Characterising Cellulolytic and Biofuel Generating Bacteria using Proteomics for Future Implementation as Biofuel Producers

Mahendra Parasram Raut

(BSc, MSc)

Thesis submitted to the University of Sheffield,
Department of Chemical and Biological Engineering
for the degree of Doctor of Philosophy

December 2013

Dedication

I dedicate this thesis to my parents because whatever I am today is due to my parent's hardwork, prayers and love. Aai! I hope that this achievement will complete the dream that you had for me. I also dedicate this thesis to my brothers (Dharmendra, Jayandra), sisters (Sangeeta, Premlata, Vidya) and sister in-law (Shubhangi), Neice (Gloria), friends, teachers and all those who help me in every aspects of my life.

Acknowledgements

There are number of people without whom it would not have been possible to complete this thesis. I am greatly indebted to the following people. At the outset, I wish to express my deepest gratitude to the immense contribution of my parents and my PP in my life. My deepest appreciation goes to my supervisor Prof. Phillip C. Wright, whose encouragement, mentorship and inspiration enabled me to develop the knowledge about the subject and confidently carry out this research. I am greatly thankful to him for his trust in my scientific abilities and his constant encouragement which pushed me to achieve many things.

I wish to thank Nishikant Wase for introducing me to this wonderful Department of Chemical and Biological engineering and his moral support. I would like to express my deepest gratitude to, Esther Karunakaran, Josselin Noirel, Narciso Couto, Trong Khoa Pham and Caroline Evans for their enthusiastic guidance, support and inspiration from the beginning of my research. I would like pay my gratitude to Rahul Kapoore for his moral support and for the comfort he has provided me during my PhD. I also acknowledge Dr. Saw-Yen Ow and Dr. Malinda Salim for their guidance and support in the beginning of my research.

A massive thanks to all my friends in chemical and biological engineering department, Wen Qiu, Yimin Chen, Faith Robert, Stephen Jaffe, Joseph Longworth, Danying Wu, Sakaram Waghmare, Joy Mukherjee, Bharathi Sekar, Ana Pereira-Medrano, Jagroop Pandhal, Ioanna Dimitriou, Mark Wells, Gregory Fowler and Sekar Raju for the support, laughter and many memorable moments. I would also like to thank Prof. Catherine Biggs for her support, encouragement and allowing me to use of the surface characterisation equipment in the lab. I would also like to thank to Leonardo Gomez, Andrew Hunt, and Veronica Ongaro from The University of York for providing me interesting samples to analyse. I also thank all technical and administrative staff, Maria, Dave, Marine, Keith, Usman, Mark, Oz. You are all very helpful. I am also deeply indebted to all scientists and friends from NEERI, India whom I have worked with previously. They have always been kind and supportive, guiding me during my early days of research. It really means a lot to me. I also would like to thank, from the bottom of my heart, to Dawn and Gina for their friendship and support. There are too many of you to name but I love you all.

Finally I would like to extend my sincere thanks to the Ministry of Social Justice and Empowerment, Govt. of India for providing me financial support under the National Overseas Fellowship (Grant No. 11015/22/2008-SCD-V) to carry out this research. I also thank ChELSI and EPSRC (EP/E036252/1, EP/E053556/1) and SUNLIBB (SUNLIBB—Sustainable Liquid Biofuels from Biomass Biorefining, Project 251132, FP7-ENERGY-2009-BRAZIL) programme for their support in this work.

Declaration

This is a declaration to state that this thesis is an account of the author's work which was conducted at the University of Sheffield, UK. This work has not been submitted for any other degree of qualification.

Abstract

In natural ecosystems, microbes are mostly found in diverse and complex communities or consortia that can live symbiotically and fulfil most important global biogeochemical cycles. These processes are very difficult or impossible to achieve by a single bacterium. At the global level, scientists have come to know the innate capacity of natural microbial consortia and are starting to understand natural communities and to develop recombinant synthetic consortia for future biotechnology application. In order to overcome key challenges arising due to fossil fuel depletion and contribution to global warming, consolidated bioprocessing (CBP) is thought to be a low cost processing scheme for lignocellulosic biofuel production. Consortia of cellulolytic and biofuel producing microorganisms could be an attractive alternative to single organism approaches. However, proper understanding of the biology of native microbes and their implementation in the development of consortia needs rigorous research study at the system-wide level.

Given the immense potential in the conversion of lignocellulosic biomass to biofuels, anaerobic microorganisms are of great interest to researchers. Therefore, this research is focussed on two different anaerobic bacteria: *Fibrobacter succinogenes* S85, which is an efficient cellulose degrader, but cannot produce biofuels; and *Clostridium acetobutylicum* ATCC 824, a promising solvents (acetone, ethanol, butanol) producing bacterium that cannot degrade cellulose. The study of these microbes at the systems level will help to understand the biological complexity of these microbes and provide valuable information for future CBP development. Based on capabilities of these microbes, two individual aspects have been proposed and investigated.

In this thesis, an investigation of the surface colloidal properties and surface-membrane associated proteins of *F. succinogenes* involved in cellulose degradation by biotin labelling method using two substrate conditions cellulose and glucose (control) is carried out. Further analysis of the *F. succinogenes* membrane using high throughput quantitative proteomics using isobaric tag for relative and absolute quantitation (iTRAQ) is presented. This iTRAQ study reveals many novel proteins associated with cellulose degradation, adding valuable information on the mechanism of cellulose degradation in this bacterium.

In this thesis, a preliminary technical study comparing two digestion systems (in-gel and in-solution) of soluble proteins from *C. acetobutylicum* and two peptide separation techniques (SCX and HILIC) is presented. Results reveal that in-gel digestion with HILIC separation is superior to

SCX for soluble proteomics from this system. Results are further used in a quantitative proteomics study in the presence of cellobiose and lignin and elucidate the effect of lignin on solvent production and various metabolic processes.

This thesis demonstrates that both *F. succinogenes* and *C. acetobutylicum* can potentially be used in co-culture to utilise cellulose and pre-treated lignocellulosic waste for bio-augmented bioalcohol production in consolidated bioprocess (CBP) development framework.

List of publications

Journal articles

R1 Mahendra P. Raut, Esther Karunakaran, Joy Mukherjee, Catherine A. Biggs and Phillip C. Wright, Influence of substrates on the surface characteristics and membrane proteome of *Fibrobacter succinogenes* S85 (submitted to Journal of Bacteriology)

R2 Mahendra P. Raut, Narciso Couto, Trong Khoa Pham, Caroline Evans, Josselin Noirel and Phillip C. Wright, Influence of Lignin on *Clostridium acetobutylicum* ATCC 824 Proteome: a Quantitative Proteomics Analysis (Manuscript under preparation)

R3 Mahendra P. Raut, Narciso Couto, Esther Karunakaran, Caroline Evans, Josselin Noirel and Phillip C. Wright, Quantitative membrane proteome of *Fibrobacter succinogenes* S85: A comparison between enrichment methods and biological interpretation (Manuscript under preparation)

R4 Mahendra P. Raut, Trong Khoa Pham, Ioanna Dimitriou and Phillip C. Wright Alcoholic fermentation of carbon sources in biomass obtained from chemical and biological hydrolysis by *Clostridium acetobutylicum* ATCC 824 (Manuscript under preparation)

R5 An experimental comparison between tryptic digestion approaches and HPLC peptide separation techniques for proteome LC-MS/MS analysis (Technical note under preparation)

Conference presentations

(P1) Mahendra P. Raut and Phillip C. Wright, (2011) Can lignin be degraded by *C. acetobutylicum* ATCC 824?. SUNLIBB/CeProBIO meeting, Magdalene Collage, Cambridge, UK., 21stto 23rd September 2011

(P2) Mahendra P. Raut and Phillip C. Wright, (2012) Lignin influencing biofuel production in *C. acetobutylicum* ATCC 824.SUNLIBB/CeProBIO meeting, INRA Centre de Recherche de Versailles, France, 26th to 28thSeptember 2012

(P3) Mahendra P. Raut and Phillip C. Wright, (2013) Using innate microbial capacity for lignocellulosic biofuel generation, SUNLIBB/CeProBIO meeting, Wageningen, The Netherlands, 25th to 27thSeptember 2013

(P4) Mahendra P. Raut and Phillip C. Wright, (2013) Can lignin be degraded by *Clostridium acetobutylicum* ATCC 824?-A piece of the puzzle, SUNLIBB/CeProBIO meeting, Wageningen, The Netherlands, 25th to 27thSeptember 2013

Lists of Abbreviations

µg	Microgram
µL	Microlitre
µm	Micrometre
1-D	One dimensional
2DE	Two dimensional electrophoresis
AB	Ammonium bicarbonate
ABE	Acetone butanol ethanol
CAN	Acetonitrile
ATP	Adenosine triphosphate
ATR	Attenuated total reflectance
BP	British Petroleum
BSA	Bovine serum albumin
CAZy	Carbohydrate active enzymes
CBMs	Cellulose binding modules
CBP	Consolidated bioprocessing
CBPs	Cellulose binding proteins
CID	Collision induced dissociation
Da	Dalton
DNA	Deoxyribo nucleic acid
EG	Endoglucanases
EI	Electron impact
EPM	Electrophoretic mobility analysis
ESI	Electrospray ionization
FA	Formic acid
FDR	False discovery rate
FTICR	Fourier transform ion cyclotron resonance
FTIR	Fourier transform infra-red
GC	Gas Chromatography
GC-FID	Gas chromatography-Flame ionization detector
GC-MS	Gas chromatography-mass spectrometry
GC-TCD	Gas chromatography- Thermal conductivity detector
GH	Glycoside Hydrolase
GHnc	Glycoside Hydrolase not yet assigned
GT	Glycosyltransferase
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High performance liquid chromatography
ICAT	Isotope-coded affinity tags
IEA	International Energy Agency
IEF	Isoelectric focussing
iTRAQ	Isobaric tags for relative and absolute quantitation
L	Litre
LC-(2D)	Two dimensional liquid chromatography
LC-(MD)	Multi-dimensional Liquid chrmatography
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	Liquid chromatography tandem mass spectrometry
m/z	mass/charge
MALDI	matrix assisted laser desorption/ionization
MATH	Microbial adhesion to hydrocarbons

MeCATs	Metal-coded affinity tags
mL	mililitre
Mm	Milimetre
MMTS	methyl methanethiosulfonate
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular weight
NHS	N-hydroxysuccinimidyl
NP-40	nonyl phenoxypolyethoxylethanol-40
OD	Optical density
OM	Outer membrane
ORF	Open reading frames
PAGE	Polyacrylamide gel electrophoresis
PBS	phosphate buffer solution
PCA	Principal component analysis
pI	isoelectric point
PL	Polysaccharide lyase
QIT	Quadrupole ionTrap
QqTOF	quadrupole-quadrupole-time-of-flight
q-TOF	Quadrupole Time-of-flight
RF	Radio frequency
RNA	Ribonucleic acid
RP	Reverse Phase
RPLC	reverse phase liquid chromatography
RP-LC (2-D)	Two dimensional reverse phase liquid chromatography
SCX	Stong cation exchange
SDS	Sodium dodesyl sulphate
SILAC	stable isotope labeling by/with amino acids in cell culture
TCA	tricarboxylic acid cycle
TCA	tricloroacetic acid
TCEP	tris 2-carboxyethyl phosphine hydrochloride
TEAB	Triethyl ammonium bicarbonate
TEM	Transmission electron microscopy
TFA	Trifluoroacetic acid
TMT	Tandem mass tags
TNT	2,4,6 trinitrotoluene
UV	Ultraviolet
ZIC	zwitterionic

Table contents

List of Figures

List of Tables

Appendices

Chapter 1

Introduction

1.1 Background

In order to secure a reliable and constant supply of energy and mitigate environment threatening problems due to existing fossil fuels, generation of alternative fuels from renewable resources is mandatory. Lignocellulosic biomass has a great potential as a prime candidate for future biofuel generation (so called second generation biofuel) since it is economical and environmentally friendly due to its wide availability and reproducibility. Furthermore, unlike first generation biofuels, it does not directly compete with food stocks. However, lignocellulosic feed stocks still require resources such as land, water and fertilizers to grow them. Biological conversion of lignocellulosic biomass into fuels is a promising alternative process. Consolidated bioprocessing (CBP) is one of the most cost effective approaches for second biofuel generation compared to first biofuel generation [1], it involves saccharification of complex polymers and their subsequent conversion into value-added products in a single step and can be achieved either by native or recombinant microbes. However, introduction of both capabilities in a single microbe still remains a challenge. A detailed study of native microbes capable of producing biofuel from lignocellulosic biomass is required for commercial implementation of future lignocellulosic biofuel generation.

Anaerobes have great ancient history in bioenergy generation [2-4] and are of great interest to researchers, since the most efficient lignocellulose degraders and biofuel producers are anaerobes [5].

Therefore, the outlook of this thesis is to mainly focus on the study of the metabolic systems of two natural microorganisms: *Fibrobacter succinogenes* S85, a efficient cellulose degrader and *Clostridium acetobutylicum* ATCC 824, a

promising solvent producer. The PhD study explores both microbes at the omic level and adds novel information for future biofuel generation.

Considering that proteins are main functional component of all biological systems and biochemical processes taking place within cells, a global quantitative proteomics study was undertaken to identify metabolic pathways involved in cellulose degradation and biofuel generation.

The research described in this thesis provides vital information regarding *F. succinogenes* and *C. acetobutylicum* and their individual capabilities for cellulose degradation and biofuel generation for use in future lignocellulosic biofuel generation studies.

1.2 Hypothesis and Objectives

The overall objective is to understand the mechanism of cellulose degradation in Fibrobacter succinogenes S85 *and bio-alcohol production in Clostridium acetobutylicum* ATCC 824 *at the proteomic level that can help in further implementation and process design for future consolidated bioprocessing development (CBP)*

The key driving force behind undertaking this PhD is the need to find alternatives for existing fuels, a solution to tackle global energy and environmental issues. Native microbes with commercial value for biofuel generation need to be studied at the proteomic level to re-discover their functional capabilities for further enhanced implementation in biofuel generation.

Within this thesis, the author used proteomics tools to shed light on poorly studied cellulose degradation mechanism in *F. succinogenes* and also lignin-induced metabolic changes in *C. acetobutylicum.*

Chapter 2 provides comprehensive review of the literature relating to anaerobic microbiology in lignocellulosic biofuel generation followed by the description of the bacteria examined in this research. Following from this, the thesis is divided

into two major parts and broken down into eight subsections (summarised in Figure 1-1).

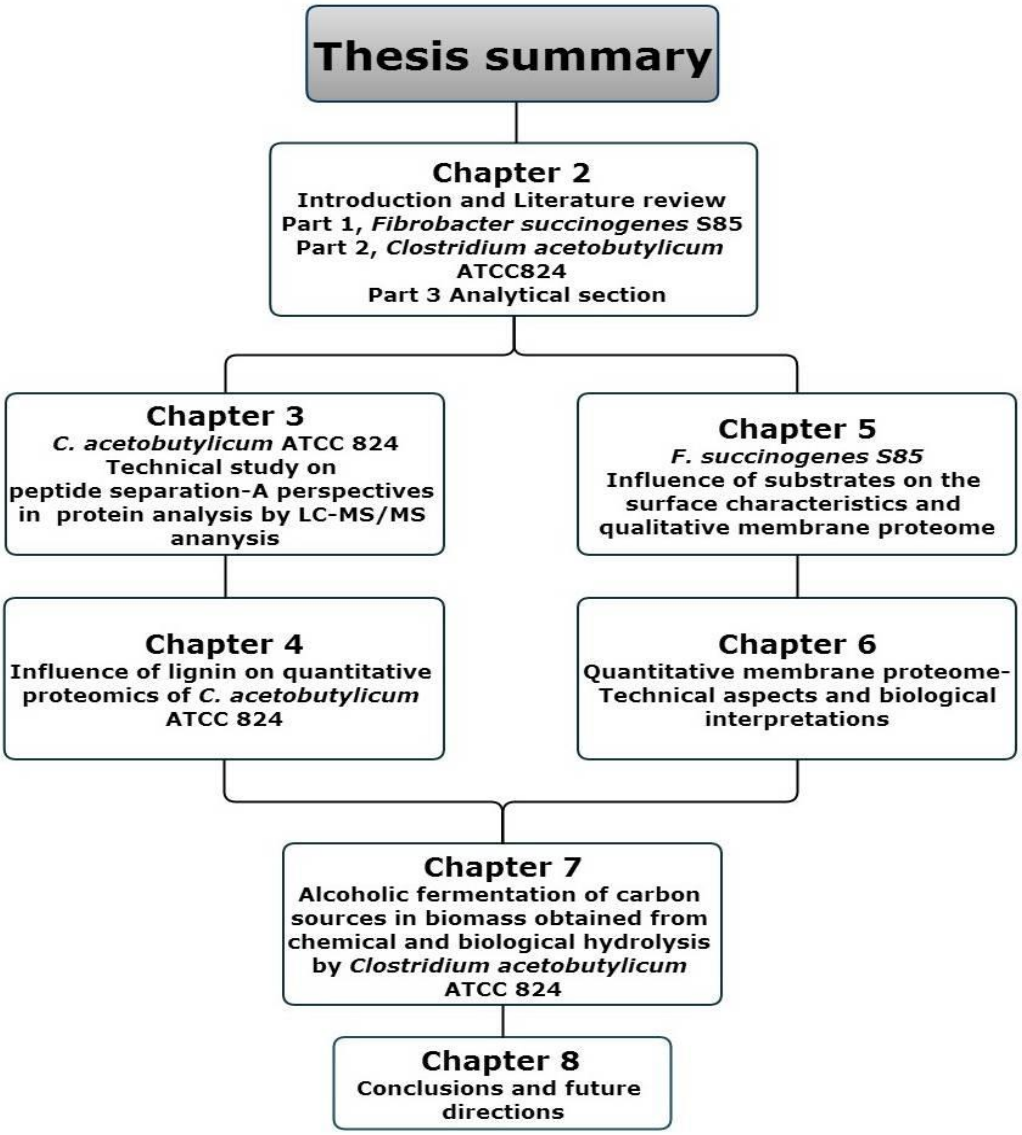

Figure1-1 Flow diagram of the tasks undertaken in this thesis (relevant chapters are titled)

Chapter 3 seeks to implement separation techniques used for fractionation of peptides in order to improve the coverage of the proteome. Fractionation techniques affect peptide recovery from trypsin digestion and subsequent protein identification and characterisation. The proteins were obtained from *C. acetobutylicum* at the exponential and stationary phase to obtain a complex protein mixture. The chapter compares HILIC and SCX based separation

workflows. The information obtained from this study was further implemented in the following Chapters 4, 5, 6 study.

Chapter 4 involves a quantitative proteomics analysis of *C. acetobutylicum* in the presence of lignin. Considering the future utilisation of this bacterium for biomass to biofuel generation and to check its response to lignin, it was thought to check the influence of lignin on metabolic process and on solvent production.

Chapter 5 examines the surface-membrane of *F. succinogenes* in the presence of cellulose using colloidal surface characterisation techniques and membrane protein biotin labelling. The idea was to understand the mechanism of cellulose degradation in this bacterium and find out the role of surface-membrane proteins in cellulose degradation. The proteomic technique used in this study is a membrane biotinylation workflow which has not previously been used for this bacterium in membrane proteomics.

Chapter 6 extends work carried out in Chapter 5 by using a quantitative membrane proteomic analysis (iTRAQ based quantitation). The study compares two techniques of membrane protein enrichment and separation: 1) membrane isolation by ultracentrifugation followed by quantitative analysis and 2) biotin-streptavidin protein enrichment (biotinylation) followed by quantitative analysis. In particular, the work here demonstrates a comparison between two techniques in terms of membrane proteins coverage that will further help to understand the mechanism of cellulose degradation and other related membrane activities. This is the first study so far where quantitative proteomics has been performed on the membrane of this bacterium. Results are compared against previous literature and results obtained from Chapter 5.

Chapter 7 examines the utilisation of sugar components for biofuel production obtained from biological and chemical hydrolysis of different biomass sources. *C.*

acetobutylicum possesses diverse metabolic pathways for sugar utilisation. Therefore, half a section of this chapter describes the use of miscanthus hydrolysate to produce biofuel by *C. acetobutylicum*. Another half section of this chapter seeks to use both bacteria as is a co-culture for bio-augmented biofuel production using cellulose and pre-treated lignocellulosic biomass. The products obtained from both studies are compared.

Chapter 8 summarises the findings and overall conclusions. Further, this chapter also directs future work for using these microbes to develop a consolidated bioprocessing (CBP) route.

Chapter 2

Literature Review

2.1 Background

There has recently been widespread insecurity about the supply and cost of transportation fuels, as well as fossil fuel related environmental concerns [6]. It is forcing to generate alternative biofuels from renewable resources which should compatible with existing fossil fuels. Worldwide, governments are promoting use of renewable energy sources, particularly utilisation of biomass feedstocks [7]. According to the International Energy Agency (IEA), global energy consumption will increase by 56 % between 2010 and 2040, half of which will be attributed to developing countries particularly China and India [8]. Biomass can provide a substantial contribution to supplying future energy demand in a sustainable way [9].

Biofuels can be made by processing food crops and other plants and animal products or wastes [10]. Biofuels directly derived from the easily fermentable sugars and vegetable oils, are referred to as “first generation biofuels”, have always been a major part of debate as they directly compete with food supply (the “food vs fuel” conundrum). On the other hand, a fuels derived from lignocellulosic biomass, referred to as “second generation biofuels”, can be a promising alternative in future biofuel generation since they can be derived from cell walls of the plant biomass such as lignocelluloses [11]. Lignocellulosic biomass possesses advantages over other alternative candidates because of its high energy content, significant compatibility, renewability, geographical availability, and eco friendliness [12, 13]. It is thus considered as a prime source for future biofuel generation.

However, despite these advantages, the development of lignocellulose mediated bio-fuels is hampered by the lack of genetic data in high-yielding strains and

difficulties in optimizing their metabolic pathways in recombinant microbes [14]. According to a IEA report, plant biomass can be used directly as fuel or can be transformed into useful by-products such as fuel, electricity and heat [15]. However, in order to utilise biomass in a sustainable manner, one needs to develop strategies to extract energy from plant biomass [15].

Recently, the term third generation biofuel is been introduced to the mainstream and it refers to biofuel derived from algae. Despite of its tremendous potential as future energy resource, the technical and economical issues need to be overcome before industrial level production [16].

2.2 Lignocellulose: a complex of polymers

Cellulosic biomass is sometimes referred to as lignocellulosic biomass [12]. Lignocellulose is a naturally occurring complex of polymers produced in large quantity by plants [17]. A plant captures solar energy and stores in the form of this highly specialized complex of polymers in plant cell walls (refer to Figure 2-1). It is typically composed of cellulose, hemicellulose, and lignin polymers [18]. The plant cell wall usually contains 35 to 50 % cellulose, 20 to 35 % hemicellulose and 10 to 25 % lignin of dry weight of cell wall [19]. However, the composition of lignocellulosic biomass varies from species to species, but the ratio of lignin to cellulose and hemicelluloses rarely deviates from the ratiometric stoichiometry of 1:4 [20]. Other than lignocellulose, pectin is another complex plant polymer found in primary cell walls of dicotyledonous plants. It can also be found in small amount in secondary cell walls of dicots and both type of cell walls of monocots [21].

Cellulose is a major constituent of plant cell walls, made up of long chain of glucose molecules linked with β-(1→4) glycosidic bonds [22] that builds linear polymeric chain of about 800 to 12000 glucose molecules [15]. Cellulose chains

that are packed together by hydrogen bonding thus form highly insoluble structures (crystalline) called microfibrils which are further embedded with amorphous regions (non-crystaline) [23].

Hemicellulose, the second major constituent of lignocellulosic biomass is comprised of heterogeneous polysaccharides [24]. It is composed of three major hemicellulose polymers; 1) xylan, composed of β-1-4 linked D-xylose units that can be display side groups such as D-galactose, L-arabinose, glucuronic acid, acetyl, feruloyl, and *p*-coumaroyl units, 2) galacto (gluco) mannans, a backbone of β-1-1 linked of D-mannose and D-glucose with side chain of D-galactose and 3) xyloglucans that consist of β-1-4 linked of D-glucose backbone substituted with D-xylose [15, 23].

In some of plant species, pectins, a structural heteropolysaccharide are also found especially in primary cell walls and non woody terrestrial plants. They consist of D-galacturonic acid residues (smooth region) and D-galacturonic acid residues interrupted with α-1,2 linked L-rhamnose residues [15].

These constituents are then further cross linked with phenolic lignin polymers; thus creating a complex network of bonds that provide structural strength to the plant cell wall and making it resistant to mechanical degradation and to chemical hydrolysis [22, 25, 26]. Lignin is a three dimensional polymer of phenyl-units (namely, *p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol) that provides compressive strength and the cell wall stiffness [24, 27] in order to resist microbial attack. However, some microorganisms happen to be able to efficiently degrade these polymers in order to survive and they have developed different strategies for microbial degradation [28]. Figure 2-1 represents schematic of the three major polymers in lignocelluloses.

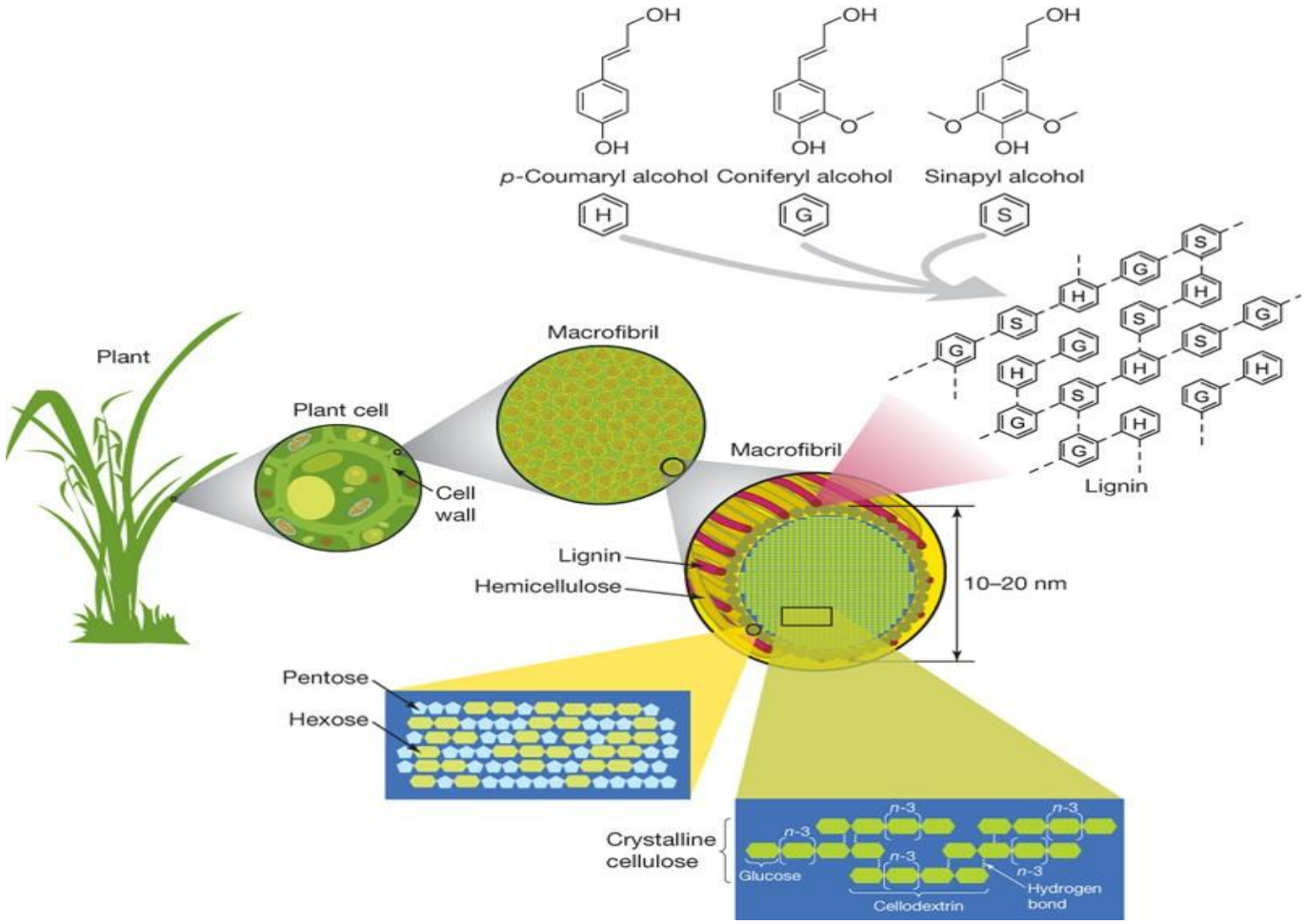

Figure 2-1 Lignocellulose structure consists of cellulose, hemicelluloses and lignin. Reproduced from [12]

2.3 Enzyme involved in lignocelluloses degradation

Cellulose polymers get hydrolyzed through the combined action of several enzymes belonging to the class of glycosyl hydrolases, which are expressed by a variety microorganisms [29, 30]. Three major classes of hydrolytic enzymes involved in cellulose degradation are 1) endo-1,4 β-glucanases (EC 3.2.1.4), 2) exo-1,4-cellobiohydrolases (EC 3.2.1.91), and 3) 1-glucosidases (EC 3.2.1.21) [31]. Endoglucanases break the cellulose randomly into oligosaccharides of different lengths which act as substrates for exoglucanases. Exoglucanases then cuts oligosaccharides from reducing and non-reducing ends resulting into cellobiose (disaccharide) as the end product which is in turn solubilised to glucose monomers by β-glucosidases [5]. Hemicellulose mainly consists of arabinoxylan (81 % w/w) in wheat straw and enzymes such as endoxylanases and β-xylosidases are involved in its hydrolysis [32].

However, lignin has a highly complex and relatively diverse structures, which makes it comparatively resistant to microbial degradation [33]. Enzymes are specific to breaking down the specific bonds within the lignin complex. The effectiveness of enzymes also varies from enzyme to enzyme [34]. The following enzymes have been characterized: peroxidases and phenol oxidases in aerobic systems [35] and phenylphosphate synthases [36] and phenylphosphate carboxylases in anaerobic systems [37].

Similarly, the insolubility and the heterogeneity of cellulose makes it impervious to enzymatic degradation resulting into its accumulation in the terrestrial environments. Therefore, the biodegradation of cellulose to fermentable sugar

requires the development of efficient and cost effective technology [38]. Genomics insights and understanding can bring much technology about.

In this PhD study, the idea is to develop native microbial strategies for consolidated bioprocessing, two anaerobic species *Fibrobacter succinogenes* S85 (cellulose degrader) and *Clostridium acetobutylicum* ATCC 824 (solvent producer) are selected and studied. The structure of this chapter is as follows; Part 1 presents a literature review of cellulose degradation systems in *F. succinogenes*, Part 2 discusses biofuel generating microbes, *C. acetobutylicum,* a solventogenic bacteria and bioprocess development for simultaneous saccharification and fermentation (CBP) and Part 3 discuses experimental methods used in this PhD project.

Part 1

2.4 Lignocellulose degradation

2.4 .1 Multiple microbial strategies for lignocellulosic biomass degradation

Due to the recalcitrant nature of the lignocellulosic biopolymers, pre-treatment is a requirement for either enzymatic hydrolysis or direct (such as chemical) conversion into fermentable sugars. The various pre-treatment technologies are under development and are being tested at an industrial scale such as stream explosion, acid/ alkaline, organosolv pre-treatments [39]. However, as the process is rate limited by saccharification, it is necessary to optimize the process efficiency and high throughput for biofuel generation [12]. The major concern in pre-treatment technologies development is the energy balance [40], since most of these techniques rely on heating the materials at high temperature (100-

200^{o}C). This makes them energy intensive. On the other hand, enzymatic hydrolysis is a relatively new concept and can be achieved by using pure enzymes (discussed in section 2.3) [39] or using whole microorganisms itself as a source of enzymes [41]. Figure 2-2 represents key steps in biomass to biofuel generation.

However, as the key enzymatic interactions in this degradation process remain elusive, researchers are still in the process of determining the optimal platform organisms. The candidates include model bacteria such as *Clostridium.thermocellum* or the fungus *Trichoderma reesei* [42].

Some microbial communities comprise several key groups of anaerobes capable of degrading lignocellulosic biomass. Recent studies have suggested that *Clostridium* and *Fibrobacter* species are the most predominant lignocellulose degrading strains [43-45]. *Clostridium* flourishes in anoxic environments such as landfills and fresh water sediments, while *Fibrobacter*s are most active in the rumen ecosystem [46-48].

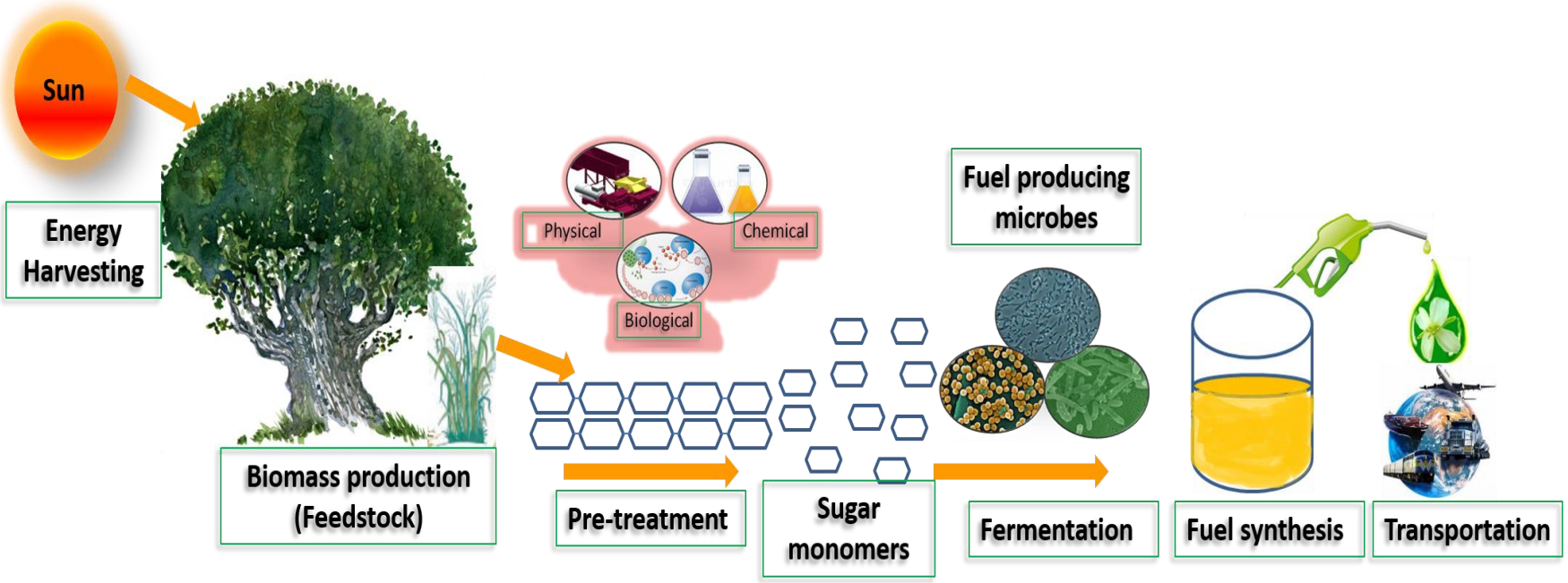

Figure 2-2 Biological route of energy transmission from sun to fuels (Modified from: [12, 49]).

2.4.2 Anaerobes in lignocelluloses degradation

During the course of evolution, a variety of microorganisms have developed cellulolytic capacity and exist in virtually every niche and clime [18, 46, 50, 51]. Saprophytism is one of the most commonly adopted lifestyle of microbes, whereby they extract energy from the dead or decaying organic matter mainly derived from the plant biomass [15]. In this context, organisation of cellulolytic enzymes varies among the microorganisms, ranging from free and soluble enzyme systems to high molecular weight protein complexes present on the cell surface (cellulosome) [5, 30, 52].

The innate capacity to utilise lignocelluloses is mainly attributed by fungi and bacteria. Cellulolytic bacteria found in different genera such as *Clostridium, Ruminococcus, Caldicellulosiruptor, Butyrivibrio, Acetivibrio, Cellulomonas, Erwinia, Thermobifida, Fibrobacter, Cytophaga and Sporocytophaga* [15]. Fungal species in cellulose degradation are entirely distributed within the kingdom from protist-like *Chytridomycetes* to the advanced *Basidiomycetes* [15].

Aerobic fungi, especially *Tricoderma reesei* and aerobic bacteria produce numerous individual extracellular enzymes in high concentration and these act in concert [53]. Therefore, most of the aerobes use free enzyme systems [54], while anaerobes use large cellulase complex systems (cellulosome), such as *C. thermocellum* [55]. Other systems such as brown rot fungi (basidiomycetes) appear to use less well studied oxidative mechanisms [56].

Two cellulolytic bacteria, *Cytophaga hutchinsonii* and the rumen anaerobic bacterium *Fibrobacter succinogenes* do not appear to use either of these two mechanisms, but are the best specialists in cellulose degradation [57]. This

suggests that they have to use a novel mechanism for cellulose degradation that can help us develop an efficient biofuel bioprocess.

2.4.3 Rumen microbiome and plant biomass degradation

The herbivores belonging to the order Artiodactyla such as cattle, deer, sheep and goat, possess a unique digestion system [58]. These organisms have specialised compartments of the stomach for lignocellulosic biomass digestion. The digestive system includes; ingestion (eating), chewing (mastication), swallowing (deglutition), absorption of nutrients and elimination of solid wastes (defecation). The first two compartments of stomach (reticulum and rumen) are commonly called the rumen. In the rumen, there are microbes which symbiotically live and digest the plant materials and make nutrients available for the ruminants, which are themselves unable to digests lignocelluloses [58].

The rumen ecosystem of ruminant animals is one of the most interesting, old and highly evolved cellulose digesting systems in nature [59, 60]. It is by far, the world's largest commercial fermenter, with a net volume of about $2x10^{11}$L since most of the diet of ruminants depends on cellulosic biomass [61]. Researchers believe that, microbiomes in the gastrointestinal tract of herbivores are well specialised for rapid polymers degradation and fermentation [62-64]. The study of the gut microbiomes started in the second half of the 20^{th} century, and the field attracted interest in the late 1980s and 1990s due to emerging biotechnological tools [62]. The rumen microbial ecosystem comprises 30 predominant bacterial species, 40 protozoa species, and 5 species of fungi. Bacterial species are considered the most important component in feed degradation (Figure 2-3) [65-67].

The study of efficient mechanisms of lignocellulose degradation in the rumen has been the subject of extensive research in the field of energy, but there is still a lack of knowledge on how the major fibrolytic bacteria and fungi degrades lignocellulosic biomass and interact in the rumen [68]. Therefore, it is necessary to study the symbiotic microbial digestion system in the rumen biochemically, genetically and proteomically [69].

Three main rumen bacteria including *F. succinogenes, Ruminococcus albus, and Ruminococcus flavefaciens* have been extensively studied, and they play a major role in the biological degradation of dietary fibre [70, 71]. These strains have been found to efficiently degrade cellulose at rates similar to that in the rumen itself ($0.1 gh^{-1}$) [61]. Notably, these rates comparatively exceed most cellulolytic microbes, except for a few thermophilic bacteria which exhibits even higher conversion rates (e.g. 0.16 gh^{-1} for *C. thermocellum*) [5]. The lists of major polysaccharide degrading bacteria with the phylum they belong to are shown in Figure 2-3. Due to the rapid rates of microbial utilisation of lignocellulose by these microbes, there is considerable interest in the utilization of these microbial enzyme systems using biotechnological applications for renewable bioenergy generation [5]. As a result, we wish to study the most efficient cellulose degrading strains particularly *F. succinogenes* at proteomic level for in-depth knowledge of their biology. Proteomics reflects functional status of the cell in response to environmental stimuli [72].

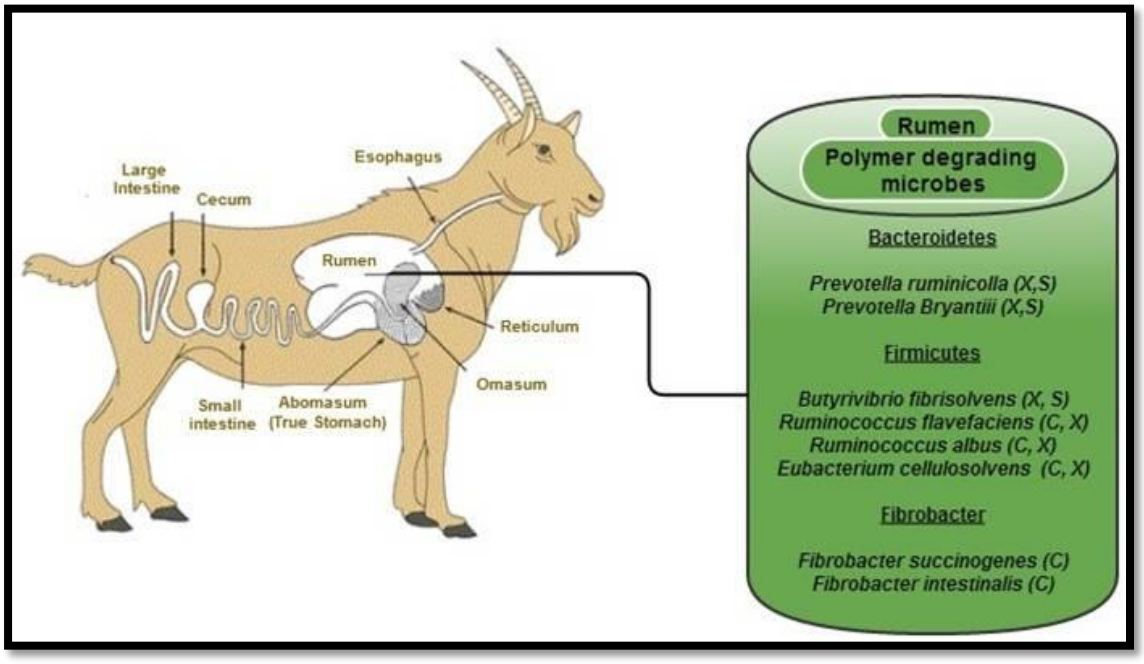

Figure 2-3 The principal polymer degrading bacteria in ruminants (modified from: [64])

2.4.4 Fibrobacter succinogenes S85, a specialised cellulose degrader

F. succinogenes is a rod shaped gram negative, obligatory anaerobic bacterium. It is believed to be one of the major plant cell wall degrading bacteria present in the rumen [73-75], and most specialised in plant cell wall digestion [76, 77]. However, recent studies have also proved that strains belonging to these species can also be found in other parts of the digestive system [78], termite guts [79], in anaerobic landfills [80] and in oceans, flax retting ponds and fresh water lakes [80-82]. Due to its diversity of the habitats, researchers believe that it may contain a variety of novel cellulases and hemicellulases that can significantly contribute to future bio-fuel generation.

Recent 16S RNA based taxonomy has revealed that *F. succinogene*sis not related closely to other Eubacteria [83]. Thus, it has been classified into a separate phylum branch, Fibrobacter [83]. There are two primary species, *F. succinogenes S85* and *F. intestinalis* NR9 which showed a 93 % similarity and 20 % homology between each other based on 16S RNA based analysis [83-87]. The genome of *F. succinogenes* was also recently completed and the full annotated genome sequence [43] and their predicted protein coding open-reading-frames (ORF) is available at: http://www.tiger.org/tdb/rumenomics/genomes.shtml.

F. succinogenes can be classified into four phylogenetic groups (groups 1-4) on the basis of sequenced data [88]. However, the contribution of each group to plant fibre digestion is still poorly understood. A recent phylogenic study has suggested that there may be an ecological predominance of the strains of group 1 given the higher diversity in their fibrolytic capabilities (such as the ability to degrade cellular and extracellular-based cellulose) [89]. This group contain *F. succinogenes*. On this basis, *F. succinogenes* has been considered to be one of the better model organisms to understand *in vitro* cell wall degradation [87, 90].

Compared to other fibrolytic species, such as *R. flavefaciens* and *R. albus*, *F. succinogenes* digests fibers faster and to a greater extent. In fact, this species even digests crystalline cellulose more actively than other ruminal species [91, 92]. The study of the sugar metabolism of *F. succinogenes* S85 [93] suggested that this bacterium only uses glucose and cellobiose as an energy source using glycolysis metabolic pathway, i.e. the Embden-Meyerhof pathway and succinate, acetate and glycogen [94]. *F. succinogenes* S85 can also degrade hemicelluloses, particularly xylose sugar oligomers xylan by expressing

xylanases [95]. However, it is unable to utilize pentose or xylooligosaccharides released from xylan metabolism [95]. *F. succinogenes S85* also showed high levels of acetylxylan esterase and arabinofuranosidase activities when incubated with wheat straw [95].

2.4.5 Fibrobacter succinogenes and proposed mechanism of cellulose degradation

Unlike other cellulolytic microbes, *F. succinogenes* does not degrade cellulose by using a cellulosome or an extracellular free enzyme system [96]. Some studies proposed that the initial attachment of *F. succinogenes*to cellulose [44, 97] and an unusual orientation of cells along the crystallographic axis of cellulose fibres may facilitate subsequent degradation of crystalline cellulose by *F. succinogenes* [98, 99]. It has been hypothesised that the degradation of cellulose occurs at the cell surface-membrane of *F. succinognenes* by the following three steps: i) adhesion of cells to cellulose fibres via an outer membrane protein complex, ii) disruption of cellulose fibres by carbohydrate active enzymes and transfer of individual cellulose chains to the periplasmic space and iii) cleavage of chain into the sugar molecules [57, 100], a portion of which is released into the medium [101, 102]. It possesses a unique mode of polymer digestion that can produce various sugar components like glucose, cellodextrin, xylose [93, 102]. These properties of *F. succinogenes* makes it a suitable candidate for consolidated bioprocessing development [6]. The proposed mechanism of cellulose degradation is given in Figure 2-4.

