approach 1 and 3) was compared. X-axis represents different quenching protocols with their respective control sample where: 33M = 33% methanol; 33A = 33% methanol + AMBIC; 60M = 60% methanol; 60A = 60% methanol + AMBIC; 60H10 = 60% methanol + 10 mM HEPES; 60H70 = 60% methanol + 70 mM HEPES; 70M = 70% methanol; 70A = 70% methanol + AMBIC; 100M = 100% methanol; 60-41 = 60% methanol with solvent to sample ratio 4:1; 100-21 = 100% methanol with solvent to sample ratio 2:1 and SN = supernatant. Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 197 metabolites (See Appendix 7.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

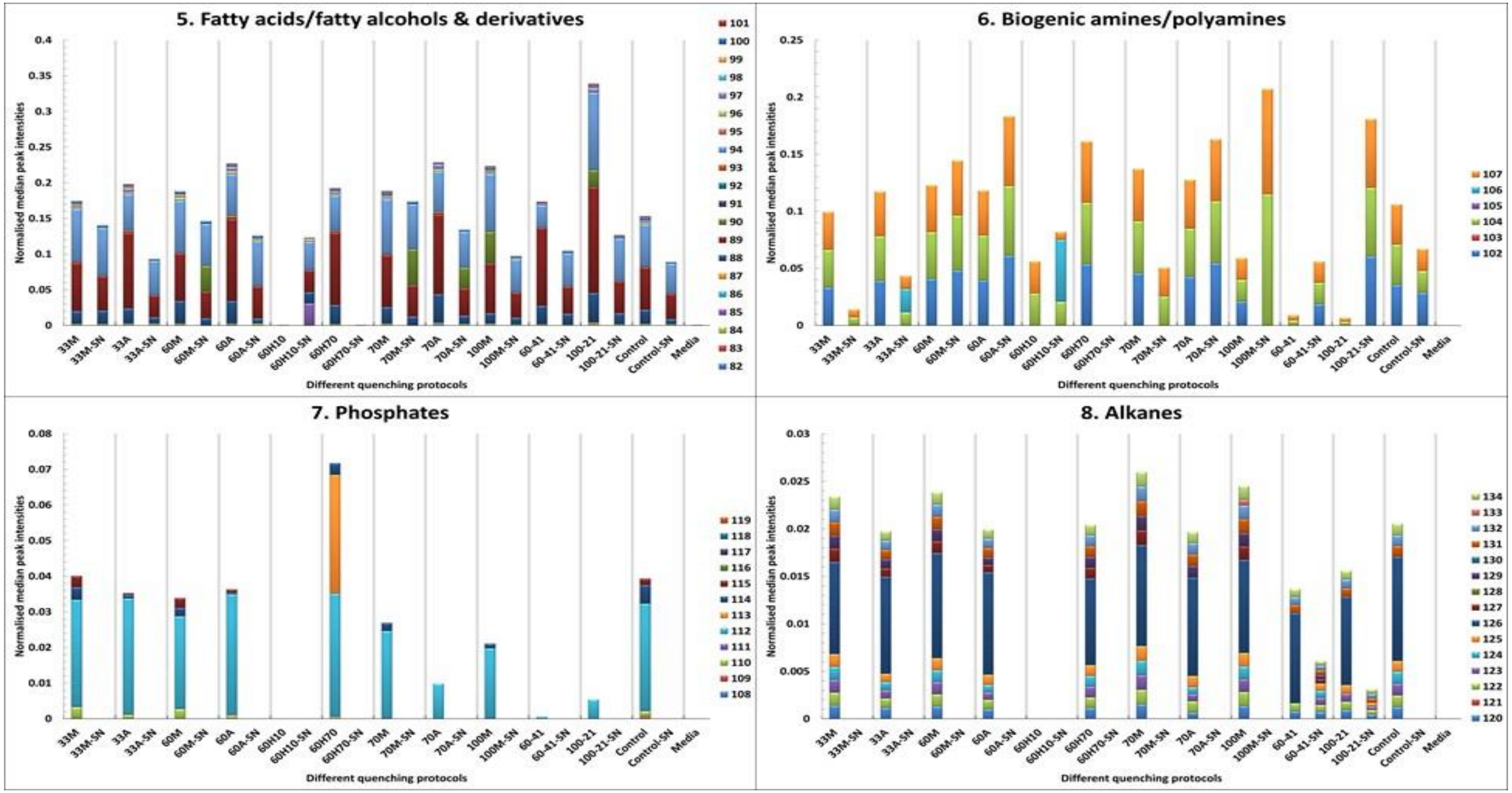


Figure 7.4 (continued) Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis across various quenching treatments (from approach 1 and 3) was compared. X-axis represents different quenching protocols with their respective control sample where: 33M = 33% methanol; 33A = 33% methanol + AMBIC; 60M = 60% methanol; 60A = 60% methanol + AMBIC; 60H10 = 60% methanol + 10 mM HEPES; 60H70 = 60% methanol + 70 mM HEPES; 70M = 70% methanol; 70A = 70% methanol + AMBIC; 100M = 100% methanol; 60-41 = 60% methanol with solvent to sample ratio 4:1; 100-21 = 100% methanol with solvent to sample ratio 2:1 and SN = supernatant. Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 197 metabolites (See Appendix 7.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

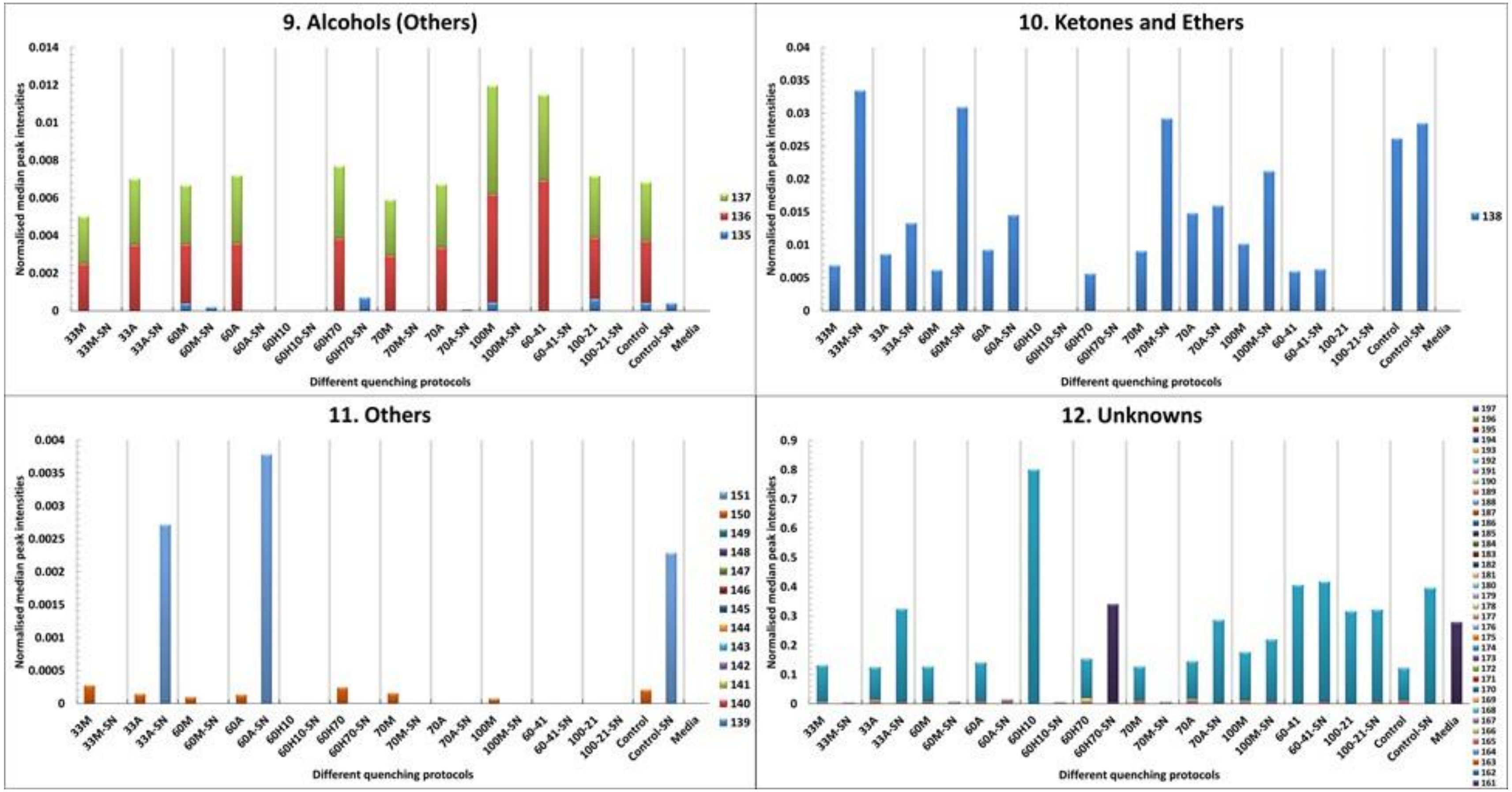


Figure 7.4 (continued) Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis across various quenching treatments (from approach 1 and 3) was compared. X-axis represents different quenching protocols with their respective control sample where: 33M = 33% methanol; 33A = 33% methanol + AMBIC; 60M = 60% methanol; 60A = 60% methanol + AMBIC; 60H10 = 60% methanol + 10 mM HEPES; 60H70 = 60% methanol + 70 mM HEPES; 70M = 70% methanol; 70A = 70% methanol + AMBIC; 100M = 100% methanol; 60-41 = 60% methanol with solvent to sample ratio 4:1; 100-21 = 100% methanol with solvent to sample ratio 2:1 and SN = supernatant. Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 197 metabolites (See Appendix 7.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

Recoveries of organic acids (Fig 7.4 (1))

Among all treatments, 33% aqueous methanol showed higher recoveries for organic acids in cell extracts (33M) and correspondingly lower leakage in supernatant (33M-SN). Compared to recoveries in the cell extract (33A), 33% aqueous methanol supplemented with AMBIC resulted in major loss of indole-2-carboxylic acid indicating higher leakage due to higher recoveries in supernatant (33A-SN). However it is important to note that higher recoveries of indole-2-carboxylic acid in culture media due to possible high carbon source concentration in the medium might have contributed to the higher recoveries in 33A-SN. Quenching with 100% methanol managed to recover only one organic acid with less intensity in cell extract (100M) and showed higher recoveries for many organic acids in supernatant (100M-SN) indicating higher leakage of organic acids. On the other hand 60% aqueous methanol supplement with 10 mM HEPES showed no recoveries in cell extract (60H10) and higher in supernatant (60H10) indicating major leakage of organic acids, whereas with 70 mM HEPES managed to recover few organic acids in cell extract (60H70) with none identified in supernatant (60H70-SN) indicating no leakage at all. Overall, superior recoveries of organic acids were observed with 33% aqueous methanol (33M) with minimal leakage.

Recoveries of sugars/sugar alcohols and derivatives (Fig. 7.4 (2))

Among all quenching treatments (except with 100M) lower recoveries were observed in cell extract compared to recoveries in supernatant. Possible reason might be higher concentration of sugars in media might have contributed to the higher recoveries in supernatant among all the quenching treatments. Moreover, the cell wall of *C. reinhardtii* is made up of complex multi-layered extracellular coat composed of carbohydrates, insoluble hydroxyproline-rich proteins and several chaotropic soluble glycoproteins without cellulose or hemicellulose, arranged into at least seven distinct layers (Goodenough and Heuser, 1985, Imam et al., 1985, Monk, 1988, Schwede et al., 2011). It is likely that the cell wall could have broken/disintegrated due to cold shock phenomenon during quenching treatments resulting in release of intermediate sugars/sugar alcohols into extracellular media contributing to the higher recoveries in supernatant. Another strong possibility might be excretion of sugar/sugar alcohols in extracellular media. The increased recovery of sugars/sugar alcohols in media compared to that recovered in cell extracts suggests that sugars/sugar alcohols might have been excreted from cells into extracellular media during metabolism.

Recoveries of amino acids and derivatives (Fig. 7.4 (3))

Among all treatments, 60% aqueous methanol (60M) followed by 33% aqueous methanol (33M) recovered higher number of amino acids and derivatives in cell extracts with good intensities. Whereas, the observed increase of amino acids in supernatant (100M-SN and 60H10-SN)) but not in cell extract (100M and 60H10) with 100% methanol and 60% aqueous methanol supplemented with 10 mM HEPES indicates lower recoveries in cell extract and higher leakage of amino acids. Our finding with 100% methanol treatment was in agreement with that of Britten and McClure (Britten and McClure, 1962) where authors demonstrated leakage of all free amino acids in E. *coli* upon variation in the osmolality of the surrounding medium. Remaining treatments managed to recover considerable amount of amino acids in cell extracts.

Recoveries of nucleotides/nucleosides/nucleobases (Fig. 7.4 (4))

Only few treatments showed recovery of one or two nucleotides/nucleosides/nucleobases in cell extracts which includes 33% (33M), 60% (60M), 70% (70M) aqueous methanol and 33% methanol supplemented with AMBIC (33A) and 60% aqueous methanol supplemented with 70 mM HEPES (60M70). Among all the applied treatments in total, only 4 metabolites were recovered within this class with GC-MS analysis. The possible reason for lower recoveries in both the cell extracts and the supernatant might be that GC-MS is not an ideal technique for analysing nucleotide. This has been previously suggested in a few studies. In such cases LC-MS or CE-MS (Ramautar et al., 2009, Soga et al., 2003) might be a suitable alternative to GC-MS in order to improve the metabolome coverage with respect to nucleotides/nucleosides/nucleobases. Our results were in agreement with the results reported elsewhere (Gonzalez et al., 1997, Villas-Bôas et al., 2005) where no detectable leakage of nucleotides were observed with the cold methanol solution (in our case this corresponds to the 60M treatment) in case of yeast sample.

Recoveries of fatty acids/fatty alcohols and derivatives (Fig. 7.4 (5))

With 100% methanol higher recoveries were observed in cell extract (100M) compared to the supernatant (100M-SN) indicating minimal leakage of fatty acids/fatty alcohols with this treatment. Apart from that, all the treatments where methanol was supplemented with 0.85% AMBIC (33A, 60A and 70A) showed considerably good recoveries in cell extract compared to that of observed in supernatants. No recovery was observed in cell extract and higher recoveries were observed in supernatant with the 60% aqueous methanol supplemented with 10 mM HEPES indicating higher leakage of fatty acids/fatty alcohols with this treatment. Superior recovery was observed with 60% aqueous methanol supplemented with 70 mM HEPES in cell extract (60H70) with very lower recoveries in supernatant (60H70-SN) indicating negligible leakage of fatty acids/fatty alcohols with this treatment. Somewhat similar conclusions were reported in previously published report (Zhao et al., 2014) where higher peak intensities were reported for fatty acids with glycerol-saline quenching solution and extraction with 100% methanol using freeze-thaw cycles. However, as discussed earlier in section 7.1, we did not consider quenching solvent that includes glycerol due its disadvantages. Furthermore, in the above study, the analysis was restricted to quantification of only specific classes of metabolites. On the other hand, in our study, we have adopted a broader approach involving quantification of nearly all the identified metabolites in 12 different metabolite classes. In addition, our results are contradictory to that of Lee and Fiehn (Lee and Fiehn, 2008a), where less leakage was reported with the use of 70% aqueous methanol whereas in our case minimal leakage was observed with the 60% aqueous methanol supplemented with 70 mM HEPES. However, it is important to note that authors simply adopted previously published quenching protocol with minor modification without critically evaluating and validating them with other quenching treatments.

Recoveries of biogenic amines/polyamine (Fig. 7.4 (6))

Among all the treatments, superior recovery for three major biogenic amines/polyamines was observed in either the cell extract or in the supernatant which includes ethanolamine (102), putrescine (104) and tryptamine (107). Higher recoveries were observed in supernatant than in cell extract with quenching treatments 60M-SN, 60A-SN, 60H10-SN, 70A-SN and 100M-SN indicating leakage of biogenic amines with these treatments, whereas higher recoveries in cell extract (60H70) were observed while no recoveries in supernatant (60H70-SN) indicating no leakage occurred with this treatment.

Recoveries of phosphates and alkanes (Fig. 7.4 (7 and 8))

Good recoveries of phosphates and alkanes were observed in the cell extract with minimal leakage into supernatant with all the treatments except with 60H10 treatment, where negligible amounts of phosphates and alkanes were recovered in both the cell extract and the supernatant. Among all the treatments higher recoveries were observed with 60% aqueous methanol supplemented with 70 mM HEPES (60H70). In case of phosphates, our results were in agreement with the original results reported elsewhere (Gonzalez et al., 1997, Villas-Bôas et al., 2005) where no detectable leakage of sugar phosphates were observed with the cold methanol solution (In our case 60M) in case of yeast sample. The possible explanation behind no leakage might be due to similar levels of such sugar phosphates as found in the culture media. Moreover, we found that gradual increase in the methanol concentration of quenching solvent resulted in gradual decrease in their corresponding intracellular levels which was in further agreement with a previously published report (Canelas et al., 2008).

Recoveries of alcohols and ketones and ethers (Fig. 7.4 (9 and 10))

Among all the applied treatments, higher recoveries were observed for alcohols in the cell extract with no recoveries in the supernatant indicating no leakage with 100% methanol (100M) treatment. Similarly, good recoveries were observed in the cell extract with all other treatments except with 60M-SN and 60H70-SN where minor leakage was observed for one metabolite in the supernatant whereas no recoveries were observed for alcohols and ketones in both the cell extract 60H10) and in the supernatant (60H10-SN) with 60% aqueous methanol supplemented with 10 mM HEPES. In case of ketones, all the treatments (except with 60H70) showed higher recoveries in the supernatant than in the cell extract indicating leakage of ketones and ethers. Only 60% aqueous methanol supplemented with 70 mM HEPES managed to recover higher intensities in the cell extract with no recoveries in the supernatant indicating no leakage.

Recoveries of others and unknown classes (Fig. 7.4 (11 and 12))

Only few other classes of metabolites were recovered among all the treatments in the cell extracts, except with 60H10 and 70A treatments where no recoveries were observed in both the cell extract and in the supernatant. On the other hand large numbers of unknown metabolites were recovered in both the cell extract and supernatant with all the applied treatments. Unknown class stands for metabolites or mass spectral tags (MSTs) which were not fully identified in the database due to the absence of authenticated pure reference standards however their other properties resembles to those metabolites which were identified very well and their MSTs were properly documented in the database. Variable recoveries were observed for unknowns with all the applied treatments in both the cell extract and in the supernatant leading to mixed conclusions.

In summary, recovery of intracellular metabolites decreases with corresponding increase in extracellular levels as the methanol concentration increases. Our results were in agreement with that of de Jonge and co-workers (de Jonge et al., 2012, Wittmann et al., 2004) whereas in contrast to that of previously published reports where authors reported increased leakage from *S. cerevisiae* (Canelas et al., 2008) and *E. coli* (Schädel et al., 2011) with the corresponding increase in methanol concentration. Conclusions drawn with a bacterial model were primarily based on results from ATP assay, flow cytometry and OD based assay and does not include highly sensitive MS based hyphenated techniques. Moreover, the results of all the applied techniques does not correlates with each other. On the other hand, conclusions drawn from study with *S. cerevisiae*, involved quantification of only 35 metabolites (all were hydrophilic compounds) compared to our present study (151 identified metabolites). Contradictory conclusions to our findings were also reported by Tredwell and co-workers where no major difference in leakage from *Pichia pastoris* were observed with the varying concentration of methanol and inclusion of various buffer additives in quenching solvent (Tredwell et al., 2011). Furthermore, Canelas and co-workers tested influence buffer additives (HEPES, AMBIC and/or tricine) in methanol on metabolite leakage from yeast, and reported no significant benefit in buffering or increasing the ionic strength of the quenching solvent, in contrary authors reported slightly lower intracellular recoveries (Canelas et al., 2008), which was completely contradictory to our finding with *C. reinhardtii*. The possible reasons behind this contradiction might be differences in cell/sample type. Moreover it is important to note that conclusions were drawn from estimation of only two specific metabolites which were being used as a representative examples for the respective class. Similarly, in contrast to our findings where 60% aqueous methanol supplemented with 70 mM HEPES yielded highest intracellular recoveries with minimal leakage, Schädel and co-workers (Schädel et al., 2011) reported higher leakage in *E. coli* with the inclusion of HEPES to methanol compared to conventional 60% methanol. These contradictory observations further support our conclusions derived in Chapter 3 and 4, that sampling and quenching techniques are highly sample/cell dependent and needs critical evaluation and validation for every organism under investigation before being adopted for a quantitative metabolomics study.

7.3.2 Effect of prolonged exposure to quenching solvent on metabolite leakage (Approach 2)

Ideally, quenched cells should be processed as quickly as possible in order to avoid the leakage of intracellular metabolites into the extracellular environment, as prolonged contact time of cells with the quenching solvents might increase the chances of metabolite losses via diffusion of small metabolites through the cell membrane. In 1992, De Koning and Van Dam (Koning and Dam, 1992) demonstrated no leakage of intracellular metabolites from yeast sample after exposure of quenching solvent to sample for 30 minutes. This has been evaluated in past for *S. cerevisiae* (Canelas et al., 2008) and for fungus *Penicillium chrysogenum* (de Jonge et al., 2012), however there are no reports of such studies carried out on microalgal samples. Therefore, to test this theory for microalga, cells of *C. reinhardtii* were exposed to the quenching solvent (60% v/v aqueous methanol) for a prolonged period of various time intervals to evaluate the extent of metabolite leakage. Briefly, 1mL of cell suspension was rapidly plunged into 1mL of quenching solvent as described in section 7.2.3. To evaluate the extent of metabolite leakage in response to prolonged exposure to the quenching solvent, broader range of time intervals were selected including 0, 15, 30, 45, 60 and 90 min. In case of 0 min treatment, samples were processed immediately by centrifugation, whereas for other treatments prolonged exposure was achieved by leaving the quenched samples at -40°C for the above specified time intervals prior to centrifugation. Parallel set of samples were processed as control for each time intervals. For example in case of control sample for 15 min time interval, 1 mL of cell suspension was harvested and allowed to stand for 15 min followed by addition of 1mL of water and subsequently centrifuged. The addition of water was done in order to account for the variations caused by dilution. Water was purposely selected in this case instead of any other buffer solution such as PBS, as addition of PBS might result in extremely higher concentration of phosphates and will interfere in subsequent GC-MS analysis or will result in overestimation of phosphates in intracellular pools. For an overall comparison, the harvested cell extracts, cell-free supernatant of both quenched (samples) and non-quenched cells (controls) for all the above specified time intervals along with the blank sample (culture medium) were analysed to determine the extent of metabolite leakage in response to prolonged exposure to the quenching solvent. After monitoring cell-free supernatant of both quenched, non-quenched cells and the blank sample, the necessary correction was done for appropriate calculation of intracellular metabolites as shown in figure 7.2. A summary of the unique recovery efficiency of all the applied treatments within this approach is shown in Fig. 7.5, where the metabolite class and median numbers detected for various treatments are plotted.

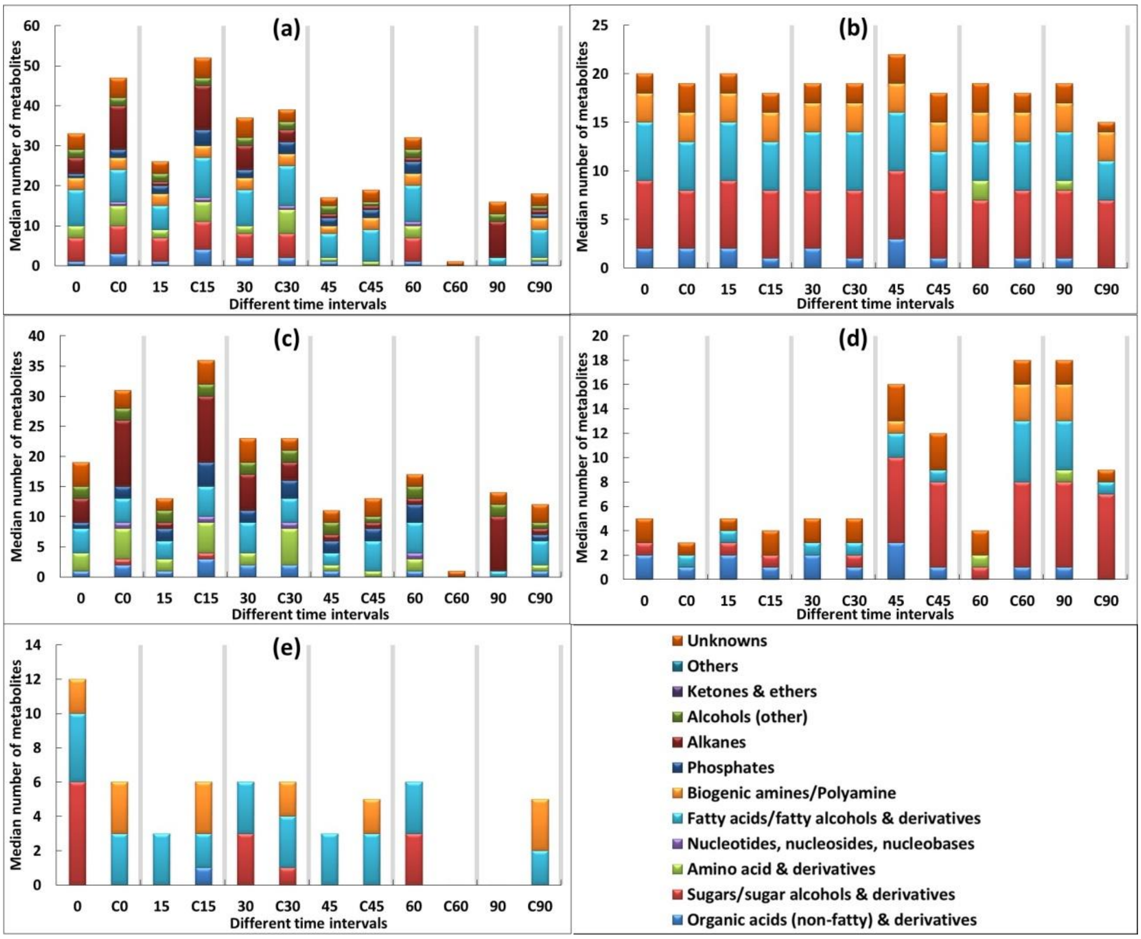


Figure 7.5 A summary of the unique recovery efficiency of all the six quenching treatments involved in Approach 2. X-axis represents different time intervals along with their control samples (C). Y-axis represents median number of metabolites of each class. After all the treatment the extracted metabolites from cell extracts, cell-free supernatant post quenching and blank samples were analysed by GC-MS. (a) metabolites identified in cell extracts only, (b) metabolites identified in supernatant only, (c) metabolites present only in cell extract (and not in the supernatant) - unique to cells; (d) metabolites present only in supernatants (and not in cell extract) - unique to supernatants, (e) metabolites present in both the cell extract and supernatant – (common to both).

As can be seen from figure 7.5c, samples processed at 0 time interval yielded higher recoveries of metabolites compared to all other treatments except for 30 min interval where surprisingly higher numbers of metabolites were detected than that of 0 min time interval. A similar trend was observed with the control samples where higher recoveries were observed with C0 treatment compared to other treatments except for C15 where again surprisingly higher recoveries were observed than that of C0. The possible reason behind higher recoveries might be degradation or inter-conversion of metabolites. Correspondingly higher number of metabolites unique to the supernatant (Fig 7.5d) were detected as the contact time of cell suspension was prolonged for more than 30 min with that of quenching solvent. Approximately four fold increase in recoveries of extracellular metabolite numbers was observed with 45, 60 and 90 min treatments compared to that of 0 min treatment (Fig 7.5d). No variations in recoveries of metabolites classes and numbers were observed between 0, 15 and 30 min time intervals whereas small increase in the recoveries of metabolites were observed with control treatments as contact time was increased from 0 to 15 and then to 30 min indicating increased metabolite leakage.

With respect to analysis of metabolite classes recovered and the median intensities of individual metabolites identified within each class, all the identified metabolites with the different quenching treatments were initially classified into twelve different classes based on their physicochemical properties. The results of the investigation are summarised in Fig. 7.6, where within individual metabolite class, the median intensities of individual metabolites were plotted against various quenching treatments. Within each class the median intensities of individual metabolites identified in cell extracts and cell free quenching supernatant with various quenching treatments were plotted separately adjacent to each other. Among all the applied treatments within approach 2, in total 165 unique metabolites were identified in both the cell extract and the supernatant which are listed in Appendix 7.2.

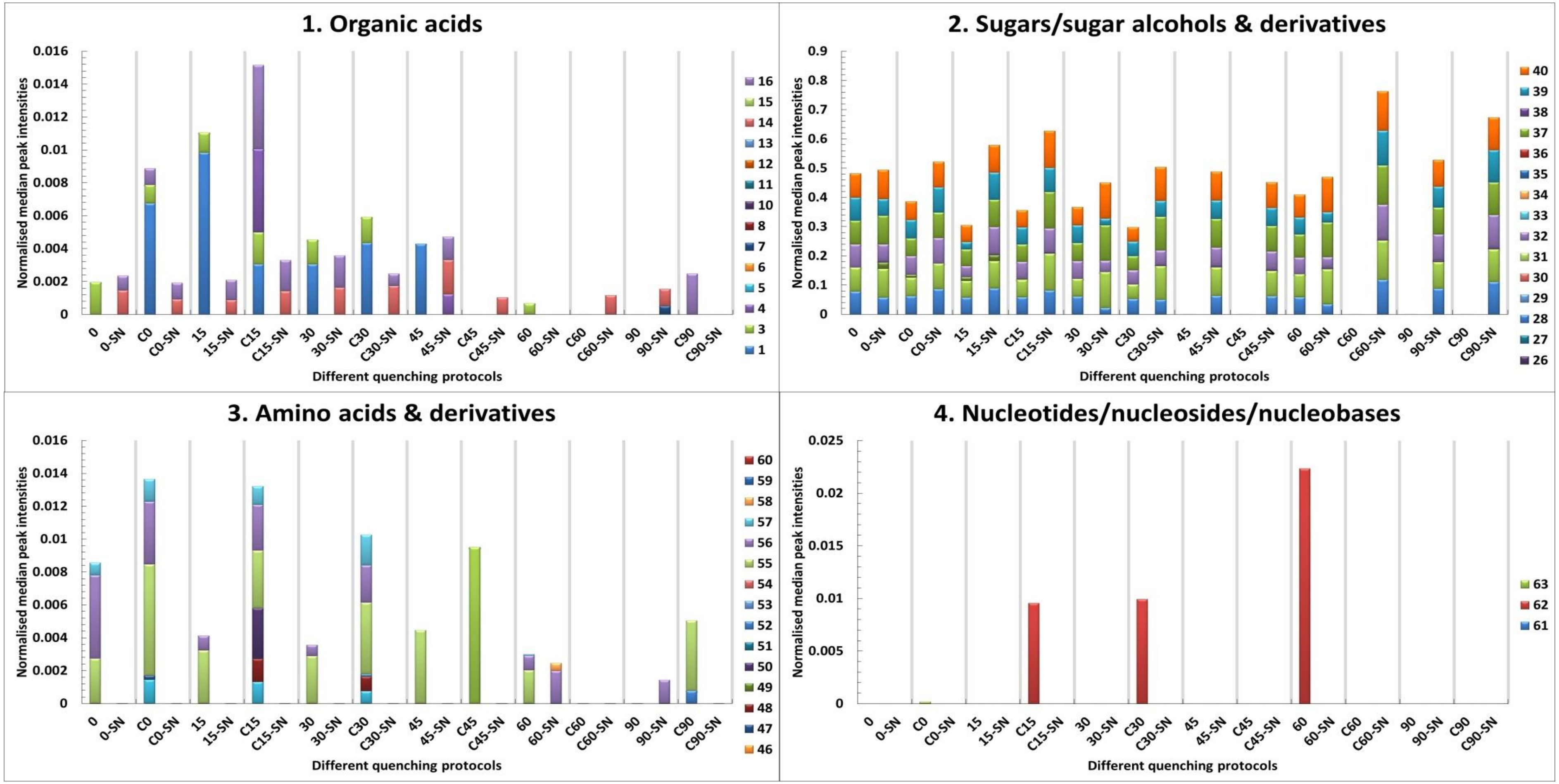


Figure 7.6 Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis across various quenching treatments (from approach 2) was compared. X-axis represents different time intervals in minutes along with their control samples (C) whereas, SN = supernatant. Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 165 metabolites (See Appendix 7.2) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

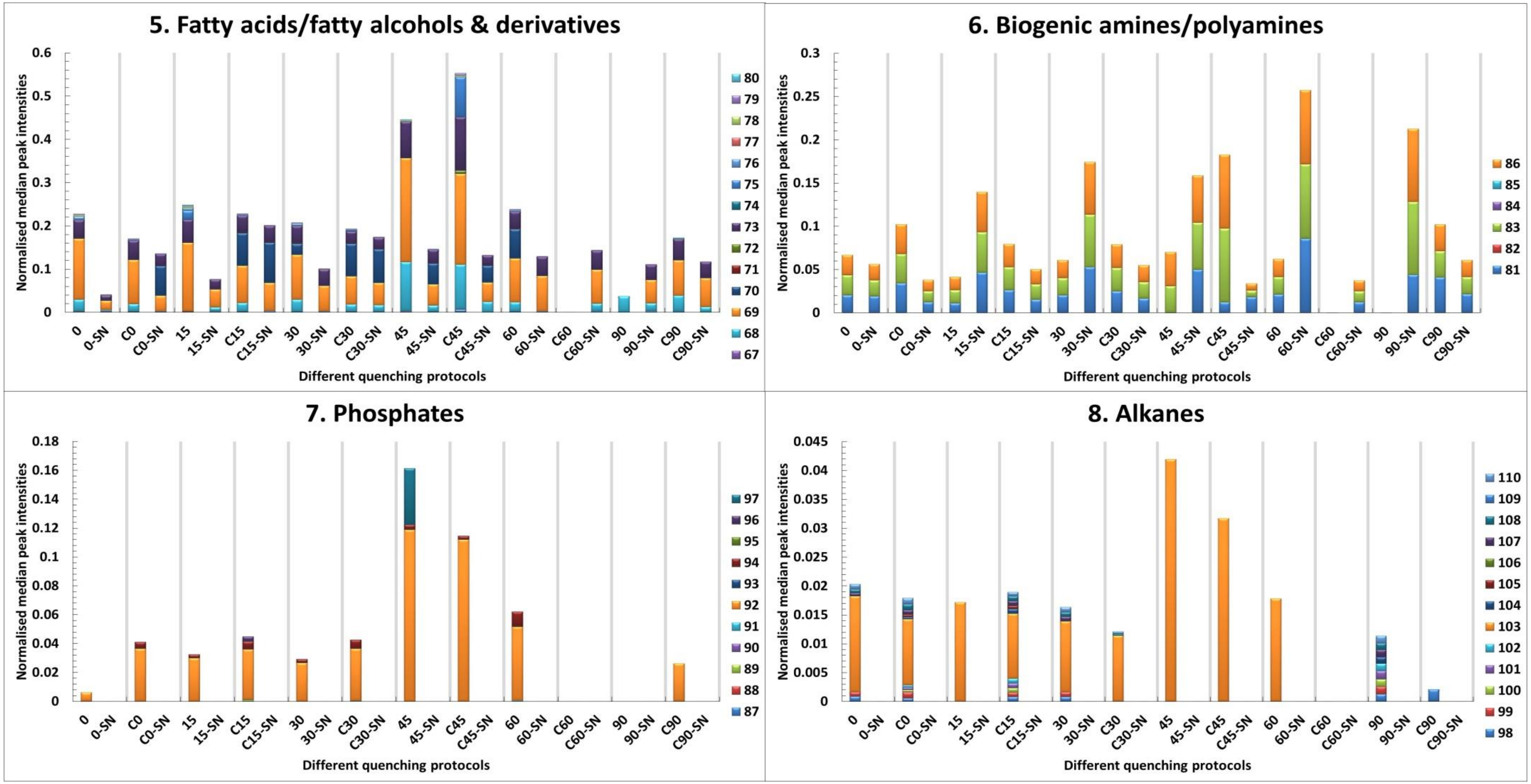


Figure 7.6 (continued) Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis across various quenching treatments (from approach 2) was compared. X-axis represents different time intervals in minutes along with their control samples (C) whereas, SN = supernatant. Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 165 metabolites (See Appendix 7.2) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

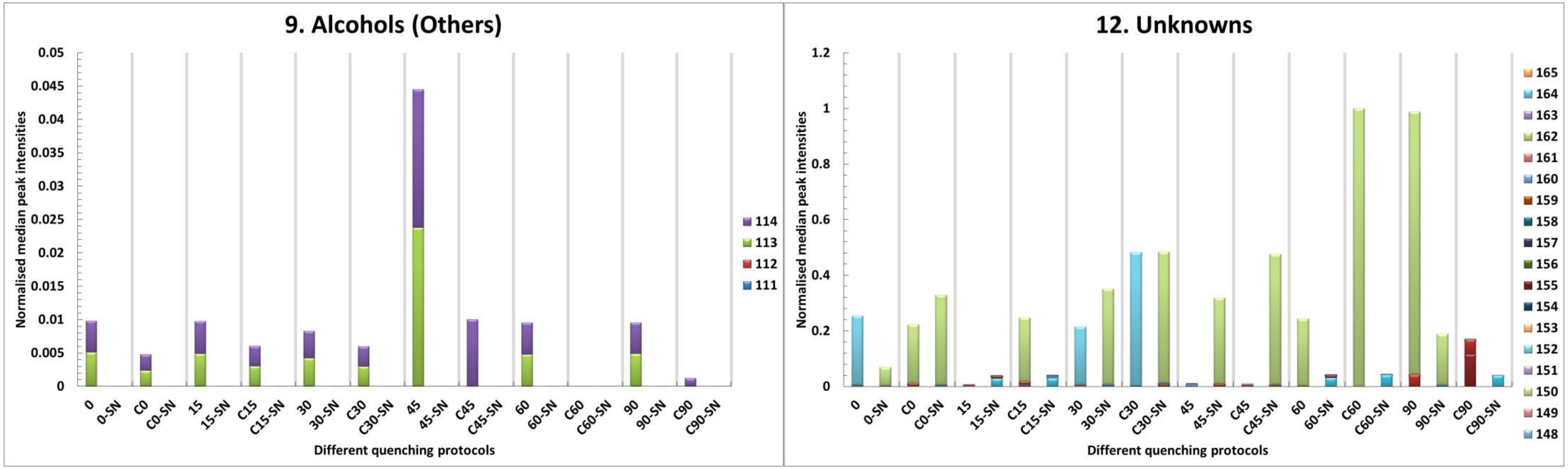


Figure 7.6 (continued) Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis across various quenching treatments (from approach 2) was compared.. X-axis represents different time intervals in minutes along with their control samples (C) whereas, SN = supernatant. Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 165 metabolites (See Appendix 7.2) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

Recoveries of organic acids (Fig 7.6 (1))

Variable median intensity for fewer organic acids was observed across all the treatments where no significant conclusions can be drawn from such observations. However, total median intensities of all organic acids showed gradual decrease in cell extract recoveries from 15 min onwards up to 45 min hold time with corresponding increased recoveries in supernatant indicating leakage, whereas minor recoveries were observed in cell extracts with 45, 60 and 90 min interval suggesting complete degradation or inter-conversion of all organic acids. With respect to analysis of specific metabolites, intracellular levels of 2-piperidinecarboxylic acid (1) was not observed with 0 time point compared to control (C0) indicating leakage, whereas surprisingly the leakage was minimal at 15 min time point compared to control (C15). However 15 min time point showed marked leakage for erythronic acid (4) and threonic acid (16), indicating correlation for specific metabolites between their levels in the supernatant and in the cell extracts, whereas there was no such correlation found for overall recovery of all organic acids between 0 and 15 min time points and their respective controls.

Recoveries of sugars/sugar alcohols and derivatives (Fig. 7.6 (2))

In case of sugars/sugar alcohols and derivatives lower recoveries for intracellular levels were observed as the exposure of cell suspension to the quenching solvent was increased from 0 to 30 minutes, indicating time dependant slow release of sugars into the extracellular medium. At 45 and 90 min exposure time absolutely no recoveries were observed in the cell extracts whereas higher recoveries were observed in the supernatants indicating complete leakage of these metabolite class at given hold time.

Recoveries of amino acids and derivatives (Fig. 7.6 (3))

In case of amino acids and derivatives no leakage was observed up to 30 min exposure with good intracellular levels for fewer amino acids, whereas at 45 min time interval only single amino acid was recovered in cell extract. While at 60 and 90 min hold time increased recoveries in supernatant indicates slower time dependant leakage of fewer amino acids from cells. Our findings were in agreement with previously published report where leakage of amino acids was increased with the increase in contact time of sample with the quenching solvent (Villas-Bôas et al., 2005). In case of control samples marked decrease in intracellular levels of amino acids were observed with increase in exposure time.

Recoveries of nucleotides/nucleosides/nucleobases (Fig. 7.6 (4))

Among all the applied treatments recoveries of nucleotides/nucleosides/nucleobases was negligible in both the cell extract and in the supernatant, this might be due to possible reasons as discussed in section 7.3.1.

Recoveries of fatty acids/fatty alcohols and derivatives (Fig. 7.6 (5))

Gradual increase in extracellular levels for this class of metabolites were observed as the hold time increased from 0 to 45 min (0-SN, 15-SN, 30-SN and 45-SN) indicating slower leakage of these less polar or semi-polar metabolites via diffusion. At 90 min hold time, higher recoveries in numbers and median intensities were observed for extracellular levels than intracellular levels indicating higher leakage. In case of control samples, among all the treatments except that of C45 min hold time, gradual increase in extracellular levels were observed compared to intracellular levels as exposure to quenching solvent was increased from C0 time point up to C30 min time point, whereas absolutely no intracellular recoveries were observed with C60 with correspondingly higher extracellular levels suggesting complete leakage of fatty acids at this time point. Possible reason behind leakage might be prolonged exposure to organic solvent (methanol in this case) which facilitates non-specific membrane lipid extraction as suggested elsewhere (Schädel et al., 2011).

Recoveries of biogenic amines/polyamine (Fig. 7.6 (6))

Major polyamines in *C. reinhardtii* include putrescine (number 83) which is a multifunctional molecule responsible for normal growth of the cell, gene expression, chromatin structure and precursor of the two other polyamines namely spermidine and spermine (Theiss et al., 2004, Theiss et al., 2002, Voigt et al., 2000). Putrescine and other biogenic amines were detected with all the applied quenching treatments. The results suggest slower time dependant leakage of all biogenic amines/polyamines, as extracellular levels were gradually increased with corresponding decrease in the intracellular levels as the exposure to quenching solvent was increased from 0 min to 45 min (except for 45 min treatment). In case of 60 and 90 min exposure no recoveries were observed in cell extracts while higher extracellular levels indicated major leakage of these metabolites.

Recoveries of phosphates and alkanes (Fig. 7.6 (7 and 8))

In case of phosphates, absolutely no recoveries were observed in the quenching supernatant among all the applied treatments. Possible reason might be larger and more polar nature of these metabolites making their diffusion through the cell wall difficult as suggested elsewhere in case of yeast and fungi samples (Canelas et al., 2008, de Jonge et al., 2012), leading to less leakage compared to other classes such as amino acid and organic acids which are smaller compounds and might be released more easily than the larger ones. However, despite their non-polar nature, similar trends as that of phosphates were observed with alkanes where no recoveries were observed in the quenching supernatant. The possible reason might be their larger size making their diffusion through the cell wall difficult as most of the alkanes that were recovered have MW ranging from 212 to 352.

Recoveries of alcohols and ketones and ethers (Fig. 7.6 (9 and 10))

In case of alcohols similar trend was observed as that of with phosphate and alkanes, where no recoveries were obtained with the quenching supernatant among all the applied quenching treatments justifying similar reasons as discussed above for phosphates. In case of intracellular levels obtained negligible variations were observed with 0, 15 and 30 min treatment indicating there was no metabolite leakage despite prolonged exposure to the quenching solvent up to 30 min. No recoveries for ketones and ethers were observed in both the cell extracts and in the supernatant among all the applied treatments.

Recoveries of others and unknown classes (Fig. 7.6 (11 and 12))

We did not observed any recoveries for other classes of metabolites among all the applied treatments whereas in case of unknowns larger intra- and extracellular levels were observed among all the applied treatments. As the exposure time to the quenching solvent was increased from 0 min to 45 min, gradual decrease in the intracellular levels with corresponding increase in the extracellular levels clearly indicates metabolite leakage.

In summary, nearly all the metabolite classes showed gradual decrease in the intracellular levels with corresponding increase in the extracellular levels, as the contact time of sample to quenching solvent was increased. Our findings were in agreement with previously published reports, where similar conclusions were drawn with yeast and bacterial models (Canelas et al., 2008, de Jonge et al., 2012, Villas-Bôas et al., 2005).

7.3.3 Effect of quenching solvent to culture ratio (temperature influence) (Approach 3)

Influence of sample to quenching solvent ratio was studied in past for *S. cerevisiae* and *E. coli* samples where quenching with pure methanol at sample to solvent ratio of 1:5 at -40°C showed reduced leakage of intracellular metabolites compared to use of 60% aqueous methanol. In order to test this parameter with microalgal sample and to study the influence of quenching solvent and sample mixture, temperature and final concentration of methanol after quenching, 1 mL culture of *C. reinhardtii* was harvested and rapidly quenched with either 100% methanol with quenching solvent to sample ratio of 2:1 or with 60% aqueous methanol with quenching solvent to sample ratio of 4:1. Both the quenching solvents were pre-chilled to -50°C prior to quenching. Addition of the cells to the quenching solution increased the temperature by no more than 15˚C, thereby keeping the resulting mixture temperature below -20°C sufficient to stop the metabolism as demonstrated in past (Schädel et al., 2011, Wellerdiek et al., 2009b, Weuster-Botz, 1999). For an overall comparison, the harvested cell extracts, cell-free supernatant of both quenched and non-quenched cells along with the blank sample (culture medium) were analysed and the necessary correction was done for appropriate calculation of intracellular metabolites and to determine the extent of leakage of intracellular metabolites during quenching. The results of the investigation are summarised in figure 7.3.

With both the treatments 60-41 and 100-21 higher metabolite numbers in the supernatant (Fig 7.3b), relatively high numbers detected in both the cells and the supernatants (Fig 7.3e), and corresponding low numbers unique to the cell pellets (Fig 7.3c) were observed suggesting high metabolite leakage. Higher proportion of metabolites detected in Figure 7.3e, compared to that detected in Figure 7.3c or Figure 7.3d confirms that there is an increased chance of metabolite leakage with both the treatments. With this approach we did not observe any significant improvements in preserving the cell integrity by altering the final methanol concentration in resultant mixture after quenching. Moreover, increment in quenching solvent to sample ratio in order to keep the temperatures of the resulting mixture below -20°C for effective halting of metabolism seems to have no influence on preserving the cell integrity and in minimising the metabolite leakage. In contrast, higher leakage was observed with this approach compared to approach 1.

With respect to analysis of metabolite classes recovered and median intensities of individual metabolites identified within each class, all the identified metabolites were initially classified into twelve different classes based on their physicochemical properties. The results of the investigation are summarised in Fig. 7.4, where within individual metabolite class, median intensities of individual metabolites were plotted against various quenching treatments.

In case of amino acids, organic acids, biogenic amine/polyamines, sugars/sugar alcohols and ketones and ethers very low recoveries were obtained in both the cell extracts and in the supernatant with both the quenching treatments compared to those in approach 1, where higher median peak intensities were observed in the supernatant compared to that in cell extracts, indicating higher leakage. With both the treatments no recoveries were observed for nucleotides/nucleosides/nucleobases and other classes of metabolites. Superior recoveries of fatty acids/fatty alcohols and derivatives were obtained in the cell extract with the 100-21 treatment with corresponding decrease in extracellular levels indicating minimal leakage. Moreover the median peak intensities recovered for all the metabolites within this class was superior compared to all the applied treatments in approach 1. A similar trend was observed in recoveries of alcohols with 100-21 treatment.

7.4 Conclusions

In summary, we have performed the comprehensive investigation of appropriate sampling and quenching methods for *C. reinhardtii* using GC-MS, which involved untargeted quantitative analysis of several intracellular and extracellular metabolites. Our optimised miniaturised quenching method requires only 1 mL of microalgal culture, thereby subsequently reducing the volume of quenching solvent required, suitable for processing larger number of samples within shorter period of time, ideal for metabolomics investigation where a large number of samples need processing. To date, there is no effective alternative to cold methanol quenching method, where quenching can only be achieved by sharp temperature drift which usually leads to leakage caused by cold shock phenomenon. Hence, for reliable metabolome analysis it is essential to measure the metabolite levels in all possible fractions and establish mass balance to trace the fate of metabolites during quenching treatment. Our results clearly showed higher losses of intracellular metabolites with the use of conventional 60% methanol for quenching. Analysis with respect to influence of varying methanol concentration in quenching solvent on the extent of metabolite leakage clearly showed higher leakage with the increase in methanol concentration. On the other hand, analysis with respect to inclusion of various buffer additives to quenching solvent, 60% aqueous methanol supplemented with 70 mM HEPES managed to recover higher intracellular levels for nearly all metabolite classes with minimal leakage, supporting our previous findings with adherent mammalian cells. Increment in quenching solvent to sample ratio in order to keep the temperatures of the resulting mixture below -20°C for effective halting of metabolism seems to have no influence on preserving the cell integrity and in minimising the metabolite leakage. In contrast, higher leakage was observed with this approach. The study of prolonged exposure of sample to quenching solvent has shown higher amount of leakage as the contact time of sample to that of quenching solvent increases. Hence for quantitative metabolomics studies in *C. reinhardtii*, we strongly recommend quenching with 60% aqueous methanol supplemented with 70 mM HEPES (-40°C) at 1:1 sample to quenching solvent ratio, as it resulted in higher recoveries for intracellular metabolites with subsequent reduction in the metabolite leakage for all classes of metabolites. We believe the outcome of this research and the optimised quenching method can be directly applied and extended to other similar freshwater microalgal cultures or indeed cells with a similar cell envelope architecture studies and provide a standardised approach to metabolomics validations in other organisms.

Chapter 8

Influence of extraction solvents on metabolome recovery in the model microalga *Chlamydomonas reinhardtii*

8. Influence of extraction solvents on metabolome recovery in the model microalga *Chlamydomonas reinhardtii*

8.1 Introduction

The metabolite extraction is a fundamental step in most of the metabolomics workflows required for analysis of intracellular metabolites (Kell et al., 2005). For an efficient extraction of all intracellular metabolites, it is essential to permeabilise the cell wall (where it exists) and/or cell membrane and allow for extraction of all the intracellular metabolites with the aid of an extracting agent. To date, a range of extraction solvents have been proposed such as acids, bases, organics, alcohols, water and the usage of hot or cold temperatures. However, the selection of an optimal extraction solvent and method seems to be based on the metabolite classes of interest and more importantly the biological material under investigation (Duportet et al., 2012). A comprehensive summary of such key extraction solvents and techniques applied so far for the metabolome analysis of various prokaryotic and eukaryotic samples are summarised in table 8.1.

From Table 8.1, it is clear that the analysis of compatibility between the extraction solvent and the subsequent analytical platform is gaining more attention and reveals a trend towards selection of more mild extraction conditions such as use of cold organic solvents. In the past, harsh buffers containing strong acids such as perchloric acid (PCA) or alkali such as potassium hydroxide (KOH) were commonly employed as routine methods for the extraction of acid and alkali stable compounds from animal and plant tissues (Kopka et al., 1995), microorganisms (Buchholz et al., 2002, Prasad Maharjan and Ferenci, 2003b) and filamentous fungi (Hajjaj et al., 1998). Acidic extraction was commonly employed for extraction of nucleotides and hydrophilic metabolites. However recent findings suggest that extractions with PCA and KOH results in lower number of data points with poor reproducibility compared to mild extraction solvents, as most of the metabolites are unstable at low or high pH conditions (Hiller et al., 2007b, Madla et al., 2012, Prasad Maharjan and Ferenci, 2003b, Winder et al., 2008). Moreover, these methods are time consuming, as they require neutralisation of the sample at later stage (Dietmair et al., 2010). In addition, these hard methods tend to focus on specific chemical classes of metabolites (acids or amines) (Lee and Fiehn, 2008a), thus requiring additional methods for covering other classes of metabolites and cause huge interference with a range of analytical techniques. Dietmair and co-workers reported poor recoveries of intracellular metabolites from mammalian samples with alkali extraction due to formation of a viscous precipitate after centrifugation, making separation of supernatant from the cells extremely difficult.

Table 8.1 Summary of the main different extraction solvents and techniques applied so far for the metabolome analysis of various prokaryotic and eukaryotic samples

Sample	***Solvents used***	***Ratio***	***Volume added***	***Extraction Technique***	***Analytical Technique***	***Recommended***	***Reference***
Human blood plasma	*Methanol*	*DOE*	*800 uL*	*Vibration mill*	*GC-ToF-MS*	*Methanol:Water (8:1)*	*(Trygg et al., 2005)*
	Acetonitrile						
	Ethanol						
	Chloroform						
	Acetone						
	Combination of above						
Various mice tissue types	*Methanol*		*3 times of sample*	*Homogenisation*	*HPLC-QQQ-MS*	*100% Methanol*	*(Römisch-Margl et al., 2012)*
	Ethanol						
	10 mM Phosphate buffer (pH 7.5) (4°C)						
	Methanol + Phosphate buffer (4°C)						
	Ethanol + Phosphate buffer (4°C)	*85:15*					
	Ethanol:Dichloromethane (4°C)	*1:1*					
Erythrocytes	*M:C:W (Varying pH)*	*4:2:3*	*1.35 mL*	*Centrifugation*	*LC-ESI-QToF-MS*	*Varying observations*	*(Sana et al., 2008)*
HepG2 cells (Adherent mammalian)	*M:C:W (Mono-phasic)*	*50:20:20*	*1 mL*	*Ultrasonication*	*GC-MS*	*M:C:W (50:20:20)*	*(Bai et al., 2011)*
Panc-1 cells (Adherent mammalian)	*75% M:C*	*9:1*	*900 uL*	*Vortex & Centrifugation*	*UPLC-ESI-QToF-MS (HILIC & RPLC)*	*75% M:C (90:10)*	*(Bi et al., 2013)*
	Acetonitrile						
	Methanol						
	M:C:Acetonitrile						
Macrophages (Adherent cells)	*75%,50% & 25% Methanol*		*1 mL*	*Vortex & Centrifugation*	*GC-ToF-MS*	*50% Acetonitrile*	*(Cheng et al., 2013)*
	75%,50% & 25% Acetonitrile						
Clonal-β-cells (Adherent mammalian)	*82% Methanol*		*1 mL*	*Ball mill*	*GC-ToF-MS*	*82% Methanol*	*(Danielsson et al., 2010)*
Adherent mammalian cell lines	*M:C:W*	*1:1:0.1*	*800 uL*	*Freeze-Thaw*	*GC-ToF-MS*	*M:C:W (1:1:0.1) or 80% Aqueous methanol*	*(Dettmer et al., 2011)*
	M:iP:W						
	80% Aqueous methanol						
	Acetone						
	80% Aqueous acetone						
	Acid/Base methanol						
	Methanol						
MCF-7 cells (Adherent mammalian)	*25% Ethanolic-KOH(0.3M)*		*700 uL*	*Freeze-Thaw*	*GC-MS*	*Methanol (70°C)*	*(Hutschenreuther et al., 2012)*
	M:C:W (70°C)			*Sonication*			
	Methanol (70°C)						
Adherent mammalian cell lines	*Methanol*		*11.6 mL*		*NMR*	*M:C:W*	*(Martineau et al., 2011)*
	2% PCA in water	*2%*					
	M:W	*4:1*					
	Acetonitrile:Water	*4:1*					
	M:C:W (Bi-phasic)	*2:2:1.8*					
MDCK cells (Adherent mammalian)	*75% Boiling ethanol (80°C)*		*2 mL*	*Ultrasonication*	*Luminescence, UV measurement & Anion-exchange chromatography*	*M:C (4°C)*	*(Ritter et al., 2008)*
	Boiling water (95°C)						
	1 M Formic acid						
	Cold acetonitrile (4°C)						
	Cold methanol (4°C)						
	Cold ethanol (4°C)						
	M:C (4°C)						
CHO Cells (Suspension mammalian)	*50% Aq. M & Chloroform*		*1.5 mL*	*Vortex & Incubation*	*HPLC*	*Cold 50% Acetonitrile or Acetonitrile*	*(Dietmair et al., 2010)*
	Methanol Freeze (-40°C)		*1 mL*				
	Cold 50 % Methanol (-40°C)						
	Acetonitrile (Ice-cold)						
	Hot 80 % Methanol (70°C)						
	Cold 100% Methanol (-40°C)						
	Hot Ethanol (80°C)						
	Hot Ethanol HEPES (80°C)						
	Cold Ethanol (Ice-cold)						
	Hot Water (95°C)						

	0.05 M KOH						
	0.5 M PCA						
Bacterial (E. coli)	Methanol (-40°C)		500 uL	Freeze-Thaw	NMR	M:C (2:1)	(Bertini et al.)
	M:C (-40°C)	2:1	750 uL				
Gram + & - bacteria, Yeast, Filamentous fungi	75% Boiling Ethanol with 0.25M HEPES pH 7.5		x	Water Bath	GC-MS	Contradictory observations	(Duportet et al., 2012)
	50% Cold Methanol Freeze-thaw (3 Cycles)		2.5 mL	Freeze-Thaw			
	Cold Methanol			Vortex			
	Cold Methanol Sonication			20KhZ Sonication			
Bacterial (Lactobacillus Plantarum)	Cold Methanol	x	1 mL	Thaw	ATP Bioluminescence Assay & Fluorometric analysis	Cold Methanol	(Faijes et al., 2007)
	35% PCA (-20°C)						
	Boiling Ethanol (90°C)			Hot water bath			
	M:C						
	C:W						
Bacterial (E. Coli)	43% PCA (-20°C)		2 mL	Freeze-Thaw	LC-ESI-Iontrap-MS	Hot water 30 mM TEA pH 7.5	(Hiller et al., 2007)
	Buffered Boiling 75% Ethanol 10 uM HEPES pH 7.5	75:25	10 mL	Water Bath			
	Hot water 30mM TEA pH 7.5		4 mL	Thermomixer			
Bacterial (Mycobacterium Bovis)	C:M	x	10 mL	Sonication	NMR, Gravimetry & Electron Microscopy	Sonication with C:M	(Jaki et al., 2006)
				Bead beating			
				LN2 grinding			
				Microwave			
Bacterial (E. coli)	Methanol & Water	Variable	300 uL	Vortex & centrifugation	LCMS/MS	Methanol:Water 80:20	(Kimball and Rabinowitz, 2006)
Bacterial (E. coli)	Hot Ethanol (90°C)			Incubation	TLC	100% Methanol	(Prasad Maharjan and Ferenci, 2003)
	Hot M:W (70°C)	2:1	500 uL				
	PCA	50%		Sonication			
	KOH	0.1N					
	Cold 100% M & re-extract with 50%Methanol (-20°C)						
	M:C (-20°C)						
Bacterial (Klebsiella oxytoca)	Cold Methanol (-50°C)		1 mL	Freeze-Thaw	LC-ESI-Q-ToF-MS	Cold Methanol (-50°C)	(Park et al., 2012)
	Boiling ethanol (90°C)						
	M:C	2:1					
	Hot water (95°C)						
	KOH (0.25 M) (80°C)						
	PCA (0.25 M)						
Bacterial (E. coli)	M:W (-80°C)	80:20	300 uL	Centrifugation	LC-MS/MS (HILIC)	Acidic Acetonitrile:Water	(Rabinowitz and Kimball, 2007)
	Acetonitrile:Water (4°C)	80:20					
	Ethyl acetate (4°C)						
	Acidic & Basic M:W (-80°C)	80:20					
	Acidic & Basic Acetonitrile:M:W (-20°C)	40:40:20					
	Acidic & Basic Acetonitrile:Water (4°C)	80:20					
Bacterial (Saccharophagus degradans)	Cold Methanol (0°C)		500 uL	Mixer mill	GC-ToF-MS	AMW & WiPM	(Shin et al., 2010)
	50% Acetonitrile:Water (0°C)	1:1					
	AMW (0°C)	2:2:1					
	WiPM (0°C)	2:2:5					
Bacterial (E. coli)	Hot methanol (80°C)	x	200 uL	Freeze-Thaw	LC-ESI-Q-ToF-MS	Ethanol:Water	(Yanes et al., 2011)
	Hot 60% Aq. Ethanol/5 mM Ammonium acetate (80°C)		400 uL				
	Cold 60% Aq. Ethanol/5 mM Ammonium acetate (4°C)		300 uL				
	Boiling water in 1 mM EDTA & 1 mM HEPES		300 uL				
	Acetone/Methanol (-20°C)		400 uL				

Another classical example of a harsh extraction method involves, the use of extraction at elevated temperatures with buffered boiling solvents, which was reported as fast, simple, accurate and reliable method with superior recoveries of intracellular metabolites from yeast (Gonzalez et al., 1997) and *E. coli* (Hiller et al., 2007b). However recent studies with mammalian samples demonstrate significant loss of heat labile metabolites and a build-up of degradation products with this method (Araujo et al., 2013). Use of other solvents such as ethyl acetate or acidic/basic methanol-water combinations resulted in extremely poor recoveries for nearly all the metabolite classes compared to cold organic solvents, as demonstrated previously with a bacterial model (Rabinowitz and Kimball, 2007). Similar conclusions were drawn with the use pure acetone for mammalian samples, which appears to be completely inadequate for global metabolome studies by GC-MS (Dettmer et al., 2011).

Monophasic solvent systems seem to be a more preferred choice for extraction in microbial and yeast samples, as discussed in Chapter 2 - section 2.2.2. With the use of monophasic solvent systems, polar metabolites can be extracted with either methanol, ethanol, isopropanol, acetonitrile, water, methanol/water or suitable combinations of these solvents, whereas non-polar metabolites can be extracted with chloroform or ethyl acetate.

Biphasic solvent systems such as methanol/chloroform/water mixtures offers several unique advantages over monophasic solvent systems. With such systems, one phase, for example the aqueous methanol-water polar phase can be used to extract polar metabolites, whilst the other phase (e.g., chloroform) can be used to extract non-polar metabolites. Both the phases can be extracted simultaneously and each fraction can be analysed separately with better resolution, following centrifugation (Madla et al., 2012). Analysis of both polar and non-polar metabolites from the same sample will be advantageous for metabolomics investigations, as it will avoid much of the variations caused by the analysis of both polar and non-polar metabolites from separate samples (Tambellini et al., 2013). Moreover the use of chloroform in biphasic solvent systems, ensures denaturation of enzymes, thereby halting the metabolism and preventing further degradation or interconversion of metabolites, which might result in false underestimation or overestimation of intracellular pools (Madla et al., 2012). In contrast to these advantages, Gullberg and co-workers, reported that although biphasic solvents can be advantageous in yielding purer extract for both hydrophilic and lipophilic compounds, the procedures are time consuming, difficult to automate thereby decreasing the scope for high-throughput analyses, and overall less suitable for metabolomics investigations (Gullberg et al., 2004b). Moreover considerable loss of metabolites might occur, as some of the metabolites might be associated/leftover with the cell debris, which is usually located at the interphase between the polar and non-polar solvents (in this case between water and chloroform phases). However, this problem can be overcome by employing solvent mixtures such as M:C:W at

1:1:0.1 or 6:2:2 solvent ratios, which results in a single phase (Dettmer et al., 2011, Gullberg et al., 2004b). Most of the metabolomics studies with the mammalian samples reported higher extraction efficiencies with the use of methanol/chloroform/water mixtures at various solvent ratios as can be seen from Table 8.1. Exceptions to these findings are studies conducted with CHO cells, where the authors compared 12 different extraction methods and concluded cold extraction using 50% acetonitrile is the best extraction solvent (Dietmair et al., 2010). Similarly extraction with methanol/chloroform and 82% methanol showed higher extraction efficacy in adherent MDCK cells and clonal-β-cells respectively. However, it is important to note that all the above reports with the mammalian samples did not carry out an investigation on biphasic solvent systems while comparing the various extraction solvents. To date, only a single report included biphasic solvent system (MCW mixture) for comparison against monophasic solvent systems and reported higher extraction efficiency with use of pure methanol (Hutschenreuther et al., 2012a).

As discussed above and in Chapter 2, metabolite extraction from bacterial, yeast and mammalian samples has been widely studied, whereas similar reports on microalgal samples are just starting to appear in the literature and are very few in numbers. Bölling and co-workers reported MCW at a ratio of 10:3:1 as the best extraction solvent for *C. reinhardtii* (Bölling and Fiehn, 2005). High amount of water is usually present in interstitial spaces of *C. reinhardtii* cells, hence to account for this water, lower water content was selected in the extraction buffer (Veyel et al., 2014). Recently, (Lee and Fiehn, 2008a) tested five different extraction solvent systems in microalga *C. reinhardtii*. Authors reported minor quantitative differences with all the extraction methods evaluated. However, based on analytical precision, the authors concluded MCW (5:2:2) as the best extraction solvent.

The objective of this investigation is to improve the GC-MS based metabolome coverage of *C. reinhardtii* by comparing and evaluating the efficiency of carefully selected extraction methods from previously published reports. In total, we have evaluated the efficacy of 15 extraction methods amongst both the monophasic and biphasic solvent systems, which were initially evaluated and optimised separately and later compared together. The samples were analysed by GC-MS and all the identified metabolites were classified into 12 different metabolite classes based on their physicochemical properties. The evaluation criterias used includes: the median and the average number of peaks detected within each metabolite class for each extraction solvent, the normalised median peak intensity of structurally identified compound, and the reproducibility of extraction method (n=5).

8.2 Material and Methods

8.2.1 Chemicals and analytical reagents

All chemicals and reagents were obtained from Sigma-Aldrich (Dorset, U.K.), unless stated otherwise.

8.2.2 Microalgal cultivation

The *C. reinhardtii* strain (CC4323) was grown under similar conditions as described in Chapter 7 (section 7.2.2).

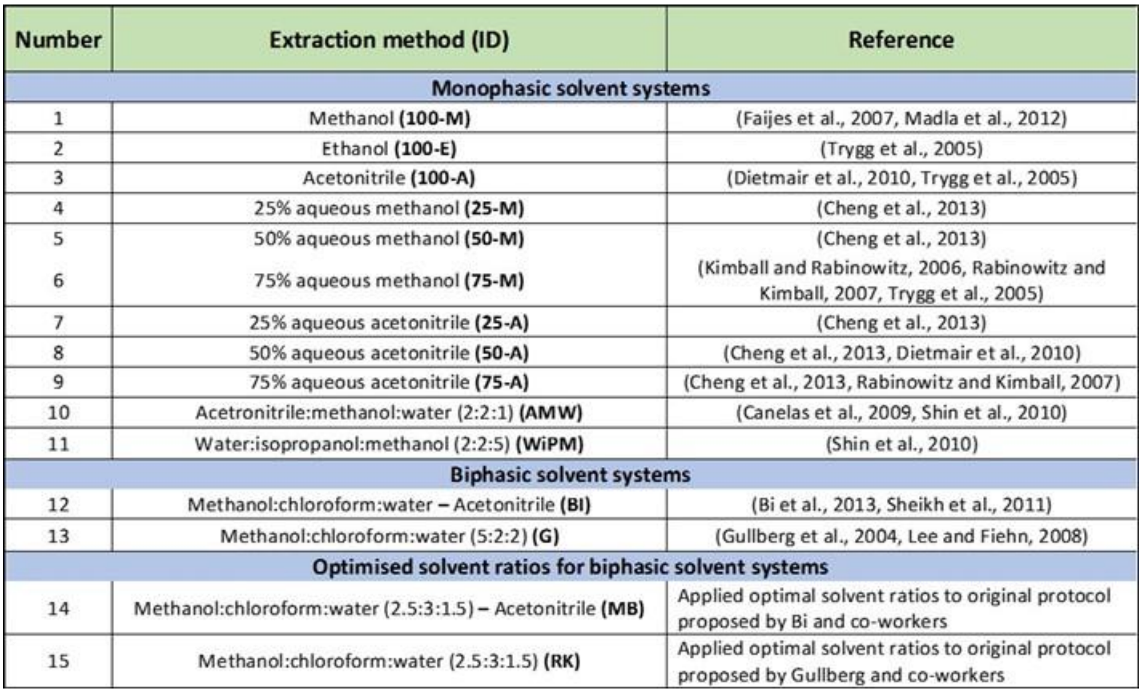
8.2.3 Sampling and quenching

At the incubation site 1 mL of cell suspensions were rapidly plunged into a 2 mL pre-chilled Eppendorf containing 1 mL of pre-chilled 60% aqueous methanol supplemented with 70 mM HEPES (-50°C) as optimised previously in Chapter 7. Addition of the cells to the quenching solution increased the temperature by no more than 15°C. The centrifuge was set at -9°C and the rotor was pre-chilled at -24°C. The quenched biomass was then centrifuged for 2 min at 2500*g* at -9°C. The supernatant was removed rapidly and discarded. The pellets were snap frozen in liquid nitrogen and stored at -80°C for 3 weeks for further analysis.

8.2.4 Metabolite extraction

Various monophasic and biphasic solvent systems were evaluated for the extraction efficiency in order to improve the metabolome coverage. In total 15 extraction protocols were selected from literature, which were optimised and evaluated for microalgal samples as described in table 8.2. After 3 weeks of storage in -80°C, the snap frozen pellets were lyophilized at -50°C overnight prior to extraction. Briefly, to lyophilized cells, 500 µL of extraction solvent (pre-chilled at -48°C) was added, along with an equal volume of glass beads (425-600 µm i.d., acid washed, from Sigma). Cells were then disrupted using (Cell disruptor, from Genie), 5 cycles of disruption was performed with 2 min disruption, with an interval of 2 min on ice. The homogenisation step was carried out in cold room (4°C). The sample was then centrifuged at 13,000 rpm, at -9°C for 15 min to remove any cell debris. The supernatant was transferred to new pre-chilled Eppendorf tube (-20°C) and the remaining cell debris was subjected to re-extraction (5 Cycles) with 500 µL of extraction solvent & centrifuged at 13,000 rpm, at -9°C for 15 min. The supernatant was then evaporated to dryness using vacuum concentrator (5301 vacuum concentrator, from Eppendorf). The dried extract was then stored at -80°C for further analysis.