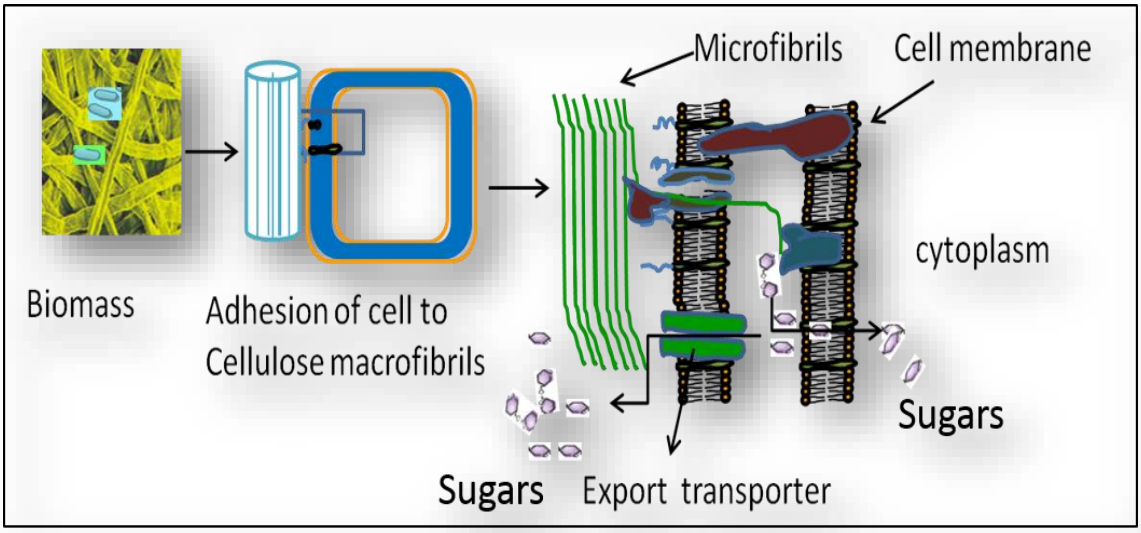

Figure 2-4 Possible mode of cellulose metabolism in *F. succinogenes*. Figure represents adhesion of cells to cellulose fibres, disruption of cellulose fibres by carbohydrate active enzymes on cell surface, transfer of individual cellulose chain to periplasmic space and cleavage of chain into the sugar molecules [57, 100], a portion of which is released into the medium [102, 103].

Previous work has demonstrated that the glycosidic residues located on the surface cellulose binding proteins (CBPs) (especially CBP 180-kDa) now known as cellulose binding modules (CBMs), cellulose binding domains of endoglucanses (EG2, EGF) and chloride-stimulated cellobiosidases are crucial for adhesion to cellulose [97]. Two dimensional electrophoresis (2DE) based proteomics study of the outer membrane (OM) by Jun *et al.* [44] identified 25 OM proteins in cellulose grown cells. A total 16 of these were up-regulated by growth on cellulose compared to glucose grown cells. Out of those, 19 proteins were down regulated or absent when *F. succinogenes* S85 was grown on glucose, which hints at a potential role in cellulose degradation. Fourteen proteins having unknown functions were also identified in the same study. Proteins with an unknown function are not surprising since 50 % of open reading frames (ORF) in *F. succinogenes* have unknown function [44]. The same study also demonstrated

the importance of CBMs in adhesion to cellulose. The efficient adherence ability of this bacterium, encouraged researchers to investigate the mechanism of adhesion and digestion of plant cell walls. Moreover, the presence of 104 glycosyl hydrolases in the genome annotation [104] suggests that a more rigorous investigation is necessary to understand the mechanism of cellulose degradation by *F. succinogenes*.

The enzymatic system of *F. succinogenes* S85, has also been studied at the biochemical and molecular level [105]. Twenty enzymes have been characterized *in vitro* (i.e. many different cellulases, xylanases, ferulic acid and acetyxylan esterases, alpharabinofuranosidases, and alpha-glucuronidases) in this organism [106]. In addition, five different endoglucanases, one cellobiosidase, one cellodextrinase, four xylanases, two acetylxylan esterases were identified [69]. However, the genome of *F. succinogenes* S85 lacks certain essential cellulosomal components, particularly the *cipA* gene that encodes the scaffoldin protein, the primary structural feature of the cellulosome [85].

Recently, two cellulose binding modules (CBMs) (previously known as cellulose binding protein, CBP), a essential component in substrate enzymes interaction of *F. succinogenes* were purified and cloned [107]. CBMs plays important role in assisting the enzymes to hydrolyse cellulosic materials [108, 109]. Further, there were five CBPs isolated and identified in the rumen fluid [69]. Fibrolytic enzymes endB and CelF are considered as important cellulolytic enzymes in *F. succinogenes* [87].

It was believed that the outer membrane proteins are closely associated with adhesion and cellulose hydrolysis as observed by Gong and Forsberg when they studying two adherent mutant strains of *F. succinogenes* [110]. These studies

have provided valuable information about cellulose degradation, but the route of mechanism in *F. succinogenes* not yet been resolved [44].

Therefore, to extend our knowledge about the functional role of these proteins in cellulose degradation and for proper use of *F. succinogenes* in future CBP development, a further high through-put proteomics based study is required. This would generate a dataset and hopefully lead to a deeper understanding.

Part 2

2.5 Biofuel production

2.5. 1 Anaerobes in biofuel generation

Current biofuel research aims to generate energy by-products: mainly advanced alcohols (such as ethanol, propanol, butanol), propanediol and butanediol, diesel, hydrogen and biogas from biological resources [111].

Biofuel-producing microorganisms have wide applications in the field of energy generation from biomass. These microbes can be either aerobes or anaerobes that concentrate energy in the form of biomass and by-products. Particularly, anaerobes always have been an interest of researchers.

Anaerobic digestion has a great ancient history in renewable energy generation (such as biogas) [2, 4, 112, 113]. It has been also widely studied in microbiology for alcohol production, lactic acid fermented food and biogas generation from organic waste [4]. Thus, the anaerobic bioprocessing is a naturally occurring process, which produces useful by-products in the form of biogas and liquid biofuel from biomass.

We are starting to understand the diverse and functional properties of anaerobes. This understanding will be used for biofuel production and could provide vital clues for future energy requirements [114]. Recently, anaerobic digestion of biomass has received considerable attention among the governments of various countries like the United Kingdom [115, 116], Germany and Denmark [117].

Particularly, obligate anaerobes possess unique metabolic features that may have tremendous potential for the development of novel biotechnological tools for biofuel generation. Anaerobic biotechnological approaches for production of liquid energy carriers from biomass are based on two major biological processes: 1) hydrolysis which can convert biomass to simple sugars and 2) fermentation which produces acetone, n-butanol and ethanol (ABE).

However, unlike aerobic microorganisms whose anabolic features are already largely exploited and characterised by biotechnology, anaerobes still need further systematic study for a better understanding of their biology and of their metabolic dynamics prior to their proper utilisation in future biofuel generation.

2.5.2 Microbial biofuel production

Research has focused on renewable biofuels derived using microorganisms in the form of alcohols, hydrocarbons, fatty acids derived from lignocellulosic plant materials. This scientific and technological agenda for the development of renewable biofuel technologies has been a priority in the US and worldwide over the last 35 years. *Saccharomyces cerevisiae* and *Zymomonas mobilis* are the most exploited microbes for glucose fermentation to produce ethanol [118]. However, to make the process economical, it is necessary to utilise non-edible carbohydrates such as lignocellulose and hydrolyse them to sugar monomers in either *Z. mobilis* or *S. cerevisiae* [118]; However, there has been limitations and

subsequent fermentation. Researchers believed that such ideal microbe with both capabilities can be achieved by heterologously expressing cellulases and related enzymes with the development of genetically modified microbes.

2.5.3 Clostridium species and biofuel generation

Clostridium is a large genus and belongs to the firmicutes. It consists of around 100 species [119] that includes some potential pathogens, which cause diseases such as botulism and tetanus [120]. Nevertheless, most obligate anaerobes that belong to this genus have industrial applications [121]. They possess very diverse metabolic pathways, allowing them to grow on a wide range of substrates such as glucose, cellobiose, arabinose, mannose, xylose and component moiety contained in lignocellulosic hydrolysate [122, 123] including polymers starch, hemicelluloses and cellulose [124]. Some of the species of the genus Clostridium are gaining particular interest because of their industrial importance in high energy content alcohols production (such as isopropanol, n-butanol and n-hexanol) [118]. Host species used industrially are those performing acetone-butanol-ethanol (ABE) fermentation [125]. This type of fermentation is one of the oldest large scale industrial fermentation processes for chemical production using microbes. It has become a model for subsequent development of modern fermentation and bioprocessing technology [126]. The concentration of the solvents varies from species to species, substrate to substrate and with culture conditions [127].

In response to the oil crisis in the 1970s, research activities increased in both academia and industry on various aspects of solventogenic clostridia have taken off. Further, in 2006, DuPont and British Petroleum (BP) has announced the industrial-scale set up of ABE fermentation in the United Kingdom. China, has been practicing ABE fermentation at industrial level for several decades. Due to

strong competition with petrochemical industries, ABE industries were closed in late 1990s. However, China re-established all ABE plants in 2006 [128]. As a result, solventogenic clostridia attracted renewed attention and were exploited for ABE fermentation, production, product tolerance, toxicity, technical process developments, sporulation and cellulolytic activity [129]. Several microbial strains from the genus clostridia are able to produce ABE such as *C. acetobutylicum, Clostridium beijerinkii, Clostridium saccharobutylicum, Clostridium saccharoperbutylacetonicum* etc [130]. However, the selection of the species depends on nutrient conditions and substrate availability. *C. acetobutylicum and C. beijerinkii* are the most exploited species from this genus [131].

2.5.4 *Clostridium acetobutylicum* and lignocellulosic solventogenesis

Clostridium acetobutylicum constitutes a model organism since it can produce acetone ethanol and butanol (ABE) by fermentation. It was first discovered by the Israel's first president Chaim Weizmann in early 1920s [129]. *C. acetobutylicum* was then used to produce only acetone which was used to produce the propellant cordite during World War I and butanol was just an unwanted by-product [129]. However, because of the recent interest in butanol as a promising alternative to existing fuels, *C. acetobutylicum kept* being continually exploited at the genetic and physiological level. The most advantageous feature of *C. acetobutylicum* is that it can utilise various carbohydrates sugars, which makes it most suitable to ferment different agricultural, industrial and waste products [132, 133]. Interestingly, *C. acetobutylicum* produces and expresses cellulosomal gene clusters but is unable to utilize cellulose as its sole carbon source [134, 135]. However, a little cellulolytic activity has been observed on carboxymethyl cellulose, and acid swollen cellulose [136]. Recently, a genomic study of *C.*

acetobutylicum revealed strong similarities with the cellulosomal gene clusters of *C. cellulolyticum* [136].

Due to metabolic diversity in substrate utility and product formation, the study is more challenging in *C. acetobutylicum*. Therefore, for the successful utilisation of *C. acetobutylicum* at industrial scale and the increase at its ABE fermentation efficiency, an in-depth knowledge of the organism at the genetic, proteomic and physiological level is essential [137].

2.5.5 Life cycle and metabolic pathway regulation

Since *C. acetobutylicum* possess metabolic pathway diversity in substrate utilisation, the fermentation mechanism is more and more complex. The life cycle of *C. acetobutylicum* is divided into three major phases based on growth: 1) exponential growth phase (acidogenesis, acids formation), 2) transition to stationary growth phase (re-assimilation of acids and simultaneous formation of solvents) and 3) spore formation [129]. During the exponential growth phase, acetic acid and butyric acids are produced (acidogenesis) [138, 139]. The production of the acids results in a lower pH of the medium, thereby affecting intracellular pH and proton gradient dissipation. As a result, to avoid lethal effects of acids, cells shift their metabolism to solvent formation. Some of these acids (e.g. butyric and acetic acid) are utilised by the cells as substrates for solvent production and converted to butanol, acetone and ethanol [140]. The onset of solvent production initiates another complex mechanism of sporulation. Solventogenesis and sporulation both are simultaneously regulated by the transcription factor Spo0A, a master regulator of sporulation. As seen in spore forming *Bacillus substilis,* Spo0A is an initiator of sporulation in solventogenic clostridia [141, 142]. Although acidogenesis, solventogenesis and sporulation are

relatively well understood, the regulatory mechanism at the cell signalling, physical and chemical level still remains to be elucidated [140]. Recently, cell-cell communication (also called quorum sensing system) (agr-dependant) has been observed in the *C. acetobutylicum.* It plays major role in metabolic networks including solventogenesis and sporulation mechanism [140]. A schematic diagram showing the complex life cycle and fate of glucose metabolism in *C. acetobutylicum* is shown in Figure 2-5. Using modern tools, E. T. Papoutsakis and his co-workers studied the life cycle of *C. acetobutylicum* using genomics, transcriptomics, morphology and sporulation [143, 144] and concluded that studies partly understood the regulatory network in solventogenesis and sporulation. Although, in practical terms, *C. acetobutylicum* follows diverse pathways and produces different products at different stages of the life cycle. It makes it challenging to sum up all the reactions for fermentation considering carbon and redox balances. It is proposed that ATP is always associated with acidogenesis and high NAD(P)H levels with solventogenesis [145]. Considerable efforts have been made towards redox balance management by diverting electron flow from hydrogen generation to butanol production for the increase in the production of $NAD(P)^+$ [146-149], as hydrogen generation significantly reduces butanol production [149].

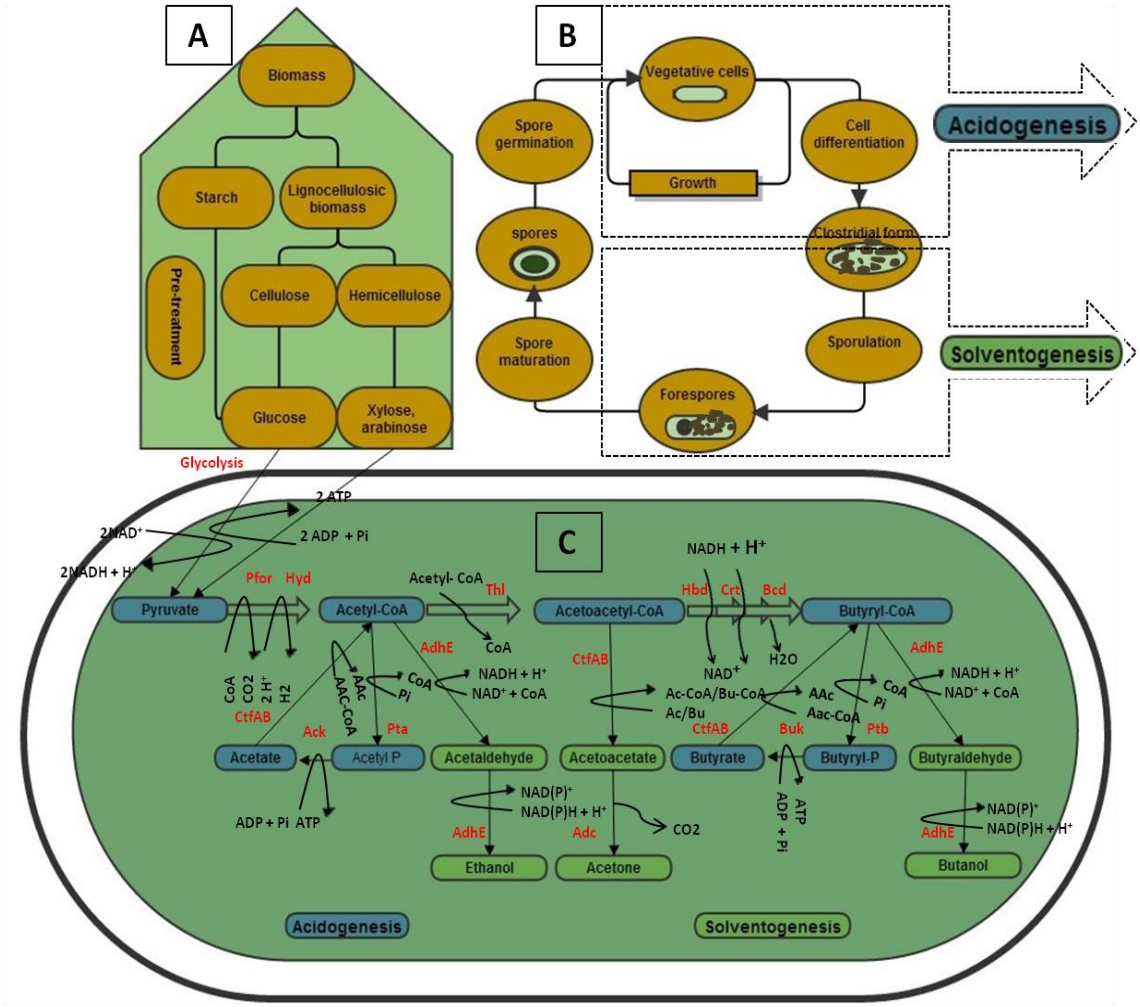

Figure 2-5 Life cycle and glucose metabolism in *C. acetobutylicum* (Modified from [129, 150]. A) Preatreatment of lignocelluloses, B) Life cycle of *C. acetobutylicum* and C) Fate of glucose metabolism through acidogenesis and solventogenesis. The red letters show the enzymes involved in fermentative pathways: Hyd, dehydrogenase; Pfor, pyruvate:ferredoxin oxidoreductase; Thl, thiolase Hbd, 3-hydroxybutyryl-CoA; Crt, crotonase; Bcd, butyryl-CoA dehydrogenase; AdhE, aldehyde/alcohol dehydrogenase; CtfAB, acetoacetyl-CoA:acyl-CoA transferase; Ack, acetate kinase; Pta, phosphotransacetylase; Buk, butyrate kinase; Ptb, phosphotransacetylase; Adc, Acetoacetate decarboxylase

2.5.6 Metabolic engineering of C. acetobutylicum and different study strategies

One of the major benefits of *C. acetobutylicum* is that it can carry out the conversion of a wide range of sugar monomers into biofuels. Although

solventogenic clostridia are able to utilise a diverse range of substrates, the production of biofuels remains limited due to the toxicity of end products (e.g. butanol) and lignocellulose degradation products [151].

This is why researchers worldwide are currently studying *C. acetobutylicum* to improve various aspects of fermentation such as productivity, products toxicity and tolerance, technical process development, cellulolytic clostridia and consolidated bioprocessing development. A review by Eversloh and Bahl summarising the metabolic engineering tools and strategies for *C. acetobutylicum* to improve biofuel production has been published recently [129]. A DNA microarray study on *C. acetobutylicum* under oxygen stress conditions provided details about the detoxification end products and the redox balance [152]. Proteomics investigations on the other hand were initiated in 2002 using two dimensional (2D) gel electrophoresis [153]. Soon after the *C. acetobutylicum* genome was sequenced, mass spectrometry analysis was carried out on soluble proteins compared to transcriptomics data. However, the data set from both studies on butanol tolerance only partially correlated [154, 155]. In 2010, Janssen *et al.* conducted a more systematic proteomics study of acidogenesis and solventogensis in chemostat culture conditions and provided differential protein expressions and correlated with transcriptomics data [156]. Proteomics provides a more detailed understanding of biology than genomic studies since it varies from cell to cell and time to time while genome is almost constant.

Recent developments in systems biology methods can provide vital clues about metabolic pathways and fluxes. Difficulties in genetic accessibility of clostridia are a major problem in metabolic engineering of clostridial species. However, in 2007 a group of researchers developed a targeted gene knock out system called “The ClosTron system” for clostridial species. It can be used for gene silencing. This

method uses a Group II intron that directly causes mutagenesis [157]. This system can be an important platform for metabolic engineering of clostridial butanol production. The various efforts that have been made in last 8 years to improve butanol production in *C. acetobutylicum* using metabolic engineering are summarised in the review published by Eversloh and Bahl [129]. Due to branched fermentative pathways and a complex life cycle, rational metabolic engineering of solventogenic clostridia for improved butanol production is hampered and remains challenging.

Another aspect where researchers are taking considerable interest is in combinatorial metabolic engineering strategies, such as the screening of desirable phenotypes. Screening is one of the oldest methods and proved a successful technique for isolating butanol producing, butanol tolerant strains of *C. acetobutylicum* until now. In recent studies, the mutant Rh8 of *C. acetobutylicum* DSM1731 [154], a butanol tolerant strain and mutant *C. beijerinckii* BA101 [158], an improved butanol producing strain were isolated.

Other attempts on recombinant strategies have also been made recently and the butanol pathway of clostridia was reconstructed in various species such as *E. coli* [159, 160], *S. cerevisiae* [161], *B. Subtilis* and *Pseudomonas putida* [162], *Lactobacillus brevis* [163] and *Clostridium Ijungdahlii* [164]. However, butanol production was found to be significantly low when compared to solventogenic clostridia (without by-product removal). So far, only clostridia can produce butanol, therefore if the process is to be optimised and brought to an industrial level, we need to think about the major issues governing redox balance, energy equilibrium, and potential inhibitors produced during pre-treatment.

2.5.7 Multi-organism approach for direct lignocellulosic biofuel generation

Despite all the above findings, the race for a single organism solution to biofuel production remains complicated and uncertain [6]. Thus, a mixed culture consortium might provide the required breakthrough to develop a commercial level solution [6]. The microbes that bear cellulose degrading as well as biofuel generating capacities are the main focus of researchers for process development for lignocellulosic biomass conversion. Many of these qualities can be found in native microorganisms, therefore the innate power of natural microbes cannot be underestimated.

Biomass processing through biological route involving transformations of carbohydrate polymers into sugars and fermentations of these sugars into valuable by-products is called consolidated bioprocessing (CBP). CBP is an alternative approach with outstanding potential, offering low cost and high efficiency compared with the process needed to produce cellulases [1]. To utilise this potential from natural microbes, microorganisms have to be developed for high cellulose conversion and to produce the desired products at high yield and titre. Both capabilities are present in native microorganisms, but only microbial systems have yet to be improved [1]. CBP comprises four transformations in a single step; 1) hydrolytic enzyme production (cellulases and hemicellulases), 2) hydrolysis of corresponding carbohydrate polymers into fermentable sugars (glucose, mannose and galactose), 3) fermentation of hexose sugars and 4) fermentation of pentose sugars. CBP of cellulosic biomass has been reviewed by Lynd *et al.* [1].

Natural microbial ecosystems are the best examples of CBP, where microbes never live isolated but rather in complex and diverse communities (called consortia) which completes most of the biogeochemical cycles [165]. As researchers are particularly interested in anaerobes [5], CBP requires the combining of native and recombinant microbes that possess both efficient properties of substrate utilisation and product formation [166-170]. However, the development of stable and productive microbial communities is still challenging and for successful implementation in large scale CBP of lignocellulosic materials, we have to develop stable consortia with the necessary functionality, process control and efficiency [170]. Therefore, continued exploration of microbes with cellulose utilisation and biofuel generation capacities will be required.

Microbial bioprocessing seems a promising approach to energy conversion, in particular for lignocellulosic biofuel production [1]. A combination of microbes with desirable abilities such as substrate utilisation and product (biofuel) formation can provide a major breakthrough as a low cost technology. The utilisation of combinations of cultures (co-culture) with distinct capabilities is one of the alternatives for CBP.

Fibrobacter succinogenes is the most efficient cellulose degrader found in the rumen, while on the other hand, *Clostridium acetobutylicum* has the best capability to ferment a diverse range of sugar components (glucose, cellobiose, manose, arabinose, xylose and galactose) into acetone ethanol and butanol [171, 172]. Thus using these two mesophilic anaerobes in a co-culture has potential for CBP development.

A similar multi-organism approach was tested for bioenergy production from cellulosic biomass: *C. acetobutylicum* and *Clostridium cellulolyticum*

demonstrated that the rate of cellulose utilization in the co-culture is improved compared to the mono-culture of *C. cellulolyticum* [173], *Clostridium cellulolyticum and Rhodopseudomonas palustris* were syntrophically grown as co-culture. The increased in cellulose degradation were observed by *C. cellulolyticum* because of removal of inhibitory byproduct (pyruvate) by R. palustris [174], *C. acetobutylicum* and *Ethanoigenens harbinense* were tested for biohydrogen production using microcrystalline cellulose as substrate. The improved cellulose hydrolysis and hydrogen production were observed as compared to that of monoculture conditions [175], and *S. cerevevisiae* and *Scheffersomyces stipites* were used to utilised glucose/xylose mixture to enhanced ethanol production [176].

Part 3

2.6 Analytical section

2.6.1 Proteome

The term proteome was first introduced by Wilkins in 1996 [177]. It deals with the study of entire set of proteins expressed by given cell at given time. Unlike genomics and transcriptomics, proteomics is an advanced step in the study of the biological system, and is relatively more dynamic and inherently complex. The proteome varies from cell to cell, depending on biochemical interactions with genome, and environmental conditions.

The field of proteomics is growing at an amazing rate [178], due to its accuracy, sensitivity and advancement in analytical tools. The proteomic techniques have become widespread, expanding from mere protein profiling to accurate and high-

throughput protein quantification across different samples [178]. Peptide based analysis of proteomics using mass spectrometry is widely accepted. After qualitative proteomics, quantitative proteomics is gaining global interest since quantitative changes in proteins reflect the genuine state of the biological system [179]. Various separation techniques combined with mass spectrometric (MS) analysis have been developed recently to achieve better proteomic analysis [180].

In order to increase throughput and access to proteins that are difficult to resolve by gel electrophoresis, a alternative “Shotgun” proteomics approach was developed [181-183]. Bottom-up or so-called “shotgun” proteomics is the most widely used analytical technique for protein analysis [184, 185], and relies on three main steps: 1) protein digestion (trypsin proteolysis) into peptides, 2) separation by high performance liquid chromatography (HPLC) and 3) MS/MS identification and quantification of parent proteins [186]. Every step in bottom-up proteomics plays an important role and has influence on the downstream MS/MS analysis.

The following sub-sections provide brief overview of steps in bottom-up proteomics as they are used in this PhD project.

2.6.1.1 Protein digestion

Protein digestion into peptides is one of the key step in proteomics, usually followed by LC-MS analysis. There are two major strategies for converting proteins from a biological source to peptides for MS/MS based proteomic analysis.

In-gel digestion of proteins was first introduced by Shevchenko *et al*. in 1996 [187] and has continuously been used by researchers since. The procedure relies on destaining, alkylation, reduction, trypsin digestion and peptide extraction. In-gel digestion allows proteins to become separated and denatured within a gel matrix and trypsin digested. It is most suitable for complex and hydrophobic proteins, which are very difficult to undergo proteolysis in solution. The peptides generated from in-gel digestion are devoid of salts and detergents due to extra washing and cleaning steps. This prevents subsequent contamination during mass spectrometry analysis. In-gel digestion allows one to separate proteins based on their molecular weight, thus increasing the dynamic range of analysis of proteins [188].

In-solution digestion, on the other hand, is another useful protein digestion alternative. It is relatively simple workflow in terms of sample handling and speed, but needs constant maintenance of the LC-MS instrumentation [189]. It is most favourable for low or medium complexity protein samples. However, folding of proteins in solution may prevent proteolysis and therefore some denaturing agents are used prior to digestion (such as urea, surfactant, detergent etc). The procedure relies on reduction, alkylation, and trypsin digestion. Thus, each procedure has its own advantages and disadvantages and can be applied on the basis of nature of protein samples.

2.6.1.2 Chromatographic separation of peptides

HPLC based two or multidimensional peptide fractionation is most applied chromatographic technique in proteomics to reduce the complexity of the peptide mixture obtained from tryptic digestion before injection to the MS analysis. First-dimensional separation of peptides (pre-fractionation) plays important role in

comprehensive proteome analysis strategies [190]. To further improve MS/MS performance, two-dimensional LC (2-D) or multi-dimensional LC (M-D) are often used [191]. Thus to reduce complexity of samples, reverse phase liquid chromatography (RPLC) is the method of choice as a 2-D separation prior sample subjected to MS/MS analysis. Off-line multidimensional mode is one of the most popular for peptide separation/fractions and uses either strong cation exchange (SCX) or hydrophilic interaction liquid chromatography (HILIC) separation followed by online RP-LC (2-D) and subsequent MS/MS analysis [192]. The success of the M-D separation technique purely depends on the orthogonality of the separation dimensions and efficiency of the individual separation technique applied [191].

Pre-fractionation (1-D separation) can be achieved by number of chromatographic techniques. Most popularly used techniques are; ion exchange, hydrophilic interaction, isoelectric focussing (IEF) and mixed mode pH reverse phase (RP) [191].

SCX is an ion exchange chromatography, where the peptides are separated based on charge and eluted using a high gradient salt buffer or by changing the pH of the eluents [193]. However, recent observations suggests that the SCX-RP mode of 2-D separation is probably not ideal option in terms of the orthogonality because of peptide cluster formation and narrow retention window [194, 195].

Another important separation mode is HILIC. HILIC is based on hydrophilic interactions. It was first introduced by Linden *et al*. in 1975 [196]. HILIC can be an excellent alternative to SCX [197]. HILIC uses a hydrophilic stationary phase and organic mobile phase reverse to RP-chromatography. In this mode, elution

of the peptides occurs with increasing hydrophilicity (water content) of the mobile phase. The performance also depends on partitioning and electrostatic interactions or hydrogen bonding with the stationary phase [198]. Different stationary phases have recently been developed for HILIC separation including; 1) underivatized silica stationary phase with functional groups siloxanes, silanols with/without small amount of metals, 2) derivatized silica, a cation exchanger Polysulfoethyl A, 3) the weak cation exchanger PolyWAX, 4) TSK gel amide 80, 5) zwitterionic (ZIC)-HILIC and click saccharides [197].

2.6.1.3 Ionization

A mass spectrometer mainly consists of three major parts: an ion source, a mass analyzer, and a detector. The ionisation is particularly necessary to transform biomolecules in the liquid phase to gaseous phase ions in order to analyse by MS. Ionisation resulting in loss or gain of charge from neutral species, produces peptide ions. There are two soft ionization techniques in order to volatilize peptides, an electrospray ionization (ESI) [199] and matrix assisted laser desorption/ionization (MALDI) [200]. In ESI, charged bio-molecules are accelerated by passing samples through a high voltage needle at atmospheric pressure[199]. In MALDI, the UV laser beam bombards matrix crystals to produce ionized matrix molecules that transfers protons to the analytes and charges analyte molecules [201]. ESI is the most common method coupled with q-TOF and ion-trap type instruments. An evaporated droplet of positively charged ions further enters into the analysers. ESI and MALDI interfaces are being used with different analysers [202].

2.6.1.4 Analysers

A mass analyser is the next compartment of the mass spectrometer that separates charged analytes on the basis of charge-to-mass ratio and transmitted to the detector. Mainly, four types of mass analysers are commonly used in mass spectrometry analysis of bio-molecules, which include; Quadrapole mass analyzer, Time Of Flight (TOF), Quadrapole ion traps and Fourier transform ion cyclotron resonance (FTICR) [203].

In this PhD, the instrument used was: a QTOF instrument (QSTAR-XL, Applied biosystems AB/MDS Sciex) coupled with ESI source (Figure 2-6A) for quantitative proteomics of *C. acetobutylicum* ATCC 824 in response to lignin and membrane proteomics of *F. succinogenes* S85 for the discovery of cellulose degradation mechanism. Similarly, qualitative proteomics was carried out for technical comparison between in-gel/in-solution/SCX/HILIC using an ion-Trap instrument (Brucker HCT Ultra PTM Discovery, Bruker Daltonics, Bremen, Germany) (Figure 2-6B).

2.6.1.5 Quadrapole Time-of-Flight (QTOF) mass analyzer

A quadrapole mass analyzer separates ions using an electric field. It consists of 4 parallel rods/pole with opposite polarity which generates radio frequency (RF) fields when voltage is applied to each rod [204]. The electric force on the ions results in oscillation of ions between 4 rods. The QqTOF tandem MS consists of 3 aligned quadrapoles, Q0, Q1, and Q2 followed by a TOF analyser [205]. Quadrapoles can be used as mass analyzer as well as ion transmitters. Q0 and Q1 are normally used in RF mode. The ionized samples from the ESI source

travel to the Q0 where ion cooling and selection takes place. The focused ions are then transmitted to Q1. The selected ions based on their trajectory transmit stability (precursor ions) are further transmitted to the Q2 where they undergo another treatment called collision induced dissociation (CID). In Q2, precursor ions accelerated and collide with inert gas molecules resulting in fragment ionswhich further enter into the time-of-flight tube and are subsequently detected by detector (Fig 2-6A).

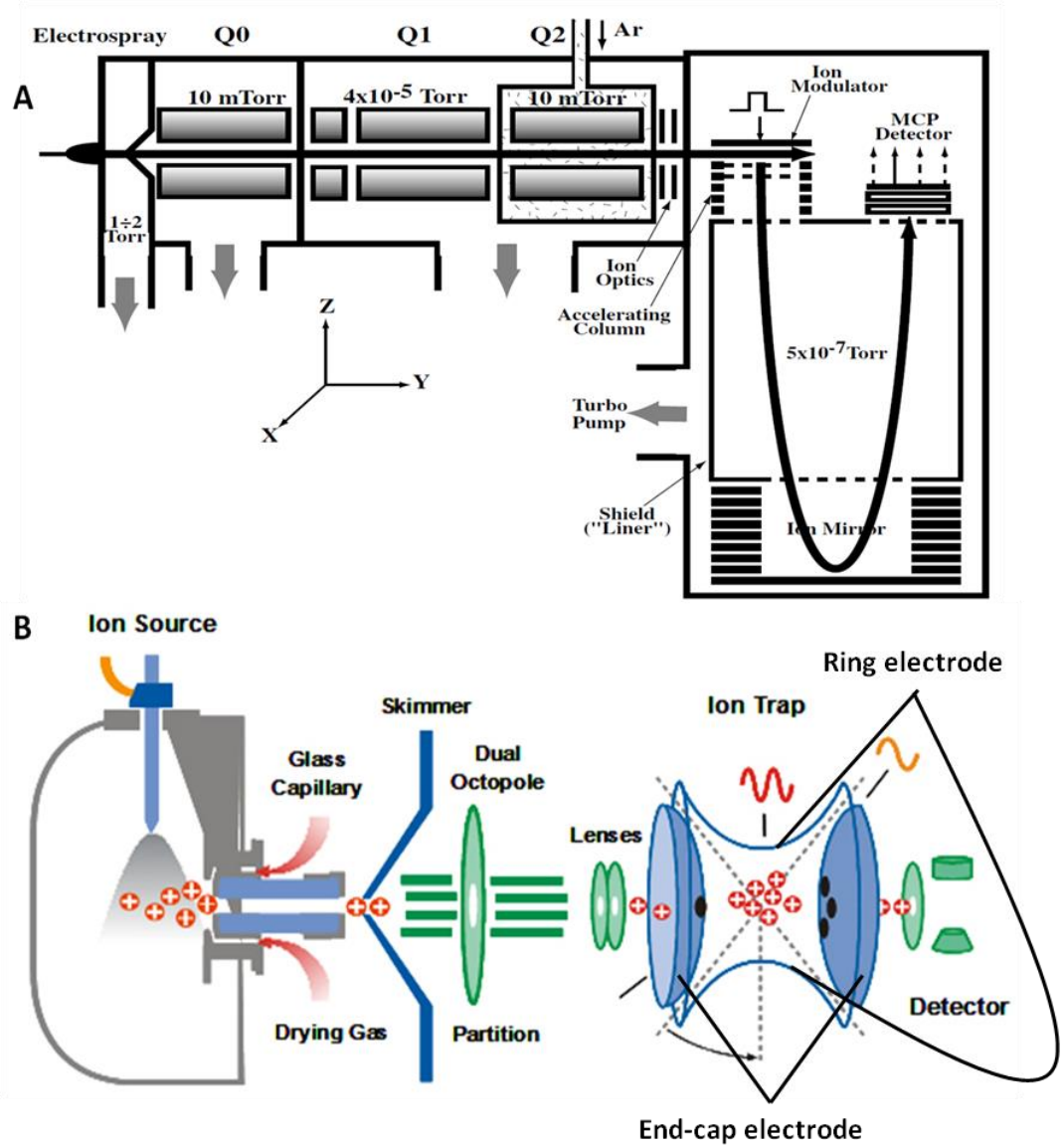

Figure 2-6 A) Schematic presentation of tandem Qq-TOF mass spectrometer consisting of electro spray tips, curtain gas, quadrapoles Q0, Q1 and Q2 and TOF mass analyser (Reproduced from [205]). B) An experimental view of Ion-Trap MS (Figure adopted from Bruker Daltonics HCTUltra PTM discovery manual).

2.6.1.6 Q-Ion Trap (QIT) MS

In 3D QIT MS, RF voltages are applied to the ring electrodes. The end cap electrodes are purposely kept at ground potential. Ions above a selected value of m/z ratio are trapped in the mass analyzer. The RF field in the trap makes a potential well. Ions roll down and get trapped in the well. The ions are further collided with helium and lose further energy, slowed down and focused in the centre of the trap. To transmit the ions from the trap, an additional auxillary RF (1/3rd of the frequency of the main frequency) is applied to the end cap. The main RF amplitude from the ring electrode and the auxillary RF are increased simultaneously, each *m/z* starts resonating and ejected out which is further detected by the electron multiplier/detector. A continual increase in the RF amplitude applied to the ring electrode results in the ejection of ions viathe endcap electrodes, towards an ion detector. The ions having a smaller m/zare ejected and detected first [206] (Fig 2-6B).

2.6.1.7 Quantitation

Protein quantitation in complex biological systems is most exciting but still a challenging part of proteomics. The small changes at the protein/peptide level in response to biological/physiological conditions can be measured by quantitative proteomics [207]. Over the last decades, tremendous progress has been made in the development of proteomics methods in terms of high throughput techniques, designing of algorithms and software for quantitative analysis [207, 208]. Quantitative proteomics can be divided into two forms: absolute quantitation that determines exact mass concentration of proteins and relative quantitation that determines relative abundance of specific proteins in samples. MS

quantitation is normally based on relative intensity of m/z ratio of peptides. Relative quantitation methods include: Isotope-coded affinity tags (ICAT) [209, 210], Isobaric labelling (Tandem mass tags (TMT) [211], and Isobaric tags for relative and absolute quantitation (iTRAQ)) [212, 213], Label free quantification [214], Metal-coded tags (MeCATs) [215], and Stable isotope labelling with amino acids in cell culture (SILAC) [216]. In this project thesis, the iTRAQ quantitation method was used.

2.6.1.7.1 iTRAQ (Isobaric Tags for Relative & Absolute Quantitation)

In iTRAQ chemical tags attach to the N-termini and side chain amines of peptides and allows relative quantitation of proteins from multiple (upto 8) phenotypes in single experiment [217]. The tags consist of three main component parts: a reporter ion (N-methylpiperazine), a mass balance group (carbonyl), and a peptide-reactive group (NHS-ester) (Figure 2-7). Since all iTRAQ tags are of equal masses they described as isobaric tags. However, during fragmentation by MS/MS, carbonyl moieties are released as neutral loss and produces isotope-encoded reporter ions of different masses that helps to quantitate the proteins from which peptides originates [217]. The piperazine groups range in mass from 114-117Da (in case of 4-plex) and 113-121 Da (excluding 120) in the case of 8-plex tags. Tags are covalently linked to peptides through the amine reactive N-hydroxysuccinimidyl (NHS) ester moiety [217]. In a normal 4/8-plex experiment, each of the tags is used to label a different pool of peptides, derived from a different set of samples. After the labeling, the peptide samples are combined together. As the iTRAQ tags are physio-chemically identical, peptides from each sample are co-elute from the liquid chromatography step and are introduced to the MS simultaneously.

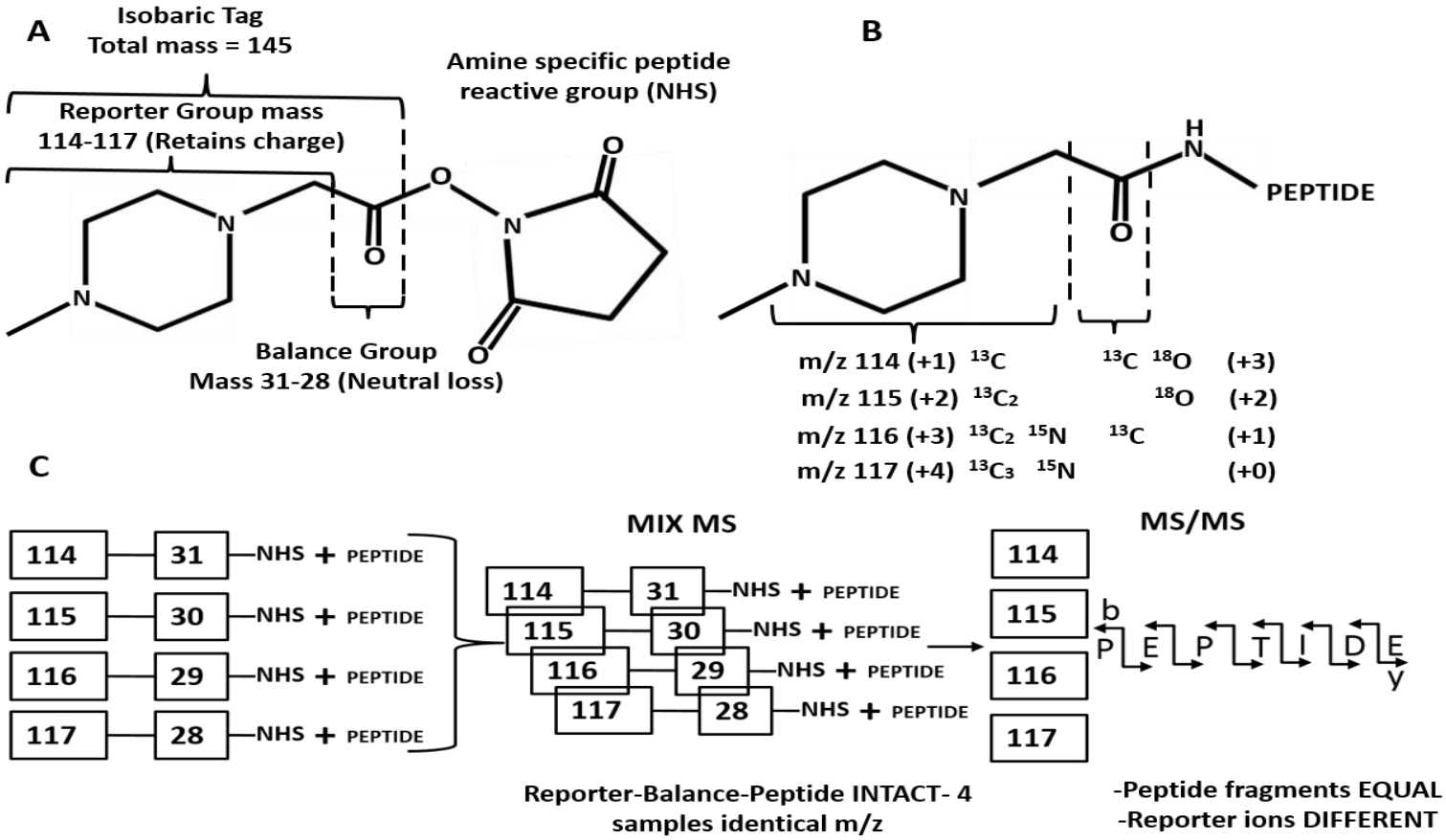

Figure 2-7 A) Diagram showing different components of an iTRAQ tag B) Formation of amide link between tag and peptide amine group (N-terminal or amine group of lysine side chain), C) isotopic tagging (multiplex) of 4 identical peptides and subsequent MS/MS analysis (reproduced from [217]).

2.6.2 Membrane proteomics and challenges

The physical properties of the proteins play an important role in protein structure and protein-protein interactions. The amino acid composition and polarity of its side chain incorporates hydrophilic/hydrophobic characters in the protein. Hydrophilic/hydrophobic amino acids determines protein structure and their physical location in the system. The outer surface of the soluble proteins are rich in polar amino acids particularly serine and threonine whereas membrane proteins possess an outer ring with hydrophobic amino acids like lysine and arginine which trap them in a lipid layer and avoid to solubilisation in water.

Membrane proteins constitute 20-30 % of the cellular proteome and are mainly involved in various physiological processes including cell adhesion, signal transduction, nutrient uptake, transportation, and endocytosis [218, 219]. There has recently been a great increase in the interest in membrane proteomics analysis.

Mass spectrometry based approaches have been widely used in proteomics studies [220] as they provide detailed information about qualitative and quantitave measurement. However, unlike soluble proteins and other organelle proteins, these approaches cannot be directly applied to the membrane proteomics because of the difficulties in membrane protein extraction/solubilisation and also of subsequent protease (trypsin) digestion for shotgun proteomics [221].

2.6.2.1 Membrane biotinylation

Biotinylation is the covalent attachment of biotin (Vitamin H) to biomolecules such as protein, DNA and RNA. In 1942, it was observed that a biotin molecule binds

non-covalently to egg white glycoprotein (avidin) and bacterial protein (streptavidin) with high affinity and specificity [222]. Since then the chemical modification of the biomolecules with biotin has extensively been used as an affinity based tool in many biochemical and biomedical research areas [223]. The biotin-avidin complex forms extremely rapidly and, remarkably, it is unaffected by pH, temperature, organic solvents and other denaturing agents.