Table 8.2 Summary and IDs used for various extraction solvent systems evaluated for their extraction efficiencies in this study which includes: Monophasic and biphasic solvent systems used in the past literature, and in addition the biphasic solvent systems in which the solvent ratios were optimised prior to their use



Number	Extraction method (ID)	Reference
Monophasic solvent systems		
1	Methanol **(100-M)**	(Faijes et al., 2007, Madla et al., 2012)
2	Ethanol **(100-E)**	(Trygg et al., 2005)
3	Acetonitrile **(100-A)**	(Dietmair et al., 2010, Trygg et al., 2005)
4	25% aqueous methanol **(25-M)**	(Cheng et al., 2013)
5	50% aqueous methanol **(50-M)**	(Cheng et al., 2013)
6	75% aqueous methanol **(75-M)**	(Kimball and Rabinowitz, 2006, Rabinowitz and Kimball, 2007, Trygg et al., 2005)
7	25% aqueous acetonitrile **(25-A)**	(Cheng et al., 2013)
8	50% aqueous acetonitrile **(50-A)**	(Cheng et al., 2013, Dietmair et al., 2010)
9	75% aqueous acetonitrile **(75-A)**	(Cheng et al., 2013, Rabinowitz and Kimball, 2007)
10	Acetronitrile:methanol:water (2:2:1) **(AMW)**	(Canelas et al., 2009, Shin et al., 2010)
11	Water:isopropanol:methanol (2:2:5) **(WiPM)**	(Shin et al., 2010)
Biphasic solvent systems		
12	Methanol:chloroform:water – Acetonitrile **(BI)**	(Bi et al., 2013, Sheikh et al., 2011)
13	Methanol:chloroform:water (5:2:2) **(G)**	(Gullberg et al., 2004, Lee and Fiehn, 2008)
Optimised solvent ratios for biphasic solvent systems		
14	Methanol:chloroform:water (2.5:3:1.5) – Acetonitrile **(MB)**	Applied optimal solvent ratios to original protocol proposed by Bi and co-workers
15	Methanol:chloroform:water (2.5:3:1.5) **(RK)**	Applied optimal solvent ratios to original protocol proposed by Gullberg and co-workers

8.2.5 Metabolite derivatization, GC-MS analysis, metabolite identification and data analysis

Metabolite derivatization, GC-MS analysis, metabolite identification and data analysis was conducted as described in Chapter 3.

8.3 Results and Discussion

In the previous Chapter, we have optimised and evaluated appropriate sampling and quenching method for *C. reinhardtii*. In this Chapter, the approach was directed towards evaluation of selected extraction methods for *C. reinhardtii* and the main objective is to improve the metabolome coverage. To achieve our objective, we designed the experiments and categorised them into two approaches as illustrated in figure 8.1 Firstly, we evaluated and compared previously published monophasic solvent systems for the extraction efficiency. Secondly, the metabolite recovery efficiencies using biphasic solvent systems were investigated. Finally, in order to propose the best extraction solvent system for *C. reinhardtii*, both the monophasic and biphasic solvent systems with various treatments involved within them were compared, assessed, evaluated and validated with respect to the recoveries of different classes and numbers of metabolites. Further, to compare the relative performance of all the applied extraction methods, data normalisation to the sum of peak intensities of all the identified metabolites for each sample was performed. Due to obvious disadvantages associated with the acidic, alkaline and hot extraction methods discussed in section 8.1, we excluded these procedures in the current study.

The evaluation and comparison within different treatments were based on three response variables, where only features that were present in at least three biological replicates out of five were considered for further analysis:

- Response variable 1: Median/average number of intracellular metabolites recovered in cell extracts
- Response variable 2: Recoveries of metabolites within twelve different classes of metabolites with respect to their normalised median peak intensities.
- Response variable 3: The reproducibility of the extraction methods (n=5)

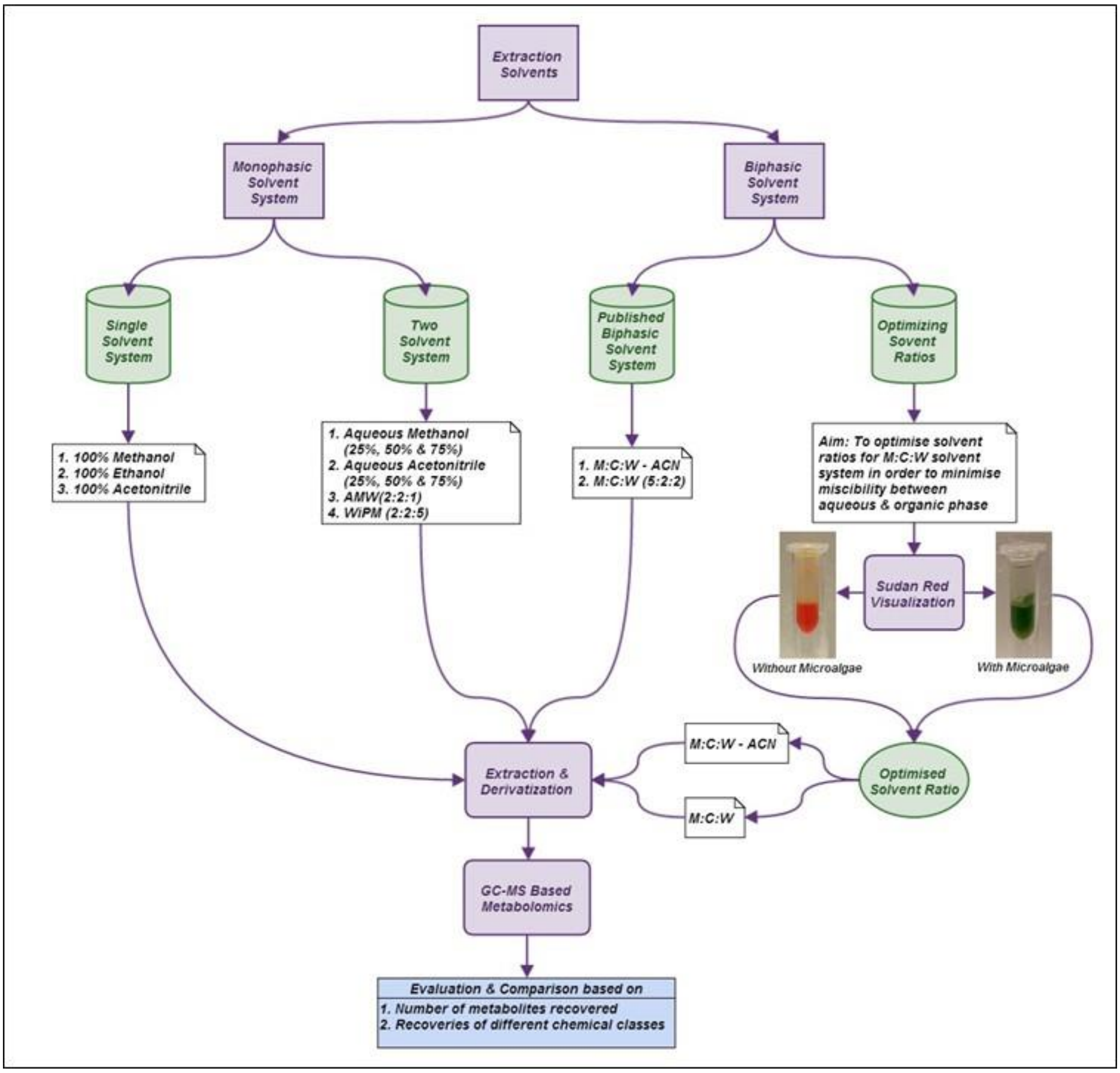

Extraction Solvents
Monophasic Solvent System
Biphasic Solvent System
Single Solvent System
Two Solvent System
Published Biphasic Solvent System
Optimizing Sovent Ratios
1. 100% Methanol
2. 100% Ethanol
3. 100% Acetonitrile
1. Aqueous Methanol (25%, 50% & 75%)
2. Aqueous Acetonitrile (25%, 50% & 75%)
3. AMW(2:2:1)
4. WIPM (2:2:5)
1. M:C:W - ACN
2. M:C:W (5:2:2)
Aim: To optimise solvent ratios for M:C:W solvent system in order to minimise miscibility between aqueous & organic phase
Sudan Red Visualization
Without Microalgae
With Microalgae
Optimised Solvent Ratio
M:C:W - ACN
M:C:W
Extraction & Derivatization
GC-MS Based Metabolomics
Evaluation & Comparison based on
1. Number of metabolites recovered
2. Recoveries of different chemical classes

Figure 8.1 Experimental design and general workflow adopted for the evaluation and optimisation of extraction protocols for *C. reinhardtii*.

8.3.1 Monophasic solvent systems for extraction

For a quantitative metabolomics investigation, it is essential to quantify all the intracellular metabolites. Taking this into consideration, it is essential to select an optimal extraction solvent having a polarity suitable for solubilizing both the polar and non-polar substances in order to improve the metabolome coverage. To achieve our objective and to compare and optimise the extraction efficiency of selected monophasic solvents, cells of *C. reinhardtii* were harvested, quenched and extracted as described in section 8.2 with eleven different extraction solvents. For an overall extraction efficiency comparison, the harvested cell extracts were analysed by GC-MS.

8.3.1.1 Analysis based on recoveries of different classes of metabolites recovered

All the identified metabolites with the eleven different extraction treatments were initially classified into twelve different classes based on their physicochemical properties. The results of this investigation are summarised in Fig. 8.2, where within individual metabolite class, the normalised median intensities of individual metabolites were plotted against various extraction treatments. In total 162 unique metabolites were identified among all the applied extraction treatments which are listed in Appendix 8.1.

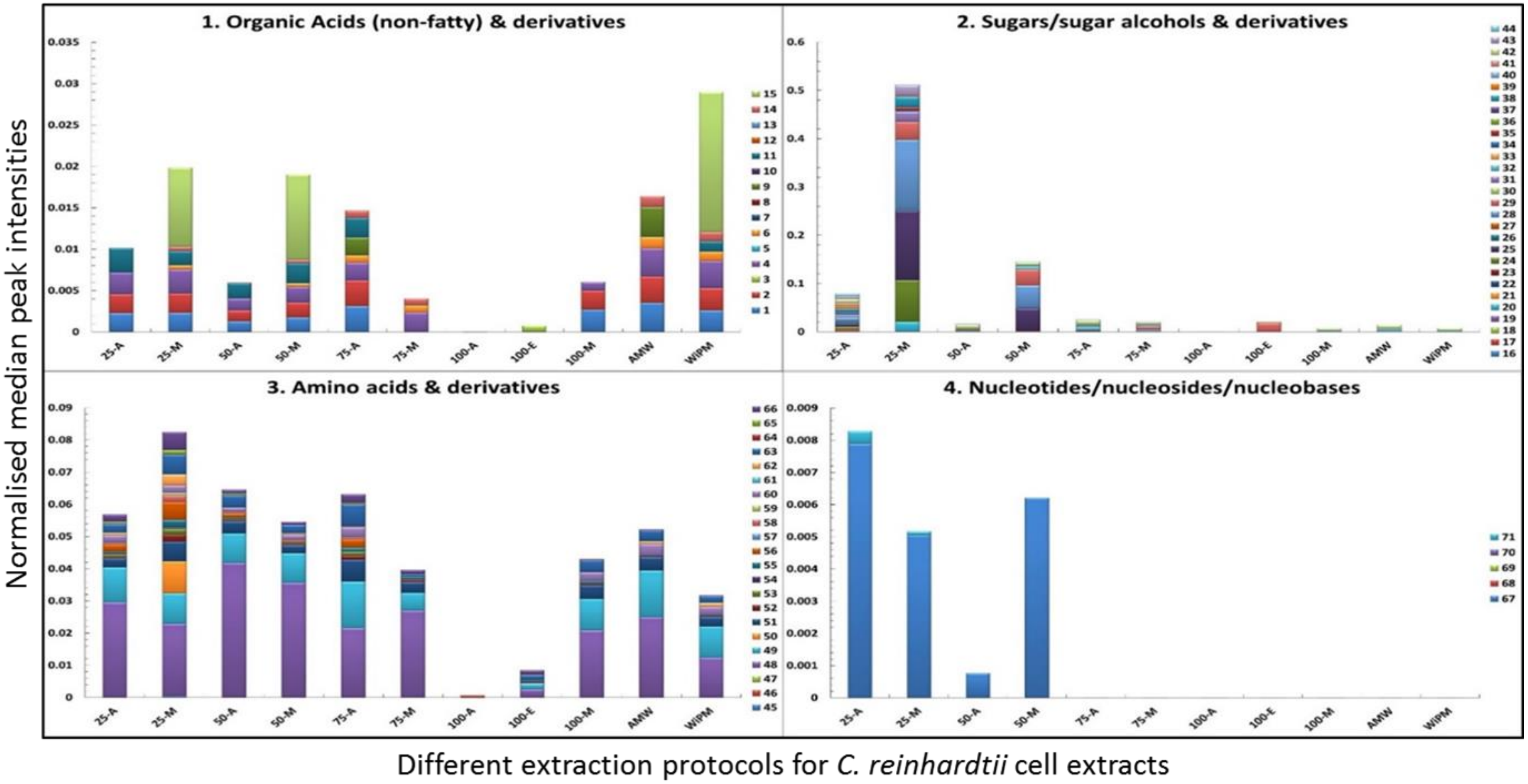


Figure 8.2 Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis using various monophasic extraction solvent systems was compared. X-axis represents different extraction protocols where: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol; 75-A = 75% aqueous acetonitrile; 75-M = 75% aqueous methanol; 100-A = pure acetonitrile; 100-E = pure ethanol; 100-M = pure methanol; AMW = acetonitrile/methanol/water (2:2:1) and WiPM = water/isopropanol/methanol (2:2:5). Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 162 metabolites (See Appendix 8.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

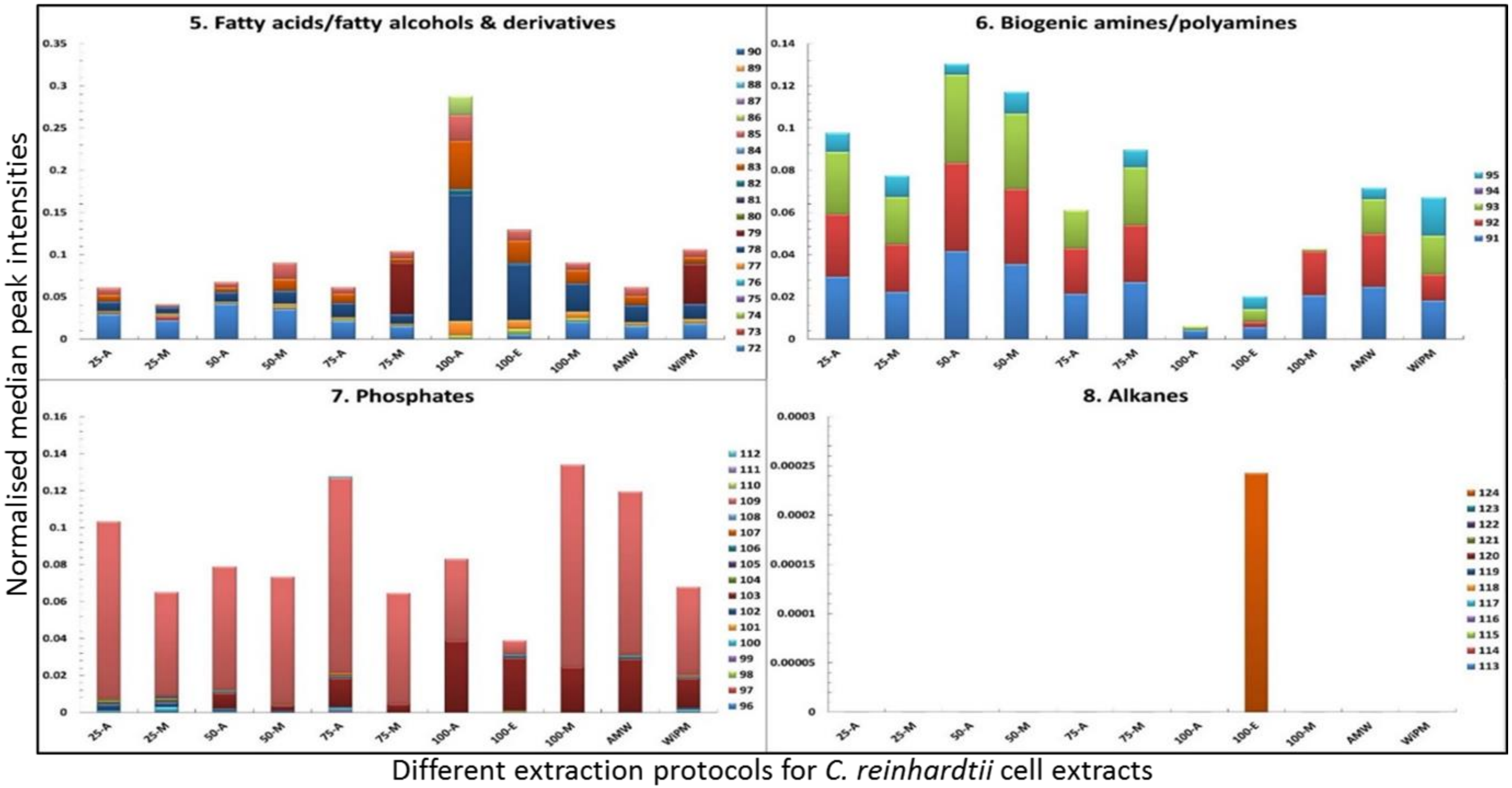


Figure 8.2 (continued) Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis using various monophasic extraction solvent systems was compared. X-axis represents different extraction protocols where: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol; 75-A = 75% aqueous acetonitrile; 75-M = 75% aqueous methanol; 100-A = pure acetonitrile; 100-E = pure ethanol; 100-M = pure methanol; AMW = acetonitrile/methanol/water (2:2:1) and WiPM = water/isopropanol/methanol (2:2:5). Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 162 metabolites (See Appendix 8.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

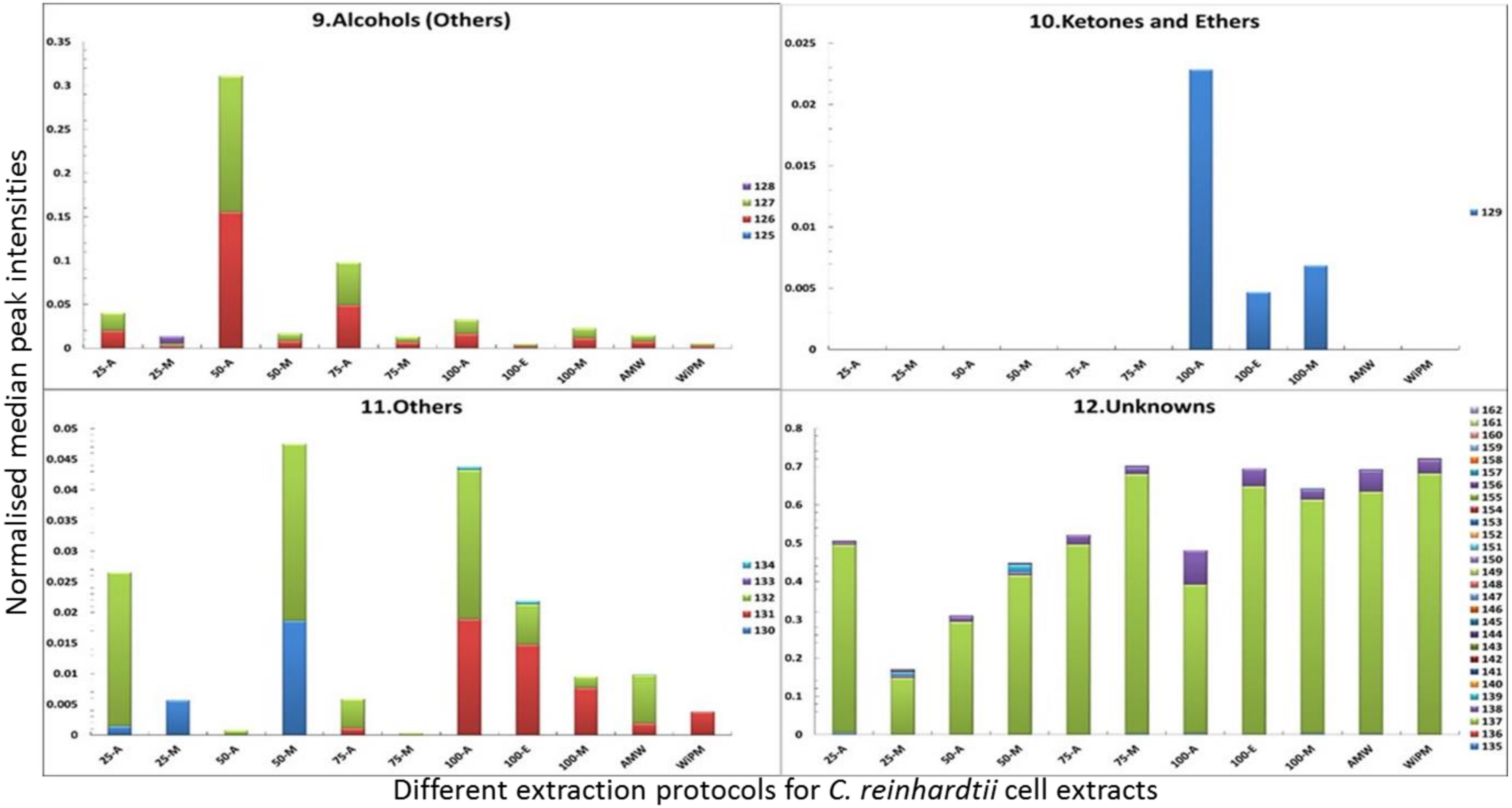


Figure 8.2 (continued) Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis using various monophasic extraction solvent systems was compared. X-axis represents different extraction protocols where: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol; 75-A = 75% aqueous acetonitrile; 75-M = 75% aqueous methanol; 100-A = pure acetonitrile; 100-E = pure ethanol; 100-M = pure methanol; AMW = acetonitrile/methanol/water (2:2:1) and WiPM = water/isopropanol/methanol (2:2:5). Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 162 metabolites (See Appendix 8.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

Recoveries of organic acids (Fig 8.2 (1))

In total, 15 organic acids were identified with all the applied treatments, among which highest numbers of organic acids (7) were recovered with 25-M, 50-M, 75-A and WiPM methods compared to other methods. For all the identified organic acids, WiPM followed by 25-M showed recoveries with good intensities for all metabolites, whereas lower median intensities were observed with 75-A. On the other hand, AMW treatment showed good recovery for six organic acids with good intensities. With 25-A and 50-A, only four organic acids were recovered where higher intensities for them was obtained with 25-A. Only three organic acids were recovered with 75-M, 100-E and 100-M, whereas absolutely no recoveries were observed with 100-A.

Recoveries of sugars/sugar alcohols and derivatives (Fig. 8.2 (2))

In total, 29 sugars/sugar alcohols and derivatives were identified with all the applied treatments, among which higher number of metabolites were recovered with 25-M (15) and 25-A (14) compared to other methods. For all the identified sugars/sugar alcohols, superior intensities were observed with 25-M. Although 25-A recovered similar number of metabolites as 25-M, the intensities recovered with this method was very poor for all the identified metabolites. With rest of the methods, 50-M showed fair recoveries in numbers and intensities for the metabolites. However, very poor recoveries were observed with rest of the applied methods. Absolutely no recoveries were observed with 100-A.

Recoveries of amino acids and derivatives (Fig. 8.2 (3))

In total, 22 amino acids and derivatives were identified with all the applied treatments, among which highest numbers (18) with higher intensities were obtained with 25-M compared to other methods. Good numbers (~14) with good intensities was also recovered with 25-A, 50-A and 75-A methods, whereas lower numbers with lower intensities observed with other methods in which 100-A and 100-E showed poor recovery with lower intensities.

Recoveries of nucleotides/nucleosides/nucleobases (Fig. 8.2 (4))

In case of nucleotides/nucleosides/nucleobases class of metabolites, only adenosine and uridine were identified with selected methods, which includes 25-A, 25-M, 50-A and 50-M. Among which only method 25-A and 25-M showed recovery of both metabolites, whereas with the other two methods only adenosine recovery was observed. Other methods showed absolutely no recoveries for this class of metabolites. The possible reason for lower recoveries in both cell extracts and the supernatant might associated with the subsequent GC-MS analysis employed, as it is known to be poor for analysing nucleotides (Dietmair et al., 2010, Kimball and Rabinowitz, 2006, Rabinowitz and Kimball, 2007). In such cases LC-MS or CE-MS might be more suitable alternatives.

Recoveries of fatty acids/fatty alcohols and derivatives (Fig. 8.2 (5))

In total, 19 fatty acids/fatty alcohols and derivatives were identified with all the applied treatments. All the methods showed good recovery of metabolite numbers. However, with respect to metabolite intensities only 100-A showed good yield. Other methods showed variable selectivities for different metabolites with lower intensities obtained for all of them compared to 100-A. Both the lowest number and intensities were observed with 25-M. With the increase in acetonitrile concentration from 25% (25-A) to 75% (75-A) negligible variations were observed. Increase in recoveries were observed with the corresponding increase in the methanol concentration (from 25% (25-M) to 75% (75-M). In contrast, with use of pure organic solvents recoveries were lowest with methanol (100-M) and highest with acetonitrile (100-A). It seems that these less polar metabolites appear to be better extracted with 75-M, compared to 50-M and 100-M. Our findings are in contrast to that of Duportet and co-workers who reported higher recoveries with 100-M compared to 50-M (Duportet et al., 2012).

Recoveries of biogenic amines/polyamine (Fig. 8.2 (6))

In total, 4 biogenic amines/polyamines were recovered with all the applied treatments which includes, ethanolamine (91), phenethylamine (92), putrescine (93) and tryptamine (95). All the four metabolites were recovered with all the applied methods except with 75-A and 100-M which showed no recoveries for tryptamine and 100-A which failed to recover tryptamine and phenethylamine.

Recoveries of phosphates and alkanes (Fig. 8.2 (7 and 8))

In total, 17 phosphates were recovered with all the applied treatments. Higher numbers were recovered with 25-M and 75-A method, which showed recovery of in total nine phosphates. Both methods recovered different phosphates except for four phosphates (100,102,103 and 112), which were found to be common with both methods. Recovery of metabolite indices 100 and 102 was found to be greater with 25-M, whereas with 75-A exceptionally higher recovery was observed for glycerol-3-phosphate compared to 25-M. Increase in metabolite numbers and corresponding intensities was observed with increase in the acetonitrile concentration (from 25% (25-A) to 75% (75-A)) in the extraction solvent. The exact opposite trend was observed in case of methanol, where decrease in numbers and corresponding intensities were observed with the increase in the methanol concentration (from 25% (25-M) to 75% (75-M). Only recovery for glycerol-3-phosphate was observed with 75-M, 100-A and 100-M methods. No recoveries for alkanes was observed with all the applied treatments, except 100-E method, which managed to recover only tridecane.

Recoveries of alcohols and ketones and ethers (Fig. 8.2 (9 and 10)

In total only four alcohols were recovered with all the applied treatments, where only 25-M showed good recovery for all of them with lower intensities. With rest of the methods only menthol (126) and phytol (127) was recovered, in which higher intensities were observed with 50-A and negligible variation in recoveries were observed with other methods. Amongst ketones and ethers, only hexadecanal was recovered with 100-A, 100-M and 100-E methods, whereas with rest of the methods no recoveries was observed.

Recoveries of others and unknown classes (Fig. 8.2 (11 and 12))

Only few other classes of metabolites were recovered among all the treatments with negligible variations, except with 50-A and 75-M treatments, where low recoveries were obtained. On the other hand large numbers of unknown metabolites were recovered with all the applied treatments. Unknown class (Appendix 8.1) stands for metabolites or mass spectral tags (MSTs) which were not fully identified in the database due to absence of authenticated pure reference standards. However, their other properties resemble those of metabolites which can be identified with proper documentation of MSTs in the database. No meaningful conclusions can be drawn for this class of metabolites as variable recoveries were observed with all the applied treatments.

8.3.1.2 Analysis based on number of metabolites recovered

Results of analysis based on mean recoveries of total number of metabolites extracted and identified among different monophasic extraction solvents are summarised in figure 8.3.

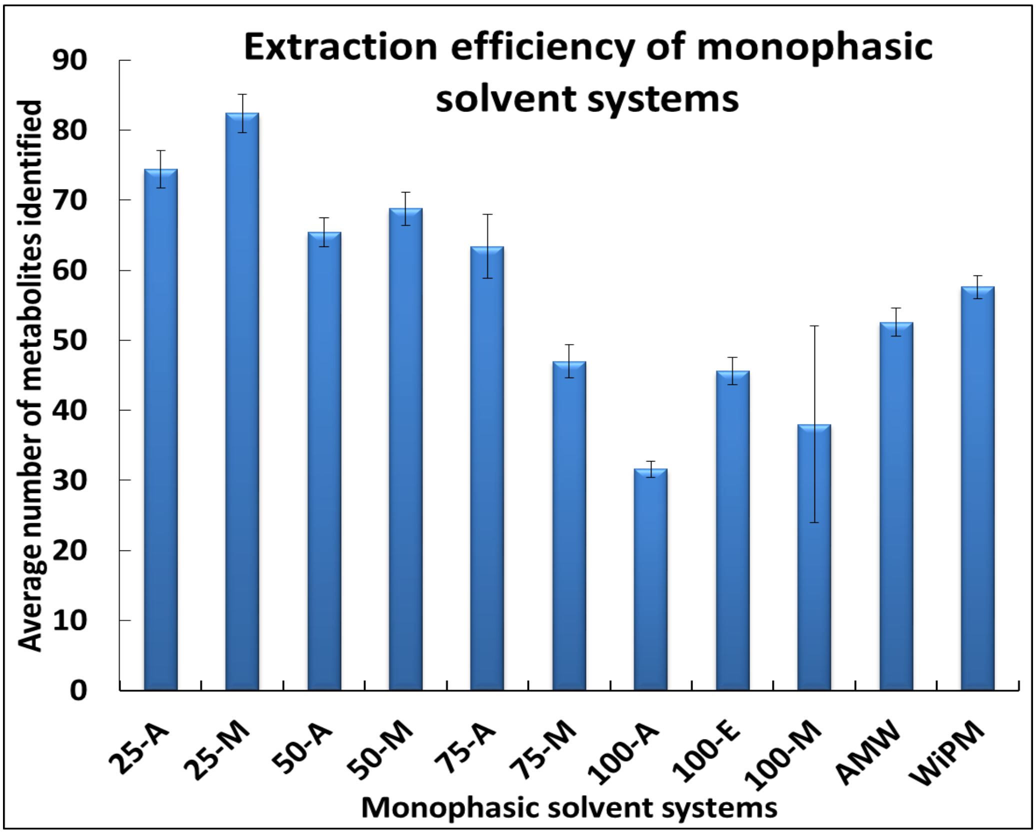


Figure 8.3 Extraction efficiency of various monophasic extraction solvents was compared in recovery of average numbers of metabolites from *C. reinhardtii* cell extracts, based on GC-MS analysis. X-axis represents different extraction protocols where: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol; 75-A = 75% aqueous acetonitrile; 75-M = 75% aqueous methanol; 100-A = pure acetonitrile; 100-E = pure ethanol; 100-M = pure methanol; AMW = acetonitrile/methanol/water (2:2:1) and WiPM = water/isopropanol/methanol (2:2:5). Y-axis represents average number of metabolites recovered.

As can be seen from figure 8.3, the highest number of metabolites were extracted with 25% aqueous methanol (25-M, 82 metabolites), 25% aqueous acetonitrile (25-A, 74 metabolites) and 50% aqueous methanol (50-M, 69 metabolites) compared to all other applied treatments. Similar recoveries in numbers were observed with 50% and 75 % aqueous acetonitrile (50-A and 75-A). In contrast, lowest recoveries in numbers were obtained with 100% methanol (100-M), acetonitrile (100-A) and ethanol (100-E), among which slightly higher recoveries were obtained with 100% ethanol (100-E) compared to pure methanol (100-M) and acetonitrile (100-A). In case of monophasic solvents where mixing of more than two solvents resulting into a single phase solvent system, higher recoveries were obtained with both the solvent systems (AMW and WiPM) compared to that obtained with pure methanol (100-M), acetonitrile (100-A), ethanol (100-E) and 75% aqueous methanol (75-M). Comparison between such solvent systems showed higher recoveries with WiPM solvent system than the AWM solvent system.

Pure methanol and various methanol-water combinations were selected for evaluation, due to reported advantages of methanol as an extraction solvent which usually results in good recovery for most of the metabolite classes with good reproducibilities (Madla et al., 2012). As reported elsewhere (Gullberg et al., 2004b), methanol as an extraction solvent might introduce interference in the derivatization reactions required for GC-MS analyses by causing trans methylation of sugar esters. Owing to this risk behind use of methanol, we have introduced alternatives such as ethanol and acetonitrile. On the other hand, the basis for selection of isopropanol is its superior ability to extract lipophilic constituents as reported elsewhere (Lee and Fiehn, 2008a).

Overall it can be seen that, recoveries in numbers decrease with the corresponding increase in the methanol and/or acetonitrile concentration from 25% to 100%. Best recoveries were obtained with 25-M and then with 25-A method compared to other methods, whereas contradictory findings were reported by Cheng and co-workers, where they reported better recoveries with 50-A from human macrophages compared to 25-M, 50-M, 75-M, 25-A and 75-A (Cheng et al., 2013). The use of pure solvent systems such as methanol, acetonitrile and ethanol alone has shown poor recoveries compared to other methods. Our results were in agreement with that of Yanes and co-workers (Yanes et al., 2011), who reported similar findings and suggested that, pure polar solvents like methanol or acetonitrile or ethanol and pure non-polar solvents like acetone/methanol should not be used for simultaneous extraction of both polar and non-polar metabolites. In contrast, Villas-Boas and co-workers reported higher metabolite recoveries from yeast with 100-M than 50-M (Villas-Bôas et al., 2005). The likely explanation behind these variations might be differences in the cell types. The higher recoveries with acetonitrile-water combinations compared to pure acetonitrile, might be due to inclusion of water, as recoveries with pure acetonitrile was very poor. Despite higher polarity of acetonitrile, good recovery of non-polar metabolites such as fatty acids was observed with 100-A. Comparison within pure organic solvents showed that recovery with ethanol and methanol was almost similar with minor variations and was higher than that obtained with pure acetonitrile. Our findings were in partial agreement with Trygg and co-workers, who demonstrated poor recoveries for all metabolite classes from human blood plasma with acetonitrile (data not shown) compared to that obtained with methanol and ethanol alone (Trygg et al., 2005). Authors finally concluded methanol as the best extraction solvent, however not all methods as investigated in our study were considered for evaluation. In addition, the reproducibility of recovered metabolites with 100% methanol (100-M) was not as great compared to other solvent systems as seen by the larger error bars in figure 8.3. Most of the intracellular metabolites are polar and hydrophilic, which can be likely explanation for poor recoveries obtained with use of pure organic solvents, as solubility of polar compounds (especially nucleotide triphosphates) is very limited. Our findings are in agreement with that of Dietmair and co-workers (Dietmair et al., 2010). Inclusion of treatments involving mixing of two or more solvents at fixed proportions as suggested and optimised previously in case of *Saccharophagus degradans* (Shin et al., 2010), seems to have no significant effect in improving the metabolite recoveries in *C. reinhardtii* compared to other treatments evaluated in our study. Shin and co-workers demonstrated excellent recoveries in numbers with AMW and WiPM compared to that obtained with pure methanol and 50% aqueous acetonitrile (Shin et al., 2010). In our study similar results were seen with higher recoveries in AMW and WiPM compared to use of pure methanol. However contrast findings were observed to that of Shin and co-workers (Shin et al., 2010) where higher recoveries were obtained with 50% aqueous methanol compared (which was not included in other study) to that obtained with AMW and WiPM solvent systems.

Overall better coverage with acceptable compromise between recovery of both polar and non-polar metabolites from *C. reinhardtii* can be achieved by using 25-M. Likely causes behind variations reported in previously published reports might be due to either different cell types under investigation or use of different analytical platform (LC-MS) or might be due to comparison based on few extraction methods, not covering wide range of extraction protocols as considered in our case or might be due to conclusions drawn from studies based on analysis of only small subset of metabolome (such as ATP or only metabolites having high turnover rate) (Canelas et al., 2009, Ritter et al., 2008) rather than broad range of metabolites as considered in our study or combination of above.

8.3.2 Biphasic solvent systems for extraction

Biphasic solvent systems (liquid-liquid extraction) are often used to improve the coverage of both polar and non-polar metabolites and offers obvious advantages of differential solvent solubility and solvent immiscibility. In biphasic solvent systems, the primary requirement is to drive the phase separation into polar and non-polar phases, which is usually achieved by pairing polar aqueous solutions with non-polar organic solvents such as chloroform. The inclusion of aqueous phase co-solvents such as methanol, ethanol and isopropanol which are miscible with water, are often employed in order to enhance the solubility of less-polar metabolites in biphasic solvent systems. To improve the metabolome coverage, the most commonly reported optimal biphasic system for various sample types (yeast, bacteria and mammalian samples) includes methanol/chloroform/water extraction solvent at solvent ratios (5:2:2) (Canelas et al., 2009, Gullberg et al., 2004b, Lee and Fiehn, 2008a, Martineau et al., 2011). Methanol/chloroform/water extraction solvent system was recently suggested as optimal extraction solvent also adherent cancer cell metabolomics with inclusion of acetonitrile at later stage for protein precipitation (Araujo et al., 2013, Sheikh et al., 2011). In case of biphasic solvent systems, it is vital to carefully consider the nature of aqueous and organic solvents and the volumes and solvent ratios employed to drive the phase separation, as all these parameters have significant influence on reproducibility and recovery of both polar and non-polar metabolites and ultimately affects the overall metabolome coverage. As discussed in section 8.1, most of the extraction methods evaluated for *C. reinhardtii* consisted of MCW solvent mixture with various solvent ratios, where with all them contradictory conclusions were reported.

Taking this into account, in our study, we adopted previously reported biphasic solvent systems with their proposed solvent ratios and evaluated them for their extraction efficiencies from *C. reinhardtii*. The solvent systems tested are methanol/chloroform/water (5:2:2) and methanol/chloroform/water solvent system with inclusion of acetonitrile at later stage as originally used by Bi and co-workers (Araujo et al., 2013). Subsequently these solvent systems were optimised with respect to their solvent ratios in order to achieve the optimal phase separation, results of which are then compared within and with the results obtained with their respective original solvent ratios.

8.3.2.1 Optimisation of solvent ratio for optimal phase separation

In this preliminary investigation, the approach was directed towards achieving optimal phase separation by finding the appropriate ratio of aqueous to organic solvent for methanol/chloroform/water solvent system. Initially, equal volumes (0.8 mL) of 80% aqueous methanol (80% selected as at lower methanol concentration there will be a risk of freezing at lower temperatures) was placed in ten 2 mL Eppendorfs. Increasing volumes of chloroform from 0.1 mL to 1 mL were added to 0.8 mL aliquotes of 80% aqueous methanol to simulate the extraction step as displayed in table 8.3. Subsequent addition of 200 µL of water was carried out to each Eppendorf in order to help drive phase separation. As suggested elesewhere (Sana et al., 2008), a lipid soluble diazo dye (Sudan I) was added to each tube to aid visualisation of phase separation. Sudan I had excellent solubility in non-polar organic phase such as chloroform and upon its solubility it imparts a bright yellow or orange colour to chloroform which helps in identifying the miscibility between the polar and non-polar phases. The results of this investigation are summarised in figure 8.4 (a). As can be seen from figure 8.4 (a) at 0.1 and 0.2 ml addition of chloroform, no phase separation occurred, whereas upon 0.7 mL addition of chloroform a light orange shade was imparted to the upper aqueous phase, due to partial miscibility between chlorofom and methanol. The addition of 0.8 mL addition of chloroform resulted in optimal phase separation between the upper polar aqueous phase (methanol/water) and the lower non-polar organic phase (chloroform), as highlighted by green dotted lines. Thus the final concentration of methanol/chlorofom/water at solvent ratios (2.5:3:1.5) resulted in optimal phase separation. To further assess whether the calculated solvent ratios were optimal or not, increasing volumes of chloroform from 0.7 mL to 0.95 ml were added to each of six 2 mL Eppendorfs containing 0.8 mL of 80% methanol and 0.2 mL of water as decribed in table 8.4. However, this time the experiments were conducted on actual cell pellets of *C. reinhardtii*, in order to evaluate its effectiveness and practicality. The results of this investigation are summarised in figure 8.4 (b), where the highlighted green dotted lines again confirms that addition of 0.8 mL of chloroform to the Eppendorf resulted in optimal phase separation. This supports our previous conclusion derived from figure 8.4 (a). Therefore, for optimal phase separation methanol/chloroform/water at solvent ratios (2.5:3:1.5) was selected as the optimised solvent system for further experiments. In addition, similar solvent ratios were applied to that of methanol/chloroform/water solvent system originlly used by Bi and co-workers with inclusion of acetonitrile at later stage for protein precipitation (Araujo et al., 2013). The results of both these solvent systems with optimised solvent ratios were compared with respect to their extraction efficiency with that of the originally used methanol/chloroform/water (5:2:2) solvent system (Gullberg et al., 2004b, Lee and Fiehn, 2008a) and that used by Bi et al (Araujo et al., 2013) (methanol/chloroform/water solvent system with inclusion of acetonitrile at later stage). Therefore within biphasic solvent systems, in total, four extracton protocols were evaluated with respect to their extraction efficiencies in *C. reinhardtii*.

The extraction efficiency of low molecular weight compounds can be greatly enhanced by deproteination tretements by addition of oraganic solvents or salts or by changing the pH using strong acids (Trygg et al., 2005). As discussed in section 8.1, addition of salts introduces intereference with the subsequent analytical platform whereas acidic conditions favour extraction of only specific classes of compounds. Therefore protein precipation using acetonitrile was selected in this study, as it is also being employed previously in Chapter 4. As can be seen from figure 8.5 (a), samples extracted with BI method with inclusion of acetonitrile for protein precipitation step, the precipitated proteins can be visually observed between the polar and non-polar interphase, whereas with the RK method (figure 8.5 (b), optimised solvent ratios of MCW (2.5:3:1.5) mixtures resulting in equal distribution of upper polar aqueous phase (methanol-water) and lower non-polar phase (chloroform).

Table 8.3 Increasing the volumes of chloroform by 100 µL to 800 µL aliquots of 80% aqueous methanol to simulate the extraction step. Subsequent addition of 200 µL of water was done in order to help drive phase separation

Sr No.	*80 % Methanol*	*Chloroform*	*Add Sudan I Granules*
			Add water & Centrifuge
1	*800 µL*	*100 µL*	*200 µL*
2	*800 µL*	*200 µL*	*200 µL*
3	*800 µL*	*300 µL*	*200 µL*
4	*800 µL*	*400 µL*	*200 µL*
5	*800 µL*	*500 µL*	*200 µL*
6	*800 µL*	*600 µL*	*200 µL*
7	*800 µL*	*700 µL*	*200 µL*
8	*800 µL*	*800 µL*	*200 µL*
9	*800 µL*	*900 µL*	*200 µL*
10	*800 µL*	*1000 µL*	*200 µL*

Table 8.4 Increasing the volumes of chloroform by 50 µL to 800 µL aliquots of 80% aqueous methanol to simulate the extraction step. Subsequent addition of 200 µL of water was done in order to help drive phase separation

Sr No.	*80 % Methanol*	*Chloroform*	*Add Sudan I Granules*
			Add water & Centrifuge
1	*800 µL*	*700 µL*	*200 µL*
2	*800 µL*	*750 µL*	*200 µL*
3	*800 µL*	*800 µL*	*200 µL*
4	*800 µL*	*850 µL*	*200 µL*
5	*800 µL*	*900 µL*	*200 µL*
6	*800 µL*	*950 µL*	*200 µL*

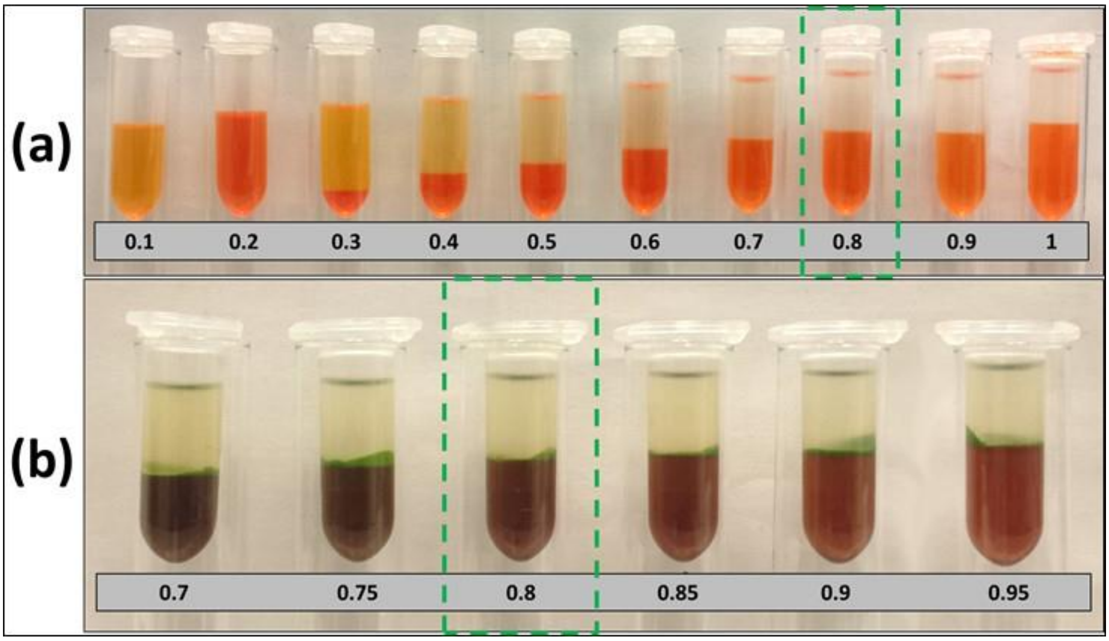

(a)
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
(b)
0.7
0.75
0.8
0.85
0.9
0.95

Figure 8.4 Addition of Sudan I granules (A lipid soluble diazo dye used to aid visualisation of phase separation) to Eppendorf tube containing volumes of aqueous and organic components as described in table 8.3, in order to visulise the chloroform phase and estimate the extent of mixing with the methanol/water. Where (a) = without *C. reinhardtii* culture and (b) = with *C. reinhardtii* culture. Colour banding within the aqueous phase (upper) in fig (a) is due to rediffusion of chloroform into the aqueous phase.

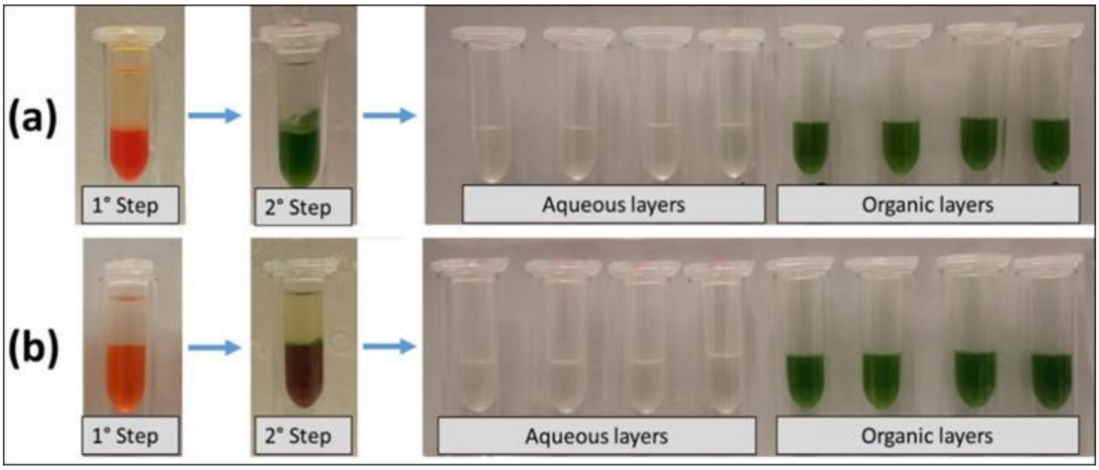

(a)
1° Step
2° Step
Aqueous layers
Organic layers
(b)
1° Step
2° Step
Aqueous layers
Organic layers

Figure 8.5 Snapshot of representative samples processed with BI and RK method. (a) = M:C:W – Acetonitrile solvent system (BI) as used by Bi and co-workers (Bi et al., 2013); whereas (b) = M:C:W (2.5:3:1.5) solvent system (RK) with optimised solvent ratios leading to optimal phase separation.

8.3.2.2 Analysis based on recoveries of different classes of metabolites recovered

With respect to analysis of metabolite classes recovered and median intensities of individual metabolites identified within each class, all the identified metabolites with the four different extraction treatments as listed in table 8.2 (biphasic solvent systems and optimised solvent ratios for biphasic solvent systems) were initially classified into twelve different classes based on their physicochemical properties. The results of the investigation are summarised in Fig. 8.6, where within individual metabolite class, median intensities of individual metabolites were plotted against various extraction treatments. In total 162 unique metabolites were identified among all the applied extraction treatments which are listed in Appendix 8.1.

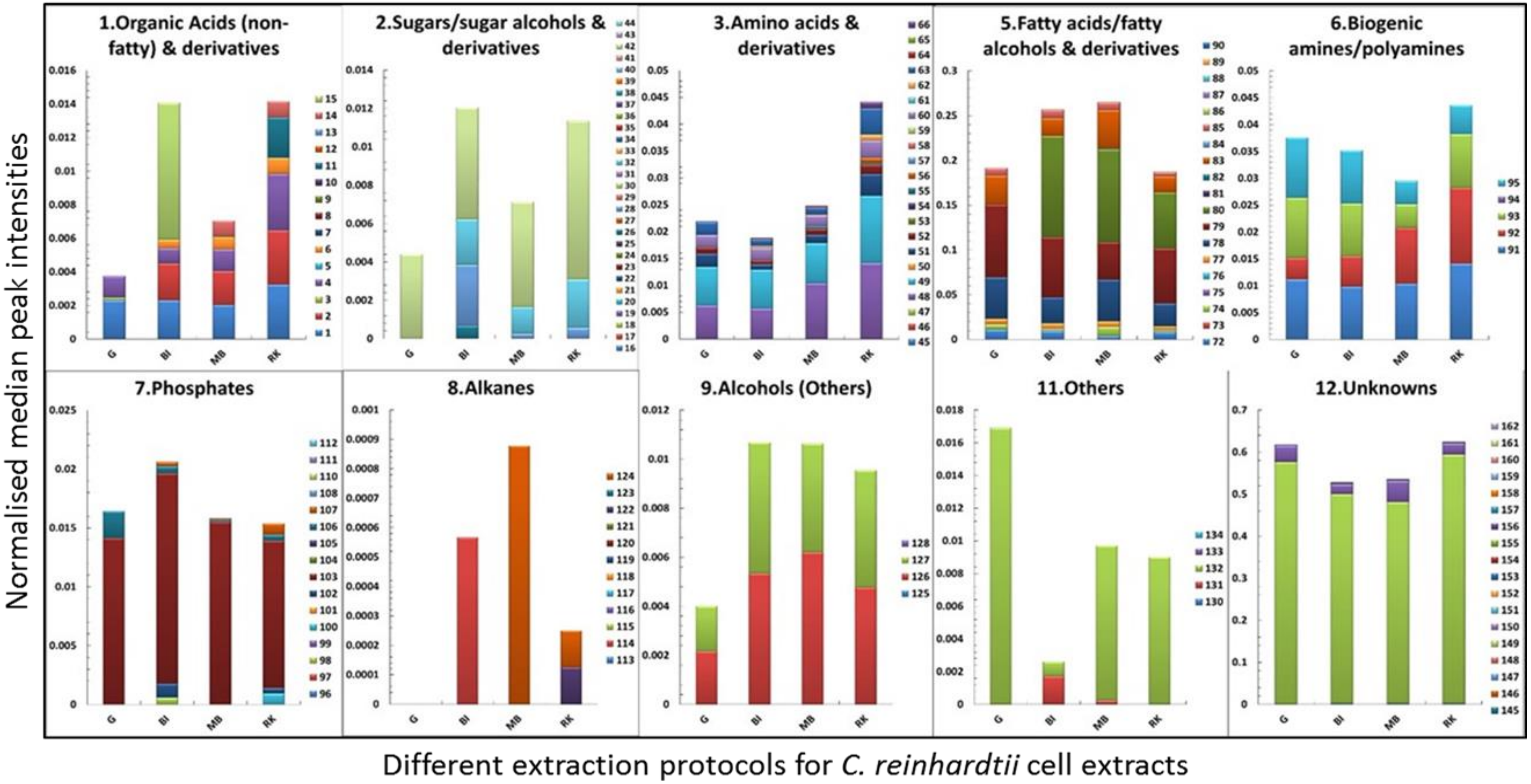


Figure 8.6 Median intensities of metabolites recovered in individual metabolite class from *C. reinhardtii* cell extracts based on GC-MS analysis using various biphasic extraction solvent systems was compared. X-axis represents different extraction protocols where: G = methanol/chloroform/water (5:2:2); BI = methanol/chloroform/water – With inclusion of acetonitrile for protein precipitation; MB = Modified BI protocol, methanol/chloroform/water (2.5:3:1.5) – With inclusion of acetonitrile for protein precipitation; RK = Optimised solvent ratios for MCW mixture, methanol/chloroform/water (2.5:3:1.5). Y-axis represents normalised median intensities of individual metabolite recovered within each metabolite class. In total 162 metabolites (See Appendix 8.1) were identified across all the treatments which were classified into 12 different metabolite classes and plotted separately.

Recoveries of organic acids (Fig 8.6 (1))

In total, 15 organic acids were identified with all the applied treatments, amongst which highest numbers (6) with higher intensities were obtained with RK method compared to other methods. On the other hand, in the BI method recoveries of five organic acids were obtained, with lower intensities for those metabolites which were also identified with the RK method. Whilst exceptionally higher recoveries were obtained for valeric acid (15) with Bi method, this metabolite was not detected with the RK method. Similar numbers to that of BI method were recovered with MB method with negligible variations in recoveries for metabolites found to be common with both methods. Least numbers (3) with lower recoveries were observed with the G method.

Recoveries of sugars/sugar alcohols and derivatives (Fig. 8.6 (2))

In total, 29 sugars/sugar alcohols and derivatives were identified with all the applied treatments, amongst which only four sugars/sugar alcohols with higher intensities were recovered with BI method compared to other treatments, whereas only three recovered with that of MB (with lower intensities) and RK method (with higher intensities). Only one sugar was recovered with G method with lower intensities compared to other treatments. Only trehalose (42) was recovered with all the applied treatments with good intensities.

Recoveries of amino acids and derivatives (Fig. 8.6 (3))

In total, 22 amino acids and derivatives were identified with all the applied treatments, among which highest numbers (10) with higher intensities were obtained with RK method. Good numbers (~8) were also recovered with rest methods, however recovered intensities for those metabolites was found to be lower than that obtained with the RK method. Slightly higher intensities were recovered with the MB method for those metabolites found to be common with the G and BI methods.

Recoveries of nucleotides/nucleosides/nucleobases (Fig. 8.6 (4))

Among all the applied treatments absolutely no recoveries were observed for this class of metabolites, might due to possible reasons as discussed in section 8.3.1.1.

Recoveries of fatty acids/fatty alcohols and derivatives (Fig. 8.6 (5))

In total 19 fatty acids/fatty alcohols and derivatives were identified with all the applied treatments. All the methods managed recover good numbers (7–9), however, with respect to metabolite intensities only MB showed good yield, whereas other methods showed variable selectivities for different metabolites with variable intensities obtained for all of them. Both numbers recovered and intensities were the lowest with the G method, except for caproic acid (72), hexadecanoic acid (78) and hexadecanol (79), which showed higher intensities compared to the BI and RK methods.

Recoveries of biogenic amines/polyamine (Fig. 8.6 (6))

In total, only 4 biogenic amines/polyamines were recovered with all the applied treatments which includes, ethanolamine (91), phenethylamine (92), putrescine (93) and tryptamine (95). All the four metabolites were recovered with all the applied methods but higher intensities for all of them were observed with only RK method.

Recoveries of phosphates and alkanes (Fig. 8.6 (7 and 8))

In total, 17 phosphates were recovered with all the applied treatments. Higher numbers were recovered with the RK method, which showed recoveries of in total six phosphates. Only inositol-2-phosphate (106) and glycerol-2-phosphate (103) were recovered with all the applied treatments where higher intensities were observed with the G and BI method respectively. Mannose-6-phosphate (107) was recovered with all the methods except with G method whilst higher intensities were recovered with the RK method. Only the RK and BI methods showed recoveries of glyceric acid-2-phosphate (102) with higher intensities observed with the BI method. Selective recovery for gluconic acid-6-phosphate and ethanolamine phosphate was observed with only RK and BI method respectively. Overall, the RK and BI methods appear to be suitable for recoveries of this class of metabolites, whereas G and MB method showed poor recovery with recoveries of only 2 and 3 metabolites respectively.

In case of alkanes, very poor recoveries were observed with all the applied treatments, where the G method showed absolutely no recovery for this class of metabolites. Only RK method showed recoveries of 2 alkanes (tetradecane (122) and tridecane (124)) compared to BI and MB. These latter methods showed selective recovery for only one alkane, dodecane (114) and tridecane (124), respectively.

Recoveries of alcohols and ketones and ethers (Fig. 8.6 (9 and 10)

In total, only two alcohols, menthol (126) and phytol (127), were recovered with all the applied treatments, where negligible variations in recoveries were observed with BI, MB and RK methods. Lower intensities for both the metabolites was observed with G method. Among class ketones and ethers, absolutely no recoveries were observed with all the applied treatments.

Recoveries of others and unknown classes (Fig. 8.6 (11 and 12))

Only lumichrome (132), a metabolite reportedly secreted by *C. reinhardtii* (Veyel et al., 2014), was recovered with all the applied treatments where a particularly high recovery was obtained with the G method and low recovery was obtained with the BI method. On the other hand, ergocalciferol (131) was only recovered with the BI and MB method, where higher intensities was obtained with BI method. Only three unknown class of metabolites were recovered with all the applied treatments with negligible variations in their recoveries.

8.3.2.3 Analysis based on number of metabolites recovered

Results of analysis based on mean recoveries of total number of metabolites extracted and identified among different biphasic extraction solvents are summarised in figure 8.7.

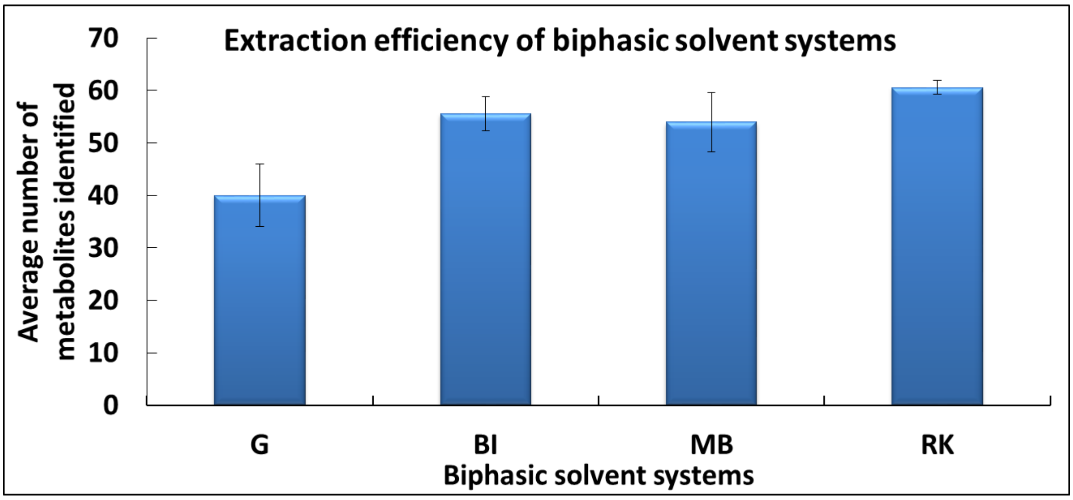


Figure 8.7 Extraction efficiency of various biphasic extraction solvents was compared in recovery of average numbers of metabolites from *C. reinhardtii* cell extracts, based on GC-MS analysis. X-axis represents different extraction protocols where: G = methanol/chloroform/water (5:2:2); BI = methanol/chloroform/water – With inclusion of acetonitrile for protein precipitation; MB = Modified BI protocol, methanol/chloroform/water (2.5:3:1.5) – With inclusion of acetonitrile for protein precipitation; RK = Optimised solvent ratios for MCW mixture, methanol/chloroform/water (2.5:3:1.5). Y-axis represents average number of metabolites recovered.

As can be seen from figure 8.7, highest number of metabolites (64) were extracted with RK method compared to all other applied treatments. Similar recoveries in numbers were observed with BI (57) and MB (56) method with minor variations. In contrast, lowest recoveries in numbers were obtained with G method. Among all the applied treatments superior reproducibility was observed with only RK method compared to all other treatments as represented by small error bars, whereas poor reproducibility's were observed with rest methods as represented by comparatively larger error bars.

From the overall results with respect to all evaluation criterias as discussed in section 8.1, RK method seems to be a preferred choice among all evaluated biphasic extraction solvents. The higher recoveries and intensities for nearly all classes of metabolites might be attributed to careful optimisation and tailoring of the solvent ratios as performed in section 8.3.2.1.

8.3.3 Comparison of extraction efficiencies of monophasic and biphasic solvent systems

In this section, all the applied treatments so far within both the monophasic and biphasic solvent systems are compared, assessed, evaluated and validated with respect to the recoveries of different classes and numbers of metabolites.

8.3.3.1 Analysis based on number of metabolites recovered

The results of this investigation are summarised in figure 8.8, where the average number of metabolites recovered within each metabolite class was plotted against the 15 extraction methods evaluated within both the monophasic and biphasic solvent systems.

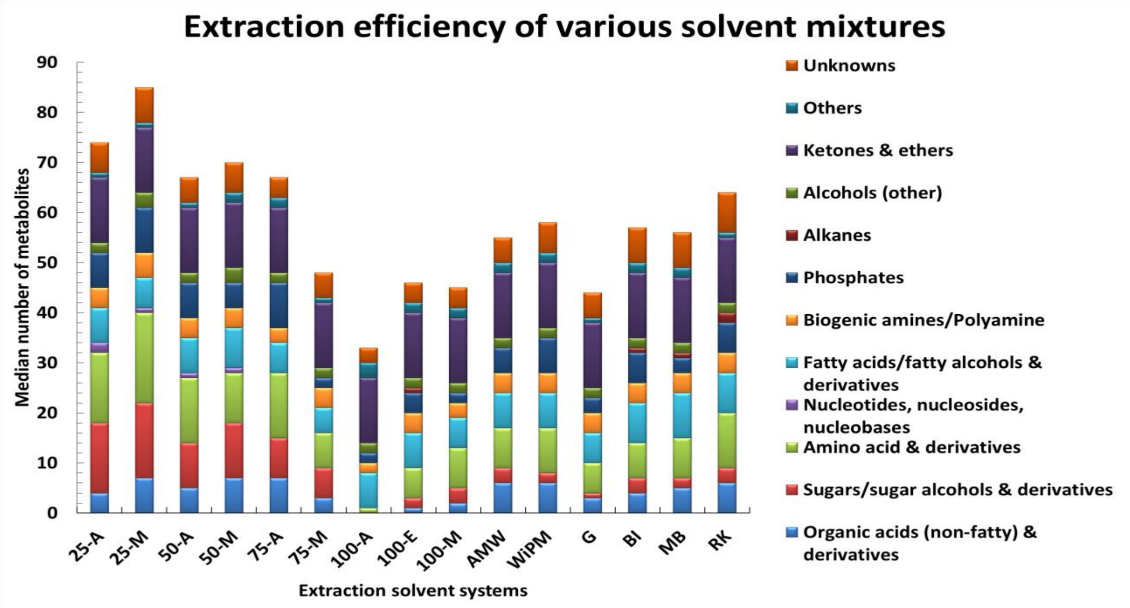
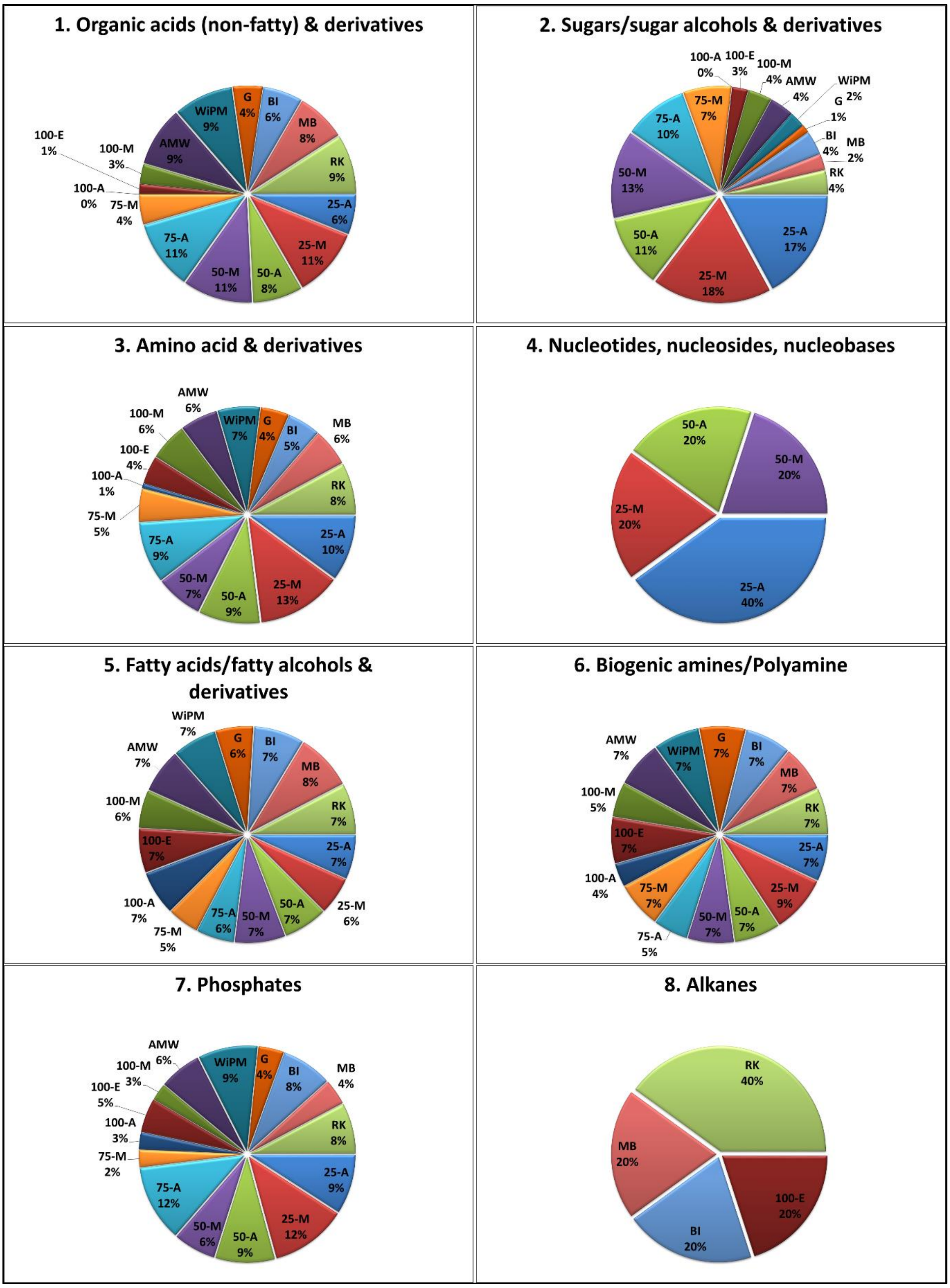


Figure 8.8 Extraction efficiency of various monophasic and biphasic extraction solvents was compared in recovery of median numbers of metabolites within each chemical class from *C. reinhardtii* cell extracts, based on GC-MS analysis. X-axis represents different extraction protocols where: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol; 75-A = 75% aqueous acetonitrile; 75-M = 75% aqueous methanol; 100-A = pure acetonitrile; 100-E = pure ethanol; 100-M = pure methanol; AMW = acetonitrile/methanol/water (2:2:1); WiPM = water/isopropanol/methanol (2:2:5); G = methanol/chloroform/water (5:2:2); BI = methanol/chloroform/water – With inclusion of acetonitrile for protein precipitation; MB = Modified BI protocol, methanol/chloroform/water (2.5:3:1.5) – With inclusion of acetonitrile for protein precipitation; RK = Optimised solvent ratios for MCW mixture, methanol/chloroform/water (2.5:3:1.5). Y-axis represents median number of metabolites.

Overall among monophasic solvent systems, 25-M (82) and 25-A (74) and among biphasic solvent systems, RK method (61) showed recoveries of higher number of metabolites within each class. Amongst both the solvent systems good recoveries were also obtained with the 50-A (65), 50-M (69) and 75-A methods (63).

8.3.3.2 Analysis based on recoveries of different classes of metabolites recovered

All the identified metabolites with the 15 different extraction treatments were initially classified into twelve different classes (Chapter 4 - Table 4.2) based on their physicochemical properties. The results of the investigation are summarised in Fig. 8.9, where values for numbers of metabolite detected within each class with the applied treatments are reported as the number percentage of metabolites identified.



1. Organic acids (non-fatty) & derivatives

2. Sugars/sugar alcohols & derivatives

3. Amino acid & derivatives

4. Nucleotides, nucleosides, nucleobases

5. Fatty acids/fatty alcohols & derivatives

6. Biogenic amines/Polyamine

7. Phosphates

8. Alkanes

Figure 8.9 Pie charts representing extraction efficiency of various monophasic and biphasic extraction solvents in recovery of average percentage numbers of metabolites within each of the twelve metabolites classes from *C. reinhardtii* cell extracts.

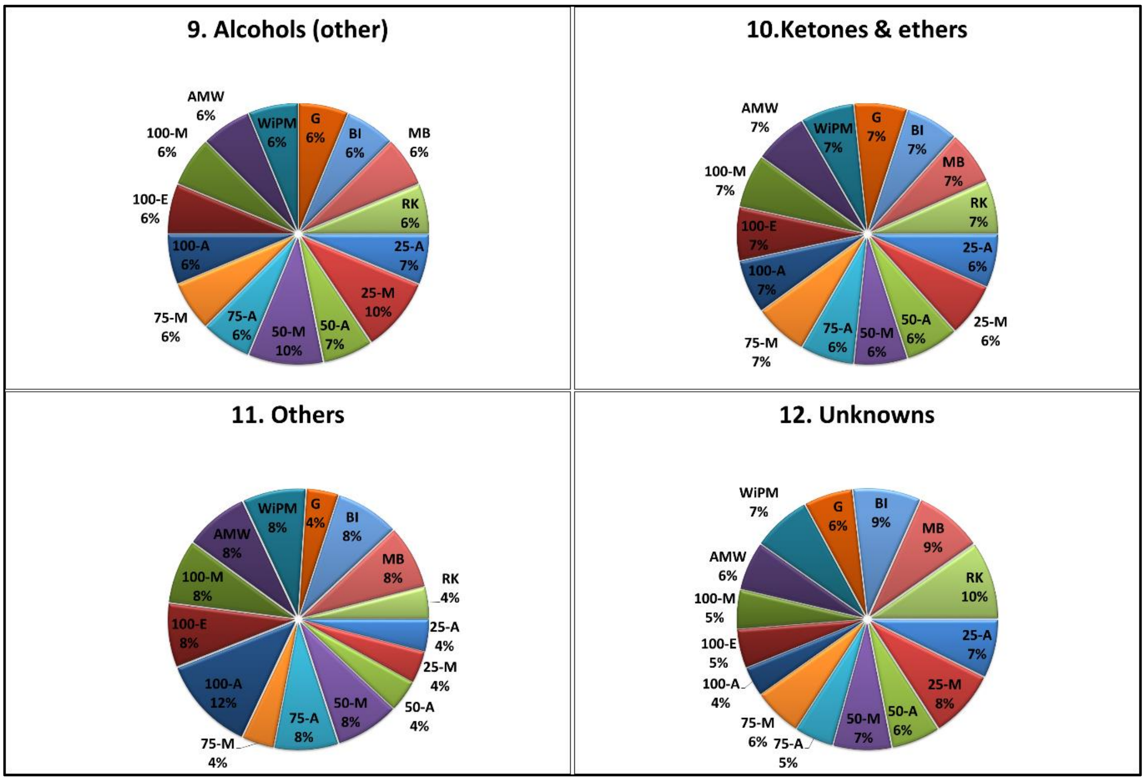


Figure 8.9 (continued) Pie charts representing extraction efficiency of various monophasic and biphasic extraction solvents in recovery of average percentage numbers of metabolites within each of the twelve metabolites classes from *C. reinhardtii* cell extracts.

Higher proportion of organic acids and derivatives were recovered with 25-M, 50-M and 75-A methods (11%) whereas good recoveries (9%) were also obtained with methods such as RK, AMW and WiPM. In case of sugars/sugar alcohols and derivatives, superior recoveries were obtained with 25-M (18%) and 25-A method (17%). A similar trend was observed in recovery of amino acids and derivatives, where 25-M (13%) and 25-A (10%) managed to recover higher proportions. For metabolites belonging to classes such as organic acids, sugars/sugar alcohols and amino acids very poor recoveries were obtained with 100-M, 100-E, 100-A, 75-M and G method. Similar recoveries for amino acids were observed with 100-M and WiPM, which was in agreement with a previously published report (Dettmer et al., 2011). In case of nucleotides/nucleosides/nucleobases only four methods (25-M, 25-A, 50-M and 50-A) managed to recover metabolites from this class, where exceptionally higher recoveries were obtained with 25-A (40%), whereas other methods yielded similar recoveries (20%). Among all the applied treatments, recovery of fatty acids/fatty alcohols and derivatives was found to be similar (~7%) with minor variations, except with MB method which yielded slightly higher recoveries (8%) and 75-M which resulted in lower recoveries (5%) compared to other methods. In case of biogenic amines/polyamines, higher recoveries were obtained with 25-M (9%), whereas lower recoveries (~5%) were obtained with 100-M, 100-A and 75-A. Similar recoveries (7%) were obtained for metabolites of this class with other methods. Higher proportion of phosphates were recovered with 25-M and 75-A methods (12%), whereas good recoveries (~8-9%) were also obtained with methods such as RK, WiPM, BI and 50-A. Only biphasic solvent systems such as RK, MB and BI methods were managed recover few alkanes, whereas only 100-E method among monophasic solvent system managed to recover one alkane. Overall the recovery of this class of metabolites was extremely poor with all the applied treatments. Higher recoveries for alcohols (10%) were obtained with 25-M and 50-M, whereas similar recoveries with minor variations (~6-7%) was observed with all the other methods. Similarly minor variations in recovery of ketones and ethers (6-7%) was observed with all the applied methods with no significant differences. In case of metabolites belonging to other and unknown class, variable recoveries with the applied methods were observed in which 100-A (12%) and RK method (10%) managed to recover higher proportions for the respective class.

8.3.3.3 Further validation of selected optimal extraction solvents with previously employed extraction solvents

Overall, 25% aqueous methanol (25-M) and acetonitrile (25-A) resulted in higher number of metabolites with higher intensities observed for most of the metabolite classes. In order to further validate the raw data obtained from GC-MS analysis, the raw chromatograms obtained with 25-M, 25-A, 100-M and G method were compared with respect to analysis of the number of peaks and their corresponding intensities recovered. The 100-M and G methods were purposely included in these comparative analysis, as 100% methanol (100-M) was suggested as the best extraction solvent for bacterial models (Faijes et al., 2007, Park et al., 2012, Prasad Maharjan and Ferenci, 2003b, Winder et al., 2008), whereas methanol:chloroform:water (5:2:2) (G method) as originally used for *C. reinhardtii* (Gullberg et al., 2004b, Lee and Fiehn, 2008a) and therefore adopted in our preliminary work with microalgal samples (Chapter 6).

Figure 8.10 represents GC-MS profiles of intracellular metabolites extracted with four different extraction solvents from *C. reinhardtii*. Comparison between four different extraction solvents with respect to total number of peaks and their intensities observed, showed major differences between them. As can be seen from figure 8.10, the 25-M method yielded highest number of peaks with higher intensities. Slightly lower numbers were recovered with 25-A method with similar intensities as obtained with 25-M method. On the other hand, compared to 25-M and 25-A method, very less numbers of peaks with extremely lower intensities for them was recovered with the 100-M and G method as represented by green and blue regions highlighted by green and blue dotted lines respectively. In total ,82 ± 3 and 74 ± 3 unique metabolites can be identified with higher relative peak areas for 25-M and 25-A method respectively, whereas only 38 ± 14 and 40 ± 6 unique metabolites were identified with lower relative peak areas with

100-M and G method respectively. This analysis clearly reveals that, the extraction efficiency and reproducibility of results obtained with 25-M and 25-A method was superior compared to that of previously adopted 100-M and G method. Nearly two fold increase in numbers and 2.5 fold increase in intensities recovered was observed with 25-M and 25-A method compared to 100-M and G method.

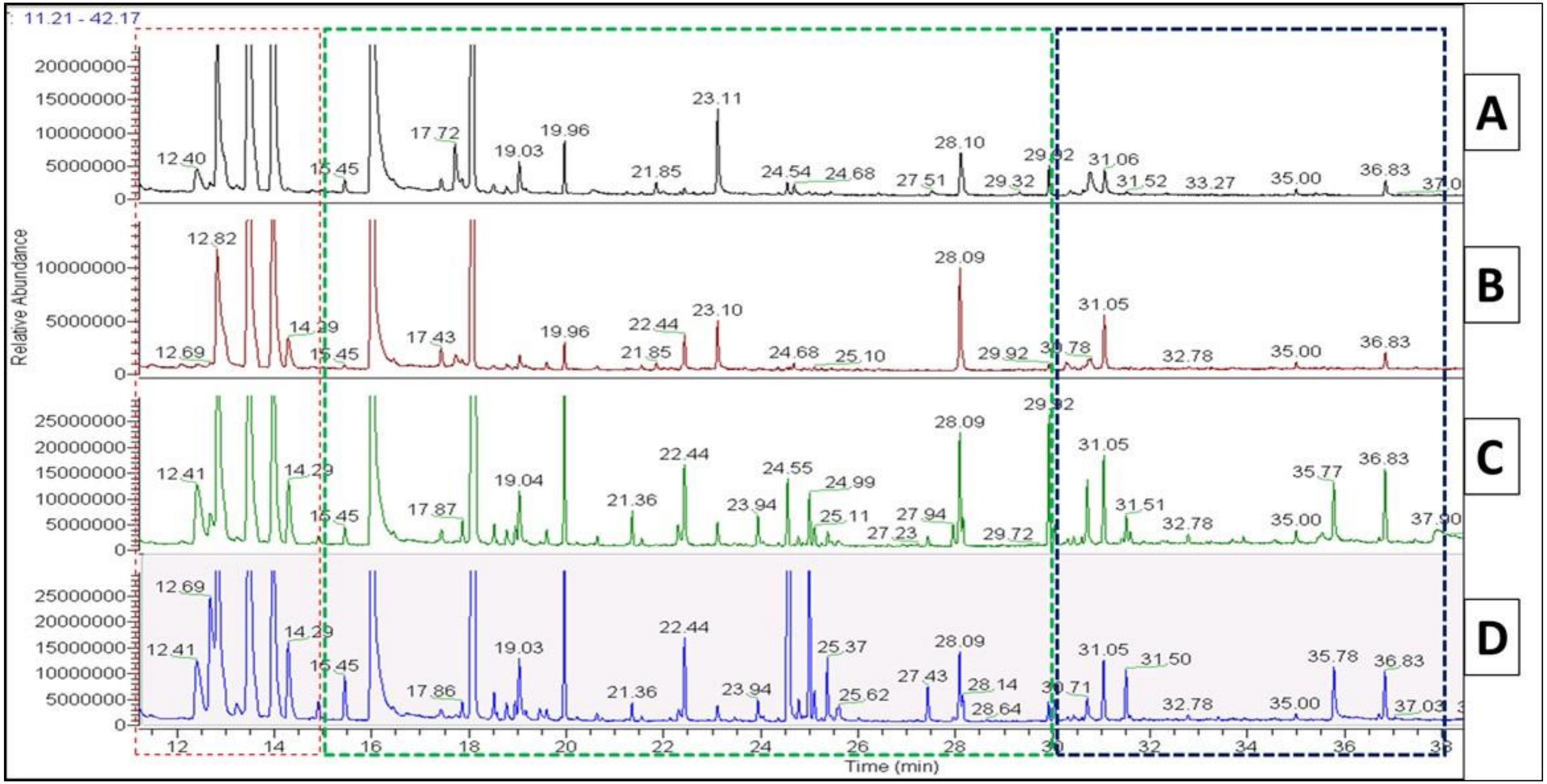


Figure 8.10 GC-MS based analysis of intracellular metabolites extracted from *C. reinhardtii* with four different extraction solvents. Cells were quenched with 60% aqueous methanol supplemented with 70 mM HEPES prior to extraction of metabolites with various extraction solvents. X-axis represents retention time in minutes, whereas Y-axis represents relative abundance. The raw chromatograms obtained from GC-MS analysis with various extraction solvents were labelled where: A = 100% methanol (100-M method); B = methanol/chloroform/water (5:2:2) (G method); C = 25% aqueous acetonitrile (25-A) and D = 25% aqueous methanol (25-M).

In addition, for further validation of GC-MS data, XCMS online, a web based platform was also employed for feature detection, retention time correction, alignment, annotation and visualisation of the results using cloud plot and non-metric multidimensional scaling (MDS) plots. Only selected extraction methods were considered for further analysis, which included the 25-M, 25-A, 50-M, 50-A, 100-M and G methods. Smaller changes in the retention time of a GC method could largely affect the quality of the data obtained. Therefore, to ensure the reliable and accurate identification and quantification of metabolites from microalgal samples, it is essential to determine that, there is no large drift in retention times across all the samples analysed. To determine deviations in the retention time across all the samples, XCMS online platform was used where retention time corrections were done by OBI-Warp and peak groups method (Prince and Marcotte, 2006). The results of this investigation are summarised in figure 8.11.

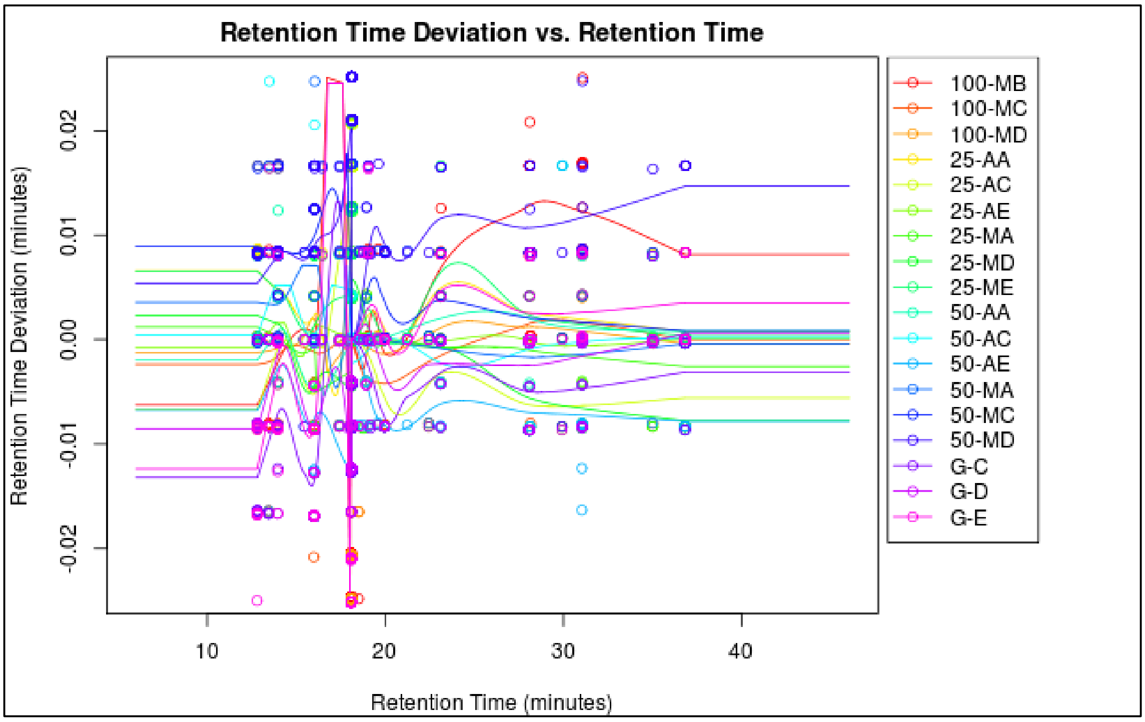


Figure 8.11 Deviations of retention time of all samples are shown in relation to the retention time of the first sample in the legend (black line) (OBI-Warp method) and in relation to the median retention time of features from all samples (peak groups).

As can be seen from figure 8.11, there was a negligible retention time deviation (±0.02 min) across all the samples, thus ensuring the accuracy and reliability of the data obtained and thereby conclusions drawn from such data.

For accurate and reliable interpretation of metabolomics results, it is essential to obtain information about data characteristics such as *p*-value, fold change, retention time, *m/z* ratio and signal intensity of all detected features. Numerous visualisation techniques such as PCA, heat maps, volcano plots and scatter plots have been employed so far to gain insight into these data characteristics, however none of these tools have the capabilities to simultaneously represent all these data characteristics. Recently, cloud plot visualisation tool also known as differential feature plot was introduced, which can simultaneously represent all of these data characteristics. Therefore, the data generated by GC-MS analysis from the above mentioned extraction methods was processed further and used to generate cloud plot using the XCMS online platform. The results of these investigations are summarised in figure 8.12.

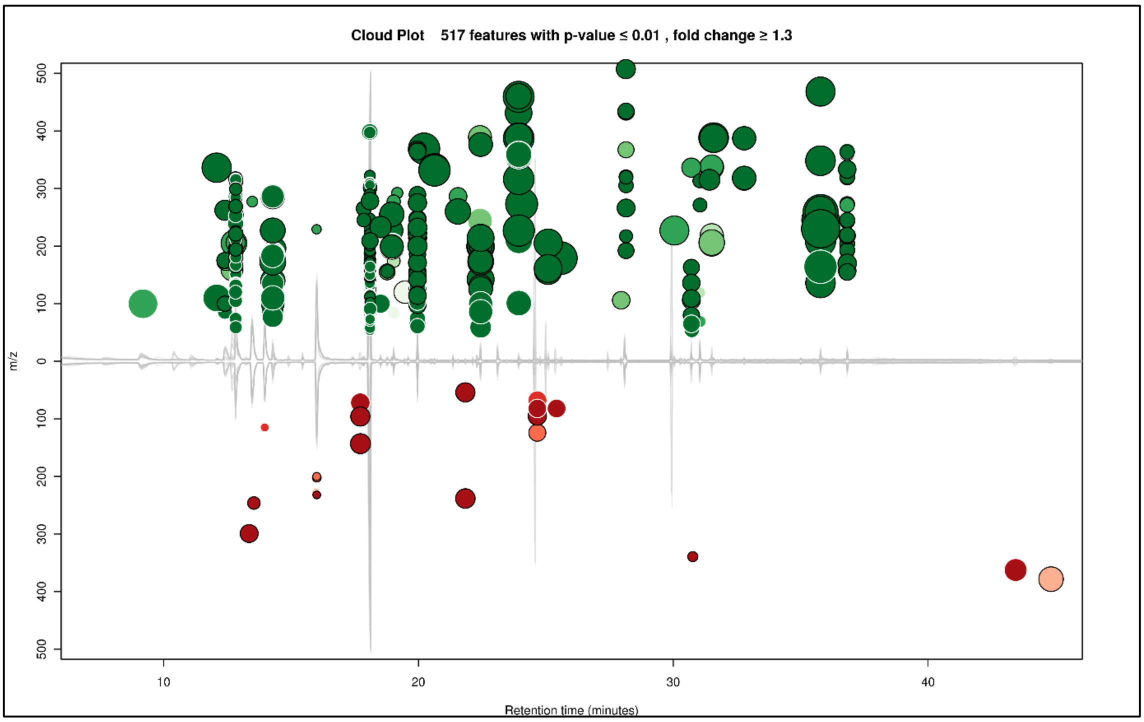


Figure 8.12 Cloud plot of extraction methods data set displaying 517 features with p-value ≤ 0.01 and fold change ≥ 1.3. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value. Darker the colour, the smaller the p-value.

As can be seen from figure 8.12, in total 517 dysregulated features were detected whose intensities were altered between sample groups with p-value ≤ 0.01 and fold change ≥ 1.3, when 100-M method (Dataset 1 as control) was compared against other extraction methods (Dataset 2), which included 25-M, 25-A, 50-M, 50-A and G method. Up-regulated features are shown in green, whereas down-regulated features were shown in red bubbles. The size of each bubbles corresponds to the log fold change of that feature. The shade of the bubble corresponds to the magnitude of the p-value, darker the colour, the smaller the p-value. The statistical significance of the fold change, as calculated by a Welch *t* test with unequal variances, is represented by the intensity of the features colour, where features with low *p*-values are brighter compared to features with high *p*-values. The Y- co-ordinate of each feature corresponds to the *m/z* ratio of the compound as determined by the MS. The features that were identified with the METLIN database are being colour coded with the black outline surrounding the bubble, whereas those not identified in the database are displayed without a black outline.

It is important to note that, the main objective of employing the XCMS platform was not to identify the metabolites with the METLIN (human) database, as we have already done this above with the NIST and GOLM (plant) databases but to effectively validate our previous conclusion drawn from our analysed datasets and further supporting them by taking advantages of these visualisation tools to display significant differences across the extraction methods employed in this study.

In addition to the cloud plot, we have also analysed the results by non-metric multidimensional scaling (MDS) in order to visualise the high-dimensional datasets obtained with various extraction methods, results of which are summarised in figure 8.13. The MDS plots were generated based on the median intensities of all the aligned features in order to depict similarities between the samples from various extraction methods and to assess the reproducibilities of each method evaluated in this study.

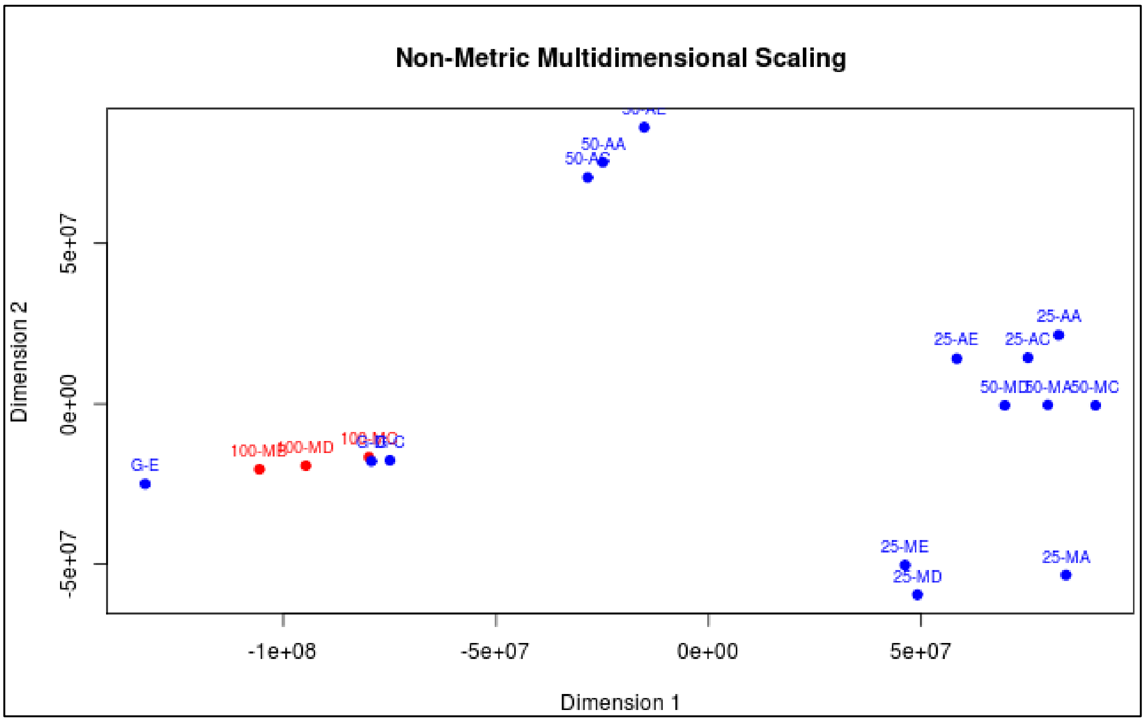


Figure 8.13 Non-metric multidimensional scaling (MDS) plot was generated based on the median intensities of all aligned features detected across all the applied extraction methods, in order to explore similarities or dissimilarities between the sample data sets. The colours (red/blue) are assigned based on the extraction solvent system employed where: Red = 100-M (pure methanol) and Blue = other solvents systems which includes: 25-A= 25% aqueous acetonitrile; 25-M = 25% aqueous methanol; 50-A = 50% aqueous acetonitrile; 50-M = 50% aqueous methanol and G = methanol/chloroform/water (5:2:2). A, B and C = biological replicates.

As can be seen from figure 8.13, all the three technical replicates within each method were placed closely on the plot for all the methods except for that of G method, thereby displaying poor reproducibility of methanol/chloroform/water (5:2:2) extraction solvent (G method). On the other hand, it can be clearly seen that the median intensities obtained with 25-M was significantly higher than all other applied methods. Significantly lower intensities were obtained with 100-M and G method, compared to 25-M method. This finding was in agreement with our previous conclusion as discussed in section 8.3.3.2.

8.3.3.4 Agreement and dis-agreement with respect to outcome of our findings

As can be seen from table 8.1, contradictory reports were published concerning the best extraction solvent for specific cell types. Considering this our results were also in partial agreement and dis-agreement with previously published reports. In such case, it is worthwhile to clarify the causes for such apparent contradictions. To our knowledge, there is no evidence in the literature where monophasic and biphasic solvents systems have been systematically compared together with respect to the recoveries of intracellular metabolites from *C. reinhardtii*. Although few reports compared and evaluated some extraction methods (Bölling and Fiehn, 2005, Lee and Fiehn, 2008a), the analysis was restricted to evaluation of very limited extraction solvents and more specifically concentrated on various solvent ratios used for MCW solvent system. Due to these reasons our findings are in contrast with the two earlier reports on *C. reinhardtii*, where the authors reported minor variations across all the evaluated extraction solvents. We have clearly demonstrated that the choice of extraction solvent has a significant impact on the recovery of intracellular metabolites from *C. reinhardtii.* Similar conclusions were drawn previously with *Arabidopsis thaliana* (Gullberg et al., 2004b) and with yeast sample (Canelas et al., 2009). However, conclusions drawn from studies carried on yeast samples (Canelas et al., 2009) were based on only a subset of the metabolome, which cannot be extrapolated to other wide range of metabolites in order to derive any meaningful conclusions. In contrast, conclusions drawn in our study were based on broader range of metabolite classes. Similarly, conclusions drawn from studies on mammalian samples (Araujo et al., 2013), were based on analysis of selected 20 metabolites, where authors reported higher recoveries in MCW-with inclusion of acetonitrile (BI) compared to pure methanol (100-M) and MCW mixture (G), which was in agreement with our findings.

Ideally, factors such as extraction efficiency and analysis with respect to metabolite degradation or inter-conversion due to enzymatic activities should be considered whilst optimising extraction methods. However, at this stage, we have focused on determining the number of peaks detected for each extraction solvent, the peak intensity of structurally identified compound, and the reproducibility of metabolite quantification with minimal analytical variations. In case of filamentous fungi, Madla and co-workers reported higher recoveries with 100-M compared to MCW mixture (Madla et al., 2012), whereas in our case nearly similar recoveries were obtained with 100-M and G methods. Despite advantages of biphasic solvent systems, lower than expected recoveries observed with G method compared to 100-M might be due to improper solvent ratios being employed. As in case of RK method, with optimised solvent ratios, the recoveries were greatly improved for all metabolite classes. Another possible reason behind lower recoveries may be inappropriate volumes of chloroform being employed in the MCW mixture which will increase risk of co-precipitation of lipophilic metabolites and other compounds closely associated with protein complexes (Lee and Fiehn, 2008a). Inclusion of treatments involving mixing of two or more solvents at fixed proportions as suggested and optimised previously for *Saccharophagus degradans* (Shin et al., 2010), appears to have no significant effect on improving metabolite recoveries in *C. reinhardtii* when compared to other treatments evaluated in our study. Shin and co-workers demonstrated excellent recoveries in numbers with AMW and WiPM compared to that obtained with pure methanol and 50% aqueous acetonitrile (Shin et al., 2010). In our study, similar results were obtained, where higher recoveries were obtained with AMW and WiPM compared to use of pure methanol, however contrasting findings were observed where higher recoveries were obtained with 50% aqueous methanol compared (which was not included in other study) to that obtained with AMW and WiPM solvent systems. In case of mammalian metabolomics, few studies have evaluated biphasic solvent systems. However, results have been based on analysis of only the aqueous fraction (Martineau et al., 2011), in such cases direct comparison with our study cannot be made as we have analysed both the aqueous and the organic fractions prior to drawing any conclusions. As mentioned in section 8.3.1, causes behind variations reported in previously published reports might be due to either different cell types under investigation or use of different analytical platform (LC-MS) or might be due to comparison based on few extraction methods (Dettmer et al., 2011, Hutschenreuther et al., 2012a), not covering a wide range of extraction protocols as considered in our case or might be due to conclusions drawn from studies based on analysis of only small subset of metabolome (such as ATP or only metabolites having high turnover rate) (Faijes et al., 2007) rather than broad range of metabolites as considered in our case or combination of above.

8.4 Conclusions

In summary, we have performed a comprehensive evaluation of 15 extraction solvents for untargeted GC-MS based metabolome analysis of *C. reinhardtii*. For comparative analysis, we have initially evaluated and compared previously published monophasic solvent systems for the extraction efficiency. We have also evaluated and optimised the biphasic solvent systems. Finally, in order to propose the best extraction solvent system for *C. reinhardtii*, both the monophasic and biphasic solvent systems with various treatments involved within them were compared, assessed, evaluated and validated with respect to the recoveries of different classes and numbers of metabolites. The evaluation criteria adopted to define extraction efficacy of different extraction methods involved estimation of number of peaks detected for each extraction solvent, the peak intensity of structurally identified compound and the reproducibility of metabolite quantification with minimal analytical variations. Based on our results, we have successfully demonstrated that the choice of extraction solvent has a significant impact on recovery of intracellular metabolites from *C. reinhardtii.* Our results clearly indicate that 25 % aqueous methanol is the best suited extraction solvent for untargeted metabolomic analysis of *C. reinhardtii*, as the highest number of metabolites belonging to various chemical classes were recovered with good intensities and reproducibility compared to other evaluated methods. In case of biphasic solvent system, we propose MCW mixture at a solvent ratio of (2.5:3:1.5) as a best extraction sovent as optimal solvent ratios resulted in equal distribution of upper polar aqueous phase (methanol-water) and lower non-polar phase (chloroform), therby improving the metabolome coverage compared to other biphasic solvent systems evaluated. The use of pure solvents such as methanol, acetonitrile and ethanol alone has shown poor recoveries compared to other methods, therefore not recommended.

Chapter 9

Optimisation and comparative evaluation of conventional and direct trans-esterification methods for reliable quantification of FAMEs from micro-algal biofuel strains

9. Optimisation and comparative evaluation of conventional and direct trans-esterification methods for reliable quantification of FAMEs from micro-algal biofuel strains

9.1 Introduction

Microalgae are attractive alternatives to terrestrial oleaginous specie and represent a promising feedstock for the production of biofuels. Several microalgal strains have been reported to be significantly rich in oil which can be converted into biodiesel, however parameters such as chain length and level of unsaturation of FAMEs obtained are crucial for fuel properties and are largely dependent on employed algal strain and applied growth conditions (Daroch et al., 2013). Chain length of between C14 to C22 and low unsaturation level are considered to be the most ideal for biodiesel production.

Fatty acid methyl esters (FAMEs) can be produced from the microalgal lipids through trans-esterification process, however overall workflow involves different stages which includes: selection of suitable species with higher lipid contents, optimising the parameters for cultivation and harvesting of microalgae, selection of appropriate lipid extraction technique and solvent, trans-esterification of extracted lipids into FAMEs and reliable quantification of the lipids obtained at various stages. For microalgal biodiesel production, an ideal extraction method should be more selective towards extraction of specific lipid classes and simultaneously minimise the co-extraction of non-lipid contaminants. Several methods have been used for extracting lipids from algae in which the first step is to disrupt the microalgal cell wall for which selection of appropriate extraction method will depend primarily on the rigidity of microalgal cell wall and the lipid classes of interest. A large number of protocol were based on conventional method in which lipids were extracted using solvent or solvent mixture followed by their transmethylation, whereas on the other hand in case of direct transesterification methods lipids were extracted and methylated in a single step by addition of methylating agent directly to the algal biomass avoiding the lipid extraction step.

Several authors compared both conventional methods and suggested a few modifications to them (Griffiths et al., 2010, Iverson et al., 2001, Kumari et al., 2011), however it is interesting to note that the samples under investigation had affected lipid compositions leading to inconsistencies in reported methodologies and variable conclusions. Recently Iverson and co-workers applied both methods to marine samples containing >2% lipid and concluded that the Bligh and Dyer method resulted in significant lower estimates of lipid contents than that of the classical Folch method (Iverson et al., 2001). Smedes and co-workers suggested modification to the original Bligh and Dyer method with respect to the solvent ratios used and proposed that, higher methanol proportions improves the recovery of lipids (Smedes, 1999b). Recent studies evaluated both methods for the recovery of lipids from microalgal samples, and concluded that the Folch extraction method was better than the Bligh and Dyer method for algal lipid estimation (Griffiths et al., 2010, Kumari et al., 2011). Recently, chloroform has been evaluated by U.S EPA for its carcinogenic properties and suggested that it is likely to be carcinogen (Andersen et al., 2000). On the other hand, methanol poisoning in primates and non-primates is known (Clay et al., 1975, Tephly, 1991). Therefore, the use of toxic solvents on an industrial scale will have a great impact on the environment and will pose a threat to health. In order to this threat, it is essential to minimise their use or safer alternatives should be used, which are less toxic and less harmful.

TE is a multiple step reversible reaction in which triglycerides are converted to diglycerides, then diglycerides are converted to monoglycerides and then monoglycerides are converted to fatty acid methyl esters (FAME) and glycerol (Mata et al., 2010). TE reactions are generally carried out using short-chain alcohol (usually methanol) and in presence of acid or base or sequential use of both acid and base catalyst to enhance the TE reaction rates. Direct transesterification (DT) methods offer several advantages over two step TE methods (extraction followed by TE). In DT methods, a separate extraction step is avoided and FAs are simultaneously extracted and transesterified in presence of a catalyst. The esterification or TE reactions are carried out either after extraction of lipids or else they can be effected while the lipids are present inside the sample matrix. As discussed in Chapter 2, there are several challenges involved in *In situ* lipids extraction and TE, among which the most important one is, the selection of appropriate solvent for lipid solubilisation. The presence of specific small amount of water results in precipitation of the long chain SFAs and TAGs and, undergoes very slow reaction with most of the frequently used solvents (Carrapiso and García, 2000). Moreover, presence of water in *in situ* TE reaction may cause strong interference as water is a stronger electron donor than methanol. Therefore reactions must be performed devoid of water or only on samples with low moisture content. Lyophilisation is the preferred way to remove water from sample matrix, as oven drying can bring out alterations in unsaturated FAs. Removal of water facilitates penetration of solvents however in some samples it may decrease the extraction yield due to alteration of the physical structure of the sample matrix. Several researchers have suggested the use of water scavenger such as 2,2-dimethoxypropane which can also be used as methylating agent, however it has been shown to result in appearance of spurious peaks in the chromatograms because of the presence of unreacted reagent (Carrapiso and García, 2000). However, inclusion of such pre-treatments (lyophilisation or addition of an antioxidant) should be carefully considered for large scale biodiesel production.

Various methods have been developed for the quantification of lipids and fatty acids. The commonly used conventional methods include extraction of lipids by using single or mixed solvents followed by lipid quantification by gravimetry (weighing) or by chromatographic methods such as thin layer chromatography (TLC), High performance liquid chromatography (HPLC) and gas chromatography (GC). However growing research in microalgal biodiesel application demands for development of an analytical method for reliable quantification of fatty acids.

Given this, the objective of this investigation is to address the major challenges involved in biodiesel production from microalgae which includes:

- Selection of suitable microalgal strain for biodiesel application.
- Evaluation of various disruption techniques using design of experiment approach
- Development of GC-FID based method for reliable quantification of fatty acids
- Comparative evaluation of conventional (2 step TE) and direct TE methods

9.2 Material and Methods

9.2.1 Chemicals and analytical reagents

All the chemicals and analytical reagents used were of chromatography standard. For the identification and quantification of fatty acids the 37 component FAME mix (C4–C24) with varied concentration in dichloromethane was purchased from Supelco USA. All the internal standards, other analytical reagents and HPLC grade solvents used, were purchased from Sigma-Aldrich (Dorset, U.K.), unless stated otherwise.

9.2.2. Selection of microalgal species, cultivation and harvest

Three different microalgal species from two different phyla were analysed in this study. All strains were obtained from the culture collection of Algae and Protozoa (CCAP, UK). The *C. reinhardtii* strain (CC4323) was grown under similar conditions as detailed in Chapter 7 – section 7.2.2. *N. salina* strain (CCAP 849/2) and *D. salina* strain (CCAP 19/18) were grown under similar conditions as detailed in Chapter 6 – section 6.2.1. The cells were harvested for all the three species at an OD of 1.0 at a wavelength of 680nm.

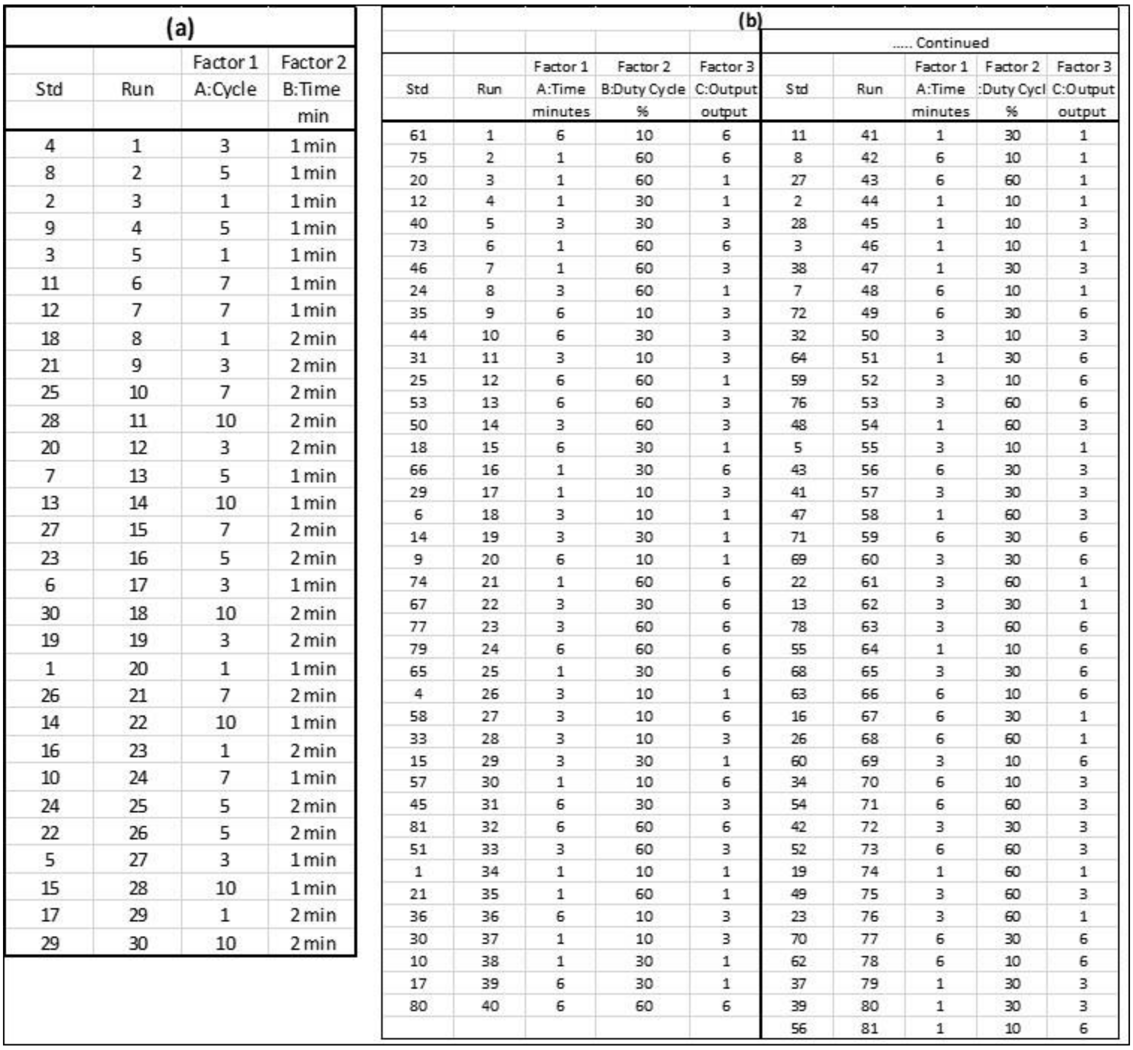
The cells were harvested in 5mL aliquots in 15mL Falcon tubes and centrifuged at 4500 rpm. The 4 mL supernatant was discarded and cells were re-suspended in 1mL of supernatant, transferred to the new Eppendorf (pre-weighed) and centrifuged further at 8500 rpm for 3minutes at 4°C.

The supernatant was discarded and the pellets obtained were weighed to estimate the wet algal biomass and frozen at -80°C until further analysis. Dry cell weight (DCW) determinations were performed in triplicate for all the three strains, where pellets were dried at 105°C for 2 h and re-weighed. The harvesting sample volume for all the three species was kept constant for all the treatments investigated in this study.

9.2.3 Evaluation of cell disruption techniques using experimental design

We have investigated most commonly employed two cell disruption techniques on selected two microalgal species. To test the robustness of combining different parameters within each cell disruption technique across two different species, we employed Design expert software Version 7.0.b1.1 (stat-Ease, Inc., Minneapolis, MN, www.statease.com) to design the experiments and evaluate the results. In case of bead beating experiments, we have generated two factorial design containing three replicates for each condition, which involves cycle (factor 1, range 1-10 cycles) and relaxation interval (factor 2, range 1-2 minutes), which resulted in total number of 30 samples for analyses as shown in table 9.1 (a). Each cycle represents disruption followed by relaxation interval on ice for specified time. Whereas, in case of sonication experiments, we have generated three factorial design containing three replicates for each condition, which involves time (factor 1, range 1-6 minutes), duty cycle (factor 2, range 10-60 %) and output (factor 3, range 1-6), which resulted in total number of 81 samples for analyses as shown in table 9.1 (b). All the sonication treatments were carried out on ice with an interval of 1 minute.

Table 9.1 Bead beating (a) and sonication (b) treatments generated by two and three factorial design respectively using design of expert software



(a)

Std	Run	Factor 1 A:Cycle	Factor 2 B:Time min
4	1	3	1 min
8	2	5	1 min
2	3	1	1 min
9	4	5	1 min
3	5	1	1 min
11	6	7	1 min
12	7	7	1 min
18	8	1	2 min
21	9	3	2 min
25	10	7	2 min
28	11	10	2 min
20	12	3	2 min
7	13	5	1 min
13	14	10	1 min
27	15	7	2 min
23	16	5	2 min
6	17	3	1 min
30	18	10	2 min
19	19	3	2 min
1	20	1	1 min
26	21	7	2 min
14	22	10	1 min
16	23	1	2 min
10	24	7	1 min
24	25	5	2 min
22	26	5	2 min
5	27	3	1 min
15	28	10	1 min
17	29	1	2 min
29	30	10	2 min

(b)

Std	Run	Factor 1 A:Time minutes	Factor 2 B:Duty Cycle %	Factor 3 C:Output output
61	1	6	10	6
75	2	1	60	6
20	3	1	60	1
12	4	1	30	1
40	5	3	30	3
73	6	1	60	6
46	7	1	60	3
24	8	3	60	1
35	9	6	10	3
44	10	6	30	3
31	11	3	10	3
25	12	6	60	1
53	13	6	60	3
50	14	3	60	3
18	15	6	30	1
66	16	1	30	6
29	17	1	10	3
6	18	3	10	1
14	19	3	30	1
9	20	6	10	1
74	21	1	60	6
67	22	3	30	6
77	23	3	60	6
79	24	6	60	6
65	25	1	30	6
4	26	3	10	1
58	27	3	10	6
33	28	3	10	3
15	29	3	30	1
57	30	1	10	6
45	31	6	30	3
81	32	6	60	6
51	33	3	60	3
1	34	1	10	1
21	35	1	60	1
36	36	6	10	3
30	37	1	10	3
10	38	1	30	1
17	39	6	30	1
80	40	6	60	6

..... Continued

Std	Run	Factor 1 A:Time minutes	Factor 2 :Duty Cycl %	Factor 3 C:Output output
11	41	1	30	1
8	42	6	10	1
27	43	6	60	1
2	44	1	10	1
28	45	1	10	3
3	46	1	10	1
38	47	1	30	3
7	48	6	10	1
72	49	6	30	6
32	50	3	10	3
64	51	1	30	6
59	52	3	10	6
76	53	3	60	6
48	54	1	60	3
5	55	3	10	1
43	56	6	30	3
41	57	3	30	3
47	58	1	60	3
71	59	6	30	6
69	60	3	30	6
22	61	3	60	1
13	62	3	30	1
78	63	3	60	6
55	64	1	10	6
68	65	3	30	6
63	66	6	10	6
16	67	6	30	1
26	68	6	60	1
60	69	3	10	6
34	70	6	10	3
54	71	6	60	3
42	72	3	30	3
52	73	6	60	3
19	74	1	60	1
49	75	3	60	3
23	76	3	60	1
70	77	6	30	6
62	78	6	10	6
37	79	1	30	3
39	80	1	30	3
56	81	1	10	6

9.2.3.1 Beat beating extraction protocol

To the frozen pellets harvested from 5 mL aliquots, 1.2 mL of a methanol:chloroform (1:2, *v/v*) mixture was added along with an equal volume of glass beads (425-600 µm i.d., acid washed, from Sigma). Cells were then disrupted using Cell disruptor (from Genie) according to the run order listed in table 9.1 (a). After cell disruption, the sample was then centrifuged at 14000 rpm, at 4°C for 10 min to remove any cell debris and beads. The supernatant was carefully transferred to new pre-chilled Eppendorf tube (4°C) to which 400uL of each chloroform and water was added. The resultant mixture was then centrifuged again at 8500 rpm at 4°C for 10 min. The organic phase (bottom layer) was recovered, measured and transferred to the new Eppendorf tube (pre-weighed) followed by evaporation to dryness under inert nitrogen gas using six port mini-vap evaporator (Sigma-Aldrich, Dorset, U.K.). The dried extract was then stored at -80°C until further analysis.

9.2.3.2 Sonication extraction protocol

The similar protocol was adopted as described in section 3.2.3.1, except instead of using cell disruptor, cells were disrupted using sonicator (450 Sonifier, from Branson), according to the run order listed in table 9.1 (b).

9.2.3.3 Preparation of FAMEs for cell disruption techniques

The extracted lipids were then converted into fatty acid methyl esters (FAMEs) as described by Vandenbrouck and co-workers with little modifications (Vandenbrouck et al., 2010a). Briefly 250uL of chloroform:methanol (1:1, *v/v*) was added to the dried extract and then transferred to the 2 mL amber colour glass vials, followed by addition of 100uL of 10% BF_3/methanol (*w/v*). The mixture was then incubated at 80°C for 90 minutes. The samples were cooled to room temperature for 10 minutes prior to addition 300uL of water, 600uL of hexane. The contents were then transferred to the new Eppendorf, vortexed and centrifuged at 13000 rpm at 4°C for 10 min. The organic phase (top layer) was separated from aqueous phase and transferred to the new Eppendorf. 500uL measured volume of organic phase was then evaporated to dryness under inert nitrogen gas using six port mini-vap evaporator. The resulting dried FAMEs were then reconstituted in 100uL of hexane prior to their identification and quantification by GC-FID as described in section 9.2.8.

9.2.4 Comprehensive evaluation of conventional two step TE methods

The lipid extraction efficiency of conventional methods like Bligh and Dyer (Bligh and Dyer, 1959), Folch (Folch et al., 1957) and Smedes (Smedes, 1999b) were compared to those of recently suggested modified methods (Cequier-Sánchez et al., 2008, Matyash et al., 2008, Ryckebosch et al., 2012, Sheng et al., 2011) to test the effect of different solvent mixtures and ratio. In addition, we have applied the optimized bead beating cell disruption parameters across all the methods as described in section 9.2.3 to that of the conventional methods as a part of modification and optimization. All the investigated methods were listed in table 9.2 and described below.

Table 9.2 List of all the methods investigated along with their codes used in this study

Codes	Method details	Method type
1-A	*Bligh & dyer*	**Conventional methods**
1-B	*Bligh & dyer with Bead beating*	
1-C	*Bligh & dyer with Sonication*	
1-D	*Modified BD by Smedes*	
2-A	*Folch*	
2-B	*Folch with Bead beating*	
2-C	*Folch with Sonication*	
2-D	*Modified Folch method*	
3	*Methanol:MTBE (1.5:5) (Matyash & Sheng)*	**Other conventional methods**
4	*Methanol:Chloroform (1:1) (Ryckebosch)*	
5	*Methanol:DCM (1:2) (Cequire-Sanchez)*	
6-A	*5% Methanolic HCl; Toluene for lipids solubilisation(Garcias: Original protocol)*	**DT Acid Catalyst**
6-B	*5% Methanolic HCl; Chlorofrom:Methanol (2:1) for lipid solubilisation (Laurens: Little modification)*	
6-C	*10% BF3 in methanol (Laurens: Little modification)*	
6-D	*5% Acetyl chloride in methanol*	
7-A	*0.5M Sodium methoxide*	**DT Base Catalyst**
7-B	*1:4 v/v TMG in methanol (Laurens: Modified)*	
8-A	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Original Protocol)*	**DT Both Acid & Base Catalyst**
8-B	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Modified Protocol)*	
8-C	*Sodium meth./BF3 in methanol. Incubated together (Laurens: Original Protocol)*	
8-D	*Sodium meth. & Acetyl chloride. Incubated separately*	
8-E	*Sodium meth. & M. HCl Incubated separately*	

Method 1: Bligh and dyer method (BD)

1-A: Bligh and Dyer method

The original Bligh and Dyer method (Bligh and Dyer, 1959) suggests that the protocol can only be applied to those samples which contain at least 80 ± 1% of water and about 1% lipid. Moreover, it is essential to keep the volumes of chloroform, methanol and water, before and after dilutions in the proportion of 1:2:0.8, *v/v/v* and 2:2:1.8, *v/v/v* respectively. (Ratio's considering the amount of water present in the wet algal biomass). So in order to mimic the original protocol as closely as possible, the water content of the wet algal biomass was first determined and adjusted to 80%. All the extraction steps were performed in the same way as (Bligh and Dyer, 1959); except that the sample and solvent amounts were reduced drastically and instead of homogenization the samples were vortexed at maximum speed for 3 minutes. The total lipid contents were determined gravimetrically as described in section 9.2.6. The extracted lipids were then subjected to the trans-esterification process for FAME analysis as described in section 9.2.3.3.

1-B: BD method with bead beating

Lipids were extracted in the same way as described in method 1A, except that instead of vortexing the samples, the optimized bead beating parameters were used.

1-C: BD method with sonication

Lipids were extracted in the same way as described in method 1A, except that instead of vortexing the samples, the optimized sonication parameters were used.

1-D: Modified BD method by Smedes with bead beating

Lipids were extracted using non-chlorinated solvents as suggested by (Smedes, 1999a), where methanol was replaced by propan-2-ol and chloroform was replaced by cyclohexane. The proposed solvent ratios used in this method are propan-2-ol-cyclohexane-water (8:10:11, *v/v/v*) considering the amount of water present in the wet algal biomass. All the extraction steps were performed in the same way as described by Smedes et al (Smedes, 1999a); except that instead of homogenization, samples were bead beaten with the optimized parameter settings. The total lipid contents were determined gravimetrically as described in section 9.2.6. The extracted lipids were then subjected to the trans-esterification process for FAME analysis as described in section 9.2.3.3.

Method 2: Folch method

2-A: Folch method

The original Folch method (Folch et al., 1957) was based on the assumption that the tissue/sample has specific gravity of water, so 1 g of tissue/sample is computed as 1 ml. Moreover, (Folch et al., 1957) implies on initial homogenization of the sample with chloroform-methanol mixture (2:1, *v/v*) to a final dilution, 20 fold the volume of wet sample and final ratios of chloroform-methanol-water should be kept as 8:4:3, *v/v/v*. So in order to mimic the original protocol as closely as possible, the wet weight of algal biomass was first determined. All the extraction steps were performed in the same way as (Folch et al., 1957); except that instead of homogenization, samples were bead beaten with the optimized parameter settings. The total lipid contents were determined gravimetrically as described in section 9.2.6. The extracted lipids were then subjected to the trans-esterification process for FAME analysis as described in section 9.2.3.3.

2-B: Folch method with bead beating

Lipids were extracted in the same way as described in method 2A, except that instead of vortexing the samples, the optimized bead beating parameters were used.

2-C: Folch method with sonication

Lipids were extracted in the same way as described in method 2A, except that instead of vortexing the samples, the optimized sonication parameters were used.

2-D: Folch method with use of non-chlorinated solvents and bead beating

Lipids were extracted in the same way as described in method 2A, except that instead of vortexing the samples, the optimized bead beating parameters were used. Moreover, methanol was replaced by propan-2-ol and chloroform was replaced by cyclohexane while keeping the final solvent ratios of propan-2-ol/cyclohexane/water to (8:10:11, *v/v/v*). All other conditions were kept constant as that of method 2-A. The total lipid contents were determined gravimetrically as described in section 9.2.6. The extracted lipids were then subjected to the trans-esterification process for FAME analysis as described in section 9.2.3.3

Method 3: Matyash and Sheng method

Lipids were extracted as suggested by (Matyash et al., 2008, Sheng et al., 2011), where chloroform was replaced by methyl-*tert*-butyl ether (MTBE) while keeping the ratio of methanol-MTBE to 1.5:5 (*v/v)*. Briefly, 300uL of methanol was added to the wet algal biomass along with an equal volume of glass beads (425-600 µm i.d., acid washed, from Sigma). The samples were vortexed for 2 minutes followed by addition of 1000uL of MTBE. Instead of incubation at room temperature as suggested (Matyash et al., 2008, Sheng et al., 2011); samples were bead beaten with the optimized parameters. 250uL of water was added and sample were vortexed again for 1 minute, followed by centrifugation for 10 minutes at 8500 rpm. The upper organic layer of MTBE was recovered, measured and transferred to the new Eppendorf (pre-weighed), while the bottom aqueous phase was discarded. All the extract was evaporated to dryness under inert nitrogen gas using six port mini-vap evaporator (Sigma-Aldrich, Dorset, U.K.). The total lipid contents were determined gravimetrically as described in section 9.2.6. The extracted lipids were then subjected to the trans-esterification process for FAME analysis as described in section 9.2.3.3.

Method 4: Ryckenbosch method

Lipids were extracted using methanol-chloroform mixture (1:1, *v/v*) as suggested by (Ryckebosch et al., 2012) with little modifications. Briefly, 1.2 mL of a methanol-chloroform (1:1, *v/v*) mixture was added to the wet algal biomass along with an equal volume of glass beads (425-600 µm i.d., acid washed, from Sigma). Cells were then disrupted using Cell disruptor (from Genie) with the optimized parameters. After cell disruption, the samples were then centrifuged at 14000 rpm, at 4°C for 10 min to remove any cell debris and beads. The supernatants were carefully transferred to new pre-chilled Eppendorf tube (4°C) to which 400uL of each chloroform and water were added. The resultant mixture was centrifuged again at 8500 rpm at 4°C for 10 min. The organic phase (bottom layer) was recovered, measured and transferred to the new Eppendorf tube (Pre-weighed) followed by evaporation to dryness under inert nitrogen gas using six port mini-vap evaporator (Sigma-Aldrich, Dorset, U.K.). The total lipid contents were determined gravimetrically as described in section 9.2.6. The extracted lipids were then subjected to the trans-esterification process for FAME analysis as described in section 9.2.3.3.

Method 5: Cequire-Sanchez method

Lipids were extracted using methanol-dichloromethane mixture (1:2 *v/v*) as suggested by (Cequier-Sánchez et al., 2008) otherwise all the steps were same as described in method 4.

9.2.5 Comprehensive evaluation of direct trans-esterification (DT) methods

After thorough literature review, in total eleven direct-transesterification (DT) protocols were selected or constructed which were then compared within and with the conventional methods, to test the efficiency of FAME production from microalgal species. Primarily, all the DT protocols were categorized based on the nature of the catalyst used to catalyse the esterification or TE reactions. Briefly, they were classified into acid catalysed, base catalysed and combined (acid and base) catalysed DT methods. All the DT methods were modified scaled down procedure on biomass quantities between 1 to 2 mg of dry cell weight for *N. salina* and *D. salina* and between 15 to 16 mg of dry cell weight for *C. reinhardtii*. However, all the protocols were carried out on wet algal biomass, where the wet cell weight was 8 to 10 mg for *N. salina* and *D. salina* and between 50 to 60 mg for *C. reinhardtii.* Across all the methods, the volume of organic phase recovered was measures and only fixed volume of organic phase was processed further for drying to avoid intra- and inter-method variations caused due to variable solvent volumes.

Method 6: Acid catalysed DT methods

6-A: 5% Methanolic hydrogen chloride (*v/v*) (Garcia's method)

The DT protocol was modified from that of Garcia (de La Cruz Garcia et al., 2000). Briefly, 350uL of toluene was added to the wet algal biomass. The samples were mixed thoroughly by vortexing, followed by the addition of 550ul of 5% methanolic hydrogen chloride. The samples were then incubated at 80°C for 1 h. After cooling to room temperature, 730uL of aqueous 6%

K_2CO_3 and 350uL of toluene were added. The mixture was vortexed for 1 minute at maximum speed and centrifuged at 13000 rpm at 4°C for 10 min. The organic phase (upper toluene layer) was recovered, measured and transferred to the new Eppendorf tube (Pre-weighed) followed by evaporation to dryness under inert nitrogen gas. The resulting FAMEs were then re-suspended in 100uL toluene prior to their identification and quantification by GC-FID as described in section 9.2.8.

6-B: 5% Methanolic hydrogen chloride (Lauren's method)

The DT protocol was modified from that of Laurens (Laurens et al., 2012b). Briefly, 200uL of chloroform-methanol mixture (2:1, *v/v*) was added to the wet algal biomass. The samples were mixed thoroughly by vortexing, followed by the addition of 300ul of 5% methanolic hydrogen chloride. The samples were then incubated at 80°C for 1 h. After cooling to room temperature, 300uL of water and 600uL of hexane were added. The mixture was vortexed for 1 minute at maximum speed and centrifuged at 13000 rpm at 4°C for 10 min. The organic phase (upper hexane layer) was recovered, measured and transferred to the new Eppendorf tube (Pre-weighed) followed by evaporation to dryness under inert nitrogen gas. The resulting FAMEs were then re-suspended in 100uL hexane prior to their identification and quantification by GC-FID as described in section 9.2.8.

6-C: 10% Boron trifluoride in methanol (*w/v*) (Lauren's method)

All the steps were carried out in the same way as described above for method 6-B, however 10% BF_3 in methanol was used as an acid catalyst instead of 5% methanolic hydrogen chloride.

6-D: 5% Acetyl chloride/methanol (*v/v*)

All the steps were carried out in the same way as described above for method 6-B, however freshly prepared 5% acetyl chloride/methanol was used as an acid catalyst instead of 5% methanolic hydrogen chloride.

Method 7: Base catalysed DT methods

7-A: DT using 0.5M Sodium methoxide (Lauren's method)

All the steps were carried out in the same way as described above for method 6-B, however 200uL of toluene was used instead of 200 µL of chloroform-methanol mixture (2:1, *v/v*) to solubilize the lipids and in case of catalyst, 0.5M sodium methoxide was used as a base catalyst instead of 5% methanolic hydrogen chloride.

7-B: DT using 1:4 (*v/v)* tetramethylguanidine (TMG) in methanol (Modified Lauren's method)

All the steps were carried out in the same way as described above for method 7-A, however, freshly prepared 1:4, *v/v* TMG in methanol was used as a base catalyst instead of 0.5M sodium methoxide and samples were incubated for 30 minutes instead of 1 h.

Method 8: Combined (acid and base) catalysed DT methods

A representative snapshot of the processed samples using base and combined catalyst is displayed in fig. 9.1.

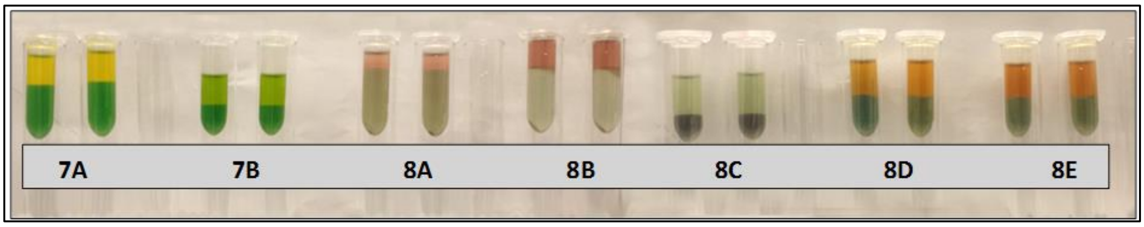


Figure 9.1 A representative snapshot of the samples processed by DT methods using base catalyst alone and sequential use of both base and acid catalyst. Visual differences on derived biodiesel (top organic layer) across different applied methods (summarised in Table 9.2) can be seen clearly.

8-A: 0.5M Sodium methoxide/10% Boron trifluoride in methanol (Griffith's method)

The DT protocol was modified from that of (Griffiths et al., 2010). Briefly, 300uL of toluene was added to the wet algal biomass. The samples were mixed thoroughly by vortexing, followed by the addition of 600ul of 0.5M sodium methoxide. The samples were then incubated at 80°C for 20 minutes. After cooling to room temperature, 600uL of 10% BF_3/methanol was added and mixture was incubated further at 80°C for 20 minutes. After cooling to room temperature, 250uL water and 250uL of hexane were added. The mixture was vortexed for 1 minute at maximum speed and centrifuged at 13000 rpm at 4°C for 10 min. The organic phase (upper hexane-toluene layer) was recovered, measured and transferred to the new Eppendorf tube (Pre-weighed) followed by evaporation to dryness under inert nitrogen gas. The resulting FAMEs were then re-suspended in 100uL toluene prior to their identification and quantification by GC-FID as described in section 9.2.8.

8-B: 0.5M Sodium methoxide/10% Boron trifluoride in methanol (Modified Griffith's method)

The DT protocol was modified from that of (Griffiths et al., 2010). Briefly, 300uL of toluene was added to the wet algal biomass. The samples were mixed thoroughly by vortexing, followed by the addition of 300ul of 0.5M sodium methoxide. The samples were then incubated at 80°C for 20 minutes. After cooling to room temperature, 300uL of 10% BF_3 in methanol was added and mixture was incubated further at 80°C for 20 minutes. After cooling to room temperature, 300uL water and 600uL of hexane were added. The mixture was vortexed for 1 minute at maximum speed and centrifuged at 13000 rpm at 4°C for 10 min. The organic phase (upper hexane-toluene layer) was recovered, measured and transferred to the new Eppendorf tube (Pre-weighed) followed by evaporation to dryness under inert nitrogen gas. The resulting FAMEs were then re-suspended in 100uL toluene prior to their identification and quantification by GC-FID as described in section 9.2.8.

8-C: 0.5M Sodium methoxide/10% Boron trifluoride in methanol (Lauren's method)

The DT protocol was modified from that of (Laurens et al., 2012b). Briefly, 300ul of 0.5M sodium methoxide and 300uL of 10% BF_3 in methanol were added to the wet algal biomass simultaneously. The samples were mixed thoroughly by vortexing and then incubated at 80°C for 55 minutes. After cooling to room temperature, 1000 µL of hexane was added. The mixture was vortexed for 1 minute at maximum speed and centrifuged at 13000 rpm at 4°C for 10 min. The organic phase (upper hexane-toluene layer) was recovered, measured and transferred to the new Eppendorf tube (Pre-weighed) followed by evaporation to dryness under inert nitrogen gas. The resulting FAMEs were then re-suspended in 100uL hexane prior to their identification and quantification by GC-FID as described in section 9.2.8.

8-D: 0.5M Sodium methoxide/5% Acetyl chloride in methanol

All the steps were carried out in the same way as described above for method 8-B, however 5% acetyl chloride in methanol was used as an acid catalyst instead of 10% BF_3 in methanol.

8-E: 0.5M Sodium methoxide/5% Methanolic hydrogen chloride

All the steps were carried out in the same way as described above for method 8-B, however 5% methanolic hydrogen chloride was used as an acid catalyst instead of 10% BF_3 in methanol.

9.2.6 Gravimetric analysis

The total lipid contents were determined gravimetrically (n=3) as mg per dry weight of algal biomass as suggested by (Burja et al., 2007) with little modification using the following equation; Total lipid content (mg/DW of algal biomass) =$(E_{Dry} - E_{Empty})\ x\ V_{Extract}/V_{Dry}\ x\ W_s$
where, E_{Dry} is the weight (mg) of the Eppendorf with dried residue, E_{Empty} is the weight (mg) of the empty Eppendorf, $V_{Extract}$ is the volume (mL) of final extract, V_{Dry} is the volume of extract transferred to the Eppendorf and W_s is the weight of the sample (g) extracted. All the weight measurements were carried out with Sartorius precision weighing microbalance CPA2P (Fisher Scientific, U.K.) at a precision of 0.001 mg.

9.2.7 Scanning electron microscopy (SEM) analysis

SEM analysis was performed on samples of untreated wet algal biomass (as a control), bead beating treated algae and sonication treated algae. All the specimens were prepared for SEM analysis as detailed in Chapter 4 – section 4.2.5.

9.2.8 Gas chromatography

The prepared FAMEs were identified and quantified using a Thermo Finnigan TRACE 1300 GC-FID System (Thermo Scientific, Hertfordshire, UK) onto a TR-FAME capillary column (25 m x 0.32 mm x 0.25 µm). The derivatized sample volume of 1µL was injected in split injection mode at 250°C. The split flow was 75 mL/min and purge flow was 5 mL/min. The GC was operated at a constant flow of 1.5 mL/min helium. The temperature program was started at 150°C for 1 min, followed by temperature ramping at 10°C/min to final temp of 250°C and held constant at 250°C for 1 min. The total analysis time was 12 minutes with detector temperature at 250°C.

9.2.9 Statistical analysis of the results

Initially peak identities were ascertained using external standard 37 component FAME mix (Supelco) and peak areas were integrated using chromatography data system (Chromeleon 7 software). Based on the known amount values of 37 FAME components, a ratio was established between the area and the amount. The amount of unknown component in the microalgal extract was then determined by their peak areas by generating calibration curve for individual component and quantified by reference external standard when expressing the amounts in mg/L. Data for identified FAMEs, retention time, peak areas and amounts in mg/L was collected manually from Chromeleon 7 software and exported in *.xlsx format to Microsoft office Excel 2013 for further data analysis. In total, for cell disruption techniques evaluation n=3 and for method evaluation n=5 replicates (biological) were ran, among which FAMEs identified only in 2 or more and 3 or more replicates respectively were considered as true hits.

9.2.10 Accuracy, precision, LOD and LOQ

The accuracy, precision, limits of detection (LOD) and quantification (LOQ) were evaluated with microalgal extract spiked with different concentration of three internal standards (n=5).

9.2.11 Biomass water content

The biomass water content was determined (n=3) using HB43_S halogen moisture analyser (Mettler Toledo, U.K.). To avoid losses of wet algal biomass during transfer from Eppendorf to moisture analyser plate, aluminium foil (pre-weighed) was placed in an Eppendorf prior to centrifugation of algal biomass, media was decanted and aluminium foil was placed in moisture analyser plate for water content estimation.

9.3 Results and Discussion

9.3.1 Selection of microalgal strains

The prime constituents of microalgal cell wall are polysaccharide and glycoprotein matrix acting as a defence against its environment (Gerken et al., 2013). Briefly, the microalgal cell wall contains wide variety of substances like chitin, protein, cellulose, murein, $CaCO_3$ and silica (D'Oca et al., 2011). Chemical complexity and structural robustness of algal cell wall varies across different species and hence requires different extraction techniques for efficient weakening of cell wall by liberation of cell wall mono or polysaccharides or by complete disruption of cell wall prior to lipid extraction. The optimal extraction technique and parameters should be valid for all microalgae and could tremendously reduce the amount of extraction solvent and energy input required for lipid extraction which will ultimately result in higher lipid extraction efficiencies. The microalgal species belonging to different taxonomic class and having a different cell wall structure were selected for evaluation and optimisation of cell disruption techniques.

C. reinhardtii: Unicellular, bi-flagellate, freshwater green alga (Chlorophyta, Chlorophyceae).

The cell wall of *C. reinhardtii* is a complex multi-layered extracellular coat composed of carbohydrates, insoluble hydroxyproline-rich proteins and several chaotrope soluble glycoproteins without cellulose or hemicellulose, arranged into at least seven distinct layers (Goodenough and Heuser, 1985, Imam et al., 1985, Monk, 1988, Schwede et al., 2011). The central domain is differentiated into three layers, the outer crystalline, a medial granular and an inner dense layer. Some of the cell wall component are fibrous whereas others are globular (Goodenough and Heuser, 1985). In addition to the hydroxyproline rich glycoproteins, the insoluble glycoprotein framework of the *C. reinhardtii* cell wall also contains minor amounts of 14-3-3 proteins which are involved in cross-linking of hydroxyproline rich glycoproteins (Voigt and Frank, 2003).

N. salina: Unicellular, marine green alga, (Heterokontophyta, Eustigmatophyceae).

Trilaminar outer cells walls (TLS) of *N. salina* contains insoluble, non-hydrolysable aliphatic bio-macromolecules termed as Algaenans, which exhibit very high resistance to enzymatic, chemical and bacterial degradation (Grossi et al., 2001, Schwede et al., 2011). Alkyl diols are the building blocks of algaenan usually linked to extractable or non-extractable polar lipids through ester and presumably amide and/or sugar and/or sulphate linkages. Similarly, unsaturated alcohols are also believed to be present, bound to the polar lipids (Grossi et al., 2001). Recently (Gerken et al., 2013) suggested additional presence of N-acetylglucosamine containing polymer in the cell wall.

D. salina: Unicellular, bi-flagellate, marine naked green alga (Chlorophyta, Chlorophyceae).

D. salina is morphologically similar to *Chlamydomonas* except it lacks a rigid polysaccharide cell wall and the cell is enclosed by a thin elastic plasma membrane covered by a mucus surface coat (Ben-Amotz and Avron, 1990, Borowitzka et al., 1984, Lynch and Thompson, 1982, Oren, 2005).

9.3.2 Selection of disruption techniques

A wide range of cell disruption techniques were reported in the literature for efficient extraction of lipids from microalgal and other biological samples. Among them, bead beating and sonication were the most popular cell disruption techniques frequently reported in the past literature. Several recent reports favour bead beating over sonication (Chen et al., 2011, Cheng et al., 2011, Halim et al., 2012, Ryckebosch et al., 2012, Serive et al., 2012, Tang et al., 2011) while the others favours sonication over bead beating (Adam et al., 2012, Araujo et al., 2013, Jaki et al., 2006, Kumari et al., 2011, Paik et al., 2009) as the best disruption technique. Each of the method has its own constraints and efficiency advantages. In contrast, some reports suggest combined use of both the techniques (Breuer et al., 2013) or combined use with the other cell disruption techniques like grinding and microwave (Lee et al., 2010, Šoštarič et al., 2012).

Henceforth, the second objective of our research was to assess the efficiency of bead beating and sonication techniques in extracting the lipids across two selected microalgal strains. For all the treatments, lipid extraction was carried out with 1.2 mL of methanol-chloroform mixture (1:2 *v/v*) as suggested elsewhere (Folch et al., 1957) whereas the extracted lipids were trans-esterified as suggested elsewhere (Vandenbrouck et al., 2010a). The amount of total lipids extracted and the number of FAMEs detected from both the species were used as an indication of the efficiency of the cell disruption technique. In an effort to optimize the suitable parameter combinations and to test the robustness of combining those parameters within each cells disruption technique for lipid extraction, we have employed design expert software to design the experiments. The screened factors evaluated within each cell disruption techniques were different and accordingly generated factorial designs were described in detail in section 9.2.3. Only FAMEs that were present in at least two biological replicates out of three were considered for further analysis. Furthermore, the microalgal cell integrity was studied using scanning electron microscopy (SEM), for untreated wet algal biomass, bead beating treated and sonicated biomass following extraction.

9.3.2.1 Bead beating experimental design and results

After thorough literature review as discussed in section 9.1.4.2, we have carefully selected the parameters and directed our approach towards using small diameter beads 425-600 µm i.d. and design of experiments with fixed milling time as 2 minutes but variable number of cycles (1–10 cycles) and relaxation time (1-2 minutes). In total ten different treatments with different parameters were applied for both the species (n=3), generating in total 30 samples for analyses for each species. The treatment details with parameters tested are listed in table 9.1.