Chemically-modified biotin derivatives can be used for surface protein biotinylation, they are therefore useful in membrane protein analysis. Basically, the protein analysis relies on four steps: biotinylation, separation, purification and protein analysis (Figure 2-8). Separation and purification of the biotinylated proteins can be achieved by avidin/streptavidin coated with solid supports such as resin, magnetic beads, microtitre plates and chips. Biotin affinity technology has already been used to study the membrane proteomics of the intact bacterial cells [224-229].

Biotinylation can be achieved chemically or enzymatically. However, the chemical method is more popular among the researchers, since it provides greater flexibility than enzymatic method and can be performed both in *vitro* and *vivo* conditions [230].

N-Hydroxysuccinimide (NHS) ester activated biotin is the most popular reagent since it contains an extended spacer arm to reduce steric hindrances associated with avidin/streptavidin binding [223]. The N-hydroxysulfosuccinamide (NHS) ester group of this complex efficiently reacts with primary amine groups (-NH2) on the protein moieties in aqueous solutions of neutral to basic pH. Since it is modified with a sodium sulfoxide group on the succinimidyl ring, Sulfo-NHS-LC-biotin is water soluble which, makes it suitable for labelling surface and

membrane proteins and as it cannot permeate the membrane one can avoid soluble protein contamination [228]. The extraordinary affinity of the biotin to the avidin/streptavidin allows for protein biotin complexes to separate from the solution. The separated proteins then can be eluted out from the streptavidin-biotin complex using reducing agents and isolated proteins used for further proteomics analysis.

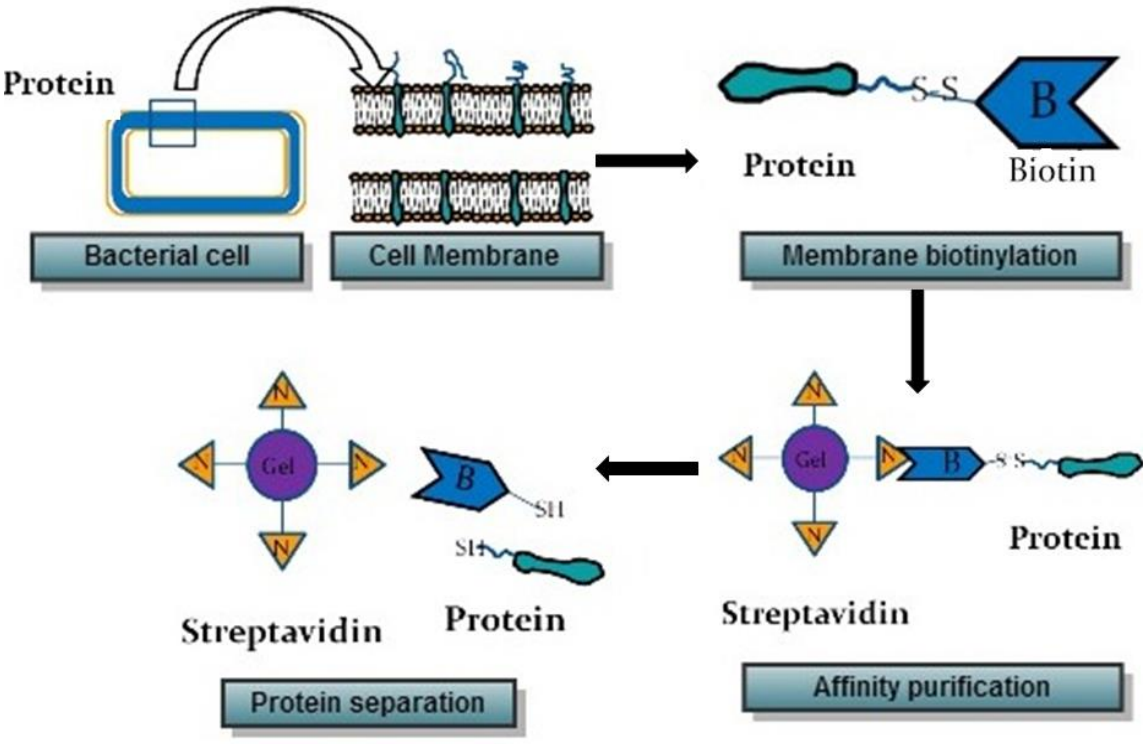

Figure 2-8 Schematic workflow of membrane biotinylation

2.6.3 By-product measurements

2.6.3.1 Extracellular metabolites analysis by gas chromatography- mass spectrometry

Gas chromatography (GC) is an analytical instrument used for the separation and identification of chemical compounds. In this chromatography samples are extracted and dissolved in a mobile phase (it can be a gas, a liquid or a

supercritical fluid). The mobile phase passes through an immobile, immiscible stationary phase in column [231].

In this PhD, three analytical and separation GC-based techniques were used: including, Gas chromatography-Flame ionization detector (GC-FID), Gas chromatography- Thermal conductivity detector (GC-TCD) and Gas chromatography-mass spectrometry (GC-MS) for the analysis of fermentation by-products and H_2 gas analysis depending on nature of samples detection limits in each of the each techniques.

GC-FID

In this type of GC, the detector consisting of a hydrogen/air flame and collector plate. The sample passes through column and reaches the flame which ionise molecules and produces ions. These ions we further collected on biased electrodes and produce electromagnetic signals. GC-FID is one of the most extremely sensitive and widely used techniques in the field of chemistry [231]. It is relatively less expensive and requires less maintainance. GC-FID use to measure organic compounds with high linearity and detectivity. The only limitation of GC-FID is that it destroys samples during the combustion. It can not detect inorganic compunds because of its nonselective nature. In this study this technique used to analyse fermentation organic by-products (ethanol, butanol acetic acid and butyric acid).

GC-TCD

This instrument fitted with a detector which consists of an electrically-heated wire or thermistor.The temperature of the sensing element depends on the thermal conductivity of the gas flowing around it. Changes in thermal conductivity, during organic molecules displace some of the carrier gas molecules, resulting in

temperature rise in the element which is sensed as a change in resistance. The TCD is possibly not that sensitive like other dectectors but it is non-specific and non-destructive. Normally, gas chromatograph has two pairs of TCDs. One pair is placed in the column effluent to detect the separated components as they leave the column, and another pair is placed before the injector or in a separate reference column. The resistances of the two sets of pairs are then arranged in a bridge circuit [232]. GC-TCD is reatively simple technique that provides large linear dynamic range and detects organic and inorganic compounds. Due to its non-destructive nature, sample can be collected after detection. It is widely use to detect natural gas in samples. The main limitation with TCD is its low sensitivity. In this study we used this technique to detect the H_2 gas in the sample.

GC-MS

GC coupled with MS is a robust system with excellent separation capabilities and higher throughput for metabolomics workflows [233]. Figure 2-9 shows the schematic construction of a GC-MS instrument. Separation depends on the boiling point of metabolite and its interaction with stationary phase of the column used. As we analysed our samples for volatile compound, we skipped derivatization and the sample preparation steps. Procedure for analysis of fermentation products was used as previously described by Pham *et al*. [234] where the supernatant was removed from fermentation broth, centrifuged at 16000 x g for 5 min, and directly injected to the GC. The sample is introduced into the inlet of the GC instrument via an injection port kept at high temperature. Generally two different injection methods are used splitless (whole sample is introduced into the column) and split injection (only small portion is introduced for trace analysis). The polarity of the column can also be changed by using different stationary phase such as DB-5, DB-50 and CPSil-8. Metabolites passing through

the column are transmitted to GC and ionized by Electron-Impact (EI) ionization [235]. In EI-mode, the vaporized metabolites are impacted by beam of electrons to fragment and ionize those molecules. The ionized fragments are then pushed into the mass analyzer. Finally, analytes are identified using database such as NIST or plant specific metabolome database Glom (http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/ gmd.html). The advantages of the GC-MS technique used in this study is its reproducibility. However, common disadvantages of this techniques are low resolution and mass discrimination. In chapter 4, this technique was used to analyse fermentation organic by-products (ethanol, butanol acetic acid and butyric acid).

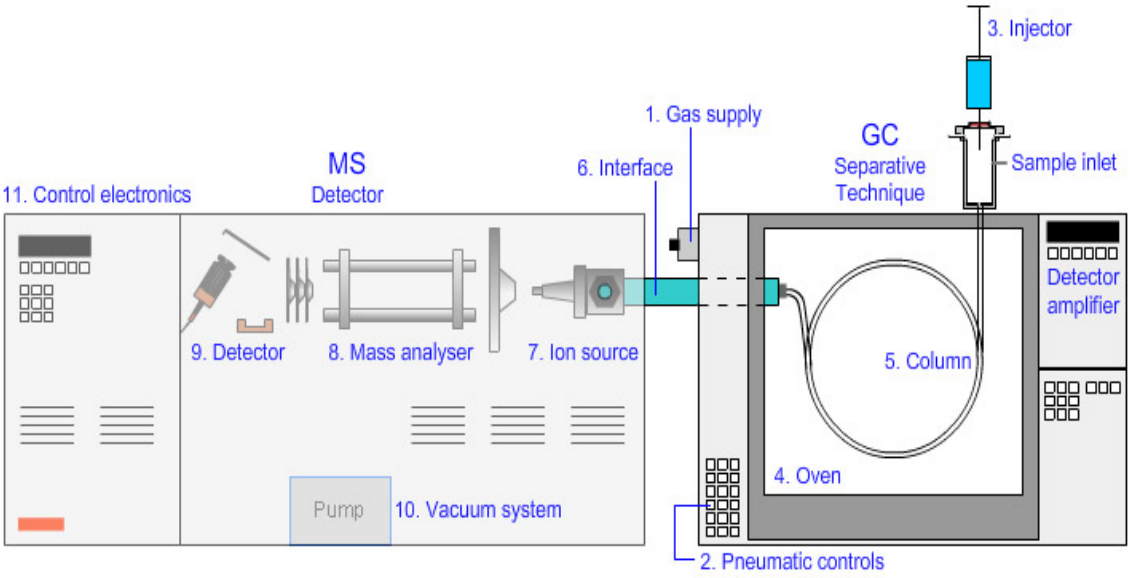

Figure 2-9 Schematic of gas chromatogram interface to ion-trap mass spectrometer system (Reproduced from http://www.chromacademy.com)

2.7 Thesis aims

This chapter detailed the fundamental aspects of second generation biofuels production and the existing literature about two important anaerobes with the

distinct capabilities for saccharification and fermentation of cellulosic materials. These organisms have been studied at genomic, transcriptomic, proteomic and metabolomic level. Finally, the importance of co-culture in consolidated bioprocessing development was discussed.

Consequently, the focus of this thesis comprises two main areas namely cellulose degradation by *F. succinogenes* and biofuel generation by *C. acetobutylicum*. The first part will focus on the detailed study of the effect of lignin on biofuel production in *C. acetobutylicum* (Chapter 4) using shotgun proteomics. We shall seek to characterise the influence of lignin on various metabolic processes in *C. acetobutylicum*. Isobaric mass tagging (iTRAQ) technique will be employed for quantitative assessment of global changes in the proteome of *C. acetobutylicum*.

The second part will focus on the detailed study of the proposed mechanism of cellulose degradation in *F. succinogenes* using selective membrane (Chapter 5) and quantitative proteomics method (Chapter 6). This information would be helpful to establish better cellulose degradation systems for consolidated biofuel generation. Further a separate chapter (Chapter 7) focuses on utilisation of chemically and biologically pre-treated biomass for fermentation into ABE using *C. acetobutylicum.* The first part of this chapter comprises chemically treated Miscanthus biomass directly converted into ABE production using *C. acetobutylicum* and second part comprises simulatanious sachharification and fermentation using co-culture (*F. succinogenes and C. acetobutylicum*) condition for consolidated bioprocessing development for ABE production.

Chapter 3

An experimental comparison between tryptic digestion approaches and HPLC peptide separation techniques for proteome LC-MS/MS analysis

Abstract

Effective protein digestion and subsequent peptide separation by HPLC techniques are decisive steps in successful mass spectrometry (MS) analysis of proteins. There is a need to improve them due to issues of relative low proteome coverage often experienced. For this purpose, a complex proteome lysate were extracted from the *Clostridium acetobutylicum* ATCC 824. This study compares in-gel and in-solution digestion of proteins and two commonly used separation techniques in proteomics; hydrophilic interaction liquid chromatography (HILIC) and strong cation exchange (SCX) chromatography. The separation techniques used for fractionation of peptides affect peptide recovery from trypsin digestion. Our results showHILIC separation provides better protein/peptide recovery than SCX.

3.1 Introduction

Proteomics is defined as the set of techniques used to characterise proteins in complex biological systems. A proteomic study can deal with the identification, the quantification, the posttranslational modification, and the interaction of proteins within a system [236]. Ultimately, the aim of proteomics experiment is to extract information about protein structure, function and biological adaptation. It attempts to do this ultimately for the whole protein complement of a cell. Protein expression changes with the cellular and environmental conditions, thus differs from cell to cell and time to time [237]. As compared to the genome, the proteome is highly dynamic and inherently more complex, since one gene can give rise to different protein isoforms which can be further modified post translationally. Each of these forms can have a specific function and activity, therefore a better understanding of the biochemical processes that take place in the cell are needed [238].

To identify the setsof protein content that genome encodes of one organism is obviously important. It is also important to know the quantity of proteins in the cell. Quantification is more relevant from the biological point of view because comparison between different cellular states and response of the cell to internal or external stimulus can only be understood when quantification strategies are employed [179]. Protein quantification, however, still represents a huge challenge for researchers since they have to develop efficient and rapid methods for identifying and quantifying a large number of proteins from the complex biological samples.

As previously mentioned, the proteome is a very complex system where proteins are present at concentrations varying from very low copy number per cell (less than 10 molecules per cell) to high copy numbers (more than 100,000), this dynamic range has been shown [239-242] to limit the number of proteins which it is possible to identify and quantify in a single experiment. One-dimensional (1-D) and two-dimensional (2-D) polyacrylamide gel electrophoretic separations are relatively old protein separation methods [243]. They have several disadvantages, such as detection limits with the concentration and hydrophobicity of proteins, and they are time consuming and tedious processes [244, 245].

Due to this limitations associated with 1-D and 2D gel separation of proteins, researchers continuously sought to develop alternative, gel-free liquid chromatographic and advanced electrophoretic approaches [243, 246, 247]. Newer gel-free separation approaches used for peptide identification and quantification. Protein information is then inferred from the proteolytic peptides resulting from protein digestion.

Various gel-free based separation techniques combined with mass spectrometry (MS) analysis have been recently developed to achieve better proteomic analysis [180]. Among these different strategies for proteomic analysis, the bottom-up so called “shotgun” proteomics approach is the most widely used [184, 248]. This analytical technique relies on complex protein samples being enzymatically digested into peptides followed by separation by 2D and subsequent MS analysis. Pre-fractionation is usually applied (1-D) using high performance liquid chromatography (HPLC) using an off-line separation strategy such as strong cation exchange (SCX) or hydrophilic interaction liquid chromatography (HILIC). Collected fractions then further subjected to an on-line separation using reverse-

phase chromatography which is typically connected to a MS interface with an electrospray ionisation source.

Every step in bottom-up proteomics plays an important role and has an influence on the downstream analysis of proteins. Protein digestion and peptide separation are equally important steps in shotgun proteomics, since these overwhelms the problems with direct mass spectral analyses of proteins and complex mixture of peptides obtained from digestion.

Protein digestion is the process of conversion of proteins to peptides. In most cases, trypsin enzymes are used to digest the proteins. It cleaves proteins at specific sites (at carboxyl side of lysines and arginines) and produces peptides of suitable lengths for MS sequencing [189]. Following digestion, HPLC based multidimensional peptide fractionation is used to reduce the complexity of the peptide mixture obtained from tryptic digestion before injection to the MS analysis platform [249]. Therefore, there is wide interest in high resolution fractionation of peptides in shotgun proteomics for both qualitative and quantitative proteomics [250]. First-dimensional separation of peptides (pre-fractionation) plays critical a role in comprehensive proteome analysis strategies [190]. SCX is a most popular separation [251, 252] than HILIC in shotgun proteomics as a first dimension. To further improve MS/MS performance, two-dimensional LC (2-D) or multi-dimensionalLC (M-D) are often used [191]. Reverse phase liquid chromatography (RPLC) is the method of choice as a 2-D separation prior sample to MS/MS analysis.

Our technical report compares two digestion systems (in-gel and in-solution) of a complex protein mixture and two subsequent peptide separation techniques

(SCX and HILIC) to obtain a better understanding of these two crucial steps in shotgun proteomics.

3.2 Experimental section

3.2.1 Cell culture procedures, harvest and proteome extraction

All solvents, such as acetonitrile and water used in this experiment were HPLC grade (Fisher Scientific Loughborough, UK). All chemicals and reagents were purchased from Sigma Aldrich (UK) unless otherwise detailed.

Duplicate culture sets of *Clostridium acetobutylicum* ATCC 824 (hereby denoted as *C. acetobutylicum*) were cultivated under anaerobic conditions in 125 mL serum bottles containing media described by Lopez-Contreras *et al*. [253]. The medium included 0.75 g L^{-1} KH_2PO_4, 0.75 g L^{-1} K_2HPO_4, 0.348 g L^{-1} $MgSO_4$, 0.01 g L^{-1} $MnSO_4.H_2O$, 0.01 g L^{-1} $FeSO_4.7H_2O$, 1 g L^{-1} NaCl, 5 g L^{-1} yeast extract and 2 g L^{-1} $(NH_4)_2SO_4$. In all experiments, 5 g L^{-1} cellobiose was added as a carbon source (instead of glucose) and 1 g L^{-1} cysteine hydrochloride was added as a reducing agent (instead of asparagines). 100 mL of medium was prepared anaerobically in 125 mL serum bottles. It was heated to boiling and flushed with 100 % N_2 gas for 10 minutes following by cooling with continuous flushing of N_2 gas for another 10 min. Finally, bottles were closed with butyl rubber, crimped sealed and sterilised at 121^{o}C for 15 min. The sterilised medium was inoculated with freshly prepared inoculum (OD at 600nm equal to 1.6) from the stock culture and incubated at 38^{o}C for 14 and 48 hrs. Cells were harvested from two culture broths (100 mL each), one was harvested at the mid exponential phase (OD at 600nm equal to 1.43) and the second was taken at the late stationary (OD at 600nm equal to 2.6) phase by centrifugation at 10,000 g for 5 min at 4^{o}C and

mixed together to obtain a complex protein lysate sample. The complex protein lysate was produced to increase the dynamic range of detectable peptides during the fractionation.

3.2.2 Cell harvesting and protein extraction

Harvested cell pellets were washed twice with phosphate buffer solution (PBS) (pH 8 containing 8 g L^{-1} NaCl, 0.35 g L^{-1} KCl, 1.32 g L^{-1} Na_2HPO_4, and 0.081 g L^{-1} NaH_2PO_4) and finally once with protein extraction buffer, triethyl ammonium bicarbonate (TEAB) (0.5M, pH 8.5). Cells were then re-suspended in 600 µL extraction buffer (pH 8.5 containing 0.95 g L^{-1} Sodium dodesyl sulphate (SDS) and 300 mg sterilized glass beads (425-600µm)). Cells were disrupted using bead beating (Disruptor Genie, Scientific Industries Ltd, USA) 20 times for 1 minute, with 1minute on ice between each run. Unbroken cell and debris were pelleted by centrifugation at 3,000g for 5 minute. Supernatant-contained soluble proteins were transferred to new low-protein binding Eppendorf tubes (Eppendorf, Cambridge UK). This fraction was clarified by centrifugation at 21,000 g for 90 min at 4^{o}C.

Proteins were acetone precipitated using a ratio of 1:4 (v/v) of sample to acetone. The mixture was incubated (approximately 16 hrs) overnight at -20^{o}C. Finally, protein pellets were re-dissolved in 0.5 M TEAB buffer containing 1 g L^{-1} RapiGest, a protein solubilising reagent (Waters, Milford, MA). Protein concentration was determined by Bradford assay according to the manufacturer's protocol (Sigma Aldrich). Briefly, 20 µL of protein sample (1:1 dilution) was mixed with 980 µL of Bradford reagent mixed and incubated for 10 min at room

temperature. Intensity was recorded at 580 nm and concentration determined using a standard calibration curve obtained from standard BSA.

3.2.3 Protein digestion

Four hundred microgram protein samples were equally divided into four low-protein binding Eppendorf tubes. Two sets of 100 μg of each were considered for in-gel and in-solution digestion respectively.

3.2.3.1 In-gel digestion of proteins

In-gel digestion allows proteins to become denatured within the gel matrix. These denatured proteins then get digested by trypsin. It is most suitable for complex and hydrophobic proteins which are very difficult to proteolyse. SDS-PAGE was performed on duplicate samples as per the protocol described by Sambrook and Russel [254]. Briefly, 100 μg of each sample was mixed with sample buffer and boiled at 95^{o}C for 5 min and loaded to the gel and allowed to run until reaching the resolving gel. After reaching the resolving gel, a single band for each sample was sliced carefully from the gel and placed in a new Eppendorf tube. The digestion procedure was performed as previously described by Karunakaran *et al.* [255] with few modifications. Instead of ammonium bicarbonate (AB) we used TEAB buffer for the subsequent steps. The protein bands were distained twice with 400 μL of 200 mM TEAB in 40 % acetonitrile (ACN) in water by incubating at 37^{o}C for 30 min. The supernatant was discarded and gel pieces dried in a vaccum concentrator (Vacuum concentrator 5301, Eppendorff, UK) for approximately 5 min at 30^{o}C. Entrapped gel proteins were reduced using 400 μL of HPLC grade water containing 1μL of 50 mM tris 2-carboxyethyl phosphine hydrochloride (TCEP) by incubating at 60^{o}C for 1 hour. Gel pieces were briefly

centrifuged at 13,000 g for 10 s and all liquid was discarded. In the subsequent step, proteins were alkylated using 400 µL alkylation buffer containing 1µL of 200 mM methyl methanethiosulfonate (MMTS) at room temperature for 30 min. Gel pieces were washed twice with 400 µL of 50 mM TEAB solution by incubation at room temperature for 15 min. Finally samples were washed once with 400 µL of 50 mM TEAB in 50 % ACN for 15 min by incubation at 37^{o}C. After incubation, samples were briefly centrifuged at 13,000 g for 10 min and all liquid was discarded. Gel pieces were subsequently dried in a vacuum concentrator for approximately 15-30 min at 30^{o}C. In the next step, proteins were digested with trypsin with a trypsin/protein ratio 1:50 (w/w) (Applied Biosystems, USA) in 200 µL of 40 mM TEAB in 9 % (v/v) ACN for approximately 16 hours by incubation at 37^{o}C. At this stage, 0.1 % (v/v) RapiGest (protein solubilising agent) was added. After digestion, samples were centrifuged briefly at 13,000 g for 10 s and the supernatant was collected in a new Eppendorf tube. Peptides were extracted twice with 100 µL of 5 % (v/v) formic acid and once with 50 µL of 100 % ACN. Finally, all the supernatants were combined together and vacuum dried (Scanvac; module speen 40, Lynge, Denmark) and stored at -20^{o}C until further LC-MS/MS analysis.

3.2.3.2 In solution digestion

Simultaneously, another set with 100 µg of duplicate samples was transferred into new low-protein binding tubes. Protein samples in 0.5 M TEAB buffer (pH 8.5) containing 0.1 % (v/v) RapiGest were reduced with 1 µL of TCEP (50 mM) and samples were incubated at 60^{o}C for one hour. Subsequently samples were alkylated using 1 µL of 200 mM MMTS. In the next step, samples were digested

with a 1:50 (w/w) trypsin (Promega, UK) to protein ratio by overnight incubation for 16 hours at 37^{o}C. Digested peptides obtained by centrifugation at 13,000 g for 5 min and supernatant were transferred to a new Eppendorf tube and vacuum dried.

3.2.4 Peptide enrichment and fractionation

Peptide fractionation was carried out as described by Ow *et al.* [249] with a few modifications. Each sample from in-gel and in-solution digestion was fractionated using two different techniques SCX-HPLC and HILIC-HPLC. Enrichment and fractionation of complex digested samples were performed on an Agilent 1100 series HPLC (Agilent Berkshire UK), coupled to a 200mm PolyHYDROXYETHYL-A (5μm, 4.6mm ID, 200 Å, PolyLC) HILIC column and on a BioLC HPLC unit (Dionex, Surrey UK) coupled to a 200mm PolySULFOETHYL-A (5μm, 4.6mm ID, 200 Å, PolyLC) SCX column.

Samples were re-dissolved in respective transfer buffer A depending on the technique (100 μL of HILIC buffer A: 80 % ACN,10mM ammonium formate, pH 3; and 70 μL of SCX buffer A: 25 % ACN, 0.1 % formic acid), vortexed and spun down at 13000 g for 5 min. Supernatant was transferred to new Eppendorf tube and loaded to the respective columns. Enriched peptides either by hydrophilic or ionic interaction on respective columns were then eluted using solution buffer B (HILIC: 5 % ACN, 10mM ammonium formate, pH 5.0; SCX: 25 % ACN, 0.5 M KCl, 0.1 % formic acid) using a 95 min linear gradient at the flow rate of 500 μL min^{-1}. In both techniques, the linear gradient ramp was as follows: 0 % B (0-10 min), 20 % B (10-15 min), 20-60 % B (15-65 min), 60-100 % B (65-75 min), 100 % B (75-85 min), 0 % B (85-95 min).

Chromatograms were monitored using an Agilent 1100 ultraviolet detector module at 280 nm for HILIC and using UV detector UVD170U at 214 nm for SCX. Thirty five fractions were collected for each sample and vacuum dried. Fractions were collected every minute from 10-45 min for HILIC and 13-47 min for SCX respectively. To remove the high salts from SCX fractions, C18 clean up was carried out with collected fractions, whereas this additional step was not required for HILIC fractions [256] since buffers are compatible with MS. Finally, fractions were vacuum dried and kept at -20^{o}C until further analysis.

3.2.5 ESI mass spectrometry and identification of proteins

Peptides obtained from in-gel and in-solution digestion were re-suspended in reverse phase (RP) buffer (3 % acetonitrile and 0.1 % formic acid) and submitted to the electrospray ionisation-ion trap mass spectrometer HCT Ultra (Bruker, Daltonics, Bremen, GmbH, Germany) coupled with an online capillary liquid chromatography system (Famos, Switchos and ultimate from Dionex/LC Packings, Amsterdam, The Netherlands). Peptides were separated on a Pepmap C-18 RP capillary column (LC packagings) at a constant solvent flow rate of 0.3 µL/min. The elution of peptides was performed using a solvent gradient of buffer A (3 % ACN and 0.1 % FA) and buffer B (97 % ACN and 0.1 % FA): 5 min 97 % buffer A, buffer B was increased to 35 % in 38 min. This was then held at 90 % for 6 min and finally decreased to 3 % in 5 min. Data acquisition was set to the positive ion mode with a mass range of 300-2000 m/z. Tandem mass spectrometry was performed on peptides with +2, +3, and +4 charge states.

3.2.6 Peptide identification

Data obtained from MS analysis were converted to Mascot generic peaklist files (MGF) using Data Analysis software ver. 4.0 (Bruker Daltonics, Coventry, UK). Converted peaklists were then submitted to the in-house software Phenyx algorithm cluster (Binary version 2.6; Genebio Geneva) for peptide identification. The searches were performed against the UniProt database for *C. acetobutylicum* ATCC 824. Search parameters were set at MS tolerance of 0.4 Da and MS/MS tolerance of 0.4 Da. Peptide level filters were set to a z-score of 5.0, max *p*-value significance of 1.0E-5 and AC score of 5. Search space was also confined to trypsin peptides with a maximum of 1 missed cleavage.The overall workflow is shown in Figure 3-1.

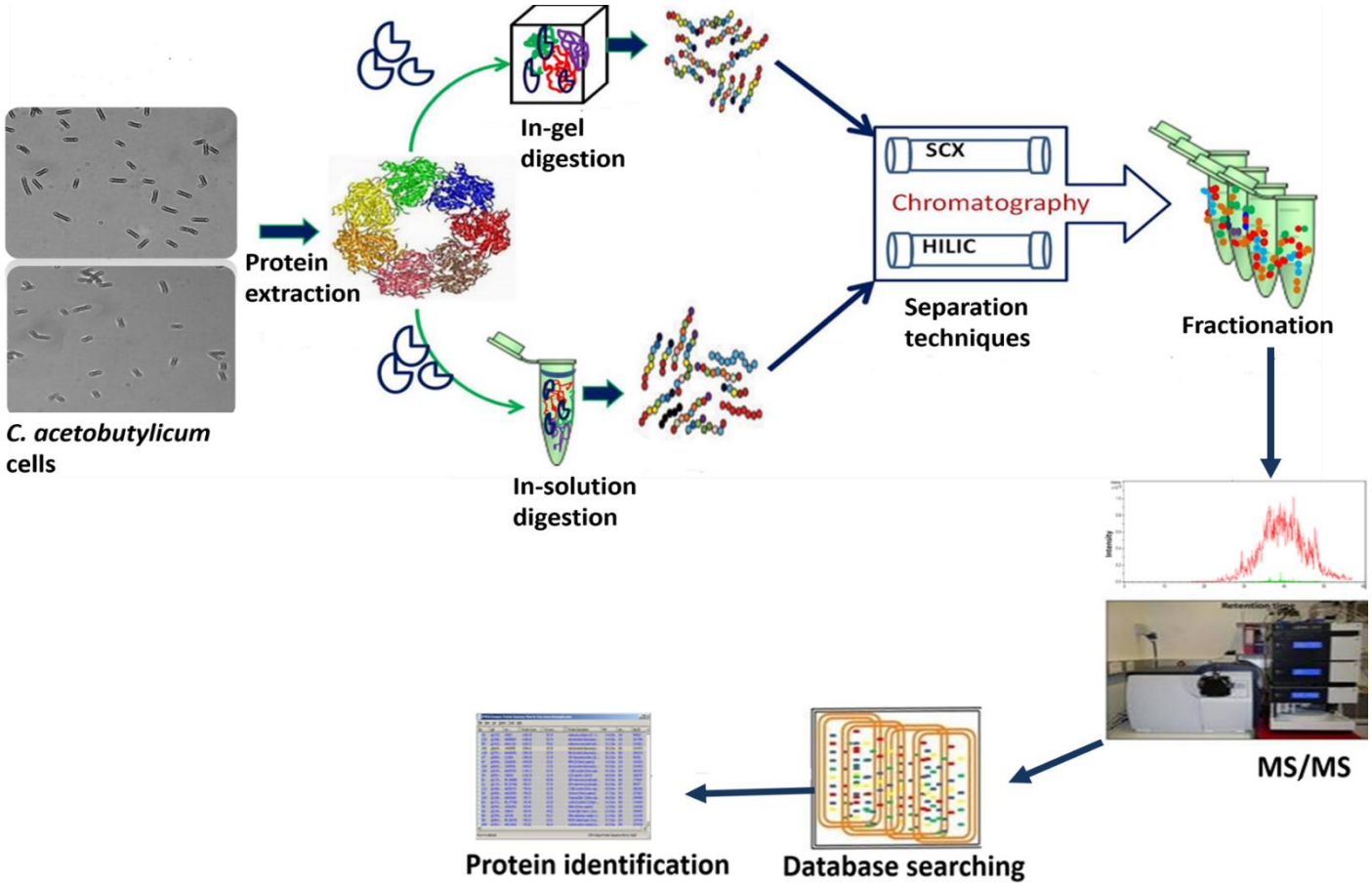

Figure 3-1 Shot-gun proteomics workflow. Protein lysate extracted and digested into peptides (in-gel and in-solution). Peptide complexity reduced by HILIC and SCX separation (fractionation) and then analysed by MS/MS where peptides are separated according to their m/z ratio to yield the precursor ion spectrum. Selected peptides are fragmented and assigned peptide sequences based on database against *C. acetobutylicum* and the proteins are identified.

3.3 Results and Discussions

Tryptic digestion of proteins and subsequent fractionation of peptides are two prerequisite steps in shotgun proteomics. The protein digestion usually achieved by in-gel or in-solution digestion. Two-dimentational LC separation of peptides reduces the spatial and temporal complexity of peptides resulting in increase in measurable peptide numbers and widening the overall dynamic range thus increase in proteome coverage [191].

To check the effect of well known digestion and separation techniques on peptide distribution and protein recovery, we analyzed in-gel and in-solution digestion and subsequent fractionation using SCX and HILIC and identified by mass spectrometry. HILIC separation is based on hydrophilic interaction between peptides and neutral hydrophilic stationary phase (hydrogen bonding), where peptides elute out with increasing hydrophilicity (opposite to reverse phase high performance liquid chromatography) [257, 258]. In SCX, on the other hand, the separation is based on net charges on peptides (ion exchange chromatography) [259]. The interactions of peptides with stationary phases are shown in Figure 3-2.

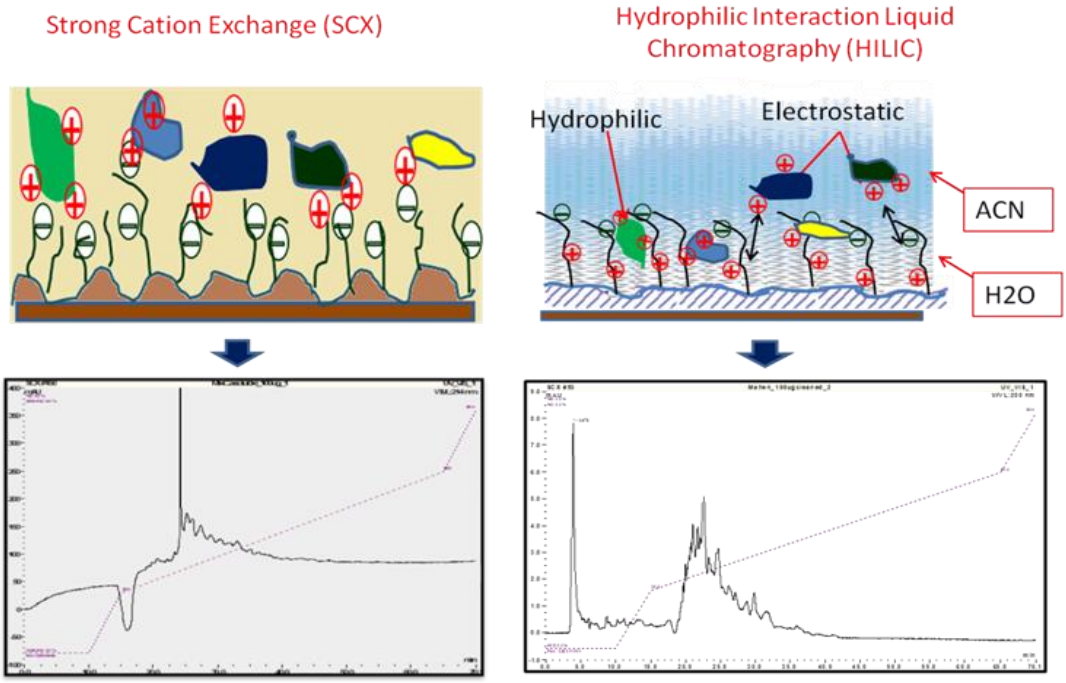

Figure 3-2 Chemistry behind HILIC and SCX separation of peptides

This report showed the effects of digestion and separation techniques on protein/peptide distribution, MS/MS identification, unique peptide distribution, peptide length, isoelectric point, precursor charges and hydrophobicity, and amino acid composition of peptides are discussed in the following sections.

3.3.1 Protein recovery using HILIC/SCX fractionation

In total, 431 proteins were identified and 105 were found common to all techniques. They were distributed in in-gel/HILIC, in-solution/HILIC, in-gel/SCX and in-solution/SCX as 296, 287, 211 and 208 respectively.

The protein recovery from HILIC separations showed a considerably higher number of proteins than SCX. The overlapping of the proteins among the treatments is given in Figures 3-3 & 3-4. In-gel/in-solution digestion and

subsequent separation by either HILIC or SCX did not show much difference in protein recovery in terms of numbers. However, when we compared in-gel/insolution/SCX with in-gel/in-solution/HILIC, unique proteins significantly varied from digestion to digestion and showed greater protein recovery via in-gel/in-solution/HILIC separation. Unique proteins in in-gel digestion via SCX and HILIC are 22, 65 and for in-solution digestion of SCX and HILIC are 13 and 47 respectively (Figure 3-3). The results show that the digestion steps have considerable impact on subsequent downstream processing proteins.

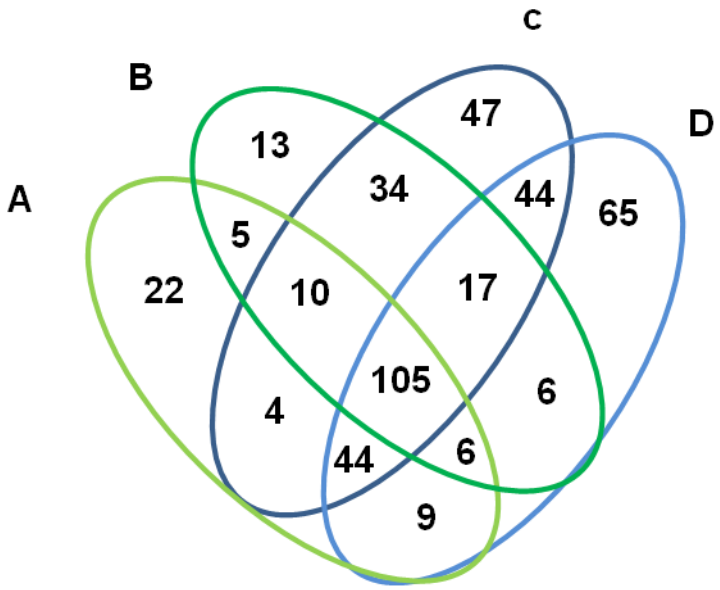

Figure 3-3 Venn diagram of identified proteins distribution; A) In-gel/SCX, B) In solution/SCX, C) In-solution/HILIC and D) In gel/HILIC.

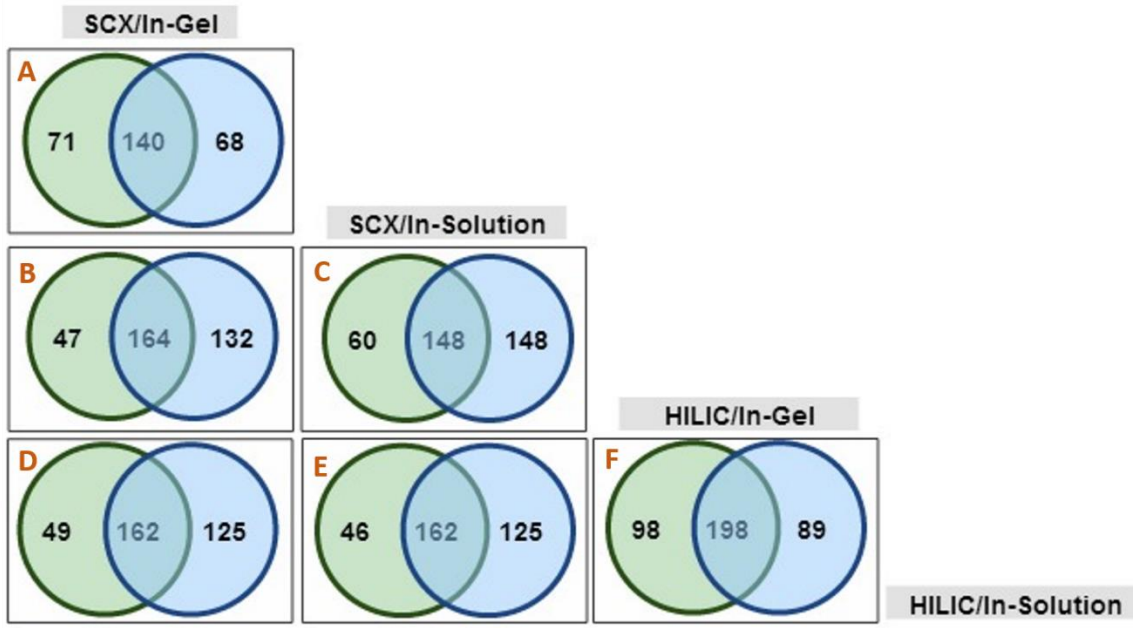

Figure 3-4 Distribution of identified proteins among the treatments. Comparing In-gel/In-solution/SCX vs In-gel/In-solution/HILIC. Data represents overlapping of proteins among the treatments. A) SCX/In-gel (green colour circle) vs SCX/In-solution (blue colour circle), B) SCX/In-gel vs HILIC/In-gel, C) SCX/In-solution vs HILIC/In-gel, D) SCX/In-gel vs HILIC/In-solution, E) SCX/In-solution vs HILIC/In-solution, F) HILIC/In-gel vs HILIC/In-solution.

We also observed differences in the isoelectric point (pI) and molecular weight (MW) of the identified proteins (Figure 3-5). Most of the identified unique proteins had MW in the range of 7.06-176.98, 5.44-76.206, 12.96-245.77, and 11.01-91.86 (kDa) for in-gel/HILIC, in-solution/HILIC, In-gel/SCX and in-solution/SCX respectively. Whereas pI ranged from 4.18-10.72, 4.21-12.31, 4.91-10.14 and 5.2-9.13 for in-gel/HILIC, in-solution/HILIC, in-gel/SCX and in-solution/SCX respectively. pI is the pH at which there is no net charge on bio-molecules which is one of the most important characteristics of the bio-molecules. Although, we didn't find much difference in pI and MW range of proteins among the treatment, but an average pI of unique proteins from in-solution digested samples was found to be higher than the in-gel digested samples. The reason behind this is unknown.

On the other hand, MW of proteins was comparatively higher in in-gel digestion than the in-solution digested samples of both SCX and HILIC separation. This indicates that in-gel digestion is perhaps most suitable for higher MW proteins. The efficiency of in-gel digestion for high MW proteins was previously noted [260, 261]. The observations may suggest that feasibility of the digestion techniques can be particularly important in the study of desirable proteins with known pI and MW.

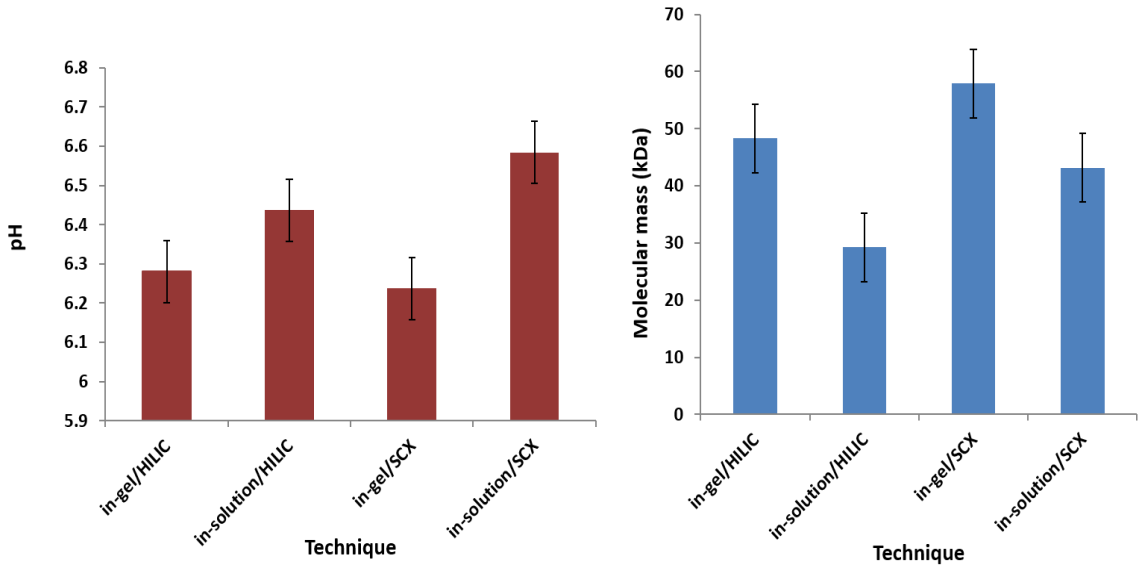

Figure 3-5 An average of molecular weight (MW) and pI of the unique proteins identified in different techniques.

Similarly, 740, 670, 938, and 884 peptides were identified in in-gel/SCX, in-solution/SCX, in-gel/HILIC and in-solution/HILIC techniques respectively. The number of peptides identified in in-gel/in-solution/HILIC, are more than in-gel/in-solution/SCX. In this case as well, we could not see much difference in digestion systems, but considerably higher numbers of peptides were obtained by HILIC separation than SCX. This indicates that either in-gel or in solution digestion and

subsequent fractionation by HILIC has much better performance when compared to SCX in terms of peptides. Presumably this was because the peptide separation was based on different properties and principles, and the number and types of peptides varies from technique to technique. The unique peptide distribution and the overlap among techniques is shown in Figure3-6.

It is assumed that in-solution digestion is more efficient than the in-gel digestion [262]. However, in this study, the number of proteins/peptides obtained from in-gel digestion was considerably higher than proteins/peptides obtained from in-solution digestion following both separation methods, indicating that the efficiency of in-gel digestion cannot be underestimated.