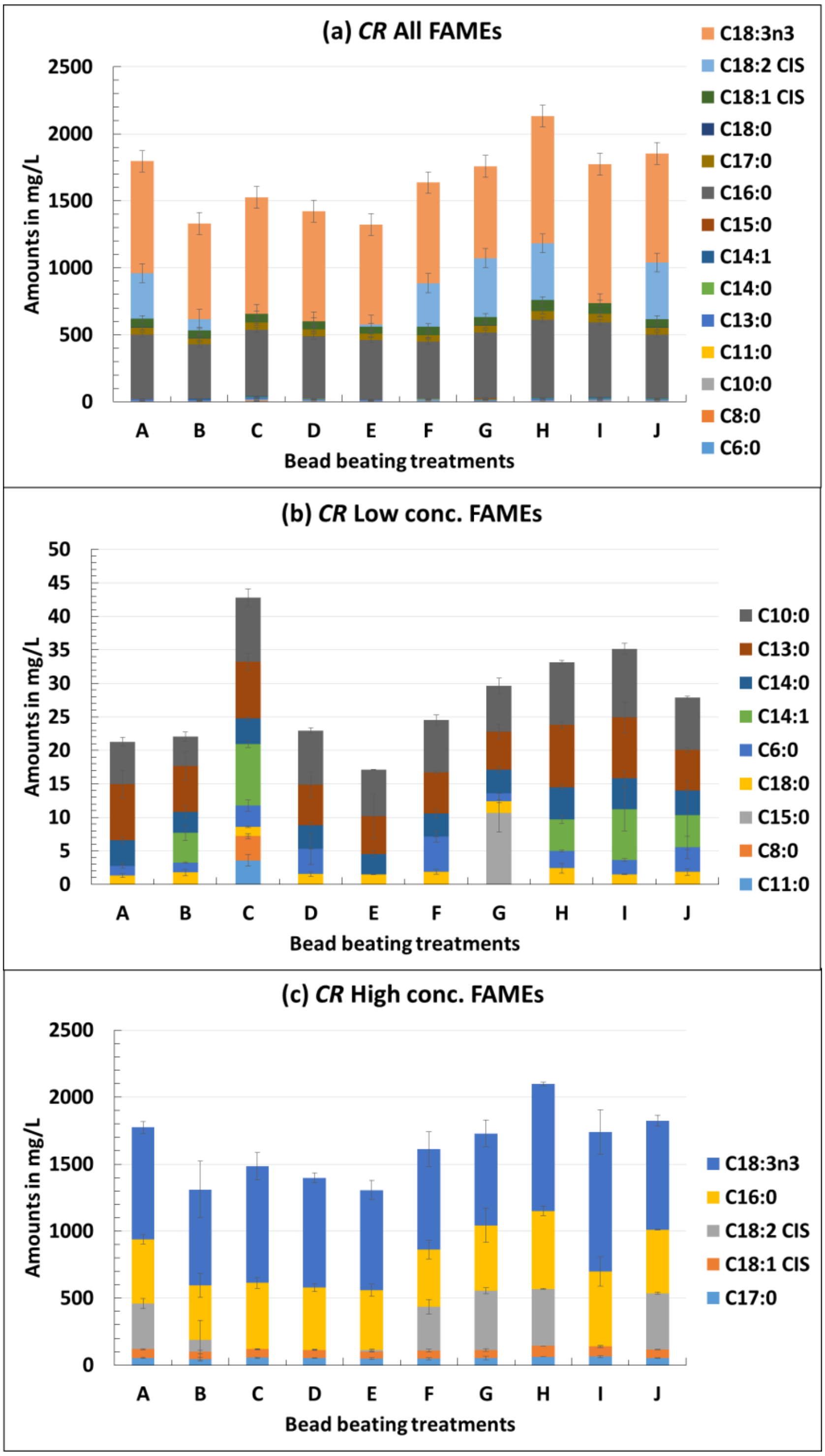


Figure 9.2 Effect of 10 different bead beating treatments on recovery of FAMEs was compared for *C. reinhardtii*. X-axis represents different applied treatments. Each treatment comprise of 2 factor combination generated by design expert software using 2 factorial design (Factor1: Cycle / Factor2: Time), where A = 1/1; B = 1/2; C = 10/1; D = 10/2; E = 3/1; F = 3/2; G = 5/1; H = 5/2; I = 7/1 and J = 7/2. Y-axis represents amount of FAMEs recovered in mg/L per DCW. After all the treatments the extracted lipids from cell extracts were analysed by GC-FID: (a) = Overall FAME recovery; (b) = recovery of low concentration FAMEs and (c) = recovery of high concentration FAMEs.

With the bead beating treatments in total 14 unique FAMEs were identified in *C. reinhardtii*. The results of investigation for overall coverage of FAMEs for *C. reinhardtii* are summarized in fig. 9.2 (a). Fig. 9.2 (b) summarises, 9 out of 14 FAMEs found to be present in lower concentration range (1-11 mg/L per DCW), whereas fig. 9.2 (c) indicates remaining 5 out of 14 FAMEs found to be present in higher concentration range (52-1040 mg/L per DCW). The recovery of low concentration FAMEs seems to be variable with the different bead beating treatments applied (fig. 9.2 (b)). Caprylic (C8:0) and undecanoic acid (C11:0) are only detected with treatment C, whereas pentadecanoic acid (C15:0) is detected only with treatment G. Myristic (C14:0) and stearic acid (C18:0) were detected with all the treatments with similar recovery, while capric (C10:0) and tridecanoic acid (C13:0) were also detected with all the treatments, however recovery was slightly less for capric acid (C10:0) with treatment B and recovery of tridecanoic acid (C13:0) was slightly higher with H and I treatments. The recovery of myristoleic acid (C14:1) was found to be selective with only few treatments like B, C, H, I and J. Overall treatment C revealed superior recovery of low concentration FAMEs followed by I and H treatments.

On the other hand the recovery of high concentration FAMEs was pretty much consistent with all the treatments applied (slightly higher recoveries obtained with treatment H), except for linoleic acid (C18:2 CIS) which was not detected with C, D, E and I treatments (fig. 9.2(c)). Overall superior recovery was obtained with treatment H for both low and high concentration FAMEs, exception to this was treatment C which showed superior recovery for low concentration FAMEs only.

In case of *N. salina*, similar trend was observed with recovery of low and high concentration FAMEs as that of *C. reinhardtii*. Briefly, in total 16 unique FAMEs were identified in *N. salina*. The results of investigation for overall coverage of FAMEs are summarized in fig. 9.3 (a). Fig. 9.3 (b) summarises the 10 out of 16 FAMEs found to be present in lower concentration range (1-31 mg/L per DCW), whereas fig. 9.3 (c) indicates remaining 6 out of 16 FAMEs found to be present in higher concentration range (53-704 mg/L per DCW).

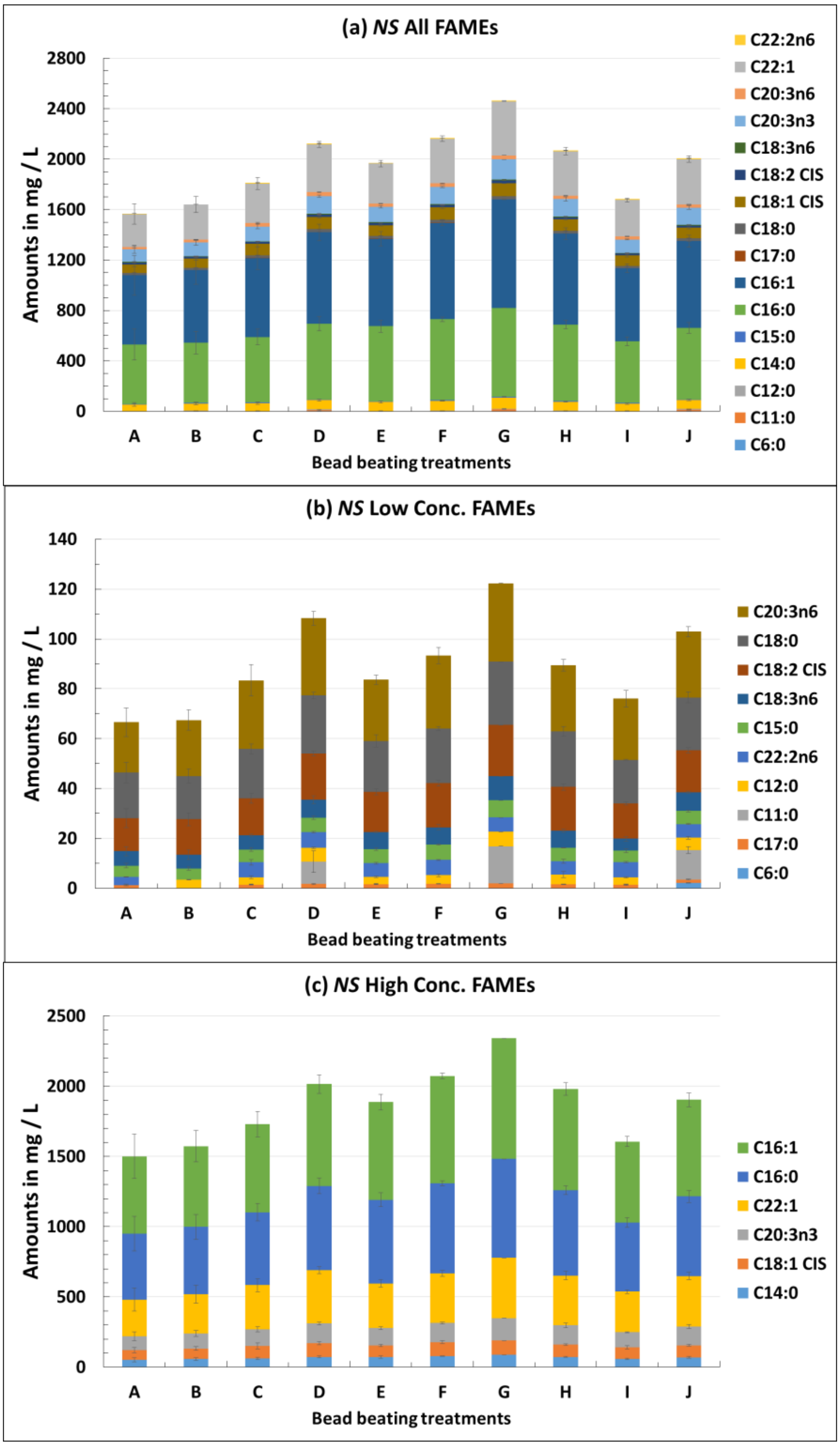


Figure 9.3 Effect of 10 different bead beating treatments on recovery of FAMEs was compared for *N. salina*. X-axis represents different applied treatments. Each treatment comprise of 2 factor combination generated by design expert software using 2 factorial design (Factor1: Cycle / Factor2: Time); A = 1/1; B = 1/2; C = 10/1; D = 10/2; E = 3/1; F = 3/2; G = 5/1; H = 5/2; I = 7/1 and J = 7/2. Y-axis represents amount of FAMEs recovered in mg/L per DCW. After all the treatments the extracted lipids from cell extracts were analysed by GC-FID: (a) = Overall FAME recovery; (b) = recovery of low concentration FAMEs and (c) = recovery of high concentration FAMEs.

Caproic acid (C6:0) was detected only with treatment J, whereas undecanoic acid (C11:0) was only detected with treatments D, G and J. Heptadecanoic acid (C17:0) and cis-13, 16-

docosadienoic acid (C22:2n6) were detected in all the treatments except with B, whereas lauric acid (C12:0) was detected in all except with the treatment A. The remaining FAMEs like pentadecanoic acid (C15:0), stearic acid (C18:0), linoleic acid (C18:2 CIS), g-Linolenic acid (C18:3n6) and cis-8, 11, 14-eicosatrienoic acid (C20:3n6) were recovered with negligible differences in their concentration with all the treatments. Overall treatment G revealed an superior recovery of low concentration FAMEs. On the other hand, all the high concentration FAMEs present in *N. salina,* showed superior recovery with all the applied treatments with slightly higher recovery with treatment G. Overall superior recovery of both low and high concentration FAMEs were obtained with treatment G. In total 8 unsaturated fatty acids were identified in *N. salina,* all of which showed superior recovery with all the treatments except for cis-13, 16-docosadienoic acid (C22:2n6) which was not identified with the treatment A and B.

9.3.2.2 Sonication experimental design and results

In this study, we have directed our approach to evaluate and optimize the sonication parameters and finally compare them to that of optimized bead beating parameters in terms of efficient extraction of lipids from two microalgal species. As mentioned above in section 9.2.2, we have employed design of experiment software, where we have used three factorial design, which involves time (factor 1, range 1-6 minutes), duty cycle (factor 2, range 10-60 %) and output (factor 3, range 1-6), which resulted in total number of 27 treatments (n=3), generating in total 81 samples for analyses for each species. The treatment details with parameters tested are listed in table 9.1.

With the sonication treatments in total 14 unique FAMEs were identified in *C. reinhardtii*. The results of investigation for overall coverage of FAMEs for *C. reinhardtii* are summarized in figure 9.4 (a). Figure 9.4 (b) summarises the 8 out of 14 FAMEs found to be present in lower concentration range (1.5-145 mg/L per DCW), whereas figure 9.4 (c) indicates remaining 6 out of 14 FAMEs found to be present in higher concentration range (47-1250 mg/L per DCW).

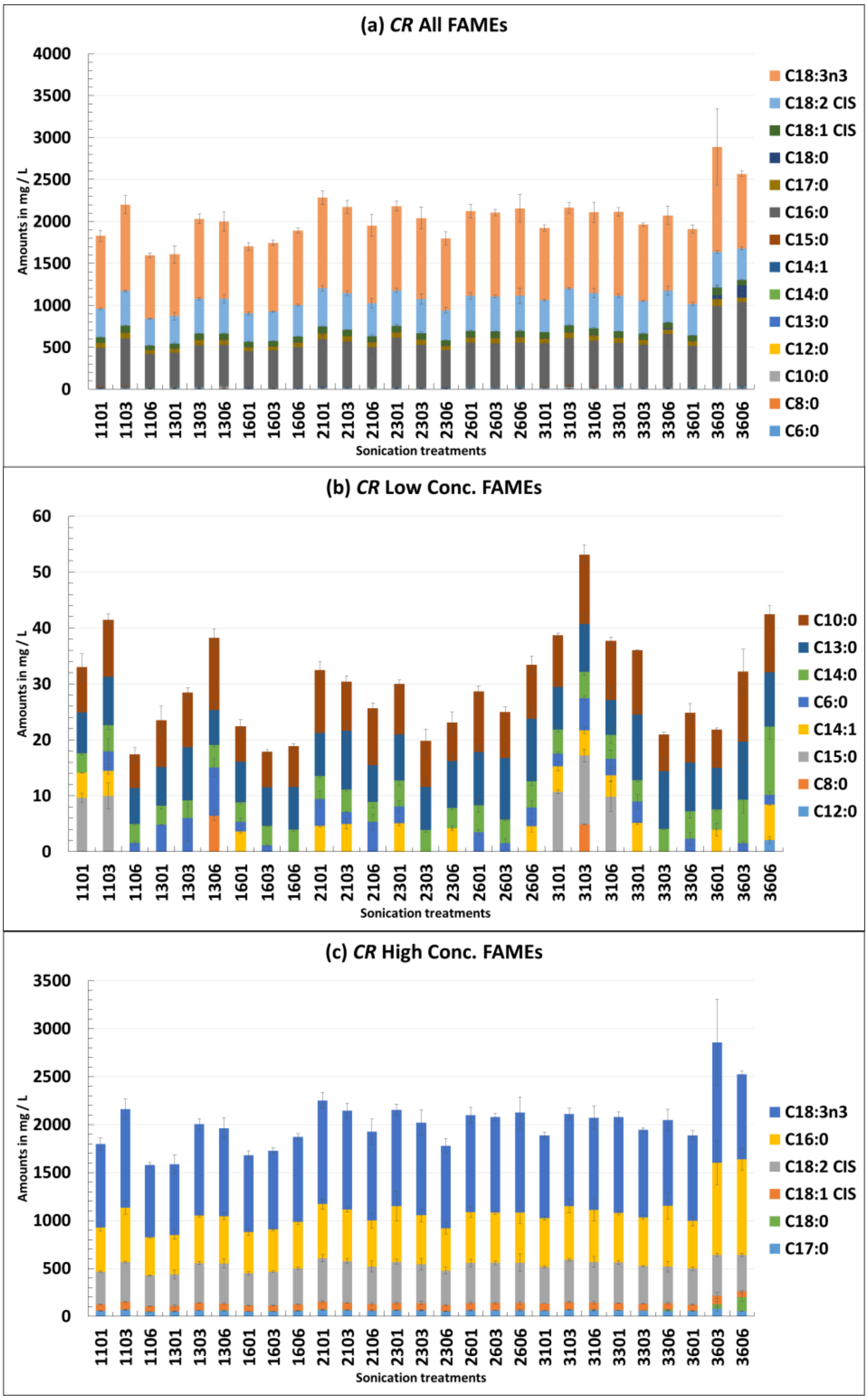


Figure 9.4 Effect of 27 different sonication treatments on recovery of FAMEs was compared for *C. reinhardtii*. X-axis represents different applied treatments. Each treatment comprise of 3 factor combination generated by design expert software using 3 factorial design (Factor1: Time / Factor2: Duty cycle / Factor3: Output power). In 1101, first 1 digit stands for time factor (range 1; 2 and 3 minutes); 10 digit stands for duty cycle factor (range 10; 20 and 30%) and last 1 digit stands for output factor (range 1; 3 and 6). Similar trend was followed for all the coding system for further 26 treatments; Y-axis represents amount of FAMEs recovered in mg/L per DCW. After all the treatments the extracted lipids from cell extracts were analysed by GC-FID: (a) = Overall FAME recovery; (b) = recovery of low concentration FAMEs and (c) = recovery of high concentration FAMEs.

The recovery of low concentration FAMEs seems to be variable with the different sonication treatments applied (figure 9.4 (b)). Lauric acid (C12:0) was only trapped with one treatment (3606) whereas caprylic acid (C8:0) was recovered only with two treatments (3103 and 1306(slightly higher)). On the other side recovery of caproic acid (C6:0), myristoleic acid (C14:1) and pentadecanoic acid (C15:0) were found to be selective with only few treatments. Briefly, caproic acid (C6:0) was unable to recover with 1606, 2303, 2306, 3303 and 3601 treatments, whereas it was recovered with all other treatments with highest recovery yielded with 1306 treatment. Similarly myristoleic acid (C14:1) and pentadecanoic acid (C15:0) were only captured with few treatments (figure 9.4 (b)) with highest recovery achieved with 3606 and 3103 treatments respectively. Whereas capric acid (C10:0), tridecanoic acid (C13:0), myristic acid (C14:0) were successfully recovered with all the applied treatments with highest recovery achieved with 1306, 3301 and 3606 respectively.

The recovery of high concentration FAMEs were pretty much consistent with all the treatments applied (slightly higher with treatment 3603 and 3606), except for stearic acid (C18:0) which was surprisingly not detected with any other applied treatments except for two (3603 and 3606) (figure 9.4 (c)). Overall recovery of FAMEs were superior with treatments 3603 and 3606, however the reproducibility was not too great as indicated by higher standard deviation values. In contrast to this, 3103 treatment showed better coverage for all low and high concentration FAMEs with superior reproducibility, exception to this was lauric acid (C12:0) which was not detected with this treatment.

In case of *N. salina*, similar trend was observed with recovery of low, medium and high concentration FAMEs as that of *C. reinhardtii*. Briefly, in total 20 unique FAMEs were identified in *N. salina*. The results of investigation for overall coverage of FAMEs are summarized in figure 9.5.

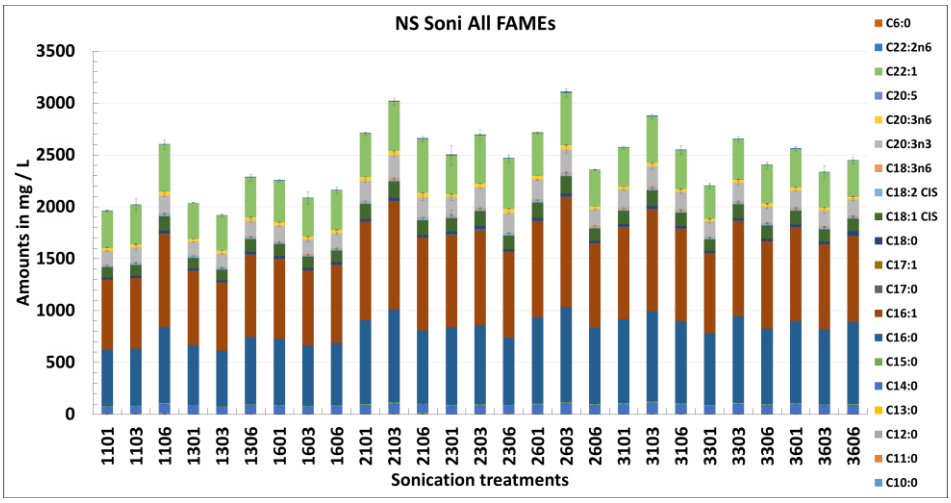
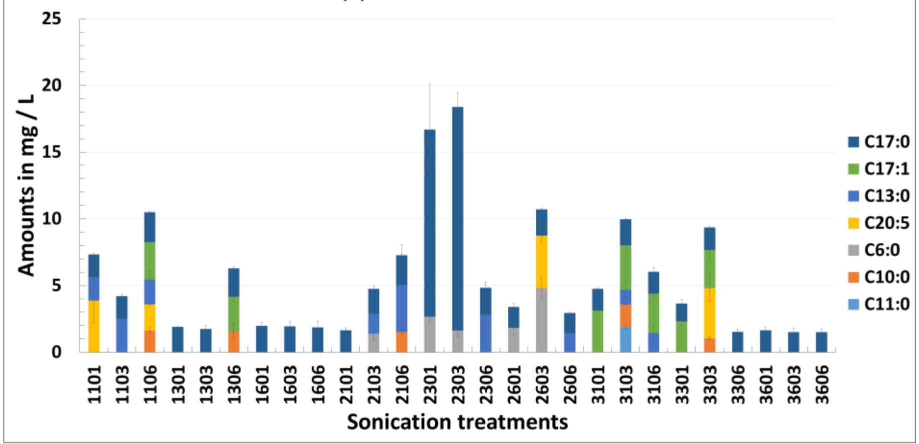


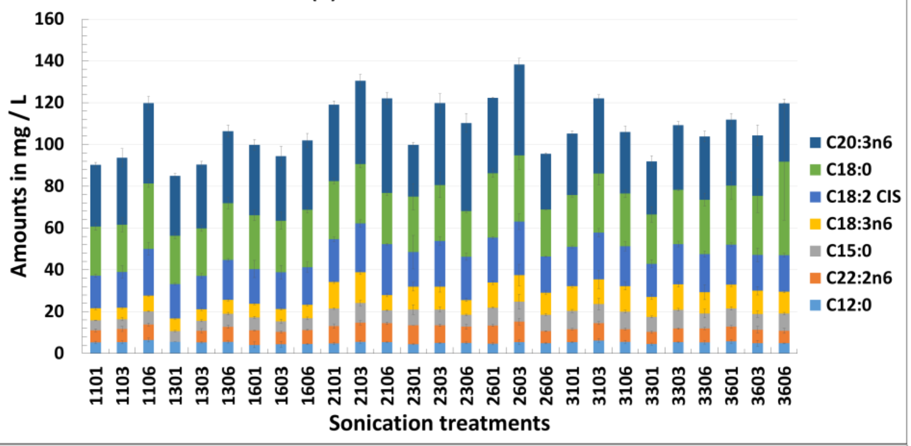
Figure 9.5 Effect of 27 different sonication treatments on recovery of all FAMEs was compared for *N. salina*. X-axis represents different applied treatments. Each treatment comprise of 3 factor combination generated by design expert software using 3 factorial design (Factor1: Time / Factor2: Duty cycle / Factor3: Output power). In 1101, first 1 digit stands for time factor (range 1; 2 and 3 minutes); 10 digit stands for duty cycle factor (range 10; 20 and 30%) and last 1 digit stands for output factor (range 1; 3 and 6). Similar trend was followed for all the coding system for further 26 treatments; Y-axis represents amount of FAMEs recovered in mg/L per DCW.

Figure 9.6 (a) summarises 7 out of 20 FAMEs found to be present in lower concentration range (2-17 mg/L per DCW), Figure 9.6 (b) summarises other 7 out of 20 FAMEs found to be present in medium concentration range (4-45 mg/L per DCW) whereas figure 9.6 (c) indicates remaining 6 out of 20 FAMEs found to be present in higher concentration range (65 - 1060 mg/L per DCW).

(a) *NS* Low Conc. FAMEs



(b) *NS* Medium Conc. FAMEs



(c) *NS* High Conc. FAMEs

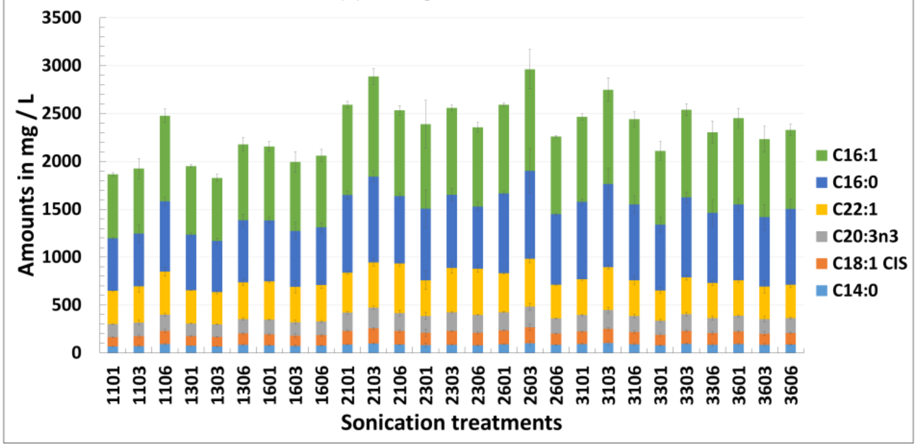


Figure 9.6 10 Effect of 27 different sonication treatments on recovery of FAMEs was compared for *N. salina*. X-axis represents different applied treatments. Each treatment comprise of 3 factor combination generated by design expert software using 3 factorial design (Factor1: Time / Factor2: Duty cycle / Factor3: Output power). In 1101, first 1 digit stands for time factor (range 1; 2 and 3 minutes); 10 digit stands for duty cycle factor (range 10; 20 and 30%) and last 1 digit stands for output factor (range 1; 3 and 6). Similar trend was followed for all the coding system for further 26 treatments; Y-axis represents amount of FAMEs recovered in mg/L per DCW. After all the treatments the extracted lipids from cell extracts were analysed by GC-FID: (a) = Recovery of low concentration FAMEs; (b) = recovery of medium concentration FAMEs and (c) = recovery of high concentration FAMEs.

The recovery of low concentration FAMEs were found to be extremely selective with the different sonication treatments applied (figure 9.6 (a)). Undecanoic acid (C11:0) was found to be detected with only 3103 treatment. The recoveries of caproic acid (C6:0), capric acid (C10:0), tridecanoic acid (C13:0), cis-10-heptadecenoic acid (C17:1) and cis-5, 8, 11, 14, 17-eicosapentaenoic acid (C20:5n3) were found to be variable with selective treatments as can be seen in figure 9.6 (a), whereas heptadecanoic acid (C17:0) was recovered with all the applied treatments however the recovery was found to be extremely high with 2301 and 2303 treatments. Similarly recovery of tridecanoic acid (C13:0) with the 1206 and 2306 treatments was higher but with very poor reproducibility. Only two treatments namely 1106 and 3103 seem to have recovered most of the low concentration FAMEs (5 out of 7) with good reproducibility with respect to overall coverage.

Negligible variance was found in recovery of medium and high concentration FAMEs with all the applied treatments. The superior recovery was achieved with 2603 and 2103 treatment for both the medium and high concentration FAMEs, however recovery of low concentration FAMEs was not too great, on the other hand treatment 3103 also showed superior recovery (slightly less than that of 2603 and 2103) for all low, medium and high concentration FAMEs with good reproducibility results.

9.3.2.3 Comparison between two cell disruption techniques

In case of *C. reinhardtii,* both the cell disruption techniques recovered equal number of overall FAMEs (14), whereas in case of *N. salina*, bead beating recovered 16 while sonication treatments recovered 20 FAMEs. In *C. reinhardtii*, similar number of low concentration FAMEs were identified with both the techniques, however concentration of individual FAMEs recovered was found to be slightly higher with the sonication treatment (3103) compared to bead beating treatment (C), whereas in *N. salina*, similar trend was observed, where slightly higher recovery of low/medium concentration FAMEs were recorded with sonication parameters compared to bead beating treatments. Moreover it was interesting to note that, sonication treatments have recovered additional five low concentration FAMEs which were not detected with the bead beating parameters.

In case of high concentration FAMEs, both techniques recovered similar number of FAMEs with approximately similar recovery of individual FAMEs concentration in *C. reinhardtii*, however the case was not the same for *N. salina,* where both the treatments detected same numbers of FAMEs, however concentration of individual FAMEs recovered varied greatly. Sonication treatments (2103, 2603, and 3103) showed much higher recovery in terms of individual FAME concentration as compared to the bead beating treatments (G) as can be seen in figure 9.6 (c) and 9.2 (c). Overall, for both the species, sonication recovered slightly higher concentration for all the identified FAMEs compared to that of bead beating treatments.

The results of % total fatty acid content per DCW obtained for both the species with all the treatments applied within both the cell disruption techniques are summarised in figure 9.7

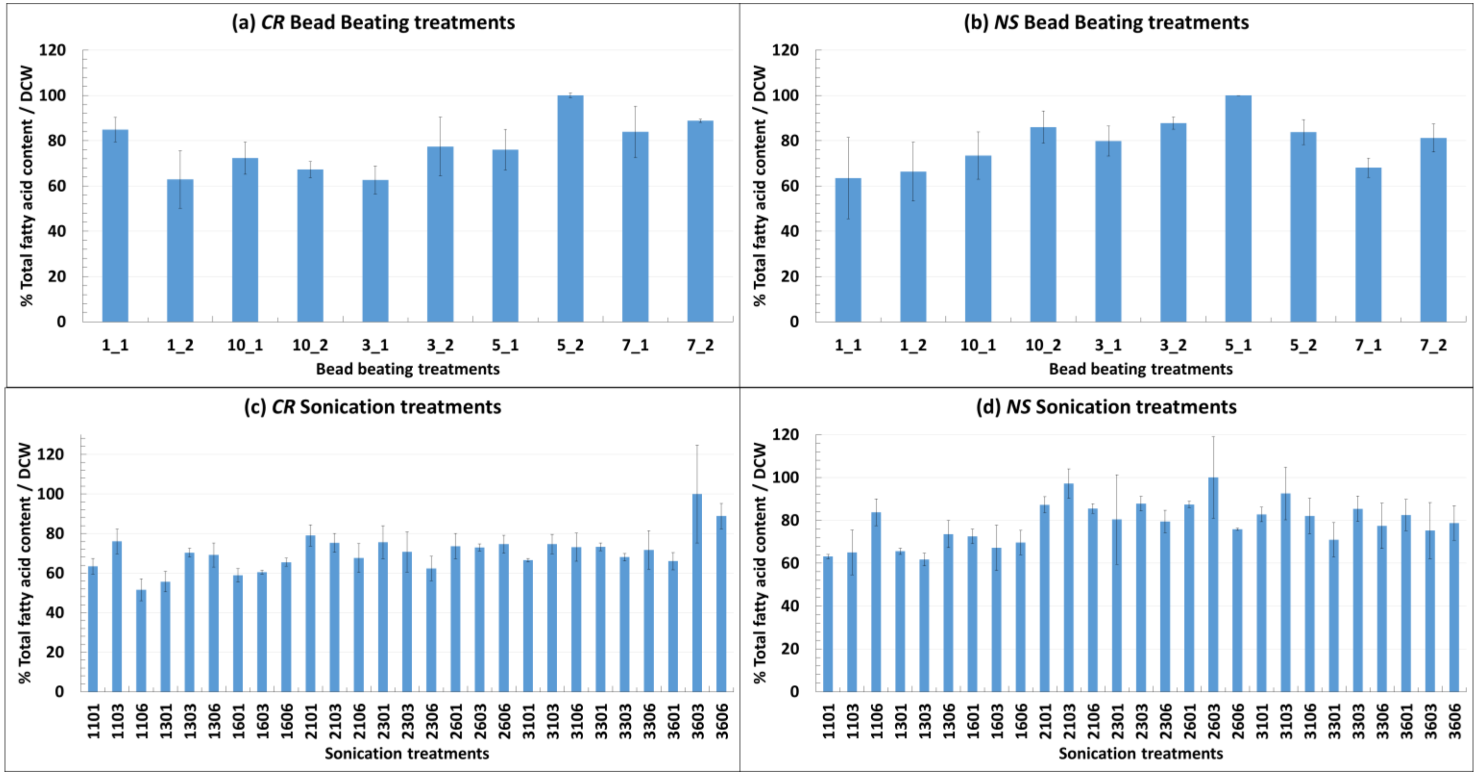


Figure 9.7 Effect of bead beating and sonication treatments on recovery % total FA content per dry cell weight was compared for *C. reinhardtii* and *N. salina*. X-axis in (a) and (b) represents different applied treatments. Each treatment comprise of 2 factor combination generated by design expert software using 2 factorial design (Factor1: Cycle / Factor2: Time); A = 1/1; B = 1/2; C = 10/1; D = 10/2; E = 3/1; F = 3/2; G = 5/1; H = 5/2; I = 7/1 and J = 7/2, whereas X-axis in (c) and (d) represents different applied treatments. Each treatment comprise of 3 factor combination generated by design expert software using 3 factorial design (Factor1: Time / Factor2: Duty cycle / Factor3: Output power). In 1101, 1 digit stands for time factor (range 1; 2 and 3 minutes); 10 digit stands for duty cycle factor (range 10; 20 and 30%) and last 1 digit stands for output factor (range 1; 3 and 6). Similar trend was followed for all the coding system for further 26 treatments; Y-axis represents % total FA content per DCW. After all the treatment the extracted lipids from cell extracts were analysed by GC-FID: (a) = % total FA content per DCW by bead beating treatments in *C. reinhardtii*; (b) = % total FA content per DCW by bead beating treatments in *N. salina*; (c) = % total FA content per DCW by sonication treatments in *C. reinhardtii* and (d) = % total FA content per DCW by sonication treatments in *N. salina*.

Within bead beating technique, as compared to other treatments, the recovery was higher (100%) with 5_2 and 5_1 parameters in *C. reinhardtii* and *N. salina* respectively. It is important to note that for both the species, 5 cycles of bead beating with fixed milling time of 2 minutes for each cycle was found to be the common factor which resulted in highest recovery total fatty acids. Within sonication technique, as compared to other treatments, the recovery was higher (100%) with 3603 and 2603 parameters in *C. reinhardtii* and *N. salina* respectively. However, with both the parameters the reproducibility was really poor as showed by large error bars (figure 11 (c) and 11 (d)). Speaking with regards to the coverage of both low and high concentration FAMEs with good reproducibility, 3103 seems to be the most suitable combination of parameters resulting in highest number of FAMEs recovered, 12 out of 14 FAMEs in *C. reinhardtii* and 18 out of 20 FAMEs in *N. salina* with an superior reproducibility (figure 9.4 (b) and (c) and figure 9.6 (b) and (c).

Overall it can be seen that with the optimised parameters both the techniques showed superior recoveries of FAMEs and total fatty acid content. Stearic acid (C18:0) is considered as an important candidate for biodiesel production and is found to be present in smaller concentration in case of *C. reinhardtii.* Despite higher concentration of FAMEs recovered with sonication treatments compared to bead beating treatments, loss of stearic acid (C18:0) with all the sonication treatments (except 3603 and 3606) in *C. reinhardtii*, was considered as an inappropriate solution for biodiesel production. In contrast, all the bead beating treatments recovered equal concentrations of stearic acid (C18:0). Moreover at laboratory scale, with sonication only a single sample can be processed at one time whereas with bead beating technique, twelve samples can be processed simultaneously. As mentioned earlier, for microalgal biodiesel production, an ideal disruption technique should be one, which can selectively extract the lipid classes of interest without causing any form of damage to them. In the context of this scenario, bead beating was selected as a suitable disruption technique and optimised parameters within this technique were further evaluated by SEM analysis.

9.3.2.4 Scanning electron microscopy (SEM) analysis on disrupted algal biomass

Furthermore, the microalgal cell integrity was studied with SEM analysis for all the three microalgal species under investigation. The above optimised parameters within each cell disruption technique with the two microalgal species were adopted and applied for all the three microalgal species under investigation for SEM analysis. The microalgal cell integrity was studied on samples of untreated wet algal biomass (as a control), bead beating treated algae and sonication treated algae. From both the cell disruption techniques, two parameters were selected for SEM analysis, one which resulted in highest and the other which resulted in lowest recovery of FAMEs. From bead beating technique, 5_2 (higher recovery) and 1_2 (lower recovery) treatments were selected, whereas from sonication technique, 3103 (higher recovery) and 1106 (lower recovery) treatments were selected for SEM analysis. The results of SEM analysis for *C. reinhardtii, D. salina and N. salina* are summarised in fig. 9.8, 9.9 and 9.10 respectively. From the SEM observation of untreated wet algal biomass, bead beaten and sonicated cells from all the three species, it was found that with 5_2 and 3103 treatments, cell wall structure was totally disturbed and fragmented, leading to formation of homogenous cell paste. However, SEM observations of cells treated with parameters 1_2 and 1106, clearly suggests that there was only partial disruption of the cells, leaving some part of the cell integrity intact. All the SEM observations correlates and support our above finding where 5_2 and 3103 treatments were resulted in higher total FAME recovery compared to that 1_2 and 1106 treatments which recovered less number of FAMEs.

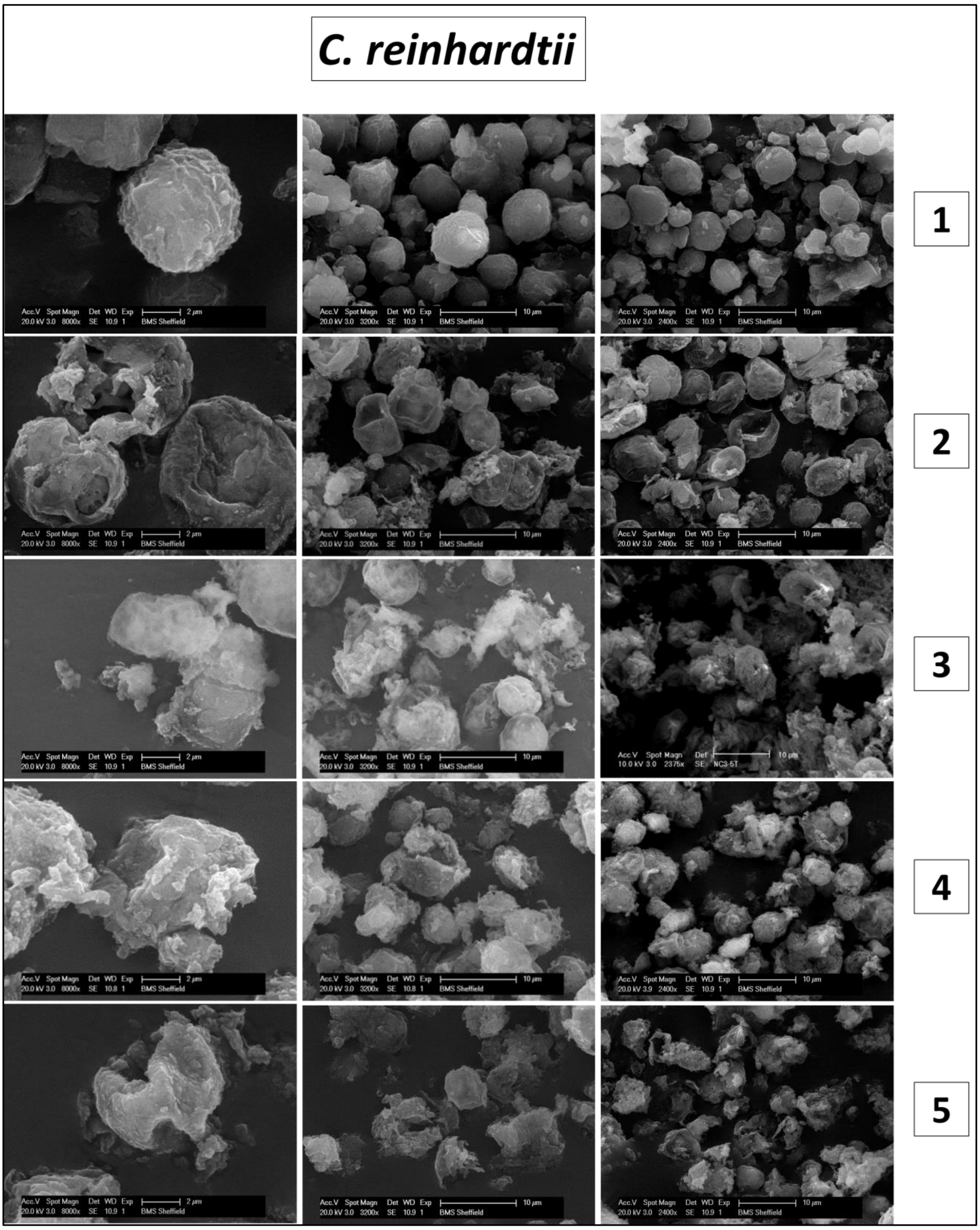


Figure 9.8 The cell disruption efficiency of bead beating and sonication treatments on *C. reinhardtii* cell integrity was visualised and compared using SEM observations, where: 1 = untreated wet algal biomass; 2 = bead beating 1_2 treatment with lowest FAME recovery; 3 = bead beating 5_2 treatment with highest FAME recovery; 4 = sonication 3103 treatment with highest FAME recovery and 5 = sonication 1106 with lowest FAME recovery.

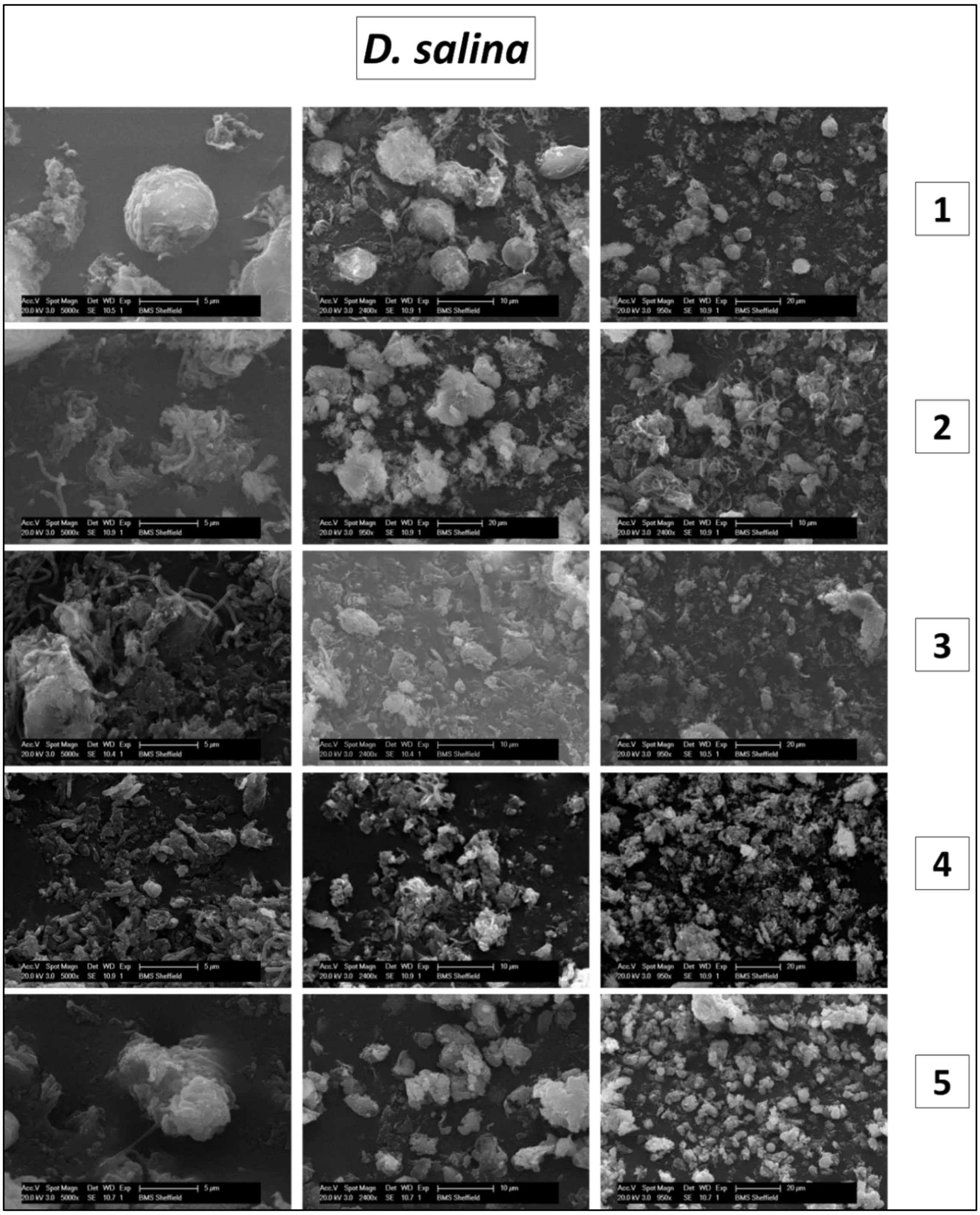


Figure 9.9 The cell disruption efficiency of bead beating and sonication treatments on *D. salina* cell integrity was visualised and compared using SEM observations, where: 1 = untreated wet algal biomass; 2 = bead beating 1_2 treatment with lowest FAME recovery; 3 = bead beating 5_2 treatment with highest FAME recovery; 4 = sonication 3103 treatment with highest FAME recovery and 5 = sonication 1106 with lowest FAME recovery.

N. salina

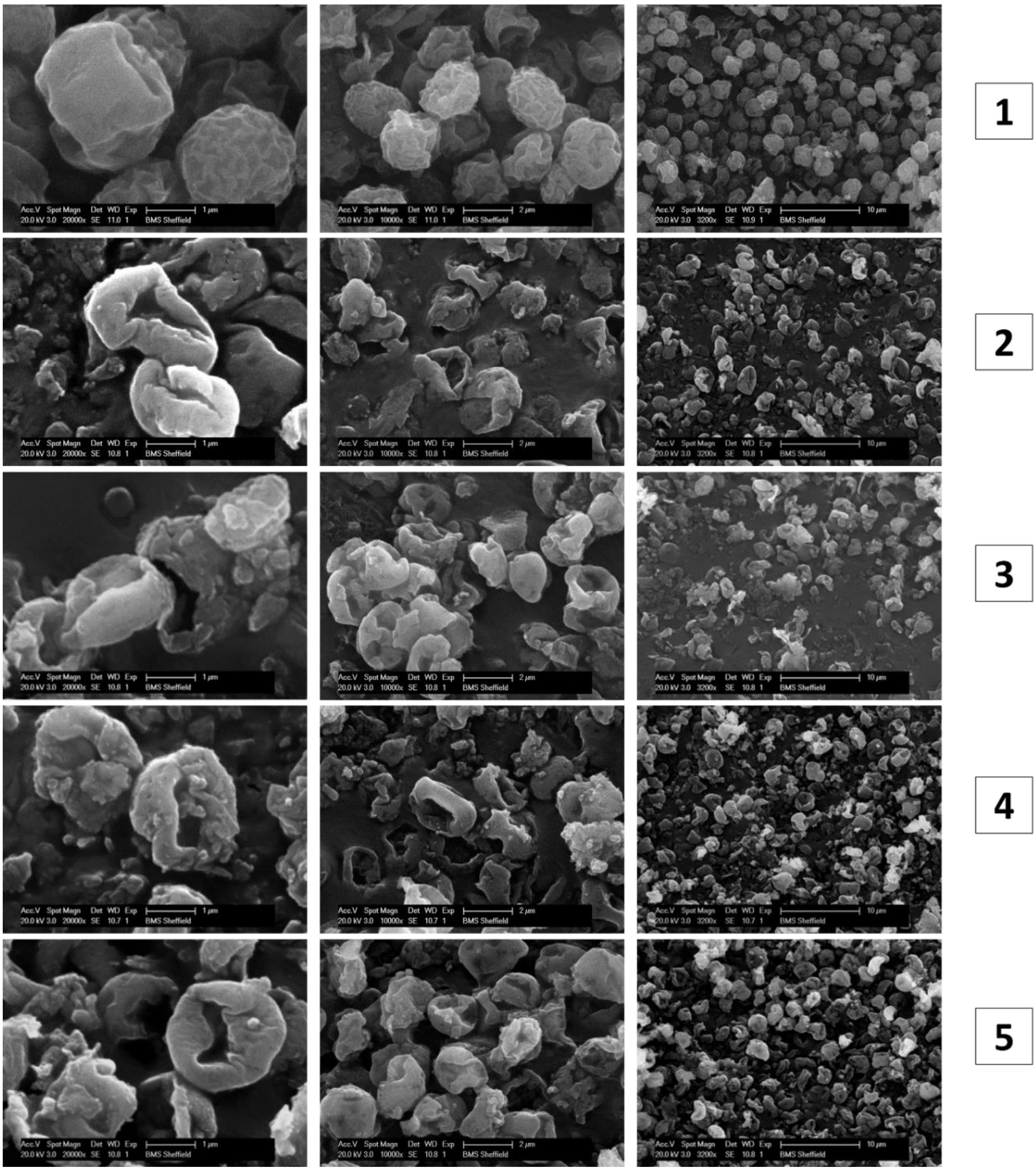


Figure 9.10 The cell disruption efficiency of bead beating and sonication treatments on *N. salina* cell integrity was visualised and compared using SEM observations, where: 1 = untreated wet algal biomass; 2 = bead beating 1_2 treatment with lowest FAME recovery; 3 = bead beating 5_2 treatment with highest FAME recovery; 4 = sonication 3103 treatment with highest FAME recovery and 5 = sonication 1106 with lowest FAME recovery.

9.3.3 Optimization of GC method and parameters

On the instrument side, the reliable and accurate quantification of fatty acids primarily depends on optimised GC parameters and calculation of identified FAMEs concentration based on internal and external standards. It is most important to make sure that the instrument is properly calibrated for every single sample analysis and analyst must ensure and have knowledge about after how many number of samples instruments needs further calibration with external/internal standard to ensure reliable and accurate quantification or else there would be risk of obtaining erratic and inaccurate conclusions from results.

To ensure reliable identification and quantification of FAMEs from microalgal samples, within each sequence, the external standard (37 FAME mix) was run before and at the end of every sequence as a calibration and check standard respectively. The results of check standards are then compared to the concentration of the calibration standard to evaluate the validity of calibration and amount calculation. Initially, we ran 24 algal sample (sequence 1) and later 72 samples (sequence 2) and the results of check standards were compared to the calibration standard within each sequence and calibration plot was generated for randomly selected 5 components from 37 FAME mix. The results of comparison between sequence 1 and 2 for which retention time deviation plot and calibration curve plots were generated are summarised in figure 9.11;

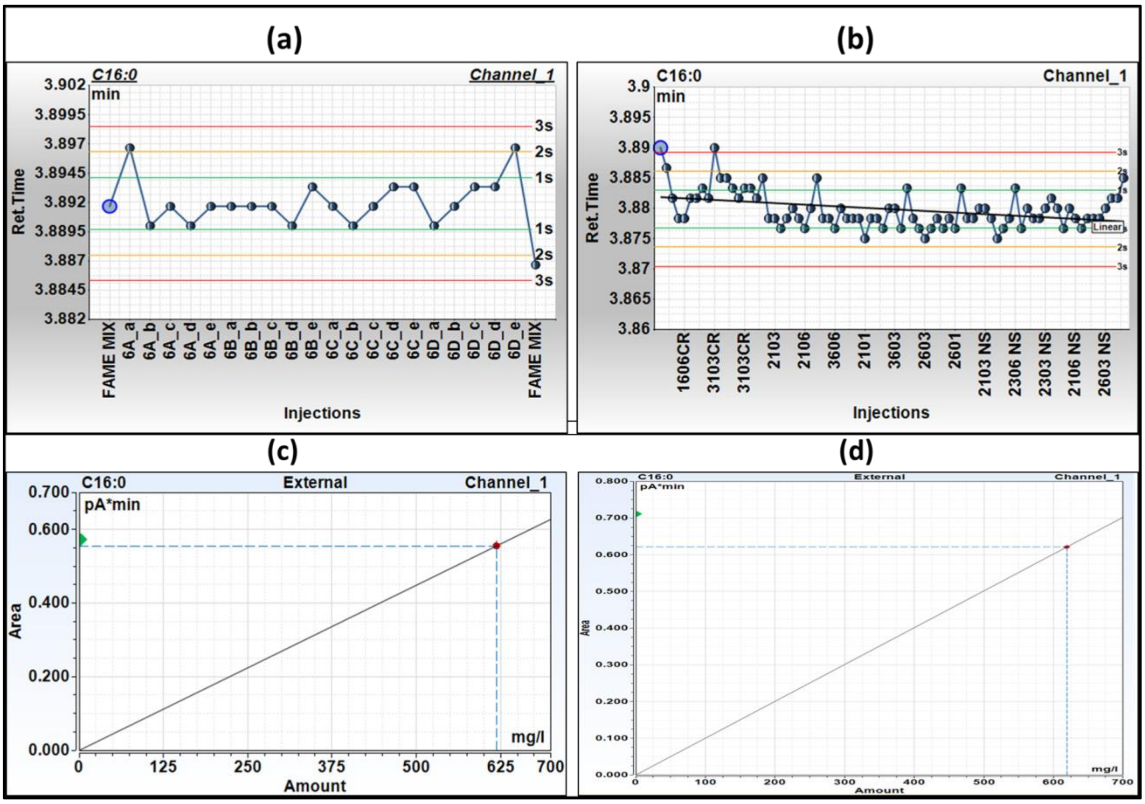


Figure 9.11 Retention time deviations for sequence 1 and 2 are summarised in figure (a) and (b) respectively; whereas calibration curve plots generated for sequence 1 and 2 are displayed in figure (c) and (d)n respectively. Both the plots were generated for C16:0 FAME from *N. salina*, where in sequence 1, 24 samples were ran wheras in sequence 2, 72 samples were ran in total as can be seens in circular froms in figure (a) and (b). The first circle represents calibration standard run; then all samples and the last circle represents the check standard run. The green triangular sign in the calibration curve plots indicate check standard run in both the sequences.

As can be seen from figure 9.11 (a) and (b) there was a negligible retention time deviation between the calibration and check standard (2s to 3s) for both the sequences. Similarly, there was negligible variation between calibration and check standard for sequence 1 where 24 samples were ran (figure 9.11 (c)). On the other hand, huge variation as indicated by green triangle (sign for check standard, ran at the end of the sequence) was observed in calibration curve plot for sequence 2 where 72 samples were ran (figure (d)). This clearly implies that results from sequence 2 could be highly erratic and inaccurate with respect to individual FAME concentration obtained and conclusions drawn from it could be seriously misleading. Henceforth, it is mandatory to evaluate the validity of the calibration and amount calculation from external/internal standard. We strongly recommend running calibration and check standard within each sequence having no more than 20 samples run within each sequence, in order to obtain accurate and reliable data set.

On the other hand, GC is well known for quantifying individual fatty acids as well as total fatty acid content in a very short analysis time. Despite this fact, the current analysis time required for FAME profiling was around 35 to 90 minutes per sample, not suitable for research studies which needs hundreds of samples analysed each day. Considering this, we optimised our GC parameters with 37 FAME mix and sample analysis time was reduced drastically to 12 minutes with good resolution, where normal analysis time reported in recent publications (Griffiths et al., 2010, Kumari et al., 2011, Laurens et al., 2012a, Laurens et al., 2012b, Slocombe et al., 2013, Xu et al., 2010) was between 35 to 90 minutes. Method details are described in detail in section 9.2.8. Excellent resolution for all 37 FAMEs were obtained within a very short chromatographic run. The representative chromatogram for 37 FAME mix with total analysis time of 9 minutes was shown in figure 9.12.

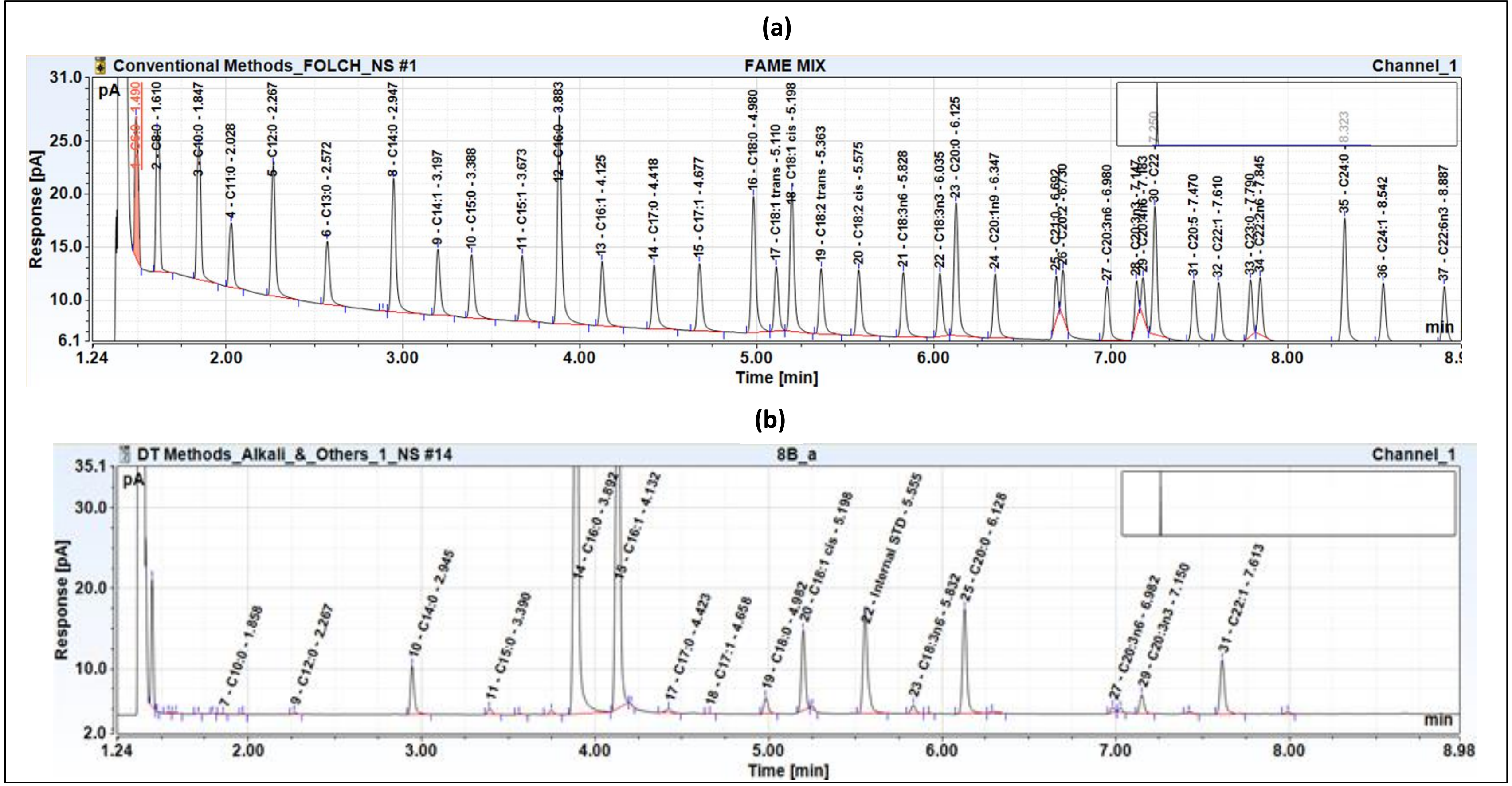


Figure 9.12 Representative GC-FID chromatogram of 37 FAME Mix (C8-C24) standard from Supelco (a) and of *N. salina* sample (b). The *N. salina* cells were extracted with toluene and directly transesterified with 0.5M sodium methoxide (alkali catalysts) and later with 5% acetyl chloride in methanol (acid catalyst). The trans-esterified sample volume of 1 µL was injected in split mode at 250°C and separated on a TR-FAME capillary column (25 m x 0.32 mm x 0.25 µm).

9.3.4 Conventional methods of 2 step trans-esterification

Effect of solvent and solvent ratios

The use of non-chlorinated solvents have potential to be an alternative to the more toxic chloroform in recovery of lipids from different sample matrices, however there were mixed conclusions from various authors. One possible reason might be that, evaluation performed on different sample matrices or in some cases only limited solvent systems were considered for evaluation, leaving more potentially useful ones behind. To minimise discrepancies and for a clear picture while selection of appropriate solvent system for microalgal lipid extraction, we have carefully considered the optimal extraction solvent systems and their ratios and compared them within and with the conventional solvent systems. The solvents systems and their ratios include methanol-chloroform (2:1), methanol-chloroform (1:1), propane-2-ol-cyclohexane-water (8:10:11), methanol-MTBE (1.5:5) and methanol-DCM (1:2) which were evaluated across three different microalgal species (*C. reinhardtii, N. salina and D. salina*).

Effect of pre-treatments

The FAME composition of *Schizochytrium limacinum SR21* from the wet and lyophilised algal biomass was found to be the same as investigated by Johnson using both extraction-TE and DT methods, however total FAME yield from the wet biomass was found to be less than that from the dry biomass (Johnson and Wen, 2009). In addition, (Ryckebosch et al., 2012) regarded lyophilisation as an advantageous process, as it improves the lipid extraction by increasing the surface area of the sample, makes weighing simpler and does not require special storage conditions afterwards. However as previously reported, lyophilisation also had many adverse effect on recovery of short chain FAs, as freezing and thawing has higher hydrolysis rate than the heating process for TGs, by causing disruption of fat globules (Lepage and Roy, 1984). Further a recent study (Griffiths et al., 2010) suggests that water content up to 10% of the reaction volume resulted in no adverse effect on recovery of FAMEs, which was in agreement with another report (Lepage and Roy, 1984). Another study (O'fallon et al., 2007) concluded that the water content up to 33% has no adverse effect on recovery of FAMEs from fish oil, henceforth water is a part of the FAME production process and not antagonist to it. Another report also arrived on similar conclusion (Patil et al., 2012), where author used microwave for both extraction-TE and DT methods and stated that residual water in wet algae becomes an excellent organic solvent at near-critical conditions and is non-toxic.

Another issue to consider is the addition of antioxidants to prevent lipid oxidation during sample treatment. As is well known, algae itself contains large amount of natural antioxidants like carotenoids, tocopherol and polyphenols, henceforth there is no need to add any synthetic antioxidant (Goh et al., 2010, Hu et al., 2008, Li et al., 2007, Murthy et al., 2005). Recently (Ryckebosch et al., 2012) evaluated the effect of addition of a synthetic antioxidant on FAME recovery of exogenous EPA added to the microalgal extract and confirmed that addition of antioxidant is not really required for lipid analysis. In case of DT method, addition of antioxidant is also not required to protect oxidation of MUFA and PUFA, as they were not prone to oxidation once methylated. In contrast, use antioxidant like BHT or 2, 2-dimethoxypropane resulted in an additional problems like an extra peak in chromatograms (Griffiths et al., 2010, Lepage and Roy, 1984).

Based on the above observations, we evaluated all our methods on wet algal biomass and skipped the process of lyophilisation and addition of a synthetic antioxidant. Moreover, these pre-treatments are more time consuming and will incur more cost for industrial scale application, and potentially introduce undesired artefacts in the analytical protocol.

Gravimetric analysis

The total lipid contents in mg/DCW were estimated with gravimetric analysis with five conventional methods (two step TE) across three microalgal species are summarised in figure 9.13. Overall, for all the species highest yield was obtained with the classical Folch (2B and 2D) method compared to the BD method. These findings were in agreement with (Kumari et al., 2011). The total lipid contents are largely affected by the solvent system, solvent ratio and disruption techniques employed for the extraction, as demonstrated by huge variation in recovery of total lipids between conventional (BD and Folch) and method 3, 4 and 5. However, it is important to note that, the gravimetric determination are highly variable and inaccurate for total lipid estimations as along with lipids, co-extraction of non-lipid material (like proteins, pigments, glycerol, phosphates, polyphenols etc…) contribute to the total weight and might lead to variable interpretations from different researchers.

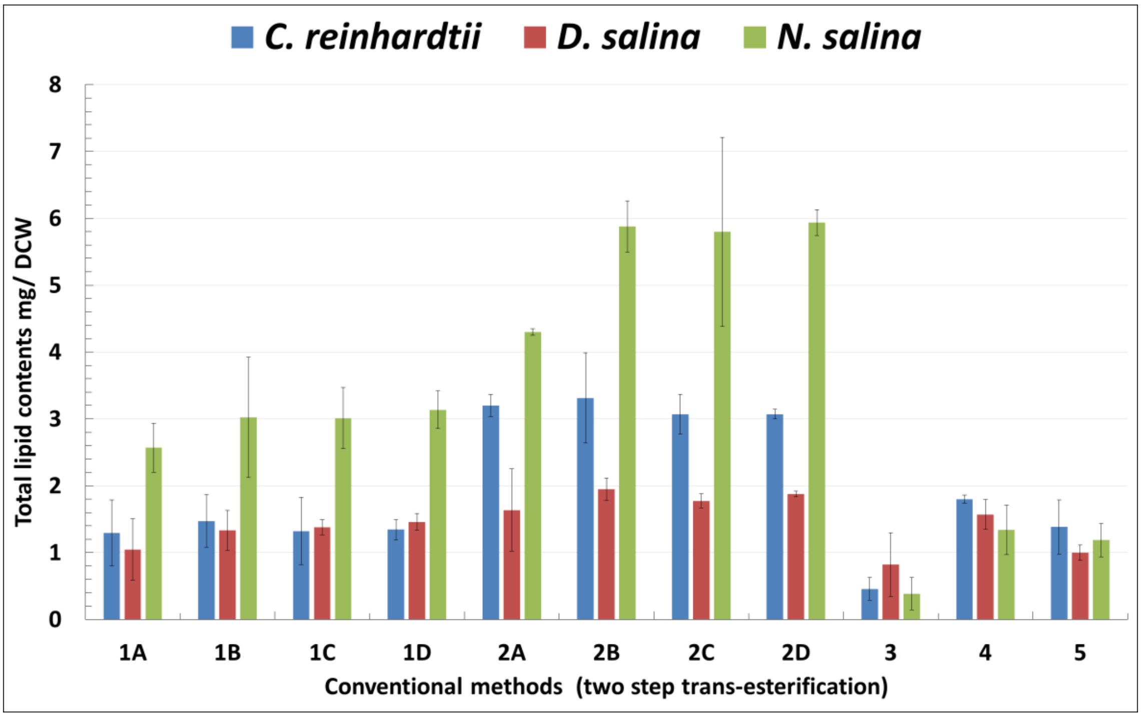
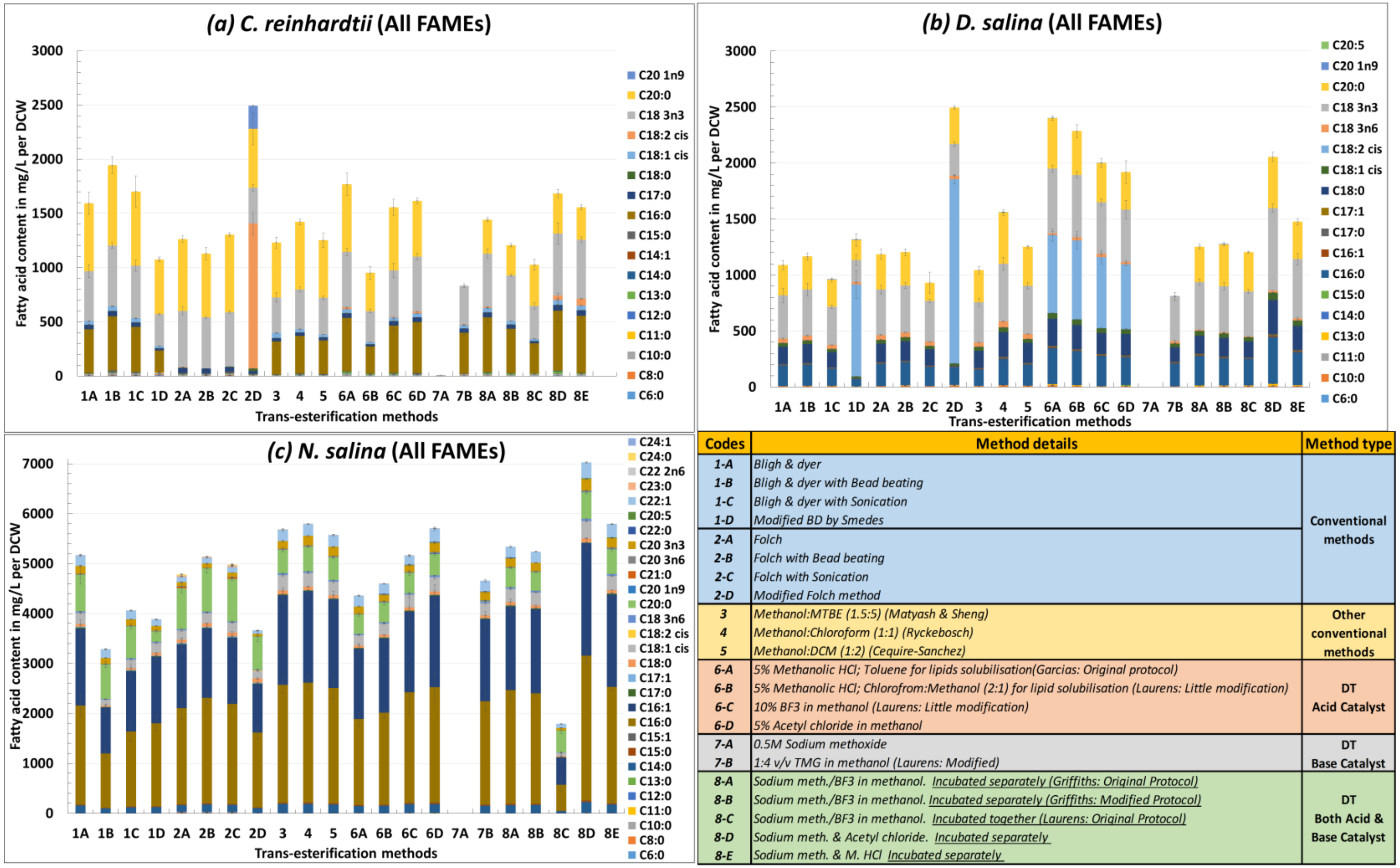


Figure 9.13 Total lipid contents estimated by conventional methods using gravimetric determination were expressed in mg/DCW of biomass in *C. reinhardtii*, *D. salina* and *N. salina*. The details of codes from 1A to 5 are belonging to conventional methods the details of which are described in Table 9.2.