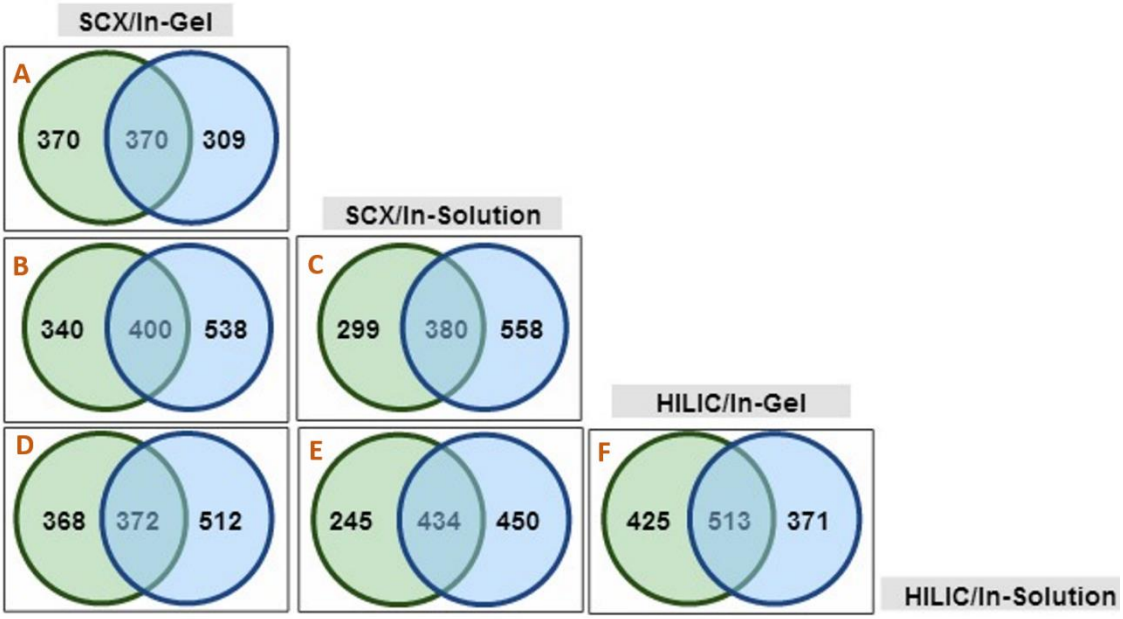

Figure 3-6 Peptide overlapping among the techniques examined. Comparing In-gel/In-solution/SCX vs In-gel/In-solution/HILIC. A) SCX/In-gel (green colour circle) vs SCX/In-solution (blue colour circle), B) SCX/In-gel vs HILIC/In-gel, C) SCX/In-solution vs HILIC/In-gel, D) SCX/In-gel vs HILIC/In-solution, E) SCX/In-solution vs HILIC/In-solution, F) HILIC/In-gel vs HILIC/In-solution.

3.3.2 Separation efficiency of SCX and HILIC

3.3.2.1 Total MS/MS and unique peptide distribution

The separation efficiency of peptides by SCX and HILIC were assessed by measuring the total number of MS/MS and unique peptides across the 35 fractions (Figure 3-7). The number of MS/MS and unique peptides were analysed using (Mathematica v9, Wolfram). The number of MS/MS and unique peptides were found to be significantly higher in in-gel/in-solution/HILIC as compared to in-gel/in-solution/SCX. MS/MS and unique peptides distribution for the HILIC separation workflow were more evenly across the fractions than in the SCX workflow. It seems that in in-gel/in-solution/SCX fractionation, peptides tended to elute in clusters resulting in the accumulation of peptides in only a few fractions. The clustering of peptides with similar charges has been observed previously in SCX workflows [263, 264]. Compared to SCX, more peptide were identified in more fractions of HILIC fractionated samples. That reflects the ability of HILIC to separate peptides more uniformly throughout the gradient, thereby further reducing the complexity of the samples. The advances of HILIC over SCX have been recently reviewed by Boersema *et al.* [197].

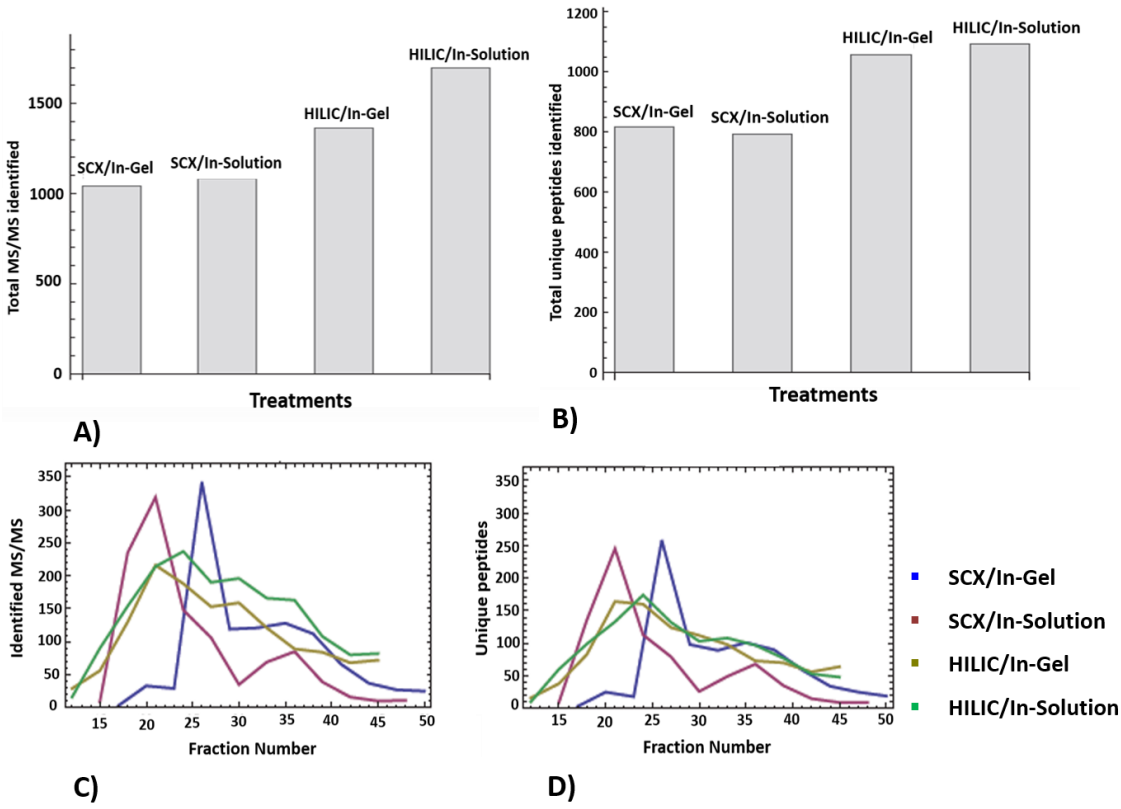

Figure 3-7 Peptide analysis. Number of MS/MS and unique peptides identified in in-gel/insolution/SCX/HILIC. A) & C) Total MS/MS, B) & D) Total unique peptides.

3.3.2.2 The pI and Hydrophobicity distribution of peptides

Physico-chemical properties of proteins/peptides play a crucial role in the separation of proteins/peptides from complex biological mixture and thereby helps us to design proteomics experiments for further identification [265]. The main parameters include, precursor charge, isoelectric point (pI), hydrophobicity, and molecular weight (MW). Precursor charges distribution of peptides are shown in Figure 3.8. The charge state of peptides obtained from tryptic digestion is likely to be +2 to +4 depending on number of basic residues in the peptide [266]. Tryptic peptides usually have a charge of +2 due to the basic N-terminus and lysine and arginine at the C-terminus. Most of the peptides identified in this study showed

+2 precursor charges, indicating that these were peptides obtained from trypsin digestion (Figure 3-8). We also observed +3 precursor charges in all cases with comparatively higher values for in In-gel/HILIC workflows. However, there were no significant differences observed among the techniques.

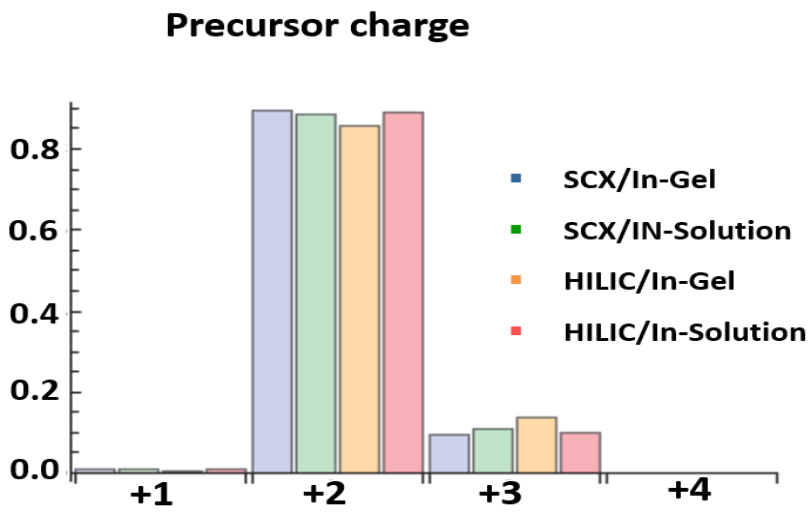

Figure 3-8 Peptide analysis. Precursor charges distribution among the treatments

In SCX (in-gel, in –solution) experiments as well as HILIC (in-gel, in-solution) experiments, the pI values seemed to spread wider and to increase on average as fractionation number increased (Figure 3-9). However, in both treatments of SCX, the pI value showed a gradual and uniform increase with the progression of fraction numbers. In SCX separation, the pI increases with increase in retention time [267]. However, no trend was observed in the case of HILIC (in-gel, in-solution).

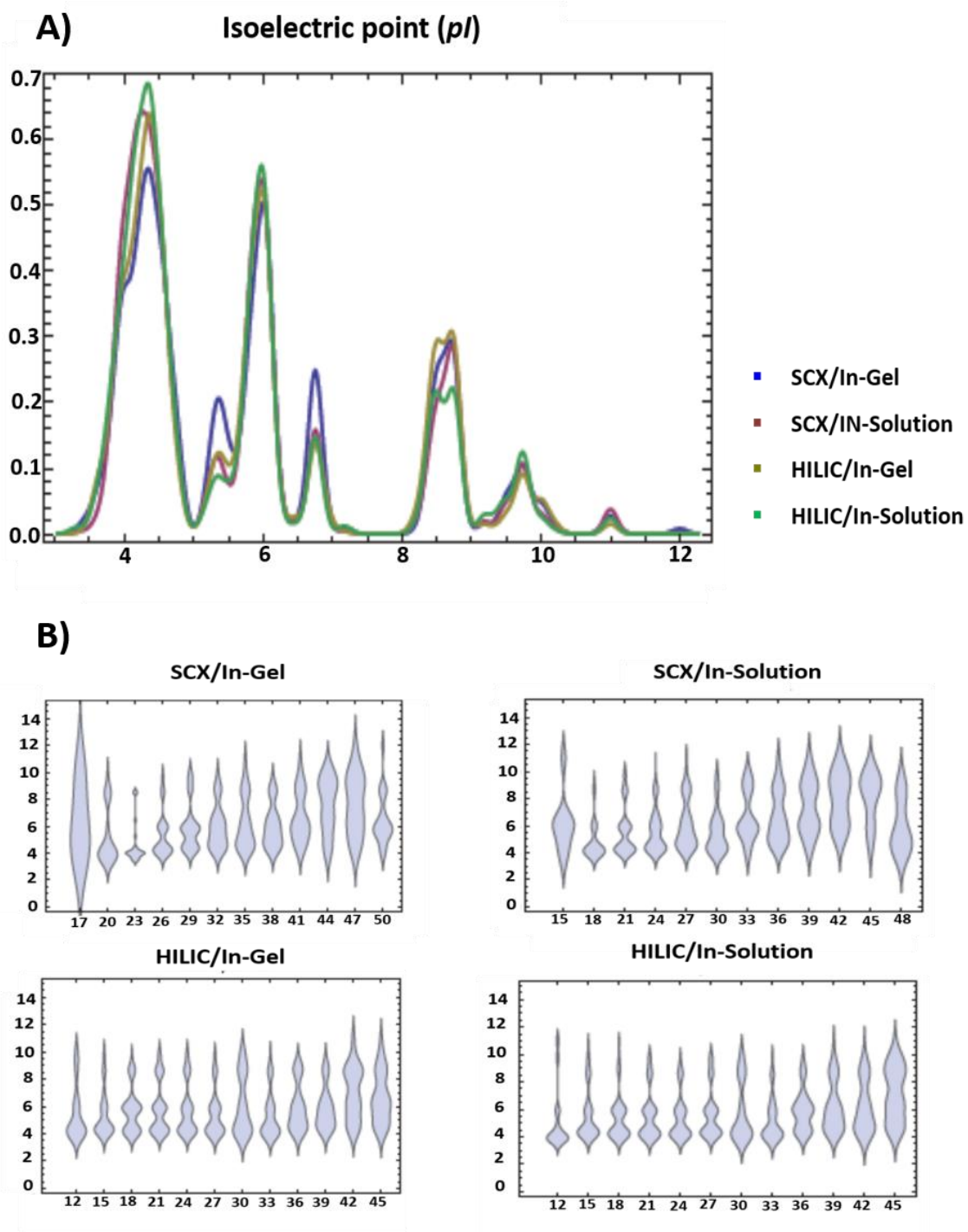

Figure 3-9 Peptide analysis. A) Distribution of unique peptides identified according to pI values in each fractions. B) Distribution of unique peptides based on their pI in each technique.

The average value of hydrophobicity, on the other hand, showed gradual decreases in late fractions in both the SCX and HILIC (Figure 3-10). The trend is obvious for HILIC separation because retention time increases with increase in hydrophilicity [268] but not for SCX because the separation is based on pI. However, a gradual decrease in hydrophobicity was more uniform in HILIC workflows as compared to SCX workflows.

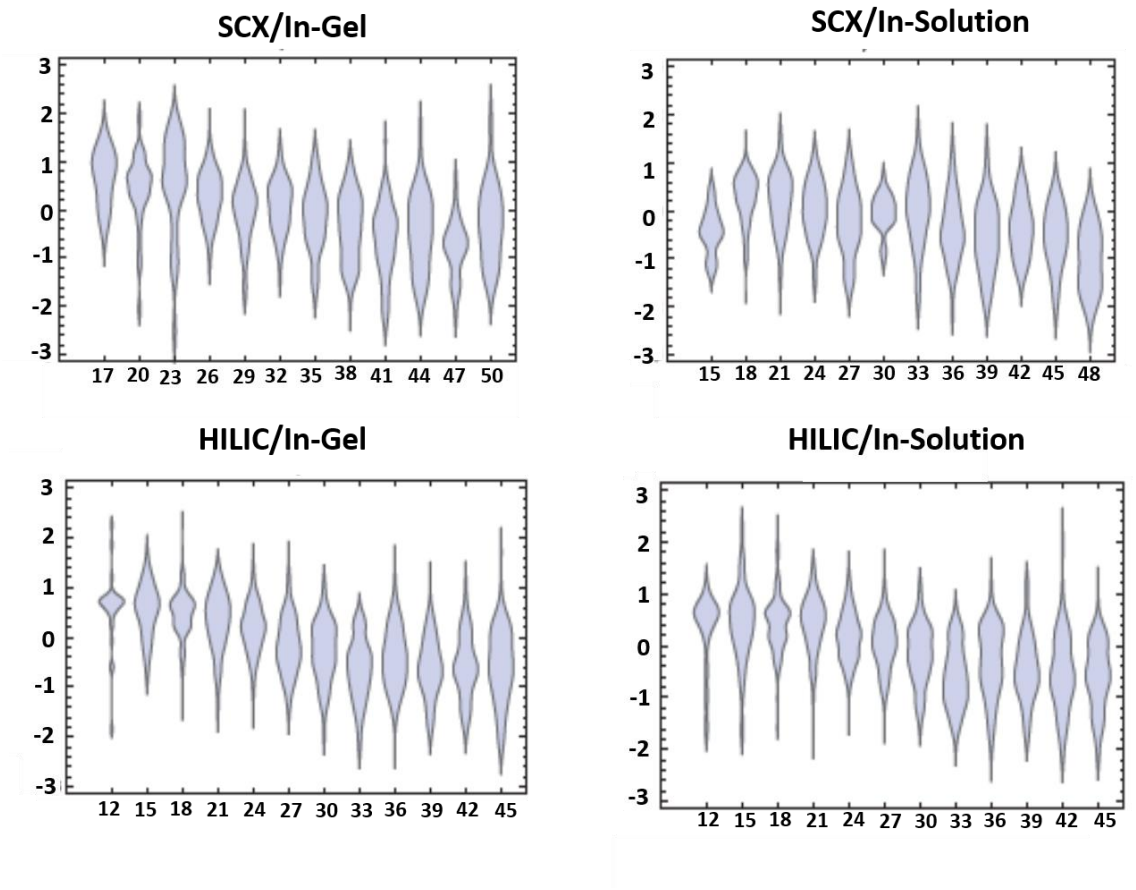

Figure 3-10 Peptide analysis. Distribution of unique peptides according to hydrophobicity in each fractions

In SCX, because the majority of peptides are doubly and triply charged they will elute as compact clusters of ions, therefore little separation is achieved. On the other hand, HILIC has a different distribution pattern, peptides are more efficiently distributed across all fractions, therefore a better orthogonality is achieved. In terms of identification and or quantification, HILIC will provide better results

because better separation will generate a more uniform distribution of peptides per fraction. SCX will generate a very compact distribution of all peptides in a narrow window of the elution profile. Each sample will be highly complex and less information is obtained by MS. Quantification will be also compromised due to more interferences from close ions during MS and MS/MS isolation.

Our results suggested that the two well known methods of tryptic digestion are certainly affected by pI and the molecular weight of the proteins (Figure 3-5). In-gel digestion can be the better option for proteins with high molecular weight and In-solution digestion can be better technique for the proteins with high pI value. For 1-D separation of peptides, SCX is the current choice in proteomics. However, total MS/MS and unique peptide distribution (Figure 3-7) across the fractionation in HILIC, indictates that HILIC has better uniform distribution thus has highest degree of orthogonality to RP than SCX separation. The narrow window of fractionation of petides by SCX (Figure 3-4 & Figure 3-6) resulting in lower protein/peptides coverage was observed in this study. The distribution of peptides separation from SCX is purely based on solution charges therefore bulk of peptides with charge +2 and +3 were eluted in early fractions (fractions 15-30). The gradual and uniform decrease in the hydrophobicity (Figure 3-10) of the peptides further indicated superiority of the HILIC fractionation over SCX. Especially in-gel digestion/HILIC separation showed better separation than the all other comparisions.

3.4 Conclusion

Unlike some other omics techniques such as transcriptomics, proteomics is more complex and dynamic because it changes from cell to cell and time to time. That forces researchers to try to optimise proteomics techniques to achieve a maximum protein recovery. The improvement in protein digestion and the technique employed for the separation of peptides can greatly contribute to the overall development of proteomics. Here, we compared HILIC to SCX separation and in-gel vs in-solution digestion workflows.

From our results, in-gel digestion and subsequent HILIC separation appear to be the most useful technique. However, our results also reflect the advantages and disadvantages of the techniques in terms of unique protein/peptide loss. Therefore, based on the type of biological protein samples, suitable chromatographic techniques or combinations of techniques selection can be used.

Chapter 4

Influence of lignin on *Clostridium acetobutylicum* ATCC 824 Proteome: a quantitative proteomics analysis

Abstract

Clostridial species such as *Clostridium acetobutylicum* are known for their ability to produce biofuels from a diverse range of sugar components derived from lignocellulosic biomass. The influence of the lignin component on biomass fermentation is poorly studied in *Clostridium acetobutylicum*. In this study, *C. acetobutylicum* was grown in medium containing either cellobiose only or cellobiose plus lignin in order to investigate the influence of lignin on metabolic behaviour and on the efficiency of biofuel production. The metabolic perturbations were analysed using isobaric tags for relative and absolute quantitation (iTRAQ) and liquid chromatography tandem mass spectrometry. We quantified 579 proteins comprising eighteen cellular functional groups. Comparing cellobiose and cellobiose plus lignin conditions, differential expression affected carbohydrate metabolism, amino acid metabolism, the TCA cycle, energy metabolism, the cell cycle/cell division processes, signal transduction/chemotaxis, stress response, transcription and translation. Although enzymes directly involved in the degradation of lignin could not be detected, our study provides valuable insights into the metabolic flux response of *C. acetobutylicum* at the proteomic level. This can pave the way for the metabolic optimisation of biofuel production in these organisms.

4.1 Introduction

Due to growing uncertainties regarding the supply and cost of transportation fuels and concerns about their related environmental impact, the sustainable production of clean energy has become a strategic priority. The biological degradation of lignocellulosic biomass, which is a key step in global carbon cycling, has great potential as a prime feedstock for future biofuel generation [5]. For that reason,anaerobic *Clostridia* have received much attention in recent years because of their ability to produce alternative biofuels from renewable biomass and agricultural waste material [150]. Thus, extensive studies have been carried out using *Clostridia* for direct solvent production using biomass such as hardwood [269], domestic organic waste [270], starch based waste packing peanuts [271], agriculture waste [150], corn fibres [272], palm oil waste [273], sludge waste [274], whey [275] and sago starch [276].

Particularly, *C. acetobutylicum* is a most promising candidate for future biofuel generation since it can ferment a wide range of biomass sugars into acetone, butanol, ethanol (ABE) [277] and biohydrogen [278]. However, major bottlenecks still hamper the economics of the ABE production from biomass such as the inhibitory effect of many compounds [277], including lignin degradation by-products [279], produced during fermentation. The inhibitory effects of phenolic compounds and furfural on ABE fermentation have also been noticed in *Clostridial* species [277, 280-283]. Thus, research efforts are still needed to understand the biology of this solvent-producing bacterium to achieve economically-viable yields.

Lignin is the second most abundant aromatic backbone of lignocellulosic biomass after cellulose. Lignin cannot be fermented, but its significant energy content can be nonetheless used to enhance fermentation processes [284]. The anaerobic degradation of aromatic compounds is a relatively new concept, and over the past decades, several metabolic pathways involved in anaerobic degradation of aromatic compound have been elucidated and are reviewed elsewhere [285]. A few studies have shown that *Clostridial* species are involved in aromatic compound degradation and adopt different degradation strategies to do so [286-289]. 2,4,6 trinitrotoluene (TNT) degradation is extensively studied in *C. acetobutylicum* [290-295] and it was shown that *C. acetobutylicum* might possess a unique metabolic pathway for the degradation of aromatic compounds, including lignin.

Currently, a difficulty arises because of the inhibitory effect of lignin on the metabolic processes in *C. acetobutylicum*. To understand this effect and to alleviate it, we sought to investigate the metabolic response of *C. acetobutylicum* ATCC 824 to the presence of lignin using iTRAQ-based quantitative proteomics and metabolite analysis approaches.

Therefore, based on the preliminary metabolite analysis experiments, we conducted iTRAQ based quantitative proteomics with two substrate conditions: Cellobiose and Cellobiose plus lignin. Our results reveal the first high throughput investigation of *C. acetobutylicum* in presence of lignin and describe the changes that occur during exponential and stationary phases.

4.2 Experimental methods

4.2.1 Bacterial strain and growth conditions

All chemicals and reagents were purchased from Sigma-Aldrich (UK) unless otherwise specified. *C. acetobutylicum* ATCC 824 was procured from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). *C. acetobutylicum* was maintained anaerobically on medium as previously described by Lopez-Contreras *et al*. [270]. Briefly, the growth media contained 0.75 g L^{-1} KH_2PO_4, 0.75 g L^{-1} K_2HPO_4, 0.348 g L^{-1} $MgSO_4$, 0.01 g L^{-1} $MnSO_4.H_2O$, 0.01 g L^{-1} $FeSO_4.7H_2O$, 1 g L^{-1} NaCl, 1.0 g L^{-1} cysteine chloride, 5 g L^{-1} yeast extract, 2 g L^{-1} $(NH_4)_2SO_4$. The cellobiose-only growth medium contains 5 g L^{-1} cellobiose (hereafter denoted as C) whereas the cellobiose/lignin medium contains 5 g L^{-1} cellobiose plus 1 g L^{-1} lignin (alkali carboxylated) (hereafter denoted as CL). Media were prepared anaerobically in the presence of 100 % N_2 in 125 mL serum bottles and autoclaved (as described in Chapter 3 section 3.2.1). The culture media were seeded with 1 mL of 18-hour-long cultures (corresponding late-log phase, OD at 600_{nm} equal to 1.3), grown in media containing cellobiose as carbon source. Cultures were incubated at 38^{o}C and growth curves were monitored at OD_{600nm} using an Ultraspec spectrophotometer (Model 2100 pro, Amersham Bioscience). Cellobiose concentration was estimated by the Anthrone method [296].

4.2.2 Fermentation by-products analysis

Fermentation products were identified and quantified as previously reported Pham *et al.* [234]. Briefly, ethanol, butanol, acetic acid and butyric acid were detected and quantified using a Finnigan Trace DSQ single Quadrupole GC-MS coupled with an auto-sampler model AS3000 (Thermo Electron Corporation, USA) and a 30 m × 0.25 mm ID × 0.25 µm df Stabbilwax fused silica column (Thames Restek, UK). Approximately 50 µL aliquots were extracted, centrifuged at 17,000 g for 2 min and transferred to a MS vial, and then 1µL of sample was withdrawn for GC-MS analysis. The total GC-MS analysis running time was 14 min and temperature gradient was performed with a hold at 45^{o}C for 3 min, followed by a ramp at a rate of 15^{o}C/min to 120^{o}C, then 30^{o}C/min to 210^{o}C and finally a hold for 1 min at 210^{o}C. Helium was used as the carrier gas at a flow rate of 1 mL/min, MS transfer line and injection port temperature were set at 250^{o}C and 210^{o}C respectively. The MS detector gain was used at 1x10^{5} eV with a full scan (positive polarity mode) ranging from 20 to 65 m/z, while electron ionization temperature was operated at 230^{o}C.

4.2.3 Cell harvest and proteome extraction

Cells grown in C or CL media were harvested at mid-exponential (16 h) and late stationary phase (48 h) by centrifugation at 10,000 g for 5 min at 4^{o}C. Harvested cell pellets were washed twice with phosphate buffer saline (PBS) at pH 7.4 and finally once with the protein extraction buffer (0.5 M triethyl ammonium bicarbonate (TEAB), pH 8.5). The cells were re-suspended in 600 µL extraction buffer (TEAB pH 8.5 containing 0.095 % SDS) and 300 mg sterilized glass beads

(425-600µm) were also added in a 1.5 mL eppendorf tube (Eppendorf, Cambridge UK). At this stage, 5 µL of protease inhibitor cocktail set II was added. Proteins were extracted using a disruptor (Disruptor Genie, Scientific Industries Ltd, USA) with 20 cycles alternatively 1 min vortexing and 1 min incubating on ice. Unbroken cells and cells debris were discarded firstly by centrifugation at 3000 g for 5 min. Five microlitres of benzonase® nuclease (1:100) was added to supernatant containing soluble proteins to degrade DNA and RNA from the sample. Subsequently, the supernatant was further centrifuged at 21,000 g for 90 min at 4^{o}C. The supernatant was collected and proteins were precipitated overnight (approximately 16 h) using acetone at -20^{o}C (ratio sample: acetone = 1:4 v/v). Precipitated proteins were recovered by centrifugation at 17,000 g for 20 min at -9^{o}C and air dried. Finally, the pellet was dissolved in 0.5 M TEAB (pH 8.5) buffer containing 0.1 % RapiGest SF (Waters, Milford, MA). The total protein concentration was quantified by the Bradford assay according to the manufacturer's protocol (Sigma Aldrich).

4.2.4 iTRAQ Labelling and LC-MS/MS analysis of proteome

iTRAQ 8-plex labelling was performed for samples according to the manufacturer's protocol (8-plex iTRAQ reagent Multiplex kit, ABSciex, USA). Briefly, 100 µg of proteins from each sample was firstly reduced with 1 µL of 50 mM Tris (2-carboxyethyl) phosphine hydrochloride (TCEP) and incubated at 60^{o}C for one hour. Samples were then alkylated using 1 µL of 200 mM methyl methanethiosulfonate (MMTS) at room temperature for 10 min. Subsequently, proteins were digested with trypsin with a ratio of 1:50 (trypsin:proteins)

(Promega, UK) overnight at 37ºC. A biological duplicate was used for each phenotype.

Each biological phenotype was labelled with relevant iTRAQ reagents as shown in Figure 4-1**.** Labelled samples were acidified with TFA (final concentration of solution 0.5 %) to precipitate out RapiGest and the supernatant was obtained by centrifugation at 17,000 g for 5 min at 4ºC and dried using a vacuum concentrator (Scanvac; module speen 40, Lynge, Denmark).

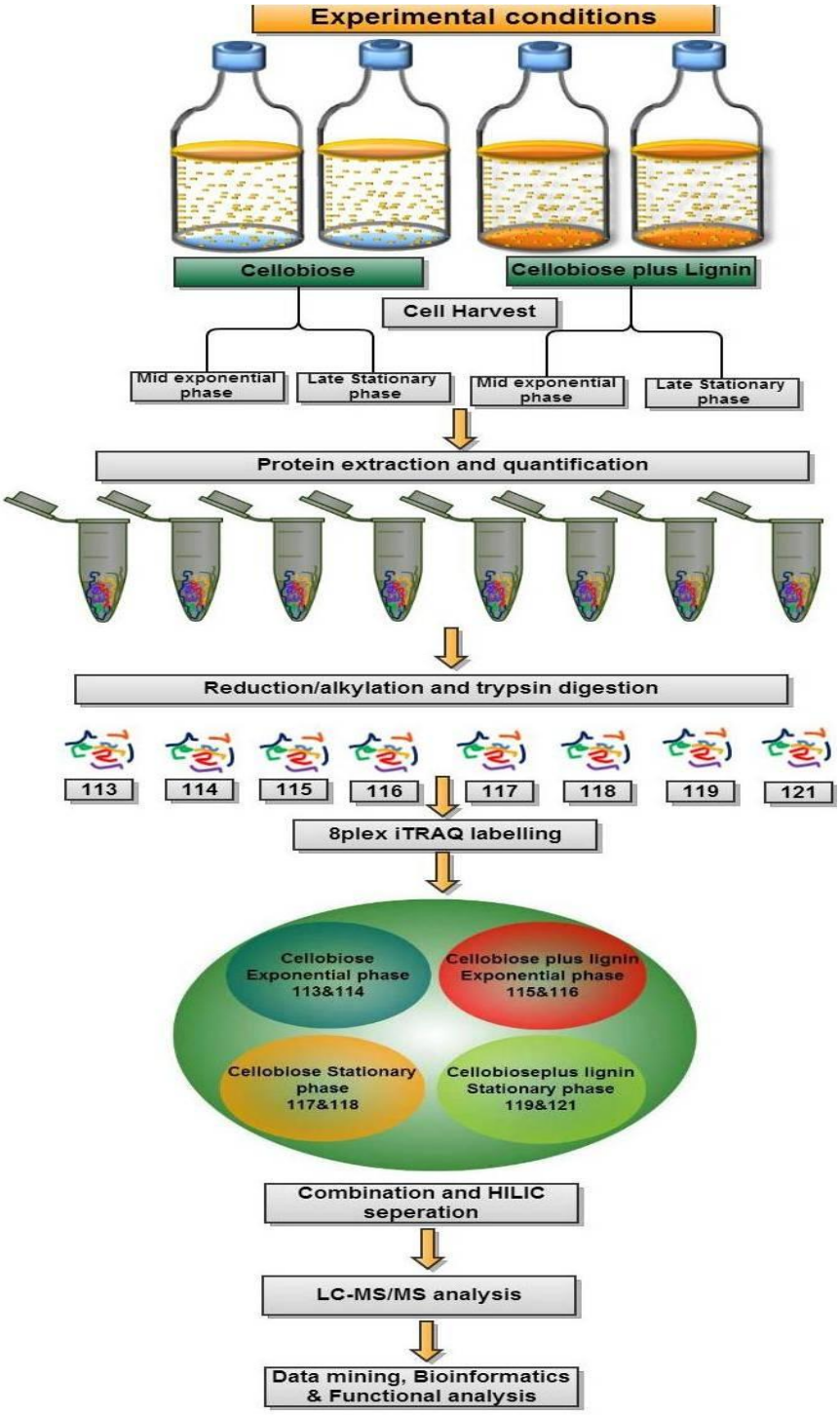

Figure 4-1 Quantitative proteomics workflow used in the current experiment (2 independent biological replicates per condition)

4.2.5 HILIC fractionation of peptides

Dried iTRAQ-labelled peptides were re-suspended in 100 µL of HILIC buffer A (80 % acetonitrile (ACN), 10 mM ammonium formate, pH 3 for fractionation. The sample was centrifuged at 17,000 g for 5 min and injected into an Agilent 1100 series HPLC(Agilent Berkshire, UK) coupled with a HILIC column (PolyHYDROXYETHYL-A, column, 5 µm pore size, 100 mm length, 4.6 mm ID, PolyLC Columbia, MD, USA) at a flow rate of 0.5 mL/min using an UV detector at 280nm (Dionex, UVD170U). The following buffers were used: buffer A (80 % ACN, 10 mM ammonium formate, pH 3) and buffer B (5 % ACN, 10 mM ammonium formate, pH 5). The 90 min gradient consisted of 0 % Bfor 10 min, 0-20 % B for 15min, 20-40 % for 30 min, 40-60 % for 15 min, 60-100 % B for 5 min, and 100 % A for 15 min. Fractionation and chromatogram was monitored through Chromeleon software (Dionex/LC packing Netherlands). Collected fractions were then dried in a vacuum concetrator (Genevac Ltd Suffolk, UK) (Brand name, country) and subjected to further LC-MS analysis. Further C_{18} desalting of samples was skipped for HILIC samples [256].

4.2.6 Mass Spectrometry analysis

The selected fractions were re-dissolved in 22 µL Reverse phase (RP) buffer A (3 % ACN, 0.1 % FA) for further LC-MS/MS analysis before submitting to a QStar XL Hybrid ESI Quadrupole time-of-flight Tandem mass spectrometer (ABSciex, Concord, Ontario Canada) coupled with an online nano high performance liquid chromatography HPLC system (Ultimate 3000, Dionex, Surrey UK). 10 µL of each selected fraction was injected into the nanoLC-ESI-MS/MS system, peptide

separation was performed by an Acclaim® PepMap100 column (C-18, 3 µm, 100Å, 15 cm) at a constant flow rate of 300 nL/min. The buffers used in the liquid chromatograph were RP buffer A (3 % ACN with 0.1 % FA), and RP buffer B (97 %ACN with 0.1 %FA) and the 120 min gradient was used as follows: 0-3 % B for 5 min, 3-35 % B for 90 min, 35-90 % of B for 0.5 min, 90 % of B for 6.5 min, finally 3 % of buffer Bfor 18 min. The MS detector was set to scan 350-1800 m/z in the positive mode using Analyst® QS 2.0 software (Applied Biosystems, USA) and data were acquired in the data-dependent acquisition mode. Peptides of charge +2, +3, +4 (intensity binning) for each TOF-MS scan (400-1250 *m/z*) were dynamically selected and isolated for MS/MS fragment ion scans (100-1600 *m/z*). Two RP-HPLC-MS runs were performed.

4.2.7 Data interpretation and protein identification

The generated tandem MS data files (wiff) from the QSTAR XL were converted into generic MGF format *via* the mascot.dll embedded script (V1.6) coupled with Analyst QS v. 1.1.1 (Applied Biosystems). Peptide identification was performed using an in-house Phenyx algorithm cluster (Binary version 2.6; Genebio Geneva). The database search was performed within the *C. acetobutylicum* ATCC 824 (taxon ID: 272562) database containing 3825 protein sequences downloaded from Uniprot (June 2012). Searches were conducted against a forward/reverse concatenated database to determine the false positive discovery rate (FDR). Protein identifications were accepted as positive based on a probability filter cut-off of 95 %. Mass tolerances for peptide identification were set to 0.6 Da and 0.1 Da for MS and MS/MS respectively. Peptide level filters were set to a minimal *Z*-score of 5.0, a maximal *p*-value of 10^{-4}, and an AC score of 5. Trypsin was used with a maximum of 1 missed cleavage site. The

modifications were performed as follows: eight-plex iTRAQ mass shifts (+304 Da) of lysin (K) and N terminus as fixed modification, cys CAM (+57 Da) as fixed modification on the cysteine (C) residue and oxidation of metheonine (M) (+16 Da) as a variable modification on the M residue. The reporter ions' intensities were then exported to Microsoft Excel for further analysis. Isotopic and median corrections were applied using an in-house automated method as described in Ow *et al*. [212] and proteinquantification values were obtained in log space. Further, these data were analysed using a method described by Pham *et al.* [297] with significant changes ($\alpha = 0.05$).

4.3 Results and discussion

4.3.1 Cell growth, cellobiose consumption and extracellular metabolites production

The OD curves of cells grown on C and CL followed a similar trend up to the late exponential phase, at which point the two curves significantly diverged. The maximal density (OD_{600nm}) of cells grown on CL was 3.5 approximately at 31 h post inoculation compared to 2.4 at 24 h for cells grown on C (Figure 4-2). The cellobiose consumption profile showed a comparatively slow start of utilisation of cellobiose in the presence of lignin (Figure 4-2). Abnormal morphology of the bacterium was observed, with asymmetric and filamentous phenotypic cell division in the presence of lignin as shown in Figure 4-3. This could be the reason for the difference in growth density pattern between C and CL grown cells. The presence of lignin in the growth medium induced a negative impact on extracellular metabolite production (acetate, butyrate, ethanol and butanol)

(Figure 4-4). Concentrations of these products (acetate, butyrate ethanol, and butanol) were lower in the presence of lignin compared to that in the C (control) condition.

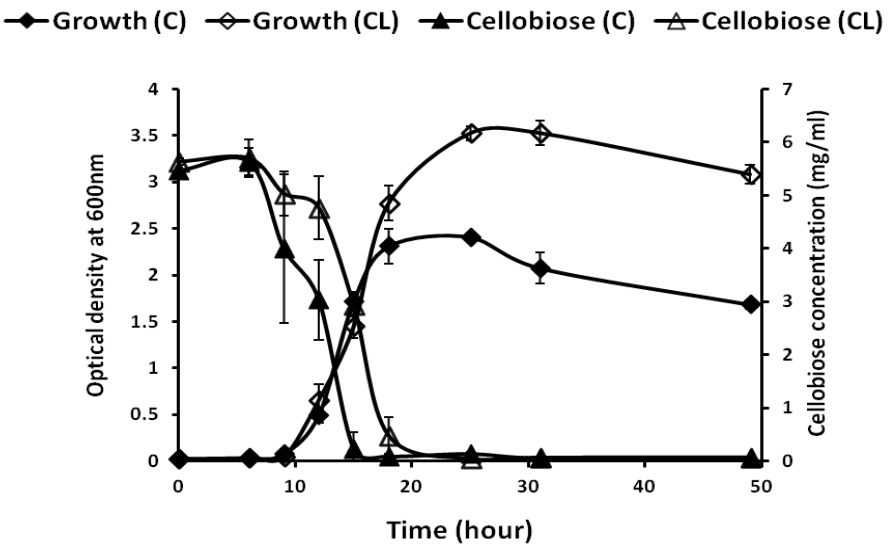

Figure 4-2 Growth and cellobiose concentration profiles of *C. acetobutylicum* grown on C and CL. Data were taken from biological triplicates.

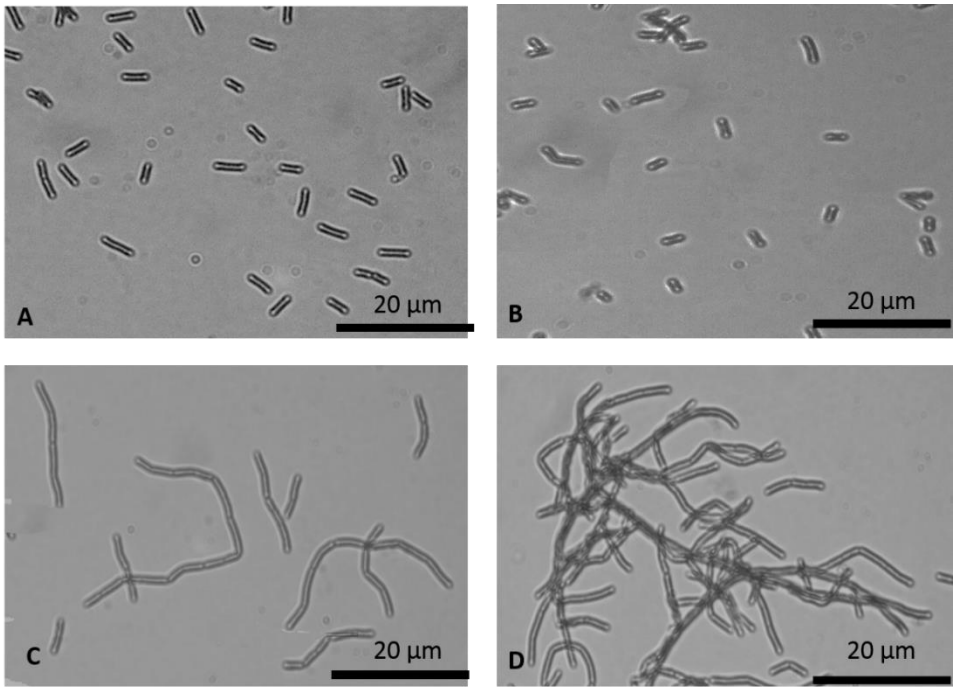

Figure 4-3 Morphological changes of cells grown on C (normal cells) (A&B) and CL (filamentous phenotype) (C&D) at exponential (14 h) and stationary phase (48 h) at 100X magnification.

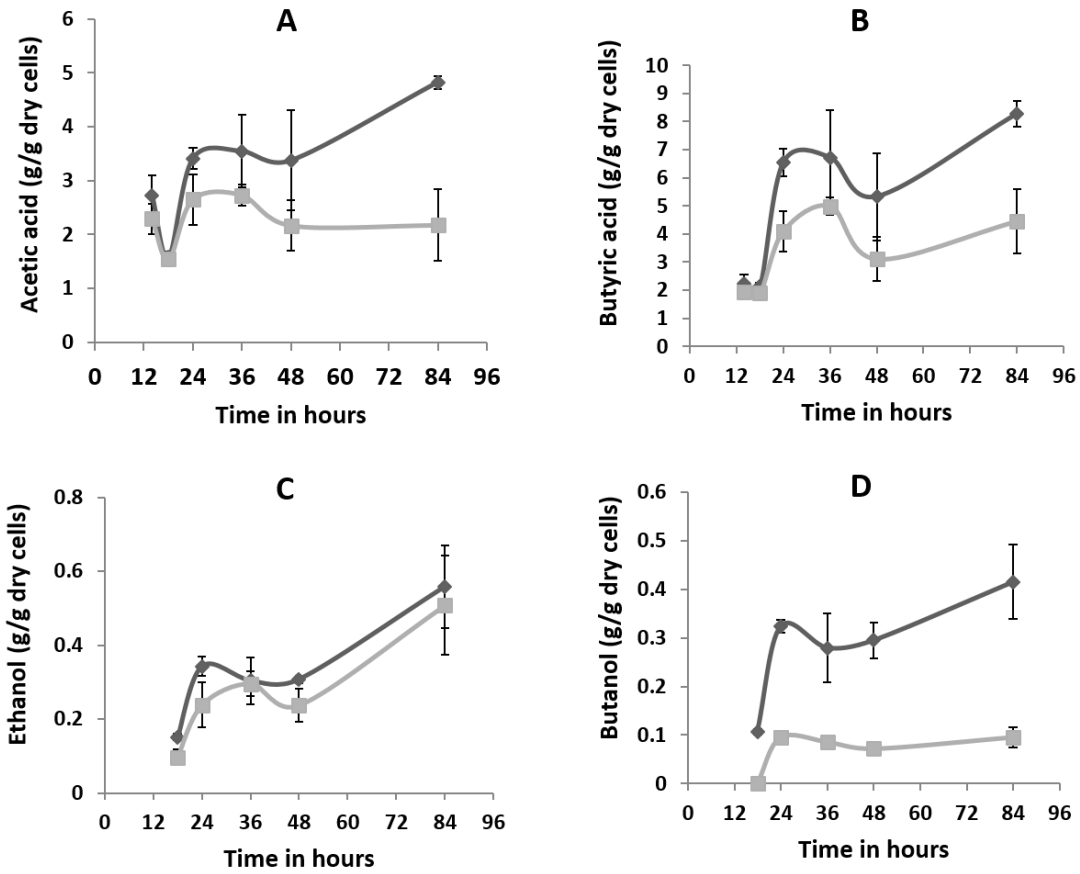

Figure 4-4 Fermentation products of *C. acetobutylicum* under C (—◆—) and CL (—■—) conditions: (A) acetic acid, (B) butyric acid, (C) ethanol, and (D) butanol. Data were taken from biological triplicates.

The results presented in this study shows that, in presence of lignin, the concentration of metabolites (acetic acid, butyric acid, ethanol and butanol) were significantly reduced. The concentration of acetic acid reached a maximum of 3.5 gm/gm dry cells at 36 hour for C and 2.7 gm/gm dry cells for CL. At the same time, production of butyric acid, reached 6.7 and 4.98 gm/gm of dry cells at 36 hours for the C and CL condition respectively. Our results show simultaneous production of ethanol and butanol. Onset of ethanol and butanol production started at 18 hours for both C and CL conditions, which is early stationary phase.