Comparison of conventional methods

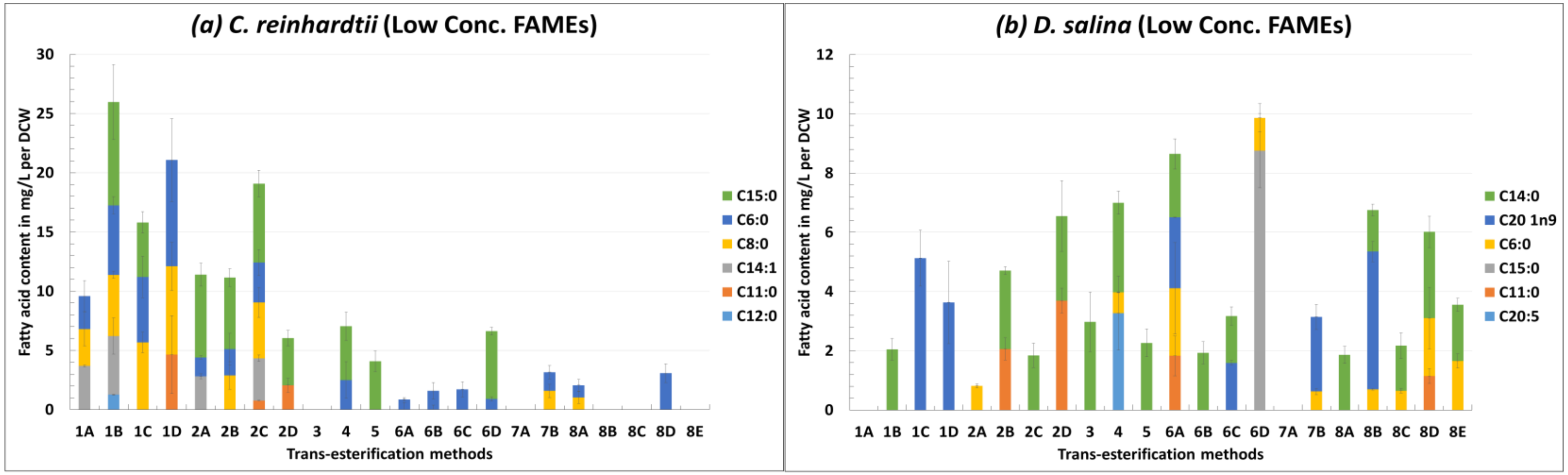
The individual total fatty acid content values were reported in mg/L per DCW. While comparing all the method across all the three species, equal volume of cultures were harvested and the volume of extraction solvents and trans-esterification reagents were kept constant for all the methods to avoid any variations in the recovery of FAMEs. In addition, after phase separation, the volume of organic phase containing FAMEs were measured prior to further processing and the volumes were kept constant across for all the methods to avoid further variations. Only FAMEs that were present in at least three biological replicates out of five were considered for further analysis.

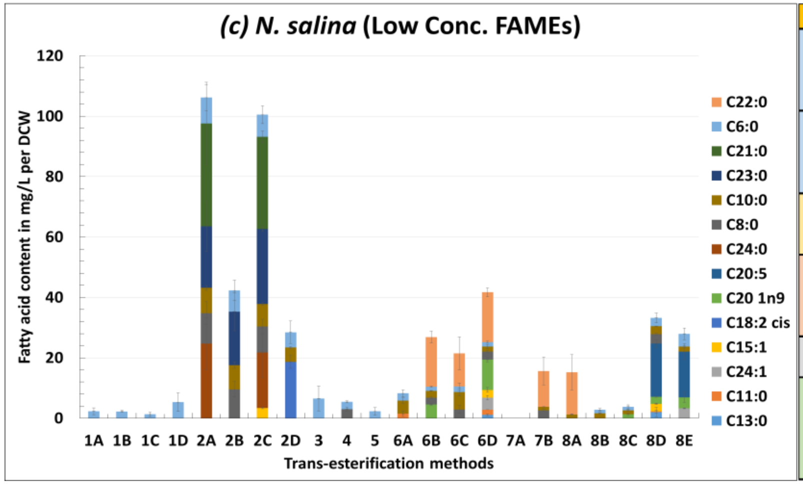
The results of conventional methods like Bligh and Dyer method (1A) with its modification (1B,1C and 1D) and Folch method (2A) with its modifications (2B, 2C and 2D) were compared within and with the recently suggested 2 step trans-esterification methods as suggested by Matyash and sheng (Matyash et al., 2008, Sheng et al., 2011)(Method 3) , Ryckebosch (Ryckebosch et al., 2012) (Method 4) and Cequire-sanchez (Cequier-Sánchez et al., 2008) (Method 5). Results were summarised in figure 9.14, 9.15, 9.16 and 9.17.



Codes	Method details	Method type
1-A	*Bligh & dyer*	Conventional methods
1-B	*Bligh & dyer with Bead beating*	
1-C	*Bligh & dyer with Sonication*	
1-D	*Modified BD by Smedes*	
2-A	*Folch*	
2-B	*Folch with Bead beating*	
2-C	*Folch with Sonication*	
2-D	*Modified Folch method*	
3	*Methanol:MTBE (1.5:5) (Matyash & Sheng)*	Other conventional methods
4	*Methanol:Chloroform (1:1) (Ryckebosch)*	
5	*Methanol:DCM (1:2) (Cequire-Sanchez)*	
6-A	*5% Methanolic HCl; Toluene for lipids solubilisation(Garcias: Original protocol)*	DT Acid Catalyst
6-B	*5% Methanolic HCl; Chlorofrom:Methanol (2:1) for lipid solubilisation (Laurens: Little modification)*	
6-C	*10% BF3 in methanol (Laurens: Little modification)*	
6-D	*5% Acetyl chloride in methanol*	
7-A	*0.5M Sodium methoxide*	DT Base Catalyst
7-B	*1:4 v/v TMG in methanol (Laurens: Modified)*	
8-A	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Original Protocol)*	DT Both Acid & Base Catalyst
8-B	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Modified Protocol)*	
8-C	*Sodium meth./BF3 in methanol. Incubated together (Laurens: Original Protocol)*	
8-D	*Sodium meth. & Acetyl chloride. Incubated separately*	
8-E	*Sodium meth. & M. HCl Incubated separately*	

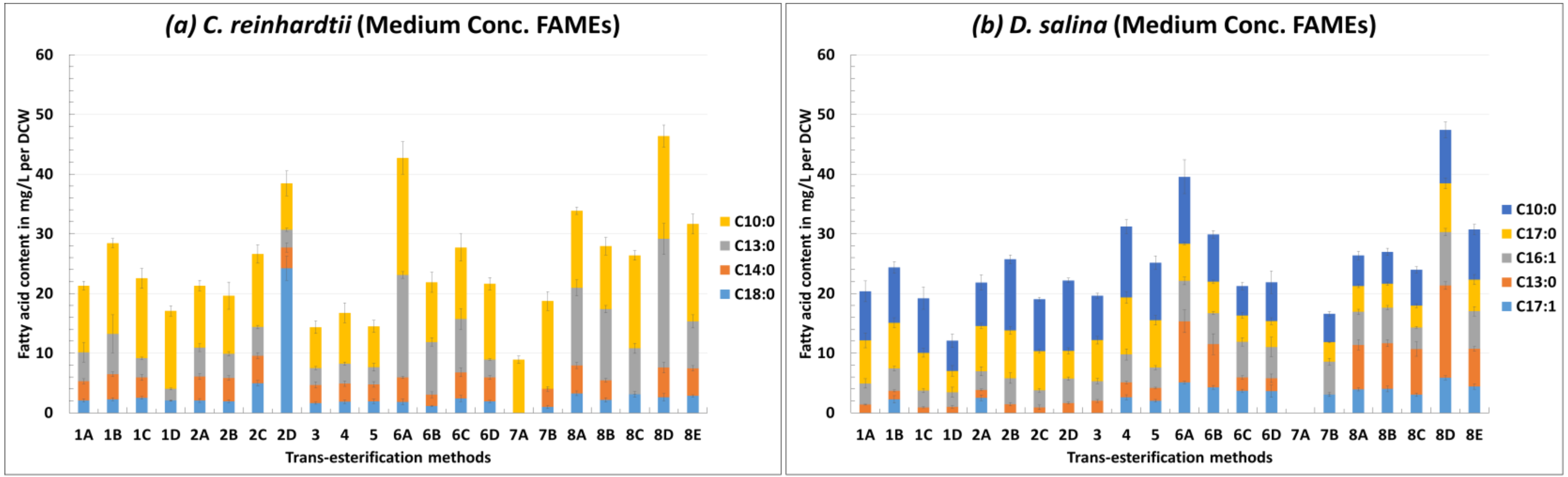
Figure 9.14 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery total FAs content and composition were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total FA content per DCW and individual FA composition expressed in mg/L per DCW.

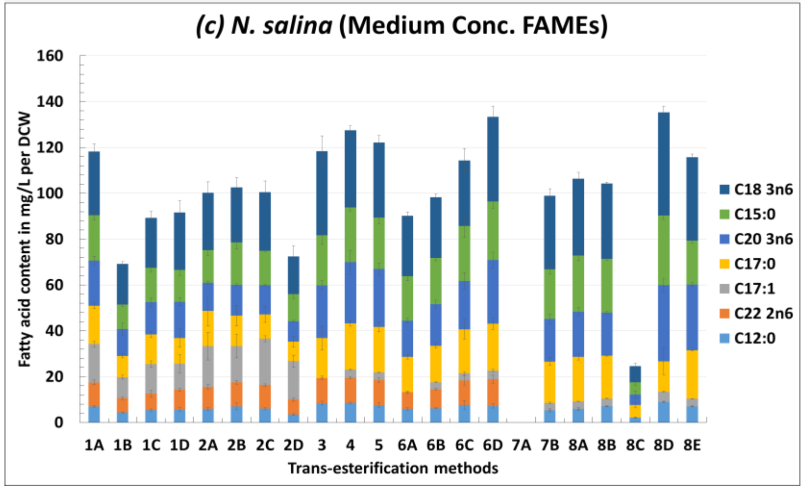




Codes	**Method details**	**Method type**
1-A	*Bligh & dyer*	**Conventional methods**
1-B	*Bligh & dyer with Bead beating*	
1-C	*Bligh & dyer with Sonication*	
1-D	*Modified BD by Smedes*	
2-A	*Folch*	
2-B	*Folch with Bead beating*	
2-C	*Folch with Sonication*	
2-D	*Modified Folch method*	
3	*Methanol:MTBE (1.5:5) (Matyash & Sheng)*	**Other conventional methods**
4	*Methanol:Chloroform (1:1) (Ryckebosch)*	
5	*Methanol:DCM (1:2) (Cequire-Sanchez)*	
6-A	*5% Methanolic HCl; Toluene for lipids solubilisation(Garcias: Original protocol)*	**DT Acid Catalyst**
6-B	*5% Methanolic HCl; Chlorofrom:Methanol (2:1) for lipid solubilisation (Laurens: Little modification)*	
6-C	*10% BF3 in methanol (Laurens: Little modification)*	
6-D	*5% Acetyl chloride in methanol*	
7-A	*0.5M Sodium methoxide*	**DT Base Catalyst**
7-B	*1:4 v/v TMG in methanol (Laurens: Modified)*	
8-A	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Original Protocol)*	**DT Both Acid & Base Catalyst**
8-B	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Modified Protocol)*	
8-C	*Sodium meth./BF3 in methanol. Incubated together (Laurens: Original Protocol)*	
8-D	*Sodium meth. & Acetyl chloride. Incubated separately*	
8-E	*Sodium meth. & M. HCl Incubated separately*	

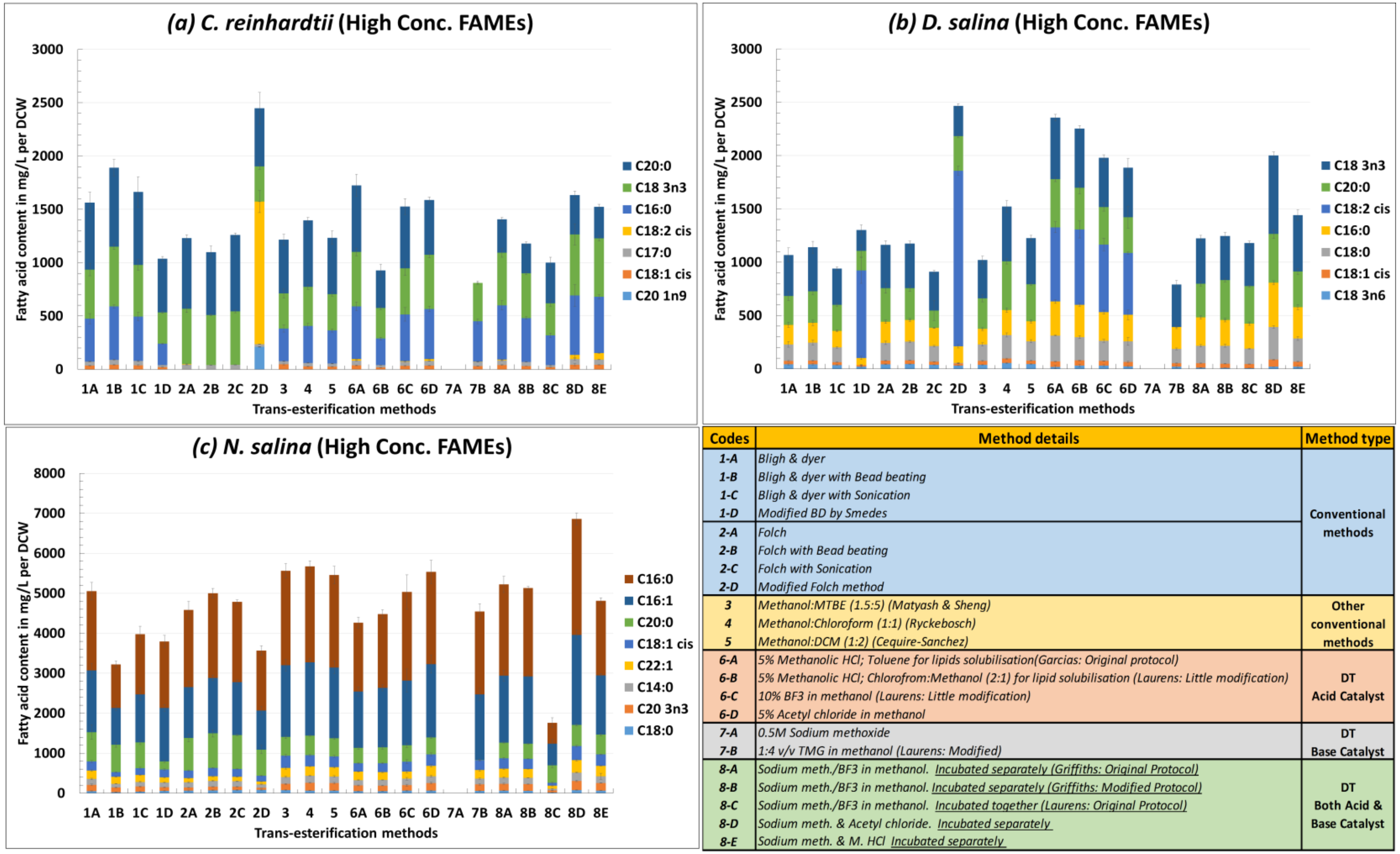
Figure 9.15 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery low concentration FAMEs were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total FA content per DCW and individual FA composition expressed in mg/L per DCW.





Codes	Method details	Method type
1-A	*Bligh & dyer*	Conventional methods
1-B	*Bligh & dyer with Bead beating*	
1-C	*Bligh & dyer with Sonication*	
1-D	*Modified BD by Smedes*	
2-A	*Folch*	
2-B	*Folch with Bead beating*	
2-C	*Folch with Sonication*	
2-D	*Modified Folch method*	
3	*Methanol:MTBE (1.5:5) (Matyash & Sheng)*	Other conventional methods
4	*Methanol:Chloroform (1:1) (Ryckebosch)*	
5	*Methanol:DCM (1:2) (Cequire-Sanchez)*	
6-A	*5% Methanolic HCl; Toluene for lipids solubilisation(Garcias: Original protocol)*	DT Acid Catalyst
6-B	*5% Methanolic HCl; Chlorofrom:Methanol (2:1) for lipid solubilisation (Laurens: Little modification)*	
6-C	*10% BF3 in methanol (Laurens: Little modification)*	
6-D	*5% Acetyl chloride in methanol*	
7-A	*0.5M Sodium methoxide*	DT Base Catalyst
7-B	*1:4 v/v TMG in methanol (Laurens: Modified)*	
8-A	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Original Protocol)*	DT Both Acid & Base Catalyst
8-B	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Modified Protocol)*	
8-C	*Sodium meth./BF3 in methanol. Incubated together (Laurens: Original Protocol)*	
8-D	*Sodium meth. & Acetyl chloride. Incubated separately*	
8-E	*Sodium meth. & M. HCl Incubated separately*	

Figure 9.16 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery medium concentration FAMEs were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total FA content per DCW and individual FA composition expressed in mg/L per DCW.



Codes	Method details	Method type
1-A	*Bligh & dyer*	**Conventional methods**
1-B	*Bligh & dyer with Bead beating*	
1-C	*Bligh & dyer with Sonication*	
1-D	*Modified BD by Smedes*	
2-A	*Folch*	
2-B	*Folch with Bead beating*	
2-C	*Folch with Sonication*	
2-D	*Modified Folch method*	
3	*Methanol:MTBE (1.5:5) (Matyash & Sheng)*	**Other conventional methods**
4	*Methanol:Chloroform (1:1) (Ryckebosch)*	
5	*Methanol:DCM (1:2) (Cequire-Sanchez)*	
6-A	*5% Methanolic HCl; Toluene for lipids solubilisation(Garcias: Original protocol)*	**DT Acid Catalyst**
6-B	*5% Methanolic HCl; Chlorofrom:Methanol (2:1) for lipid solubilisation (Laurens: Little modification)*	
6-C	*10% BF3 in methanol (Laurens: Little modification)*	
6-D	*5% Acetyl chloride in methanol*	
7-A	*0.5M Sodium methoxide*	**DT Base Catalyst**
7-B	*1:4 v/v TMG in methanol (Laurens: Modified)*	
8-A	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Original Protocol)*	**DT Both Acid & Base Catalyst**
8-B	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Modified Protocol)*	
8-C	*Sodium meth./BF3 in methanol. Incubated together (Laurens: Original Protocol)*	
8-D	*Sodium meth. & Acetyl chloride. Incubated separately*	
8-E	*Sodium meth. & M. HCl Incubated separately*	

Figure 9.17 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery high concentration FAMEs were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total FA content per DCW and individual FA composition expressed in mg/L per DCW.

Comparison within the BD method and its modifications with respect to use of disruption technique employed, suggested that in case of *C. reinhardtii* and *D. salina*, bead beating method (1B) resulted in highest recoveries of total FAs which was complying with our previous findings as mentioned in section 9.3.2.3, In case of *N. salina*, sonication method (1C) gave better results. On the other hand, comparison within the F method and its modifications, suggested that in case of *N. salina* and *D. salina* bead beating method (2B), whereas in case of *C. reinhardtii*, sonication method (2C) resulted in highest recoveries of total FAs. Comparison within recently suggested conventional methods (3, 4 and 5) employing different solvent ratios and use of non-chlorinated solvents, suggested that method 4 involving use of methanol-chloroform ratio 1:1 (*v/v*) for lipid extraction resulted in highest recoveries of total FAs across all the three species. Our findings were in agreement with (Ryckebosch et al., 2012) where author compared his results with that of (Matyash et al., 2008, Sheng et al., 2011) who suggested use of methanol:MTBE (1.5:5).

Overall, 2D method involving use of non-chlorinated solvents like propan-2-ol, cyclohexane and water mixture (8:10:11 *v/v/v*) rendered the highest total FAs recoveries in *C. reinhardtii* and *D. salina*. Whereas in case of *N. salina*, method 3 and 5 in which chloroform was replaced by non-chlorinated solvent like MTBE (Matyash et al., 2008, Sheng et al., 2011) and DCM (Cequier-Sánchez et al., 2008) and method 4 involving use methanol-chloroform mixture (1:1 *v/v*) as suggested by (Ryckebosch et al., 2012) resulted in higher total FAs recoveries with negligible differences. In conclusion, use of non-chlorinated solvents for lipid extraction resulted in higher FAs recoveries as compared to that of conventional methods across all the three species. Recently, (Kumari et al., 2011) compared and evaluated several methods, however author didn't involve important solvent mixture such as use of Propane-2-ol and cyclohexane (method 2D), which yielded higher recoveries for total FAs in our study compared to other methods. Extremely higher recoveries of linoleic acid (C18:2 CIS) were obtained with 2D methods in case of *C. reinhardtii* and *D. salina*, an important candidate for biodiesel production.

9.3.5 One step direct trans-esterification (DT) methods

9.3.5.1 Evaluation of acid catalyst

Higher total FA content was observed with the use of 5% methanolic HCl (method 6A) in case of both *C. reinhardtii* and *D. salina*, whereas higher recoveries were obtained with 5% acetyl chloride in methanol (method 6D) in case *N. salina*. In case of *C. reinhardtii* and *D. salina*, higher recoveries were obtained with 5% methanolic HCl (method 6A), primarily due to better trans-esterifying efficiency of 5% methanolic HCl towards SFAs like capric acid (C10:0), tridecanoic acid (C13:0), palmitic acid (C16:0), stearic acid (C18:0) and arachidic acid (C20:0) compared to that of

5% acetyl chloride in methanol. Our findings were contradictory to that of (Harmanescu, 2012), where the authors reported higher recoveries with the use BF_3/methanol over that of methanolic HCl for DT for FA analysis in vegetable matrix with low fat content.

The recovery of medium and high concentration FAs with all the catalysts were found to be the same across all the three species (fig. 9.16 and 9.17), except for 5% methanolic HCl with which, cis-10-heptadecenoic acid (C17:1) was not recovered in case of *N. salina* (method 6A) and only trace amounts of linoleic acid (C18:2 CIS) was recovered in case of *C. reinhardtii.* In contrast to this, huge variations in recovery of low concentration FAs (most of them were SFA in case *C. reinhardtii* and *D. salina*) were observed with the use of different acid catalysts. In case of *C. reinhardtii* and *D. salina,* pentadecanoic acid (C15:0) was only recovered with 5% acetyl chloride in methanol and that too in fairly high amount compared to other catalysts. However in case of *D. salina*, undecanoic acid (C11:0), myristic acid (C14:0) and cis-11-eicosenoic acid (C20:1n9) were only recovered with 5% methanolic HCl (method 6A) compared to that of with 5% acetyl chloride in methanol (method 6D).

Across all species BF3/methanol and 5% acetyl chloride in methanol showed equal trans-esterification efficiency with minor variations for recovery of SFAs, MUFAs and PUFAs as can be seen in fig. 9.18, 9.19 and 9.20, however it was worth to note that higher number of FAMEs were trapped with 5% acetyl chloride in methanol (method 6D). As demonstrated in this study, it can be clearly seen that, with respect to the number unique FAs recovered with each catalyst, acetyl chloride has the highest TE efficiency and henceforth recovered maximum number of FAMEs as compared to other catalysts across all the three species, especially in case *N. salina* where superior recovery was obtained for low concentration FAMEs. Overall it can be clearly concluded that within methods involving use of acid catalyst, method 6D seems to be the best in yielding higher recovery of total FA content and all SFA, MUFA and PUFA across all the three species.

9.3.5.2 Evaluation of base catalyst

DT with a 0.5 M sodium methoxide alone as an alkaline catalyst resulted in no TE reaction at all across all the three species. Possible reason might be the presence of water in the wet algal biomass which inhibited the catalyst completely. This finding was in agreement with (Griffiths et al., 2010). In contrast to this use of 1:4, *v/v* TMG in methanol resulted in superior recovery of FAs, compared to none with that of 0.5 M sodium methoxide.

9.3.5.3 Evaluation of combined use of catalyst

The amount of both the catalysts were reduced to half in method 8B to that of method 8A, however only minor variation in the recovery of total FAs were observed suggesting such high amounts of catalyst are not really required to carry out TE reactions, where the available biomass amounts on laboratory scales are ranging from 1 to 16 mg DCW arising from approximately 15 mL of culture. DT with combined use of SM and 5% acetyl chloride in methanol, incubated separately (method 8D) followed by method E involving use SM and M. HCl, resulted in highest recoveries of total FAs across all the DT methods involving use of combined catalysts. The lowest recovery of total FA content was obtained with method C in which SM and BF3/methanol were incubated together. The possible reason for the lowest recovery might be exclusion of pre-soaking step essential for prior solubilisation of algal lipids, as in all the other methods toluene was added for prior solubilisation of various lipids. Another potential reason might be incubation of both the acid and base catalysts together, where efficiency of both the catalysts might have been reduced.

The recovery of PUFAs (fig. 9.20) were not affected greatly across different treatments except for cis-5, 8, 11, 14, 17-eicosapentaenoic acid (C20:5n3) which was not recovered in method 8D and 8E in *N. salina.* In contrast, huge variations were observed in recovery of SFAs (fig. 9.18) across all the species. Caproic acid (C6:0) was not recovered in *N. salina* and *D. salina* with method 8A, while in case of *C. reinhardtii* no traces were detected with method 8B, 8C and 8E. Apart from that, myristic acid (C14:0) was not recovered with method 8C in *C. reinhardtii,* whereas undecanoic acid (C11:0) was only recovered with method 8D in *D. salina.* Similarly recovery of caprylic acid (C8:0), undecanoic acid (C11:0) and tridecanoic acid (C13:0) were only obtained with method 8D and that of behenic acid (C22:0), only with method 8A in case of *N. salina.*

In case of MUFAs (fig.9.19), not much variations were observed in case of *C. reinhardtii*, as only one MUFA was recovered with all the applied methods, however in case *D. salina,* recovery of cis-11-eicosenoic acid (C20:1n9) was only obtained with method 8B. In *N. salina,* cis-10-pentadecenoic acid (C15:1) and nervonic acid (C24:1n9) were only recovered with method 8D and 8E respectively, whereas cis-10-heptadecenoic acid (C17:1) was detected with all the methods except with 8C and similarly cis-11-eicosenoic acid (C20:1n9) was detected with all the methods except with method 8A and 8B.

Overall it can be clearly concluded that within methods involving use of both the catalyst, use SM followed by 5% acetyl chloride in methanol (method 8D) seems to be the best method yielding higher recoveries of total FA content and all SFA, MUFA and PUFA across all the three species.

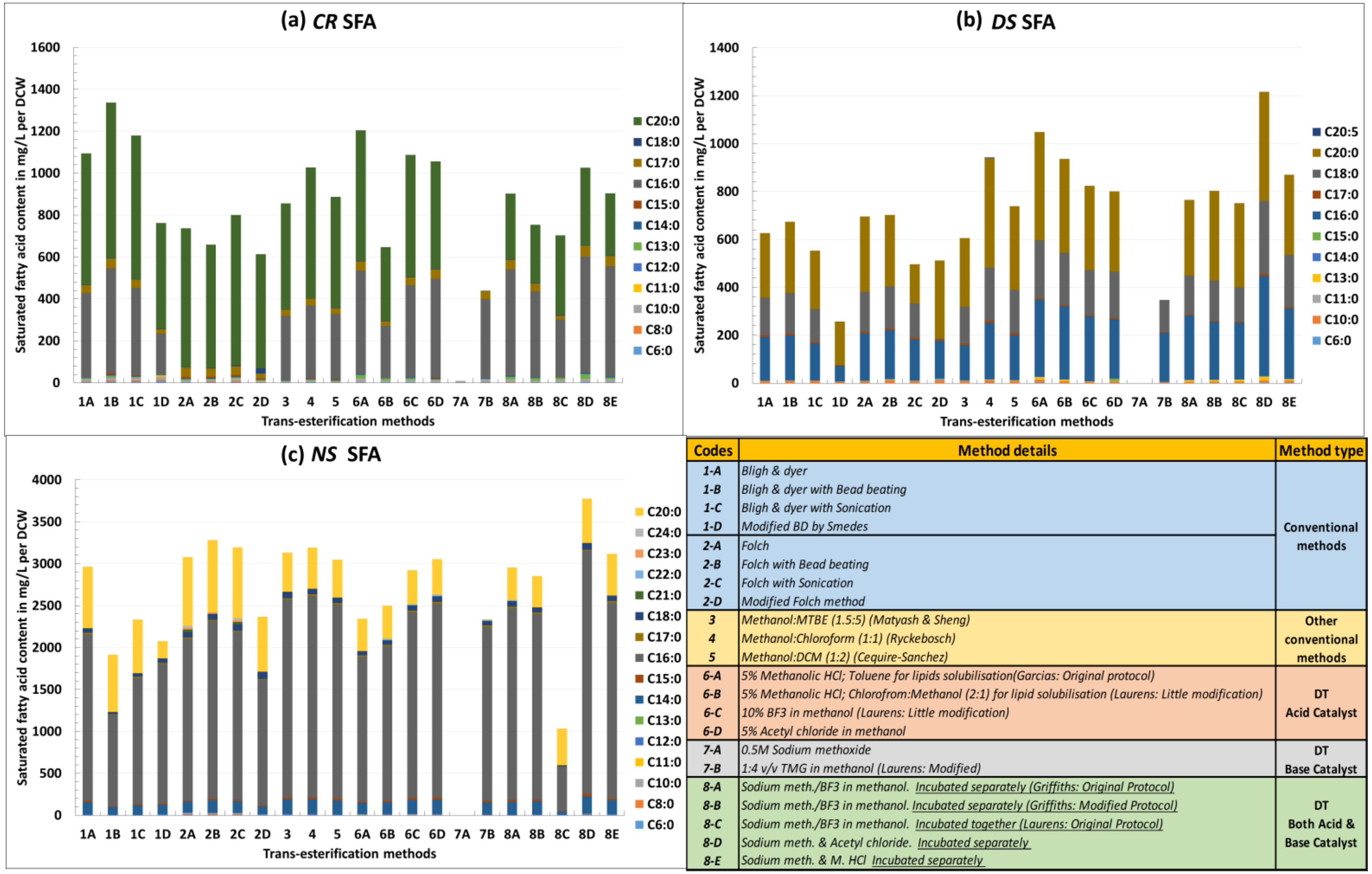
9.3.5.4 Comparison between acid, base and combined use of both catalysts.

As discussed above, no trans-esterification occurred with DT using alkaline catalyst like sodium methoxide, whereas total recovery of FAs obtained with use 1:4, *v/v* TMG in methanol was far better compared to that of SM but not too great compared to other methods. The obvious reason is the high content of FFAs in algae (as high as 35.1%) which reduces the efficiency of base catalyst, as FFAs content >3% results in formation of soap. The visual differences can also be noticed in figure 9.1. Henceforth use of alkaline catalyst alone for DT should be avoided whenever possible.

In contrast to this, DT using acid catalyst resulted in superior recovery especially with the use of 5% methanolic HCl (6A) in case of *C. reinhardtii* and *D. salina* and with the use of 5% acetyl chloride in methanol (6D) in case of *N. salina*.

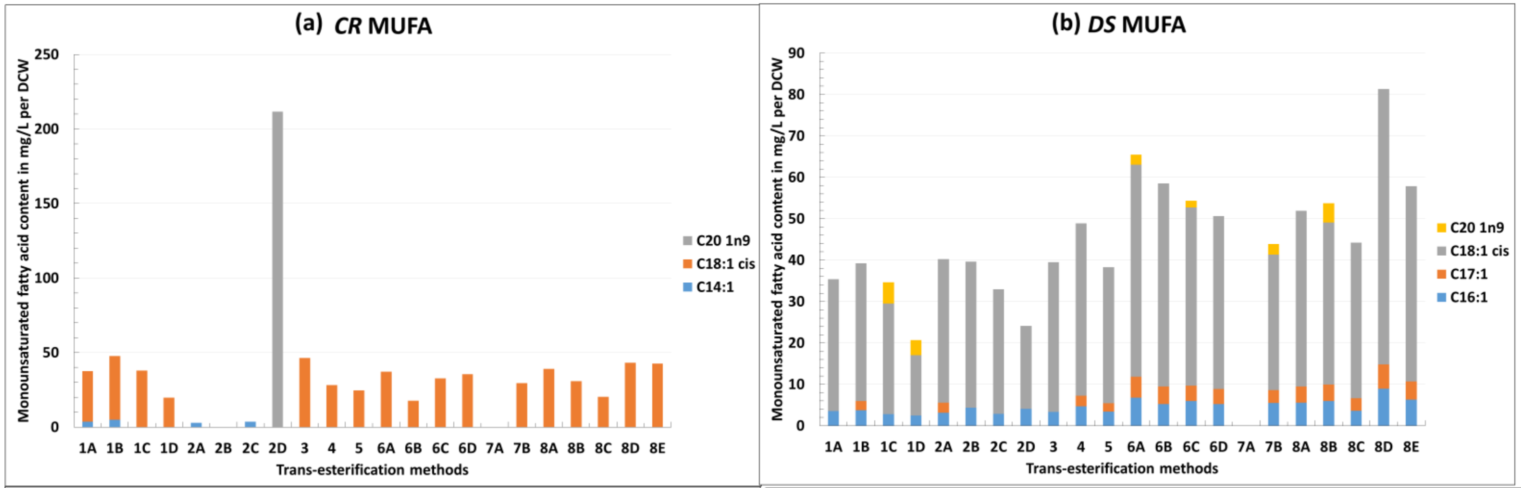
While, in case of sequential use of base catalyst SM followed by an acid catalysis with 5% acetyl chloride in methanol (8D), managed to obtain highest recoveries of total FAs content across all the species.

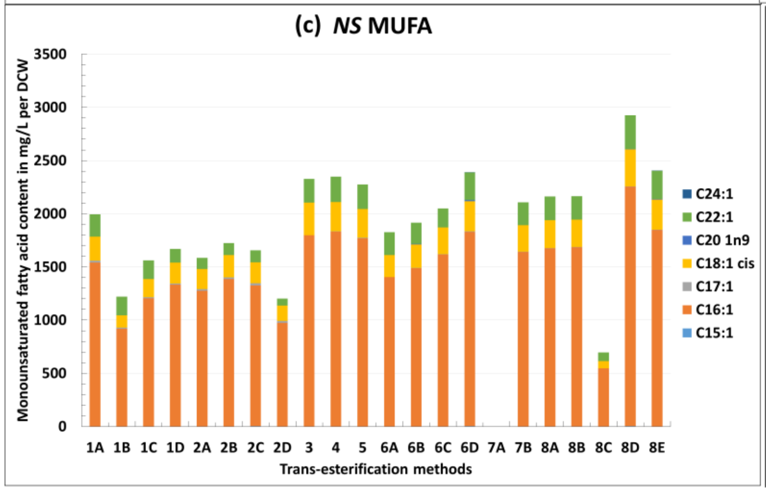
Comparison between all the applied DT methods (6A to 8E) across all the three species, suggests that combined use of base catalyst SM, followed by acid catalysis with 5% acetyl chloride in methanol (8D), was superior method, in terms of number of individual FAMEs (SAFs, MUFAs and PUFAs) and total FA content recovery and henceforth most recommended.

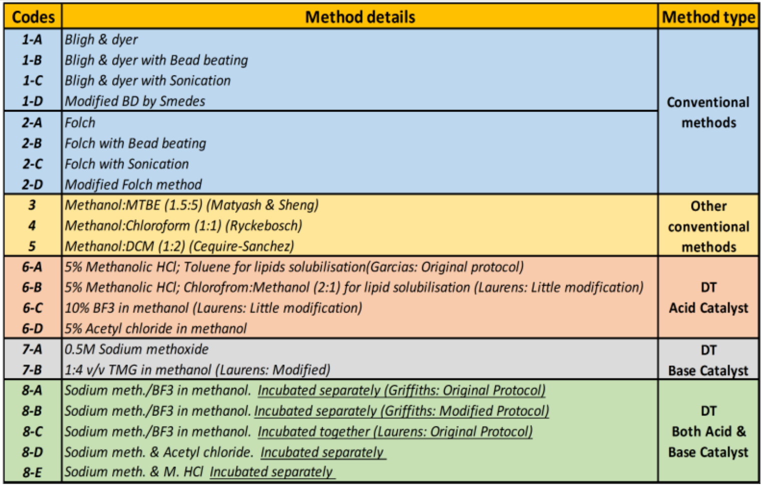


Codes	Method details	Method type
1-A	*Bligh & dyer*	Conventional methods
1-B	*Bligh & dyer with Bead beating*	
1-C	*Bligh & dyer with Sonication*	
1-D	*Modified BD by Smedes*	
2-A	*Folch*	
2-B	*Folch with Bead beating*	
2-C	*Folch with Sonication*	
2-D	*Modified Folch method*	
3	*Methanol:MTBE (1.5:5) (Matyash & Sheng)*	Other conventional methods
4	*Methanol:Chloroform (1:1) (Ryckebosch)*	
5	*Methanol:DCM (1:2) (Cequire-Sanchez)*	
6-A	*5% Methanolic HCl; Toluene for lipids solubilisation(Garcias: Original protocol)*	DT Acid Catalyst
6-B	*5% Methanolic HCl; Chlorofrom:Methanol (2:1) for lipid solubilisation (Laurens: Little modification)*	
6-C	*10% BF3 in methanol (Laurens: Little modification)*	
6-D	*5% Acetyl chloride in methanol*	
7-A	*0.5M Sodium methoxide*	DT Base Catalyst
7-B	*1:4 v/v TMG in methanol (Laurens: Modified)*	
8-A	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Original Protocol)*	DT Both Acid & Base Catalyst
8-B	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Modified Protocol)*	
8-C	*Sodium meth./BF3 in methanol. Incubated together (Laurens: Original Protocol)*	
8-D	*Sodium meth. & Acetyl chloride. Incubated separately*	
8-E	*Sodium meth. & M. HCl Incubated separately*	

Figure 9.18 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery saturated fatty acids (SFAs) content were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total SFA content per DCW and individual FA composition expressed in mg/L per DCW.

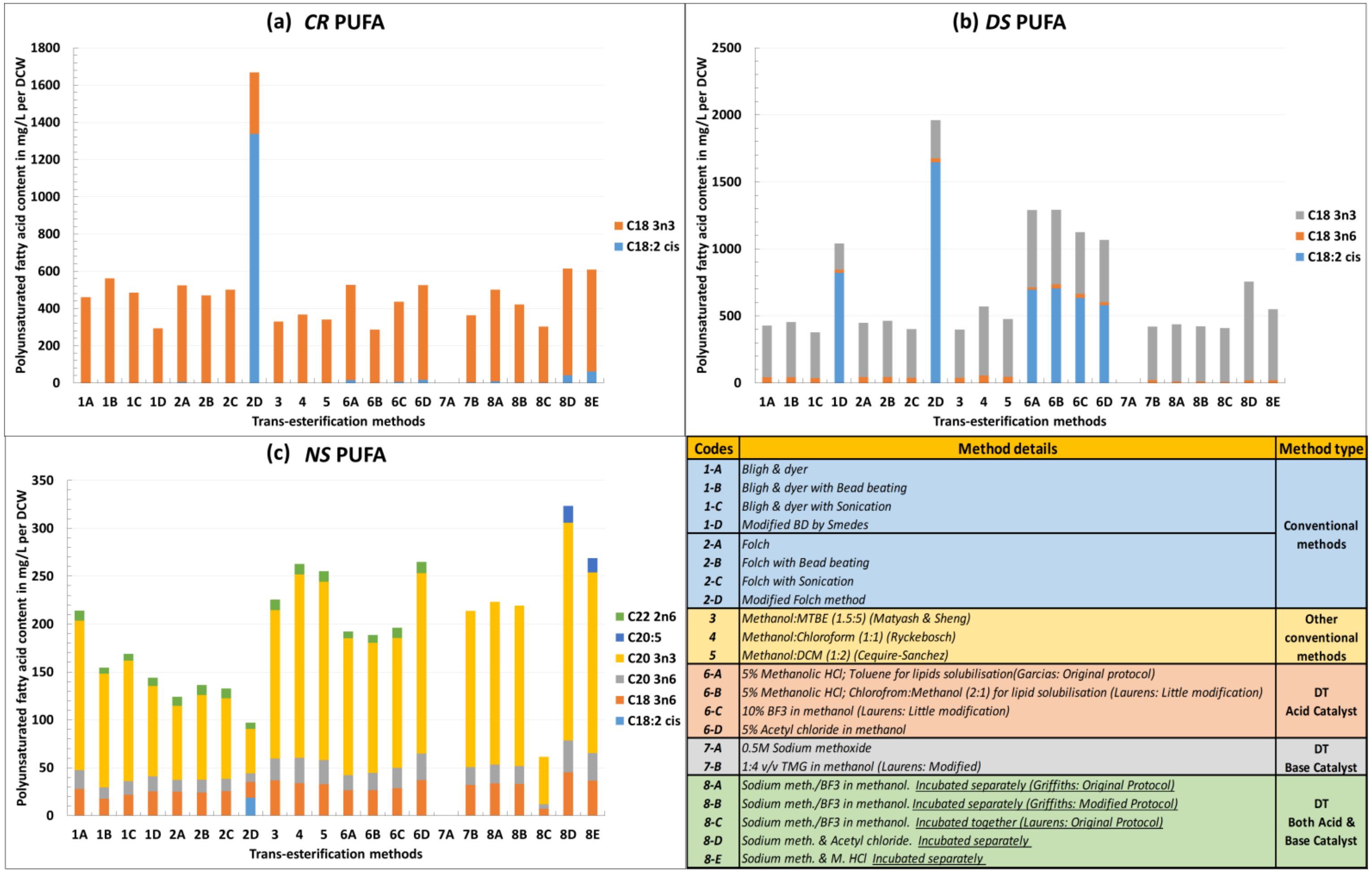






Codes	Method details	Method type
1-A	*Bligh & dyer*	**Conventional methods**
1-B	*Bligh & dyer with Bead beating*	
1-C	*Bligh & dyer with Sonication*	
1-D	*Modified BD by Smedes*	
2-A	*Folch*	
2-B	*Folch with Bead beating*	
2-C	*Folch with Sonication*	
2-D	*Modified Folch method*	
3	*Methanol:MTBE (1.5:5) (Matyash & Sheng)*	**Other conventional methods**
4	*Methanol:Chloroform (1:1) (Ryckebosch)*	
5	*Methanol:DCM (1:2) (Cequire-Sanchez)*	
6-A	*5% Methanolic HCl; Toluene for lipids solubilisation(Garcias: Original protocol)*	**DT Acid Catalyst**
6-B	*5% Methanolic HCl; Chlorofrom:Methanol (2:1) for lipid solubilisation (Laurens: Little modification)*	
6-C	*10% BF3 in methanol (Laurens: Little modification)*	
6-D	*5% Acetyl chloride in methanol*	
7-A	*0.5M Sodium methoxide*	**DT Base Catalyst**
7-B	*1:4 v/v TMG in methanol (Laurens: Modified)*	
8-A	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Original Protocol)*	**DT Both Acid & Base Catalyst**
8-B	*Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Modified Protocol)*	
8-C	*Sodium meth./BF3 in methanol. Incubated together (Laurens: Original Protocol)*	
8-D	*Sodium meth. & Acetyl chloride. Incubated separately*	
8-E	*Sodium meth. & M. HCl Incubated separately*	

Figure 9.19 Figure 9.20 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery monounsaturated fatty acid contents (MUFAs) concentration FAMEs were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total MUFA content per DCW and individual FA composition expressed in mg/L per DCW.



Codes	Method details	Method type
1-A	Bligh & dyer	Conventional methods
1-B	Bligh & dyer with Bead beating	
1-C	Bligh & dyer with Sonication	
1-D	Modified BD by Smedes	
2-A	Folch	
2-B	Folch with Bead beating	
2-C	Folch with Sonication	
2-D	Modified Folch method	
3	Methanol:MTBE (1.5:5) (Matyash & Sheng)	Other conventional methods
4	Methanol:Chloroform (1:1) (Ryckebosch)	
5	Methanol:DCM (1:2) (Cequire-Sanchez)	
6-A	5% Methanolic HCl; Toluene for lipids solubilisation(Garcias: Original protocol)	DT Acid Catalyst
6-B	5% Methanolic HCl; Chlorofrom:Methanol (2:1) for lipid solubilisation (Laurens: Little modification)	
6-C	10% BF3 in methanol (Laurens: Little modification)	
6-D	5% Acetyl chloride in methanol	
7-A	0.5M Sodium methoxide	DT Base Catalyst
7-B	1:4 v/v TMG in methanol (Laurens: Modified)	
8-A	Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Original Protocol)	DT Both Acid & Base Catalyst
8-B	Sodium meth./BF3 in methanol. Incubated separately (Griffiths: Modified Protocol)	
8-C	Sodium meth./BF3 in methanol. Incubated together (Laurens: Original Protocol)	
8-D	Sodium meth. & Acetyl chloride. Incubated separately	
8-E	Sodium meth. & M. HCl Incubated separately	

Figure 9.20 Figure 9.21 Effect of various conventional (Two step TE: Extraction followed by TE) and DT methods on recovery polyunsaturated fatty acid contents (PUFAs) concentration FAMEs were compared for *C. reinhardtii* (a), *D. salina* (b) and *N. salina* (c). X-axis represents different applied methods, description of each method is described in table adjacent to fig. (c). Y-axis represents total PUFA content per DCW and individual FA composition expressed in mg/L per DCW.

9.3.6 Comparison of conventional and DT methods

Recovery of SFAs

Overall with all the applied methods, in total 12, 10 and 16 SFAs were detected in *C. reinhardtii, D. salina* and *N. salina* respectively (figure 9.18). The common SFAs (most suitable for biodiesel productions) recovered across all the species were caproic acid (C6:0), capric acid (C10:0), undecanoic acid (C11:0), tridecanoic acid (C13:0), myristic acid (C14:0), pentadecanoic acid (C15:0), palmitic acid (C16:0), heptadecanoic acid (C17:0), stearic acid (C18:0) and arachidic acid (C20:0). Among which capric acid (C10:0), palmitic acid (C16:0), heptadecanoic acid (C17:0) and arachidic acid (C20:0) were found to be present in higher concentration which were detected efficiently with all the methods across all the three species except for capric acid (C10:0) which was not recovered with BD methods (1A to 1D) and with the other conventional methods like 3, 4 and 5 in case of *N. salina.*

On the other hand low concentration SFAs showed huge variations across different methods applied. In *C. reinhardtii,* only Folch methods (2A to 2D) were unable to recover palmitic acid (C16:0), whereas pentadecanoic acid (C15:0) was recovered well with all the conventional methods compared to DT methods where exception was method 6D. The recovery of medium chain fatty acids like caprylic acid (C8:0), undecanoic acid (C11:0) and lauric acid (C12:0) was only recovered well with the conventional methods and lost with the DT methods.

In case of *D. salina,* pentadecanoic acid (C15:0) was only recovered with method 6D, while caproic acid (C6:0) was only recovered with all DT methods and lost with all the conventional methods. Undecanoic acid (C11:0) was only recovered with 2 Folch (2B, 2D) and 2 DT (6A, 8D) methods.

In case of *N. salina,* tricosanoic acid (C23:0) and lignoceric acid (C24:0) were only recovered with Folch methods, whereas undecanoic acid (C11:0), tridecanoic acid (C13:0) were only recovered with the acid catalyst methods (6A, 6D) and with method 8D. Similarly, behenic acid (C22:0) was only recovered with all acid catalysts methods and with method 8A.

Overall, better recovery of SFAs were obtained with the conventional method (1B) in *C. reinhardtii* and with DT method (8D) in case of *D. salina* and *N. salina.*

Recovery of MUFAs

Overall with all the applied methods, in total 3, 4 and 7 MUFAs were detected in *C. reinhardtii, D. salina* and *N. salina* respectively (figure 9.19). The common MUFAs detected across all the three species were oleic acid (C18:1n9c) and cis-11-eicosenoic acid (C20:1n9), whereas palmitoleic acid (C16:1) and cis-10-heptadecenoic acid (C17:1) were detected only in *N. salina* and *D. salina.* Oleic acid (C18:1n9c) was recovered with all the applied methods except with the

Folch methods (2A – 2D) in case of *C. reinhardtii,* whereas very few methods were able to recover cis-11-eicosenoic acid (C20:1n9), in which exceptionally high recovery with only method 2D was observed in *C. reinhardtii*. On the other hand, all the methods successfully recovered palmitoleic acid (C16:1) in case of *N. salina* and *D. salina*, among which higher recovery was obtained with the method 2D, however the recovery of cis-10-heptadecenoic acid (C17:1) was variable with the conventional methods in case of *D. salina* and in both the species method 3 was unable to recover it at all. Highest number of MUFAs were detected in *N. salina,* in which trace amounts of nervonic acid (C24:1n9) was detected only with the method 6D and 8E and similarly cis-10-pentadecenoic acid (C15:1) was detected in trace amounts only with the method 2C, 6D and 8D. The recovery of cis-11-eicosenoic acid (C20:1n9) was found to be superior across all the methods.

Overall, better recovery of MUFAs were obtained with the conventional method (2D) in *C. reinhardtii* and with DT method (8D) in case of *D. salina* and *N. salina.*

Recovery of PUFAs

Overall with all the applied methods, in total 2, 4 and 6 PUFAs were detected in *C. reinhardtii, D. salina* and *N. salina* respectively (figure 9.20). Only linoleic acid (C18:2 CIS) was detected in all the three species, in which only method 2D was able to recover it in higher concentration compared to other PUFAs. Additionally, with the DT methods using acid catalysts (method 6A to 6D) trace levels were recovered in *C. reinhardtii* and good recovery was obtained in *D. salina.*

On the other hand all the methods showed good recoveries with minor variations, for g-linolenic acid (C18:3n6) which was only detected in *N. salina* and *D. salina* and a-linolenic acid (C18:3n3) which was only detected in *C. reinhardtii* and *D. salina.*

Highest number of PUFAs were detected in *N. salina,* in which all the methods showed good recovery of cis-8, 11, 14-eicosatrienoic acid (C20:3n6) and cis-11, 14, 17-eicosatrienoic acid (C20:3n3), however recovery of cis-5, 8, 11, 14, 17-eicosapentaenoic acid (C20:5n3) and cis-11, 14-eicosadienoic acid (C20:2n6) was found to be selective. (C20:2n6) was only recovered with the conventional and DT using acid catalyst methods whereas, only method 8D and 8E were able to recover (C20:5n3). Overall, better recovery of PUFAs were obtained with the conventional method (2D) in *C. reinhardtii* and with DT method (6A) and (8D) in case of *D. salina* and *N. salina* respectively.

In conclusion, it can be said that the conventional methods recovered higher total FA content in case of *C. reinhardtii,* in which SFAs were well recovered with the method 1B while recovery of MUFAs and PUFAs was most favoured with method 2D. In contrast to this in cases of *D. salina* and *N. salina*, method 8D was found to be superior across all the methods, in recovery of all SFAs, MUFAs and PUFAs.

9.4 Conclusions

In this chapter, **four experiments** were carried out to optimize the FAME production workflow from selected three microalgal strains. All laboratory scaled down procedures were developed to carry out FAME production processes on a very small amounts of microalgal biomass ranging from 1 to 16 mg DCW. The overall biodiesel fuel characteristics largely dependent on fatty acid composition of microalgal species, hence strains were carefully selected which includes *C. reinhardtii, D, salina* and *N. salina.* We have evaluated all our methods on wet algal biomass and skipped the process of lyophilisation, as lyophilisation is known to have many adverse effect on recovery of short chain SFAs (important for biodiesel production). Furthermore, addition of a synthetic antioxidant is also not required, as algae itself contains large amount of natural antioxidants. Moreover, these pre-treatments are more time consuming and will incur more cost for industrial scale application.

The optimal extraction technique and parameters should be valid for all microalgae and could tremendously reduce the amount of extraction solvent and energy input required. Furthermore, an ideal disruption technique should be one which can selectively extract the lipid classes of interest (from biodiesel point of view) without causing any form of damage to them. Henceforth, **in the first experiment** we have used design of experiments to assess the lipid extraction efficiencies of widely cited bead beating and sonication techniques across two selected microalgal strains *C. reinhardtii* and *N. salina*, which differ in their cell wall structure. The microalgal cell integrity was further studied using scanning electron microscopy (SEM), for untreated wet algal biomass, bead beating treated and sonicated biomass following extraction. Within bead beating technique, 100% recoveries were obtained with 5_2 and 5_1 parameters in *C. reinhardtii* and *N. salina* respectively. Whereas within sonication technique, 100% recoveries were obtained with 3603 and 2603 parameters in *C. reinhardtii* and *N. salina* respectively. Despite slightly higher concentration of FAMEs recovered with sonication treatments compared to bead beating treatments, loss of stearic acid (C18:0) with all the sonication treatments (except 3603 and 3606) in *C. reinhardtii*, was considered as an inappropriate solution for biodiesel production. In contrast, all the bead beating treatments recovered equal concentrations of stearic acid (C18:0). Moreover at laboratory scale, using sonication, only a single sample can be processed at one time whereas with bead beating technique, twelve samples can be processed simultaneously which will save tremendous amount of time and energy. Hence based on experimental evidences and SEM observation, bead beating was selected as a suitable disruption technique with optimised parameters (5 bead beating cycle with 2 minutes relaxation time on ice), for all further experiments.

To ensure the reliable and accurate quantification of fatty acids is primarily depending on optimised GC parameters and calculation of identified FAMEs concentration based on internal and external standards. Hence **in the second experiment** of GC parameters optimisation, we have evaluated our results by generating calibration curve and retention time deviation plots for the external standards. From results we strongly recommend running of external standard as a calibration and check standard within each sequence having no more than 20 samples in order to obtain accurate and reliable data set.

In case of conventional two step TE methods (extraction followed by TE), use of non-chlorinated solvents have potential to be an alternative to more toxic chloroform in recovery of lipids from different sample matrices. Furthermore, use of different solvent ratios within the same solvent system have a large effect on extraction efficiency. To explore these potentials of carefully selected solvent mixtures and their ratios, **in the third experiment,** we have evaluated use of non-chlorinated solvent mixtures and compared them within and with the conventional solvent systems. Results obtained by gravimetric estimations were highly variable and inaccurate for total lipid estimations as along with lipids, co-extraction of non-lipid material also contributing to the total weight and might leads to variable conclusions from different researchers. Despite this fact, gravimetric analysis for lipid content is still being widely used in lipid research community, we strongly advise not to rely on conclusion drawn from gravimetric determinations. Analysis of total FAs content by GC-FID is an appropriate alternative to gravimetric analysis, hence we have evaluated extraction and TE efficiency all our methods using GC-FID. Based on our results we have successfully demonstrated use of non-chlorinated solvents for microalgal lipid extraction across all the three species, with higher FAs recoveries as compared to that of toxic and carcinogenic conventional solvents mixtures (methanol and chloroform). The best yields were obtained with propane-2-ol, cyclohexane and water solvent mixture.

The final fourth experimental approach was directed towards careful selection of different catalysts and evaluation of their TE efficiency for biodiesel production from microalgal lipids. In total eleven DT protocols were carefully selected or constructed which were then compared within and with the conventional methods. Primarily, all the DT protocols were categorized based on the nature of the catalyst used; acid catalysed, base catalysed and combined (acid and base) catalysed DT methods. For thorough investigation, all the recovered FAMEs across different applied methods were categorised according to their concentration (low, high and medium) and based on their degree of saturation (SFAs, MUFAs and PUFAs). Among three acid catalysts evaluated, acetyl chloride showed highest TE efficiency and henceforth recovered maximum number of FAMEs as compared to other catalysts across all the three species. Among two base catalysts, SM showed no detectable FAME, whereas recovery obtained by TMG in methanol was not too great compared to acid catalyst methods. The obvious reason behind no or lower yield with base catalysts is higher FFAs content of algae, as FFAs content >3% results in formation of soap. Henceforth use of alkaline catalyst alone for DT of microalgal biomass should be avoided whenever possible. Among sequential use of both the catalyst, use of base catalyst SM, followed by acid catalysis with 5% acetyl chloride in methanol (8D), showed superior results, in terms of number of individual FAMEs (SAFs, MUFAs and PUFAs) and total FA content recovery and henceforth most recommended. It is important to note that, so far sequential use of SM with acetyl chloride in methanol or methanolic HCl has not been applied to microalgae for evaluating FAME recoveries for biodiesel productions.

In conclusion, it can be said that for biodiesel applications the conventional BD method with bead beating (method 1B) recovered higher total FA content in case of *C. reinhardtii,* (specially SFAs were well recovered), while recovery of MUFAs and PUFAs was most favoured with method 2D. In contrast to this in cases of *D. salina* and *N. salina*, use of base catalyst SM, followed by acid catalysis with 5% acetyl chloride (method 8D) was found to be superior across all the methods, in recovery of all SFAs, MUFAs and PUFAs. In summary, DT methods can be applied to a small amount of microalgal biomass (1-16 mg dry cell weight (DCW)), thus requiring small volume of solvent and reagents and so it is less expensive, less time consuming, the concern about post-extraction waste disposal and use of non-environment friendly solvents (for extraction purpose) are avoided, thereby making them more suitable for sustainable large scale biodiesel production.

Hence for biodiesel applications, we strongly recommend the use of DT method with sequential use of SM followed by acetyl chloride in methanol, however where different classes of lipid are of interest, use of non-chlorinated solvent mixture (propane-2-ol, cyclohexane and water) based on Folch principle (method 2D) is recommended.

Chapter 10

Evaluation of HILIC and IP-RP-HPLC for targeted analysis of amino acids, organic acids, water soluble vitamins, nucleotides, nucleosides and nucleobases towards improving metabolome coverage.

10. Evaluation of HILIC and IP-RP-HPLC for targeted analysis of amino acids, organic acids, water soluble vitamins, nucleotides, nucleosides and nucleobases towards improving metabolome coverage.

10.1 Introduction

The application of LC-MS in metabolomics has been growing over the past few years. The recovery of the analyte by collecting fractions is possible with LC, whereas this which seems to be a major challenge with GC separations. The coupling of LC to MS increases the number of analytes detected by decreasing the ionization suppression that occurs in direct infusion mass spectrometry (DIMS). The most commonly used soft ionization techniques are ESI to a greater extent and APCI to a lesser extent. The LC-MS platform in metabolomics offers several advantages over GC-MS, such as operation at lower analysis temperature, thus enabling analysis of thermo labile metabolites which often degrade during GC analysis. LC-MS analysis does not require chemical derivatization, thus simplifying sample preparation steps and identification of metabolites. Ionization in the positive ion mode (amines and amino acids) and negative ion mode (sugars and lipids) leads to the detection of two sets of analytes that can differ significantly. Hence, depending upon the mass analyser, detection in both positive and negative ion mode simultaneously, in a single run is possible with LC-MS. This reduces the time required for analysis and avoids possible bias due to injection errors. Although when compared to DIMS ion suppression is decreased by coupling LC and MS, ion suppression continues to still be a major disadvantage of LC-MS when compared to GC-MS. Ion suppression can further be overcome to some extent by miniaturization of ESI to nanospray ionization and by a better separation of metabolites using LC. Apart from ion suppression another issue in LC-MS based metabolomics is contamination of the MS source and adduct formation. These have significant consequences for the robustness of the method and the lack of transferable LC-MS libraries for metabolite identifications (Coulier et al., 2006). In addition, most of the compounds that are unstable or have low ionisation efficiency or are difficult to fragment, and are not suitable for LC-MS analysis. However, coupling LC-MS with derivatization reaction for such classes of metabolites might enhance their coverage by improving retention of polar compounds on RP columns, increasing HPLC separation selectivity, improved MS selectivity by increasing analyte MW and enhancing analyte nebulization and ionisation efficiency at MS interface (Xu et al., 2011). Despite these advantages, derivatization based LC-MS suffers from major issues such as formation of by-products, degradation and inter-conversion of metabolites. Therefore, derivatization is not recommended for global metabolomic studies.

The most commonly used columns in LC separation are reverse-phase columns, such as octodecyl silica (C18) or C8, for efficient separation of non-polar/hydrophobic metabolites such as aromatics, lipids, phospholipids and bile acids. However, retention of polar metabolites on reverse-phase columns is not sufficient and these compounds often co-elute close to the column void volume, thus making their detection difficult with MS. This problem can be overcome by use of HILIC which offers a metabolic profile complementary to RPLC by providing greater retention of hydrophilic metabolites. It offers complementary performance to that of RP-LC and similar to that of NP-LC, where instead of non-aqueous mobile phase, an eluent with high content of organic solvent is used, which further increases the sensitivity when coupled to ESI-MS (Cubbon et al., 2010). In HILIC, a stagnant water layer is established at the interface of the stationary phase and the separation is achieved by partitioning the analyte between that water layer and the organic mobile phase. The problems associated with the use of HILIC columns include slower re-equilibration time after gradient elution and lower resolution than that of C18 columns. Hence HILIC is not suitable for analyses requiring faster gradient conditions and re-equilibration times (Callahan et al., 2009). However recent advances with the stationary phases for HILIC such as ZIC-*p*HILIC columns have provided better separation and reproducible retention times for most of the polar metabolites including those which are poorly retained on RP columns. Some reports suggest the combined use of both HILIC and RP-LC for better metabolome coverage (Lu et al., 2008b, Spagou et al., 2010). Although it seems a better approach, in reality, the data generated from both the platforms are extremely difficult to fuse in metabolomic studies (Theodoridis et al., 2012).

An alternative option for the analysis of hydrophilic metabolites with reverse phase columns is the use of ion pair reagents containing non-polar moieties. The ion pair reagents have the ability to retain hydrophilic metabolites on a hydrophobic stationary phase. Thus, non-covalent interactions between the ion pair reagent with the metabolite present in ionized form, allows greater retention of polar metabolites on reversed phase column by the interaction of the hydrophobic moiety of the ion pair reagent with the stationary phase. This approach has been widely applied in microbial metabolomics to study glycolysis, pentose phosphate and TCA cycle metabolites (Coulier et al., 2006, Luo et al., 2007a). Another alternative to RP-LC is mixed mode chromatography which was recently being suggested for better coverage of metabolome specifically for polar metabolites. Multi-mode ODS columns such as Scherzo SM-C18, in which ODS + cation + anion ligands were embedded within alkyl functional groups, offer better retention and separation for highly polar metabolites as well as non-polar metabolites in a single run, therefore offers both RP and NP separations (Patti, 2011, Yanes et al., 2011). However, very limited evidence of such multi-mode chromatographic separation has been demonstrated, to date, for global metabolite profiling. This requires further investigation to test its applicability. The separation of polar analytes can be further improved with the use of stationary phases having polar functional groups imbedded in the C18 chain for better interaction with polar metabolites. Alternatively, long capillary monolithic columns can also be used as seen with metabolomics study of *Arabidopsis thaliana,* which resulted in detection of several hundred peaks (Tolstikov et al., 2003).

Recently, UPLC system which uses a polar mobile phase, typically consisting of water and acetonitrile has been introduced. UPLC operates at relatively higher flow rate and higher pressure around 15,000 psi and involves the use of a column packed with sub-2-µm stationary phase. Reduction in the particle size of the stationary phase greatly enhances the chromatographic resolution and efficiency and therefore can be applied to resolving complex biological mixtures in non-targeted metabolic profiling (Zhang et al., 2012). Miniaturization, such as reduced column dimension and reduced flow rates, increases the sensitivity in ESI-MS, which offer several application in metabolomics and proteomics as demonstrated with S*hewanella oneidensis* (Bedair and Sumner, 2008a). Recently, Kinetex core-shell columns were introduced by Phenomenex with 2.6 µm core shell particle size which offers reduced backpressure by a factor of around 40% compared to that of 1.7 µm fully porous particles, thereby allowing faster UPLC runs with higher peak capacity and improved peak intensities (Horváth et al., 2010, Kortz et al., 2011).

Very few reports described an optimised approach for efficient retention and separation of polar anionic metabolites, such as organic acids (Burgess et al., 2011). Many organic acids can be easily trapped with negative ion mode. However, in complex mixtures their sensitivity and selectivity is very poor (Xu et al., 2011). Some reports suggest the use of metal chelating agents such as EDTA for efficient separation of polar anionic organic acids using nano-LC-MS technique (Myint et al., 2009). Analysis of highly polar metabolites such as nucleotides and sugar phosphates and metabolites with varying polarity such as amino acids and water soluble vitamins are the most difficult to analyse with LC-MS due to their highly ionic nature. Recent reports suggest use of ion pairing reagents for successful retention and separation of nucleotides using RP-HPLC (Luo et al., 2007b, Park et al., 2011). Recently, diode array detector was employed at 254 nm for the efficient separation of 16 nucleotides using ZIC HILIC column (Chen et al., 2012). Similarly, separation of 12 purine and pyrimidine bases using HILIC was demonstrated with diode array detector (Marrubini et al., 2010).

Most of the optimised methods reported in the literature are specifically effective for separation of hydrophobic and polar metabolites, however no efforts have been taken to optimise the parameters within a specific separation approach to maximise the coverage for major metabolite classes that could be separated under similar isocratic or gradient condition using single stationary phase column. To achieve such kind of comprehensive approach it is essential to optimise chromatographic parameters with either HILIC or RP-HPLC or IP-RP-HPLC based separations using UV detector. We aimed to address this approach in this chapter for selected metabolite classes.

10.2 Material and Methods

10.2.1 Equipments and columns

Equipment's and columns	Supplier
HPLC Agilent 1100 (Coupled with RI and UV detector)	Agilent
HPLC Ultimate 3000	Dionex, LC packing's
Capillary HPLC Ultimate 3000	Dionex, LC packing's
HTC Ultra PTM Discovery system	Bruker Daltonics
ZIC-HILIC (4.6mm X 150mm, 5um 200A)	Merck
ZIC-HILIC Guard column (Peek 20 x 2.1)	Merck
ZIC-HILIC Capillary column (0.3 mm X 150mm, 3.5um 200A)	VWR
Scherzo SM-C18 (2.0mm X 150mm, 3um)	Scherzo
Scherzo SM-C19 (0.3mm X 150mm, 3um)	Scherzo
Kinetex (150 x 2.10 mm 2.6u C18 100A)	Phenomenex

10.2.2 Chemicals and reagents

All chemicals and reagents were obtained from Sigma-Aldrich (Dorset, U.K.), unless stated otherwise.

10.2.3 Preparation of metabolite standards and cocktails

In total 96 metabolites belonging to various metabolite classes were selected and purchased from Sigma-Aldrich (UK). A 100 mM stock of all selected 96 metabolites were prepared individually in appropriate solvent. In total 20 metabolites cocktails at stock concentration of 500 µM, containing varying number of individual metabolites were prepared based on their solvent compatibility and solubility.

10.2.4 Buffer compositions

10.2.4.1 ZIC-HILIC (Isocratic)

Buffer A for ZIC-HILIC	
Acetonitrile	70%
20 mM Ammonium formate in water pH 6.4	30%

10.2.4.2 RP HPLC (Gradient)

Buffer A = 0.1% formic acid in water

Buffer B = 0.1 % formic acid in acetonitrile

10.2.4.3 IP RP HPLC (Gradient)

Buffer 1:

Buffer A for IP RP HPLC	
Sodium-1-hexanesulfonate monohydrate	7 mM
Triethylammonium acetate	1 mM
Buffer B for IP RP HPLC	
Methanol	100%

Buffer 2:

Buffer A for IP RP HPLC	
Sodium-1-hexanesulfonate monohydrate	1 mM
Triethylammonium acetate	1 mM
Buffer B for IP RP HPLC	
Methanol	100%

Buffer 3:

Buffer A for IP RP HPLC (pH 4)	
Sodium-1-hexanesulfonate monohydrate	1 mM
Triethylammonium acetate	1 mM
Buffer B for IP RP HPLC	
Methanol	100%

Buffer 4:

Buffer A for IP RP HPLC (pH 3)	
Sodium-1-hexanesulfonate monohydrate	1 mM
Triethylammonium acetate	1 mM
Buffer B for IP RP HPLC	
Methanol	100%

Buffer 5:

Buffer A for IP RP HPLC (pH 2.5)	
Sodium-1-hexanesulfonate monohydrate	1 mM
Triethylammonium acetate	1 mM
Buffer B for IP RP HPLC	
Methanol	100%

10.2.5 Hydrophilic interaction liquid chromatography (HILIC)

All samples were analysed by HILIC on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was mentioned above and isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was optimised and kept constant, which was 10 µL. Chromatograms were obtained for further analysis using UV detection at a wavelength of either 206, 240 or 254 nm. The obtained chromatograms were processed with an Agilent Chemstation software for further data analysis. The identification of metabolites was achieved by comparing their retention time and mass spectra against known standards.

10.2.6 Reversed phase liquid chromatography (RP HPLC)

All samples were analysed by RP HPLC on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). RP HPLC buffers used were mentioned above and gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Starting at 100% buffer A , the gradient was extended to 25% buffer B in 20 min, followed by a linear increase to 90% buffer B over 1 min at a flow rate of 0.5 mL min^{-1}. Sample volume was kept constant for all the samples which was 8.9 µL controlled by selecting full loop injection, where calculated loop volume was 8.9 µL. Chromatograms were obtained for further analysis using UV detection at a wavelength of either 206, 240, 254, 260 or 270 nm. The identification of metabolites was achieved by comparing their retention time and mass spectra against known standards.

10.2.7 Ion pair reversed phase liquid chromatography (IP RP HPLC)

All samples were analysed by IP RP HPLC on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). IP RP HPLC buffers used were mentioned above and gradient conditions were used as described below,

Gradient condition (1): Starting at 100% buffer A the gradient was increased to 25% buffer B in 20 min, followed by a linear increase to 90% buffer B over 1 min at a flow rate of 0.5 mL min^{-1}.

Gradient condition (2): Starting at 100% buffer A the gradient was increased to 47% buffer B in 12 min, followed by a linear increase to 98% buffer B over 3 min at a flow rate of 0.75 mL min^{-1}. Sample volume was kept constant for all the samples which was 8.9 µL controlled by selecting full loop injection where calculated loop volume was 8.9 µL. Chromatograms were obtained for further analysis using UV detection at a wavelength of either 206, 254, 260 or 270 nm.

10.2.8 Direct infusion mass spectrometry

Purified fractions of metabolites from metabolite cocktails were collected from an Agilent 1100 HPLC using fraction collector and infused directly on an ion trap mass spectrometer (HCT Ultra PTM Discovery System, Bruker Daltonics) at a flow rate of 180 µL min^{-1} for metabolite identification based on mass spectrum. Mass spectrometer was set to acquire the data in both the positive and negative ion mode, with a selected mass range of 50 – 600 m/z. The capillary voltage was kept at -2500 V to maintain capillary current between 30 – 50 nA. The nebulizer gas used was nitrogen at a flow rate of 10 L min^{-1} and at a gas pressure 0.4 bar.

10.3 Results and Discussion

As discussed and demonstrated in Chapter 6, excellent coverage for most of the metabolite classes was observed with GC-MS based analysis. However, recovery of nucleotides/nucleosides/nucleobases, pigments and sugar phosphates was severely compromised, requiring development of optimized protocol with alternative analytical platform such LC-MS for coverage of these classes of metabolites, which we propose to address in this Chapter. The aim of this investigation was to evaluate which of the available HPLC approaches were most suitable for the specific classes of metabolites which will aid in efficient retention and separation. Based on the available literature, we have specifically directed our efforts towards optimisation, evaluation and comparison of HILIC and IP-RP-HPLC separation for the better retention and separation of selected metabolite classes which includes amino acids, organic acids, water soluble vitamins, nucleotides, nucleosides and nucleobases.