Ethanol and butanol production was relatively low in CL conditions. As the concentration of acetic acid and butyric acid ceased, the production of ethanol and butanol increased in both conditions and reached a maximum ethanol concentration of 0.55 gm/gm of dry cells in C and 0.50 gm/gm of dry cells in CL conditions at 84 hours. Butanol production reached maximum level to 0.41 mg/mg of dry cells in C and 0.07 mg/mg of dry cells in CL. Our results showed lower production of acids that subsequently reduced solvent production in lignin condition. The solvent production kick starts as soon as culture enters late exponential phase/early stationary phase (metabolic shift) [298]. There was initial decrease in acids production on the onset of solventogensis (Figure 4-4A, B). However, interestingly, the trend showed that acids production was observed during solventogenesis. Previous studies have noticed that simultaneous production of acids and solvents during solventogenesis can be possible and suggested that acidogenesic and solventogenic cells may coexist in the culture [150, 298-302]. Metabolic changes are quite sensitive to pH and the concentration of metabolites present in the culture broth. During solvengenesis there is shift in pH due to utilisation of a portion of acids for solvent production that may allow cells to reboot the production of acids during solventogenesis.

4.3 2 Quantitative proteomics in response to lignin

Since lignin had a significant effect on growth, morphology and extracellular metabolite production, one could expect alterations due to lignin on *C. acetabutylicum*'s metabolism. We sought to identify such alterations using high-throughput proteomics. Out of 7126 peptides, a total of 3062 unique peptides were identified. These mapped to 653 proteins from which 579 proteins were

quantified with >=2 peptides at 2% False discovery rate (FDR) [303]. These proteins were assigned to 21 functional categories according to their cellular functions as depicted in Figure 4-5 (http://www.uniprot.org/). Table 4-1 indicates numbers of significant up/down regulated proteins among the functional categories. We have compared exponential (acidogenic) phase and stationary (solventogenic) phase under C and CL conditions.

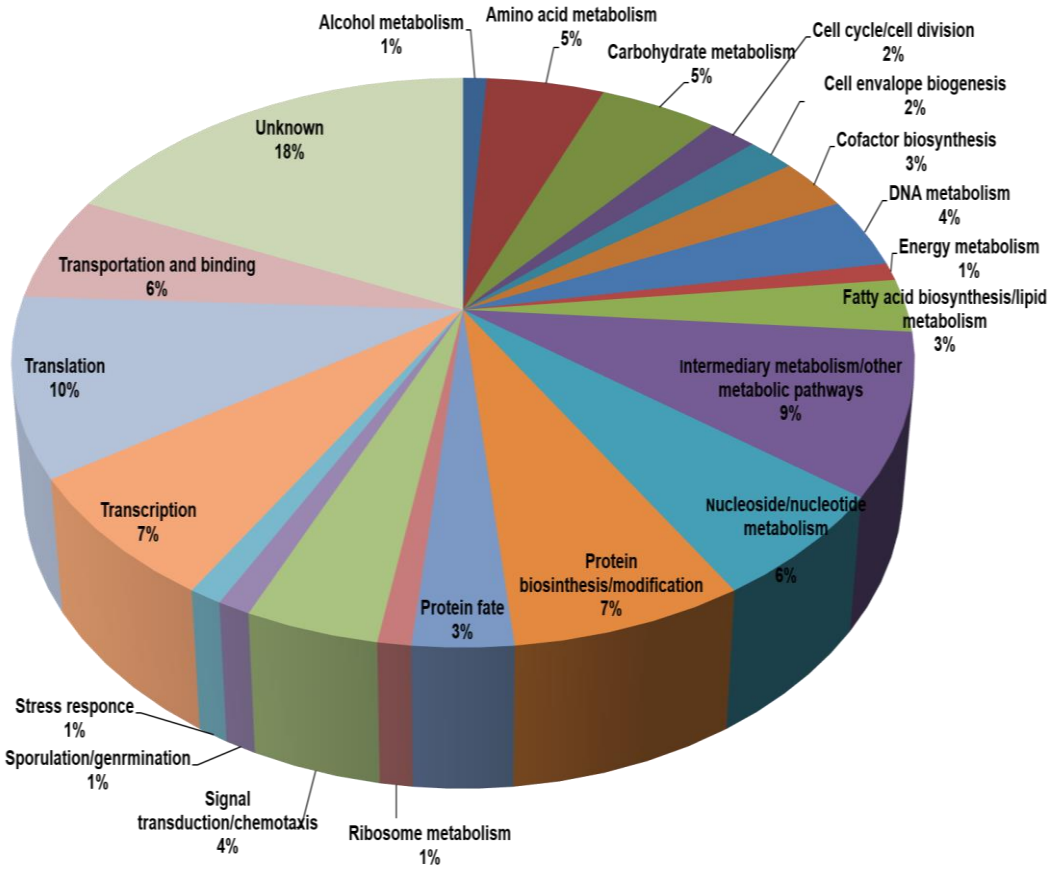

Figure 4-5 Functional classification of iTRAQ quantified proteins belongs to various metabolic pathways under C and CL at exponential (14hrs) and stationary phase (48hrs).

Table 4-1 Global comparison of proteins identified by iTRAQ between (C and CL at exponential (Exp) and stationary (Sta) phases. Red and green arrows represent up- and down-regulated proteins respectively.

Functional categories	Global comparison between phases of cellobiose (C) and cellobiose plus lignin (CL)			
	ExpC/ExpCL (159)	*StaC/StaCL (130)*	*ExpC/StaC (211)*	*ExpCL/StaCL (257)*
Alcohol metabolism	1↑ 1↓	1↑ 1↓	1↑ 2↓	2↑ 2↓
Amino acid metabolism	3↑ 4↓	- 4↓	5↑ 7↓	3↑ 10↓
Carbohydrate metabolism	8↑ 2↓	4↑ 5↓	10↑ 1↓	7↑ 8↓
Cell cycle/cell division	3↑ 2↓	1↑ 1↓	1↑ 2↓	2↑ 5↓
Cell envalope/cell wall biogenesis	- 4↓	3↑ -	1↑ 4↓	5↑ 2↓
Cofactor biosynthesis	1↑ -	- -	4↑ 3↓	4↑ 3↓
DNA metabolism	1↑ 5↓	4↑ 6↓	1↑ 5↓	5↑ 6↓
Energy metabolism/electron transport	- 1↓	- 1↓	1↑ 2↓	- 2↓
Fatty acid biosynthesis/lipid metabolism	1↑ 6↓	1↑ 5↓	1↑ 9↓	3↑ 8↓
Intermediary metabolic pathways	6↑ 6↓	4↑ 4↓	6↑ 13↓	6↑ 11↓
Nucleoside/nucleotide metabolism	2↑ 3↓	3↑ 1↓	2↑ 10↓	5↑ 12↓
Protein biosynthesis/modification	7↑ 8↓	2↑ 7↓	9↑ 11↓	7↑ 8↓
Protein fate	2↑ 1↓	- 4↓	2↑ 4↓	3↑ 6↓
Signal transduction/chemotaxis	3↑ 3↓	2↑ 1↓	2↑ 5↓	3↑ 5↓
Sporulation/genrmination	- -	- 1↓	- 1↓	- 1↓
Stress response	- 3↓	- 2↓	- 3↓	2↑ 2↓
Transcription	7↑ 1↓	2↑ 9↓	7↑ 6↓	4↑ 12↓
Translation	16↑ 7↓	2↑ 14↓	24↑ 6↓	19↑ 18↓
Transportation and binding	8↑ 8↓	6↑ 6↓	5↑ 9↓	10↑ 9↓
Unknown	16↑ 9↓	16↑ 8↓	13↑ 13↓	22↑ 15↓
Total proteins	86↑ 74↓	86↑ 44↓	95↑ 116↓	104↑ 153↓

4.3.3 Metabolic pathways

The identified proteins and their differential expressions in C and CL conditions are discussed in following sections based on their functional annotations.

4.3.3.1 Amino acid metabolism

Three proteins gamma-glutamyl phosphate reductase (CA_C3254), glutamine synthatase type III (CA_2658), and cysteine synthese (CA_2235) belonging to the glutamine, proline and cysteine metabolism were up-regulated in the exponential phase of CL when compared to C (ExpC/ExpCL). An enzyme aspartyl/glutamyl-tRNA amidotransferase B-1 (CA_C2669), which transfers NH_2 groups to convert glutamate to glutamine and aspartate to aspargine, was also up-regulated in the exponential phase of CL (ExpC/ExpCL). Interestingly, lignolytic activity positively correlated with the enzymes involved in NH_2 metabolism in previous studies [304-306]. The γ-glutamyl phosphate reductase (CA_C3254) was up-regulated in both exponential and stationary phases of cells grown on CL (ExpC/ExpCL, ExpCL/StaCL). In a previous study on *Pseudomonas putida* KT2440 [307], this enzyme under phenol (aromatic compound) stress conditions was observed to be up-regulated. However, up-regulation of this enzyme in the C condition in the stationary phase (ExpC/StaC) indicated that the γ-glutamyl phosphate reductase might be involved in the oxidative stress response induced during solventogenesis. The enzymes from arginine and proline biosynthesis; a pyrroline-5-carboxylate reductase (CA_C3252) was up-regulated in the stationary phase of C-fed (C) cells compared to the exponential phase (ExpC/StaC). In a previous study, a higher production of glutamate and

proline was noticed in the presence of phenol in *Corynebacterium glutamicum* [308, 309]. Other proteins including Zn-dependant hydrolase (CA_C2723) and O-acetylhomoserine sulfhydrylase (CA_C2783) were also up-regulated in the stationary phase of lignin-fed (CL) cells (ExpCL/StaCL). A butyryl-CoA dehydrogenase (CA_C2711), which is involved in fatty acid metabolism, butyrate metabolism and amino acid (valine, leucine, and isoleucine) degradation was upregulated in the stationary phase of both C and CL [310]. In *C. acetobutylicum,* it also plays a central role in both acid and solvent production [311]. In this study, proteins from fatty acid synthesis and valine leucine and isoleucine degradation were down-regulated and only butanoate metabolism was up-regulated in the stationary phase of both conditions, indicating metabolic flux diverted towards the synthesis of butanoate, a substrate for butanol production (ExpC/StaC, ExpCL/StaCL). However, it was comparatively down regulated in the stationary phase of CL (StaC/StaCL), which reflects in our results (Figure 4-4).

The thiamine synthesis enzyme (ThiH) (CA_1356) was up-regulated in the stationary phase of CL grown cells when compared to the exponential phase (StaCL/ExpCL). It is involved in thiamine production in anaerobic conditions [312]. However, ThiH was only observed to be up-regulated in the presence of lignin. Thiamine plays a vital role in benzene ring degradation though benzaldehyde lyase [313]. However, we didnot identify benzyldehyde lyase in this study. Two enzymes, Serine hydroxymethyltransferase (CA_C2264) and formate-tetrahydrofolate ligase (CA_C3201) were up-regulated in the presence of lignin when compared to C conditions (ExpC/ExpCL), StaC/StaCL) and ExpCL/StaCL). These enzymes are involved in amino acid metabolism (particularly serine, glycine and threonine), one-carbon metabolism and

tetrahydrofolate interconversion in *E. coli* [314]. Serine hydroxymethyltransferase enzyme induction were previously correlated to cell protection from oxidative drought stress [315], osmotic stress in *Lactococcus lactis* [316] and salt stress conditions [317]. Enzymes involved in one-carbon metabolism were significantly expressed due to sudden metabolic changes in the presence of methylated compounds such as lignin, alkaloids and betaines [318], lignin-derived compounds [319], and xenobiotic compounds in bacteria [320]. Selenocysteine lyase, Nifs family (CA_C3291) and D-ananyl-alanine synthetase A(CA_C2895) were up-regulated in the stationary phase (with and without lignin) hinting at an involvement of these enzymes in solventogenesis [321].

4.3.3.2 Carbohydrate metabolism

In the exponential phase, lignin induced the early up-regulation of enzymes of the glycolysis and pentose phosphate pathways (ExpC/ExpCL): glucose-6-phosphate isomerase (CA_C2680), 2-keto-3-deoxy-6-phosphogluconate aldolase (CA_C2973), transketolase (CA_C0944), and glycolysis pathway: 2-3 biphosphoglycerate independent phosphoglycerate mutase (CA_C0712), enolase (CA_C0713), pyruvate kinase (CA_C0518) and 6-phosphofructokinase (CA_C517). The up-regulation of the pentose phosphate pathway was observed in previous work in furfural-induced stress conditions in *Saccharomyces cerevisiae* [322] and *Zymomonas mobilis* [323]. It was also observed that the oxidative pentose phosphate pathway was expressed in lignin-producing cells to produce NADPH [324] to prevent oxidative stress [325]. The pentose phosphate pathway is also source of reducing power associated with NADPH production, can be used to combat oxidative stress [326]. The enzyme 2-keto-3-deoxy-6-

phosphogluconate aldolase (CA_C2973), which is involved in the Entner-Doudroff pathway, was up-regulated in the stationary phase of cells grown on CL (ExpCL/StaCL). The accumulation of this enzyme in the cell results into bacteriostatis and subsequent blockage of the pentose phosphate pathway [327]. Notably, glycolysis and pentose phosphate pathways were comparatively down-regulated in CL conditions in the stationary phase (StaC/StaCL), whereas in the C treatment, both the pathways were up-regulated (ExpC/StaC). This could possibly explain why the production of acid as well as solvents was lower in the presence of lignin.

The majority of other carbohydrate degradation-related enzymes were up-regulated during the stationary phase in the presence of lignin including triosephosphate isomerase (CA_C0711), possible pectin degradation protein (CA_C3376), galactose mutarotase related enzyme (CA_C3032), phosphomanomutase (CA_C2337) and phosphor β glucosidase. In the other comparison between the exponential phase and the stationary phase of C-grown cells, the following enzymes were up-regulated: triosephosphate isomerase (CA_C0711), galactose mutarotase related enzyme (CA_C2337) and phosphor-β glucosidase (CA_C1408).

Interestingly, we can report the significant up-regulation of possible pectin degradation protein (CA_C3376) in presence of lignin, (ExpCL/StaCL, 16.27 fold and StaC/StaCL, 5.34 fold) and phosphomannomutase (CA_C2337), (2.64 fold, StaC/StaCL). Phosphomannomutase is essential for synthesis of extrapolysaccharides [328]. Interestingly, the pectin degradation protein possesses a cupin domain and belongs to the diverse cupin superfamily [329]. The role of cupin with dioxygenases in aromatic ring degradation have been

studied in soil bacteria [330]. Therefore, the significant expression of this enzyme in this study hints at a possible have role in lignolytic activity too in *C. acetobutylicum*.

4.3.3.3 TCA cycle and Energy metabolism

Interestingly, most of the enzymes involved in the TCA cycle were observed to be down-regulated in CL when compared ExpC/ExpCL, ExpC/StaC, StaC/StaCL and ExpCL/StaCL conditions. During the transition from acidogenesis to solventogenesis, the expression of most of the core pathways, including pyruvate carboxylase, the TCA cycle and amino acid synthesis, was greatly reduced. Changes in the metabolic pathways may redirect energy towards solvent production [331]. ATP synthase subunit b (CA_C2869), an enzyme involved in ATP synthesis, was up-regulated in both conditions during solventogenesis (stationary phase). However, in *C. acetobutylicum*, high ATP generation is associated with acidogenesis and high NADP(H) with solventogeneis [145]. It was found to be even more up-regulated during the exponential phase of cells grown on CL (ExpC/ExpCL). This may be because of the cell requirement for more energy to overcome the high energy barrier to power the pathways that are induced due to the presence of lignin. Recently, significant up-regulation of ATP synthase was observed in lignin amended cells of *Enterobacter lignolyticus SCF1*, where they correlated ATP synthase activity with lignin ring reduction [332]. However, it can also be induced in adaptive response to the stress conditions [333].

4.3.3.4 Transport and binding

The bacterial transport system can be used to control uptake and efflux of molecules across the membrane [334]. Our results indeed suggest significant changes in the transport system proteins in the presence of lignin. Up-regulated proteins during the exponential phase of CL grown cells (ExpC/ExpCL) include three amino acid binding proteins (CA_C1590, CA_C0880, CA_C0111), two phosphotransferases (CA_C1705, CA_C1353) and a predicated permease (CA_C2255). Interestingly, we found several ABC transporter proteins up-regulated during the stationary phase of CL (ExpCL/StaCL, StaC/StaCL) including phosphotransferase system IIC (CA_C1353) and IID (CA_C0068), ATPase component of ABC transporters (CA_C3012, CA_C0147, CA_C3288, CA_C3012, CA_C2982), permeases (CA_C0146, CA_C0139), periplasmic phosphate binding protein (CA_C1705), 2-oxyglutarate/malate translocator (CA_C1590). The biological mechanisms behind the changes in various transportation proteins are still unclear. However, our results indicate that, in presence of lignin, more compounds were likely to have been exchanged with the environment.

4.3.3.5 DNA metabolism, transcriptional and translational regulation

In this study, several proteins involved in transcriptional and translational regulation were differentially expressed in presence of lignin. Interestingly, comparing the exponential phase to the stationary phase of C grown cells, we found that most of the quantified transcriptional and translational proteins were

up-regulated. Our results agreed with previous transcriptomics and proteomics studies where transcriptional, translational and protein stability proteins were expressed during the stationary phase (solventogenesis) [144, 156, 335-338]. Interestingly, during the stationary phase of CL (ExpCL/StaCL, StaC/StaCL), most of the transcriptional proteins (16) were observed to be down-regulated. Thirty-five translational regulatory proteins were differentially expressed (including 17 up-regulated proteins in the presence of lignin): The study clearly indicates that the presence of lignin has a negative influence on transcriptional and translational regulatory proteins.

DNA metabolism was regulated too: 8-oxoguanine-DNA-glycosylase (CA_C2707), recombination protein RecR (CA_C0127), nucleoid associated proteins (CA_C0126) and Superfamily I DNA helicase (CA_C3036) were over expressed in the stationary phase of cells grown on CL (ExpCL/StaCL, StaC/StaCL). The enzyme 8-oxoguanine-DNA-glycosylase removes mutagenic base by-product that results from the exposure to reactive oxygen species as reported previously in anaerobic bacterium *Clostridium perfringens* and *Clostridium acetobutylicum* [339]. Expression of this enzyme directed to DNA damage can be induced by lignin. Expression of nucleoid associated protein with the recombination repair protein RecR involved in the DNA repair system [340]. The exact function of this protein remains unclear, although it is known to possess cell-signalling domains such as the GGDEF and EAL domain [341]. Our results indicate that DNA repair proteins system were activated in order to protect the cell from lignin stress. The heat shock proteins (CA_C1283 and CA_C3714), were down regulated during solventogensis in CL (StaC/StaCL). The induction of the heat shock proteins during solventogenesis has already been observed [342].

Down regulation of these proteins in the stationary phase of CL could be correlated with the lower production of ethanol and butanol in the presence of lignin (Figure 4-4).

4.3.3.6 Protein biosynthesis

We identified several regulated proteins involved in protein biosynthesis. The proteins mainly belong to the family of tRNA ligases: Asparagine-, lysine-, phenylalanine-, glutamate- and isoleucine-tRNA ligases were up-regulated in stationary phase in the presence of lignin (CL) (ExpCL/StaCL). Conversely, the tRNA ligases including Proline-, Cysteine-, Phenylalanine-, Alanine- and elongation factor P, RRNA methylase and acyl phosphatase were down-regulated in the presence of lignin compared to C at stationary phase (StaC/StaCL). We also found that the 60-kDa (CA_C2703) and 10-kDa (CA_C2704) chaperonins were down-regulated in the stationary phase compared to their respective exponential phases in both conditions (ExpC/StaC, ExpCL/StaCL). However, a 60-kDa chaperonin comparatively was down regulated in the stationary phase of CL (StaC/StaCL). These proteins play an important role in protein folding and protein assembly and are normally induced during the transition from acidogenesis to solventogenesis [143]. It is also observed to be expressed in *C. acetobutylicum* in response to various stress conditions [343] which indicates an adaptive response of this bacterium towards lignin.

4.3.3.7 Cell cycle and cell wall biosynthesis

Interestingly, many cell division and cell envelope proteins were differentially expressed in the presence of lignin. ATP dependant zinc metalloprotease (FtsH) (CA_C3202) is a quality control membrane protein involving indigestion of defective assembled protein complexes [344], it was down-regulated in the exponential phase of CL (ExpC/ExpCL). However, it was significantly up-regulated during the stationary phase in CL when compared with the exponential phase of CL (ExpCL/StaCL). FtsH is required for bacterial growth [345], gene regulatory mechanism such as heat shock response, SOS response, capsular polysachharide biosynthesis and regulators degradation [346]. The FtsA-related protein of the HSP70 family, with predicted ATPase activity, (CA_C1013) was also up-regulated in the exponential phase in CL conditions when compared to the exponential phase in C. This protein is predicted to be involved in the cell division activity. We found other cell-cycle proteins to be significantly up-regulated including cell division protein SepF (CA_C2120) and the cell division protein DivIva (CA_C2118) in CL condition. These proteins were involved in the septum formation during cell division, and the over production or mutation in these proteins, results in defective cell divisions and abnormal morphologies and filamentation phenotypes [347]. This is consistent with the morphological changes we observed for cells grown in CL (see Figure 4-3). On the other hand, the cell division protein FtsX (CA_C0498) was down-regulated in the exponential phase in CL compared to the exponential phase in C (ExpC/ExpCL). FtsX is

believed to be a part of the ABC transporter protein with a role in sporulation and the absence of this protein results in a delay of sporulation [348].

Proteins involvied in cell envelope/cell wall biogenesis UTP-glucose-1-phosphate uridylyltransferase (CA_C2335), glucose-1-phosphate thymidylyltransferase (CA_C2333), dTDP-glucose-4,6-dehydratase (CA_C2332) and UDP-N-acetylmuramoylalanine-D-glutamate ligase (CA_C3194) were down-regulated in the exponential phase of cells grown on CL (ExpC/ExpCL). However, at the stationary phase of CL these proteins were found to be comparatively up-regulated (ExpCL/StaCL, StaC/StaCL). Lignin may induce a delay or suppression of the cell envelope/cell wall biogenesis. However, at the same time, mreB/Mbl protein (CA_C1242) is thought to be involved in determining bacterial cellular structure [349] and induces sporulation [350], was up-regulated during the exponential phase of the CL treatment (ExpC/ExpCL). The presence of this protein in the mid exponential phase of the CL condition may suggests early induction of sporulation in the presence of lignin.

4.3.3.8 Signal transduction, chemotaxis and secretion

The chemotaxis signal transduction system components that mediate responses from environmental cues are highly conserved among prokaryotes [351]. We identified several chemotaxis proteins from both growth conditions: S-ribosylhomocysteine lyase (CA_C2942), HtrA-like serine protease (CA_C2433), chemotaxis histidine kinase (CA_C2220) CheA (which contains CheW-like adaptor domain) (CA_C2224), flagella motor switch protein FliG (CA_C2161), N-terminal CheY receiver domain fused to C-terminal uncharacterised CheX-like domain (CA_C0585), chemotaxis signal receiving protein (CA_C2218),

phosphocarrier protein (CA_C1820) and membrane associated signal histidine kinase like ATPase (CA_C3430).

In particular, the enzyme chemotaxis histidine kinase CheA with CheW-like adaptor domain (CA_C2220) was up-regulated in the stationary phase of cells grown in CL (ExpCL/StaCL, StaC/StaCL). Histidine kinases are part of a two-component signal transduction system, which enables bacteria to sense, respond, and adapt to a wide range of environments, stresses and growth conditions [352-354]. The S-ribosylhomocysteine lyase (CA_C2942) is an enzyme that induces the synthesis of auto-inducer AI-2, which is involved in quorum sensing activity. It was up-regulated in the exponential phase (ExpC/ExpCL) and stationary phase (StaC/StaCL)) of cells grown on CL media. It has recently been hypothesised that the Agr-dependant quorum sensing system regulates the sporulation and granulation in *C. acetobutylicum* [140], again suggesting an early onset (in exponential phase of CL) of sporulation in the presence of lignin Whereas, two response regulator proteins (CA_C3220, CA_C2939) were down-regulated during the stationary phase of cells grown on C (ExpC/StaC).

The glutathione peroxidase (CA_C1549) enzyme was up-regulated in the stationary phase of CL-fed cells with respect to the exponential phase. This enzyme protects the cell from oxidative damage [355].

4.3.3.9 Alcohol metabolism

In alcohol metabolism, acetoacetate decarboxylase (ADC) (CA_P0165) was up-regulated (1.40 fold) in the exponential phase and stationary phase of CL conditions (ExpC/ExpCL), however significant up-regulation (4.14 fold) was also

observed in the stationary phase under C conditions (ExpC/StaC). this appeared more significant than the stationary phase of CL (StaC/StaCL) indicating lignin might reduce or delay solvents production (in agreement with Figure 4-4). This protein is directly involved in solventogenesis especially for acetone production [153]. However, interestingly, we observed both ethanol and butanol production but no acetone production. Girbal and Soucaille suggested that at low pH induction of a specific metabolic operon (at high NADPH inside the cells), ethanol and butanol can be produced by *C. acetobutylicum* but not acetone [356]. Expression of acetoacetate decarboxylase and its relation with ethanol and butanol production is unclear. However, the NADH-dependent butanol dehydrogenease (CA_C3392) was significantly up-regulated during the stationary phase of cells grown on CL (ExpCL/StaCL and StaC/StaCL). It was comparatively down-regulated in exponential phase of CL (ExpC/ExpCL). This enzyme is involved in the reversible conversion of alcohol to ketone bodies and aldehydes to insure a constant supply of NAD+ for other metabolic process [357], and also plays a key role in the transition from acid production to solventogenesis [153]. Expression of this protein during the stationary phase of CL and highest production of ethanol and butanol in C conditions indicates the possibility that there was a delay in the production of solvents in the presence of lignin (CL). In addition, two more NADH-dependent butanol dehydrogenases, A and B (CA_C3298, CA_C3299), were down-regulated in the stationary phases compared to their respective exponential phases (ExpC/StaC, ExpCL/StaCL).

There were also a number of proteins identified as unknown/hypothetical in comparisons; ExpC/ExpCL, ExpC/StaC, ExpCL/StaCL and StaC/StaCL (see Appendix 4.1 4.2, 4.3, 4.4). Interestingly, the highest 35 unknown proteins were

differentially regulated in the stationary phase of cells grown CL (ExpCL/StaCL). Most of these were down-regulated. Although the functions of those proteins are unknown, their regulations might ultimately help identify their roles in metabolic flux of essential cellular pathways.

4.4 Conclusions

In this work, we report for the first time, a comprehensive iTRAQ-based proteomic analysis of *C. acetobutylicum* grown on either cellobiose only or cellobiose supplemented with lignin. We used iTRAQ-based proteomics, which is a powerful tool to understand changes in metabolic pathways. The aim was to analyze the proteomic response of *C. acetobutylicum* to lignin stress conditions with an initial view on providing insights into the mechanisms of lignocellusic biomass degradation to biofuel production. Although we could not directly identify lignin degrading enzymes due to the shotgun nature of this study, many proteins known or predicted to be associated with lignin degradation (eg, predicted pectin degradation protein) were found to be differentially expressed.

It revealed the regulation of several proteins under lignin stress conditions (in both the exponential and the stationary phases). Our data suggest that a great number of cellular functions responded to lignin: carbohydrate metabolism, amino acid metabolism, TCA cycle and energy metabolism, cell cycle/cell division, signal transduction and chemotaxis, stress response, transcriptional and translation regulation were regulated. The results shed light on the breadth of the metabolic routes involved in the lignin response in a commercially valuable bacterium. Our study also shows that *C. acetobutylicum* possesses many adaptive, stress and metabolic strategies to respond under challenging

environmental conditions. The several proteins involved in adaptive stress response of the cells in lignin stress condition were identified for the first time in this study such as serine hydroxymethyltransferase, gamma glutamyl phosphate reductase. Changes in carbohydrate metabolism were also observed in presence of lignin. Particularly, onset of enzymes involved in pentose phosphate pathway, which was previously observed to be up-regulated, to produce enough NADPH to protect cells from oxidative stress condition. The enzymes from this pathway mainly found to be up-regulated are gluocose-6 phosphate isomerase, 2 keto-3-deoxy-6-phosphogluconate aldolase, transketolase. There were up-regulation of transportation system proteins such as ABC transporters and permieases indicated that there is exchange of more compounds across the membrane in presence of lignin. In agreement to that proteins of cell signaling system particularly chemotaxis histidine kinases (CheA, CheW-like adaptor domain) were up-regulated in lignin stress condition shows adaptive response of the cells towards lignin containing environment. Our study also revealed onset of DNA repair system proteins (8-oxoguanine-DNA-glycosylase, recombination protein RecR, nucleoid associated proteins and Superfamily I DNA helicase) indicated that lignin induced mutagenesis at genomic level. In complementary to that we identified excess production of cell division proteins resulting in defective cell division (Figure 4-3) in presence of lignin. The proteins involved in alcohol production such as NADPH dependant butanol dehydrogenase was up-regulated in lignin stress condition. This protein responsible for reversible conversion of alcohol to ketone bodies and aldehydes to insure a constant supply of NAD+ for other metabolic process. The significantly up-regulation of this perticular protein shows agreement with the onset pentose phosphate pathway to produce NADPH

to protect cells from stress condition. Expression of this protein during the stationary phase of CL indicates in presence of lignin, alcohols to ketone body conversion could be possible to produce enough NADPH resulting in less production of alcohol to protect the cell from lignin (Figure 4-4 C and D).

Chapter 5

Influence of substrates on the surface characteristics and membrane proteome of *Fibrobacter succinogenes* S85

Abstract

Fibrobacter succinogenes S85 is one of the most proficient cellulose degrading bacterium among all mesophilic bacteria in the rumen of herbivores. *F. succinogenes* possesses an unclear mechanism of cellulose degradation that requires it to adhere to cellulose. To extend the understanding of the fundamental mechanism of cellulose degradation by *F. succinogenes* S85, the bacterial cell surface constituents involved in adhesion to cellulose were characterised using electrophoretic mobility analysis (EPM), microbial adhesion to hydrocarbons (MATH) assay and Fourier transform infra-red (FTIR) spectroscopy, and compared to the cell surface constituents when grown in the presence of glucose. Our results indicate that comparative changes in surface properties of *F. succinogenes* S85 occur during growth on the two different substrates, glucose and cellulose. An increase in the membrane associated proteins and their localisation was evident during the degradation of cellulose. Therefore it was concluded that the membrane proteome of *F. succinogenes* provides an important interface for cell-substrate interactions and subsequent cellulose degradation. Thus, to investigate the membrane associated proteins, we labelled the intact cell with biotin derivative (Sulfo-NHS-SS-biotin) followed by enrichment of proteins using the neutravidin affinity purification method. The identification of

several *F. succinogenes* membrane proteins provides a novel *insight* into the influence of substrates on bacterial topology during growth and reveals the crucial proteins involved in cellulose degradation mechanism as described in the previously proposed mechanism of cellulose degradation in *F. succinogenes*.

5.1 Introduction

Cellulose, an abundantly occurring organic polymer in the plant kingdom [358], has immense potential in the production of alternate fuels such as bioethanol [359]. Since cellulose is a highly stable polymer, expensive chemical hydrolysis is undertaken to ensure adequate yield of fuel from cellulose. Low cost production of fuel from cellulose necessitates the development of inexpensive pre-treatment techniques [359]. Enzymatic deconstruction of cellulose using microbes could be a promising low cost alternative to existing strategies. However, lack of in-depth understanding of cellulose degrading organisms hinders the use of these microbes for cellulose deconstruction in consolidated biofuel generation.

There are many microorganisms capable of enzymatic degradation of cellulose as reviewed by Lynd *et al*. [5]. The microbial consortia in the rumen of herbivores are well specialised for cellulose deconstruction [90]. *F. succinogenes* S85 is a dominant cellulose degrading bacterium of rumen community and actively degrades crystalline cellulose [43]. However, interestingly, unlike other cellulolytic microbes, it does not degrade cellulose by using a cellulosome or an extracellular free enzyme system [96]. Some studies proposed that the initial attachment of *F. succinogenes* to cellulose [97, 360] and unusual orientation of cells along the crystallographic axis of cellulose fibres, may facilitate subsequent degradation of crystalline cellulose by *F. succinogenes* [99, 361]. It has been hypothesised that the degradation of cellulose occurs at the cell surface-membrane of *F. succinognenes* by the following three steps, i) adhesion of cells to cellulose fibres via outer membrane protein complex ii) disruption of cellulose fibres by carbohydrate active enzymes and transfer of individual cellulose chain to

periplasmic space and iii) cleavage of chain into the sugar molecules [57, 100], a portion of which is released into the medium [102, 103].

Previous work has demonstrated that glycosidic residues of surface localised cellulose binding proteins (CBP) (especially CBP with 180-kDa), a cellulose binding domains of endoglucanses (EG2, EGF) and chloride-stimulated cellobiosidase are expressed on the surface of *F. succinogenes* and are crucial for its adhesion to cellulose [97]. Two dimensional electrophoresis (2DE) based proteomics study of outer membrane (OM) by Jun *et al.* [360] identified 25 OM proteins in cellulose grown cells where 16 of these were up-regulated by growth on cellulose as compared to glucose grown cells. The same study also demonstrated the importance of CBPs in adhesion to cellulose. Nonetheless, the proposed presence of 104 glycosyl hydrolases in the genome annotation [104] suggests that a more rigorous investigation is necessary to understand the mechanism of cellulose degradation by *F. succinogenes*.

It is of interest to investigate if the predicted glycosyl hydrolases and other proteins involved in anchoring and degradation of cellulose are localised on the cell surface-membrane of *F succinogenes*. The aim of this study was to investigate the previous suggestion that the expression of glycosyl hydrolases on the surface, membrane and periplasm (surface-membrane) of *F. succinogenes* removes the requirement of a cellulosome in this bacterium [43] and that the availability of substrates induces the expression of surface-membrane associated proteins involved in cellulose degradation.

We hypothesise that the localisation of cellulolytic enzymes on the cell surface in response to changes in substrate availability must change the cell surface

characteristics of *F. succinogenes*. In this study, we investigate the changes in cell surface chemistry of *F. succninogenes* using colloidal surface characterisation techniques such as EPM, MATH assay and FTIR. These techniques have been previously used successfully with bacteria such as *E. coli* and *B. cereus* [362, 363]. In addition, the extraction of cell surface-membrane proteins using biotinylation has been optimised for *F. succinogenes* and a large number of proteins involved in cellulose degradation have been identified, the existence of which has been only predicted via genome annotation.

5.2 Materials and methods

5.2.1 Culture condition and cultivation procedure

All chemicals and reagents were purchased from Sigma-Aldrich (UK) unless otherwise specified. The strain *F. succinogenes* S85 (ATCC 19169) was kindly provided by Prof. Paul Weimer (US Dairy Forage Research Centre, Wisconsin, USA). *F. succinogenes* S85 was cultivated under anaerobic conditions at 38^{o}C in synthetically modified Dehority medium (MDM) as described by Weimer *et al.* [364]. Medium composed of 0.9 g L^{-1} KH_2PO_4, 3.2 g L^{-1} Na_2CO_3, 1 g L^{-1} cysteine·HCl, and 0.06 g L^{-1} each of isobutyric acid, isovaleric acid, n-valeric acid and 2-methylbutyric acid. Final mineral concentration was obtained as follows 0.9 g L^{-1} NaCl, 0.9 g L^{-1} (NH_4) $2SO_4$, 0.084 g L^{-1} $MgCl_2{\cdot}_6H_2O$, 0.065 g L^{-1} $CaCl_2{\cdot}2H_2O$, 0.0275 g of $MnCl_2{\cdot}4H_2O$, 0.02 g L^{-1} $FeSO_4{\cdot}7H_2O$, 0.009 g $L^{-1}ZnCl_2$, and 0.0048 g of CoCl2·6H2O. 10 mL L^{-1} of Schaefer's vitamin solution was also added. Schaefer's vitamin solution was prepared as described by Callaway and Martin [365] containing 20 mg L^{-1} thiamine·HCl, 20 mg L^{-1} Ca-D-pantothenate, 20 mg L^{-1} nicotinamide, 20 mg L^{-1} riboflavin, 20 mg pyridoxine·HCl, 1 mg L^{-1}*p*-

aminobenzoic acid, 0.5 mg L^{-1} biotin, 0.2 mg L^{-1} vitamin B12, 0.125 mg L^{-1} folic acid, and 0.125 mg L^{-1} tetrahydrofolic acid.

Culture media was prepared in triplicate with three different carbon substitutes 1) 0.3 % (w/v) glucose 2) 0.3 % (w/v) microcrystalline cellulose (MC) and 3) 0.3 % (w/v) Acid swollen (AS) cellulose. AS cellulose was prepared by treating MC cellulose, in order to increase the surface area of cellulose particle and to make it more susceptible substrate for bacterial degradation. The cultures were incubated anaerobically under 100 % CO_2 at 38ºC in 125 ml serum bottle (containing 100 mL medium), each fitted with a butyl stopper and an aluminium crimp seal. A starter culture was grown on glucose substrate for 18 hours at an OD_{675} of ca. 0.42. One hundred millilitres of culture media was inoculated with 0.5 ml of starter culture. Specific growth rates of the bacterium under different substrate conditions were calculated from the growth measurement as absorbance (OD_{675}) versus time. Cells were harvested at the mid exponential phase, depending on the growth rate of the bacterial strain under different substrate conditions and processed further as per the protocol for each technique. For cultures grown on MC and AS cellulose, an additional step was performed, where residual cellulose-bounded cells (see figure 5-1) were removed by centrifugation at 500 g for 5 min, before cell pellets were harvested by centrifugation at 8000 g for 10 min [360]. Residual cellulose-bounded cells were detached from the cellulose particles using 0.1 % methyl cellulose in buffer (M8) solution as described previously by Kudo *et al*. [361, 366] (detachment of cells confirmed by microscopy; see figure 5-2). The M8 buffer composed of mineral solution I (3 g L^{-1} K_2HPO_4) and mineral solution II (3 g $L^{-1}KH_2PO_4$, 6 g L^{-1} $(NH_4)_2SO_4$, 6 g L^{-1} NaCl, 0.6 g L^{-1} $MgSO_4$, and 0.6 gL^{-1} $CaCl_2$). The final buffer

concentration was achieved by mixing mineral solution I and II (50:50 (v/v)). Cells harvested by centrifugation at 10000 g for 10 min and combined with previously harvested cell for further analysis. Since glucose is soluble in the medium these additional steps were not applied to glucose grown cells.

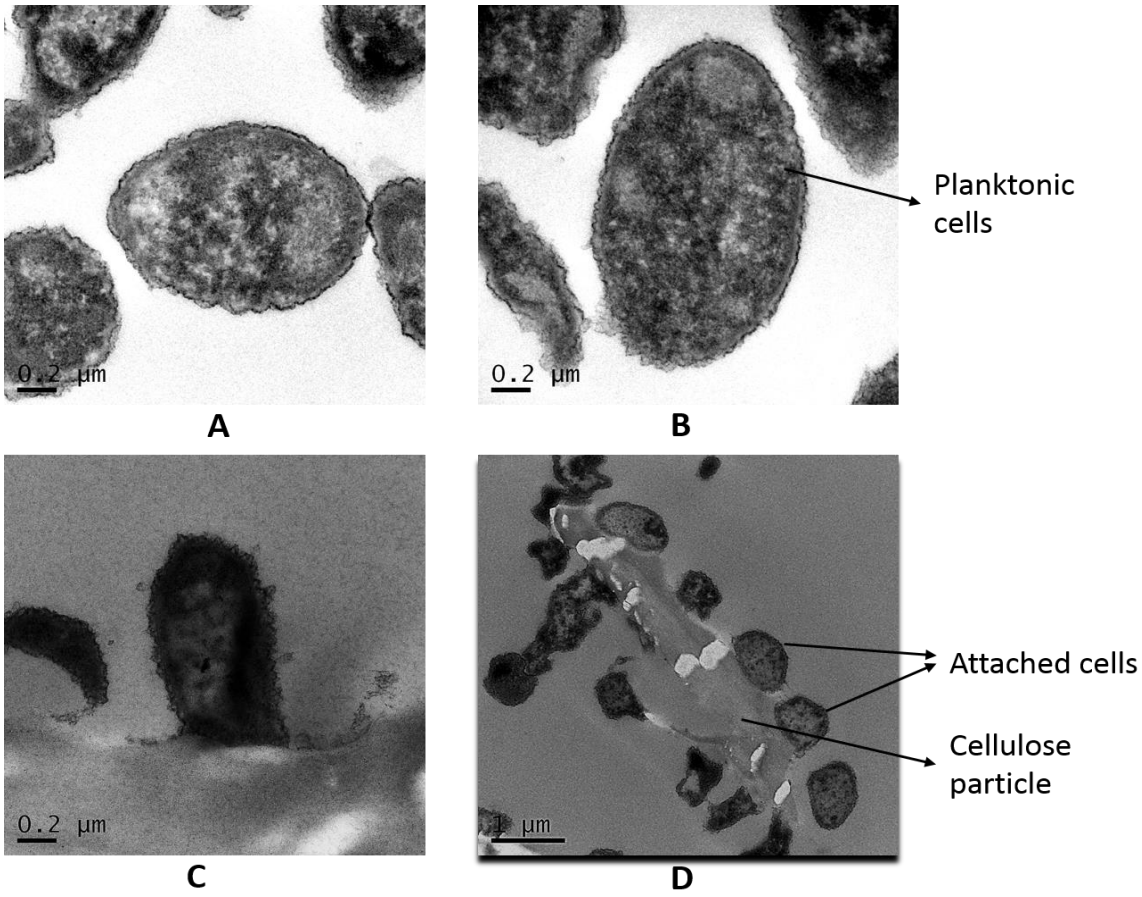

Figure 5-1 Transmission electron microscopy (TEM) images of the bacterium *Fibrobacter succinogenes* S85. (A and B) FS cells grown on glucose. (C and D) FS cells grown and attached to cellulose particles.

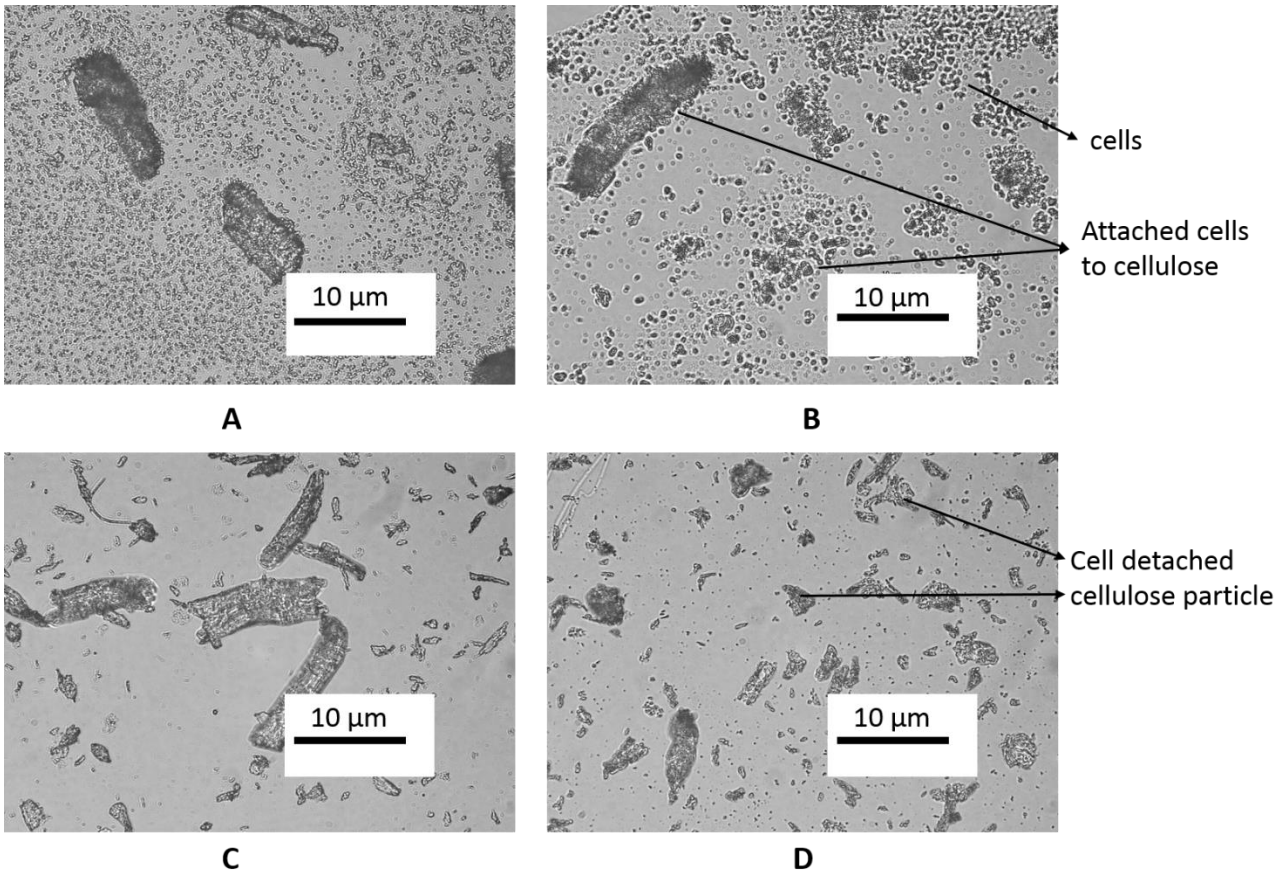

Figure 5-2 Detachment of FS cells from the cellulose particles using 0.1% methyl cellulose treatment. (A and B) Before treatment and (C and D) after treatment.

AS cellulose was prepared by a method described elsewhere [367]. Briefly, 40g of microcrystalline cellulose was mixed in 400ml of 85 % phosphoric acid solution and stored at 4^{o}C for 30 min. The solution was then suspended in 3.6 litre of pre-chilled deionised water and filtered. AC cellulose then washed twice with 2.4 litre of deionised water and resuspended in 2.4 Litre of pre-chilled distilled water and pH adjusted to 6.6 to 6.8. Finally, AC cellulose washed twice and freeze dried.

5.2.2 Carbohydrate determination

For glucose estimation, the culture broth was centrifuged at 8000 g for 10 min to remove bacterial cells and supernatant was analysed by the Nelson-Somogyi method [368]. AS cellulose and MC cellulose was separated from culture by centrifugation at 500 g for 5 min and was estimated using the method described by Updegraff [369].

5.2.3 Physicochemical properties analysis

In previous research, surface properties of bacterial cells has been extensively studied using techniques such as electrophoretic mobility measurements, microbial adhesion to hydrocarbons assay (MATH), Fourier Transform Infrared spectroscopy (FTIR) and surface protein analysis to characterise surface exposed moieties to correlate with adhesion [255, 363, 370-374].

5.2.3.1 Cell surface hydrophobicity

Microbial adhesion to hydrocarbon (MATH) was tested by the method described by Rosenberg *et al*. [375]. Briefly, bacterial cultures were harvested at the mid exponential phase as described previously. Cells obtained by centrifugation at 8000 g for 10min washed twice and resuspended in sterile 150 mM potassium chloride solution at pH 7. In this assay, a 150 mM solution of potassium chloride solution at pH 7 was used to minimise electrostatic effects, since these influences adhesion to n-hexadecane and subsequent interfere with the results [376]. The cell density was approximately adjusted to an OD of 1.0 at 675 nm. One millilitre of this suspension was transferred to new 2 ml Eppendorf tubes and 200µl of the

solvent n-hexadecane was overlaid on each sample. The mixture was vortexed briefly and incubated at room temperature for 15 min. The mixture was vortexed again for 2 min and allowed to settle for 15 min at room temperature. Finally the aqueous layer carefully separated out and the OD at675nm was measured. The hydrophobicity index (HPBI) was calculated as follows [377].

$$\mathrm{HPBI} = \left[\!\left[\frac{(\mathrm{A1} - \mathrm{A2})}{\mathrm{A1}}\right]\!\right] \times 100$$

Where A1 is the initial OD_{675} before mixing with n-hexadecane and A2 is the OD_{675} after mixing with n-hexadecane.

5.2.3.2 Electrophoretic mobility measurement

The electrophoretic mobility of cells at physiological pH was measured to determine the cell surface charge. Cells obtained from glucose and cellulose substrate conditions were washed twice with 100mM potassium chloride solution at pH 7 and OD_{675} adjusted to 1.0. Twenty microlitres of the cell suspension was mixed with 1.8 ml of 100ml KCl solution of pH range 1.5-8. The electrophoretic mobility of cells were analysed in a Zeta potential analyser (ZetaPALS, Brookhaven Instruments, UK). The measurement was conducted using an electric field of 2.5V cm at a frequency of 2.0Hz. The value reported for 3 biological replicates, is an average of 20 cycles with 6 runs conducted at 22^{o}C. The isoelectric point of the bacterial cells was determined as the point of zero electrophoretic mobility of the cell from a pH vs electrophoretic mobility graph.

5.2.3.3 Functional group analysis by Fourier transform infrared (FTIR) Spectroscopy

The FTIR analysis was carried out as described elsewhere [255]. Intact cells obtained were washed three times with 100mM potassium chloride solution at pH 7. Cells were dissolved in same potassium chloride solution for further FTIR analysis.

FTIR analysis was carried out using a Fourier transform infrared spectrophotometer (IRprestige-21 Shimadzu Corporation, UK). Intact cells obtained were mounted on the spectrophotometer using a diamond Attenuated Total Reflectance (ATR) apparatus (Pike Technologies, USA). A blank spectrum without a biological sample was run as a background and the baseline shift of the spectra was corrected using the instrument's software (IR solution). Spectra for samples were recorded in the range 600-3900cm^{-1}using the Happ-Genzel apodisation over 64 scans with a resolution of 4cm^{-1}. Characteristic absorbance peaks of macromolecules of biological origin lies between 800 and 1800 wave numbers [255, 378], thus an FTIR spectrum for this range only was considered for analysis. The spectral data processing was carried out using the IR solution software built into the Shimadzu FTIR instrument. Results obtained from the analysis were interpreted with previously published information [378, 379]. Principal component analysis (PCA) of the FTIR spectra was carried out with XLSTAT software (http://www.xlstat.com/; version 13.1.05) using the Pearson correlation.

5.2.4 Membrane proteome analysis by biotinylation

Biotinylation of *F. succinogenes S85* was performed as previously described [228] with some modification. Briefly, cells were harvested at mid exponential phase for glucose and cellulose grown cells. Cellulose grown cells were first separated from residual cellulose by centrifugation at 500 g for 2 min and further harvested by centrifugation at 8000 g for 5 min. The residual cellulose-bounded cells were detached using 0.1 % methylcellulose solution as suggested by Kudo *et al*. [361]. Cells were harvested by centrifugation and combined with previously harvested cells. Cell pellets were washed with phosphate buffer saline (PBS) containing 1 mM $MgCl_2$ PBS by centrifugation at 8000 g for 5 min at 4^oC and pellets resuspended in the 1ml PBS buffer. Final O.D at 675 nm was adjusted to a corresponding cell count of 2×10^9 cells for all substrate conditions. Cells were further centrifuged and resuspended in the 1 ml PBS containing 1mg EZ-Link® Sulfo-NHS-SS-biotin labels (Thermo, Pierce). The mixture was incubated at 4^OC for 30 minutes and excess biotin was then quenched thrice by washing with 500 mM glycine-PBS solution. Biotin labelled cell pellets resuspended in the 1 ml of radioimmunoprecipitation assay buffer (RIPA) (25 mM Tris-HCl (pH 7.6), 150 mM NaCl, 1 % NP-40, 1 % sodium deoxycholate, 0.1 % sodium dodecyl sulphate, 1.1000 dilution of protease inhibitor cocktail set II). Cell lysate was obtained by brief sonication (30 second sonication 1min on ice; 2 cycles). Cell lysate was incubated on ice for 30min with gentle occasional vortexing. At this stage additional oxidised glutathione (100µM) was added to the lysate to protect disulphide bond in the Sulfo-NHS-SS-biotin. Lysate were further centrifuged at

16000 g for 10 min at 4^{o}C and supernatant collected was stored at -80^{o}C with 10 % (v/v) glycerol until further analysis.

5.2.4.1 Neutravidin affinity purification of biotinylated proteins

Three hundred microlitres of neutravidin agarose gel was washed three times with wash buffer containing 25 mM Tris-HCl (pH 7.6), 0.15 M NaCl, 0.5 % NP40, 0.5 % sodium deoxycholate, 0.05 % SDS. The cell lysate was mixed with washed neutravidin agarose gel and incubated on ice for 2 hours. Mixture then centrifuged at 500 g for 1 min and supernatant were discarded. Gel slurry with biotinylated proteins were transferred to the column (Ultrafree-MC centrifugal filter device; Durapore polyvinylidene difluoride [PVDF], 5.0-µm pore size; millipore). Unbound proteins were removed by washing with washing buffer (25 mM Tris-HCl [pH 7.6], 0.65 M NaCl, 0.1 % NP40) twice, followed by washing with buffer (25 mM Tris HCl [pH7.6], 1.15 M NaCl, 0.1 % NP-40) and finally with Tris buffer (25 mM Tris-HCl [pH 7.6], 0.15 M NaCl) at 200 g for 15-20 second. Gel bound proteins were eluted thrice with 5 % 2-mercaptoethanol-PBS at 30^{o}C for 30 min. Proteins were precipitated by 10 % tricloroacetic acid (TCA) and centrifuged at 18000 g for 10 min at 4^{o}C [380]. The protein pellets were finally washed with ice cold acetone and air dried. The purified proteins then re-dissolved in 0.5 M TEAB buffer containing 0.1 % RapiGest (protein solubilising reagent) and further proteomics analysis was carried out.

5.2.4.2 In-gel digestion for protein identification and peptide recovery

SDS-PAGE was performed on neutravidin-agarose affinity purified proteins separated as standard procedure described elsewhere [254]. In-gel digestion of proteins was achieved as previously described by Karunakaran *et al*. [255].

Briefly, protein bands obtained by 1-D gel electrophoresis sliced in to 10 pieces and destained twice with 200 µL of 200 mM ammonium bicarbonate (AB) in 40 % acetonitrile (ACN) by incubating at 37^{o}C for 30 min. Supernatant was discarded and gel pieces dried in vaccum concentrator. Entrapped gel proteins were reduced and alkylated using 200 µL of reduction buffer (10 mM dithiothreitol) by incubating at 56^{o}C for 1 hour and 200 µL alkylation buffer (55 mM iodoacetamide) at room temperature for 30 min in dark respectively. Gel pieces were washed twice with 200 µl of 50 mM AB solution for 15 min at room temperature followed by 200µL of 50 mM AB in 50 % ACN for 15 min at 37^{o}C. In the next step, samples were digested with 1:50 (w/w) trypsin (Applied Biosystems, USA) containing 0.1 % RapiGest (protein solubilising agent) and 50 µl of 40 mM AB in 9 % ACN for approximately 16 hours by incubation at 37^{o}C. After incubation, the samples were centrifuged briefly at 13000 g for 10 seconds and supernatant was collected in the new siliconised tube. Peptides were extracted twice with 50 µL of 5 % formic acid and 50 µL of 100 % ACN. Finally, all the liquid extracted combined and peptides were vacuum dried (Vacuum concentrator 5301, Eppendorf, UK) and stored at -20^{o}C until further LC-MS/MS analysis.

5.2.4.3 ESI mass spectrometry and identification of proteins

Peptides obtained from in-gel digestion were re-suspended in reverse phase transfer buffer (3 % acetonitrile and 0.1 % formic acid) and submitted to the QStarXL Hybrid ESI-qQ-TOF-MS/MS (AB SCIEX, Concord, Ontario Canada) coupled with an online nano liquid chromatography system (Ultimate 3000, Dionex, Surrey UK). Ten microliters of each peptide sample was injected into the nanoLC-ESI-MS/MS system and then separation performed by PepMap C-18 RP capillary column (LC packing) with constant flow rate of 300 nl min^{-1}. The buffer

used in the liquid chromatography were Buffer A_{ms} (3 % ACN with 0.1 % FA), and Buffer B_{ms} (97 %ACN with 0.1 % FA), and the gradient was as follows: 0 % Buffer B_{ms} for 3 min, 3 % to 36 % Buffer B_{ms} for 90min, 36 % to 90 % of Buffer B_{ms} for 2 min, 90 % of Buffer B_{ms} for 6 min, 3 % of buffer B_{ms} for 13 min. Two precursors of charge +2 and+3 (intensity binning) for each TOF-MS scan (350-1200 *m/z*) were dynamically selected and isolated for MS/MS fragment ion scans (65-1600 *m/z*).

5.2.4.4 Peptide identification

Data obtained from tandem MS analysis were converted to Mascot generic peaklist files (MGF) using Data-Analysis software ver. 4.0 (BrukerDaltonics, Coventry UK). Converted peaklists then submitted to in-house software Phenyx algorithm cluster (Binary version 2.6; Genebio Geneva) for peptide identification. Search was performed against UniProt database for *F. succinogenes* S85 (taxon ID 59374). Search parameters were set at mass tolerance of 0.4 Da and MS/MS tolerance of 0.4 Da. Peptide level filters were set to a z-score of 5.0, max *p*-value significance of 1.0E-5. And AC score were set at 5. Search space was also limited by trypsin peptides with a maximum of 1 missed cleavage.

5.3 Results

5.3.1 Bacterial growth and substrate consumption

The substrate consumption profile and growth rate of *F. succinogenes* S85 is shown in Figure. 5-3. *F. succinogenes* S85 grew on glucose, AS cellulose and MC cellulose with growth rates of 0.20, 0.098 and 0.084h^{-1} respectively. A corresponding decrease in the concentration of substrates was seen. Glucose

was consumed at a rate of0.130 mg/mL/hr whilst AS cellulose and MC cellulose was consumed at a rate of 0.071 mg/mL/hr, and 0.057 mg/mL/hr respectively.

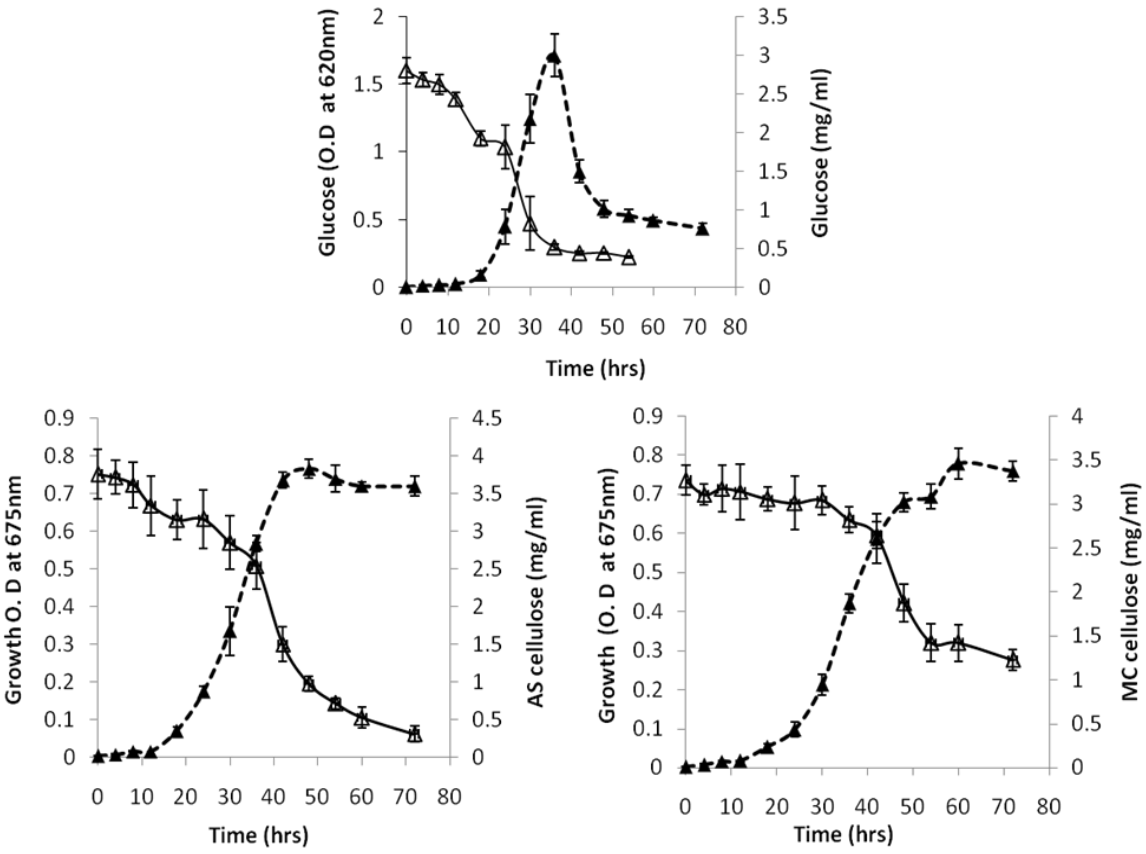

Figure 5-3 Growth and substrate consumption profile of *F. succinogenes* S85: A) Glucose B) Acid swollen (AS) cellulose and C) Microcrystalline (MC) Cellulose (∆ represents substrate utilisation and ▲ represents OD_{675})

Results indicate a rapid growth rate of *F. succinogenes* S85 when grown on glucose as opposed to cellulose substrates. The rate of mortality was also higher in glucose grown cells once cells reached the stationary phase. Previous studies noted that, in substrate depletion conditions, the cells produce extracellular proteases which causes autolysis of the cells [381]. In contrast, cells grown on

cellulose substrates were characterized by a long log phase followed by a more sustained stationary phase.

5.3.2 Hydrophobicity and Surface charge

Changes in bacterial cell surface chemistry upon exposure to different carbon sources reflect as changes in hydrophobicity and charge of the cell surface. The MATH assay is routinely used to measure the extent of hydrophobicity of bacterial cell surfaces [255, 363]. The results of the MATH assay of the cells grown in glucose and cellulose substrates condition are shown in Figure 5-4. A percentage hydrophobicity less than 30 % generally considered as the hydrophilic surface [382, 383]. Our results suggest that surface of cells grown on glucose are hydrophilic in nature. However, on exposure to the different cellulose substrates, the cell surface becomes more hydrophilic as compared to surface of glucose grown cells. The student t-test performed among the treatments shows significant difference between glucose and two types of cellulose treatments (p value = 0.0464 & 0.0409, respectively) at 95 % confidence whereas no significance difference among the two cellulose treatments (p value = 0.484).

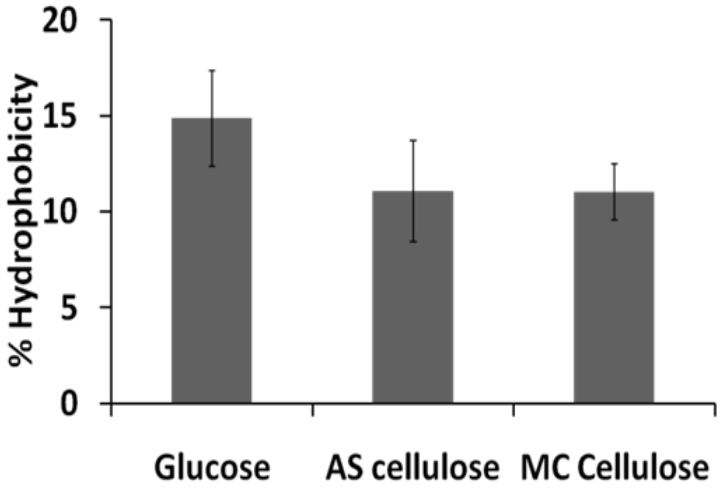

Figure 5-4 MATH assay of *F. succinogenes* S85 cells grown on different substrates. Error bars = SE value

Zeta potential is proportional to the cell surface charge which arises from the protonation/deprotonation of surface exposed functional groups [384]. The calculation of zeta potential from the measured electrophoretic mobility is based on assumptions that do not always hold true when measuring bacteria [385]. Therefore electrophoretic mobility (EPM) was used for data interpretation in this study.

The EPM data was then plotted as a function of pH (Figure 5-5). The magnitudes of the EPM become more negative with an increase in pH for cells grown in the presence of all substrates. This trend is generally observed whilst working with other bacteria such as *E. coli, Bacillus cereus, Bacillus brevis, Staphylococcus aureus, Pseudomonas putida and Alcaligenes faecalis* [255, 363, 386-389]. The trend arises due to the differences in the protonation/deprotonation state of the surface exposed functional groups.

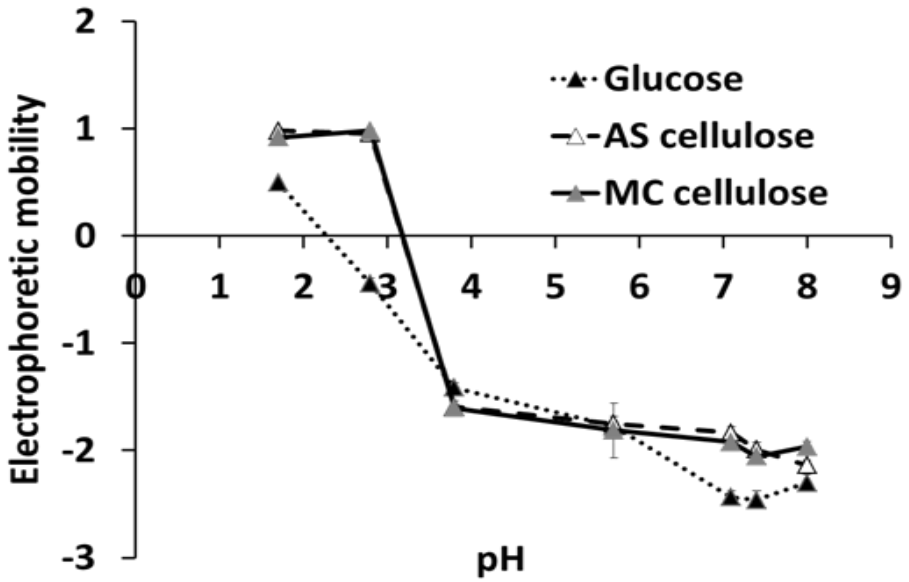

Figure 5-5 Electrophoretic mobility of *F. succinogenes* S85 cells under different carbon substrate conditions as a function of pH. Error bars = SE value.

Whilst the electrophoretic mobility of cellulose grown cells at pH 7 is only marginally less electronegative than the glucose grown cells, significant differences can be seen in the isoelectric points of the cells grown in the different substrates. The isoelectric point is the pH at which the net EPM of the cell is zero [384]. The isoelectric point for cells grown in glucose was obtained at pH 2.2, whereas for cells grown in AS cellulose and MC cellulose, the isoelectric points were between 3 and 3.5. The results demonstrate that the surface chemistry of the cells changes with changes in the presence of different carbohydrate as substrate in the culture medium.

5.3.3 Compositional changes to cell surface

The FTIR spectra of the intact cells reflects the vibrational motions of specific functional groups or bonds in biochemical components such as proteins, lipids, and carbohydrates within cell membranes [379]. The FTIR spectrum was

recorded between 600 cm^{-1} to 3900 cm^{-1}. However, the most biological information can be obtained from the spectral region between 800 cm^{-1} and 1800 cm^{-1} [378]. Therefore, the FTIR spectrum was considered for bacterial surface analysis within this range.

The ATR-FTIR spectra of the cells grown on glucose and cellulose substrates are presented in Figure 5-6. Comparison of the FTIR spectra shows that major differences were exhibited in the ring vibrations of C-O-C and C-O group of carbohydrates, C-O-P and P-O-P of polysaccharide region (1200–900 cm^{-1}) and the C-O-C group of esters region (1230cm^{1}) [378]. Considerable differences was also observed in the amide I (C=O) and amide II (N-H) regions that lie between 1700cm^{-1} and 1500cm^{-1} [378]. The results suggest a decrease in the cell surface polysaccharide display and a concomitant increase in the cell surface protein display when the cells are grown in cellulose when compared to glucose grown cells.

Furthermore, Principal component analysis (PCA) of the different spectra of cells grown on glucose and cellulose substrates was carried out (Figure 5-7). The PCA analysis reiterates the fact that the cell surface of cellulose grown cells is distinctly different from that of glucose grown cells. No significant difference is seen in the surface of the cells grown in the presence of the two different forms of cellulose.

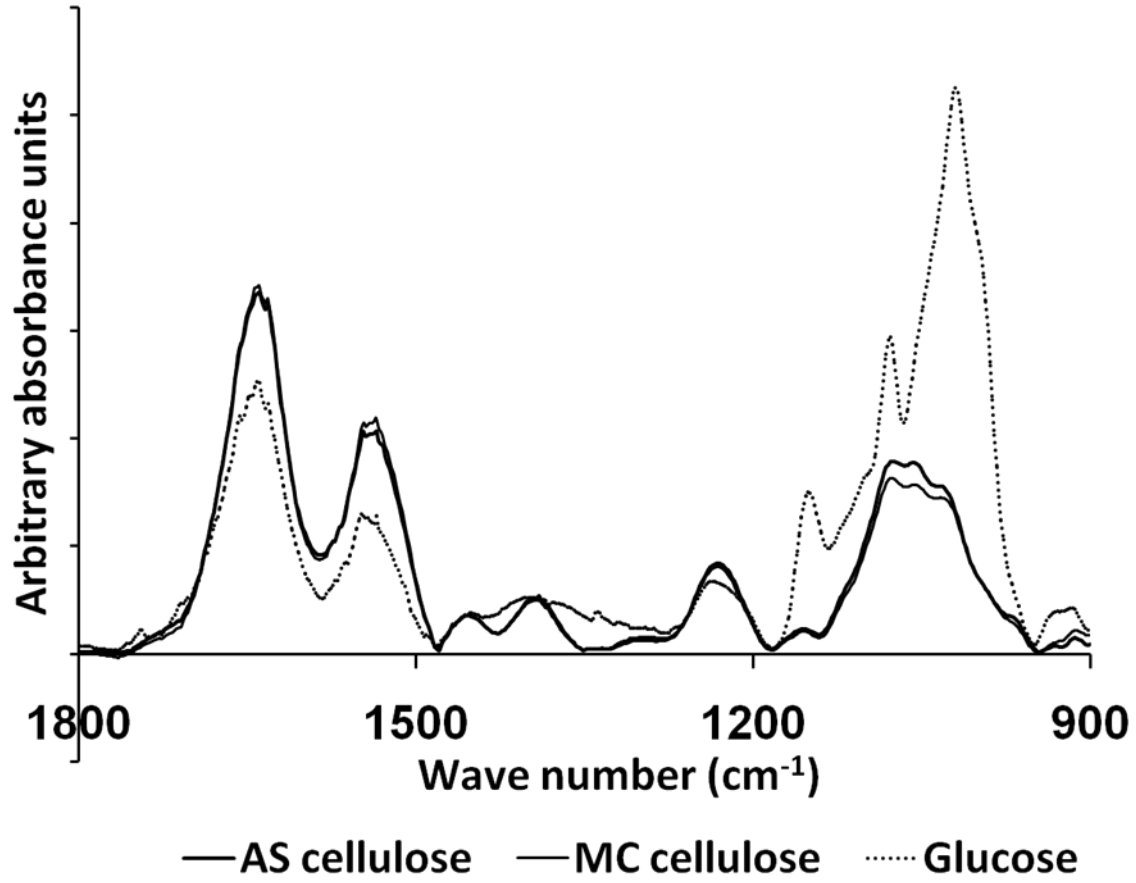

Figure 5-6 Comparative FTIR spectrum of *F. succinogenes* S85 strains grown under different carbon substrate conditions.

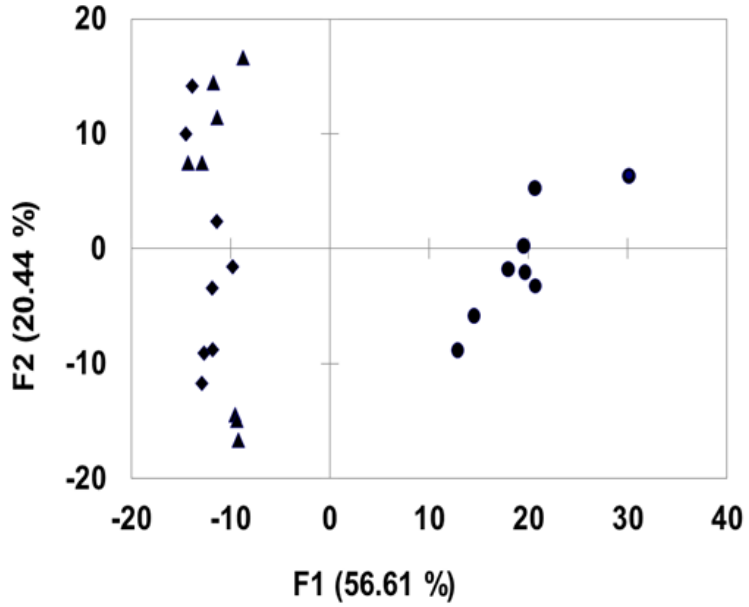

Figure 5-7 Principal component analysis (PCA) of ATR-FTIR spectra of *F. succinogenes* S85 cells grown on (●) Glucose, (▲) AS cellulose, (♦) MC cellulose

5.3.4 Neutravidin affinity purification of biotinylated surface-membrane proteins of F. succinogenes

Intact bacterial cells biotinylated with sulfo-NHS-SS-biotin and biotinylated proteins were purified by neutravidin affinity chromatography. The purified proteins were analysed by SDS-PAGE. To ensure that the proteins in the gel were truly enriched due to biotin-neutravidin affinity, a control step was included during affinity purification using proteins prepared from unbiotinylated *Fibrobacter succinogenes* cells grown on glucose (Figure 5-8). The absence of proteins in the SDS-PAGE gel of the control sample clearly demonstrated that the proteins found in the SDS-PAGE gel of the experimental samples did not arise due to inadequate wash steps or non-specific binding of proteins to neutravidin. The proteins found

in the experimental samples were digested using trypsin and were identified by MS/MS analysis.

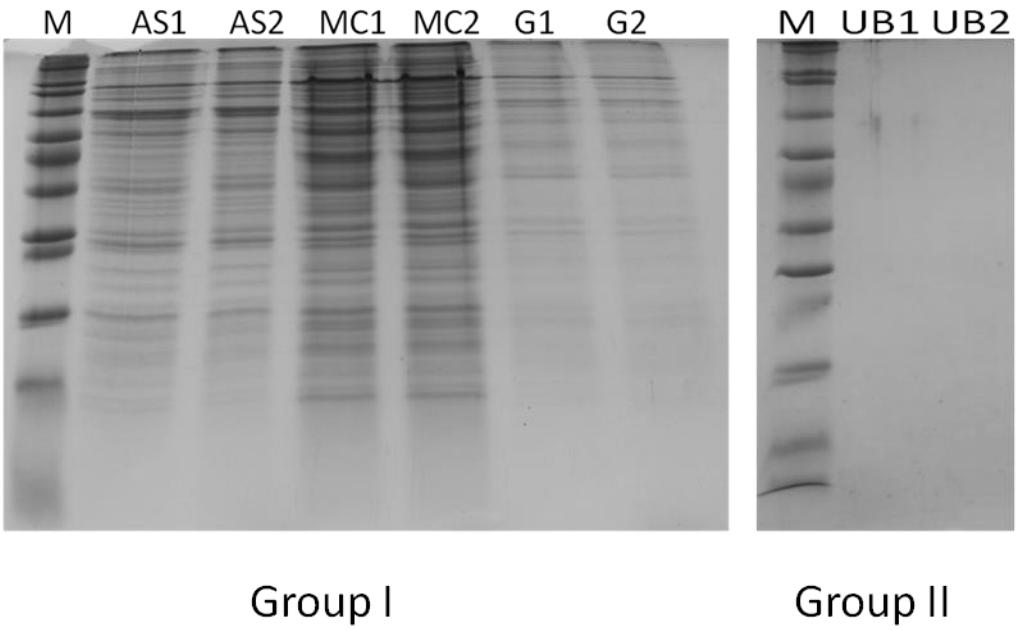

Figure 5-8 SDS PAGE of biotinylated samples from different substrate conditions Group I; (M) Marker, (AS1-2) AS cellulose, (MC1-2) MS cellulose and (G1-2) Glucose, Group II; (M) Marker, (UB1-2) Unbiotinylated samples

Across the three substrate conditions, a total of 340 proteins were identified. Distribution of the identified proteins among the three substrate conditions is summarised in Figure 5-9. All identified proteins were non-cytoplasmic, membrane associated proteins. The localisation of proteins were predicted by (PSORTb) [390]. Carbohydrate active enzymes were classified into families according to CAZy database (http://www.cazy.org) [391]. The membrane associated proteins whose functions have been previously characterised are divided into two categories on the basis of their roles in various metabolic processes; 1) cellulose degradation (Table 1; 33 proteins) and 2) Energy

generation, transport and protein-protein interaction (Table 2; 53 proteins out of 160 proteins identified). Total 160 proteins including this category are shown in Appendix 5.1. In addition to these, 147 proteins were identified as putative uncharacterised lipoprotein/membrane proteins with unknown functions (Appendix 5.2).

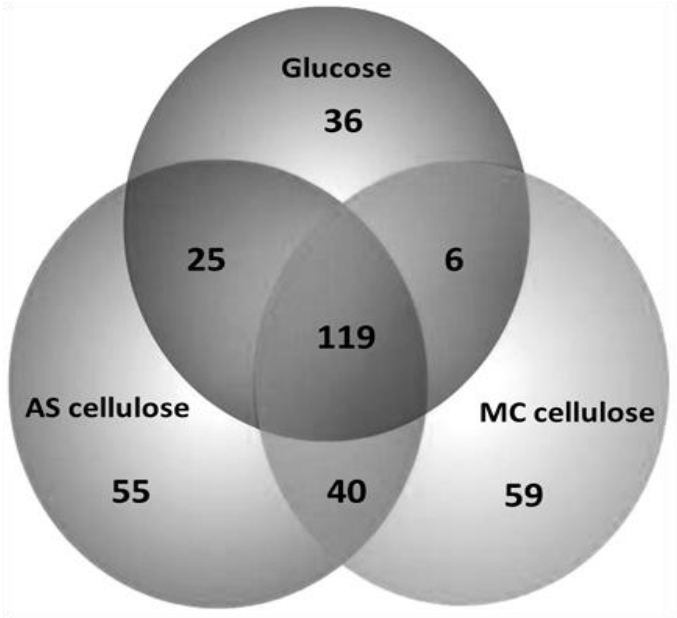

Figure 5-9 Venn diagram showing distribution of the 340 membrane-associated proteins among three different substrate conditions

Table 5-1 List of surface exposed and membrane proteins involved in carbohydrate degradation in *F. succinogenes* S85

Locus ID	Protein description	Glucose	MC cellulose	AS cellulose	Family[a]	Location[b]	Gravy index[c]	Molecular mass (kDa)[c]	pI[c]	Presence of signal peptide (amino acid position)[d]	Ref
Fisuc_0678 FSU_1114	Carbohydrate-binding protein	√	√	√	PL10, CBM6, CBM35	Extracellular	-0.456	63.41	8.8	Yes (24-25)	-
Fisuc_2363	Pectate lyase/Amb allergen	-	√	√	PL1	Extracellular	-0.423	76.728	5.96	Yes (19-20)	-
Fisuc_1252 FSU_1715	Peptidoglycan glycosyltransferase	√	√	√	GT51	Unknown	-0.35	126.864	6.76	No	-
Fisuc_0083	Glycosyl transferase family 2	-	√	-	GT2	Unknown	-0.063	35.986	8.9	No	-
Fisuc_3049 FSU_0315	Beta-galactosidase	-	-	√	GH2	Unknown	-0.304	105.982	5.18	No	[43]
Fisuc_2250 FSU_2795	O-Glycosyl hydrolase-like protein	-	-	√	GHnc	Non cytoplasmic	-0.276	77.003	5.47	Yes (17-18)	-
Fisuc_2900 FSU_0162	Cellodextrin-phosphorylase	√	√	√	GH94	Cytoplasmic membrane	-0.352	93.662	8.16	No	-
FSU_2361	Glycoside hydrolase family 9	-	√	√	GH9	Non cytoplasmic	-0.248	67.365	6.06	Yes (26-27)	[43, 392]
Fisuc_1859											
Fisuc_1860 FSU_2362	Glycoside hydrolase family 9	-	-	√	GH9	Non cytpplasmic	-0.322	71.151	6.17	Yes (21-22)	[43, 392]
FSU_2362	Cellulase	-	-	√	GH9	Non cytoplasmic	-0.309	71.378	6.1	Yes (23-24)	-
FSU_2303	Glycoside hydrolase family 8	√	√	√	GH8	Unknown	-0.219	81.39	5.55	No	[393]
FSU_2303	Glycoside hydrolase family 8	√	√	√	GH8	Non cytoplasmic	-0.24	79.81	5.63	No	[43, 393]
Fisuc_1802											
Fisuc_1219	Glycoside hydrolase family 8	√	√	√	GH8	Non cytoplasmic	-0.037	52.666	4.93	Yes (25-26)	[43]
Fisuc_1523 FSU_2005	Cellulase	-	√	√	GH5	Unknown	-0.239	42.062	4.94	Yes (19-20)	[43]
cel-3 Fisuc_2230 FSU_2772	Endoglucanase 3	-	√	√	GH5	Non cytoplasmic	-0.359	73.424	4.61	Yes (25-26)	[43, 394]
Fisuc_2364 FSU_2364	Cellulase	-	√	-	GH5	Extracellular	-0.365	98.039	5.27	Yes (18-19)	[43, 393]
FSU_2914	Cellulase	-	√	-	GH5	Extracellular	-0.359	98.317	5.27	Yes (20-21)	-

Fisuc_0786 FSU_1228	Cellulase	-	-	√	GH5	Non cytoplasmic	-0.247	78.034	5.23	Yes (21-22)	[43]
Fisuc_1473	Cellulase	-	√	-	GH45	Periplasm	-0.348	49.927	4.73	Yes (22-23)	[43]
Fisuc_0393 FSU_0809	Glycoside hydrolase family 9	√	√	√	CBP1, GH9	Non cytpplasmic	-0.335	233.01	4.97	Yes (18-19)	[43]
Fisuc_0730	Mannan endo-1,4-beta-mannosidase	√	-	-	CBMnc, GH26	Non cytoplasmic	-0.324	69.25	5.44	Yes (26-27)	[43, 395]
Fisuc_3111 FSU_0382	Carbohydrate binding family 11	√	√	√	CBM30 ,CBM11, GH51	Non cytoplasmic	-0.551	118.616	7.81	Yes (23-24)	[43, 396]
Fisuc_1525 FSU_2007	Cellulose-binding domain protein	√	√	√	CBM30	Non cytoplasmic	-0.301	29.2	6.46	Yes (35-36)	[43, 360]
FSU_3194	Fibronectin type III domain protein	√	-	-	-	Unknown	-0.607	77.87	4.78	No	-
Fisuc_2624	Fibronectin type III domain protein	√	-	-	-	Outer membrane	-0.584	88.87	4.74	Yes (27-28)	-
Fisuc_1979 FSU_2502	Fibro-slime domain protein	√	√	√	-	Non cytoplasmic	-0.312	169.38	5	Yes (32-34)	[360]
Fisuc_0377	Fibro-slime family protein	√	√	√	-	Non cytoplasmic	-0.433	98.69	5.16	Yes (21-22)	-
Fisuc_2041 FSU_2567	Putative type IV pilin	√	√	√	-	Unknown	-0.083	17.516	5.43	No	[360]
FSU_0286	Fimbriae-associated domain protein	√	√	√	-	Cytoplasmic membrane	-0.002	67.316	4.49	No	-
Fisuc_0771 FSU_1212	Pilin domain protein	-	-	√	-	Non cytoplasmic	-0.355	52.094	5.23	No	-

[a] Carbohydrate active enzymes database (http://www.cazy.org/) [391]

[b]Location of the given proteins predicted by the PSORTb subcellular localization prediction tool version 3.0 [390]

[c]Theoretical isoelectric point, molecular mass and gravy index of the given protein, as predicted by the ExPASy Compute pI/MW tool [397]

[d]Determined by SignalP v.3.0 [398] the numbers in parentheses indicates the amino acids between which cleavage is predicted to occur in the given protein

The identified proteins associated with carbohydrate degradation include: 19 glycoside hydrolases (GH) (three GH8, five GH9, seven GH5, two GH45 and each of GH45, GH26 and GH2 family); and 7 cellulases (five GH5, and each from GH45, and GH9 family. Some of these enzymes have carbohydrate binding

modules (CBM) – CBM6-11-30-35, therefore possess multiple CAZy family membership. GHnc protein (FSU_2795), consider to be a CBM that not yet been assigned to any family, which probably possesses xylanase activity involved in hemicellulose degradation based on a BLAST similarity (24 %) with a xylanase from *Aeromonas caviae*. An enzyme pectate lyase/Amb allergen (Fisuc_2363) classified as pectin lyase (PL1) is also identified in this study.

Other identified proteins are belongs to; fibro-slime family, fibronectin type III domain proteins, putative IV pilin, and cadherin. Most of these proteins are involved in adhesion, biofilm formation and cell-cell interaction.

Out of the 160 proteins, 53 selected proteins were described in Table 5-2, which are involved in various membrane processes including OmpA family/domain proteins, TonB family proteins, TPR domain proteins, extracellular solute binding proteins, substrate transporter proteins, efflux transporter proteins, proton channel proteins and capsular/surface repeat proteins.

Table 5-2 List of surface exposed and membrane proteins involved in membrane associated processes in *F. succinogenes* S85

Locus ID	Protein description	Glucose	MC cellulose	AS Cellulose	Location[b]	Gravy index[c]	Molecular mass (kDa)[c]	pI[c]	Presence of signal peptide (amino acid position)[d]	Ref
Fisuc_0344 FSU_0758	TonB family protein	√	-	-	Non cytoplasmic	-0.447	53.371	8.84	Yes (20-21)	-
Fisuc_0552 FSU_0976	Periplasmic solute binding protein	√	-	-	Unknown	-0.271	32.87	9.39	No	-
Fisuc_2918 FSU_0181	Capsular exopolysaccharide family	√	-	-	Cytoplasmic membrane	-0.165	80.26	7.97	No	-
Fisuc_0891 FSU_1339	Cell wall/surface repeat protein	√	-	-	Non cytoplasmic	-0.229	94.84	5.25	Yes (24-25)	-
FSU_2398	TPR domain protein	√	√	√	Non cytpplasmic	-0.617	146.67	8.5	Yes (20-21)	[360]
FSU_2396	OmpA family protein	√	√	√	Outer membrane	-0.408	55.97	4.86	Yes (28-29)	[360]
Fisuc_1892 FSU_2397	TPR domain protein	√	√	√	Non cytpplasmic	-0.341	83.84	5.51	Yes (23-24)	[360]
Fisuc_1891	OmpA/MotB domain protein	√	√	√	Unknown	-0.458	53.24	4.74	No	-
Fisuc_1893	TPR repeat-containing protein	√	√	√	Unknown	-0.619	146.7	8.5	Yes (20-21)	-
Fisuc_1894 FSU_2400	MotA/TolQ/ExbB proton channel	√	√	√	Cytoplasmic membrane	0.461	23.06	9.14	No	-
Fisuc_1980	Sporulation domain protein	√	√	√	Unknown	-0.476	13.35	9.56	No	-
mscL Fisuc_2074 FSU_2602	Large-conductance mechanosensitive channel	√	√	√	Cytoplasmic membrane	0.462	15.94	9.21	No	-
Fisuc_0201	MotA/TolQ/ExbB proton channel	√	√	√	Cytoplasmic membrane	0.319	24.61	7.78	No	-
Fisuc_2503 FSU_3071	Extracellular solute-binding protein family 3	√	√	√	periplasmic	-0.132	28.75	5.47	Yes (21-22)	-
Fisuc_0289 FSU_0701	Efflux transporter, RND family, MFP subunit	√	√	√	Non cytoplasmic	-0.115	37.256	9.45	No	-
Fisuc_2509 FSU_3077	OmpA family protein	√	√	√	Outer membrane	-0.381	83.917	4.75	Yes (19-20)	-
secDF Fisuc_0858 FSU_1302	Protein-export membrane protein SecD	√	√	√	Cytoplasmic membrane	0.285	93.186	8.96	No	-
Fisuc_2892	OmpA/MotB domain protein	√	√	√	Outer membrane	-0.36	32.21	5.53	Yes (19-20)	-
Fisuc_2917 FSU_0180	OmpA family protein	√	√	√	Outer membrane	-0.272	70.79	5.27	Yes (17-18)	-
Fisuc_0978	Capsular exopolysaccharide family	√	√	√	Cytoplasmic membrane	-0.242	77.96	8.94	No	-
Fisuc_1192 FSU_1653	Periplasmic solute binding protein	√	√	√	Cytoplasmic membrane	-0.125	36.391	5.03	Yes (20-21)	-
Fisuc_1226 FSU_1687	ABC transporter related protein	√	√	√	Cytoplasmic membrane	-0.323	30.503	8.6	No	-
Fisuc_1230 FSU_1691	Extracellular solute-binding protein family 5	√	√	√	Unknown	-0.336	67.554	5.69	Yes (21-22)	-
Fisuc_1591 FSU_2077	Capsular exopolysaccharide	√	√	√	Cytoplasmic membrane	-0.092	78.322	8.14	No	-
Fisuc_1592	OmpA/MotB domain protein	√	√	√	Outer membrane	-0.386	73.841	4.72	Yes (17-18)	-

FSU_0151	OmpA family protein	√	√	√	Cytoplasmic membrane	-0.566	23.937	6.65	No	-
Fisuc_1980	Sporulation domain protein	√	√	√	Unknown	-0.476	13.35	9.56	No	-
Fisuc_3033 FSU_0298	Mechanosensitive ion channel family protein	√	√	-	Cytoplasmic membrane	0.55	29.387	6.32	No	-
Fisuc_1658 FSU_2147	TPR repeat-containing protein	√	√	-	Non cytoplasmic	-0.676	27.923	7.61	Yes (22-23)	-
Fisuc_1229	Binding-protein-dependent transport systems inner membrane component	√	-	√	Cytoplasmic membrane	0.227	51.588	9.32	No	-
FSU_1690	Putative oligopeptide/dipeptide ABC transporter, permease protein	√	-	√	Cytoplasmic membrane	0.24	50.827	9.28	No	-
Fisuc_1897 FSU_2403	TonB family protein	-	√	√	Unknown	-0.393	32.452	9.84	No	-
Fisuc_0042 FSU_0435	MotA/TolQ/ExbB proton channel	-	√	√	Non cytoplasmic	-0.079	57.895	9.54	Yes (49-50)	-
Fisuc_0617	Extracellular solute-binding protein family 1	-	√	√	Unknown	-0.213	57.545	7.77	No	-
Fisuc_1151 FSU_1609	OmpA family protein	-	√	√	Non cytoplasmic	-0.274	22.491	8.74	No	-
Fisuc_1465 FSU_1938	Extracellular ligand-binding receptor	-	√	√	Non cytoplasmic	-0.212	67.212	9.25	Yes (19-20)	-
FSU_1047	Extracellular solute-binding protein	-	√	√	Unknown	-0.218	57.575	7.77	No	-
Fisuc_0288 FSU_0700	Outer membrane efflux protein	-	√	√	Outer membrane	-0.378	52.346	5.31	Yes (21-22)	[360]
Fisuc_2028 FSU_2553	Efflux transporter, RND family, MFP subunit -	-	√	-	Cytoplasmic membrane	-0.24	46.108	9.5	No	-
Fisuc_0120	Tetratricopeptide TPR_2 repeat protein	-	√	-	Unknown	-0.413	143.314	5.69	Yes (27-28)	-
Fisuc_0480 FSU_0898	Efflux transporter, RND family, MFP subunit	-	√	-	Cytoplasmic membrane	-0.036	38.448	9.52	No	-
Fisuc_0539	Type II and III secretion system protein	-	√	-	Outer membrane	-0.022	65.248	5.9	Yes (22-23)	-
Fisuc_0743 FSU_1181	ABC transporter related protein	-	√	-	Cytoplasmic membrane	-0.071	29.041	5.61	No	-
Fisuc_0885 FSU_1331	Outer membrane efflux protein	-	√	-	Outer membrane	-0.345	61.621	5.22	Yes (19-20)	-
atpD Fisuc_2839 FSU_0095	V-type ATPase, D subunit	-	√	-	Unknown	-0.448	23.981	9.87	No	-
Fisuc_1227 FSU_1688	Oligopeptide/dipeptide ABC transporter, ATP-binding protein	-	√	-	Cytoplasmic membrane	-0.138	36.622	8.33	No	-
Fisuc_1395 FSU_1863	Capsular polysaccharide biosynthesis domain protein	-	√	-	Unknown	-0.002	43.793	5.5	Yes (17-18)	-
Fisuc_1571 FSU_2056	Outer membrane efflux protein	-	√	-	Outer membrane	-0.387	47.008	5.45	Yes (20-21)	-
Fisuc_1621 FSU_2110	Extracellular solute-binding protein	-	√	-	Non cytoplasmic	-0.057	58.667	5.89	Yes (28-29)	-
Fisuc_1133 FSU_1591	Putative transporter	-	√	-	Cytoplasmic membrane	0.726	58.663	9.18	No	-
Fisuc_0149 FSU_0552	Sulfate ABC transporter, periplasmic sulfate-binding protein	-	-	√	Periplasm	-0.401	37.761	5.68	Yes (22-23)	-
Fisuc_0197 FSU_0604	OmpA/MotB domain protein	-	-	√	Unknown	-0.398	21.493	4.95	No	-
Fisuc_0555 FSU_0979	Cell wall/surface repeat protein	-	-	√	Outer membrane	-0.084	135.996	5.23	Yes (15-16)	-

In addition to these, a total of 147 proteins were identified as putative/uncharacterised proteins (Appendix 5.2). Identification of such a large number of proteins with an unknown function is not surprising since 50 % of open reading frames (ORF) identified during genomic annotation have unknown functions in *F. succinogenes* S85. To date, prediction of the function of these membrane associated proteins remains a major challenge for the researchers.