10.3.1 Hydrophilic interaction liquid chromatography (HILIC)

10.3.1.1 Separation of non-derivatized amino acids under isocratic conditions

Initially, twenty individual amino acids listed in table 10.2, at a concentration of 100uM were injected (100 µL) on to a ZIC-*p*HILIC column.

Preliminary optimisation of the concentration of the organic solvent (acetonitrile) in mobile phase was conducted. The results showed acetonitrile concentration at 70% resulted in better retention of amino acids compared to acetonitrile concentration at 60%, 80% or 90%. Therefore for all further analysis concentration of acetonitrile in the mobile phase was kept constant (70%). Our results were in contrast to that of Chen and co-workers (Chen et al., 2012). Possible reason behind this contradictory finding might be use of different detector system (Diode array) used in previous study. For HILIC separations, commonly used ammonium salts include ammonium acetate or formate due to their better compatibility with the subsequent MS analysis. For this reason ammonium formate at 20 mM concentration (pH of 6.4) was selected for further analysis, as recommended by column supplier (Merck SeQuant) for achieving best column performance. Preliminary optimisation of varying pH condition from 3 to 4.5 and 6.4 did not show much variation in the results. Therefore we selected pH 6.4 as the optimal for analysis, as it can be easily obtained with 20 mM ammonium formate solution and does not require any further modification of mobile phase for pH adjustment. We could not evaluate the influence of temperature on retention and separation of metabolites as the Agilent 1100 HPLC system did not have the column oven compartment. However, the laboratory was temperature controlled at 25°C, which was suggested as an optimal temperature for HILIC separations by Chen and co-workers (Chen et al., 2012).

Table 10.1 List of 20 amino acids prepared individually at a concentration of 100 µM

Leucine	Proline	Tyrosine	Valine
Alanine	Threonine	Cysteine	Aspartic acid
Glycine	Serine	Glutamine	Phenylalanine
Isoleucine	Histidine	Lysine	Glutamic acid
Methionine	Arginine	Tryptophan	Asparagine

Out of 20 amino acids, only 15 amino acids were well detected using preliminary optimised conditions at 206 nm with a good reproducibility and peak intensities. Phenylalanine and valine were only detected at 254 nm, whereas lysine and arginine were not detected with both the wavelengths. Representative chromatograms for few amino acids separated and detected are displayed in figure 10.1.

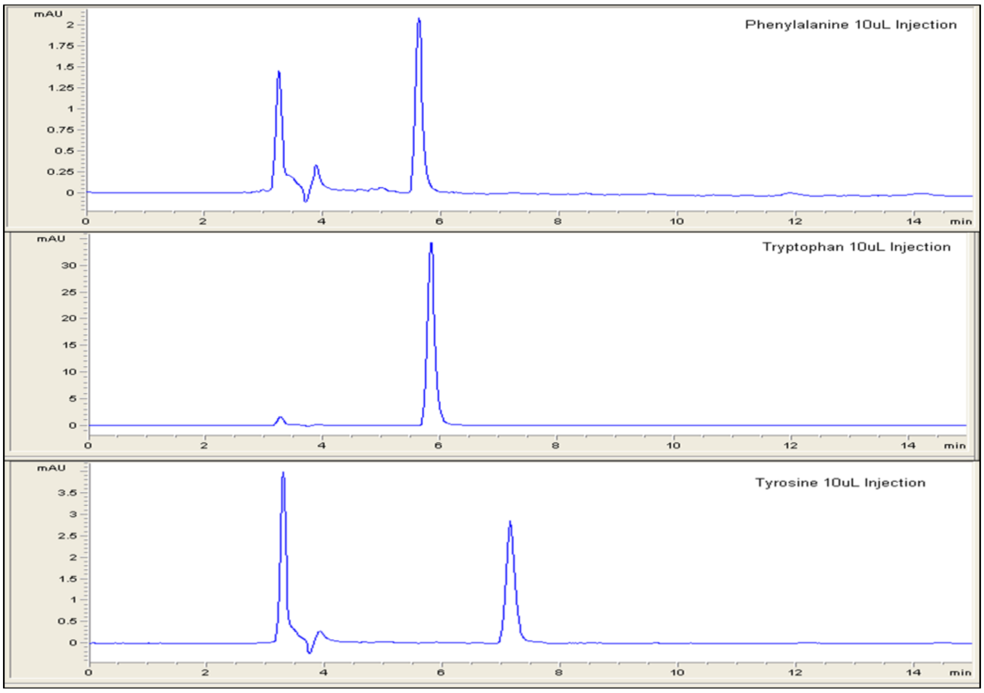


Figure 10.1 Representative individual chromatograms obtained for phenylalanine, tryptophan and tyrosine (100 µM) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of either 206, 240 or 254 nm.

Later to evaluate if we can separate these amino acids from their mixtures, cocktail of three aliphatic amino acids (leucine, alanine and glycine) was prepared at a concentration of 100 µM and injected on to HILIC column. All the three aliphatic amino acids showed better separation onto HILIC column and detected at 206 nm (figure 10.2).

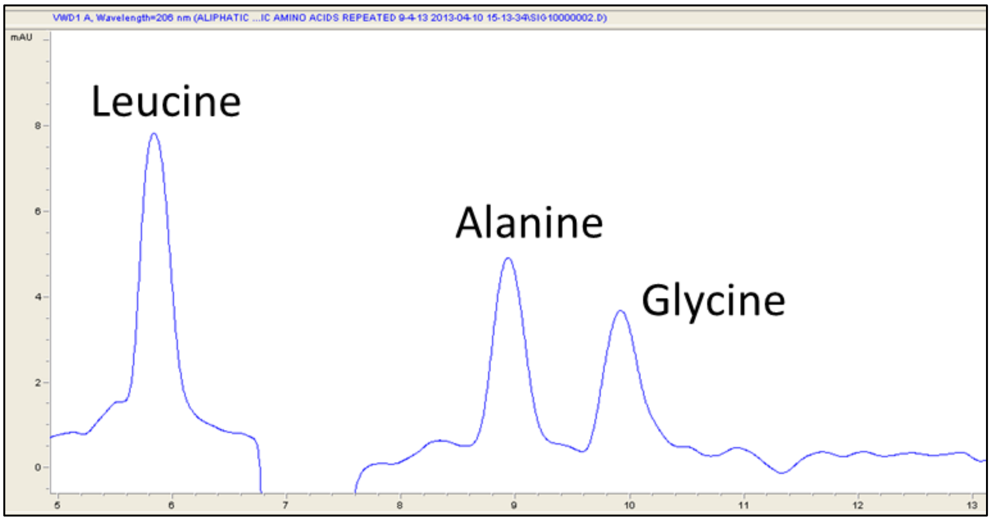


Figure 10.2 Representative chromatogram obtained for mixture of three aliphatic amino acids (100 µM) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of either 206, 240 or 254 nm.

A further cocktail of seven aliphatic amino acids (leucine, isoleucine, methionine, valine, proline, alanine and glycine) was prepared at 100 µM concentration, which resulted in separation of only four amino acids, as methionine resulted in a very broad peak which suppressed the peaks for valine, leucine and isoleucine. To test whether actually broader methionine peak is suppressing the other peaks, we ran the mixture of six aliphatic amino acids without methionine, which resulted in good separation for all the six amino acids with good intensities. The broad peak of methionine might be due to its higher polarity resulting in higher retention on to a HILIC phase. To minimise the influence of broader methionine peak, we further prepared cocktail of seven aliphatic amino acids at 100 µM concentration except that of methionine whose concentration was kept to 25 µM. The results showed better separation of all the five aliphatic amino acids, still no peaks for valine and isoleucine was seen whereas good separation and peak shape was observed for leucine. As demonstrated with individual amino acids run, valine was only detected at 254 nm, which might be the probable reason it is not being detected in the cocktail. Later, to investigate the optimal wavelength for detection of these amino acids, the above set of experiments were repeated at a wavelength of 254 nm. Except valine, none of the amino acids were detected at this wavelength. To test the optimal wavelength for the remaining amino acids, we ran the remaining individual amino acid separately at both the wavelengths, results of which further confirmed no detection for rest of the amino acids at 254 nm, except for that of phenylalanine. All the remaining amino acids were detected well at 206 nm due to the absorption of carboxyl group at this wavelength, except for lysine and arginine. Recently, arginine detection was reported at 190 nm using capillary electrophoresis in human blood plasma (Forteschi et al., 2014). The above analysis confirmed that the optimal wavelength for majority of amino acids is 206 nm except for that of valine and phenylalanine (due to presence of benzene ring in phenylalanine). A summary of retention time and wavelength at which all the individual amino acids were detected and showed a good peak shape is summarised in table 10.2,

Table 10.2 Summary of retention time and wavelength at which individual amino acids were detected

Number	Amino acid	Conc.	Wavelength	Retention time	Hydrophobicity at pH 7	Polarity
1	Lysine	100 uM	206,254	No retention	Hydrophillic	
2	Arginine	100 uM	206,254	No retention		
3	Phenylalanine	100 uM	254	5.693	Very Hydrophobic	
4	Tryptophan	100 uM	206	5.89		Less polar
5	Leucine	100 uM	206	6.25		
6	Isoleucine	100 uM	206	6.65		↓
7	Methionine	25 uM	206	6.77		
8	Tyrosine	100 uM	206	7.172	Hydrophobic	
9	Valine	100 uM	254	7.3	Very Hydrophobic	
10	Histidine	100 uM	206	7.77	Neutral	
11	Proline	100 uM	206	9.15	Hydrophillic	
12	Aspartic acid	100 uM	206	9.501		
13	Glutamic acid	100 uM	206	9.719		
14	Threonine	100 uM	206	10.02	Neutral	
15	Alanine	100 uM	206	10.17	Hydrophobic	
16	Glutamine	100 uM	206	10.74	Hydrophillic	
17	Asparagine	100 uM	206	11.247		More Polar
18	Glycine	100 uM	206	11.39		
19	Serine	100 uM	206	11.52		

As it can be seen from table 10.2, based on the working principle of HILIC, the polar amino acids were well retained on HILIC zwitterionic (both charged and neutral) phase compared to less polar amino acids which eluted earlier. As all the amino acids summarised in table 10.2 were run individually we subsequently injected a cocktail of above amino acids which resulted in poor resolution of individual amino acids as expected. We divided all 17 amino acids into 2 sets each containing 9 and 8 amino acids based on preliminary results (data not shown) as displayed in table 10.3. Arginine and lysine were excluded from this as they were not detected with the above experiments.

Table 10.3 List of amino acids included in 9 and 8 mix cocktails at a concentration of 100 µM

9 Mix		8 Mix	
1	L-phenylalanine	1	L-tryptophan
2	L-leucine	2	L-iso-leucine
3	L-methionine	3	L-tyrosine
4	L-arginine	4	L-histidine
5	L-proline	5	L-aspartic acid
6	L-glutamic acid	6	L-threonine
7	L-alanine	7	L-glutamine
8	L-asparagine	8	Glycine
9	L-serine		

The above two cocktails were then injected individually, the results of which showed better separation of all the 8 amino acids from 8 mix, whereas in case of 9 mix only 7 amino acids were separated with no separation for glutamic acid and asparagine. In order to improve the resolution, the volume of sample injected was optimised from 100 µL to 10 µL injection which improved the resolution greatly. The results of this investigation are summarised in figure 10.3.

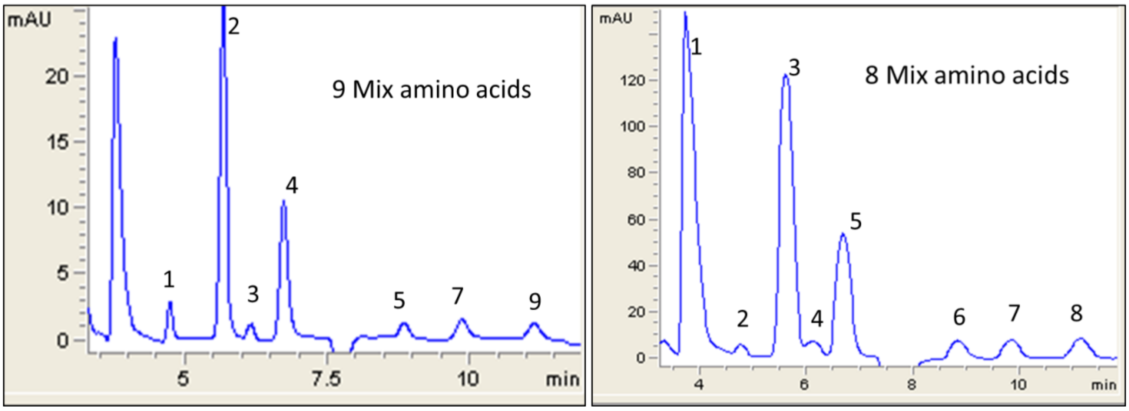


Figure 10.3 Representative chromatogram obtained for mixture of nine (left) and eight (right) amino acids (100 µM) (See Table 10.3) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of either 206, 240 or 254 nm.

As the identification of amino acids from their mixtures was purely based on reference standards, further validation with respect to metabolite identification was achieved by collecting the fractions of these peaks using fraction collector and injecting them directly into an ion trap MS (Bruker Daltonics) under conditions as described in section 10.2.8. The representative mass spectra for few amino acid fractions from 9 and 8 mix are displayed in figure 10.4 and 10.5.

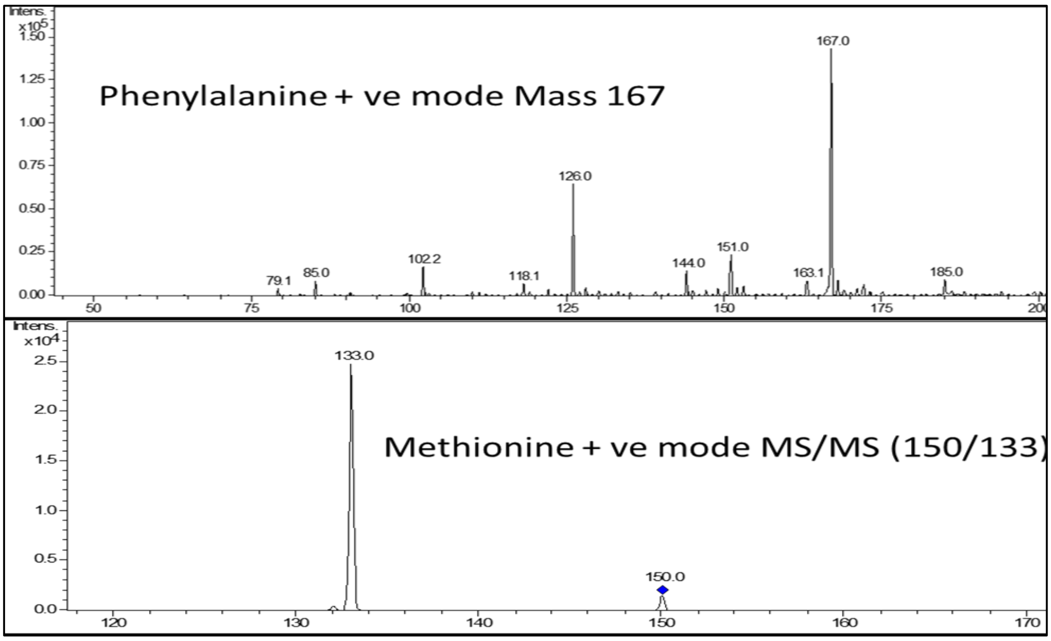


Figure 10.4 Mass spectra obtained for phenylalanine (top) and methionine (bottom), where purified HILIC fractions were collected from an Agilent 1100 HPLC using fraction collector and infused directly on an ion trap mass spectrometer (HCT Ultra PTM Discovery System, Bruker Daltonics) at a flow rate of 180 µL min^{-1} and a mass range of 50 – 600 m/z. The capillary voltage was kept at -2500 V to maintain capillary current between 30 – 50 nA.

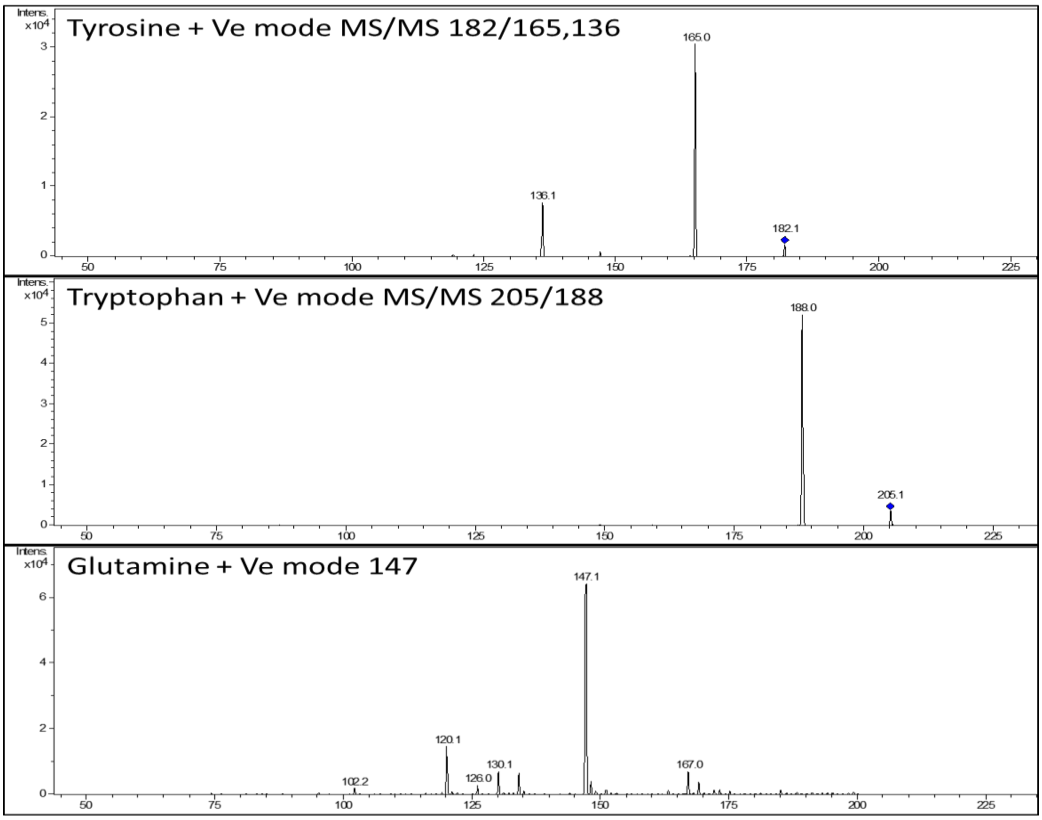


Figure 10.5 Mass spectra obtained for tyrosine (top), tryptophan (middle) and glutamine (bottom), where purified HILIC fractions were collected from an Agilent 1100 HPLC using fraction collector and infused directly on an ion trap mass spectrometer (HCT Ultra PTM Discovery System, Bruker Daltonics) at a flow rate of 180 µL min^{-1} and a mass range of 50 – 600 m/z. The capillary voltage was kept at -2500 V to maintain capillary current between 30 – 50 nA.

As can be seen from figure 10.4 and 10.5, good intensities were obtained for amino acid fractions resulting from their mixtures separated on ZIC-*p*HILIC column under isocratic conditions. In total we have separated 15 amino acids so far in two separate runs of 9 min each without requiring any derivatization. Further optimisation with respect to volume of sample injection, flow rate used and wavelength employed resulted in better resolution of these amino acids with good intensities obtained using direct infusion MS. Based on these preliminary offline LC-MS optimisation, future work will be directed towards separation of all these amino acids from mammalian and microalgal samples in a single run using online LC-MS technique.

10.3.1.2 Separation of organic acids under isocratic conditions

Similar conditions as that of applied for amino acids were employed to evaluate the efficiency of above optimised isocratic conditions for separation of 22 organic acids as summarised in table 10.4.

Table 10.4 List of 22 organic acids prepared individually at a concentration of 100 µM

	List of 22 Organic acids		
1	L-ascorbic acid	12	Isocitric acid Trisodium
2	Fumaric acid	13	Quinic acid
3	Oxalic acid	14	D(-)Iso-ascorbic acid
4	Citric acid	15	Shikimic acid
5	Malonic acid	16	Itaconic acid
6	Succinic acid	17	Phytic acid
7	Malic acid	18	Adipic acid
8	Tartaric acid	19	Nicotinic acid (Vitamin B3)
9	Gamma aminobutyric acid (GABA)	20	4-aminobenzoic acid (PABA)
10	Maleic acid	21	Benzoic acid
11	D-Lactic acid	22	Folic acid

All the organic acids were injected onto a ZIC-HILIC column using an Agilent 1100 HPLC under similar isocratic conditions as described for amino acids at different wavelengths (206, 240 and 254 nm). Results showed retention and detection for only three organic acids at 206 and 240 nm (oxalic, lactic and quinic acid) whereas eight organic acids were retained and detected at 254 nm. The representative chromatograms obtained for few organic acids at 254 nm are displayed in figure 10.6.

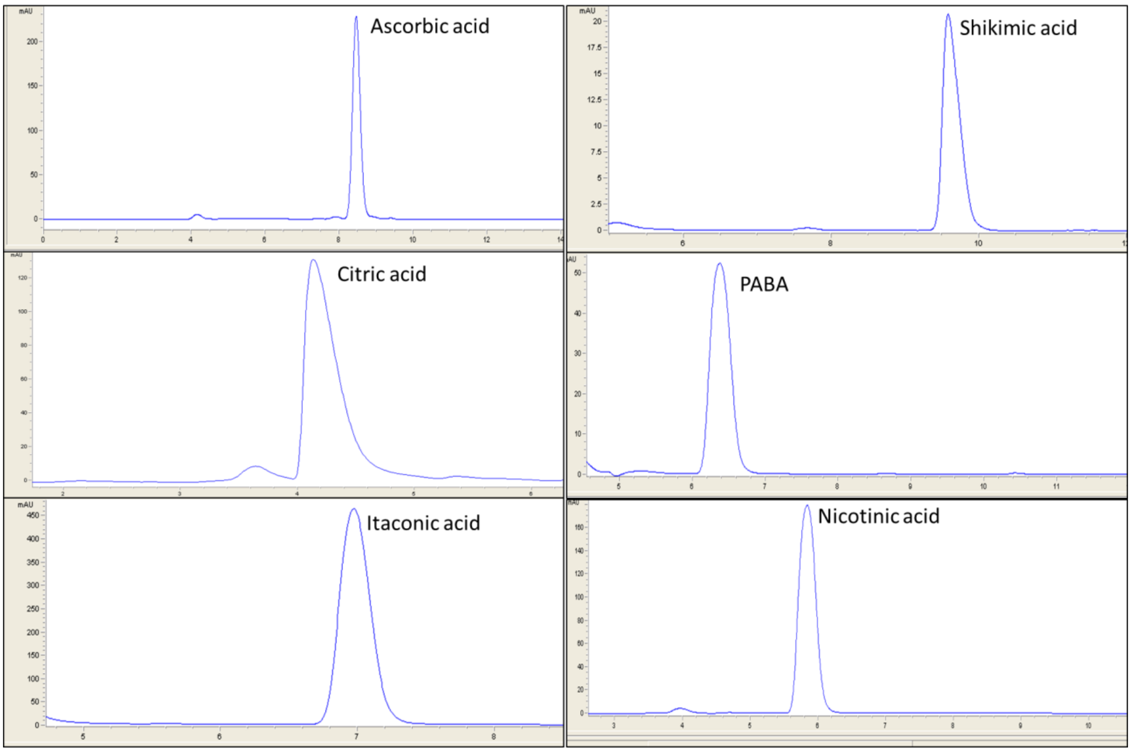


Figure 10.6 Representative individual chromatograms obtained for individual organic acids (100 µM) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

Future work involves further optimisation of HILIC conditions where efforts will be directed towards separation of all the selected 22 organic acids from their mixture in a single run.

10.3.1.3 Separation of nucleotides/nucleosides/nucleobases under isocratic conditions

Poor coverage of GC-MS for this class of metabolites demands an alternative analytical approach which can provide better coverage for these metabolite class. To evaluate the efficiency of HILIC principle in separating this class of metabolites under similar optimised isocratic condition as described above, we have initially considered 18 metabolites of this class which are summarised in table 10.5.

Table 10.5 List of 18 metabolites belonging to nucleotides/nucleosides/nucleobases class prepared individually at a concentration of 100 µM

Number	Nucleotides/nucleosides/nucleobases	Class/function
1	2-Hydroxypyridine (2-pyridone)	
2	Riboflavine	Vit B2,
3	FAD	Redox cofactor
4	Uracil	Pyrimidine
5	Adenosine	Purine nucleoside
6	Xanthine	Purine base
7	Hypoxanthine	Purine derivative
8	Cytosine	Pyrimidine
9	AMP	Nucleotides (Cell signalling)
10	Tyrosine	Amino acid
11	NADH	Coenzyme (2 nucleotides)
12	ATP	Nucleoside triphosphates
13	GMP	Nucleotides (Cell signalling)
14	GDP	Nucleoside diphosphate
15	GTP	purine nucleoside triphosphate
16	D-pantothenic acid hemicalcium salt	Vit B5, Amide
17	Guanine (Forms white suspension)	Purine
18	Adenine	Purine

Among these 18 metabolites, guanine tends to from a white suspension when mixed with other metabolites in presence of water as a solvent, therefore excluded from further investigation. All the remaining individual metabolites were run individually under similar conditions described in section 10.3.1.1. Excellent retention for all the metabolite with good resolution and intensities were obtained when ran individually, except for D-pantothenic acid and adenine. To further evaluate separation efficiency with good resolution of these metabolites from mixture, two sets of cocktails were prepared each containing 10 and 12 metabolites as summarised in table 10.6, and run under similar conditions. The results of this investigation are summarised in figure 10.7.

Table 10.6 List of 10 and 12 mix of nucleotides/nucleosides/nucleobases prepared at a concentration of 100 µM

Number	10 Mix	Number	12 Mix
1	2-Hydroxypyridine (2-pyridone)	1	2-Hydroxypyridine (2-pyridone)
2	Uracil	2	Uracil
3	Adenosine	3	Adenosine
4	Hypoxanthine	4	Xanthine
5	Cytosine	5	AMP
6	AMP	6	Hypoxanthine
7	NADH	7	Tyrosine
8	ATP	8	Cytosine
9	GDP	9	ATP
10	GTP	10	NADH
		11	GMP
		12	GDP

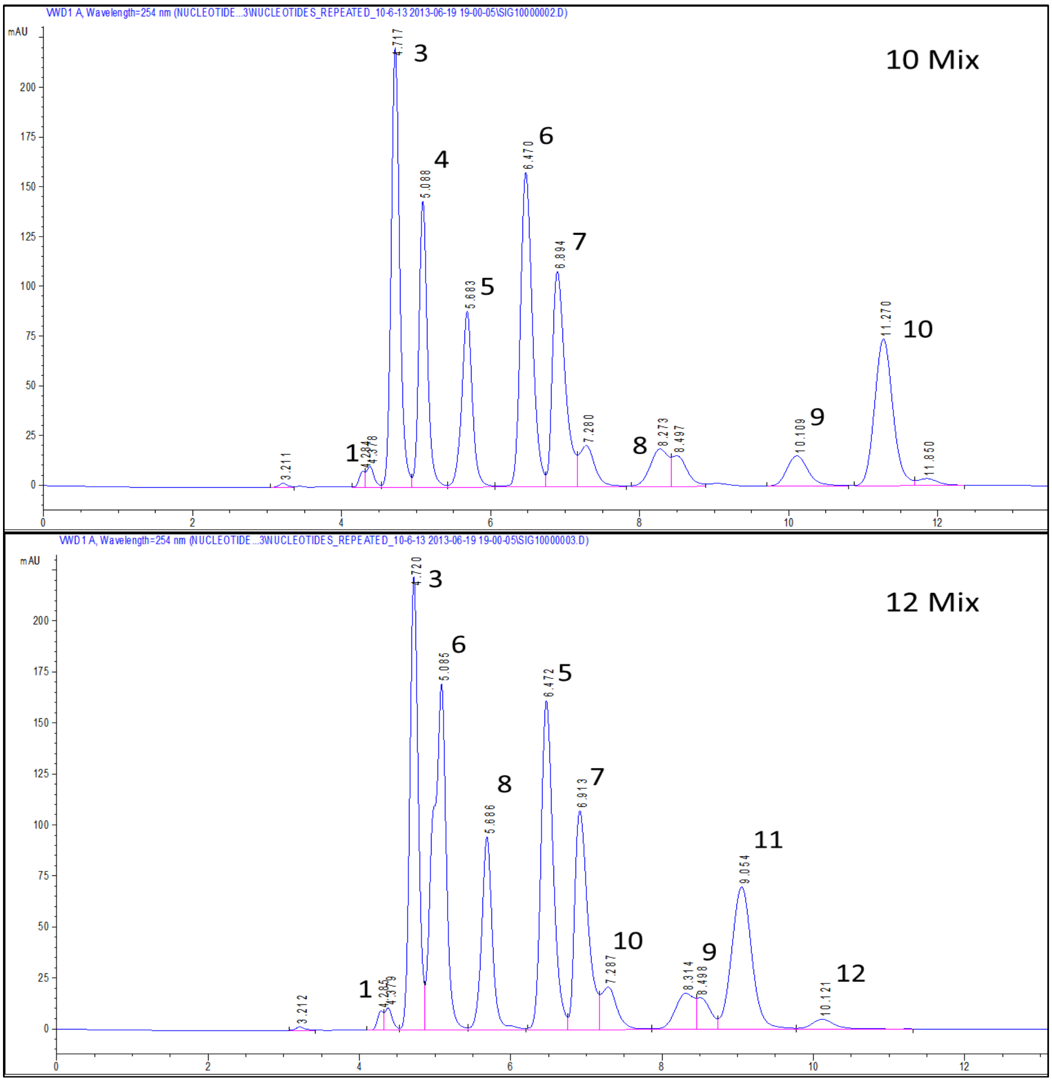


Figure 10.7 Representative chromatograms obtained for mixture of ten (top) and twelve (bottom) nucleotides/nucleosides/nucleobases (100 µM) (See Table 10.6) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

All the metabolites from both the 10 and 12 mix were well retained and separated with a good resolution and intensities except for that of uracil, where peak for uracil seems to be masked by broader peak of adenosine. Similarly the peak for xanthine seems to be overlapped with that of hypoxanthine. In both the mixtures, ATP did not produce good detectable peak on HILIC platform (8 and 9 respectively). To improve the coverage and for better resolution we carefully selected 14 metabolites and cocktail of 14 metabolites was prepared at concentration of 100 μM. 10 μL of sample was injected onto a ZIC-*p*HILIC column using an Agilent 1000 HPLC under similar isocratic conditions as described above. Guanine, adenine, D-pantothenic acid were excluded from 14 mix as no detection for these metabolites were observed with preliminary investigation, whereas AMP was lost during sample preparation, thus excluded. The results of this investigations are summarised in figure 10.8,

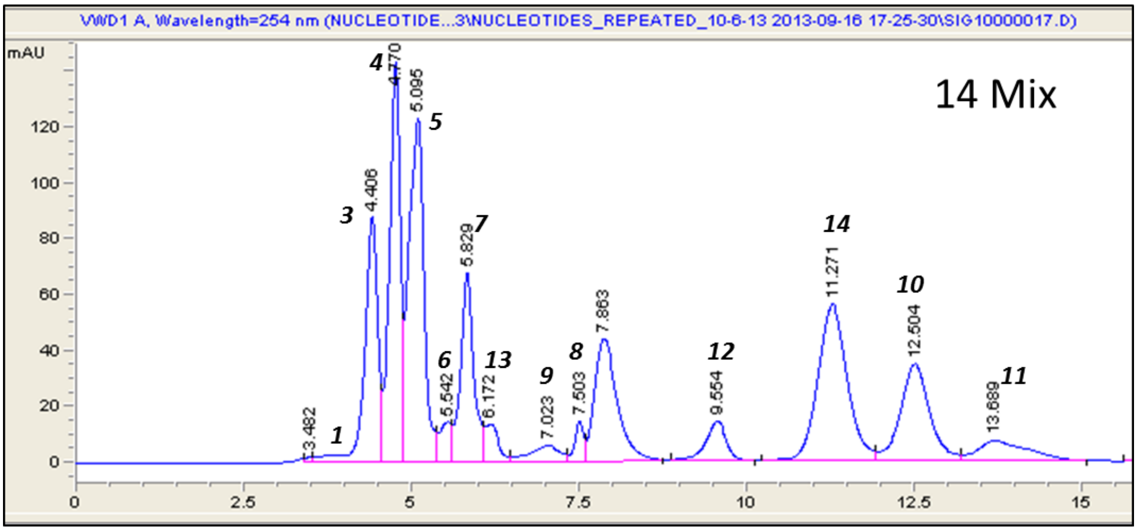


Figure 10.8 Representative chromatogram obtained for mixture of fourteen nucleotides/nucleosides/nucleobases (100 μM) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 μL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

Superior separation for all the 13 metabolites from 14 mix except for uracil was achieved as can be seen in figure 10.8, under isocratic conditions using ZIC-HILIC column. Due to time constraints, we just focused on identification of these metabolites based on reference standards and their obtained retention time under similar conditions.

10.3.1.4 Separation of water soluble vitamins under isocratic conditions

In total eight water soluble vitamins were selected for further investigations as listed in table 10.7. All the individual water soluble vitamins were initially run separately to obtain reference retention time and to test whether they can be separated well under optimised isocratic conditions. Later all the eight water soluble vitamins were mixed together (8 Mix) at 100 µM concentration and run on a ZIC-*p*HILIC column using an Agilent 1100 HPLC. The results of these investigations are summarised in figure 10.9.

Table 10.7 List of eight water soluble vitamins

Number	Water soluble vitamins
1	Ascorbic acid
2	Nicotinic acid
3	Nicotinamide
4	Pyridoxal
5	Folic acid
6	Pyridoxamine
7	Riboflavin
8	Thiamine

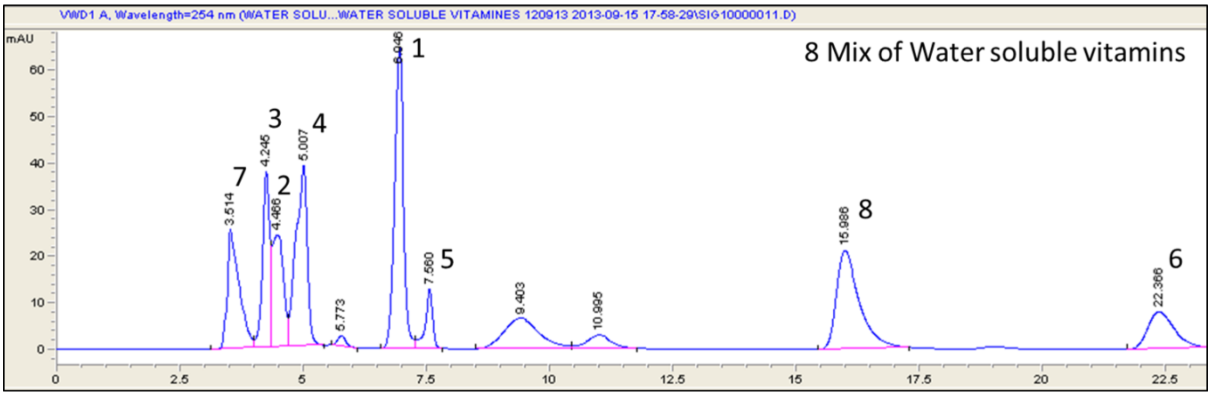
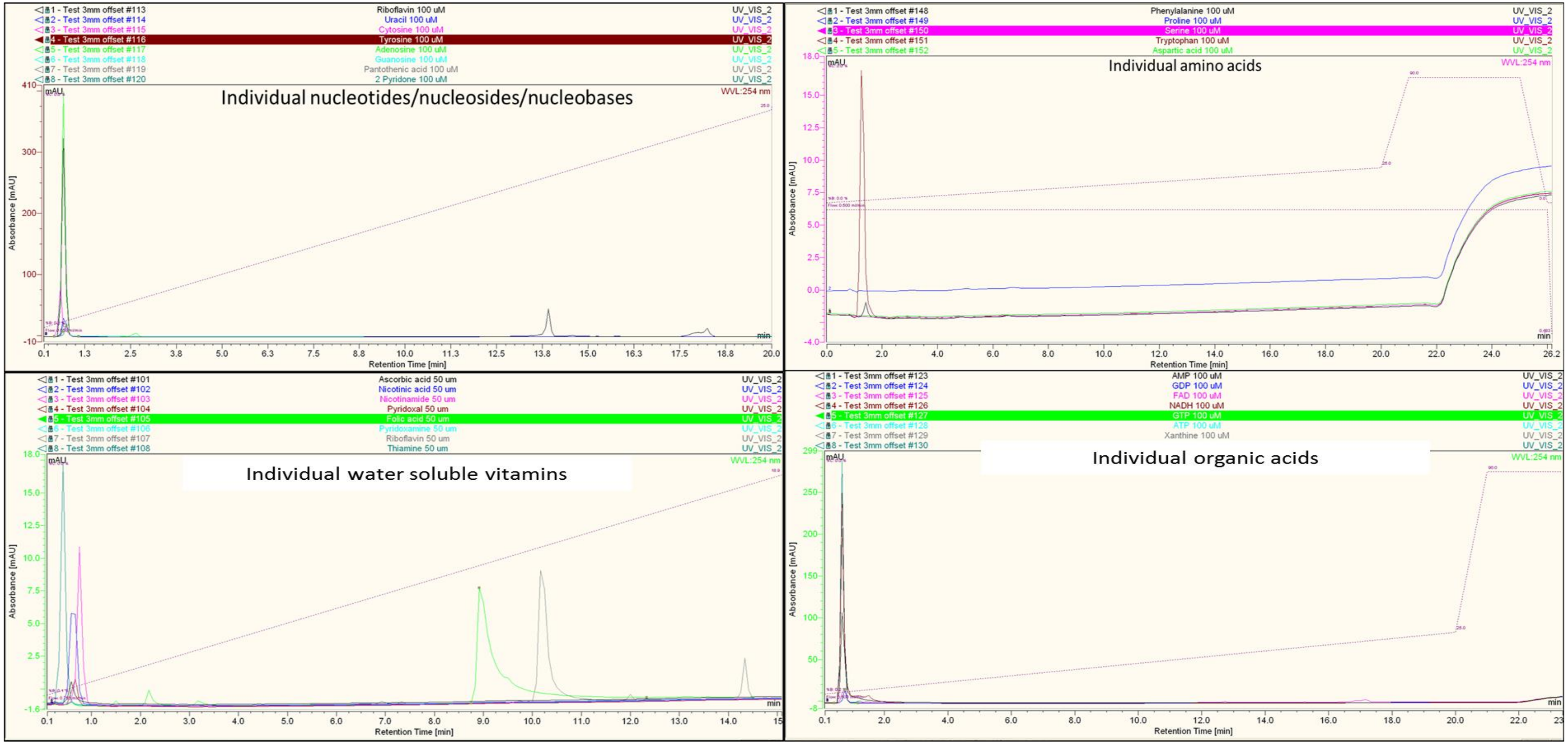


Figure 10.9 Representative chromatogram obtained for mixture of eight water soluble vitamins (100 µM) (See Table 10.7) by HILIC separation on an Agilent 1100 HPLC using ZIC-*p*HILIC column (4.6mm X 150mm, 5um 200A) coupled with the ZIC-HILIC Guard column (Peek 20 x 2.1). HILIC Buffer used was 30% 20 mM ammonium formate in water and 70% acetonitrile (pH 6.4). Isocratic conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 10 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

As can be seen from figure 10.9, superior separation for all the water soluble vitamins was achieved with ZIC-*p*HILIC column. In summary, we have optimised the HPLC conditions with ZIC-*p*HILIC column for important selected metabolite classes which forms important part of the metabolome in organisms and have significant impact in quantification of these metabolites in pathways based analysis.

10.3.2 Reversed phase HPLC (RP HPLC)

Individual metabolites belonging to four different metabolite classes as discussed above with HILIC separations were run individually. As expected, none of the metabolites belonging to four different metabolite classes were retained on RP-C18-column which requires an addition of an ion pair reagent for efficient retention and separation of these metabolite classes as discussed in section 10.1.


Individual nucleotides/nucleosides/nucleobases

Individual amino acids

Individual water soluble vitamins

Individual organic acids

Figure 10.10 Representative individual chromatograms obtained for individual nucleotides/nucleosides/nucleobases, amino acids, water soluble vitamins and organic acids (100 µM) by RP-HPLC separation on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). RP-HPLC Buffers used: 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of either 206, 240, 254, 260 or 270 nm.

10.3.3 Ion pair reversed phase liquid chromatography (IP RP HPLC)

The concentration of an ion pair reagent and pH of the mobile phase employed has greater influence on the retention and separation of metabolites using IP-RP-HPLC, by influencing solute ionization. The influence of these condition in separation of organic acids, water soluble vitamins and nucleotides/nucleosides/nucleobases was evaluated. Basic IPR (cationic), Sodium-1-hexanesulfonate monohydrate (aliphatic sulfonic acid) was purposely selected for separation of both cationic and anionic water soluble vitamins and anionic polar organic acids.

10.3.3.1 Separation of organic acid

As can be seen from figure 10.11, better retention of organic acids was observed with 7 mM IPR, whereas IPR concentration at 1 mM, all the organic acids eluted earlier at the same retention time. The straightforward reason behind increased analyte retention at increased concentration of IPR is mainly due to the fact that, the effective capacity of the column is mainly governed by the IPR, therefore decrease in IPR concentration resulted in poor retention of organic acids. However, for future MS based analysis higher concentration of IPR such as 7 mM will result in severe ion suppression and therefore not recommended. Hence further optimisation was carried out at varying the pH conditions with 1 mM IPR. Despite changing the pH of the buffer from 4 to 3 and 2.5, no improvement in separation was achieved as can be seen from figure 10.12. Possible reason might be, lowering the pH, might have reduced ionization of organic acids and thereby their interaction with the IPR, resulting in poor retention and separation.

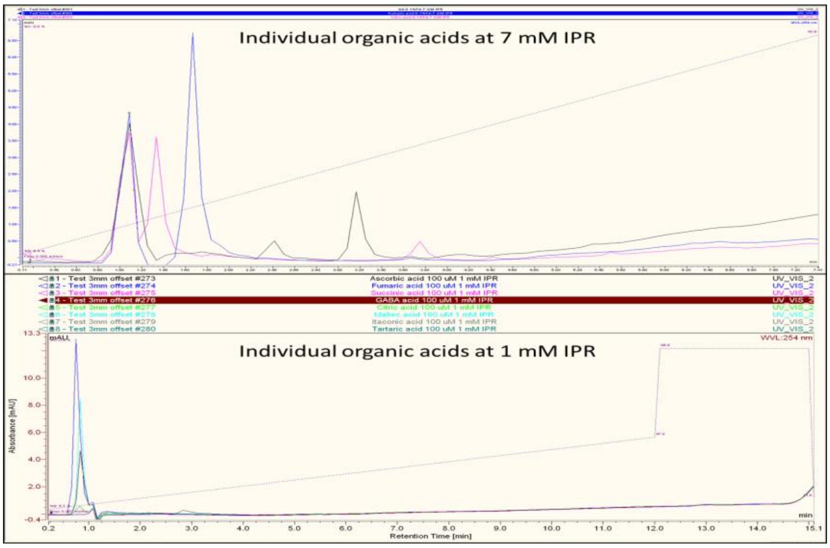


Figure 10.11 Efficiency of IP-RP-HPLC separations in retention and separation of organic acids with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at two different concentrations (7 mM and 1 mM). Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

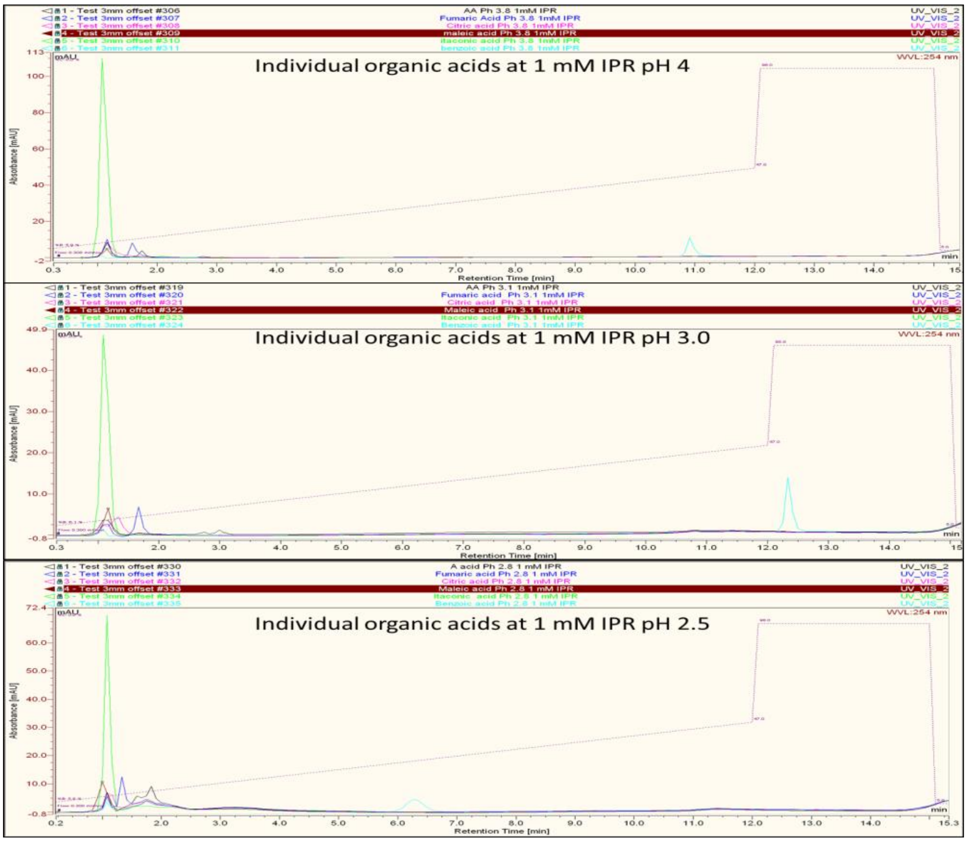


Figure 10.12 Efficiency of IP-RP-HPLC separations in retention and separation of organic acids at varying pH conditions with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at 1 mM concentration. Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

10.3.3.2 Separation of nucleotides/nucleosides/nucleobases

Initially all the individual nucleotides were run at IPR concentration of 7 mM, followed by IPR concentration of 1 mM. The use of 1mM IPR, resulted in better retention and separation for only few metabolites which include FAD and cytosine. All the remaining metabolites seemed to elute earlier between 1 to 2.5 min retention time windows. The results of this investigation are summarised in figure 10.13. Further analysis with varying pH conditions resulted in no improvement in the retention and separation of these metabolite class. At extremely lower pH all the metabolites seems to be eluting earlier compared to that of higher pH conditions (figure 10.14). From these observations it is clear that, lowering the pH reduces ionisation of these metabolites and thereby their interaction with the IPR, resulting in poor retention.

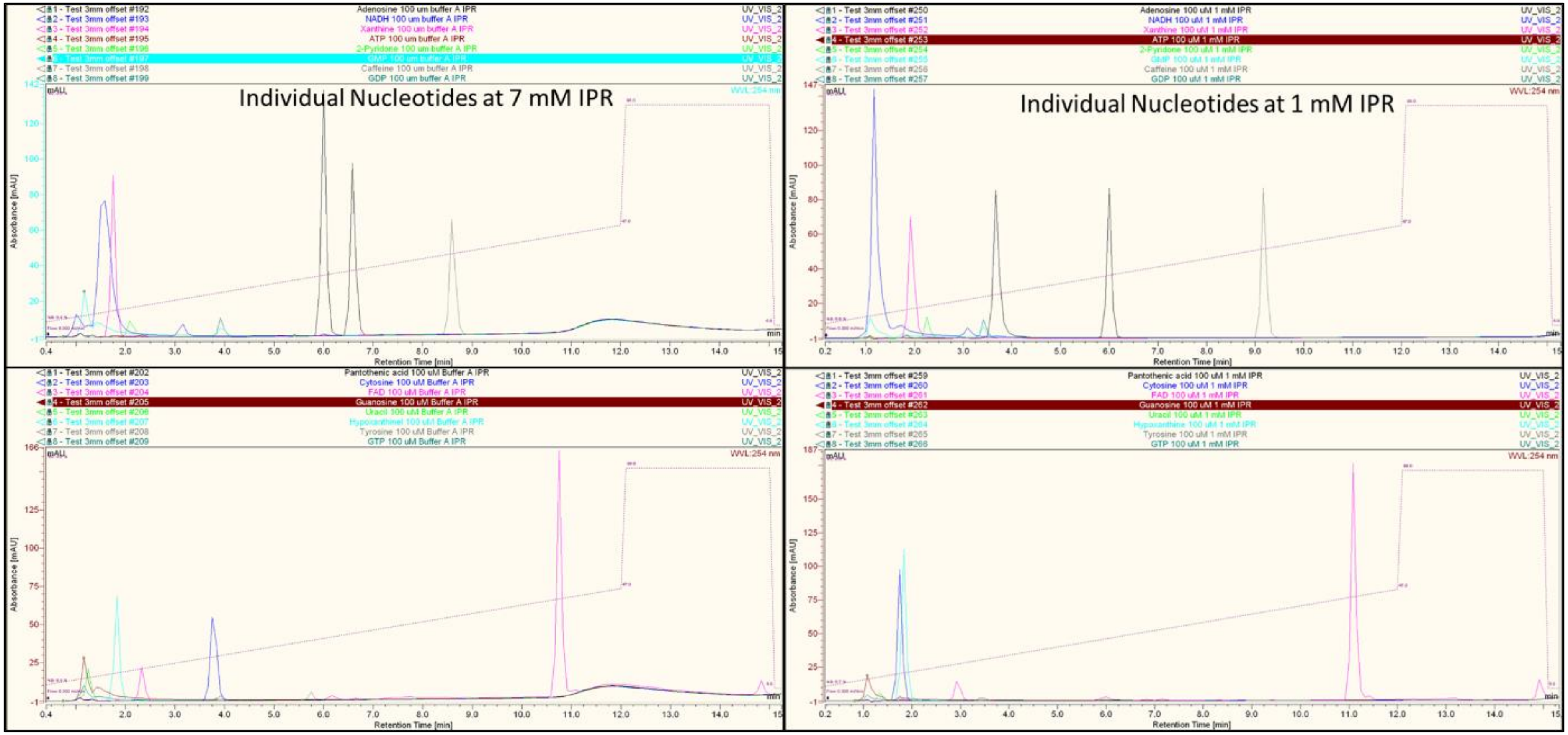


Figure 10.13 Efficiency of IP-RP-HPLC separations in retention and separation of individual nucleotides, nucleosides and nucleobases with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at two different concentrations (7 mM (left) and 1 mM (right)). Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

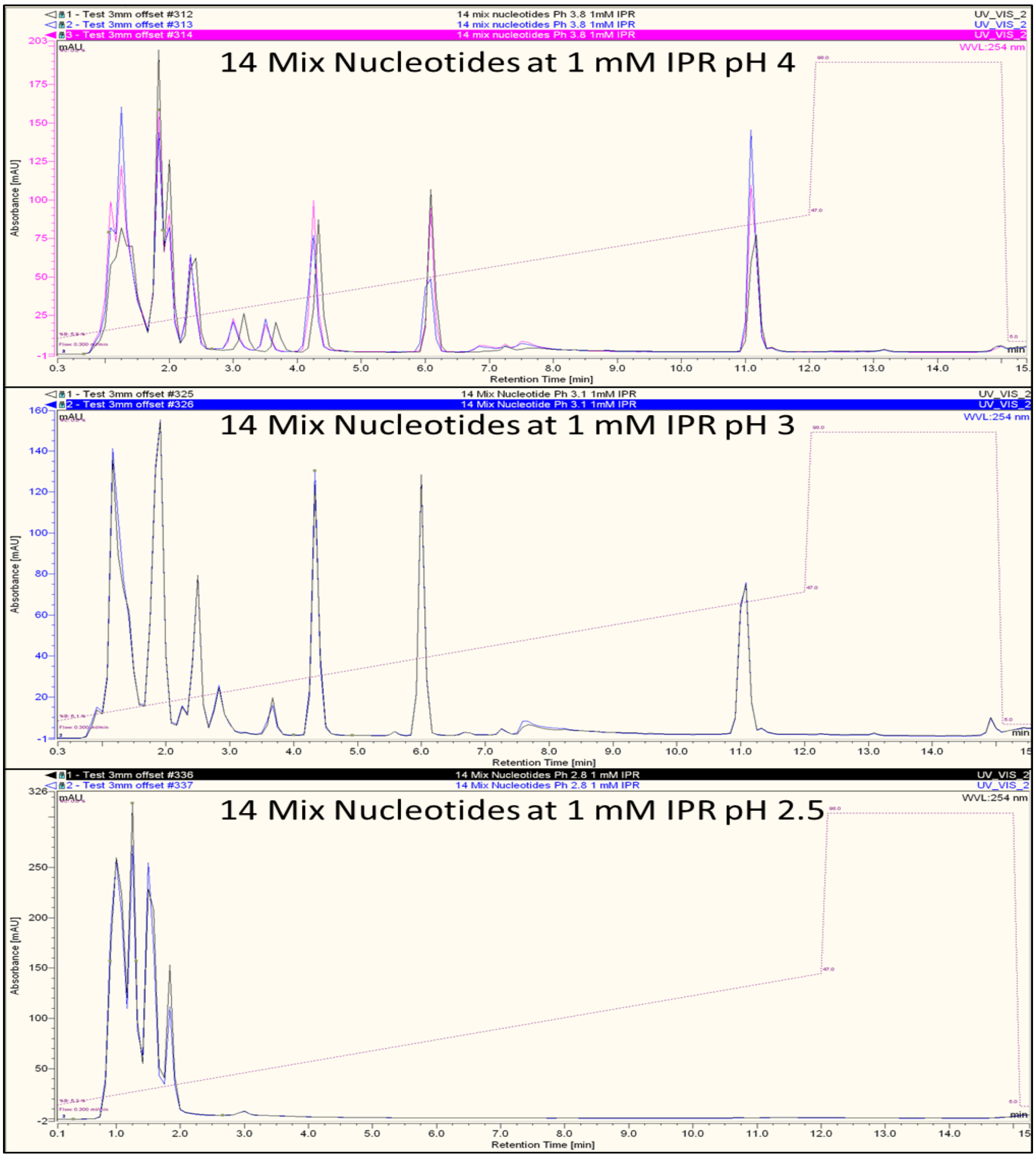


Figure 10.14 Efficiency of IP-RP-HPLC separations in retention and separation of nucleotides, nucleosides and nucleobases from their 14 mix cocktail at varying pH conditions with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at 1 mM concentration. Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

10.3.3.3 Separation of water soluble vitamins

Initially all the individual water soluble vitamins were run IPR concentration at 7 mM followed by IPR concentration at 1 mM, which resulted in better retention for all the metabolites at 7 mM IPR concentration. With 1 mM IPR all the remaining metabolites seems to elute earlier compared to that of 7 mM IPR. The straightforward reason behind increased analyte retention at increased concentration of IPR is mainly due to the fact that, the effective capacity of the column is mainly governed by an ion pair reagent, therefore decrease in IPR concentration resulted in poor retention of water soluble vitamins. The results of this investigation are summarised in figure 10.15. Further analysis with varying pH conditions resulted in no improvement in the separation of these metabolite class (figure 10.16). At extremely lower pH all the metabolites seems to be eluting earlier compared to that of higher pH conditions.

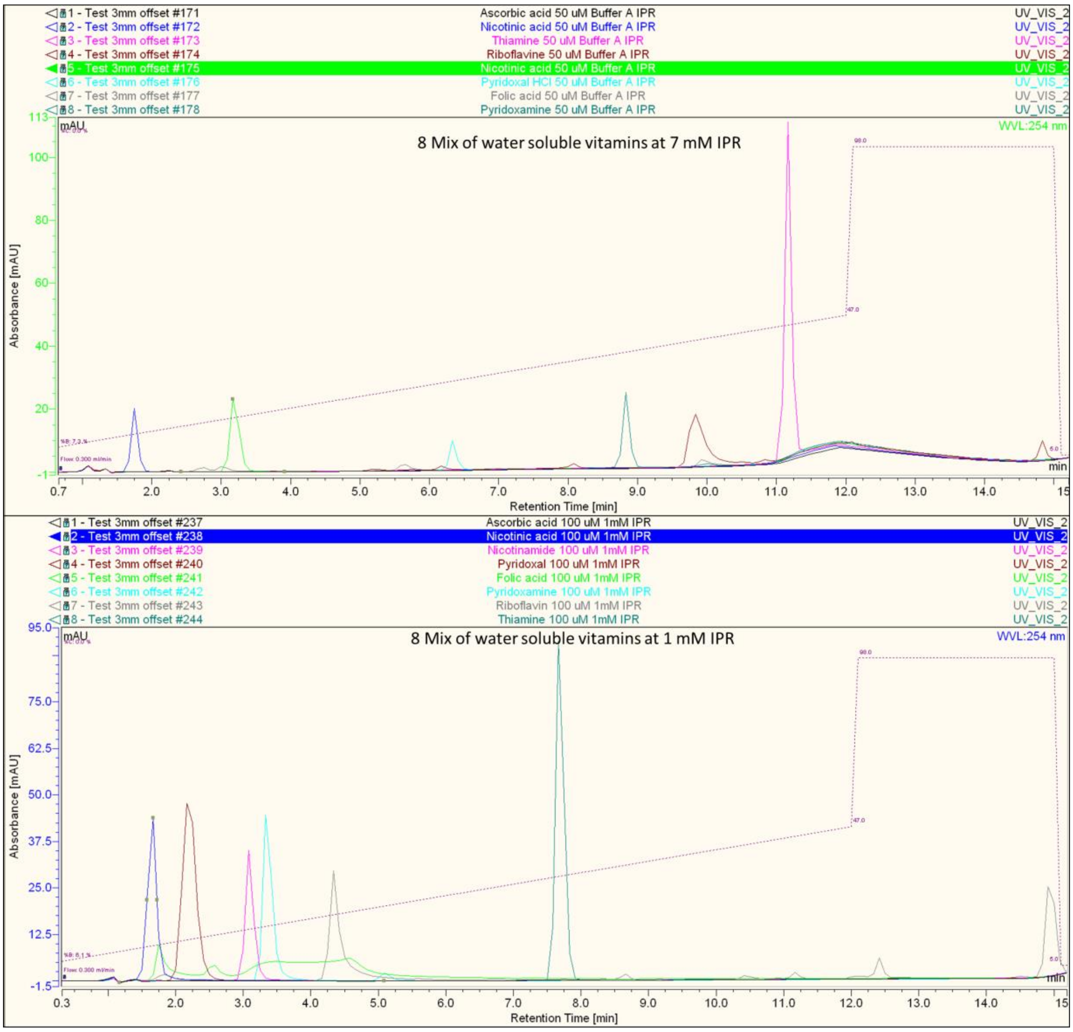


Figure 10.15 Efficiency of IP-RP-HPLC separations in retention and separation of individual water soluble vitamins with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at two different concentrations (7 mM and 1 mM). Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

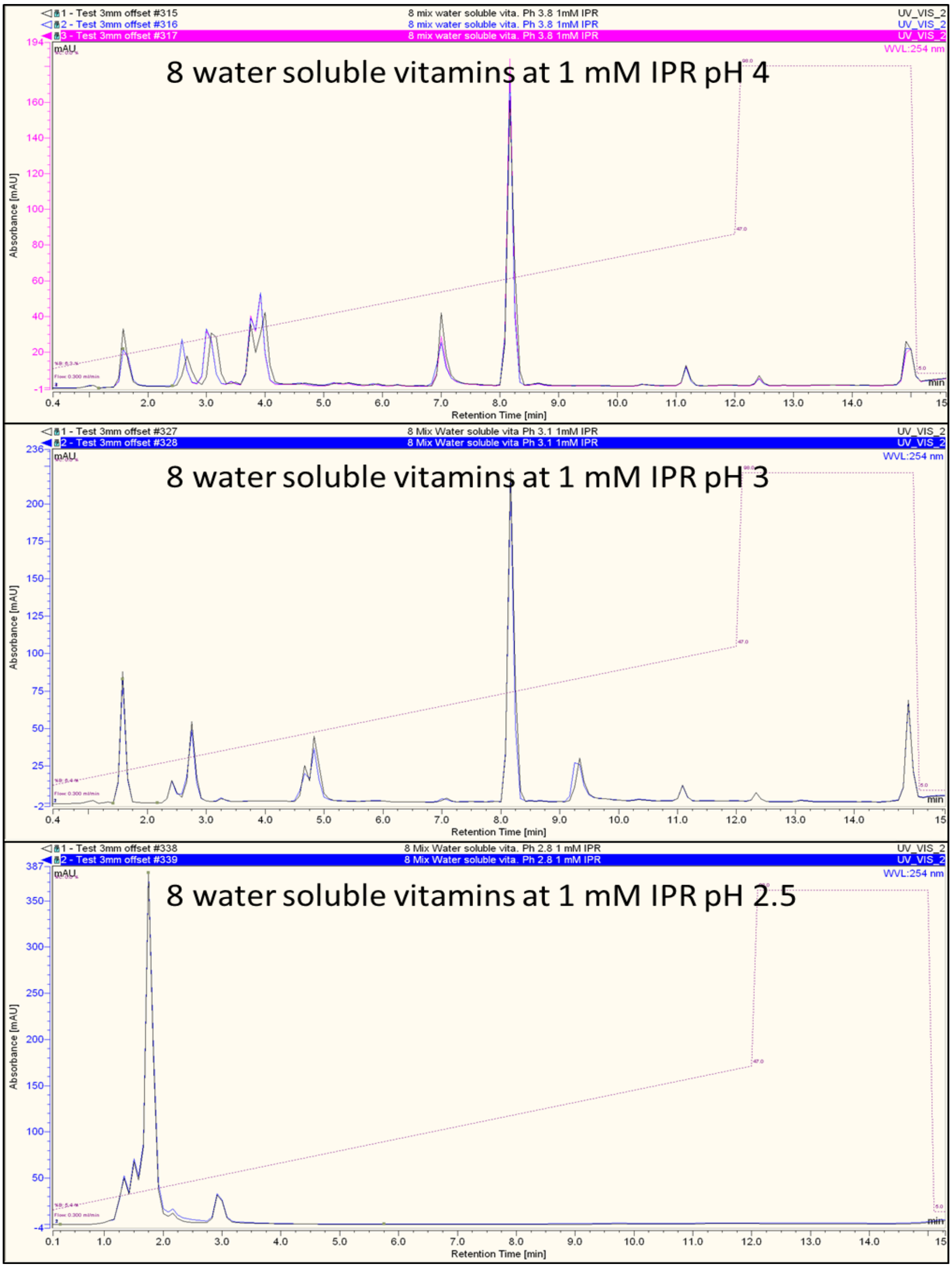


Figure 10.16 Efficiency of IP-RP-HPLC separations in retention and separation of water soluble vitamins from their 8 mix cocktail as listed in Table 10.7, at varying pH conditions with sodium-1-hexanesulfonate monohydrate as an ion pair reagent at 1 mM concentration. Separations were carried out on Dionex Ultimate 3000 HPLC using Kinetex core-shell column (150 x 2.10 mm 2.6u C18 100A). Gradient conditions were used, where flow rate was kept at 0.5 mL min^{-1}. Sample volume injected was 8.9 µL. Chromatograms were obtained using UV detection at a wavelength of 254 nm.

In summary, it can be seen that varying concentration of IPR and pH condition seems to have no influence in improving the retention and separation of amino acids (data not shown), organic acids and nucleotides/nucleosides/nucleobases metabolite class. In case of water soluble vitamins superior separation was achieved at 7mM IPR concentration, however use of such high concentration of IPR is not MS friendly and likely to cause significant interference with full scan MS analyses due to ion suppression effect.

10.3.4 Comparison between HILIC and IP RP HPLC

From above analysis with HILIC and IP RP HPLC, better separation with good resolution and intensities for all the four metabolite classes was obtained with HILIC compared to IP RP HPLC. With respect to analysis of above metabolite classes, HILIC offered several advantages over IP RP HPLC. The key points that we would like to highlight with HILIC based separations are as follows,

- Separation of all the metabolite classes was achieved under similar isocratic HPLC conditions without requirement of any derivatization step or any further treatment.
- Unlike IP RP HPLC, the separation of all the four metabolite classes were carried out without addition of an ion pair reagents which might present further analytical issues with the MS based detection and data analysis.
- With the exception for few metabolites, nearly all the metabolites classes were successfully detected at a wavelength of either 206 or 254 nm.
- Small volume of sample injection (10 µL) is required for HILIC.
- Shorter analysis time: In case of amino acids, 15 amino acids were separated in two runs lasting 9 min each, whereas 13 nucleotides from 14 mix were separated in less than 12 min in single chromatographic run. A mixture of 8 water soluble vitamins were separated in 23 min in single chromatographic run. In case of organic acids, eleven organic acids were separated individually in less than 15 min.
- Amino acids fractions can be directly infused into MS without any buffer modifications or further concentrating step which prevents further metabolite degradation or inter-conversions caused due to such additional steps.
- DI-ESI-MS analysis confirmed, good intensities were obtained for amino acids fractions separated on ZIC-*p*HILIC column.

Overall the performance of ZIC-*p*HILIC column was superior so far for the metabolite classes investigated here compared to that of IP RP HPLC. However it is important to note that this was just a preliminary investigation towards optimisation of LC-MS technique for improving the metabolome coverage. Further comprehensive research is required to develop a platform where in a single chromatographic run majority of the metabolite classes can be separated with good resolution and intensities from their mixture. We propose to address this in our future work. Later the optimised conditions will be employed to profile the metabolome of mammalian and microalgal samples to validate the efficiency of optimised protocol using capillary LC coupled to MS.

10.4 Conclusions

In summary we have performed the preliminary evaluation and comparison of HILIC and IP RP HPLC based separation for the retention and separation of four metabolites classes which includes amino acids, organic acids, water soluble vitamins and nucleotides/nucleosides/nucleobases. The preliminary optimisation of HILIC conditions with ZIC-*p*HILIC column under isocratic conditions resulted in better retention and separation of all the metabolite classes investigated in this study. The uniqueness of this method optimisation with HILIC is that, all the metabolite classes were separated under isocratic conditions and using the same buffer solvents, which was not the case in most of the targeted methods developed and reported in past literature. In all analyses, results showed that better performance was obtained with HILIC conditions compared to that IP RP HPLC. From our results, we can firmly conclude that there is no better column than ZIC-*p*HILIC so far for untargeted metabolomics studies.

Having determined that HILIC separations as the most suitable and very attractive approach for untargeted metabolomic studies, a thorough optimisation of its parameters with respect to maximising the peak efficiency and separation for remaining metabolite classes should be explored, which we propose to address in our future work. Still there are many important classes of metabolites which are poorly resolved with HILIC, therefore development of a method which can effectively capture majority of the metabolite classes for an untargeted metabolomic studies would be beneficial. Later the optimised HILIC conditions will be employed to profile the metabolome of mammalian and microalgal samples using capillary LC coupled to MS. At last it is important to note that, to date there is no single analytical method that is suitable for detection of all the metabolite classes, therefore parallel application of optimised GC-MS and LC-MS workflows for given organism would lead to a “An approach towards full metabolome coverage”.

Chapter 11

Conclusions and future work

11. Conclusions and future work

11.1 Thesis conclusions

This thesis was aimed at development and application of sample preparation workflows and analytical approaches for improving the metabolome coverage (for all metabolite classes) using GC-MS, GC-FID and offline LC-UV. This chapter summarises the success of the investigated miniaturized metabolomics workflows for mammalian and microalgal systems and investigated miniaturized workflows for biodiesel production from microalgae.

In Chapter 3, Objective 1 was tackled which was "To assess the extent of metabolite leakage with conventional approaches used for harvesting adherently growing mammalian cells. Furthermore, characterisation of combined quenching and extraction protocols for rapid and reliable metabolome analysis towards improved metabolome coverage and minimal loss of intracellular metabolites during harvesting in three adherently grown mammalian cell lines (two breast cancer cell lines MDA-MB 436, MCF7 and an endothelial cell line HMEC1)".

With the results of this investigation we provided experimental evidence demonstrating that metabolite leakage occurs with both conventional treatments (trypsinization and cell scraping), and that the cell lines are differentially influenced, suggesting any pre-treatment requires validation for each the specific cell line to allow valid analyses. Further, we examined two recently reported approaches of simultaneous quenching and extraction protocols (including cell scraping and liquid nitrogen) that showed minimal metabolome leakage. However comparison within them, showed little variations with respect to number of metabolites recovered, although recovery of different classes of metabolites appears to be altered significantly across different simultaneous quenching and extraction treatments. We also investigated the culture of cells on beads for rapid quenching and extraction, as a novel sample handling protocol which seems to be rapid, effective and an efficient approach for recovery of different classes of metabolites. For metastatic breast cancer cells MDA-MB-436, the two direct quenching approaches and the bead harvesting approach showed favourable results with respect to metabolome leakage, compared to the conventional approaches. We also examined the effect of including a protein precipitation step on the metabolite classes detected. The deproteination step did not show significant improvement in overall recoveries, and did not appear to dramatically alter the pattern of different classes of metabolites detected compared to direct quenching approaches. This methodological and thorough investigation suggests that whichever methodology is chosen, it is important to establish the level of metabolome leakage for the specific cell line investigated. Moreover, when particular class of metabolites are of interest, it is necessary to carefully consider the impact of different treatments on recovery of different classes of metabolites while selection of appropriate method.

In Chapter 4, Objective 2 was tackled, which was “To examine the influence of various washing and quenching solvents (buffered/non-buffered) in metabolomics of adherently grown mammalian cells: A case study on the metastatic cancer cell line MDA-MB-231”. In this investigation we applied the optimised modified direct cell scarping (NTcs) and LN_2 methods to evaluate the influence of various washing and quenching solvents (buffered/non-buffered) on metabolite leakage. To achieve our objective, we have employed a broader approach which involved initial quantification of intracellular ATP levels using bioluminescence assay, visualisation of cell membrane integrity using scanning electron microscopy (SEM) and recovery of eleven different metabolite classes using GC-MS technique across different quenching additives. Finally we co-related the results obtained with the ATP assay, SEM and GC-MS analysis.

The evidence obtained from different analytical platforms involving ATP assays, SEM analysis and GC-MS based metabolome analyses correlates very well and clearly indicates a single washing step with PBS and quenching with methanol supplemented with 70 mM HEPES resulted in minimal leakage of intracellular metabolites, apparently preserving the membrane integrity of adherently growing metastatic cancer cell line MDA-MB-231, as compared to other quenching solutions. In addition, we have demonstrated no interferences on derivatization reactions and GC-MS based analysis when PBS and HEPES are employed in the protocol.