5.4 Discussion

The absence of a cellulosome or extracellular free enzyme system and the absolute requirement of cellular adhesion to cellulose have prompted researchers to hypothesise that cellulolytic enzymes are harboured on the cell surface-membrane in *F. succinognenes*. Using 2DE, Jun *et al*. [360] demonstrated the occurrence of cellulose binding proteins, on the cell surface of *F. succinogenes* grown in cellulose and explained the adhesion of cells to cellulose. However, the study did not address the hypothesis that cellulolytic enzymes are also harboured on the cell surface-membrane of *F. succinogenes* as suggested by Wilson DB, and Ransom-Jones *et al*. [57, 100]. This study aims to demonstrate that exposure of *F. succinogenes* to cellulose trigger the production and localisation of cellulolytic enzymes across the membrane of *F. succinogenes*.

Our reasoning was that expression of cell surface-membrane cellulolytic enzymes should bring out a change in the membrane chemistry of the bacterial cells which could be resolved using colloidal surface characterisation techniques such as EPM, MATH assay and FTIR. These techniques have been successfully used previously to study the cell surface macromolecular display in *E. coli* and *B.*

cereus [363]. In this study, the results of MATH assay suggest that the overall hydrophobicity of the cell surface decreases during growth on cellulose when compared to glucose. Although current literature suggests that CBPs are hydrophobic, the decrease in hydrophobicity is not surprising since the MATH assay measure the net surface characteristics of the cell and does not take into account cell surface heterogenicity. Moreover, a similar decrease in cell surface hydrophobicity upon exposure to cellulose has been previously observed in *Ruminococcus albus* [399]. *R. albus* possess cellulosome for cellulose degradation and pili-protein which considered to be mediates adhesion to the cellulose [400, 401] was also found in *F. succinogenes* in this study.

The net electrophoretic mobility of a bacterium at a given pH reflects the protonation/deprotonation state of the functional group of cell surface macromolecules. A change in the distribution of the cell surface macromolecules will reflect as a change in the EPM and the isoelectric point of the cell [363]. Accordingly, an increase in the IEP from pH 3-3.5 is seen for cell grown in the presence of cellulose compared to those grown in the presence of glucose. The increase in IEP may be a consequence of the decrease in cell surface polysaccharides and increase in the protein moieties as evidenced by FTIR spectra [362]. Taken together our results suggest that an increase in *F. succinogenes* cell surface-membrane proteins content occurs upon exposure to cellulose.

Previous studies on *F. succinogenes* provide essential information about surface proteins involved in adhesion and cellulose degradation [107, 110, 360, 402]. A comparative proteomics study carried out by Jun *et al*. [360], identified 16 proteins

specific to cellulose grown cells and considered as OM proteins, however, the localisation of them not specified. Moreover, genomic sequence analysis of *F. succinogenes* revealed an unexpected high number of carbohydrate-degrading enzymes, classified into 49 different families of GHs, CBMs, carbohydrate esterase (CEs) and PLs [43]. Considering the unclear mechanism of cellulose degradation and wide range proteins possibly involved, it was essential to study the cell surface-membrane proteome further.

The present study revealed 340 membrane associated proteins, of which only 34 proteins have been observed in previous studies. Particularly, out of 33 identified carbohydrate degrading proteins, 18 proteins was previously noted in different studies, out of which 6 were previously studied [360, 392-396] and 15 proteins were predicted to be involved in carbohydrate degradation on the basis of genomic annotation [43] (Table 5-1).

Nine proteins including family: CBM6-35 and PL10 (FSU_1114), CBM1-11, GH51 (FSU_0382), CBM30 (FSU_2007), GT51 (FSU_1715), GH94 (FSU_0162), GH8 (FSU_2303, Fisuc_1802, Fisuc_1219), GH9, CBP1 (FSU_0809), were commonly identified in all treatments. The identified CBM6-11-30-and 35 are known to bind to the single chain cellulose [43]. Our results showed agreement with the previously proposed mechanism where disruption of cellulose fibres and taking up individual cellulose chains in to the periplasm via OM may possible. However, occurrence of these proteins in all treatments (including glucose) suggests that the expression of these proteins may be substrate independent. Presence of signal peptides with CBMs identified suggests, enzymes may involve in synergetic degradation of other polysaccharides such as pectin and xylan [403-405].

Cellodextrin phosphorylase (FSU_0162) belongs to GH94 was found in all substrate conditions. It is not surprising, though production of cellodextrin was observed previously in glucose and cellobiose grown cells of *Fibrobacter succinogenes* by Wells *et al*. [102]. This study further concluded that production of cellodextrin occurs by reversible phosphorylase reaction to maintained energy equilibrium and to provide carbon source to the non-adherent cells. In anaerobes, cellodextrin phosphorylases catalyse ATP independent phosphorolysis reaction and microorganism can gain energy from phosphorolytic cleavage of β-glycosidic bonds when they cutting the cellodextrin chain [5, 406]. The protein with similar kind of activity i.e. GT51 (FSU_1715) was also found in all treatments, which involved in biosynthesis of disaccharide, oligosaccharides and polysaccharides by catalysing transfer of sugar moieties from active doner molecules [407]. The localisation of this protein in inner cyctoplasmic membrane further proved possibility of cellulose degradation takes place in periplasm of this bacterium. Three GH8 proteins found in all treatments, possess several known activities such as endoglucanase (EC 3.2.1.4), lichenase (EC 3.2.1.73) and chitosanase (EC 3.2.1. 132) [408]. Expression of GH8 protein in cellobiose (glucose dimer) grown cells have been observed previously in cellulose degrading bacterium *C. thermocellum* [409] indicating expression of GH8 enzymes are possibly substrate independant.

Fimbriae associated protein (FSU_0286) was observed commonly. associated with the formation of biofilm and involved in adhesion [410, 411], *F. succinogenes* is well known for adhesion to cellulose [360]. In addition to these, we also identified 7 different kinds of cellulose adherent proteins commonly in all treatment: Fibro-slime protein (FSU_2507, Fisuc_0377) [43], IV pilin (FSU_1212,

FSU_2567) [360, 401, 412], and Fibronectin type III domain protein (FSU_3194, Fisuc_2624) (only in glucose treatment) [413, 414]. This is the first study, identified such large number of proteins involved in adhesion mechanism. It is not surprising, since cellulose binding proteins were identified in glucose grown cells previously by Jun *et al*. [360].

Interestingly, 15 proteins were specifically found only in both cellulose treatments (MC and AS) and they include; seven GH5, three GH9, two GH45 and each of GT2, GHnc and PL1. Out of 10 predicted GH5 family proteins [43], 7 different cellulases were identified in cellulose grown cells suggested that cellulases belongs to GH5 family could be the major part of cellulose degrading enzymes in *F. succinogenes* and cellulose dependant expression is possible. GH5 family possess conserved glutamic acid residues which is potentially involved in catalytic mechanism [415], which provides thermal stability to enzymes GH5 family because of that it has potential advantages in future biofuel generation [416]. Out of three, GH9 enzymes, two proteins were found in only MC cellulose grown cells. The study on rumen cellulose degraders demonstrated that GH5 and GH9 enzymes are more versatile and have remarkable capability to degrade (MC) cellulose [417], also highest activity of GH5 was observed in *Thermofida fusca* [418]. Other important GHs were found in cellulose treatments are; O-glycoside hydrolase like protein (FSU_2795), Beta-galactosidase (FSU_0315) (has been predicted previously) from GHnc and GH2. We also identified unique enzyme pectate lyase/Amb allergen (Fisuc_2363) which is involved in pectin degradation. The subcellular localisation of these enzymes (PSORTb V3.0), across the membrane supports the proposed mechanism of cellulose degradation in *F. succinogenes* that proteins involved in adhesion and cellulose

degradation must lie in the different compartments of cell membrane and not only on surface.

Total 61 proteins which are involved in various membrane associated metabolic processes are shown in Table 2. It included; seven OmpA family proteins, six TPR domain proteins two MotA/TolQ/ExbB proton channel. MotA/TolQ/ExbB proton channels are integral membrane proteins which use a proton gradient to aid the transportation of large molecules across the membrane [419]. A total of 6 TPR domain proteins have been identified in this study, three of those were specifically found in cellulose grown cells. This is the first study so far which identified distinct TPR domains in such large number in this bacterium, which plays important role in protein-protein interaction and multi-protein complex formation [420]. Two of these TPR domains (FSU_2397, FSU_2398) have been observed previously in *F. succinogenes* [69, 360]. PEGA domain protein (FSU_1783) is S-layer protein, denoted as glycoprotein part of cell envelope in many gram negative bacteria, only found in glucose grown cells [421].

Several transporter family proteins were identified in this study; including ABC transporter and efflux transporter proteins (RND family MFP subunit and putative transporter protein (FSU_1591)). ABC transporters are involved in import and export of substrates across the membrane. Particularly, efflux transporter proteins are belongs to membrane fusion protein (MFP) family, facilitates outward flow of the substances from the cell [422]. Interestingly, most of these transporter proteins were identified in cellulose treatment may suggest that degradation of cellulose occurs in periplasm and hydrolysed by-products exported out via MFPs from cell. [57, 100, 423]. The presence of several channel and transporter

proteins suggests that *F. succinogenes* possess an active transport system across the membrane.

Further, 147 proteins were identified with unknown functions (Appendix 5 2). Out of 147 proteins, only 17 (glucose), 50 (common), and 63 were identified in both cellulose (AS and MC) treatments. The number of proteins is comparatively higher in cellulose grown cells than in glucose (control) treatment, which leads us to conclude that these hypothetical proteins may be new families of proteins involved in cellulose adhesion and degradation.

In conclusion, the overall surface-membrane proteome studies revealed that different substrate conditions influence membrane protein expressions, especially those which are involved in adhesion and cellulose degradation mechanism. The results obtained from surface physicochemical analysis show a correlation with the surface-membrane protein analysis. This indicates that membrane modifications occur during growth on cellulose substrates. Decreased hydrophobicity, the suppression of the cell surface polysaccharide production, an increase in cell surface protein display and a positive EPM below pH 3, reflect the changes in surface characteristics of *F. succinogenes* S85 in cellulose grown cells. Furthermore, the identification of surface-membrane proteins of *F. succinogenes* S85 by biotinylation method will provide new insight for future studies on the Fibrobacter-cellulose interactions and subsequent cellulose degradation. By using membrane selective proteomics method (membrane biotinylation), we identified 340 membrane associated proteins. The numbers of membrane associated proteins identified in this study are far better than the proteins identified in the previous study (25 differential expressed proteins). Out of those 33 associated with cellulose degradation (Table 5-1). Seven cellulases

belongs to GH5 family were identified in only cellulose treatment. This suggested that GH5 family proteins are the major proteins in cellulose degradation in *F. succinogenes*. Interestingly some of them possess thermal stability thus can have major advantage in future biotechnological applications. Some of these proteins are localised within periplasmic membrane, provides vital clue towards proposed mechanism of cellulose degradation (degradation takes place in periplasmic space). We also identified large number of proteins involved in initial adhesion of cells to cellulose such as fimbriae associated proteins, Fibro slime proteins, IV pilin, Fibronectin type III domain proteins. Some of them also found in glucose grown cells (such as Fibronectin type III domain) suggested that expression of this protein is substrate independent. The study also identified MFP efflux proteins in cellulose treatment further provides vital clue about the degradation takes place within membrane of cells since MFP efflux proteins exports sugar to the environment.

However, to understand the adhesion mechanism of cells-cellulose and subsequent cellulose degradation in further detail, future studies in our group, will be aimed at high through-put quantitative proteomic study of surface and membrane proteins.

Chapter 6

Quantitative membrane proteome of *Fibrobacter succinogenes* S85: A comparison between enrichment methods and biological interpretation

6.1 Introduction

As discussed in chapter 5, the membrane biotinylation method identified crucial proteins involved in cellulose degradation and other related membrane activities. To gain an in-depth understanding of the mechanism of cellulose degradation and associated membrane activities, we have extended this study to a membrane quantitative proteomic level. At same time, two protein enrichment methods were compared, i.e. membrane biotinylation and membrane isolation by ultracentrifugation.

As described in the previous chapter (5.1), the microbial ecosystem in the rumen is well adapted to degrade major plant cell components such as cellulose, hemicellulose, and starch. The most cellulose-degrading genera are *Rumonococcus*, *Prevotella* and *Butyrivibrio* [64]. Among them, *F. succinogenes* S85 is a cellulose degrading bacterium of the rumen ecosystem classified into the unique phylum of Fibrobacters. It has developed the ability to degrade crystalline cellulose. Moreover, *F. succinogenes* is a symbiotically favourable bacterium for other cellulolytic, non-cellulolytic bacteria and for the host organism, since its cellulose degrading by-products such as glucose and cellodextrin are exported out for cross feeding [424, 425]. Debatably, *F. succinogenes* does not use the most well-known models for cellulose degradation systems: either a free enzyme system (eg, in *Tricoderma reesei*) or bacterial cellulosomal system (eg, in *Clostridium thermocellum*). It, consequently, has forced researchers to consider alternative models for cellulose degradation [43] that lie within the membrane of *F. succinogenes*.

To achieve this, proteomics is the most suitable approach for the study of *F. succinogenes* at the biological level. Over the last decades, proteomics has grown exponentially; it has become an established scientific discipline and a major driving force in biological research alongside genomics. Today, mass spectrometry (MS) based proteomics is a robust and extremely valuable technique in the study of complex biological systems [426].

In recent years, MS-based proteomics has first been successfully used to qualitatively characterise complex protein mixtures [207]; it has now evolved into the more advanced form of quantitative proteomics [426]. Quantitative proteomics allows us to identify differences between different physiological states of biological systems, thus key protein change reflecting cellular states [208]. At present, there are two quantitative methods that are widely used in shotgun proteomics approaches; 1) “isotopic labelling” that relies on labelling of samples with different isotopes and then combined for MS analysis and 2) “label free” that relies on spectral counting, peak area and peak intensity [427]. Isobaric tags for relative and absolute quantitation (iTRAQ) is the most commonly used isotope labelling approach in proteomics. Nonetheless, up till now, most of the approaches are applied to soluble proteomics [428], but in principle can be applied to membrane proteomics too. Membrane proteins carry out many fundamental biological processes (protein channelling, signalling, transportation and reception) [429], but their study remains challenging due to their hydrophobic nature and low abundance [430]. Consequently, there has been a considerable effort to develop the strategies applied for enriching, isolating and separating membrane proteins in order to study the membrane proteins’ function [428].

Since the cellulose degradation system of *F. succinogenes* is assumed to lie in the membrane, we sought to isolate the membrane proteins for further quantitative proteomics. We grew *F. succinogenes* on glucose and cellulose as sole carbon sources. Substrates can lead to induced protein expression within the membrane, which can be quantitatively characterised. In this study, iTRAQ based shotgun proteomics was applied to the enriched membrane proteins facilitated by ultracentrifugation and biotinylation. The results obtained from both the techniques are compared and significant changes in protein expression are discussed.

6.2 Materials and methods

6.2.1 Bacterial growth and culture conditions

All chemical were purchased from Sigma Aldrich (UK) otherwise mentioned. *F. succinogenes* S85 ATCC 19169 was kindly provided by Prof. Paul Weimer (US Dairy Forage Research Centre, Wisconsin, USA). *F. succinogenes* was anaerobically maintained as described in Chapter 5 following the procedure reported by Weimer *et al.*and Callaway and Martin [364, 365]. Briefly, the basal media was prepared by addition of the following components: 0.9 g L^{-1} KH_2PO_4, 3.2 g L^{-1} Na_2CO_3, 1 g L^{-1} cysteine·HCl, 0.9 g L^{-1} NaCl, 0.9 g L^{-1} (NH_4) $2SO_4$, 0.084 g L^{-1} $MgCl_2{\cdot}_6H_2O$, 0.065 g L^{-1} $CaCl_2{\cdot}2H_2O$, 0.0275 g of $MnCl_2{\cdot}4H_2O$, 0.02 g L^{-1} $FeSO_4{\cdot}7H_2O$, 0.009 g L^{-1} $ZnCl_2$, and 0.0048 g L^{-1} $CoCl_2{\cdot}6H_2O$. In addition, 0.06 g L^{-1} of each of the VFA (isobutyric acid, isovaleric acid, n-valeric acid and 2-methylbutyric acid) was added to the medium. Finally, 10 mL L^{-1} of Schaefer's vitamin solution (preparation is given in Chapter 5, section 5.2.1). Oxygen was removed from the media, 125 mL serum bottles, by sparging with 100 % CO_2 during and then autoclaved (as described in Chapter 3, section 3.2.1).

Media was inoculated, under anaerobic conditions, with 1 mL of culture (corresponding exponential phase, OD_{675} = 0.8) which was grown in media containing glucose as carbon source. Cultures were incubated at 38^{o}C and growth was monitored at OD_{675nm} using a Ultraspec spectrophotometer (Model 2100 pro, Amersham Bioscience). For the biological treatment conditions, 3 g L^{1} glucose and 3 g L^{-1} cellulose were used as a carbon source. Two sets of duplicate culture of glucose and cellulose grown cells were harvested at the mid exponential phase, depending on the growth rate of the bacterial strain under different substrate conditions.

6.2.2 Isolation of cells

Cells were harvested at the mid exponential phase by centrifugation at 8000 g for 10 min at 4 °C. In the next step, cell pellets were washed thrice with phosphate buffered saline (PBS) (containing $MgCl_2$) and repelleted by centrifugation at 8000 g for 5 min at 4 °C. In cellulose grown cells, additional steps were performed (as described in Chapter 5, section 5.2.1) to obtain cellulose adherent cells. For that, cellulose bounded cells were first separated by centrifugation at 500 g for 5 min and then treated with 1 g L^{-1} methyl cellulose solution in M8 buffer solution (composition is given in Chapter 5, section 5.2.1) by incubation at 38^{o}C for 30 min (repeated thrice) as suggested by Kudo *et al.* and Olsen and Mathiesen [98, 366]. Extracted cells were further centrifuged and washed again with PBS solution and mixed with the previously collected cell pellets. Cell pellets obtained from both substrate conditions were further resuspended in the PBS buffer and OD_{675} adjusted to 2.00 (total volume 20 mL).

After cell biomass harvesting, samples were processed further as per the protocol for ultracentrifugation membrane isolation and membrane protein isolation by biotinylation.

6.2.3 Isolation of membrane proteins by ultracentrifugation method

The cell pellet was resuspended in lysis buffer (20 mL) and lysis was achieved by passing cells through a French press (Thermo Electron Corporation) at 1000 psi. This process was repeated twice to improve cell lysis. Unbroken cells were separated by centrifugation at 3500 g for 5 min. The collected supernatant was ultracentrifuged (Beckman, serial no. COL97F14) at 40,000 rpm for 180 min (Temperature). Pellets (enriched membrane fraction) were washed 3-4 times with PBS buffer to reduce contamination of soluble proteins. The membrane fraction was delipidated with chloroform/methanol as described by Wessel and Flugge [431]. In the next step, membrane proteins were further purified using ice cold acetone (1:4 ratio V/V). Finally, pellets were air dried and re-suspended in the 0.5 M tryethylammonium bicarbonate (TEAB) (pH8.5) buffer containing 0.1 % RapGest and dissolved by sonicating for 5 min in an iced water bath. Proteins were quantified using a Bradford assay method as described by the manufacturer (Sigma Aldrich). Ten micrograms of each sample were run on a 1D-PAGE gel and were in-gel and in-solution digested using a (1:50 ratio trypsin to protein (w/w) to check the digestion efficiency.

6.2.4 Isolation of membrane proteins by biotinylation enrichment method

Biotinylation enrichment of proteins were performed as previously described by Ge and Rikihisa [228]. To obtained membrane proteins at bulk amount, 20 mL (OD_{675}=2.00) of cell suspension in PBS was centrifuged at 8000 g for 5 min at

4^{o}C and pellets re-suspended in the 4 mL PBS buffer (1 mM $MgCl_2$) containing 30 mg EZ-Link® Sulfo-NHS-SS-biotin labels (Thermo, Pierce). The mixture was incubated at 4^{o}C for 30 minutes and excess biotin was then quenched thrice by washing with 4 mL of 500 mM glycine-PBS solution. Biotin labelled cell pellets re-suspended in the 4 mL of radioimmunoprecipitation assay buffer (RIPA) (25 mM Tris-HCl (pH 7.6), 150 mM NaCl, 1 % NP-40, 1 % sodium deoxycholate, 0.1 % sodium dodecyl sulphate (w/v), 1.1000 dilution of protease inhibitor cocktail set II). Cell lysate was obtained by brief sonication (30 second sonication 1min on ice; 8 cycles). Cell lysates were incubated on ice for 30min with gentle occasional vortexing. At this stage additional oxidised glutathione (100µM) was added to the lysate to protect disulphide bond in the Sulfo-NHS-SS-biotin. Lysates were further centrifuged at 16000 g for 10 min at 4^{o}C and the supernatant was collected.

6.2.5 Neutravidin affinity purification of biotinylated proteins

Six millilitres of neutravidin agarose gel was washed three times with wash buffer containing 25 mM Tris-HCl (pH 7.6), 0.15 M NaCl, 0.5 % NP40, 0.5 % sodium deoxycholate, 0.05 % SDS. The cell lysate was mixed with washed neutravidin agarose gel and incubated on ice for 2 hours with gentle shaking. Protein-neutravidin agarose gel was centrifuged at 500 g for 1 min and the supernatant was discarded. Gel slurry with biotinylated proteins were transferred to the column (10 mL capacity empty Zeba™ Spin column Thermo Scientific Rockford, USA). Unbound proteins were removed by washing with buffer (25 mM Tris-HCl [pH 7.6], 0.65 M NaCl, 0.1 % NP40) (twice), followed by washing with buffer (25 mM Tris HCl [pH7.6], 1.15 M NaCl, 0.1 % NP-40) and finally with Tris buffer (25 mM Tris-HCl [pH 7.6], 0.15 M NaCl) at 200 g for 15-20 sec. Gel bound proteins were eluted thrice with 5 % 2-mercaptoethanol-PBS at 30^{o}C for 30 min. Proteins

were precipitated by 10 % tricloroacetic acid (TCA) and centrifuged at 18,000 g for 10 min at 4ºC [380]. The protein pellets were finally washed with ice cold acetone and air dried. The purified proteins then re-dissolved in 0.5 M TEAB buffer containing 0.1 % RapiGest. A quantification of membrane protein was done by the Bradford protein assay method according to the manufacturer's protocol (Sigma Aldrich). Briefly, 20 µL of protein sample (1:1 dilution) was mixed with 980 µL of Bradford reagent and incubated for 10 min at room temperature. Absorbance was recorded at 595 nm and concentration determined using a standard calibration where BSA was used.

6.2.6 In-gel digestion of proteins

In order to achieve denaturation and efficient digestion of hydrophobic proteins (membrane proteins), we chose an in-gel digestion method for proteolysis. SDS-PAGE was performed on duplicate samples of proteins obtained from both techniques as per the protocol described by Sambrook and Russel [432]. Briefly, 50 µg of each sample was mixed with sample buffer and boiled at 95ºC for 5 min and loaded to the gel and allowed to run until reaching to the resolving gel. After reaching the resolving gel, a single band (concentrated protein region) for each sample was sliced carefully from the gel and placed in new Eppendorf tubes. The following procedure was performed as previously described by Karunakaran *et al.* [255] with a few modifications in the protocol (as described in Chapter 3). Briefly, instead of ammonium bicarbonate (AB) we used TEAB buffer for the further steps. Protein bands were distained twice with 400 µL of 200 mM TEAB in 40 % acetonitrile (ACN) in water by incubating at 37ºC for 30 min. The supernatant was discarded and gel pieces dried in a vacuum concentrator (Vacuum concentrator 5301, Eppendorf, UK) for approximately 5 min at 30ºC.

Entrapped gel proteins were reduced using 400 µL of HPLC grade water containing 1µL of 50 mM tris 2-carboxyethyl phosphine hydrochloride (TCEP) by incubating at 60^{o}C for 1 hour. Gel pieces were briefly centrifuged at 13,000 g for 10 s and all liquid was discarded. In the subsequent step, proteins alkylated using 400 µL alkylation buffer containing 1µl of 200 mM methyl methanethiosulfonate (MMTS) at room temperature for 30 min. Gel pieces were washed twice with 400 µL of 50 mM TEAB solution by incubation at room temperature for 15 min. Finally they were once washed once with 400 µL of 50 mM TEAB in 50 % ACN for 15 min by incubation at 37^{o}C. After incubation, samples were briefly centrifuged at 13,000 g for 10 min and all liquid was discarded. Gel pieces subsequently were dried in a vacuum concentrator for approximately 15-30 min at 30^{o}C.In the next step, proteins were digested with trypsin at a trypsin/protein ratio of 1:50 (w/w) (Applied Biosystems, USA) in 200 µL of 40 mM TEAB in 9 % (v/v) ACN for approximately 16 hours by incubation at 37^{o}C. At this stage, 0.1 % (v/v) of RapiGest was added. After digestion, samples were centrifuged briefly at 13,000 g for 10 sec and supernatant was collected in new Eppendorf tube. Peptides were extracted twice with 100 µL of 5 % (v/v) formic acid and once with 50 µL of 100 % ACN. Finally, all the supernatants were combined together and vacuum dried (Scanvac; module speen 40, Lynge, Denmark) and stored at -20^{o}C until further iTRAQ labelling and LC-MS/MS analysis.

6.2.7 iTRAQ labelling and HILIC fractionation

iTRAQ 8-plex labelling was performed on all the samples according to the manufacturer's protocol (8-plex iTRAQ reagent Multiplex kit, ABSciences, USA). The relevant iTRAQ reagent re-suspended to a 50 µL of isopropyl alcohol and

added to each tube by taking care of which reagent added to which tube. A detailed workflow of the experimental procedure of sample preparation and iTRAQ labelling is given in Figure 6-1. After labelling samples were vortexed, spun down, and incubated at room temperature for 1 hour. In the next step, the contents of all tubes were combined in new clean protein low binding tube vortexed and spun down. Labelled samples were acidified with TFA (final concentration of solution 0.5 %) to precipitate out RapiGest and supernatant was obtained by centrifugation at 17,000 g for 5 min at 4^{o}C. These supernatant containing labelled peptides was split into two parts and dried using a vacuum concentrator (Scanvac; module speen 40, Lynge, Denmark).

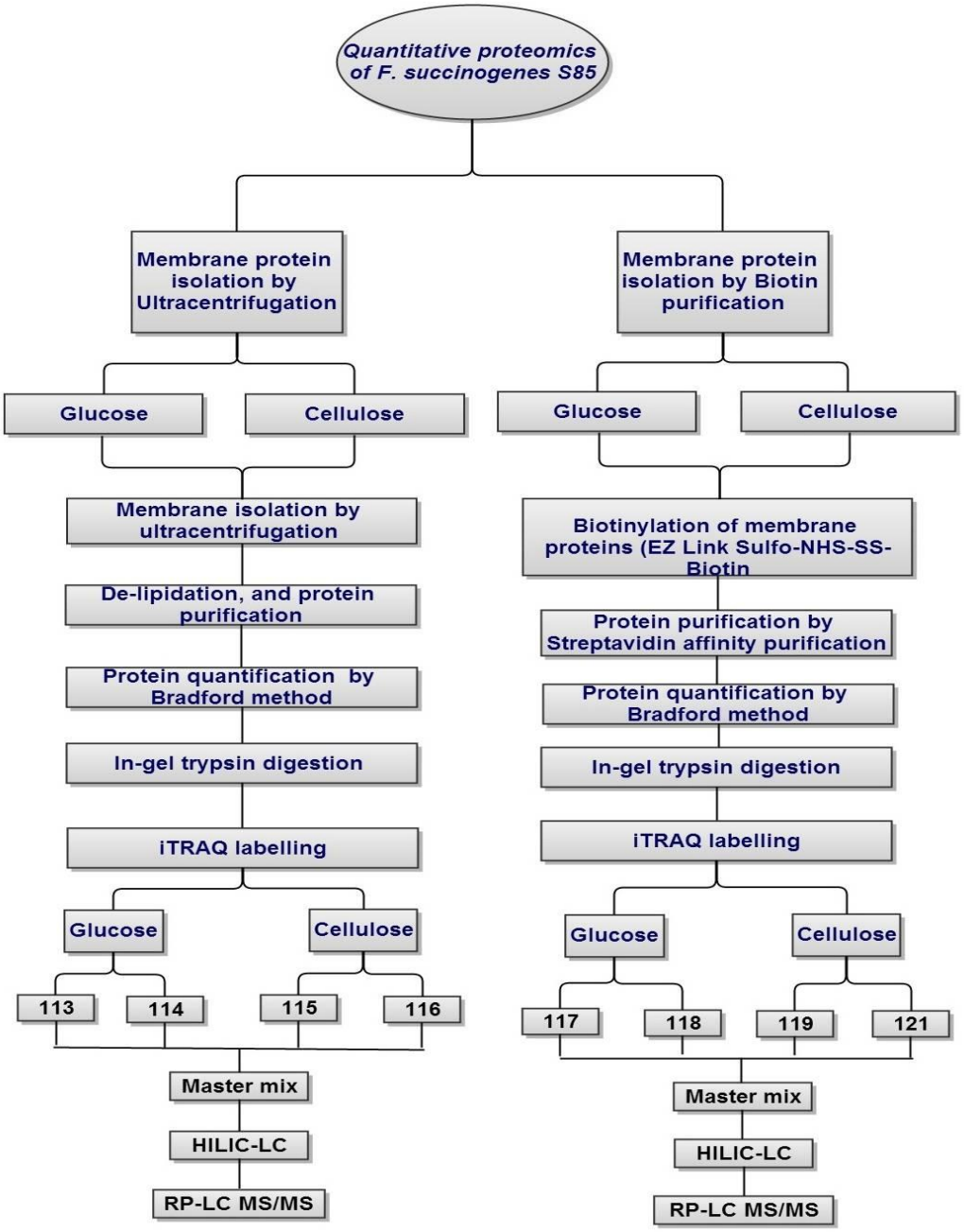

Figure 6-1 Quantitative proteomics workflow used

Once dried, iTRAQ-labelled peptides were re-suspended in 100 µL of hydrophilic interaction liquid chromatography HILIC buffer A (80 % acetonitrile (ACN), 10 mM ammonium formate, pH 3) for fractionation. The sample was centrifuged at 17,000 *g* for 5 min and injected into an Ultimate 3000 HPLC (Dionex Germering,

Germany) coupled with a HILIC column (PolyHYDROXYETHYL-A, column, 5 µm pore size, 100 mm length, 4.6 mm ID, PolyLC Columbia, MD, USA) at a flow rate of 0.5 mL/min using an UV detector at 280nm (Dionex, UVD170U). The following buffers were used: buffer A (80 % ACN, 10 mM ammonium formate, pH 3) and buffer B (5 % ACN, 10 mM ammonium formate, pH 5). The 75 min gradient consisted of 0 % Bfor 10 min, 0-60 % B for 40 min, 60-100 % for 5 min, 100 % B for 10 min, and 100 % A for 10 min. Fractionation and chromatography was monitored through Chromeleon software (Dionex/LC Packings,The Netherlands). Fractions were collected every 30 second from fractions 20-60. Collected fractions were then dried in a vacuum concetrator (Genevac Ltd Suffolk, UK) and subjected to further LC-MS analysis. Further C_{18} clean up of samples were performed to make sure that samples were free from salts.

The procedure and parameters were as described in Chapter 4 (section 4.2.6). The selected fractions were re-dissolved in 22 µL reverse phase (RP) buffer A (3 % ACN, 0.1 % FA) for further LC-MS/MS analysis before submitting to a QStar XL Hybrid ESI quadrapole time-of-flight tandem mass spectrometer (ABSciex, Concord, Ontario Canada) coupled with an online HPLC system (Ultimate 3000, Dionex, Surrey UK). 10 µL of each selected fraction was injected into the nano LC-ESI-MS/MS system, peptide separation was performed by an Acclaim® PepMap100 column (C-18, 3 µm, 100Å, 15 cm) at a constant flow rate of 300 nL/min. The buffers used in the liquid chromatography were buffer A (3 % ACN with 0.1 % FA), and RP buffer B (97 % ACN with 0.1 % FA) and the 120 min gradient was used as follows: 0-3 % B for 5 min, 3-35 % B for 90 min, 35-90 % of B for 0.5 min, 90 % of B for 6.5 min, finally 3 % of buffer B for 18 min. MS data was acquired on the Analyst® QS 2.0 software (Applied Biosystems, USA) in a

data-dependent acquisition mode. Peptides of charge +2, +3, +4 (intensity binning) for each TOF-MS scan (400-1250 *m/z*) were dynamically selected and isolated for MS/MS fragment ion scans (100-1600 *m/z*).

6.2.8 Data Interpretation and protein identification

The generated tandem MS data files (wiff) from the QSTAR XL were subjected to ProteinPilot software v4.0 (AB Sciex Famingham MA) which performed both peak picking and the database search. The search was performed within the *F. succinogenes* S85 (taxon ID: 59374) database containing 3815 protein sequences downloaded from Uniprot (October 2013). Searches were conducted against a forward/reverse concatenated database to determine the false positive discovery rate (FDR). Protein identification was accepted as positive based on a probability filter cut-off of 95 %. The reporter ions intensities were then exported to Excel for further analysis. Isotopic and median corrections were applied using an in-house automated method as described in Ow *et al.* [433] and protein quantification values were obtained in log space. Further, these data were analysed using a method described by Pham *et al*. [297] with significant changes ($\alpha = 0.05$).

6.3 Results and Discussion

iTRAQ sample preparation procedure was followed by HILIC fractionation before proceeding to RPLC-MS/MS analysis. Figure 6-2 shows that peptides effectively eluted out between 20 to 40 min. There were 6 dominant peaks observed in this period of time. A total of 44 selected fractions (30 s per fraction) between 22 to 44 minutes were injected into the LC-MS/MS.

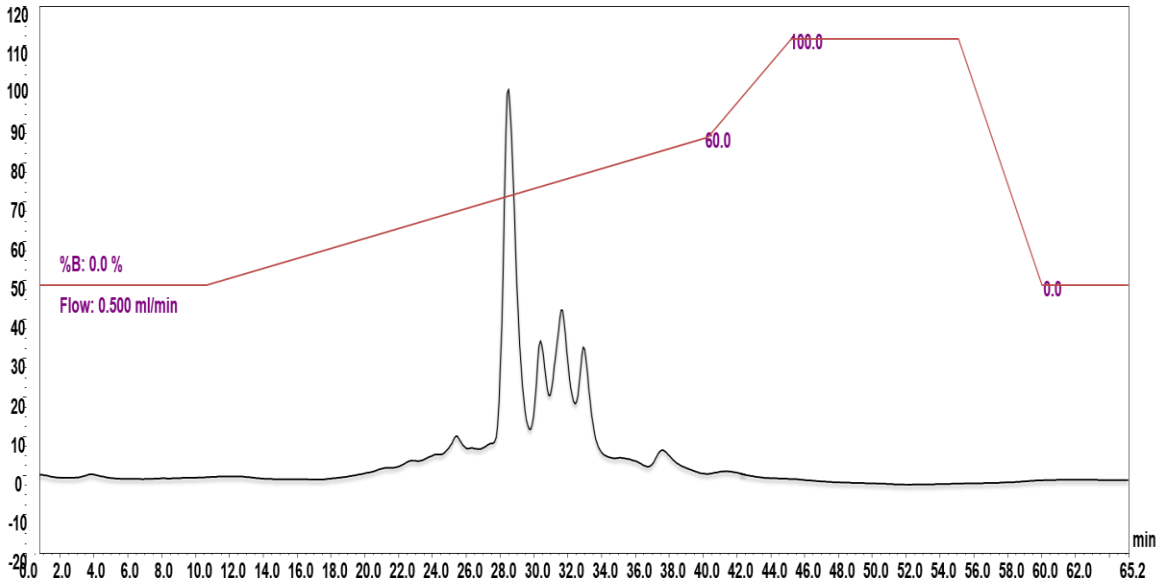

Figure 6-2 HILIC chromatogram of resultant peptide fractionation. The x axis represents the run time, and the y-axis represents peptide intensity monitored at 280 nm.

In the following sections, significant changes in the proteome obtained from two different techniques were compared on the technical and biological level.

6.3.1 Quantitative proteomics and regulation of proteins among the two membrane enrichment methods.

In this study, we identified 7801 peptides, at a 1.77 % false discovery rate (FDR). A total 3273 unique peptides sequences were identified. Out of 617 unique proteins, 437 proteins were quantified with two or more unique peptides and 472 proteins were quantified with two or more MS/MS scans [303].

The experiment was designed by supposing that the proteome profiles obtained from both methods would be comparable, the iTRAQ quantitation works when the samples are broadly similar. We checked the validity of this hypothesis: the 1D-

SDS-page gel compared the proteome profiles obtained using both methods (shown in Figure 6-3).

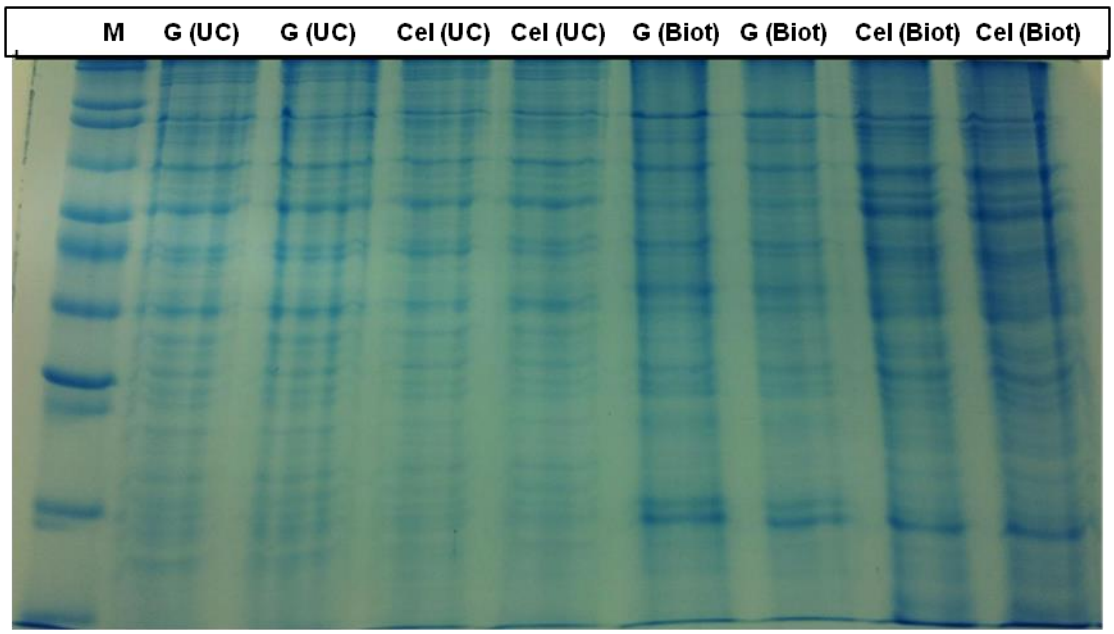

Figure 6-3 1D SDS-PAGE electrophoresis membrane protein samples (10μg each) Lane 1, Marker, Lane 2-3 Glucose (ultracentrifugation), Lane 4-5 Cellulose (ultracentrifugation), Lane 6-7 Glucose (biotinylation) Lane 8-9 Cellulose (biotinylation).

Based on our statistical analysis, significantly abundant proteins between glucose and cellulose grown cells obtained from ultracentrifugation and biotinylation were compared and discussed in following sections.

Our study revealed that significant changes were found in a total 275 proteins obtained by biotinylation, in 143 proteins obtained by ultracentrifugation and in 347 proteins in combined derived data (when cellulose grown cells were compared with glucose grown cells). A total of 70 proteins were found to be differentially regulated in common in both techniques. The protein numbers noted

in this section include cytosolic proteins. In total, 74 (51 % of 143) and 137 (49.6 % of 275) cytosolic proteins were identified in the ultracentrifugation and biotinylation enrichment methods respectively. Despite continuous improvement in membrane proteomics, cytosolic protein contamination is still the single biggest challenge in membrane proteomics that causes underrepresentation of membrane proteins in a proteome analysis [434]. Figure 6-4 shows the total number of proteins with significant changes obtained from each enrichment technique.

Therefore, the identification of cytosolic proteins in this study is not surprising. The commonly used techniques for cell membrane proteome isolation are based on sub cellular fractionation (such as ultracentrifugation). Therefore, these techniques always have been noted with cytosolic protein contamination [435-438]. On the other hand, biotinylation works on the basis of water soluble reagents that are presumably impermeable to cell membrane and only targets to membrane proteins. However, in spite of successful identification of membrane proteins, a significant amount of cytosolic proteins were also been observed previously [439, 440].

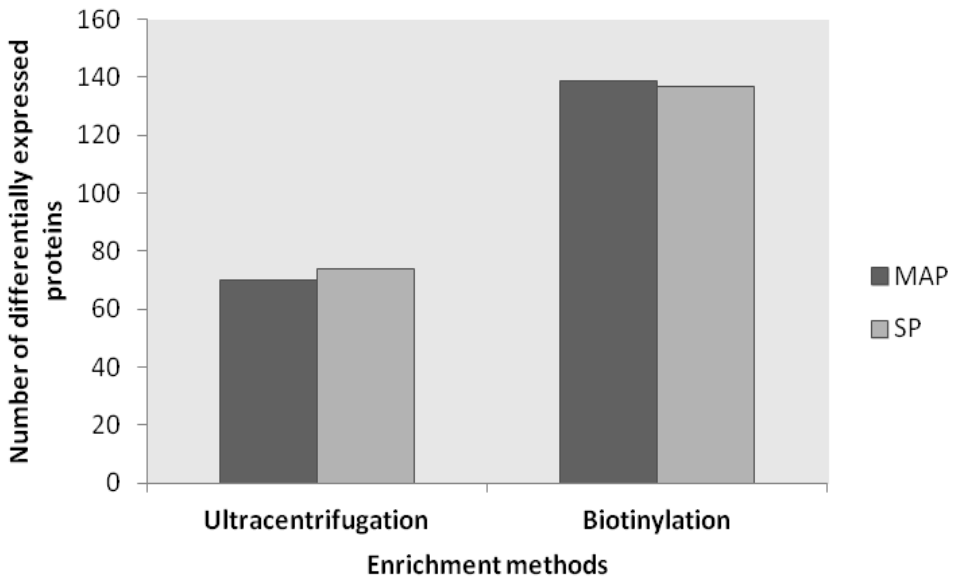

Figure 6-4 Relative abundance of proteins among the enrichment techniques and its influence on number of identified proteins. (▫ SP; Soluble proteins ▪ MAP; Membrane associated proteins)

As we are interested in membrane proteomics, in the following section, we discuss membrane proteome. In this study, biotinylation identified a promising number of membrane associated proteins (MAPs) with significant changes compared to ultracentrifugation (138 vs 70 membrane proteins). Out of those 138 MAP, only 28 proteins were found in low abundance. On the other hand, in ultracentrifugation, out of 70 MAP, 30 MAP were found in low abundance. In addition to that, 41 MAP with significant changes were commonly found in both techniques shown in Figure 6-5. A total quantified MAPs with significant changes in ultracentrifugation and biotinylation enrichment method is shown in Appendix - 6.1 &Table 6-2.