In Chapter 5, Objective 3 was tackled, which was “Further validate and compare proposed protocols for metabolomics of two metastatic TNBC cell lines: MDA-MB-231 and MDA-MB-436 with respect to pathway specific metabolome coverage”. With preliminary investigations in both the cell lines, we have concluded that both methods (NTcs and LN_2) are superior with respect to the recoveries of different metabolite classes and intensities recovered for them. Subsequent comparative analysis between methods with the aid of visualisation tools resulted in a reduced number of significantly changing features, thereby demonstrating minimum variation between both methods and justifying the importance of both methods in recovering specific metabolite classes with regard to the specific cell line under investigation.

Comparison of the metabolomes recovered in two different metastatic cancer cell lines with both methods revealed huge variations in significantly changing features, where number of metabolites identified with MDA-MB-436 were around two fold higher than that of identified in

MDA-MB-231. These observations clearly demonstrate the ability of both the proposed methods to trap the variations in the metabolome of two biologically and morphologically similar metastatic cancer cell lines. In total 154 unique metabolites were identified successfully mapped on significant pathways involved in TNBC cells, thereby demonstrating the robustness and reliability of these methods in pathways based analysis in TNBC cell lines. The observations made with the pathway-based coverage specific to cancer will surely help in generating many hypothesis based investigations in cancer community such as, for setting up a platform for classifying cancer sub-type, in defining the relative contribution of major metabolic pathways in TNBC and illustrating the underlying molecular mechanism involved in TNBC biology and finally for generating novel therapeutic regimes by identifying novel targets for anticancer therapies.

In Chapter 6, 7 and 8, Objective 4 was tackled, which was "To examine GC-MS based optimised sampling, quenching and extraction workflow for obtaining high quality metabolomic data for microalgal strain: *Chlamydomonas reinhardtii*". The results from preliminary investigation (Chapter 6) identified 216 unique metabolites across different conditions and species, thereby demonstrating the potential and suitability of GC-MS for non-targeted metabolomic investigations. However recovery of nucleotides/nucleosides/nucleotides, pigments and sugar phosphates was severely compromised which forms the basis of our Objective 6 which we have addressed in Chapter 10. Biology based metabolomics investigation identified significant changes in the metabolome of *C. reinhardtii* cultures grown under different trophic conditions and showed high co-relation and agreements with the conclusions drawn from proteomic and biochemical analysis. We evaluated the influence of two factors, carbon source and availability, on mixotrophically and photoautotrophically growing cultures of *C. reinhardtii*, within which results obtained clearly demonstrated that, the most predominant factor influencing the *C. reinhardtii* cultures is the carbon availability rather than the source.

In Chapter 7, we proposed an optimised miniaturized sampling and quenching method for microalga *C. reinhardtii*, which requires only 1 mL of microalgal culture, ideal for metabolomics investigation where a large number of samples need processing. Our results clearly showed higher loss of intracellular metabolites with the use of conventional 60% methanol for quenching. Analysis with respect to influence of varying methanol concentration clearly showed higher leakage with the increase in methanol concentration. On the other hand, analysis with respect to inclusion of various buffer additives to quenching solvent, 60% aqueous methanol supplemented with 70 mM HEPES managed to recover higher intracellular levels for nearly all metabolite classes with minimal leakage, supporting our previous findings with adherent mammalian cells. Increment in quenching solvent to sample ratio in order to keep the temperatures of the resulting mixture below -20°C for effective halting of metabolism seems to have no influence on preserving the cell integrity and in minimising the metabolite leakage. In contrast, higher leakage was observed with this approach. Moreover, with the experimental evidence we have demonstrated higher amount of leakage as the contact time of sample to that of quenching solvent increases. For quantitative metabolomics studies in *C. reinhardtii*, we strongly recommend quenching with 60% aqueous methanol supplemented with 70 mM HEPES (-40°C) at 1:1 sample to quenching solvent ratio, as it resulted in higher recoveries for intracellular metabolites with subsequent reduction in the metabolite leakage for all classes of metabolites.

In Chapter 8, we have performed a comprehensive evaluation of 15 extraction solvents (both monophasic and biphasic) for untargeted GC-MS based metabolome analysis of *C. reinhardtii*. The results obtained across various treatments were compared, assessed, evaluated and validated with respect to the recoveries of different classes and numbers of metabolites. The evaluation criteria used while proposing the best extraction solvents involved analysis with respect to number of peaks detected with each extraction solvent, the peak intensity of structurally identified compound and the reproducibility of metabolite quantification with minimal analytical variations. We have successfully demonstrated that the choice of extraction solvent has a greater impact on recovery of intracellular metabolites from *C. reinhardtii.* Our results clearly indicate that 25 % aqueous methanol is the best suited extraction solvent for untargeted metabolomic analysis of *C. reinhardtii*, as the highest number of metabolites belonging to various chemical classes was recovered with good intensities and reproducibility.

In summary, results obtained within Chapter 7 and Chapter 8, clearly indicate that 60% methanol (buffered with HEPES) and 25 % aqueous methanol are the best suited quenching and extraction solvent respectively for obtaining high quality metabolomic data for *C. reinhardtii.*

In Chapter 9, Objective 5 was tackled, which was “To optimise and make a comparative evaluation of conventional and direct trans-esterification methods for reliable quantification of FAMEs from micro-algal biofuel strains”. Bottlenecks in algal biodiesel production within the cell can be identified by metabolomics approaches, however prior to identifying and addressing these bottlenecks it is essential to optimize the FAME production workflow for biodiesel application.

Four investigations were carried out in order to optimize the FAMEs production workflow for three microalgal strains. In the first experiment, we have evaluated performance of two extraction techniques for biodiesel production for two microalgal species (*C. reinhardtii* and *N. salina* which differ in their cell wall structure) using design of experiments approach. Two analytical platforms namely SEM and GC-FID analysis were employed, this identified bead beating as the most suitable disruption technique with optimised parameters (5 bead beating cycle with 2 minutes relaxation time on ice), as better recoveries of FAMEs required for biodiesel production were achieved compared to sonication treatment. Moreover we justified exclusion of pre-treatments such as lyophilisation and addition of antioxidants, as these pre-treatments are more time consuming and will incur more cost for industrial scale application.

In the second investigation, we developed optimised GC-FID method for accurate and reliable quantification of fatty acids. Furthermore, sample analysis time was reduced drastically to 12 minutes with good resolution, where normal analysis time reported in recent publications (Griffiths et al., 2010, Kumari et al., 2011, Laurens et al., 2012a, Laurens et al., 2012b, Slocombe et al., 2013, Xu et al., 2010) was between 35 to 90 minutes. In the third investigation, we have justified use of non-chlorinated solvents instead of more toxic chloroform in recovery of lipids from different sample matrices in TE reactions. Further we have evaluated use of non-chlorinated solvent mixtures and compared them within and with the conventional solvent systems. Based on our GC-FID results, we have successfully demonstrated use of non-chlorinated solvents for microalgal lipid extraction across all the three species, with higher FAs recoveries as compared to that of conventional toxic and carcinogenic solvents mixtures (methanol and chloroform). The best yields were obtained with propane-2-ol, cyclohexane and water solvent mixture. In the fourth investigation, we evaluated the TE efficiency of different catalysts for biodiesel production from microalgal lipids. In total eleven DT protocols were carefully selected or constructed which were then compared within and with the conventional methods. For thorough investigation, all the recovered FAMEs across different applied methods were categorised according to their concentration (low, high and medium) and their degree of saturation (SFAs, MUFAs and PUFAs). Among three acid catalysts evaluated, acetyl chloride showed highest TE efficiency, whereas among base catalysts poor/no recoveries were obtained. Thus base catalyst in isolation are not recommended. Among sequential use of both catalysts, the use of base catalyst SM followed by acid catalysis with 5% acetyl chloride in methanol, showed superior results in terms of number of individual FAMEs (SAFs, MUFAs and PUFAs) and total FA content recovery and henceforth most recommended. It is important to note that, so far sequential use of SM with acetyl chloride in methanol or methanolic HCl has not been applied to microalgae for evaluating FAME recoveries for biodiesel productions. Hence for biodiesel applications, we strongly recommend use of DT method with sequential use of SM followed by acetyl chloride in methanol. However where different classes of lipid are of interest, use of non-chlorinated solvent mixture (propane-2-ol, cyclohexane and water) based on Folch principle (method 2D) is recommended.

In Chapter 10, Objective 6 was tackled, which was "To evaluate the use of HILIC and IP-RP-HPLC for targeted analysis of amino acids, organic acids, water soluble vitamins, nucleotides, nucleosides and nucleobases towards improving metabolome coverage". The preliminary optimisation of HILIC conditions with ZIC-*p*HILIC column under isocratic conditions resulted in better retention and separation of all the metabolite classes investigated in this study. The uniqueness of this method optimisation with HILIC is that, all the metabolite classes were separated under isocratic conditions and using the same buffer solvents, which is not the case in most of the targeted methods developed and reported in past literature. In all respect of analyses, results showed better performance with HILIC conditions compared to that IP RP HPLC. From our results, we can firmly conclude that there is no better column than ZIC-*p*HILIC, so far, for untargeted metabolomics studies.

11.2 Future work

Great progress in metabolomics has been achieved in last three decades, however researchers are only at the start of a long journey and further developments in all of the processes involved in metabolomics workflow are required to obtain greater metabolome coverage.

11.2.1 Analytical challenges

An approach towards full metabolome coverage

In Chapter 6, we have demonstrated excellent coverage for most of the metabolite classes with GC-MS based analysis. Coupling of GC to quadrupole MS, provides high sensitivity and larger dynamic range but it has nominal mass accuracy and slow scan speed. Further to improve the metabolome coverage with GC-MS, we propose to employ 2D GC-ToF-MS or GC-ToF-MS in future for un-targeted metabolomics studies. Recent developments in Q-TOF instruments with respect to increase in mass accuracy, sensitivity and dynamic range, makes them a more suitable choice for increased applications in metabolomics and as an cheaper alternative to FTICR mass spectrometers which provides higher mass accuracy and mass resolution. Furthermore, it offers high scan speeds essential for adequate sampling of high-resolution chromatographic peak widths in the range of 0.5 to 1 s, which also facilitates the implementation of fast GC methods, thus reducing the analysis time and increasing the productivity. In addition, some classes of metabolites such as phospho-metabolites cannot be analyzed in positive mode and require analysis to be carried out in negative mode, therefore for maximum metabolome coverage, we aim to perform future analysis in both positive and negative mode.

In Chapter 10, we have determined the HILIC separations as the most suitable and very attractive approach for untargeted metabolomic studies, a thorough optimisation of its parameters with respect to maximising the peak efficiency and separation for remaining metabolite classes should be explored, which we propose to address in our future work. Still there are many important classes of metabolites which are poorly resolved with HILIC, therefore development of a method which can effectively capture majority of the metabolite classes for an untargeted metabolomic studies would be beneficial. Later the optimised HILIC conditions will be employed to profile the metabolome of mammalian and microalgal samples using capillary LC coupled to MS. Compared with GC-MS, LC-MS is a younger technique with great possibilities for improvements. However, the major disadvantage of LC-MS is ion suppression, which can be overcome to some extent by miniaturization of ESI to nanospray ionization and by a better separation of metabolites. Apart from ion suppression another issue with LC-MS metabolomics is contamination of the MS source, and adduct formation, which have significant consequences for the robustness of the method and the lack of transferable LC-MS libraries for metabolite identifications. In LC-ESI-MS, the most-needed area for improvement to facilitate the development of new method is the proper knowledge and understanding of ionization mechanism in the LC-MS interface, which may lead to the possibility to exchange spectra libraries between various research groups. The enrichment of databases of metabolites identified by LC-MS will aid us to take full advantage of the tremendous potential application of metabolomics, as a tool for elucidation of interrelated and inter-regulated biological process.

In metabolomics research, it is essential to carry out the method validation in order to assess the performance and the fitness of a method and/or analytical system. In order to achieve this objective, we propose to take the initiative for the establishment of standards to validate the performance of hyphenated MS-based techniques, by identifying and quantifying metabolites from defined mixtures of metabolites with a wide range of relative concentration ratios, consisting of representative candidates from different classes to generate a metabolite pool. The use of such a defined metabolite pool will provide a means to quality control check on column performance, MS detection, and validation of existing methods and deconvolution softwares and to optimize the analytical parameters in MS-based metabolomics. The further study will also focus in detail, in finding the most efficient, automated, flexible and reliable data handling system in order to boost metabolomics progress by comparing different deconvolution softwares like AMDIS, ChromaTOF, XCMS, MZmine2 and AnalyzerPro. Such a comprehensive study will not only help us in finding new algorithms and a selection of suitable softwares that can be applied further in our research in order to avoid false positives and false negatives results but also will have a capability to deconvolute low-concentration components from high noise and background.

At last it is important to note that, till date there is no single analytical method that is suitable for detection of all the metabolite classes due to physicochemical diversity of the metabolites, therefore parallel application of optimised GC-MS and LC-MS workflows for given organism would lead to

“An approach towards full metabolome coverage”.

11.2.2 Challenges in microalgal metabolomics

In Chapter 7 and 8, for comprehensive evaluation of quenching and extraction methods in microalgae, model microalga *C. reinhardtii* was selected for un-targeted GC-MS based metabolomics. As *C. reinhardtii* is a freshwater species, its culture media contain trace amount of salts, whereas *N. salina* & *D. salina* are marine microalgal species and their culture media contains high concentration of salts. As discussed in Chapter 2, the presence of trace amount of salts in analytical samples causes significant interferences in MS based analysis. In addition it also introduces an additional source of variance to the experimental procedures, increases complexity of data and might affect the derivatization efficiency in GC-MS based analysis. Therefore, metabolomic studies on marine microalgal species require careful optimisation of sample preparation workflows, or else will result in poor metabolome coverage with MS based techniques. Further optimisation with respect to inclusion of washing step in sampling and quenching workflow needs to be evaluated.

The construction of metabolite-specific libraries for plant and microbial metabolomics in the past, have provided a greater ability to identify metabolites. Further development of LC-MS and GC-MS mass spectral libraries with respect to microalgae and their applicability and availability across many research laboratories is essential, but not yet achieved. As of today, there are no standard metabolite databases for these species, therefore our future research will also focus on construction of preliminary metabolite database for microalgal species.

In Chapter 9, we have developed an optimised method for accurate and reliable quantification of fatty acids from microalgal species for biodiesel application. Apart from this there are around 400 known carotenoids from microalgal extracts, among which only few like β-carotene, astaxanthin, lutein, zeaxanthin, lycopene and bixin are currently being used for commercial purposes. However there is no standardised method developed yet for accurate and reliable quantification of these pigments from numerous microalgal species. Future research will also be directed towards development of targeted LC-MS based methods for accurate and reliable quantification of these pigments.

In metabolomic studies, both the choice of analytical technique and sample preparation workflow should be carefully considered which are invariably linked to the quality and reliability of metabolomics data. Unfortunately, improvements of the analytical techniques were not followed by similar improvements in methods for sample preparation. Within Chapter 7 and 8, we have developed optimised sampling, quenching and extraction methods for obtaining high quality metabolomic data from microalga *C. reinhardtii*, however further integration of such reliable metabolomics data will help us in identifying the as-of-yet missing reactions in the metabolic network which later will allow the fine tuning of algal properties by genetic or metabolite engineering in combination with other system biology approaches (such as proteomics, genomics and transcriptomics).

11.2.3 Challenges in mammalian metabolomics

To sustain growth and proliferation of tumour cells, supplements of macromolecular precursors are constantly required and therefore tumour cells exhibit altered metabolism compared to quiescent cells. In cancer community, several researchers have employed metabolomics approach to catalogue these changes and most of the studies focused on a classification approach where healthy cells were compared against tumour cells.

In Chapter 5, we have demonstrated the ability of both the proposed methods to trap the variations in the metabolome of two biologically and morphologically similar metastatic cancer cell lines. In total 154 unique metabolites were identified and successfully mapped on significant pathways involved in TNBC cells, thereby demonstrating the robustness and reliability of these methods in pathways based analysis in TNBC cell lines. The observations made with the pathway-based coverage specific to cancer will surely help in generating many hypothesis based investigations such as, setting up a platform for classifying cancer sub-type, defining the relative contribution of major metabolic pathways in TNBC and illustrating the underlying molecular mechanism involved in TNBC biology and finally for generating novel therapeutic regimes by identifying novel targets for anticancer therapies. However, the potential of these proposed methods in illustrating the underlying molecular mechanism in cancer needs to be explored further in the future.

References

ABLETT, M. P., SINGH, J. K. & CLARKE, R. B. 2012. Stem cells in breast tumours: are they ready for the clinic? *European Journal of Cancer,* 48**,** 2104-2116.

ADAM, F., ABERT-VIAN, M., PELTIER, G. & CHEMAT, F. 2012. “Solvent-free” ultrasound-assisted extraction of lipids from fresh microalgae cells: A green, clean and scalable process. *Bioresource technology,* 114**,** 457-465.

AL AWAM, K., HAUßLEITER, I. S., DUDLEY, E., DONEV, R., BRÜNE, M., JUCKEL, G. & THOME, J. 2014. Multiplatform metabolome and proteome profiling identifies serum metabolite and protein signatures as prospective biomarkers for schizophrenia. *Journal of Neural Transmission***,** 1-12.

ALISON, M. R., LIN, W.-R., LIM, S. M. & NICHOLSON, L. J. 2012. Cancer stem cells: in the line of fire. *Cancer treatment reviews,* 38**,** 589-598.

AMARO, H. M., ESQUÍVEL, M. G., PINTO, T. S. & MALCATA, F. X. 2013. Hydrogen Production by Microalgae. *Natural and Artificial Photosynthesis: Solar Power as an Energy Source***,** 231-241.

AMELIO, I., CUTRUZZOLÁ, F., ANTONOV, A., AGOSTINI, M. & MELINO, G. 2014. Serine and glycine metabolism in cancer. *Trends in biochemical sciences,* 39**,** 191-198.

AMIN, S. 2009. Review on biofuel oil and gas production processes from microalgae. *Energy Conversion and Management,* 50**,** 1834-1840.

ANDERSEN, M. E., MEEK, M. E., BOORMAN, G. A., BRUSICK, D. J., COHEN, S. M., DRAGAN, Y. P., FREDERICK, C. B., GOODMAN, J. I., HARD, G. C. & O'FLAHERTY, E. J. 2000. Lessons learned in applying the US EPA proposed cancer guidelines to specific compounds. *Toxicological Sciences,* 53**,** 159-172.

ARANIBAR, N., BORYS, M., MACKIN, N. A., LY, V., ABU-ABSI, N., ABU-ABSI, S., NIEMITZ, M., SCHILLING, B., LI, Z. J. & BROCK, B. 2011. NMR-based metabolomics of mammalian cell and tissue cultures. *Journal of biomolecular NMR,* 49**,** 195-206.

ARAUJO, G. S., MATOS, L. J., FERNANDES, J. O., CARTAXO, S. J., GONCALVES, L. R., FERNANDES, F. A. & FARIAS, W. R. 2013. Extraction of lipids from microalgae by ultrasound application: prospection of the optimal extraction method. *Ultrasonics sonochemistry,* 20**,** 95-98.

ASBURY, G. R. & HILL, H. H. 2000. Evaluation of ultrahigh resolution ion mobility spectrometry as an analytical separation device in chromatographic terms. *Journal of Microcolumn Separations,* 12**,** 172-178.

BACHER, A., RIEDER, C., EICHINGER, D., ARIGONI, D., FUCHS, G. & EISENREICH, W. 1998. Elucidation of novel biosynthetic pathways and metabolite flux patterns by retrobiosynthetic NMR analysis. *FEMS Microbiology Reviews,* 22**,** 567-598.

BALASUBRAMANIAN, S., ALLEN, J. D., KANITKAR, A. & BOLDOR, D. 2011. Oil extraction from< i> Scenedesmus obliquus</i> using a continuous microwave system–design, optimization, and quality characterization. *Bioresource technology,* 102**,** 3396-3403.

BARBERO, G. F., PALMA, M. & BARROSO, C. G. 2006. Pressurized Liquid Extraction of Capsaicinoids from Peppers. *Journal of Agricultural and Food Chemistry,* 54**,** 3231-3236.

BEDAIR, M. & SUMNER, L. W. 2008a. Current and emerging mass-spectrometry technologies for metabolomics. *TrAC - Trends in Analytical Chemistry,* 27**,** 238-250.

BEDAIR, M. & SUMNER, L. W. 2008b. Current and emerging mass-spectrometry technologies for metabolomics. *TrAC Trends in Analytical Chemistry,* 27**,** 238-250.

BEHRENDS, V., TREDWELL, G. D. & BUNDY, J. G. 2011. A software complement to AMDIS for processing GC-MS metabolomic data. *Analytical Biochemistry,* 415**,** 206-208.

BELOUECHE-BABARI, M., JACKSON, L. E., AL-SAFFAR, N. M. S., ECCLES, S. A., RAYNAUD, F. I., WORKMAN, P., LEACH, M. O. & RONEN, S. M. 2006. Identification of magnetic resonance detectable metabolic changes associated with inhibition of phosphoinositide 3-kinase signaling in human breast cancer cells. *Molecular cancer therapeutics,* 5**,** 187-196.

BEN-AMOTZ, A. & AVRON, M. 1990. The biotechnology of cultivating the halotolerant alga< i> Dunaliella</i>. *Trends in Biotechnology,* 8**,** 121-126.

BENTHIN, B., DANZ, H. & HAMBURGER, M. 1999. Pressurized liquid extraction of medicinal plants. *Journal of Chromatography A,* 837**,** 211-219.

BI, H., KRAUSZ, K., MANNA, S., LI, F., JOHNSON, C. & GONZALEZ, F. 2013. Optimization of harvesting, extraction, and analytical protocols for UPLC-ESI-MS-based metabolomic analysis of adherent mammalian cancer cells. *Analytical and Bioanalytical Chemistry,* 405**,** 5279-5289.

BLIGH, E. G. & DYER, W. J. 1959. A rapid method of total lipid extraction and purification. *Canadian journal of biochemistry and physiology,* 37**,** 911-917.

BOBROVNIKOVA-MARJON, E. & HUROV, J. B. 2014. Targeting metabolic changes in cancer: novel therapeutic approaches. *Annual review of medicine,* 65**,** 157-170.

BÖLLING, C. & FIEHN, O. 2005. Metabolite profiling of Chlamydomonas reinhardtii under nutrient deprivation. *Plant Physiology,* 139**,** 1995-2005.

BÖLLING, C. & FIEHN, O. 2005. Metabolite profiling of Chlamydomonas reinhardtii under nutrient deprivation. *Plant Physiology,* 139**,** 1995-2005.

BOLTEN, C. J., KIEFER, P., LETISSE, F., PORTAIS, J.-C. & WITTMANN, C. 2007. Sampling for metabolome analysis of microorganisms. *Analytical chemistry,* 79**,** 3843-3849.

BOROWITZKA, L., MOULTON, T. & BOROWITZKA, M. The mass culture of Dunaliella salina for fine chemicals: from laboratory to pilot plant. Eleventh International Seaweed Symposium, 1984. Springer, 115-121.

BOYLE, N. R. & MORGAN, J. A. 2009. Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC Systems Biology,* 3.

BRACHET, A., CHRISTEN, P. & VEUTHEY, J. L. 2002. Focused microwave-assisted extraction of cocaine and benzoylecgonine from coca leaves. *Phytochemical Analysis,* 13**,** 162-169.

BREUER, G., EVERS, W. A., DE VREE, J. H., KLEINEGRIS, D. M., MARTENS, D. E., WIJFFELS, R. H. & LAMERS, P. P. 2013. Analysis of fatty acid content and composition in microalgae. *J. Vis. Exp*.

BRITTEN, R. J. & MCCLURE, F. T. 1962. The amino acid pool in Escherichia coli. *Bacteriological reviews,* 26**,** 292.

BRUCE, S. J., TAVAZZI, I., PARISOD, V., REZZI, S., KOCHHAR, S. & GUY, P. A. 2009. Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. *Analytical chemistry,* 81**,** 3285-3296.

BUCHHOLZ, A., HURLEBAUS, J., WANDREY, C. & TAKORS, R. 2002. Metabolomics: quantification of intracellular metabolite dynamics. *Biomolecular engineering,* 19**,** 5-15.

BURGESS, K., CREEK, D., DEWSBURY, P., COOK, K. & BARRETT, M. P. 2011. Semi-targeted analysis of metabolites using capillary-flow ion chromatography coupled to high-resolution mass spectrometry. *Rapid Communications in Mass Spectrometry,* 25**,** 3447-3452.

BURJA, A. M., ARMENTA, R. E., RADIANINGTYAS, H. & BARROW, C. J. 2007. Evaluation of fatty acid extraction methods for Thraustochytrium sp. ONC-T18. *Journal of agricultural and food chemistry,* 55**,** 4795-4801.

BURTON, L., IVOSEV, G., TATE, S., IMPEY, G., WINGATE, J. & BONNER, R. 2008. Instrumental and experimental effects in LC-MS-based metabolomics. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences,* 871**,** 227-235.

CALLAHAN, D. L., SOUZA, D. D., BACIC, A. & ROESSNER, U. 2009. Profiling of polar metabolites in biological extracts using diamond hydride-based aqueous normal phase chromatography. *Journal of separation science,* 32**,** 2273-2280.

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.

CANELAS, A. B., TEN PIERICK, A., RAS, C., SEIFAR, R. M., VAN DAM, J. C., VAN GULIK, W. M. & HEIJNEN, J. J. 2009. Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics. *Analytical chemistry,* 81**,** 7379-7389.

CARRAPISO, A. I. & GARCÍA, C. 2000. Development in lipid analysis: some new extraction techniques and in situ transesterification. *Lipids,* 35**,** 1167-1177.

CASTRILLO, J. I., HAYES, A., MOHAMMED, S., GASKELL, S. J. & OLIVER, S. G. 2003. An optimized protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry. *Phytochemistry,* 62**,** 929-937.

CEQUIER-SÁNCHEZ, E., RODRÍGUEZ, C., RAVELO, A. N. G. & ZÁRATE, R. 2008. Dichloromethane as a solvent for lipid extraction and assessment of lipid classes and fatty acids from samples of different natures. *Journal of agricultural and food chemistry,* 56**,** 4297-4303.

CHEN, M., TANG, H., MA, H., HOLLAND, T. C., NG, K. & SALLEY, S. O. 2011. Effect of nutrients on growth and lipid accumulation in the green algae< i> Dunaliella tertiolecta</i>. *Bioresource technology,* 102**,** 1649-1655.

CHEN, Q., LI, P., HE, J., ZHANG, Z. & LIU, J. 2008. Supercritical fluid extraction for identification and determination of volatile metabolites from Angelica dahurica by GC-MS. *Journal of Separation Science,* 31**,** 3218-3224.

CHEN, Y., BICKER, W., WU, J., XIE, M. & LINDNER, W. 2012. Simultaneous determination of 16 nucleosides and nucleobases by hydrophilic interaction chromatography and its application to the quality evaluation of ganoderma. *Journal of agricultural and food chemistry,* 60**,** 4243-4252.

CHEN, Z., LANDMAN, P., COLMER, T. D. & ADAMS, M. A. 1998. Simultaneous Analysis of Amino and Organic Acids in Extracts of Plant Leaves astert-Butyldimethylsilyl Derivatives by Capillary Gas Chromatography. *Analytical Biochemistry,* 259**,** 203-211.

CHENG, C.-H., DU, T.-B., PI, H.-C., JANG, S.-M., LIN, Y.-H. & LEE, H.-T. 2011. Comparative study of lipid extraction from microalgae by organic solvent and supercritical CO< sub> 2</sub>. *Bioresource technology,* 102**,** 10151-10153.

CHENG, J., CHE, N., LI, H., MA, K., WU, S., FANG, J., GAO, R., LIU, J., YAN, X. & LI, C. 2013. Extraction, derivatization, and determination of metabolome in human macrophages. *Journal of separation science,* 36**,** 1418-1428.

CHOI, S. P., NGUYEN, M. T. & SIM, S. J. 2010. Enzymatic pretreatment of< i> Chlamydomonas reinhardtii</i> biomass for ethanol production. *Bioresource technology,* 101**,** 5330-5336.

CHRISTIE, W. W. 1993. Preparation of ester derivatives of fatty acids for chromatographic analysis. *Advances in lipid methodology,* 2**,** e111.

CLAY, K. L., MURPHY, R. & WATRINS, W. D. 1975. Experimental methanol toxicity in the primate: analysis of metabolic acidosis. *Toxicology and applied pharmacology,* 34**,** 49-61.

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.

CRAIG, D. W., O'SHAUGHNESSY, J. A., KIEFER, J. A., ALDRICH, J., SINARI, S., MOSES, T. M., WONG, S., DINH, J., CHRISTOFORIDES, A. & BLUM, J. L. 2013. Genome and transcriptome sequencing in prospective metastatic triple-negative breast cancer uncovers therapeutic vulnerabilities. *Molecular cancer therapeutics,* 12**,** 104-116.

CRISCITIELLO, C., AZIM, H. A., SCHOUTEN, P., LINN, S. & SOTIRIOU, C. 2012. Understanding the biology of triple-negative breast cancer. *Annals of Oncology,* 23**,** vi13-vi18.

CUBBON, S., ANTONIO, C., WILSON, J. & THOMAS-OATES, J. 2010. Metabolomic applications of hilic–lc–ms. *Mass spectrometry reviews,* 29**,** 671-684.

ČUPERLOVIĆ-CULF, M., BARNETT, D. A., CULF, A. S. & CHUTE, I. 2010. Cell culture metabolomics: applications and future directions. *Drug Discovery Today,* 15**,** 610-621.

D'OCA, M. G. M., VIÊGAS, C. V., LEMÕES, J. S., MIYASAKI, E. K., MORÓN-VILLARREYES, J. A., PRIMEL, E. G. & ABREU, P. C. 2011. Production of FAMEs from several microalgal lipidic extracts and direct transesterification of the< i> Chlorella pyrenoidosa</i>. *biomass and bioenergy,* 35**,** 1533-1538.

DANIELSSON, A. P. H., MORITZ, T., MULDER, H. & SPÉGEL, P. 2010. Development and optimization of a metabolomic method for analysis of adherent cell cultures. *Analytical Biochemistry,* 404**,** 30-39.

DAROCH, M., GENG, S. & WANG, G. 2013. Recent advances in liquid biofuel production from algal feedstocks. *Applied Energy,* 102**,** 1371-1381.

DE JONGE, L. P., DOUMA, R. D., HEIJNEN, J. J. & VAN GULIK, W. M. 2012. Optimization of cold methanol quenching for quantitative metabolomics of Penicillium chrysogenum. *Metabolomics,* 8**,** 727-735.

DE LA CRUZ GARCIA, C., LOPEZ HERNANDEZ, J. & SIMAL LOZANO, J. 2000. Gas chromatographic determination of the fatty-acid content of heat-treated green beans. *Journal of Chromatography A,* 891**,** 367-370.

DENERY, J. R., DRAGULL, K., TANG, C. S. & LI, Q. X. 2004. Pressurized fluid extraction of carotenoids from Haematococcus pluvialis and Dunaliella salina and kavalactones from Piper methysticum. *Analytica Chimica Acta,* 501**,** 175-181.

DETTMER, K., ARONOV, P. A. & HAMMOCK, B. D. 2007. Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews,* 26**,** 51-78.

DETTMER, K., NÜRNBERGER, N., KASPAR, H., GRUBER, M. A., ALMSTETTER, M. F. & OEFNER, P. J. 2011. Metabolite extraction from adherently growing mammalian cells for metabolomics studies: optimization of harvesting and extraction protocols. *Analytical and bioanalytical chemistry,* 399**,** 1127-1139.

DIETMAIR, S., HODSON, M. P., QUEK, L. E., TIMMINS, N. E., CHRYSANTHOPOULOS, P., JACOB, S. S., GRAY, P. & NIELSEN, L. K. 2012. Metabolite profiling of CHO cells with different growth characteristics. *Biotechnology and Bioengineering*.

DIETMAIR, S., TIMMINS, N. E., GRAY, P. P., NIELSEN, L. K. & KRÖMER, J. O. 2010. Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol. *Analytical biochemistry,* 404**,** 155-164.

DING, J., SORENSEN, C. M., ZHANG, Q., JIANG, H., JAITLY, N., LIVESAY, E. A., SHEN, Y., SMITH, R. D. & METZ, T. O. 2007. Capillary LC coupled with high-mass measurement accuracy mass spectrometry for metabolic profiling. *Analytical chemistry,* 79**,** 6081-6093.

DUAN, X., YOUNG, R., STRAUBINGER, R. M., PAGE, B. J., CAO, J., WANG, H., YU, H., CANTY JR, J. M. & QU, J. 2009. A Straightforward and Highly Efficient Precipitation/On-pellet Digestion Procedure Coupled to a Long Gradient Nano-LC Separation and Orbitrap Mass Spectrometry for Label-free Expression Profiling of the Swine Heart Mitochondrial Proteome. *Journal of proteome research,* 8**,** 2838.

DUARTE, N. C., BECKER, S. A., JAMSHIDI, N., THIELE, I., MO, M. L., VO, T. D., SRIVAS, R. & PALSSON, B. Ø. 2007. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 1777-1782.

DUNN, W. B. 2008. Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. *Physical Biology,* 5.

DUNN, W. B. & ELLIS, D. I. 2005. Metabolomics: Current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry,* 24**,** 285-294.

DUPORTET, X., AGGIO, R. B. M., CARNEIRO, S. & VILLAS-BÔAS, S. G. 2011. The biological interpretation of metabolomic data can be misled by the extraction method used. *Metabolomics***,** 1-12.

DUPORTET, X., AGGIO, R. B. M., CARNEIRO, S. & VILLAS-BÔAS, S. G. 2012. The biological interpretation of metabolomic data can be misled by the extraction method used. *Metabolomics,* 8**,** 410-421.

DWIVEDI, P., WU, P., KLOPSCH, S. J., PUZON, G. J., XUN, L. & HILL, H. H. 2008. Metabolic profiling by Ion Mobility Mass Spectrometry (IMMS). *Metabolomics,* 4**,** 63-80.

EDWARDS, J. L. & KENNEDY, R. T. 2005. Metabolomic Analysis of Eukaryotic Tissue and Prokaryotes Using Negative Mode MALDI Time-of-Flight Mass Spectrometry. *Analytical Chemistry,* 77**,** 2201-2209.

EHIMEN, E., SUN, Z., CARRINGTON, C., BIRCH, E. & EATON-RYE, J. 2011. Anaerobic digestion of microalgae residues resulting from the biodiesel production process. *Applied Energy,* 88**,** 3454-3463.

EJIKEME, P., ANYAOGU, I., EJIKEME, C., NWAFOR, N., EGBUONU, C., UKOGU, K. & IBEMESI, J. 2010. Catalysis in biodiesel production by transesterification processes-An insight. *Journal of Chemistry,* 7**,** 1120-1132.

ERBAN, A., SCHAUER, N., FERNIE, A. R. & KOPKA, J. 2007. Nonsupervised construction and application of mass spectral and retention time index libraries from time-of-flight gas chromatography-mass spectrometry metabolite profiles. *Methods in molecular biology (Clifton, N.J.),* 358**,** 19-38.

FAHY, E., SUBRAMANIAM, S., BROWN, H. A., GLASS, C. K., MERRILL, A. H., MURPHY, R. C., RAETZ, C. R., RUSSELL, D. W., SEYAMA, Y. & SHAW, W. 2005. A comprehensive classification system for lipids. *Journal of lipid research,* 46**,** 839-862.

FAIJES, M., MARS, A. E. & SMID, E. J. 2007. Comparison of quenching and extraction methodologies for metabolome analysis of Lactobacillus plantarum. *Microbial cell factories,* 6**,** 27.

FIEHN, O. 2002. Metabolomics - The link between genotypes and phenotypes. *Plant Molecular Biology,* 48**,** 155-171.

FIEHN, O. & KIND, T. 2007. Metabolite profiling in blood plasma. *Methods in molecular biology (Clifton, N.J.),* 358**,** 3-17.

FIEHN, O., KOPKA, J., DORMANN, P., ALTMANN, T., TRETHEWEY, R. N. & WILLMITZER, L. 2000. Metabolite profiling for plant functional genomics. *Nat Biotech,* 18**,** 1157-1161.

FIEHN, O., KRISTAL, B., OMMEN, B. V., SUMNER, L. W., SANSONE, S.-A., TAYLOR, C., HARDY, N. & KADDURAH-DAOUK, R. 2006. Establishing reporting standards for metabolomic and metabonomic studies: a call for participation. *Omics: a journal of integrative biology,* 10**,** 158-163.

FIEHN, O., ROBERTSON, D., GRIFFIN, J., VAB DER WERF, M., NIKOLAU, B., MORRISON, N., SUMNER, L. W., GOODACRE, R., HARDY, N. W., TAYLOR, C., FOSTEL, J., KRISTAL, B., KADDURAH-DAOUK, R., MENDES, P., VAN OMMEN, B., LINDON, J. C. & SANSONE, S. A. 2007. The metabolomics standards initiative (MSI). *Metabolomics,* 3**,** 175-178.

FIELD, C. B., BEHRENFELD, M. J., RANDERSON, J. T. & FALKOWSKI, P. 1998. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science,* 281**,** 237-240.

FOLCH, J., LEES, M. & SLOANE-STANLEY, G. 1957. A simple method for the isolation and purification of total lipids from animal tissues. *J. biol. Chem,* 226**,** 497-509.

FORTESCHI, M., SOTGIA, S., PINTUS, G., ZINELLU, A. & CARRU, C. 2014. Simultaneous determination of citrulline and arginine in human blood plasma by capillary electrophoresis with ultraviolet absorption detection. *Journal of separation science*.

GARCIA, D. E., BAIDOO, E. E., BENKE, P. I., PINGITORE, F., TANG, Y. J., VILLA, S. & KEASLING, J. D. 2008. Separation and mass spectrometry in microbial metabolomics. *Current Opinion in Microbiology,* 11**,** 233-239.

GERKEN, H. G., DONOHOE, B. & KNOSHAUG, E. P. 2013. Enzymatic cell wall degradation of Chlorella vulgaris and other microalgae for biofuels production. *Planta,* 237**,** 239-253.

GLUZ, O., LIEDTKE, C., GOTTSCHALK, N., PUSZTAI, L., NITZ, U. & HARBECK, N. 2009. Triple-negative breast cancer—current status and future directions. *Annals of Oncology,* 20**,** 1913-1927.

GO, E. P. 2010. Database resources in metabolomics: An overview. *Journal of Neuroimmune Pharmacology,* 5**,** 18-30.

GOH, S.-H., YUSOFF, F. M. & LOH, S.-P. 2010. A Comparison of the Antioxidant Properties and Total Phenolic Content in a Diatom, Chaetoceros sp. and a Green Microalga, Nannochloropsis sp. *Journal of Agricultural Science (1916-9752),* 2.

GONZALEZ, B., FRANÇOIS, J. & RENAUD, M. 1997. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast,* 13**,** 1347-1355.

GOODENOUGH, U. W. & HEUSER, J. E. 1985. The Chlamydomonas cell wall and its constituent glycoproteins analyzed by the quick-freeze, deep-etch technique. *The Journal of cell biology,* 101**,** 1550-1568.

GOWDA, G. A. N., ZHANG, S., GU, H., ASIAGO, V., SHANAIAH, N. & RAFTERY, D. 2008. Metabolomics-based methods for early disease diagnostics. *Expert Review of Molecular Diagnostics,* 8**,** 617-633.

GRIFFITHS, M., VAN HILLE, R. & HARRISON, S. 2010. Selection of direct transesterification as the preferred method for assay of fatty acid content of microalgae. *Lipids,* 45**,** 1053-1060.

GROSS, M., BUDCZIES, J., DEMO, S., JANES, J., LEWIS, E., PARLATI, F., MACKINNON, A., RODRIGUEZ, M., YANG, J. & ZHAO, F. 2013. Abstract P2-09-03: Antitumor activity of the glutaminase inhibitor, CB-839, in triple-negative breast cancer. *Cancer Research,* 73**,** P2-09-03-P2-09-03.

GROSSI, V., BLOKKER, P. & SINNINGHE DAMSTÉ, J. S. 2001. Anaerobic biodegradation of lipids of the marine microalga< i> Nannochloropsis salina</i>. *Organic Geochemistry,* 32**,** 795-808.

GULLBERG, J., JONSSON, P., NORDSTRÖM, A., SJÖSTRÖM, M. & MORITZ, T. 2004a. Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of Arabidopsis thaliana samples in metabolomic studies with gas chromatography/mass spectrometry. *Analytical Biochemistry,* 331**,** 283-295.

GULLBERG, J., JONSSON, P., NORDSTRÖM, A., SJÖSTRÖM, M. & MORITZ, T. 2004b. Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of< i> Arabidopsis thaliana</i> samples in metabolomic studies with gas chromatography/mass spectrometry. *Analytical biochemistry,* 331**,** 283-295.

HAJJAJ, H., BLANC, P., GOMA, G. & FRANCOIS, J. 1998. Sampling techniques and comparative extraction procedures for quantitative determination of intra-and extracellular metabolites in filamentous fungi. *FEMS microbiology letters,* 164**,** 195-200.

HALIM, R., GLADMAN, B., DANQUAH, M. K. & WEBLEY, P. A. 2011. Oil extraction from microalgae for biodiesel production. *Bioresource technology,* 102**,** 178-185.

HALIM, R., HARUN, R., DANQUAH, M. K. & WEBLEY, P. A. 2012. Microalgal cell disruption for biofuel development. *Applied Energy,* 91**,** 116-121.

HALL, R. D. 2006. Plant metabolomics: from holistic hope, to hype, to hot topic. *New Phytologist,* 169**,** 453-468.

HARMANESCU, M. 2012. Comparative researches on two direct transmethylation without prior extraction methods for fatty acids analysis in vegetal matrix with low fat content. *Chemistry Central Journal,* 6**,** 1-7.

HILLER, J., FRANCO-LARA, E. & WEUSTER-BOTZ, D. 2007a. Metabolic profiling of &lt;i&gt;Escherichia coli&lt;/i&gt; cultivations: evaluation of extraction and metabolite analysis procedures. *Biotechnology Letters,* 29**,** 1169-1178.

HILLER, J., FRANCO-LARA, E. & WEUSTER-BOTZ, D. 2007b. Metabolic profiling of Escherichia coli cultivations: evaluation of extraction and metabolite analysis procedures. *Biotechnology letters,* 29**,** 1169-1178.

HIRANO, A., UEDA, R., HIRAYAMA, S. & OGUSHI, Y. 1997. CO< sub> 2</sub> fixation and ethanol production with microalgal photosynthesis and intracellular anaerobic fermentation. *Energy,* 22**,** 137-142.

HOLMES, E., RUEY, L. L., CLOAREC, O., COEN, M., TANG, H., MAIBAUM, E., BRUCE, S., CHAN, Q., ELLIOTT, P., STAMLER, J., WILSON, I. D., LINDON, J. C. & NICHOLSON, J. K. 2008. Detection of urinary drug metabolite (xenometabolome) signatures in molecular epidemiology studies via statistical total correlation (NMR) spectroscopy (Analytical Chemistry (2007) 79 (2629-2640)). *Analytical Chemistry,* 80**,** 6142-6143.

HORVÁTH, K., GRITTI, F., FAIRCHILD, J. N. & GUIOCHON, G. 2010. On the optimization of the shell thickness of superficially porous particles. *Journal of chromatography A,* 1217**,** 6373-6381.

HOTELLING, H. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology,* 24**,** 417-441.

HU, C.-C., LIN, J.-T., LU, F.-J., CHOU, F.-P. & YANG, D.-J. 2008. Determination of carotenoids in< i> Dunaliella salina</i> cultivated in Taiwan and antioxidant capacity of the algal carotenoid extract. *Food Chemistry,* 109**,** 439-446.

HUIE, C. 2002. A review of modern sample-preparation techniques for the extraction and analysis of medicinal plants. *Analytical and Bioanalytical Chemistry,* 373**,** 23-30.

HUTSCHENREUTHER, A., BIRKENMEIER, G., BIGL, M., KROHN, K. & BIRKEMEYER, C. 2013. Glycerophosphoglycerol, Beta-Alanine, and Pantothenic Acid as Metabolic Companions of Glycolytic Activity and Cell Migration in Breast Cancer Cell Lines. *Metabolites,* 3**,** 1084-1101.

HUTSCHENREUTHER, A., KIONTKE, A., BIRKENMEIER, G. & BIRKEMEYER, C. 2012a. Comparison of extraction conditions and normalization approaches for cellular metabolomics of adherent growing cells with GC-MS. *Analytical Methods,* 4**,** 1953-1963.

HUTSCHENREUTHER, A., KIONTKE, A., BIRKENMEIER, G. & BIRKEMEYER, C. 2012b. Comparison of extraction conditions and normalization approaches for cellular metabolomics of adherent growing cells with GCMS. *Anal. Methods*.

IMAM, S. H., BUCHANAN, M. J., SHIN, H.-C. & SNELL, W. J. 1985. The Chlamydomonas cell wall: characterization of the wall framework. *The Journal of cell biology,* 101**,** 1599-1607.

IVERSON, S. J., LANG, S. L. & COOPER, M. H. 2001. Comparison of the Bligh and Dyer and Folch methods for total lipid determination in a broad range of marine tissue. *Lipids,* 36**,** 1283-1287.

JAKI, B., FRANZBLAU, S., CHO, S. & PAULI, G. 2006. Development of an extraction method for mycobacterial metabolome analysis. *Journal of pharmaceutical and biomedical analysis,* 41**,** 196-200.

JAMERS, A., BLUST, R. & DE COEN, W. 2009a. < i> Omics</i> in algae: Paving the way for a systems biological understanding of algal stress phenomena? *Aquatic Toxicology,* 92**,** 114-121.

JAMERS, A., BLUST, R. & DE COEN, W. 2009b. Omics in algae: Paving the way for a systems biological understanding of algal stress phenomena? *Aquatic Toxicology,* 92**,** 114-121.

JENSEN, N. B. S., JOKUMSEN, K. V. & VILLADSEN, J. 1999. Determination of the phosphorylated sugars of the Embden-Meyerhoff-Parnas pathway in Lactococcus lactis using a fast sampling technique and solid phase extraction. *Biotechnology and Bioengineering,* 63**,** 356-362.

JOHNSON, J. P., KUMAR, P., KOULNIS, M., PATEL, M. & SIMIN, K. 2014. Crucial and Novel Cancer Drivers in a Mouse Model of Triple-negative Breast Cancer. *Cancer Genomics-Proteomics,* 11**,** 115-126.

JOHNSON, M. B. & WEN, Z. 2009. Production of biodiesel fuel from the microalga Schizochytrium limacinum by direct transesterification of algal biomass. *Energy & Fuels,* 23**,** 5179-5183.

JONES, O. A. & HÜGEL, H. M. 2013. Bridging the Gap: Basic Metabolomics Methods for Natural Product Chemistry. *Metabolomics Tools for Natural Product Discovery.* Springer.

KANEHISA, M. & GOTO, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research,* 28**,** 27-30.

KAUFMANN, B. & CHRISTEN, P. 2002. Recent extraction techniques for natural products: microwave-assisted extraction and pressurised solvent extraction. *Phytochemical Analysis,* 13**,** 105-113.

KELL, D. B., BROWN, M., DAVEY, H. M., DUNN, W. B., SPASIC, I. & OLIVER, S. G. 2005. Metabolic footprinting and systems biology: the medium is the message. *Nature Reviews Microbiology,* 3**,** 557-565.

KELL, D. B. & OLIVER, S. G. 2004. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *Bioessays,* 26**,** 99-105.

KEMPA, S., HUMMEL, J., SCHWEMMER, T., PIETZKE, M., STREHMEL, N., WIENKOOP, S., KOPKA, J. & WECKWERTH, W. 2009a. An automated GCxGC-TOF-MS protocol for batch-wise extraction and alignment of mass isotopomer matrixes from differential 13C-labelling experiments: a case study for photoautotrophic-mixotrophic grown Chlamydomonas reinhardtii cells. *Journal of Basic Microbiology,* 49**,** 82-91.

KEMPA, S., HUMMEL, J., SCHWEMMER, T., PIETZKE, M., STREHMEL, N., WIENKOOP, S., KOPKA, J. & WECKWERTH, W. 2009b. An automated GCxGC-TOF-MS protocol for batch-wise extraction and alignment of mass isotopomer matrixes from differential 13C-labelling experiments: a case study for photoautotrophic-mixotrophic grown Chlamydomonas reinhardtii cells. *Journal of basic microbiology,* 49**,** 82-91.

KHOO, S. H. G. & AL-RUBEAI, M. 2007. Metabolomics as a complementary tool in cell culture. *Biotechnology and applied biochemistry,* 47**,** 71-84.

KIM, S., LEE, D. Y., WOHLGEMUTH, G., PARK, H. S., FIEHN, O. & KIM, K. H. 2013. Evaluation and optimization of metabolome sample preparation methods for Saccharomyces cerevisiae. *Analytical chemistry,* 85**,** 2169-2176.

KIMBALL, E. & RABINOWITZ, J. D. 2006. Identifying decomposition products in extracts of cellular metabolites. *Analytical biochemistry,* 358**,** 273-280.

KITANO, H. 2002. Systems biology: a brief overview. *Science,* 295**,** 1662-1664.

KLUENDER, C., SANS-PICHÉ, F., RIEDL, J., ALTENBURGER, R., HÄRTIG, C., LAUE, G. & SCHMITT-JANSEN, M. 2009a. A metabolomics approach to assessing phytotoxic effects on the green alga &lt;i&gt;Scenedesmus vacuolatus&lt;/i&gt. *Metabolomics,* 5**,** 59-71.

KLUENDER, C., SANS-PICHÉ, F., RIEDL, J., ALTENBURGER, R., HÄRTIG, C., LAUE, G. & SCHMITT-JANSEN, M. 2009b. A metabolomics approach to assessing phytotoxic effects on the green alga Scenedesmus vacuolatus. *Metabolomics,* 5**,** 59-71.

KNOTHE, G. 2005. Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. *Fuel processing technology,* 86**,** 1059-1070.

KOEK, M. M., JELLEMA, R. H., VAN DER GREEF, J., TAS, A. C. & HANKEMEIER, T. 2010. Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives. *Metabolomics***,** 1-22.

KOEK, M. M., MUILWIJK, B., VAN DER WERF, M. J. & HANKEMEIER, T. 2006. Microbial Metabolomics with Gas Chromatography/Mass Spectrometry. *Analytical Chemistry,* 78**,** 1272-1281.

KONING, W. D. & DAM, K. V. 1992. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Analytical Biochemistry,* 204**,** 118-123.

KOPKA, J., OHLROGGE, J. B. & JAWORSKI, J. G. 1995. Analysis of in vivo levels of acyl-thioesters with gas chromatography/mass spectrometry of the butylamide derivative. *Analytical biochemistry,* 224**,** 51-60.

KORTZ, L., HELMSCHRODT, C. & CEGLAREK, U. 2011. Fast liquid chromatography combined with mass spectrometry for the analysis of metabolites and proteins in human body fluids. *Analytical and bioanalytical chemistry,* 399**,** 2635-2644.

KRONTHALER, J., GSTRAUNTHALER, G. & HEEL, C. 2012a. Optimizing high-throughput metabolomic biomarker screening: A study of quenching solutions to freeze intracellular metabolism in CHO cells. *OMICS A Journal of Integrative Biology,* 16**,** 90-97.

KRONTHALER, J., GSTRAUNTHALER, G. & HEEL, C. 2012b. Optimizing high-throughput metabolomic biomarker screening: a study of quenching solutions to freeze intracellular metabolism in CHO cells. *Omics: a journal of integrative biology,* 16**,** 90-97.

KUMARI, P., REDDY, C. & JHA, B. 2011. Comparative evaluation and selection of a method for lipid and fatty acid extraction from macroalgae. *Analytical biochemistry,* 415**,** 134-144.

KUSSMANN, M., RAYMOND, F. & AFFOLTER, M. 2006. OMICS-driven biomarker discovery in nutrition and health. *Journal of Biotechnology,* 124**,** 758-787.

LAMERS, P. P., JANSSEN, M., DE VOS, R. C. H., BINO, R. J. & WIJFFELS, R. H. 2008. Exploring and exploiting carotenoid accumulation in Dunaliella salina for cell-factory applications. *Trends in Biotechnology,* 26**,** 631-638.

LAMERS, P. P., VAN DE LAAK, C. C., KAASENBROOD, P. S., LORIER, J., JANSSEN, M., DE VOS, R. C., BINO, R. J. & WIJFFELS, R. H. 2010a. Carotenoid and fatty acid metabolism in light-stressed Dunaliella salina. *Biotechnology and bioengineering,* 106**,** 638-648.

LAMERS, P. P., VAN DE LAAK, C. C. W., KAASENBROOD, P. S., LORIER, J., JANSSEN, M., DE VOS, R. C. H., BINO, R. J. & WIJFFELS, R. H. 2010b. Carotenoid and fatty acid metabolism in light-stressed Dunaliella salina. *Biotechnology and Bioengineering,* 106**,** 638-648.

LANE, A. N. & FAN, T. W. M. 2007. Quantification and identification of isotopomer distributions of metabolites in crude cell extracts using H-1 TOCSY. *Metabolomics,* 3**,** 79-86.

LAURENS, L. M., DEMPSTER, T. A., JONES, H. D., WOLFRUM, E. J., VAN WYCHEN, S., MCALLISTER, J. S., RENCENBERGER, M., PARCHERT, K. J. & GLOE, L. M. 2012a. Algal biomass constituent analysis: method uncertainties and investigation of the underlying measuring chemistries. *Analytical chemistry,* 84**,** 1879-1887.

LAURENS, L. M., QUINN, M., VAN WYCHEN, S., TEMPLETON, D. W. & WOLFRUM, E. J. 2012b. Accurate and reliable quantification of total microalgal fuel potential as fatty acid methyl esters by in situ transesterification. *Analytical and bioanalytical chemistry,* 403**,** 167-178.

LEAVELL, M., LEARY, J. & YAMASAKI, R. 2002. Mass spectrometric strategy for the characterization of lipooligosaccharides from &lt;i&gt;Neisseria gonorrhoeae&lt;/i&gt; 302 using FTICR. *Journal of The American Society for Mass Spectrometry,* 13**,** 571-576.

LEE, D. Y. & FIEHN, O. 2008a. High quality metabolomic data for Chlamydomonas reinhardtii. *Plant methods,* 4**,** 7.

LEE, D. Y. & FIEHN, O. 2008b. High quality metabolomic data for Chlamydomonas reinhardtii. *Plant Methods,* 4.

LEE, J.-Y., YOO, C., JUN, S.-Y., AHN, C.-Y. & OH, H.-M. 2010. Comparison of several methods for effective lipid extraction from microalgae. *Bioresource technology,* 101**,** S75-S77.

LEPAGE, G. & ROY, C. C. 1984. Improved recovery of fatty acid through direct transesterification without prior extraction or purification. *Journal of lipid research,* 25**,** 1391-1396.

LI, C. I., MIRUS, J. E., ZHANG, Y., RAMIREZ, A. B., LADD, J. J., PRENTICE, R. L., MCINTOSH, M. W., HANASH, S. M. & LAMPE, P. D. 2012. Discovery and preliminary confirmation of novel early detection biomarkers for triple-negative breast cancer using preclinical plasma samples from the Women's Health Initiative observational study. *Breast cancer research and treatment,* 135**,** 611-618.

LI, H.-B., CHENG, K.-W., WONG, C.-C., FAN, K.-W., CHEN, F. & JIANG, Y. 2007. Evaluation of antioxidant capacity and total phenolic content of different fractions of selected microalgae. *Food Chemistry,* 102**,** 771-776.

LINDON, J. C., HOLMES, E. & NICHOLSON, J. K. 2004. Metabonomics: Systems biology in pharmaceutical research and development. *Current Opinion in Molecular Therapeutics,* 6**,** 265-272.

LINK, H., ANSELMENT, B. & WEUSTER-BOTZ, D. 2008. Leakage of adenylates during cold methanol/glycerol quenching of Escherichia coli. *Metabolomics,* 4**,** 240-247.

LISEC, J., SCHAUER, N., KOPKA, J., WILLMITZER, L. & FERNIE, A. R. 2006. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protocols,* 1**,** 387-396.

LIU, C., YAN, F., GAO, H., HE, M., WANG, Z., CHENG, Y., DENG, X. & XU, J. 2014. Features of citrus terpenoid production as revealed by carotenoid, limonoid and aroma profiles of two pummelos (Citrus maxima) with different flesh color. *Journal of the science of food and agriculture*.

LONGWORTH, J. 2013. Proteomics in Microalgae: A Postgenomic Approach for Improved Biofuel Production.

LORENZ, M. A., BURANT, C. F. & KENNEDY, R. T. 2011. Reducing time and increasing sensitivity in sample preparation for adherent mammalian cell metabolomics. *Analytical chemistry,* 83**,** 3406-3414.

LU, H., GAN, D., ZHANG, Z. & LIANG, Y. 2011. Sample classification of GC-ToF-MS metabolomics data without the requirement for chromatographic deconvolution. *Metabolomics,* 7**,** 191-205.

LU, W., BENNETT, B. D. & RABINOWITZ, J. D. 2008a. Analytical strategies for LC-MS-based targeted metabolomics. *Journal of Chromatography B,* 871**,** 236-242.

LU, W., BENNETT, B. D. & RABINOWITZ, J. D. 2008b. Analytical strategies for LC–MS-based targeted metabolomics. *Journal of Chromatography B,* 871**,** 236-242.

LUO, B., GROENKE, K., TAKORS, R., WANDREY, C. & OLDIGES, M. 2007a. 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.

LUO, B., GROENKE, K., TAKORS, R., WANDREY, C. & OLDIGES, M. 2007b. 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.

LYNCH, D. V. & THOMPSON, G. A. 1982. Low temperature-induced alterations in the chloroplast and microsomal membranes of Dunaliella salina. *Plant physiology,* 69**,** 1369-1375.

MADLA, S., MIURA, D. & WARIISHI, H. 2012. Optimization of Extraction Method for GC-MS based Metabolomics for Filamentous Fungi. *Journal of Microbial & Biochemical Technology,* 4**,** 005-009.

MAMAS, M., DUNN, W. B., NEYSES, L. & GOODACRE, R. 2011. The role of metabolites and metabolomics in clinically applicable biomarkers of disease. *Archives of toxicology,* 85**,** 5-17.

MANIRAKIZA, P., COVACI, A. & SCHEPENS, P. 2001. Comparative study on total lipid determination using Soxhlet, Roese-Gottlieb, Bligh & Dyer, and modified Bligh & Dyer extraction methods. *Journal of food composition and analysis,* 14**,** 93-100.

MARCINOWSKA, R., TRYGG, J., WOLF-WATZ, H., MORTIZ, T. & SUROWIEC, I. 2011. Optimization of a sample preparation method for the metabolomic analysis of clinically relevant bacteria. *Journal of microbiological methods,* 87**,** 24-31.

MARQUETTE, C. & NABELL, L. 2012. Chemotherapy-resistant metastatic breast cancer. *Current treatment options in oncology,* 13**,** 263-275.

MARRUBINI, G., MENDOZA, B. E. C. & MASSOLINI, G. 2010. Separation of purine and pyrimidine bases and nucleosides by hydrophilic interaction chromatography. *Journal of separation science,* 33**,** 803-816.

MARTIN, F. P. J., DUMAS, M. E., WANG, Y., LEGIDO-QUIGLEY, C., YAP, I. K. S., TANG, H., ZIRAH, S., MURPHY, G. M., CLOAREC, O., LINDON, J. C., SPRENGER, N., FAY, L. B., KOCHHAR, S., VAN BLADEREN, P., HOLMES, E. & NICHOLSON, J. K. 2007. A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model. *Molecular Systems Biology,* 3.

MARTINEAU, E., TEA, I., LOAËC, G., GIRAUDEAU, P. & AKOKA, S. 2011. Strategy for choosing extraction procedures for NMR-based metabolomic analysis of mammalian cells. *Analytical and bioanalytical chemistry,* 401**,** 2133-2142.

MASHEGO, M. R., RUMBOLD, K., DE MEY, M., VANDAMME, E., SOETAERT, W. & HEIJNEN, J. J. 2007. Microbial metabolomics: Past, present and future methodologies. *Biotechnology Letters,* 29**,** 1-16.

MATA, T. M., MARTINS, A. A. & CAETANO, N. S. 2010. Microalgae for biodiesel production and other applications: a review. *Renewable and Sustainable Energy Reviews,* 14**,** 217-232.

MATTHEW, T., ZHOU, W., RUPPRECHT, J., LIM, L., THOMAS-HALL, S. R., DOEBBE, A., KRUSE, O., HANKAMER, B., MARX, U. C., SMITH, S. M. & SCHENK, P. M. 2009. The metabolome of Chlamydomonas reinhardtii following induction of anaerobic H2 production by sulfur depletion. *Journal of Biological Chemistry,* 284**,** 23415-23425.

MATYASH, V., LIEBISCH, G., KURZCHALIA, T. V., SHEVCHENKO, A. & SCHWUDKE, D. 2008. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *Journal of lipid research,* 49**,** 1137-1146.

MAY, P., CHRISTIAN, J.-O., KEMPA, S. & WALTHER, D. 2009. ChlamyCyc: an integrative systems biology database and web-portal for Chlamydomonas reinhardtii. *Bmc Genomics,* 10**,** 209.

MAY, P., WIENKOOP, S., KEMPA, S., USADEL, B., CHRISTIAN, N., RUPPRECHT, J., WEISS, J., RECUENCO-MUNOZ, L., EBENHÖH, O. & WECKWERTH, W. 2008a. Metabolomics-and proteomics-assisted genome annotation and analysis of the draft metabolic network of Chlamydomonas reinhardtii. *Genetics,* 179**,** 157-166.

MAY, P., WIENKOOP, S., KEMPA, S., USADEL, B., CHRISTIAN, N., RUPPRECHT, J., WEISS, J., RECUENCO-MUNOZ, L., EBENHÖH, O., WECKWERTH, W. & WALTHER, D. 2008b. Metabolomics- and Proteomics-Assisted Genome Annotation and Analysis of the Draft Metabolic Network of Chlamydomonas reinhardtii. *Genetics,* 179**,** 157-166.

MERCK SEQUANT, A. A practical guide to HILIC a tutorial and application book.

METALLO, C. M., WALTHER, J. L. & STEPHANOPOULOS, G. 2009. Evaluation of< sup> 13</sup> C isotopic tracers for metabolic flux analysis in mammalian cells. *Journal of biotechnology,* 144**,** 167-174.

MISHUR, R. J. & REA, S. L. 2011. Applications of mass spectrometry to metabolomics and metabonomics: Detection of biomarkers of aging and of age-related diseases. *Mass Spectrometry Reviews***,** n/a-n/a.

MOESTUE, S. A., BORGAN, E., HUUSE, E. M., LINDHOLM, E. M., SITTER, B., BØRRESEN-DALE, A.-L., ENGEBRAATEN, O., MÆLANDSMO, G. M. & GRIBBESTAD, I. S. 2010. Distinct choline metabolic profiles are associated with differences in gene expression for basal-like and luminal-like breast cancer xenograft models. *BMC cancer,* 10**,** 433.

MONK, B. C. 1988. The cell wall of Chlamydomonas reinhardtii gametes: composition, structure and autolysin-mediated shedding and dissolution. *Planta,* 176**,** 441-450.

MONTON, M. R. N. & SOGA, T. 2007. Metabolome analysis by capillary electrophoresis-mass spectrometry. *Journal of Chromatography A,* 1168**,** 237-246.

MORRIS, M. & WATKINS, S. M. 2005. Focused metabolomic profiling in the drug development process: advances from lipid profiling. *Current Opinion in Chemical Biology,* 9**,** 407-412.

MORRISON, W. R. & SMITH, L. M. 1964. Preparation of fatty acid methyl esters and dimethylacetals from lipids with boron fluoride–methanol. *Journal of lipid research,* 5**,** 600-608.

MURTHY, K., VANITHA, A., RAJESHA, J., SWAMY, M. M., SOWMYA, P. & RAVISHANKAR, G. A. 2005. In vivo antioxidant activity of carotenoids from< i> Dunaliella salina</i>—a green microalga. *Life Sciences,* 76**,** 1381-1390.

MYINT, K. T., UEHARA, T., AOSHIMA, K. & ODA, Y. 2009. Polar anionic metabolome analysis by nano-LC/MS with a metal chelating agent. *Analytical chemistry,* 81**,** 7766-7772.

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. & 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.

O'FALLON, J., BUSBOOM, J., NELSON, M. & GASKINS, C. 2007. A direct method for fatty acid methyl ester synthesis: application to wet meat tissues, oils, and feedstuffs. *Journal of Animal Science,* 85**,** 1511-1521.

OLDIGES, M., KUNZE, M., DEGENRING, D., SPRENGER, G. A. & TAKORS, R. 2004. Stimulation, Monitoring, and Analysis of Pathway Dynamics by Metabolic Profiling in the Aromatic Amino Acid Pathway. *Biotechnology Progress,* 20**,** 1623-1633.

OLIVER, S. G., WINSON, M. K., KELL, D. B. & BAGANZ, F. 1998. Systematic functional analysis of the yeast genome. *Trends in Biotechnology,* 16**,** 373-378.

ONG, E. S. 2002. Chemical assay of glycyrrhizin in medicinal plants by pressurized liquid extraction (PLE) with capillary zone electrophoresis (CZE). *Journal of Separation Science,* 25**,** 825-831.

OREN, A. 2005. A hundred years of Dunaliella research: 1905–2005. *Saline systems,* 1**,** 1-14.

OSSOVSKAYA, V., WANG, Y., BUDOFF, A., XU, Q., LITUEV, A., POTAPOVA, O., VANSANT, G., MONFORTE, J. & DARASELIA, N. 2011. Exploring molecular pathways of triple-negative breast cancer. *Genes & cancer,* 2**,** 870-879.

PAIK, M.-J., KIM, H., LEE, J., BRAND, J. & KIM, K.-R. 2009. Separation of triacylglycerols and free fatty acids in microalgal lipids by solid-phase extraction for separate fatty acid profiling analysis by gas chromatography. *Journal of Chromatography A,* 1216**,** 5917-5923.

PARK, C., LEE, Y.-J., LEE, S. Y., OH, H. B. & LEE, J. 2011. Determination of the intracellular concentrations of metabolites in Escherichia coli collected during the exponential and stationary growth phases using Liquid Chromatography-Mass Spectrometry. *Bull Korean Chem Soc,* 32**,** 524-530.

PARK, C., YUN, S., LEE, S. Y., PARK, K. & LEE, J. 2012. Metabolic Profiling of Klebsiella oxytoca: Evaluation of Methods for Extraction of Intracellular Metabolites Using UPLC/Q-TOF-MS. *Applied biochemistry and biotechnology,* 167**,** 425-438.

PATIL, P. D., GUDE, V. G., MANNARSWAMY, A., COOKE, P., NIRMALAKHANDAN, N., LAMMERS, P. & DENG, S. 2012. Comparison of direct transesterification of algal biomass under supercritical methanol and microwave irradiation conditions. *Fuel,* 97**,** 822-831.

PATTI, G. J. 2011. Separation strategies for untargeted metabolomics. *Journal of separation science,* 34**,** 3460-3469.

PATTI, G. J., TAUTENHAHN, R., RINEHART, D., CHO, K., SHRIVER, L. P., MANCHESTER, M., NIKOLSKIY, I., JOHNSON, C. H., MAHIEU, N. G. & SIUZDAK, G. 2012. A view from above: cloud plots to visualize global metabolomic data. *Analytical chemistry,* 85**,** 798-804.

PEROU, C. M., SØRLIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., REES, C. A., POLLACK, J. R., ROSS, D. T., JOHNSEN, H. & AKSLEN, L. A. 2000. Molecular portraits of human breast tumours. *Nature,* 406**,** 747-752.

PHILLIPS, K. M., TARRAGÓ-TRANI, M. T., GROVE, T. M., GRÜN, I., LUGOGO, R., HARRIS, R. F. & STEWART, K. K. 1997. Simplified gravimetric determination of total fat in food composites after chloroform-methanol extraction. *Journal of the American Oil Chemists' Society,* 74**,** 137-142.

PODO, F., BUYDENS, L., DEGANI, H., HILHORST, R., KLIPP, E., GRIBBESTAD, I. S., VAN HUFFEL, S., WM VAN LAARHOVEN, H., LUTS, J. & MONLEON, D. 2010. Triple-negative breast cancer: present challenges and new perspectives. *Molecular oncology,* 4**,** 209-229.

PÓL, J., HOHNOVÁ, B. & HYÖTYLÄINEN, T. 2007. Characterisation of Stevia Rebaudiana by comprehensive two-dimensional liquid chromatography time-of-flight mass spectrometry. *Journal of Chromatography A,* 1150**,** 85-92.

PRASAD MAHARJAN, R. & FERENCI, T. 2003a. Global metabolite analysis: the influence of extraction methodology on metabolome profiles of Escherichia coli. *Analytical Biochemistry,* 313**,** 145-154.

PRASAD MAHARJAN, R. & FERENCI, T. 2003b. Global metabolite analysis: the influence of extraction methodology on metabolome profiles of< i> Escherichia coli</i>. *Analytical biochemistry,* 313**,** 145-154.

PRINCE, J. T. & MARCOTTE, E. M. 2006. Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping. *Analytical chemistry,* 78**,** 6140-6152.

PROESTOS, C. & KOMAITIS, M. 2008. Application of microwave-assisted extraction to the fast extraction of plant phenolic compounds. *LWT - Food Science and Technology,* 41**,** 652-659.

PULZ, O. & GROSS, W. 2004. Valuable products from biotechnology of microalgae. *Applied microbiology and biotechnology,* 65**,** 635-648.

PURWAHA, P., LORENZI, P. L., SILVA, L. P., HAWKE, D. H. & WEINSTEIN, J. N. 2014. Targeted metabolomic analysis of amino acid response to L-asparaginase in adherent cells. *Metabolomics***,** 1-11.

RABINOWITZ, J. D. 2007. Cellular metabolomics of Escherchia coli. *Expert Review of Proteomics,* 4**,** 187-198.

RABINOWITZ, J. D. & KIMBALL, E. 2007. Acidic acetonitrile for cellular metabolome extraction from Escherichia coli. *Analytical chemistry,* 79**,** 6167-6173.

RAMAUTAR, R., SOMSEN, G. W. & DE JONG, G. J. 2009. CE-MS in metabolomics. *Electrophoresis,* 30**,** 276-291.

RENBERG, L., JOHANSSON, A. I., SHUTOVA, T., STENLUND, H., AKSMANN, A., RAVEN, J. A., GARDESTRÖM, P., MORITZ, T. & SAMUELSSON, G. 2010. A metabolomic approach to study major metabolite changes during acclimation to limiting CO2 in Chlamydomonas reinhardtii. *Plant Physiology,* 154**,** 187-196.

RITTER, J. B., GENZEL, Y. & REICHL, U. 2008. Simultaneous extraction of several metabolites of energy metabolism and related substances in mammalian cells: optimization using experimental design. *Analytical biochemistry,* 373**,** 349-369.

ROCHA, J. M. S., GARCIA, J. E. C. & HENRIQUES, M. H. F. 2003. Growth aspects of the marine microalga Nannochloropsis gaditana. *Biomolecular Engineering,* 20**,** 237-242.

ROESSNER, U., LUEDEMANN, A., BRUST, D., FIEHN, O., LINKE, T., WILLMITZER, L. & FERNIE, A. R. 2001. Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell,* 13**,** 11-29.

ROGIERS, V. 1977. The application of an improved gas-liquid chromatographic method for the determination of the long chain non-esterified fatty acid pattern of blood plasma in children. *Clinica Chimica Acta,* 78**,** 227-233.

ROUX, A., LISON, D., JUNOT, C. & HEILIER, J. F. 2011. Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: A review. *Clinical Biochemistry,* 44**,** 119-135.

RYAN, D. & ROBARDS, K. 2006. Metabolomics: The greatest omics of them all? *Analytical Chemistry,* 78**,** 7954-7958.

RYCKEBOSCH, E., MUYLAERT, K. & FOUBERT, I. 2012. Optimization of an analytical procedure for extraction of lipids from microalgae. *Journal of the American Oil Chemists' Society,* 89**,** 189-198.

SAITO, N., ROBERT, M., KITAMURA, S., BARAN, R., SOGA, T., MORI, H., NISHIOKA, T. & TOMITA, M. 2006. Metabolomics approach for enzyme discovery. *Journal of Proteome Research,* 5**,** 1979-1987.

SANA, T. R., WADDELL, K. & FISCHER, S. M. 2008. A sample extraction and chromatographic strategy for increasing LC/MS detection coverage of the erythrocyte metabolome. *Journal of Chromatography B,* 871**,** 314-321.

SANSONE, S.-A., FAN, T., GOODACRE, R., GRIFFIN, J. L., HARDY, N. W., KADDURAH-DAOUK, R., KRISTAL, B. S., LINDON, J., MENDES, P. & MORRISON, N. 2007. The metabolomics standards initiative. *Nature biotechnology,* 25**,** 846-848.

SCALBERT, A., BRENNAN, L., FIEHN, O., HANKEMEIER, T., KRISTAL, B. S., VAN OMMEN, B., PUJOS-GUILLOT, E., VERHEIJ, E., WISHART, D. & WOPEREIS, S. 2009. Mass-spectrometry-based metabolomics: Limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics,* 5**,** 435-458.

SCHÄDEL, F., DAVID, F. & FRANCO-LARA, E. 2011. Evaluation of cell damage caused by cold sampling and quenching for metabolome analysis. *Applied microbiology and biotechnology,* 92**,** 1261-1274.

SCHENK, P. M., THOMAS-HALL, S. R., STEPHENS, E., MARX, U. C., MUSSGNUG, J. H., POSTEN, C., KRUSE, O. & HANKAMER, B. 2008. Second generation biofuels: high-efficiency microalgae for biodiesel production. *Bioenergy Research,* 1**,** 20-43.

SCHOENMAKERS, P. J., VIVÓ-TRUYOLS, G. & DECROP, W. M. C. 2006. A protocol for designing comprehensive two-dimensional liquid chromatography separation systems. *Journal of Chromatography A,* 1120**,** 282-290.

SCHUCHARDT, U., SERCHELI, R. & VARGAS, R. M. 1998. Transesterification of vegetable oils: a review. *Journal of the Brazilian Chemical Society,* 9**,** 199-210.

SCHWEDE, S., KOWALCZYK, A., GERBER, M. & SPAN, R. Influence of different cell disruption techniques on mono digestion of algal biomass. World Renewable Energy Congress, Linkoping, Sweden, 2011. 8-13.

SCHWIEBERT, E. M. & ZSEMBERY, A. 2003. Extracellular ATP as a signaling molecule for epithelial cells. *Biochimica et Biophysica Acta (BBA)-Biomembranes,* 1615**,** 7-32.

SCOTT, S. A., DAVEY, M. P., DENNIS, J. S., HORST, I., HOWE, C. J., LEA-SMITH, D. J. & SMITH, A. G. 2010. Biodiesel from algae: challenges and prospects. *Current Opinion in Biotechnology,* 21**,** 277-286.

SELLICK, C. A., HANSEN, R., MAQSOOD, A. R., DUNN, W. B., STEPHENS, G. M., GOODACRE, R. & DICKSON, A. J. 2008. Effective Quenching Processes for Physiologically Valid Metabolite Profiling of Suspension Cultured Mammalian Cells. *Analytical Chemistry,* 81**,** 174-183.

SELLICK, C. A., HANSEN, R., STEPHENS, G. M., GOODACRE, R. & DICKSON, A. J. 2011. Metabolite extraction from suspension-cultured mammalian cells for global metabolite profiling. *nature protocols,* 6**,** 1241-1249.

SERIVE, B., KAAS, R., BÉRARD, J.-B., PASQUET, V., PICOT, L. & CADORET, J.-P. 2012. Selection and optimisation of a method for efficient metabolites extraction from microalgae. *Bioresource technology,* 124**,** 311-320.

SERKOVA, N. J. & NIEMANN, C. U. 2006. Pattern recognition and biomarker validation using quantitative 1H-NMR-based metabolomics. *Expert Review of Molecular Diagnostics,* 6**,** 717-731.

SHEIKH, K. D., KHANNA, S., BYERS, S. W., FORNACE JR, A. J. & CHEEMA, A. K. 2011. Small molecule metabolite extraction strategy for improving LC/MS detection of cancer cell metabolome. *Journal of Biomolecular Techniques: JBT,* 22**,** 1.

SHENG, J., VANNELA, R. & RITTMANN, B. E. 2011. Evaluation of methods to extract and quantify lipids from< i> Synechocystis</i> PCC 6803. *Bioresource technology,* 102**,** 1697-1703.

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.

SIEW MOI, P. 2004. Handbook of Microalgal Culture. Biotechnology and Applied Phycology. *Journal of Applied Phycology,* 16**,** 159-160.

SLOCOMBE, S. P., ZHANG, Q., BLACK, K. D., DAY, J. G. & STANLEY, M. S. 2013. Comparison of screening methods for high-throughput determination of oil yields in micro-algal biofuel strains. *Journal of applied phycology,* 25**,** 961-972.

SMEDES, F. 1999a. Determination of total lipid using non-chlorinated solvents. *Analyst,* 124**,** 1711-1718.

SMEDES, F. 1999b. Revisiting the development of the Bligh and Dyer total lipid determination method. *Marine Pollution Bulletin,* 38**,** 193-201.

SMITH, C. A., WANT, E. J., O'MAILLE, G., ABAGYAN, R. & SIUZDAK, G. 2006. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical chemistry,* 78**,** 779-787.

SMITS, H. P., COHEN, A., BUTTLER, T., NIELSEN, J. & OLSSON, L. 1998. Cleanup and Analysis of Sugar Phosphates in Biological Extracts by Using Solid-Phase Extraction and Anion-Exchange Chromatography with Pulsed Amperometric Detection. *Analytical Biochemistry,* 261**,** 36-42.

SOGA, T., OHASHI, Y., UENO, Y., NARAOKA, H., TOMITA, M. & NISHIOKA, T. 2003. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *Journal of proteome research,* 2**,** 488-494.

ŠOŠTARIČ, M., KLINAR, D., BRICELJ, M., GOLOB, J., BEROVIČ, M. & LIKOZAR, B. 2012. Growth, lipid extraction and thermal degradation of the microalga< i> Chlorella vulgaris</i>. *New biotechnology,* 29**,** 325-331.

SPAGOU, K., TSOUKALI, H., RAIKOS, N., GIKA, H., WILSON, I. D. & THEODORIDIS, G. 2010. Hydrophilic interaction chromatography coupled to MS for metabonomic/metabolomic studies. *Journal of separation science,* 33**,** 716-727.

SPRATLIN, J. L., SERKOVA, N. J. & ECKHARDT, S. G. 2009. Clinical applications of metabolomics in oncology: a review. *Clinical cancer research,* 15**,** 431-440.

SPURA, J., CHRISTIAN REIMER, L., WIELOCH, P., SCHREIBER, K., BUCHINGER, S. & SCHOMBURG, D. 2009. A method for enzyme quenching in microbial metabolome analysis successfully applied to gram-positive and gram-negative bacteria and yeast. *Analytical Biochemistry,* 394**,** 192-201.

STOURNAS, S., LOIS, E. & SERDARI, A. 1995. Effects of fatty acid derivatives on the ignition quality and cold flow of diesel fuel. *Journal of the American Oil Chemists' Society,* 72**,** 433-437.

SUMNER, L. W., AMBERG, A., BARRETT, D., BEALE, M. H., BEGER, R., DAYKIN, C. A., FAN, T. W.-M., FIEHN, O., GOODACRE, R. & GRIFFIN, J. L. 2007. Proposed minimum reporting standards for chemical analysis. *Metabolomics,* 3**,** 211-221.

SUMNER, L. W., MENDES, P. & DIXON, R. A. 2003. Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry,* 62**,** 817-836.

TAMBELLINI, N. P., ZAREMBERG, V., TURNER, R. J. & WELJIE, A. M. 2013. Evaluation of Extraction Protocols for Simultaneous Polar and Non-Polar Yeast Metabolite Analysis Using Multivariate Projection Methods. *Metabolites,* 3**,** 592-605.

TANG, H., ABUNASSER, N., GARCIA, M., CHEN, M., SIMON NG, K. & SALLEY, S. O. 2011. Potential of microalgae oil from< i> Dunaliella tertiolecta</i> as a feedstock for biodiesel. *Applied Energy,* 88**,** 3324-3330.

TANG, H. R. & WANG, Y. L. 2006. Metabonomics: A revolution in progress. *Progress in Biochemistry and Biophysics,* 33**,** 401-417.

TATKE, P. & JAISWAL, Y. 2011. An overview of microwave assisted extraction and its applications in herbal drug research. *Rese J Med Plant,* 5**,** 21-31.

TAUTENHAHN, R., BÖTTCHER, C. & NEUMANN, S. 2007. Annotation of LC/ESI-MS mass signals. *Bioinformatics Research and Development.* Springer.

TAUTENHAHN, R., BÖTTCHER, C. & NEUMANN, S. 2008. Highly sensitive feature detection for high resolution LC/MS. *BMC bioinformatics,* 9**,** 504.

TAUTENHAHN, R., PATTI, G. J., RINEHART, D. & SIUZDAK, G. 2012. XCMS Online: a web-based platform to process untargeted metabolomic data. *Analytical chemistry,* 84**,** 5035-5039.

TAYMAZ-NIKEREL, H., DE MEY, M., RAS, C., TEN PIERICK, A., SEIFAR, R. M., VAN DAM, J. C., HEIJNEN, J. J. & VAN GULIK, W. M. 2009. Development and application of a differential method for reliable metabolome analysis in< i> Escherichia coli</i>. *Analytical biochemistry,* 386**,** 9-19.

TENG, Q., HUANG, W., COLLETTE, T. W., EKMAN, D. R. & TAN, C. 2009. A direct cell quenching method for cell-culture based metabolomics. *Metabolomics,* 5**,** 199-208.

TEPHLY, T. R. 1991. The toxicity of methanol. *Life Sciences,* 48**,** 1031-1041.

THEISS, C., BOHLEY, P., BISSWANGER, H. & VOIGT, J. 2004. Uptake of polyamines by the unicellular green alga< i> Chlamydomonas reinhardtii</i> and their effect on ornithine decarboxylase activity. *Journal of plant physiology,* 161**,** 3-14.

THEISS, C., BOHLEY, P. & VOIGT, J. 2002. Regulation by polyamines of ornithine decarboxylase activity and cell division in the unicellular green alga Chlamydomonas reinhardtii. *Plant Physiology,* 128**,** 1470-1479.

THEOBALD, U., MAILINGER, W., REUSS, M. & RIZZI, M. 1993. In vivo analysis of glucose-induced fast changes in yeast adenine nucleotide pool applying a rapid sampling technique. *Analytical Biochemistry,* 214**,** 31-37.

THEODORIDIS, G., GIKA, H. G. & WILSON, I. D. 2008. LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics. *TrAC - Trends in Analytical Chemistry,* 27**,** 251-260.

THEODORIDIS, G. A., GIKA, H. G., WANT, E. J. & WILSON, I. D. 2012. Liquid chromatography–mass spectrometry based global metabolite profiling: a review. *Analytica chimica acta,* 711**,** 7-16.

TIAN, F., WANG, Y., SEILER, M. & HU, Z. 2014. Functional characterization of breast cancer using pathway profiles. *BMC medical genomics,* 7**,** 45.

TIZIANI, S., EINWAS, A.-H., LODI, A., LUDWIG, C., BUNCE, C. M., VIANT, M. R. & GUENTHER, U. L. 2008. Optimized metabolite extraction from blood serum for H-1 nuclear magnetic resonance spectroscopy. *Analytical Biochemistry,* 377**,** 16-23.

TOLSTIKOV, V. V., LOMMEN, A., NAKANISHI, K., TANAKA, N. & FIEHN, O. 2003. Monolithic Silica-Based Capillary Reversed-Phase Liquid Chromatography/Electrospray Mass Spectrometry for Plant Metabolomics. *Analytical Chemistry,* 75**,** 6737-6740.

TRAINA, T. A., O'SHAUGHNESSY, J., KELLY, C. M., SCHWARTZBERG, L., GUCALP, A., PETERSON, A., TUDOR, I. C., BLANEY, M., TRUDEAU, M. & HUDIS, C. A. 2013. A Phase 2 single-arm study of the clinical activity and safety of enzalutamide in patients with advanced androgen receptor-positive triple-negative breast cancer.

TREDWELL, G. D., EDWARDS-JONES, B., LEAK, D. J. & BUNDY, J. G. 2011. The development of metabolomic sampling procedures for Pichia pastoris, and baseline metabolome data. *PloS one,* 6**,** e16286.

TRYGG, J., GULLBERG, J., JOHANSSON, A. I., JONSSON, P., ANTTI, H., MARKLUND, S. L. & MORITZ, T. 2005. Extraction and GC/MS analysis of the human blood plasma metabolome. *Analytical chemistry,* 77**,** 8086-8094.

TWEEDDALE, H., NOTLEY-MCROBB, L. & FERENCI, T. 1998. Effect of slow growth on metabolism of Escherichia coli, as revealed by global metabolite pool ('metabolome') analysis. *Journal of Bacteriology,* 180**,** 5109-5116.

VAIDYANATHAN, S. & GOODACRE, R. 2007. Quantitative detection of metabolites using matrix-assisted laser desorption/ionization mass spectrometry with 9-aminoacridine as the matrix. *Rapid Communications in Mass Spectrometry,* 21**,** 2072-2078.

VAIDYANATHAN, S., KELL, D. B. & GOODACRE, R. 2002. Flow-injection electrospray ionization mass spectrometry of crude cell extracts for high-throughput bacterial identification. *J Am Soc Mass Spectrom,* 13**,** 118-128.

VAIDYANATHAN, S., ROWLAND, J. J., KELL, D. B. & GOODACRE, R. 2001. Discrimination of Aerobic Endospore-forming Bacteria via Electrospray-Ionization Mass Spectrometry of Whole Cell Suspensions. *Analytical Chemistry,* 73**,** 4134-4144.

VAN DER WERF, M., TAKORS, R., SMEDSGAARD, J., NIELSEN, J., FERENCI, T., PORTAIS, J., WITTMANN, C., HOOKS, M., TOMASSINI, A., OLDIGES, M., FOSTEL, J. & SAUER, U. 2007. Standard reporting requirements for biological samples in metabolomics experiments: microbial and in vitro biology experiments. *Metabolomics,* 3**,** 189-194.

VAN DER WERF, M. J., JELLEMA, R. H. & HANKEMEIER, T. 2005. Microbial metabolomics: Replacing trial-and-error by the unbiased selection and ranking of targets. *Journal of Industrial Microbiology and Biotechnology,* 32**,** 234-252.

VANDENBROUCK, T., JONES, O. A., DOM, N., GRIFFIN, J. L. & DE COEN, W. 2010a. Mixtures of similarly acting compounds in< i> Daphnia magna:</i> From gene to metabolite and beyond. *Environment international,* 36**,** 254-268.

VANDENBROUCK, T., JONES, O. A. H., DOM, N., GRIFFIN, J. L. & DE COEN, W. 2010b. Mixtures of similarly acting compounds in Daphnia magna: From gene to metabolite and beyond. *Environment International,* 36**,** 254-268.

VEYEL, D., ERBAN, A., FEHRLE, I., KOPKA, J. & SCHRODA, M. 2014. Rationales and Approaches for Studying Metabolism in Eukaryotic Microalgae. *Metabolites,* 4**,** 184-217.

VIANT, M. R. 2007a. Metabolomics of aquatic organisms: the new'omics' on the block. *Marine Ecology Progress Series,* 332**,** 301-306.

VIANT, M. R. 2007b. Metabolomics of aquatic organisms: The new 'omics' on the block. *Marine Ecology Progress Series,* 332**,** 301-306.

VIJAYENDRAN, C., BARSCH, A., FRIEHS, K., NIEHAUS, K., BECKER, A. & FLASCHEL, E. 2008. Perceiving molecular evolution processes in Escherichia coli by comprehensive metabolite and gene expression profiling. *Genome Biology,* 9.

VILLAS-BÔAS, S. G. & BRUHEIM, P. 2007. Cold glycerol–saline: The promising quenching solution for accurate intracellular metabolite analysis of microbial cells. *Analytical Biochemistry,* 370**,** 87-97.

VILLAS-BÔAS, S. G., HØJER-PEDERSEN, J., ÅKESSON, M., SMEDSGAARD, J. & NIELSEN, J. 2005a. Global metabolite analysis of yeast: evaluation of sample preparation methods. *Yeast,* 22**,** 1155-1169.

VILLAS-BÔAS, S. G., MAS, S., ÅKESSON, M., SMEDSGAARD, J. & NIELSEN, J. 2005b. Mass spectrometry in metabolome analysis. *Mass Spectrometry Reviews,* 24**,** 613-646.

VILLAS-BÔAS, S. G., HØJER-PEDERSEN, J., ÅKESSON, M., SMEDSGAARD, J. & NIELSEN, J. 2005. Global metabolite analysis of yeast: evaluation of sample preparation methods. *Yeast,* 22**,** 1155-1169.

VOIGT, J., DEINERT, B. & BOHLEY, P. 2000. Subcellular localization and light-dark control of ornithine decarboxylase in the unicellular green alga Chlamydomonas reinhardtii. *Physiologia Plantarum,* 108**,** 353-360.

VOIGT, J. & FRANK, R. 2003. 14-3-3 proteins are constituents of the insoluble glycoprotein framework of the Chlamydomonas cell wall. *The Plant Cell Online,* 15**,** 1399-1413.

VOLMER, M., GETTMANN, J., SCHOLZ, S., BÜNTEMEYER, H. & NOLL, T. A method for metabolomic sampling of suspended animal cells using fast filtration. BMC proceedings, 2011. BioMed Central Ltd, P93.

WALSH, A., COOK, R. S., REXER, B., ARTEAGA, C. L. & SKALA, M. C. 2012. Optical imaging of metabolism in HER2 overexpressing breast cancer cells. *Biomedical optics express,* 3**,** 75-85.

WECKWERTH, W., WENZEL, K. & FIEHN, O. 2004a. Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics,* 4**,** 78-83.

WECKWERTH, W., WENZEL, K. & FIEHN, O. 2004b. Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics,* 4**,** 78-83.

WELLERDIEK, M., WINTERHOFF, D., REULE, W., BRANDNER, J. & OLDIGES, M. 2009a. Metabolic quenching of &lt;i&gt;Corynebacterium glutamicum&lt;/i&gt;: efficiency of methods and impact of cold shock. *Bioprocess and Biosystems Engineering,* 32**,** 581-592.

WELLERDIEK, M., WINTERHOFF, D., REULE, W., BRANDNER, J. & OLDIGES, M. 2009b. Metabolic quenching of Corynebacterium glutamicum: efficiency of methods and impact of cold shock. *Bioprocess and biosystems engineering,* 32**,** 581-592.

WERF, M. J. V. D., OVERKAMP, K. M., MUILWIJK, B., COULIER, L. & HANKEMEIER, T. 2007. Microbial metabolomics: Toward a platform with full metabolome coverage. *Analytical Biochemistry,* 370**,** 17-25.

WEUSTER-BOTZ, D. 1999. *Die Rolle der Reaktionstechnik in der mikrobiellen Verfahrensentwicklung*, Forschungszentrum Jülich, Zentralbibliothek.

WIENDAHL, C., BRANDNER, J., KÜPPERS, C., LUO, B., SCHYGULLA, U., NOLL, T. & OLDIGES, M. 2007. A microstructure heat exchanger for quenching the metabolism of mammalian cells. *Chemical engineering & technology,* 30**,** 322-328.

WIENKOOP, S., WEI, MAY, P., KEMPA, S., IRGANG, S., RECUENCO-MUNOZ, L., PIETZKE, M., SCHWEMMER, T., RUPPRECHT, J., EGELHOFER, V. & WECKWERTH, W. 2010a. Targeted proteomics for Chlamydomonas reinhardtii combined with rapid subcellular protein fractionation, metabolomics and metabolic flux analyses. *Molecular BioSystems,* 6**,** 1018-1031.

WIENKOOP, S., WEIß, J., MAY, P., KEMPA, S., IRGANG, S., RECUENCO-MUNOZ, L., PIETZKE, M., SCHWEMMER, T., RUPPRECHT, J. & EGELHOFER, V. 2010b. Targeted proteomics for Chlamydomonas reinhardtii combined with rapid subcellular protein fractionation, metabolomics and metabolic flux analyses. *Molecular BioSystems,* 6**,** 1018-1031.

WIJESEKARA, I., PANGESTUTI, R. & KIM, S.-K. 2011. Biological activities and potential health benefits of sulfated polysaccharides derived from marine algae. *Carbohydrate polymers,* 84**,** 14-21.

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.

WITTMANN, C. 2007. Fluxome analysis using GC-MS. *Microbial cell factories,* 6**,** 6.

WITTMANN, C., KRÖMER, J. O., KIEFER, P., BINZ, T. & HEINZLE, E. 2004. Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. *Analytical biochemistry,* 327**,** 135-139.

WU, L., VAN WINDEN, W. A., VAN GULIK, W. M. & HEIJNEN, J. J. 2005. Application of metabolome data in functional genomics: A conceptual strategy. *Metabolic Engineering,* 7**,** 302-310.

WURM, M. & ZENG, A. P. Estimation of protein and metabolite release rates from damaged mammalian cells by using GFP as a marker molecule. *Metabolomics***,** 1-9.

XU, F., ZOU, L., LIU, Y., ZHANG, Z. & ONG, C. N. 2011. Enhancement of the capabilities of liquid chromatography–mass spectrometry with derivatization: general principles and applications. *Mass spectrometry reviews,* 30**,** 1143-1172.

XU, Z., HARVEY, K., PAVLINA, T., DUTOT, G., ZALOGA, G. & SIDDIQUI, R. 2010. An improved method for determining medium-and long-chain FAMEs using gas chromatography. *Lipids,* 45**,** 199-208.

YANES, O., TAUTENHAHN, R., PATTI, G. J. & SIUZDAK, G. 2011. Expanding coverage of the metabolome for global metabolite profiling. *Analytical chemistry,* 83**,** 2152-2161.

YUAN, W., ANDERSON, K. W., LI, S. & EDWARDS, J. L. 2012. Subsecond Absolute Quantitation of Amine Metabolites Using Isobaric Tags for Discovery of Pathway Activation in Mammalian Cells. *Analytical chemistry,* 84**,** 2892-2899.

ZHANG, A., SUN, H., WANG, P., HAN, Y. & WANG, X. 2012. Modern analytical techniques in metabolomics analysis. *Analyst,* 137**,** 293-300.

ZHANG, D., HUANG, X., REGNIER, F. E. & ZHANG, M. 2008. Two-Dimensional Correlation Optimized Warping Algorithm for Aligning GC×GC−MS Data. *Analytical Chemistry,* 80**,** 2664-2671.

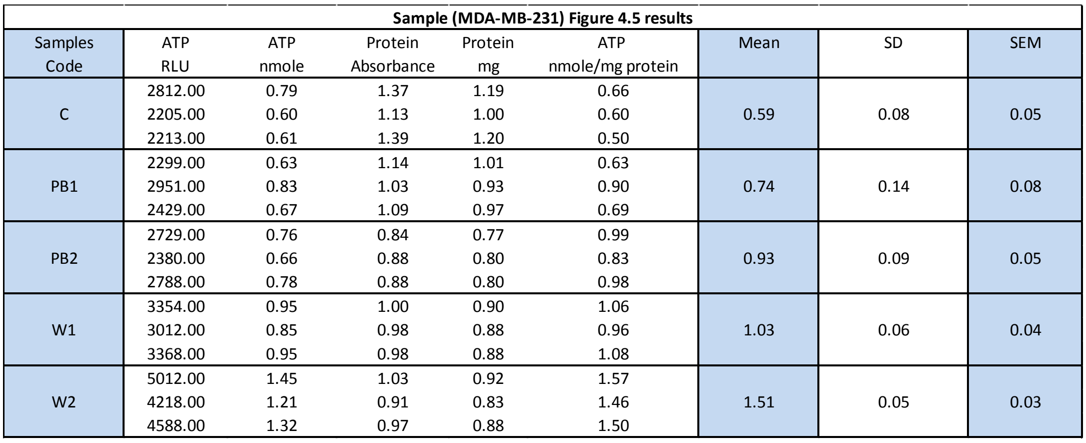
ZHAO, C., NAMBOU, K., WEI, L., CHEN, J., IMANAKA, T. & HUA, Q. 2014. Evaluation of metabolome sample preparation methods regarding leakage reduction for the oleaginous yeast< i> Yarrowia lipolytica</i>. *Biochemical Engineering Journal,* 82**,** 63-70.

Appendices

Chapter 4 Appendices

Appendix 4.1 Raw data and calculation part for estimation of nM amounts of free ATP produced in response to various washing solutions/steps. ATP levels were normalised to the protein levels. (Sample calculation part for Figure 4.5)

Initially the ATP standard curve and BCA standard curves were generated as detailed in section 4.3.2 and 4.2.4. Later, the nM amounts of free ATP produced in response to various washing steps/solutions were normalised to the protein levels as detailed below.



Sample (MDA-MB-231) Figure 4.5 results

Samples Code	ATP RLU	ATP nmole	Protein Absorbance	Protein mg	ATP nmole/mg protein	Mean	SD	SEM
C	2812.00	0.79	1.37	1.19	0.66	0.59	0.08	0.05
	2205.00	0.60	1.13	1.00	0.60			
	2213.00	0.61	1.39	1.20	0.50			
PB1	2299.00	0.63	1.14	1.01	0.63	0.74	0.14	0.08
	2951.00	0.83	1.03	0.93	0.90			
	2429.00	0.67	1.09	0.97	0.69			
PB2	2729.00	0.76	0.84	0.77	0.99	0.93	0.09	0.05
	2380.00	0.66	0.88	0.80	0.83			
	2788.00	0.78	0.88	0.80	0.98			
W1	3354.00	0.95	1.00	0.90	1.06	1.03	0.06	0.04
	3012.00	0.85	0.98	0.88	0.96			
	3368.00	0.95	0.98	0.88	1.08			
W2	5012.00	1.45	1.03	0.92	1.57	1.51	0.05	0.03
	4218.00	1.21	0.91	0.83	1.46			
	4588.00	1.32	0.97	0.88	1.50			

Concentration of free ATP (nM) in response to each treatment was calculated using equation: $y = 0.0003x - 0.571$

Concentration of protein (mg) in response to each treatment was calculated using equation: $y = -0.0934x^2 + 1.0069x - 0.0143$

The nM amounts of free ATP produced in response to various washing solutions/steps were normalised to the protein levels within each whole-cell lysate, the results are summarised in figure 4.5.

Appendix 4.2 List of 140 metabolites identified across all treatments in adherently growing metastatic cancer cell line MDA-MB-231. All the identified metabolites were classified based on their physicochemical properties into 11 different metabolite classes (class ID)

Met ID	*Corrected Met ID's*	*Class ID*
1	2-Piperidinecarboxylic acid	1
2	3-Methyl-2-oxopentanoic-acid	1
3	Adipic acid	1
4	Butanoic acid	1
5	Citric acid	1
6	Erythronic acid	1
7	Fumaric acid	1
8	Glutaric acid	1
9	Glyoxylic acid	1
10	Gulonic acid	1
11	Iminodiacetic acid	1
12	Indole-3-acetic acid	1
13	Lactic acid	1
14	Malic acid	1
15	Mandelic acid	1
16	Oxalic acid	1
17	Pyruvic acid	1
18	Threonic acid	1
19	Allose	2
20	Arabitol	2
21	Erythritol	2
22	Fructose	2
23	Galactitol	2
24	Galactose	2
25	Gentiobiose	2
26	Glucose	2
27	Inositol	2
28	Kestose	2
29	Laminaribiose	2
30	Lyxose	2
31	Mannitol	2
32	Mannose	2
33	Psicose	2
34	Ribitol	2
35	Ribose	2
36	similar to Fructose Derivate	2
37	similar to Inositol	2
38	Sorbitol	2
39	Sorbose	2
40	Sucrose	2
41	Tagatose	2
42	Threitol	2
43	Threose	2
44	Xylitol	2
45	Xylose	2
46	Alanine	3
47	Aminomalonic acid	3
48	Asparagine	3
49	Aspartic acid	3
50	Cysteine	3
51	Glutamic acid	3
52	Glutamine	3
53	Glycine	3
54	Homoserine	3
55	Isoleucine	3
56	Leucine	3
57	Lysine	3
58	Methionine	3
59	N-Carboxyglycine	3
60	Norleucine	3
61	Ornithine	3
62	Phenylalanine	3
63	Proline	3
64	Prolyl-glycine	3
65	Pyroglutamic acid	3
66	Serine	3
67	Threonine	3
68	Tryptophan	3
69	Tyrosine	3
70	Valine	3
71	Nicotinamide	4
72	Purine	4
73	Pyridoxamine	4
74	Cholesterol	5
75	Cholesterol-5beta	5
76	Dodecanoic acid	5
77	Dodecanol	5
78	Eicosan-1-ol	5
79	Eicosatetraenoic acid	5
80	Heptadecan-1-ol	5
81	Hexadecanoic acid	5
82	Isocaproic acid	5
83	Isocitric acid	5
84	Octadecan-1-ol	5
85	Octadecanoic acid	5
86	Octadecenoic acid	5
87	Pantothenic acid	5

88	Pentadecan-1-ol	5
89	Pentadecanoic acid	5
90	Tetradecanoic acid	5
91	Tridecan-1-ol	5
92	Ethanolamine	6
93	Hypotaurine	6
94	Phenethylamine	6
95	Putrescine	6
96	Sphingosine	6
97	Triethanolamine	6
98	Urea	6
99	Ethanolaminephosphate	7
100	Glycerol-3-phosphate	7
101	Phosphoric acid	7
102	Pyrophosphate	7
103	Glycerol	9
104	Propane-1,2-diol	9
105	Flavone	10
106	Glucoheptonic acid-1,4-lactone	10
107	Phenylacetaldehyde	10
108	Creatinine	11
109	Maleimide	11
110	Quinazoline	11
111	similar to Ditertbutylphenol	11
112	Valero-1,5-lactam	11
113	UK1	11
114	UK2	11
115	UK3	11
116	UK4	11
117	UK5	11
118	UK6	11
119	UK7	11
120	UK8	11
121	UK9	11
122	UK10	11
123	UK11	11
124	UK12	11
125	UK13	11
126	UK14	11
127	UK15	11
128	UK16	11
129	UK17	11
130	UK18	11
131	UK19	11
132	UK20	11
133	UK21	11
134	UK22	11
135	UK23	11
136	UK24	11
137	UK25	11
138	UK26	11
139	UK27	11
140	UK28	11

Chapter 5 Appendices

Appendix 5.1 List of 157 metabolites identified in MDA-MB-231 and MDA-MB-436 with the NTcs and LN_2 methods

Met ID	***Metabolites***	***Class ID***
1	2-Ketoglutaric acid methoxyamine	1
2	2-Piperidinecarboxylic acid	1
3	Aconitic acid	1
4	Butanoic acid	1
5	Citramalic acid	1
6	Citric acid	1
7	Erythronic acid	1
8	Fumaric acid	1
9	Glucaric acid	1
10	Glutaric acid	1
11	Glyceric acid	1
12	Glyoxylic acid	1
13	Hexonic acid	1
14	Iminodiacetic acid	1
15	Isobutanoic acid	1
16	Itaconic acid	1
17	Lactic acid	1
18	Maleamic acid	1
19	Maleic acid	1
20	Malic acid	1
21	Malonic acid	1
22	Oxalic acid	1
23	Pyrrole-2-carboxylic acid	1
24	Pyruvic acid	1
25	Succinic acid	1
26	Threonic acid	1
27	Valeric acid	1
28	1-Methyl-beta-D-galactopyranoside	2
29	beta-D-Allose	2
30	Altrose	2
31	Arabitol	2
32	D-Pinitol	2
33	Erythritol	2
34	Fructose	2
35	Galactitol	2
36	Galactonic acid	2
37	Galactose	2
38	Glucose	2
39	Glucuronic acid methoxyamine	2
40	Gulose	2
41	Inositol	2
42	Lyxose	2
43	Maltose methoxyamine	2
44	Mannitol	2
45	Melibiose methoxyamine	2
46	myo-Inositol	2
47	Raffinose	2
48	Ribitol	2
49	Ribose	2
50	Sedoheptulose	2
51	similar to Fructose Derivate	2
52	similar to Inositol	2
53	Sorbitol	2
54	Sorbose	2
55	Sucrose	2
56	Threitol	2
57	Viburnitol	2
58	Xylitol	2
59	Xylose	2
60	4-Aminobutyric acid	3
61	4-Hydroxyproline	3
62	Alanine	3
63	Alanyl-alanine	3
64	Arginine [-NH3]	3
65	Asparagine	3
66	Aspartic acid	3
67	beta-Alanine	3
68	Cysteine	3
69	Glutamic acid	3
70	Glutamine [-H2O]	3
71	Glycine	3
72	Homoserine	3
73	Isoleucine	3
74	Leucine	3
75	Lysine	3
76	Methionine	3
77	N-Acetyl-L-serine	3
78	Norleucine	3
79	Ornithine	3

80	Phenylalanine	3
81	Proline	3
82	Prolyl-glycine	3
83	Pyroglutamic acid	3
84	Sarcosine	3
85	Serine	3
86	Threonine	3
87	Tryptophan	3
88	Tyrosine	3
89	Valine	3
90	Adenine	4
91	Cytosine	4
92	Dihydroorotic acid	4
93	Orotic acid	4
94	Pipecolic acid	4
95	Purine	4
96	Thymine	4
97	Uracil	4
98	Uridine	4
99	1-Monohexadecanoylglycerol	5
100	1-Monooctadecanoylglycerol	5
101	9-(Z)-Octadecenoic acid	5
102	Caproic acid	5
103	Cholesterol	5
104	Dodecanoic acid	5
105	Dodecanol	5
106	Eicosanoic acid	5
107	Heptadecan-1-ol	5
108	Hexadecanoic acid	5
109	Nonanoic acid	5
110	Octadecan-1-ol	5
111	Octadecanoic acid	5
112	Octadecenoic acid	5
113	Panthotenic acid	5
114	Pentadecan-1-ol	5
115	Tetradecanoic acid	5
116	Tridecan-1-ol	5
117	Ethanolamine	6
118	Hypotaurine	6
119	Hypoxanthine	6
120	Phenethylamine	6
121	Putrescine	6
122	Sphingosine	6
123	Triethanolamine	6
124	Tryptamine	6
125	Tyramine	6
126	Urea	6
127	Dihydroxyacetone phosphate	7
128	Fructose-6-phosphate methoxyamine	7
129	Gluconic acid-6-phosphate	7
130	Glucose-6-phosphate methoxyamine	7
131	Glyceraldehyde-3-phosphate	7
132	Glyceric acid-2,3-diphosphate	7
133	Glycerol-2-phosphate	7
134	Glycerol-3-phosphate	7
135	Glycolic acid-2-phosphate	7
136	Mannose-1-phosphate	7
137	Monomethylphosphate	7
138	myo-Inositol-2-phosphate	7
139	Phosphomycin	7
140	Phosphoric acid	7
141	Sorbitol-6-phosphate	7
142	Glycerol	9
143	Propane-1,2-diol	9
144	Butyro-1,4-lactam	10
145	Cyclohexene-1,2,3,4-tetrol epoxide	10
146	Glucoheptonic acid-1-4-lactone	10
147	Gulonic acid-1,4-lactone	10
148	Ornithine-1,5-lactam	10
149	2,4,6-Tri-tert.-butylbenzenethiol	11
150	Creatinine	11
151	UK10	11
152	UK17	11
153	UK18	11
154	UK24	11
155	UK4	11
156	UK6	11
157	UK7	11

Appendix 5.2 List of metabolites identified in MDA-MB-231 and MDA-MB-436 using two different quenching and extraction protocols namely modified direct cell scraping (NTcs) and direct quenching and extraction with LN_2 (LN_2) classified based on their corresponding pathway involvement along with their KEGG ID

1	1-Methyl-beta-D-galactopyranoside	C03619	ABC transporters
2	Erythritol	C00503	ABC transporters
3	Glutamic acid	C00025	ABC transporters
4	Ornithine	C00077	ABC transporters
5	Sorbitol	C00794	ABC transporters
6	*beta-D-Allose*	C01487	ABC transporters
7	Mannitol	C00392	ABC transporters
8	Melibiose methoxyamine	C05402	ABC transporters
9	Raffinose	C00492	ABC transporters
10	Sucrose	C00089	ABC transporters
11	Threonine	C00188	ABC transporters
12	Xylitol	C00379	ABC transporters
13	Xylose	C00181	ABC transporters
14	Ribose	C00121	ABC transporters
15	Phenylalanine	C00079	ABC transporters
16	Maltose methoxyamine	C00208	ABC transporters
17	*Arginine [-NH3]*	C00062	ABC transporters
18	Proline	C00148	ABC transporters
19	Glycerol	C00116	ABC transporters
20	Phosphoric acid	C00009	ABC transporters
21	*Serine*	C00065	ABC transporters
22	Glutamic acid	C00025	Alanine, aspartate and glutamate metabolism
23	Aspartic acid	C00049	Alanine, aspartate and glutamate metabolism
24	*Glutamine [-H2O]*	C00064	Alanine, aspartate and glutamate metabolism
25	4-Aminobutyric acid	C00334	Alanine, aspartate and glutamate metabolism
26	Alanine	C00041	Alanine, aspartate and glutamate metabolism
27	*Asparagine*	C00152	Alanine, aspartate and glutamate metabolism
28	Panthotenic acid	C00864	beta-Alanine metabolism
29	beta-Alanine	C00099	beta-Alanine metabolism
30	Ornithine	C00077	Arginine and proline metabolism
31	*Arginine [-NH3]*	C00062	Arginine and proline metabolism
32	Proline	C00148	Arginine and proline metabolism
33	Putrescine	C00134	Arginine and proline metabolism
34	Creatinine	C00791	Arginine and proline metabolism
35	*Sarcosine*	C00213	Arginine and proline metabolism
36	*Valeric acid*	C00431	Arginine and proline metabolism
37	Aspartic acid	C00049	Arginine and proline metabolism
38	Urea	C00086	Arginine and proline metabolism
39	Methionine	C00073	Cysteine and methionine metabolism
40	Cysteine	C00097	Cysteine and methionine metabolism
41	*Homoserine*	C00263	Cysteine and methionine metabolism
42	Aspartic acid	C00049	Cysteine and methionine metabolism
43	*Lysine*	C00047	Lysine biosynthesis
44	*Valeric acid*	C00431	Lysine degradation
45	Pipecolic acid	C00408	Lysine degradation
46	2-Piperidinecarboxylic acid	C00408	Lysine degradation
47	*Homoserine*	C00263	Lysine degradation
48	Glycine	C00037	Lysine degradation
49	Threonine	C00188	Glycine, serine and threonine metabolism
50	*Sarcosine*	C00213	Glycine, serine and threonine metabolism
51	*Homoserine*	C00263	Glycine, serine and threonine metabolism
52	Glycine	C00037	Glycine, serine and threonine metabolism
53	Tryptophan	C00078	Glycine, serine and threonine metabolism
54	*Serine*	C00065	Glycine, serine and threonine metabolism
55	Aspartic acid	C00049	Glycine, serine and threonine metabolism
56	Cysteine	C00097	Glycine, serine and threonine metabolism
57	Phenethylamine	C05332	Phenylalanine metabolism
58	Phenylalanine	C00079	Phenylalanine, tyrosine and tryptophan biosynthesis
59	Tyrosine	C00082	Phenylalanine, tyrosine and tryptophan biosynthesis
60	Tyramine	C00483	Tyrosine metabolism
61	*Tryptamine*	C00398	Tryptophan metabolism
62	Tryptophan	C00078	Tryptophan metabolism

63	Isoleucine	C00407	Valine, leucine and isoleucine biosynthesis
64	Leucine	C00123	Valine, leucine and isoleucine biosynthesis
65	Valine	C00183	Valine, leucine and isoleucine biosynthesis
66	Citramalic acid	C02612	Valine, leucine and isoleucine biosynthesis
67	*Serine*	C00065	Valine, leucine and isoleucine biosynthesis
68	*Malonic acid*	C00383	Valine, leucine and isoleucine biosynthesis
69	Isoleucine	C00407	Valine, leucine and isoleucine degradation
70	Leucine	C00123	Valine, leucine and isoleucine degradation
71	Valine	C00183	Valine, leucine and isoleucine degradation
72	Ribose	C00121	Pentose phosphate pathway
73	*Glyceric acid*	C00258	Pentose phosphate pathway
74	*Glyceraldehyde-3-phosphate*	C00118	Pentose phosphate pathway
75	*Dihydroxyacetone phosphate*	C00111	Pentose phosphate pathway
76	Glucose-6-phosphate methoxyamine	C00668	Pentose phosphate pathway
77	Glucose	C00031	Pentose phosphate pathway
78	Pyruvic acid	C00022	Pentose phosphate pathway
79	Xylitol	C00379	Pentose and glucuronate interconversions
80	Xylose	C00181	Pentose and glucuronate interconversions
81	Glucuronic acid methoxyamine	C00191	Pentose and glucuronate interconversions
82	Glycerol	C00116	Pentose and glucuronate interconversions
83	Arabitol	C00532	Pentose and glucuronate interconversions
84	Lyxose	C00476	Pentose and glucuronate interconversions
85	Ribitol	C00474	Pentose and glucuronate interconversions
86	Pyruvic acid	C00022	Citrate cycle (TCA cycle)
87	2-Ketoglutaric acid	C00026	Citrate cycle (TCA cycle)
88	*Aconitic acid (cis)*	C00417	Citrate cycle (TCA cycle)
89	Citric acid	C00158	Citrate cycle (TCA cycle)
90	Fumaric acid	C00122	Citrate cycle (TCA cycle)
91	Malic acid	C00149	Citrate cycle (TCA cycle)
92	*Succinic acid*	C00042	Citrate cycle (TCA cycle)
93	Glucose	C00031	Glycolysis / Gluconeogenesis
94	*Glyceraldehyde-3-phosphate*	C00118	Glycolysis / Gluconeogenesis
95	*Dihydroxyacetone phosphate*	C00111	Glycolysis / Gluconeogenesis
96	Glucose-6-phosphate methoxyamine	C00668	Glycolysis / Gluconeogenesis
97	Pyruvic acid	C00022	Glycolysis / Gluconeogenesis
98	Lactic acid	C00186	Glycolysis / Gluconeogenesis
99	Sorbitol	C00794	Galactose metabolism
100	*Glyceraldehyde-3-phosphate*	C00118	Galactose metabolism
101	*Dihydroxyacetone phosphate*	C00111	Galactose metabolism
102	Glucose-6-phosphate methoxyamine	C00668	Galactose metabolism
103	Melibiose methoxyamine	C05402	Galactose metabolism
104	Raffinose	C00492	Galactose metabolism
105	Sucrose	C00089	Galactose metabolism
106	myo-Inositol	C00137	Galactose metabolism
107	Galactitol	C01697	Galactose metabolism
108	Galactonic acid	C00880	Galactose metabolism
109	Galactose	C00124	Galactose metabolism
110	Fructose	C00095	Galactose metabolism
111	Glucose	C00031	Galactose metabolism
112	Glycerol	C00116	Galactose metabolism
113	Sorbitol	C00794	Fructose and mannose metabolism
114	*beta-D-Allose*	C01487	Fructose and mannose metabolism
115	Mannitol	C00392	Fructose and mannose metabolism
116	*Mannose-1-phosphate*	C00636	Fructose and mannose metabolism
117	Fructose	C00095	Fructose and mannose metabolism
118	Sorbitol-6-phosphate	C01096	Fructose and mannose metabolism
119	Sorbose	C00247	Fructose and mannose metabolism
120	*Glyceraldehyde-3-phosphate*	C00118	Fructose and mannose metabolism
121	*Glutamine [-H2O]*	C00064	Purine metabolism
122	Glycine	C00037	Purine metabolism
123	Urea	C00086	Purine metabolism
124	*Glyoxylic acid*	C00048	Purine metabolism
125	*Adenine*	C00147	Purine metabolism
126	*Hypoxanthine*	C00262	Purine metabolism

Chapter 6 Appendices

Appendix 6.1 The use of AMDIS in practice

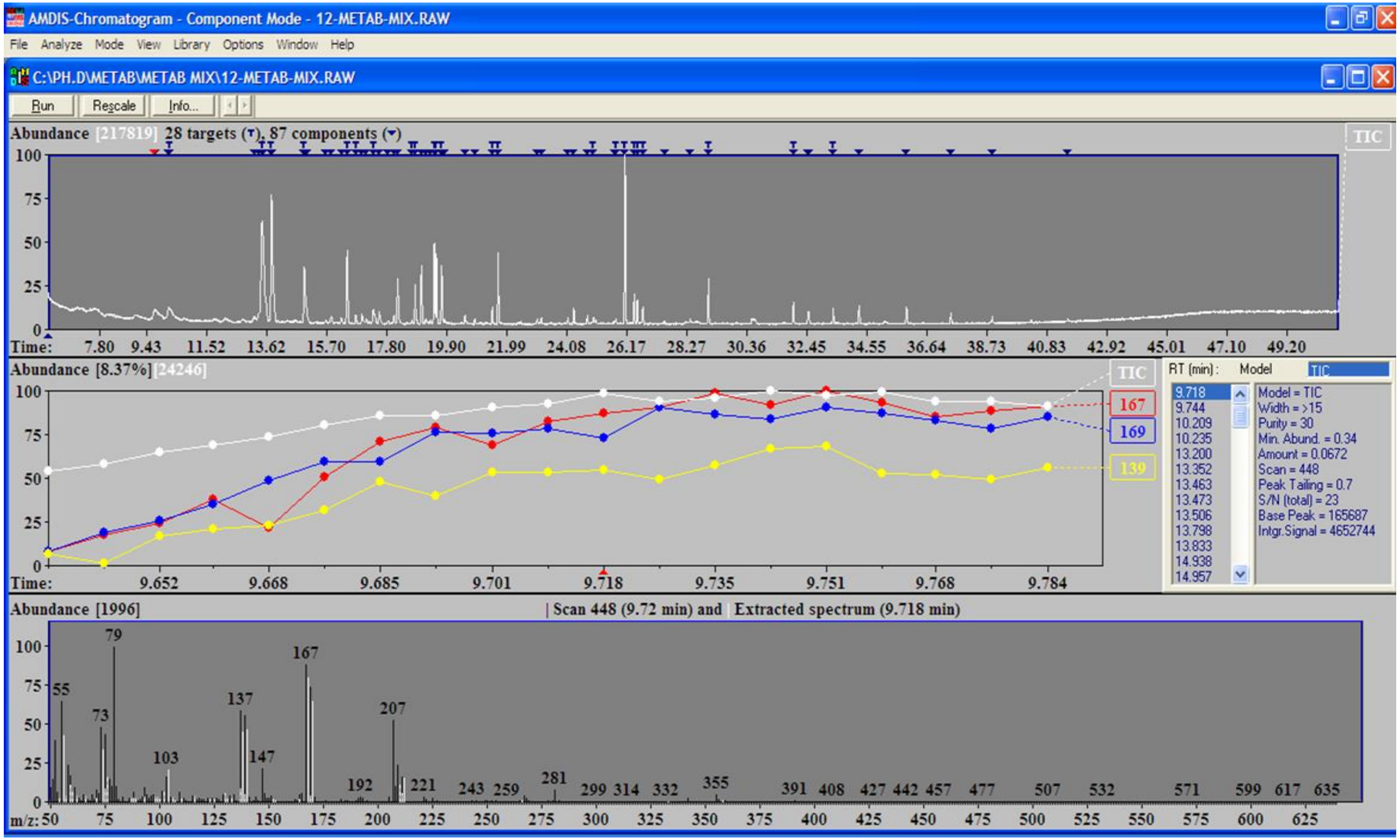
AMDIS extracts spectra for individual components in GC-MS data file and identifies compounds by matching these spectra against specialized libraries or the NIST 02 library. The GC-MS data files were subjected to analysis by AMDIS in simple mode. The result from AMDIS deconvolution software as displayed in figure 6.1; which contains confirm window and result window. The data used in identification can be obtained from the confirm window which contains the menus File, Analyse, Mode, View, Library, Options, and Windows as well as chromatographic and mass spectral displays. The result window has four sub-windows which involves chromatogram, component profile, component/target information list and lastly component/target mass spectra. The chromatogram window in the upper pane can displays the ion chromatograms for each ion as well as the total ion chromatogram.

In the middle pane, the component profile shows the total ion current and the largest of peaks over the region actually used in the deconvolution of the component, whereas the component/target information list window displays data in different formats depending upon selection of component or target. For component, window displays two list boxes, left hand side list box displays a list of retention times for each resolved component, whereas right hand list box displays information about the component, such as model, width, purity, minimum abundance, amount, scan, peak tailing and S/N for the component. For target, window displays four boxes. The top left list box contains list of retention times for each identified target. The top right list box displays list of the targets identified which are linked to the list of retention times. The bottom left list box displays data relating to the component, whereas bottom right list box displays target match factor data.

Finally in the bottom pane, component mass spectral display contains raw mass spectrum (scan) shown as black vertical lines and the deconvoluted spectrum at a given retention time shown as white lines.

To understand the influence of different parameters on the deconvolution accuracy, different parameter values were tested. While testing individual parameter, the other parameters were kept the same. Data deconvolution was performed with the specifications given: Database used : GOLM and NIST MS search software; Match factor: 60; Component width : 12; Resolution : High; Sensitivity: High; Shape requirement : Medium and Adjacent peak substraction : one.

Appendix 6.2 AMDIS chromatogram displaying deconvoluted data from 12 metabolite mixture


AMDIS-Chromatogram - Component Mode - 12-METAB-MIX.RAW
File
Analyze
Mode
View
Library
Options
Window
Help
C:\PH.D\METAB\METAB MIX\12-METAB-MIX.RAW
Run
Rescale
Info...
Abundance [217819] 28 targets (▾), 87 components (▾)
TIC
100
75
50
25
0
Time: 7.80 9.43 11.52 13.62 15.70 17.80 19.90 21.99 24.08 26.17 28.27 30.36 32.45 34.55 36.64 38.73 40.83 42.92 45.01 47.10 49.20
Abundance [8.37%] [24246]
TIC
167
169
139
Time: 9.652 9.668 9.685 9.701 9.718 9.735 9.751 9.768 9.784
RT (min):
9.718
9.744
10.209
10.235
13.200
13.352
13.463
13.473
13.506
13.798
13.833
14.938
14.957
Abundance [1996]
Scan 448 (9.72 min) and Extracted spectrum (9.718 min)
55 73 79 103 137 147 167 192 207 221 243 259 281 299 314 332 355 391 408 427 442 457 477 507 532 571 599 617 635
m/z: 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625

Model: TIC

Model = TIC
Width = >15
Purity = 30
Min. Abund. = 0.34
Amount = 0.0672
Scan = 448
Peak Tailing = 0.7
S/N (total) = 23
Base Peak = 165687
Intgr.Signal = 4652744

Appendix 6.3 List of 32 new metabolites identified in *C. reinhardtii*

ID	Metabolite name
1	1-Methyl-beta-D-mannopyranoside
2	2,4,6-Tri-tert-butylbenzenethiol
3	2-O-Glycerol-beta-D-galactopyranoside
4	4-Hydroxy-3-methoxyphenethylene glycol
5	Arabinose
6	Calystegine B2 B4
7	Cinnamic acid
8	Dodecamethylpentasiloxane
9	Eicosanol
10	Erythronic acid
11	Galactose
12	Glucopyranose
13	Gulose
14	Hexadecanoic acid
15	Hexadecanol
16	Hypoxanthine
17	Malonic acid
18	Methylmalonic acid
19	Monomethylphosphate
20	N-Acetylglycine
21	n-Docosane
22	n-Dodecane
23	n-Octacosane
24	n-Octadecane
25	n-Pentacosane
26	n-Pentadecane
27	n-Tetracosane
28	n-Triacontane
29	n-Triacosane
30	Resveratrol
31	Tetradecanoic acid
32	Xylitol

Appendix 6.4 List of 94 metabolites identified in condition A and condition D

1	2-pcarboxylic acid	1
2	Adipic acid	1
3	Aspartic acid	1
4	Citric acid	1
5	Glutamic acid	1
6	Malic acid	1
7	Pyruvic acid	1
8	Valeric acid	1
9	Indole-2-carboxylic acid	1
10	Lactic acid	1
11	Butanoic acid	1
12	Fumaric acid	1
13	Glyceric acid	1
14	Isoascorbic acid	1
15	Isobutanoic acid	1
16	Maleic acid	1
17	Propanoic acid	1
18	Threonic acid	1
19	Arabitol	2
20	Cellobiose	2
21	Erythritol	2
22	Galactitol	2
23	Galactose	2
24	Glucose	2
25	Glycerol	2
26	Ribitol	2
27	Threitol	2
28	Xylitol	2
29	Xylose	2
30	Trehalose	2
31	Galactopyranoside	2
32	Glycine	3
33	Homocysteine	3
34	Leucine	3
35	Lysine	3
36	Norleucine	3
37	Phenylalanine	3
38	Proline [+CO2]	3
39	Prolyl-glycine	3
40	Pyroglutamic acid	3
41	Serine	3
42	Threonine	3
43	Tryptophan	3
44	Isoleucine	3
45	Proline	3
46	Norvaline	3
47	Valine	3
48	Guanosine	4
49	Adenine	4
50	Caproic acid	5
51	Dodecanoic acid	5
52	Heptadecan-1-ol	5
53	Hexadecan-1-ol	5
54	Hexadecanoic acid	5
55	Hexadecenoic acid	5
56	Octadecadienoic acid	5
57	Octadecanoic acid	5
58	Octadecatrienoic acid	5
59	Octadecenoic acid	5
60	Pentadecan-1-ol	5
61	Tetradecanoic acid	5
62	Tridecan-1-ol	5
63	Heptadecanoic acid	5
64	Tetradecan-1-ol	5
65	Ethanolamine	6
66	Phenethylamine	6
67	Sphingosine	6
68	Triethanolamine	6
69	Tryptamine	6
70	Gluconic acid-6-phosphate	7
71	Glyceric acid-3-phosphate	7
72	Glycerol-2-phosphate	7
73	Glycerol-3-phosphate	7
74	Glycolic acid-2-phosphate	7
75	Inositol-2-phosphate	7
76	Mannose-6-phosphate	7
77	Phosphomycin	7
78	Phosphoric acid	7
79	myo-Inositol-1-phosphate	7
80	Xylulose-5-phosphate	7
81	Decane	8
82	Dodecane	8
83	Pentadecane	8
84	Tetradecane	8
85	Tridecane	8
86	Lumichrome	11
87	Menthol	11
88	Phytol	11
89	Putrescine	11
90	Urea	11
91	Quinazoline	11
92	Sarcosine	11
93	Iminodiacetic acid	11
94	Uracil	11

Chapter 7 Appendices

Appendix 7.1 List of identified metabolites in approach 1 & 3

Met ID	***Metabolites***	***Class***
1	2-Piperidinecarboxylic acid	1
2	3-Indoleacetic acid	1
3	Aspartic acid	1
4	Eicosanoic acid	1
5	Erythronic acid	1
6	Fumaric acid	1
7	Glutaric acid	1
8	Glyceric acid	1
9	Glyoxylic acid	1
10	Gulonic acid	1
11	Iminodiacetic acid	1
12	Indole-2-carboxylic acid	1
13	Isobutanoic acid	1
14	Lactic acid	1
15	Lyxonic acid	1
16	Malic acid	1
17	Nicotinic acid	1
18	Orotic acid	1
19	Propanoic acid	1
20	Pyruvic acid	1
21	Shikimic acid	1
22	Threonic acid	1
23	Allose	2
24	Arabitol	2
25	Cellobiose	2
26	Cellotriose	2
27	Erythritol	2
28	Fructose	2
29	Galactitol	2
30	Galactonic acid	2
31	Galactopyranoside	2
32	Galactose	2
33	Galacturonic acid	2
34	Glucose	2
35	Glucuronic acid	2
36	Glycerol	2
37	Inositol	2
38	Kestose	2
39	Lactose	2
40	Laminaribiose	2
41	Lyxose	2
42	Maltotriose	2
43	Mannitol	2
44	Mannose	2
45	Ribitol	2
46	Ribose	2
47	Sedoheptulose	2
48	Sorbitol	2
49	Sorbose	2
50	Sucrose	2
51	Threitol	2
52	Trehalose	2
53	Xylitol	2
54	Xylose	2
55	Aminomalonic acid	3
56	Asparagine	3
57	Cysteamine	3
58	Cysteine	3
59	Glutamic acid	3
60	Glutamine	3
61	Glycine	3
62	Homoserine	3
63	Isoleucine	3
64	Leucine	3
65	Lysine	3
66	Norvaline	3
67	Ornithine	3
68	Phenylalanine	3
69	Proline	3
70	Prolyl-glycine	3
71	Purine	3
72	Pyroglutamic acid	3
73	Serine	3
74	Threonine	3
75	Tryptophan	3
76	Tyrosine	3
77	Valine	3
78	Adenine	4
79	Adenosine	4
80	Guanosine	4
81	Uracil	4
82	Docosahexaenoic acid	5
83	Docosanol	5
84	Dodecanoic acid	5
85	Dodecanol	5
86	Eicosanol	5
87	Heptadecanoic acid	5
88	Heptadecanol	5
89	Hexadecanoic acid	5
90	Hexadecanol	5
91	Hexadecenoic-acid	5
92	Nonadecanoic acid	5
93	Octadecadienoic acid	5
94	Octadecanoic acid	5
95	Octadecanol	5
96	Octadecatrienoic acid	5
97	Octadecenoic acid	5
98	Pentadecanoic acid	5
99	Pentadecanol	5
100	Tetradecanoic acid	5
101	Tridecanol	5
102	Ethanolamine	6
103	Phenethylamine	6

104	Putrescine	6
105	Sphingosine	6
106	Triethanolamine	6
107	Tryptamine	6
108	Dihydroxyacetone phosphate	7
109	Ethanolaminephosphate	7
110	Gluconic acid-6-phosphate	7
111	Glycerol-2-phosphate	7
112	Glycerol-3-phosphate	7
113	Glycolic acid-2-phosphate	7
114	Inositol-2-phosphate	7
115	Mannose-6-phosphate	7
116	myo-Inositol-1-phosphate	7
117	Ribose-5-phosphate	7
118	Ribulose-5-phosphate	7
119	Xylulose-5-phosphate	7
120	Decane	8
121	Docosane	8
122	Dodecane	8
123	Eicosane	8
124	Heneicosane	8
125	Heptadecane	8
126	Hexadecanal	8
127	Nonadecane	8
128	Octacosane	8
129	Octadecane	8
130	Pentacosane	8
131	Pentadecane	8
132	Tetradecane	8
133	Tricosane	8
134	Tridecane	8
135	Heneicosanol	9
136	Menthol	9
137	Phytol	9
138	Flavone	10
139	Butylamine	11
140	Carbodiimide	11
141	Cembrene	11
142	Cyclohexene	11
143	Hydantoin	11
144	Indole-3-acetaldehyde enol	11
145	Lumichrome	11
146	Naphthalene	11
147	Piceatannol	11
148	Quinazoline	11
149	Thymine	11
150	Tocopherol	11
151	Urea	11
152	UK1	12
153	UK10	12
154	UK11	12
155	UK12	12
156	UK13	12
157	UK14	12
158	UK15	12
159	UK16	12
160	UK17	12
161	UK18	12
162	UK19	12
163	UK2	12
164	UK20	12
165	UK21	12
166	UK22	12
167	UK23	12
168	UK24	12
169	UK25	12
170	UK26	12
171	UK27	12
172	UK28	12
173	UK29	12
174	UK3	12
175	UK30	12
176	UK31	12
177	UK32	12
178	UK33	12
179	UK34	12
180	UK35	12
181	UK36	12
182	UK37	12
183	UK38	12
184	UK39	12
185	UK4	12
186	UK40	12
187	UK41	12
188	UK42	12
189	UK43	12
190	UK44	12
191	UK45	12
192	UK5	12
193	UK5	12
194	UK6	12
195	UK7	12
196	UK8	12
197	UK9	12

Appendix 7.2 List of identified metabolites in approach 2

Met ID	*Metabolites*	*Class*
1	2-Piperidinecarboxylic acid	1
2	Adipic acid	1
3	Aspartic acid	1
4	Erythronic acid	1
5	Fumaric acid	1
6	Glyoxylic acid	1
7	Gulonic acid	1
8	Iminodiacetic acid	1
9	Indole-2-carboxylic acid	1
10	Isobutanoic acid	1
11	Lactic acid	1
12	Lyxonic acid	1
13	Malic acid	1
14	Nicotinic acid	1
15	Shikimic acid	1
16	Threonic acid	1
17	Arabitol	2
18	Cellobiose	2
19	Erythritol	2
20	Erythrulose	2
21	Fructose	2
22	Galactitol	2
23	Galactopyranoside	2
24	Glucose	2
25	Glycerol	2
26	Inositol	2
27	Kestose	2
28	Lyxose	2
29	Mannitol	2
30	Mannose	2
31	Psicose	2
32	Ribitol	2
33	Ribose	2
34	Sorbitol	2
35	Sorbose	2
36	Sucrose	2
37	Threitol	2
38	Threose	2
39	Xylitol	2
40	Xylose	2
41	Alanine	3
42	Aminomalonic acid	3
43	Cysteamine	3
44	Cysteine	3
45	Glutamic acid	3
46	Glutamine	3
47	Glycine	3
48	Homoserine	3
49	Isoleucine	3
50	Leucine	3
51	Lysine	3
52	Phenylalanine	3
53	Proline	3
54	Prolyl-glycine	3
55	Pyroglutamic acid	3
56	Serine	3
57	Threonine	3
58	Tryptophan	3
59	Tyrosine	3
60	Valine	3
61	Adenosine	4
62	Guanosine	4
63	Uracil	4
64	Dodecanoic acid	5
65	Dodecanol	5
66	Eicosanol	5
67	Heptadecanoic acid	5
68	Heptadecanol	5
69	Hexadecanoic acid	5
70	Hexadecanol	5
71	Hexadecenoic-acid	5
72	Octadecadienoic acid	5
73	Octadecanoic acid	5
74	Octadecanol	5
75	Octadecatrienoic acid	5
76	Octadecenoic acid	5
77	Pentadecanoic acid	5
78	Pentadecanol	5
79	Tetradecanoic acid	5
80	Tridecanol	5
81	Ethanolamine	6
82	Phenethylamine	6
83	Putrescine	6
84	Sphingosine	6
85	Triethanolamine	6
86	Tryptamine	6
87	Erythrose-4-phosphate	7

88	Ethanolaminephosphate	7
89	Gluconic acid-6-phosphate	7
90	Glucose-6-phosphate	7
91	Glycerol-2-phosphate	7
92	Glycerol-3-phosphate	7
93	Glycolic acid-2-phosphate	7
94	Inositol-1-phosphate	7
95	Mannose-6-phosphate	7
96	myo-Inositol-1-phosphate	7
97	Xylulose-5-phosphate	7
98	Decane	8
99	Docosane	8
100	Eicosane	8
101	Heneicosane	8
102	Heptadecane	8
103	Hexadecanal	8
104	Nonadecane	8
105	Octadecane	8
106	Pentacosane	8
107	Pentadecane	8
108	Tetradecane	8
109	Tricosane	8
110	Tridecane	8
111	Ampelopsin	9
112	Heneicosanol	9
113	Menthol	9
114	Phytol	9
115	Flavone	10
116	Butylamine	11
117	Hydantoic acid	11
118	Lumichrome	11
119	Nicotinamide	11
120	Phenylpyruvic acid	11
121	Thiophene	11
122	Tocopherol	11
123	Urea	11
124	UK1	12
125	UK10	12
126	UK11	12
127	UK12	12
128	UK13	12
129	UK14	12
130	UK15	12
131	UK16	12
132	UK17	12
133	UK18	12
134	UK19	12
135	UK2	12
136	UK20	12
137	UK21	12
138	UK22	12
139	UK23	12
140	UK24	12
141	UK25	12
142	UK26	12
143	UK27	12
144	UK28	12
145	UK29	12
146	UK3	12
147	UK30	12
148	UK31	12
149	UK32	12
150	UK33	12
151	UK34	12
152	UK35	12
153	UK36	12
154	UK37	12
155	UK38	12
156	UK39	12
157	UK4	12
158	UK40	12
159	UK41	12
160	UK42	12
161	UK5	12
162	UK6	12
163	UK7	12
164	UK8	12
165	UK9	12

Chapter 8 Appendices

Appendix 8.1 List of 162 metabolites identified across all the applied extraction protocols

Mets	*Metabolites*	*Class*
1	1-Pyrroline-3-hydroxy-5-carboxylic-acid	1
2	2-Piperidinecarboxylic acid	1
3	Ascorbic acid	1
4	Aspartic acid	1
5	Citric acid	1
6	Erythronic acid	1
7	Fumaric acid	1
8	Glyoxylic acid	1
9	Iminodiacetic acid	1
10	Isobutanoic acid	1
11	Malic acid	1
12	Nicotinic acid	1
13	Pyruvic acid	1
14	Threonic acid	1
15	Valeric acid	1
16	6-deoxy-Mannopyranose	2
17	Allose	2
18	Altrose	2
19	Arabinose	2
20	Arabitol	2
21	beta-D-Allose	2
22	Cellobiose	2
23	Cellotriose	2
24	Erythritol	2
25	Galactopyranoside	2
26	Galactose	2
27	Gentiobiose	2
28	Glucose	2
28	Glucose	2
29	Glycerol	2
30	Glycinamide	2
31	Idose	2
32	Inositol	2
33	Laminaribiose	2
34	Lyxose	2
35	Mannose	2
35	Mannopyranoside	2
36	Melezitose	2
37	Ononitol	2
38	Ribitol	2
39	Ribose	2
40	Sucrose	2
41	Threitol	2
42	Trehalose	2
43	Xylitol	2
44	Xylose	2
45	Alanine	3
45	Alanine	3
46	Alanineamide	3
47	Aminomalonic acid	3
48	Cysteamine	3
49	Glutamic acid	3
50	Glycine	3
51	Homoserine	3
51	Homoserine	3
52	Leucine	3
52	Leucine	3
53	Lysine	3
53	Lysine	3
54	Methionine	3
55	Norleucine	3
56	Norvaline	3
57	Ornithine	3
58	Phenylalanine	3
59	Proline	3
59	Proline	3
60	Pyroglutamic acid	3
61	Sarcosine	3
62	Serine	3
63	Threonine	3
63	Threonine	3
64	Tryptophan	3
65	Tyrosine	3
66	Valine	3
67	Adenosine	4
68	Guanosine	4
69	Inosine	4
70	Uracil	4
71	Uridine	4
72	Caproic acid	5
73	Docosahexaenoic acid	5
74	Dodecanol	5
75	Eicosanol	5
76	Eicosatetraenoic acid	5
77	Heptadecanol	5
78	Hexadecanoic acid	5
79	Hexadecanol	5

80	Nonadecanoic acid	5
81	Nonanoic acid	5
82	Octadecadienoic acid	5
83	Octadecanoic acid	5
84	Octadecanol	5
85	Octadecatrienoic acid	5
86	Octadecenoic acid	5
87	Pentadecanoic acid	5
88	Pentadecanol	5
89	Tetradecanoic acid	5
90	Tetradecanol	5
91	Ethanolamine	6
92	Phenethylamine	6
93	Putrescine	6
94	Sphingosine	6
95	Tryptamine	6
96	Dihydroxyacetone phosphate	7
97	Erythrose-4-phosphate	7
98	Ethanolaminephosphate	7
99	Fructose-1-phosphate	7
99	Fructose-6-phosphate	7
100	Gluconic acid-6-phosphate	7
101	Glucose-6-phosphate	7
102	Glyceric acid-3-phosphate	7
102	Glyceric acid-2-phosphate	7
103	Glycerol-2-phosphate	7
103	Glycerol-3-phosphate	7
104	Glycerophosphoglycerol	7
105	Glycolic acid-2-phosphate	7
106	Inositol-2-phosphate	7
106	Inositol-1-phosphate	7
107	Mannose-6-phosphate	7
108	Phosphoenolpyruvic acid	7
109	Phosphoric acid	7
110	Ribose-5-phosphate	7
111	Ribulose-5-phosphate	7
112	Xylulose-5-phosphate	7
113	Decane	8
113	Docosane	8
114	Dodecane	8
115	Eicosane	8
116	Heneicosane	8
117	Heptadecane	8
118	Nonadecane	8

119	Octadecane	8
120	Pentacosane	8
121	Pentadecane	8
122	Tetradecane	8
123	Tricosane	8
124	Tridecane	8
125	Heneicosan-1-ol	9
126	Menthol	9
127	Phytol	9
128	Propane-1,2-diol	9
129	Hexadecanal	10
130	Cembrene	11
131	Ergocalciferol	11
132	Lumichrome	11
133	Piceatannol	11
134	Tocopherol	11
135	UK1	12
136	UK2	12
137	UK3	12
138	UK4	12
139	UK5	12
140	UK6	12
141	UK7	12
142	UK8	12
143	UK9	12
144	UK10	12
145	UK11	12
146	UK12	12
147	UK13	12
148	UK14	12
149	UK15	12
150	UK16	12
151	UK17	12
152	UK18	12
153	UK19	12
154	UK20	12
155	UK21	12
156	UK22	12
157	UK23	12
158	UK24	12
159	UK25	12
160	UK26	12
161	UK27	12
162	UK28	12