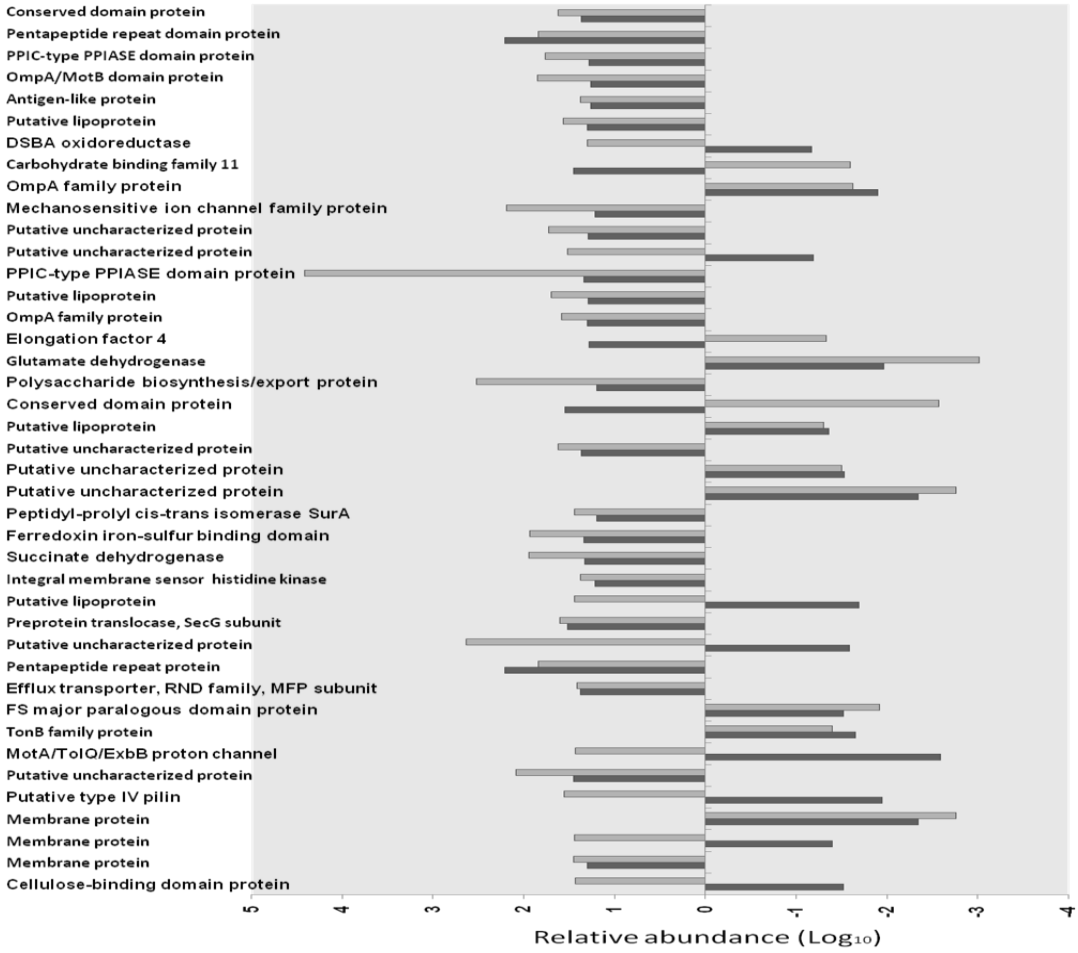

Figure 6-5 Relative abundance of identified MAP commonly obtained in both enrichment methods. (■) represents ultracentrifugation and (■) represents biotinylation.

The use of biotinylation in enrichment of membrane proteins is supported by a 41 % increase in significantly changed proteins as compared to ultracentrifugation. The difference observed between the two enrichment techniques and increased membrane proteins in biotinylation can be attributed to the affinity and targeted enrichment of membrane proteins by biotinylation [441] over traditional ultracentrifugation.

Therefore, use of biotinylation enrichment followed by quantitation supported the membrane proteomics study over the traditional membrane isolation method. Biotinylation is a well established membrane purification method therefore difference can be credited partially to the specificity of the method. Although the comparison between these two techniques is not main focus of the idea, however, the difference noted in this observation suggests how these techniques are optimised when applied within an iTRAQ proteomics workflow.

In this study, we determined cellular allocation of proteins by using PSORTb v 3.0 [390] software that helped us to identify outer membrane, cytoplasmic membrane, periplasmic, extracellular proteins and proteins with unknown location. The submembrane distributions of proteins were compared between ultracentrifugation and biotinylation (Figure 6-6). The results indicate that the outer membrane and unknown proteins (with multiple locations) were significantly expressed in biotinylation method over the traditional membrane isolation method (ultracentrifugation). The potential of biotinylation in identification outer membrane proteins were demonstrated in various studies [224-227]. However, to date, this is the first study so far which has compared the proteome by biotinylation on a quantitative level and compared to a traditional membrane protein isolation technique. Moreover, the highest number of significantly regulated unknown proteins in biotinylation probably indicates significant co-localisation of proteins across the membrane (inner membrane to surface).

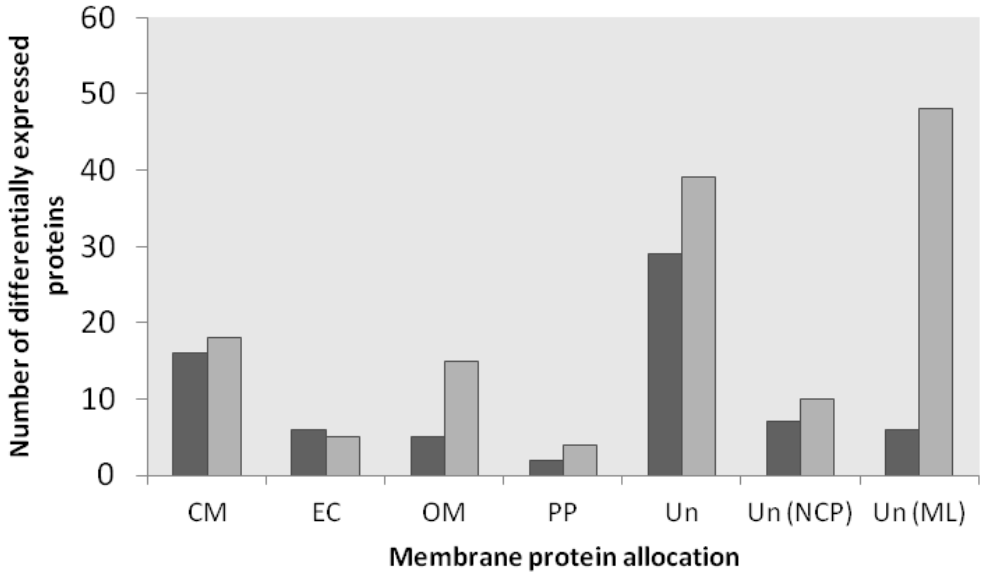

Figure 6-6 Distribution of proteins (membrane associated) with changes in relative abundance in ultracentrifugation and biotinylation enrichment method (■ represents ultracentrifugation and ▒ represents biotinylation). The sublocalisation of proteins in the membranes were predicted by tool PSORTb v 3.0: CM; Cytoplasmic membrane, EC; Extracellular, OM; Outer membrane, PP; Periplasmic membrane, Un; Unknown location, Un (NCP); Unknown non-cytoplasmic and UN (ML); Unknown multiple location.

The low/high abundance of MAPs was compared between the two enrichment techniques depending on co-localisation of proteins within membrane (Figure 6-7). The cytoplasmic membrane (CM) proteins and extra cellular (EC) proteins do not show much difference in both techniques. However, low abundance EC proteins were not observed in the ultracentrifugation enrichment. The significant recovery of high abundance proteins of the outer membrane (OM), and unknown (non-cytoplasmic) (UN NCP) proteins were observed in the biotinylation technique. The observation quite obvious for OM proteins in biotinylation since the biotin reagent preferably targets OM proteins [442]. The recovery of high abundance proteins in biotinylation is almost greater in all the sub-membrane locations. The biotin reagent can penetrate the cell membrane and label the

proteins in various sub membrane location since *F. succinogenes* posseses active transport system that helps to labele proteins through out membrane.

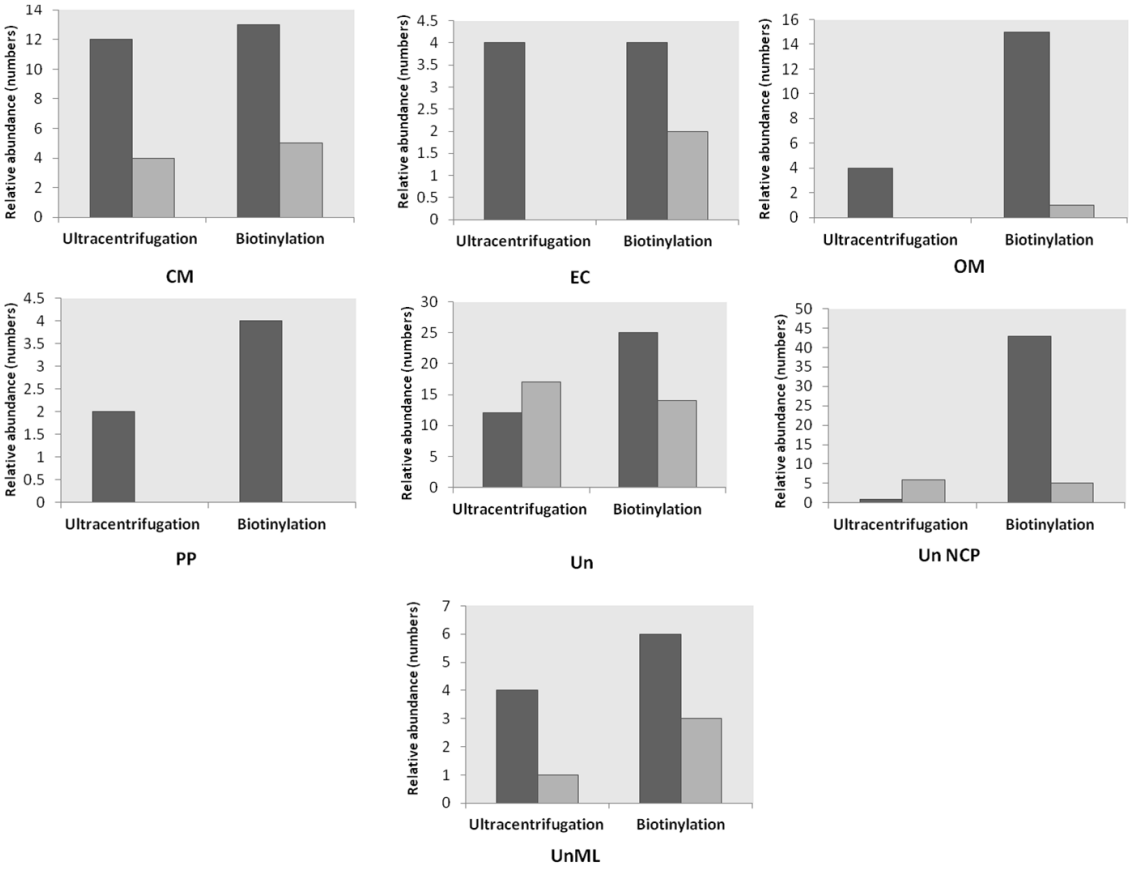

Figure 6-7 Distribution of relatively abundant protein numbers based on membrane localisation (■ represents High abundance and ■ represents Low abundance) The proteins are from different membrane locations: CM; Cytoplasmic membrane, EC; Extracellular, OM; Outer membrane, PP; Periplasmic membrane, Un; Unknown location, Un (NCP); Unknown non-cytoplasmic and UN (ML); Unknown multiple location.

6.3.2 Biological interpretations

Based on the efficiency of the biotinylation enrichment method as described in the technical comparison, we considered IDs obtained from biotinylation-iTRAQ data for biological interpretation. Further analysis was carried out based on significant changes in proteins in different substrate conditions (glucose and cellulose) based on a different biological function.

6.3.2.1 Carbohydrate degradation

It is hypothysed that a complete degradation of cellulose is a membrane associated process which is based on three main steps: 1) adhesion of cells to cellulose with the help of outer membrane proteins, 2) disruption of cellulose particles into microfibrils and transportation to the periplasmic space 3) hydrolysis of cellulose microfibrils and transportion of sugars across the membrane [57]. As proposed, various membrane associated activities are involved in cellulose degradation, and we discuss proteins are involved in cellulose adhesion/degradation, transportation, protein chanelling and proteins involved in cell signalling.

In this study, we identifed total 10 proteins involved in cellulose degradation with significant changes among the subsrate conditions. Figure 6-8 shows the relative abundance of proteins associated with carbohydrate degradation. A fibro-slime protein (Fisuc_1525) thought to be involved in adhesion mechanism [43] was found in high abundace in cellulose grown cells (compared to glucose grown cells). In previous a qualitative proteomics study, this protein was observed in glucose and cellulose conditions (Chapter 5, Table5-1). Although, genomic annotation suggests presence of 10 paralogs of this protein derived from different genes. High abundance of this perticular fibro-slime protein (Fisuc_1525) indicates that it is probably the main protein involved in a adhesion mechanism. There were three proteins belonging to the glycoside hydrolase (GH) family, namely putative glycoside hydrolase (GH98, CBM51) (Fisuc_0401), Chitinase (GH18, CBM9), (Fisuc_1530), and putative glycoside hydrolase (GH 38, 57)

(Fisuc_3030) were found in high abundance in cellulose grown cells. The GH98 proteins possess a novel putatve carbohydrate binding module (npCBM) probably operated with retention of substrate configuration [443]. According to a genome annotation of *F. succinogenes,* chitinase proteins belonging to GH18 may have different catalytic activities including cellulose binding activity [444]. The cellulose binding module (CBM), belongs to CBM30 (Fisuc_1525) and Lys domain protein belongs to CBM50 (Fisuc_0433) these were also found in high abundance in cellulose grown cells. CBM30 plays provital role in binding to single chain cellulose (microfibrils) [43]. Its ability to bind single chain cellulose and its different localisation (UN NCP) within the membrane supports cellulose chain transportation inside cell membrane and its degradation within membrane. Interstingly, CBM50 normally associated with various GH proteins belongs to GH18, GH19, GH24, GH25, and GH73 and involved in cleaving chitin or peptidoglycan [445]. The function of this protein in cellulose degradation is unclear. A cellodextrin phosphorylase (Fisuc_2900) that belongs to GH 94 was found in low abundance in cellulose grown cells. Our results show agreement with previous observations where the highest cellodextrin activity was found in glucose and cellobiose grown cells compared to cellulose grown cells [93, 102], since this enzyme only uses oligosachharides as a substrate [446].

There are three proteins related to cellulose degradation (Fisuc_3111, Fisuc_2704, Fisuc_1425) were found to be down regulated in cellulose treatment. It is not surprising, since proteins GHs and CBMs were identified in glucose grown cells in a previous study by Jun *et al*. [44]. Although, a genomic study of *F. succinogenes* suggests 49 different families of GH CBM polysachharide lyases (PL) in *F. succinogenes* [43], only a few of them have been found in this study.

This may indicates that glucose grown cells also produce GHs and CBMs and only the activity of certian proteins is triggered in the presence of cellulose.

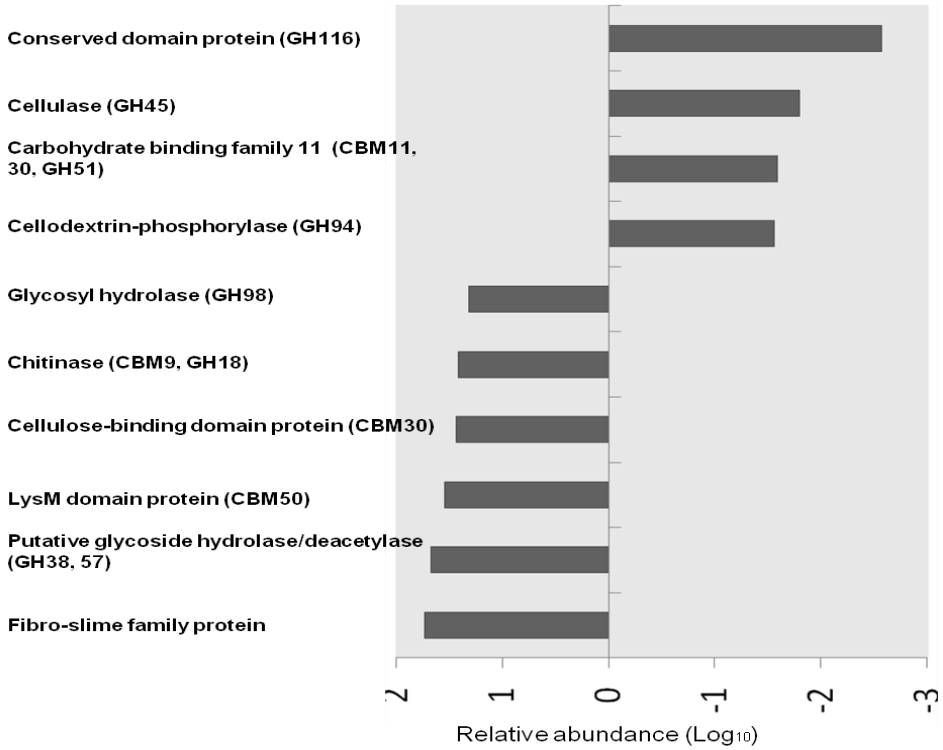

Figure 6-8 Relative abundance of identified carbohydrate degradation proteins between cellulose and glucose grown cells

6.3.2.2 Proteins involved in adhesion

There were several proteins identified in this study with adhesion capability. Seven OmpA family proteins were identified in this study and out of those, 6 proteins (Fisuc_2396, Fisuc_2509, Fisuc_2917, Fisuc_1592, Fisuc_2892, and Fisuc_2892) were found in high abundance in the cellulose treatment. The role of these proteins in structure and porin function is well studied [447]. However, a recent study proved that these proteins plays vital role in cell adhesion [448]. We also found a PEGA protein domain (Fisuc_1316) in significantly high abundance in cellulose grown cells which have adhesive ability too. The adhesion of

Lactobacillus against pathogens was also observed previously [449]. Our results suggest that proteins may have a combined effect in the cellulose adhesion mechanism in *F. succinogenes*. A putative type IV pilin protein (Fisuc_2041) was also found in high abundance in cellulose grown cells. This protein is well known for its role in adherence in *F. succinogenes* [44]. It was only observed in the cellulose treatment condition in Chapter 5, Table5-1.

Interestingly, two putative membrane associated proteins (Fisuc_1660 and Fisuc_3024) were found in significantly high abundance in cellulose grown cells (compared to glucose). The sequence similarity of uncharacterised protein (Fisuc_1660) showed 84.7 % homology with putative fusion protein (Ctha_1250) of *Chloroherpeton thalassium* strain ATCC 35110 that is involved in biofilm formation [450]. Similarly putative protein (Fisuc_3024) showed 65 % homology with putative protein (ABNIH2_09007) of *Acinetobacter baumannii* ABNIH2 and putative protein (BACEGG_00469) of *Bacteriroids eggerthii* DSM 20697 indicates that this might be involved in biofilm formation [451]. Biofilm formation is a crucial step in cellulose degradation mechanism of *F. succinogenes* S85 [43]. Therefore, these proteins might have the ability to produce biofilm in *F. succinogenes*.

Another 5 different putative uncharacterised proteins (Fisuc_0557, Fisuc_1284, Fisuc_2300, Fisuc_0209, Fisuc_0463 and FSU_0881) containing *Fibrobacter succinogenes* major paralogous domain were identified as a differentially expressed proteins among the treatments. The majority of them (4) were found in high abundance in the cellulose treatment, indicated that they probably play crucial roles in cellulose degradation mechanism. However, to date, none of them have been characterised so far. Most of the significantly regulated proteins are

putative/uncharecterised, since 50 % of the open reading frames (ORF) have unknown functions in *F. succinogenes.*

There were 5 putative proteins found at high abundance in cellulose grown cells. These proteins showed 62 % homology with Gamma-soluble NSF attachment protein (PAL_GLEAN10023235) of *Pteropus alecto* which possess a TPR domain. The TPR domaing containing proteins are involved in protein-protein interaction and protein complex formation [420]. The distinct TPR domain containing proteins were identified in Chapter 5.

6.3.2.3 Protein protein interaction

Interestingly, there were two MAPs that belong to the PPIC-type PPIASE domain proteins (FSU_0013, Fisuc_2974) that were significantly up-regulated in the cellulose treatment. These proteins are involved protein folding and protein assembly [452]. In addition, three more proteins used in protein folding (Fisuc_0872, Fisuc_1756, Fisuc_0518) were found in high abundance in cellulose grown cells. High abundance of these proteins indicates that complex protein metabolism activities takes place within the membrane during cellulose degradation.

6.3.2.4 Transportation

A MotA/TolQ/ExbB proton channel (Fisuc_1894) protein was significantly up regulated in cellulose grown cells. This protein helps translocation of proteins across the membrane [419]. Another transport protein which is significantly up-regulated in cellulose treatment was mechanosensitive ion channel family protein (Fisuc_3033). It is an energetically well tuned MS channel in microbes that allows

sequential opening of channels and protects the cell from osmotic stress [453, 454]. There were three efflux transportater proteins that were significantly changed in the cellulose treatment. Efflux transporter, RND family, MFP subunit (Fisuc_0289) and outer membrane efflux protein (Fisuc_1571) were all up-regulated and the outer membrane efflux protein (Fisuc_0288) was down-regulated in cellulose grown cells. This belongs to membrane fusion protein (MFP) family and facilitates outward flow of the substances from the cell [455].

Up-regulation of these proteins in cellulose treatment partially supports the third step in cellulose degradation mechanism, i.e. degradation of cellulose occurs in the periplasm and hydrolysed by-products are exported out via MFPs from cell [57].

Finally, there was an abundance of uncharacterised proteins significantly up-regulated in cellulose treatment however their functions are yet to be characterised. The total quantified proteins with significant changes in both techniques are shown in Appendix 6.1 and 6.2.

6.4 Conclusion

In conclusion, biotinylation-iTRAQ analysis could be a better option for membrane proteomics studies. Although, it seems that biotinylation does not help to avoid cytosolic contamination but it definitely helps to improve the membrane protein enrichment as compared to traditional membrane protein isolation methods. In this study, we compared number of significantly regulated proteins between two membrane proteomics method. The higher number of significantly regulated proteins was obtained by biotinylation (total 275 proteins) as compared to ultracentrifugation (total 143). In particular 138 MAPs were significantly changed proteins obtained by biotinylation, the numbers are quite higher than

ultracentrifugation method (28 MAPs). Our results suggested that the selectivity of biotinylation towards membrane proteomics increased their identification. The number of proteins (differentially expressed) obtained by both techniques were classified as per their colocalisation across the membrane. In particular, the proteins obtained by biotinylation from OM, UN (NCP), UN (ML)) were higher than the ultracentrifugation. The high abundance proteins were mostly obtained by biotinylation in almost every fraction of the membrane (Figure 6-7). Data suggesting that biotinylation could be the best option for membrane proteomics in terms of protein coverage.

This study suggested that there is a significant alteration in the membrane proteome that takes place during cellulose degradation. We identified various proteins that have vital role in cellulose degradation. The proteins involved in adhesion mechanism such as fibro-slime proteins, IV pilin proteins were significantly up-regulated in cellulose substrate condition indicating that those proteins could be the major proteins in initial adhesion. Another important protein named, CBM30 plays provital role in single chain cellulose adhesion was up-regulated in cellulose treatment. The surface localisation (OM) of this protein supports initial step in proposed cellulose degradation mechanism (transfer of single chain cellulose into the mebrane). The proteins involved in protein-protein interaction was higher in cellulose treatment indicating that protein metabolism within membrane was accelerated during cellulose degradation. As noticed in chapter 5, the efflux proteins belongs to MFPs were found in high abundance supports the third step in cellulose degradation mechanism (efflux of sugar monomers from the cell to environment). The large numbers of proteins with unknown functions were significantly regulated during cellulose degradation.

However, the study was limited to certain proteins due to a lack of genetic information (and annotation) and uncharacterised proteins.

Chapter 7

Alcoholic fermentation of carbon sources in biomass obtained from chemical and biological hydrolysis by *Clostridium acetobutylicum* ATCC 824

Abstract

Clostridial bacteria have increased attention as promising candidates for biofuel production from lignocellulosic biomass. This study investigated ethanol, butanol and hydrogen (H_2) productions using *Clostridium acetobutylicum* ATCC 824 from lignocellullosic biomass. The study aims to produce biofuels from chemically and biologically hydrolysed biomass. In the first approach, miscanthus hydrolysates were obtained from *Miscanthus giganteum* biomass by chemical treatment using water, and H_2O, 0.2N H_2SO_4 and 0.2N NaOH at 130^oC. Hydrolysates typically contained variable concentrations of sugar components, particularly, glucose, arabinose, xylose, and mannose. These sugars were used for producing acetone, butanol and ethanol (ABE) using *C. acetobutylicum* ATCC 824. In the second approach, we subjected cellulosic materials to biological hydrolysis using *F. succinogenes* S85 and subsequent fermentation using *C. acetobutylicum* ATCC 824. We obtained butanol, ethanol and H_2 gas productions from both approaches. Interestingly, no acetone production was observed during fermentation. This study demonstrated the great potential of *C. acetobutylicum* as a future biofuel producer generating good candidates from lignocellulosic feedstock.

7.1 Introduction

Biofuel production from lignocellulosic materials (wood, agricultural and forest residues) is promising breakthroughs over existing fuels [456]. As a result, lignocellulosic biomass has an unique place in future biofuel generation that can provide sustainable and environmentally friendly alternative fuels [457].

Lignocellulosic biofuel production is produced via two steps: 1) Deconstruction of cell wall polymers into components of sugars (pre-treatment and saccharification) and 2) conversion of sugars to biofuels (fermentation) [458]. However, the major bottleneck in conversion of biomass to biofuel conversion is the recalcitrant nature of lignocellulosic polymers that makes the hydrolysis (saccharification) step rate limiting [459, 460].

In order to convert lignocellulosic biomass to fermentable sugars and also to make it more accessible to microbial fermentation, various hydrolysis techniques are employed [41, 461, 462]. Deconstruction of lignocellulosic biomass can be achieved using either physical, chemical or biological pre-treatments [41, 463]. Various chemical and biological pre-treatments using in lignocellulosic biofuel generation have been recently reviewed by Chaturvedi and Verma [464].

Researchers have continuously been using physical and chemical pre-treatment processes for biomass deconstruction including hot water, steam explosion, CO_2 explosion, ozonolysis, solvents and acid/alkali [41, 367, 465]. However, today's industrial level biomass degradation processes are heavily depended on heat and acid/alkali treatments [22, 41, 466, 467] and thus tend to be more expensive, slow and relatively inefficient treatment [12]. On the other hand, microbial

strategies for simultaneous saccharification and fermentation is an ideal option because of low energy input and mild environmental conditions [41]. However, operation of naturally capable bacteria is necessary for biological biofuel process development.

7.1.1 Chemical strategy

7.1.1.1 Biomass hydrolysate fermentation

Ideally, chemical treatments at high temperature generate hydrolysates that are rich in sugar components. These sugar components can be further fermented into fuels using typical industrial microbes such as *Saccharomyces cerevisiae*. However, particularly, Clostridial species are of great interest over traditional ethanol producing yeast because they can naturally utilise a diverse range sugars [468] from lignocellulosic hydrolysate. They are well equipped to produce solvents using their multi-substrate utilising capacity than any other genus of the three domains (Bacteria, Archaea, Eubacteria) [468]. Moreover, they can produce high energy content alcoholic products (such as isopropanol, n-butanol and n-hexanol) [118]. Some host species are good producers of acetone-butanol-ethanol (ABE) fermentation [125] and become important industrial microbes. Butanol producing strains such as *Clostridium acetobutylicum* and *Clostridium beijerinckii,* demonstrated their potential to ferment sugars derived from hydrolysates of agriculture residues [468].

7.1.2 Microbial strategy

7.1.2.1 Lignocellulosic co-culture fermentation

Microbial strategies, on the other hand, are diverse for lignocellulose degradation and offer new avenues for the development of biological based processes for industrial scale production of biofuels [12]. However, our understanding of lignocellulose degradation is limited to the model organisms such as *Trichoderma reesei* and *Clostridium thermocellum* [42].

Microbial bioprocessing seems to be a promising approach for energy conversion, in particular for lignocellulosic biofuel production [1]. The consolidated bioprocessing (CBP) framework is an alternative microbial bioprocessing approach with outstanding potential in simultaneous saccharification and fermentation. CBP comprises four transformations in a single step; 1) hydrolytic enzyme production (cellulases and hemicellulases), 2) hydrolysis of corresponding carbohydrate polymers into fermentable sugars (glucose, mannose and galactose), 3) fermentation of hexose sugars and 4) fermentation of pentose sugars. In terms of cost for conversion of biomass to biofuel production, CBP is a most economically cheaper option [1, 173]. However, microbes for CBP development with both capabilities (efficient in cellulose conversion and biofuel generation) are lacking in a single microbe and thus need to be developed [1, 173]. Therefore, in order to get better cellulose degradation/fermentation, current CBP strategies employ natural and recombinant cellulolytic microorganisms [469]. Anaerobes with efficient cellulose degradation capability and biofuel generation are of particular interest of

researchers [5]. A combination of microbes with desirable abilities such as substrate utilisation and product (biofuel) formation can provide a major breakthrough as a low cost technology. The utilisation of combinations of cultures (co-culture) with distinct capabilities is one of the alternatives for CBP.

Fibrobacter succinogenes is the most efficient cellulose degrader found in the rumen while on the other hand *Clostridium acetobutylicum* has the best capability to ferment a diverse range of sugar components into acetone, ethanol and butanol [171, 172]. Thus, using these two mesophilic anaerobes in a co-culture may be a strong alternative for CBP development.

Similar co-culture approaches have been studied previously for bioenergy production from cellulosic biomass utilising, for example, *Clostridium acetobutylicum* and *Clostridium cellulolyticum*, *Clostridium cellulolyticum* [173] and *Rhodopseudomonas palustris* [174], *Clostridium acetobutylicum* and *Ethanoigenens harbinense* [175], and *S. cerevevisiae* and *Scheffersomyces stipites* [176].

In the present study, we attempted to use miscanthus giganteum hydrolysate for ethanol and butanol production as the main by-products. In the first approach, we achieved hydrolysis of biomass using chemical treatment followed by fermentation. An array of chemical treatments including water, alkali and acid treatments were used to extract maximum fermentable sugars from biomass in the form of hydrolysate. In the second approach, we subjected different forms of biomass (acid swollen cellulose, microcrystalline cellulose, and 1N NaOH treated miscanthus giganteum) to simultaneous microbial deconstruction and

fermentation (co-culture). The sugars released from both pre-treatments were fermented to ethanol, butanol and H_2 gas using *C. acetobutylicum* ATCC 824.

7.2 Materials and methods

7.2.1 Microorganisms used and medium preparation

All chemicals used in this study were obtained from Sigma Aldrich (UK), unless otherwise indicated.

7.2.2 *Clostridium acetobutylicum* ATCC 824

Clostridium acetobutylicum ATCC 824 was grown anaerobically in a 125 ml capacity serum bottle fitted with butyl rubber and crimp sealed containing 100ml media. The media composition was used as described by Lopez Contreras *et al.*[253] having the following composition (hereafter denoted as CA media); 0.75 g L^{-1} KH_2PO_4, 0.75 g L^{-1} K_2HPO_4, 0.348 g L^{-1} $MgSO_4$, 0.01 g L^{-1} $MnSO_4.H_2O$, 0.01 g L^{-1} $FeSO_4.7H_2O$, 1 g L^{-1} NaCl, 5 g L^{-1} yeast extract, 2 g L^{-1} $(NH_4)_2SO_4$. However, arginine was replaced with 1.0 g L^{-1} cysteine chloride as reducing agent and with 5 g L^{-1} glucose as a carbon source. The medium was heated to boiling point and cooled down by flushing with 100 % N_2 gas. The bottles were sealed with butyl rubber and crimped sealed and autoclaved for 15 min at 121^{o}C. The medium was inoculated with freshly prepared inocula and incubated at 38^{o}C for 18 to 20 hours (up to the exponential phase).

7.2.3 *Fibrobacter succinogenes* S85

The strain *F. succinogenes* S85 (ATCC 19169) was kindly provided by Prof. Paul Weimer (US Dairy Forage Research Centre, Wisconsin, USA). *F. succinogenes* S85 was cultivated under anaerobic conditions at 38^{o}C in a synthetically modified

Dehority medium (MDM) (hereafter denoted as FS media) as described by Weimer *et al.* [364, 365].

To prepare the basal media, the following stock solutions were prepared: mineral solution I (22.5 g L^{-1} KH_2PO_4), mineral solution II (11.26 g L^{-1} NaCl, 11.26 g L^{-1} $(NH_4)_2SO_4$, 1.06 g L^{-1} $MgCl_2.6H_2O$, 0.82 g L^{-1} $CaCl_2.2H_2O$, 0.344 g L^{-1} $MnCl_2.4H_2O$, 0.118 g $L^{-1}ZnCl_2$, and 0.026 g L^{-1} $CoCl_2.6H_2O$), 80 g L^{-1} Na_2CO_3 solution, 10 g L^{-1} (v/v) volatile fatty acid (VFA) (isobutyric acid, isovaleric acid, n-valeric acid and 2-methylbutyric acid) solution and 25 g L^{-1} cysteine HCl. Except mineral solution II, all stocks solutions (100 mL) were prepared by flushing with 100 % N_2 gas in 125 mL serum bottles sealed with butyl rubber plus aluminium seal and autoclaved for 15 min at 121^{o}C. Schaefer's vitamin solution (Chapter 5, section 5.2.1) was also prepared as described by Callaway and Martin [365] .

7.2.3.1 Basal medium (FS media)

One hundred mililitres of basal medium was prepared by adding 8 mL of stock solution II into 87.5 mL of distilled water and flushed with 100 % CO_2 in 125 ml bottle then closed with butyl rubber and crimp sealed, and autoclaved at 121^{o}C at 15 min. All the stocks were taken to the anaerobic chamber and added to the basal medium to achieved final basal medium concentration of components: 0.9 g L^{-1} KH_2PO_4, 3.2 g L^{-1} Na_2CO_3, 1 g L^{-1} cysteine·HCl, and 0.06 g L^{-1} each of isobutyric acid, isovaleric acid, n-valeric acid and 2-methylbutyric acid. Similarly, a final mineral concentration was obtained as following 0.9 g L^{-1} NaCl, 0.9 g L^{-1} (NH_4) $2SO_4$, 0.084 g L^{-1} $MgCl_2·6H_2O$, 0.065 g L^{-1} $CaCl_2·2H_2O$, 0.0275 g of $MnCl_2·4H_2O$, 0.02 g L^{-1} $FeSO_4·7H_2O$, 0.009 g L^{-1} $ZnCl_2$, and 0.0048 g of CoCl2·6H2O. 10 mL L^{-1} of Schaefer's vitamin solution was also added.

7.2.3.2 Medium optimisation for co-culture development

Since both bacteria have different medium and growth conditions, it was necessary to optimise the media in such way that both mesophiles can grow in a single medium. In order to achieve a modified medium, it was required to combine basal medium used for *F. succinogenes* (FS media) *and C. acetobutylicum* (CA media).

To obtain the modified media, we prepared 6 media bottles of each FS and CA media with 5 g L^{-1} glucose as a carbon source as discussed in section 2.2 & 2.3.1. Then, we combined both the media (FS to CA (v/v)) to obtain the ratio of 100 % FS, 20 % FS plus 80 % CA, 40 % FS plus 60 % CA, 60 % FS plus 40 % CA, 80 % FS plus 20 % CA and 100 % CA. There were two sets of these combinations prepared. All the combinations were prepared in an anaerobic chamber in pre-sterilized 125 mL serum bottles caped with butyl rubber and crimp sealed. These modified media were then inoculated with *F. succinogenes (*OD_{675} =0.72) and *C. acetobutylicum* (OD_{600} = 1.2), and grown on their respective media with glucose as a carbon source. The growth of both bacteria was monitored in their respective sets of media by measuring OD at 675_{nm} for *F. succinogenes* and at 600_{nm} for *C. acetobutylicum*. From the reading obtained from both bacteria, the combination of 40 % FS plus 60 % CA media was considered as a modified media for the growth of both bacteria.

7.2.4 Chemical hydrolysis of miscanthus giganteum biomass

Acid/alkali pre-treatments have been extensively studied to process agricultural biomass [470]. Miscanthus biomass hydrolysate was kindly provided by Dr. Leonardo Gomez (Centre for Novel Agricultural Products, The University of York, UK). Miscanthus hydrolysate was obtained using treatment of water, 0.2N H_2SO_4 and 0.2N NaOH at 130^oC. The supplementary salt medium was added to each bottle as suggested by Wang and Chen [471]. The supplementary salt medium contained 6 g L^{-1} $(NH4)_2SO_4$, 1.768 g L^{-1} KH_2PO_4, 2.938 g L^{-1} K_2HPO_4, 2 g L^{-1} $CaCO_3$, 10 mg L^{-1} p-aminobenzoic acid, 10 mg L^{-1} biotin and 1 mL L^{-1} mineral salt solution as described by George *et al.* [472]. The hydrolysates were then neutralised to pH 6.5 using H_2SO_4 and NaOH and centrifuged at 1000 g for 2 min. Supernatants obtained from each treatment were then sterilised using polyethersulfone steritop-GP filter paper (Millipore, 02μm). A total of 400 mL miscanthus hydrolysates were divided into two of each 500 ml capacity bottles fitted with rubber tight cap provided with inlet and outlet ports (Figure 7-1). The medium was heated to boiling and cooled (10 min each) down by continuous flushing with 100 % N_2 gas. Finally, bottles were tightened using clips. A reducing agent was added to remove remaining oxygen from the bottles using cysteine HCl (1 g L^{-1}). The pH of the media was finally re-checked to make sure that pH was maintained at 6.5. The medium was inoculated with 4 ml of freshly prepared inocula of *C. acetobutylicum* ATCC 824 to each bottle and incubated at 38^oC. The experimental set-up is shown in Figure 7-1. Finally, the supernatant was collected from the fermentation broth and subjected to ethanol, butanol and H_2 gas measurement.

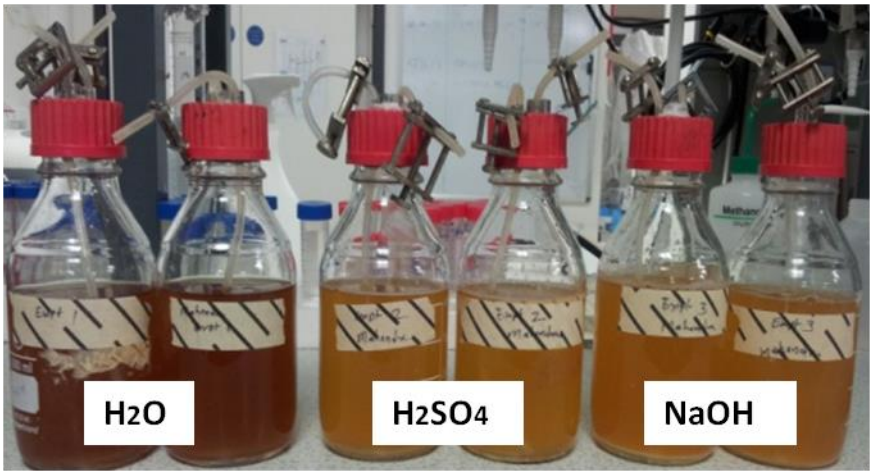

Figure 7-1 Experimental set-up of the fermentation of miscanthus biomass hydrolysate using *C. acetobutylicum*

7.2.5 Biological treatment of cellulosic biomass

For biological treatment, we selected Acid swollen (AS) cellulose, microcrystalline (MC) cellulose and 1N NaOH treated miscanthus raw (MR) biomass. 1N NaOH treatment was employed to miscanthus raw biomass in order to remove maximum lignin from the biomass [473].

One hundred millilitres of this modified media (40 % FS plus 60 % CA) was prepared with 5 g L^{-1} of AS cellulose, MC cellulose and MR biomass as a carbon source. Triplicate bottles of the media for each condition were first inoculated with *F. succinogenes* S85 to achieve hydrolysis of the substrates. *F. succinogenes S85* immediately adhered to the substrate particles and subsequently produced biofilms and released sugar into the solution [101, 102], After inoculation, bottles were incubated at 38^{o}C for 40 hours (approximately 40 hours required to achieve mid exponential phase). During this period, to avoid utilisation of released sugar monomers by planktonic cells and to achieve maximum hydrolysis, bottles were

kept stagnant to allow biofilm formation. After the 40 hour incubation, the media was then inoculated with *C. acetobutylicum* ATCC 824. Finally, supernatants were collected at 80 and 120 hours of incubation from the fermentation broth and analysed for ethanol, butanol and H_2 gas. The sampling times were selected based on the appearance of ethanol and butanol production in fermentation broth. The samples were collected after every 10 hours of incubation. Ethanol production was observed at 80hrs of incubation. Figure 7-2 shows *F. succinogenes* growth on MC cellulose and subsequent co-culture fermentation

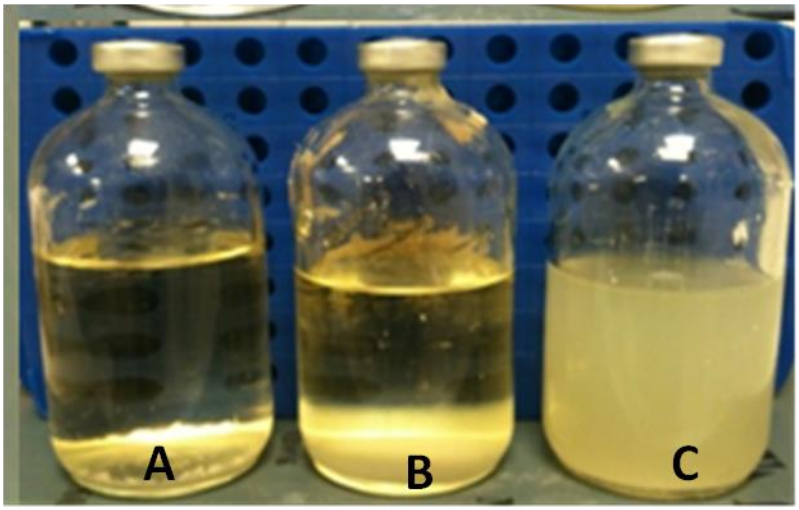

Figure 7-2 Biological hydrolysis of cellulose and fermentation. A) Modified cellulose medium B) Growth of *F. succinogenes* at 40hrs of incubation (biofilm) C) Fermentation (*F. succinogenes* plus *C. acetobutylicum*) at 120hrs

7.2.6 Dry weight of cellulosic biomass measurements

The final dry weight of AS cellulose and MC cellulose and MR biomass in fermentation broth were determined as described by Zuroff *et al.* [170].Briefly, 15 mL of broth was collected from bottles and centrifuged at 3000 g for 10 minutes, then the substrate pellet was washed twice with 1 % methylcellulose solution to

remove bound cells from the substrates. Substrate pellet were further washed with distilled water and centrifuging at 3000 g for 10 minutes. The supernatant was removed and tubes placed, open, in a drying oven at 80^{o}C. The samples were dried until a constant mass. The difference in the final and initial wt of samples was assumed to be the substrate utilised by co-culture for biofuel production.

7.2.7 Analysis of sugar concentration in miscanthus hydrolysate

The monosaccharides were separated by high performance anion-exchange liquid chromatography on a Dionex ICS-3000 using a Carbopac PA-20 column integrated with amperometry detection. The separated monosaccharides were quantified by using external calibration with an equimolar mixture of nine monosaccharides standards (arabinose, fructose, galactose, galacturonic acid, glucose, glucoronic acid, mannose, rhamnose and xylose). The separation was carried out at 0.5 mL/min flow rate in 1 % 200mM NaOH followed by 47 % H_2O_2, 22.5 % NaoH (200mM) and 30 % of 0.1M NaOH/0.5 CH_3COONa. The chromatographic separation was developed at 30^{o}C.

7.2.8 Analysis of fermentation byproducts

Fermentation products were identified and quantified as previously reported Pham *et al.* [234]. Briefly, ethanol, butanol, acetic acid and butyric acid were detected and quantified using a GC- chromatograph 7890A (Agilent Technologies) system coupled with a 30 m × 0.25 mm ID × 0.25 µm df Stabbilwax fused silica column (Thames Restek, UK). Approximately 50 µL aliquots were collected, centrifuged at 17,000 g for 2 min and transferred to a GC vial, and then 2 µL of sample was withdrawn for GC analysis. The total GC analysis running

time was 14 min and temperature gradient was performed with a hold at 45^{o}C for 3 min, followed by a ramp at a rate of 15^{o}C/min to 120^{o}C, then 30^{o}C/min to 210^{o}C and finally a hold 1 min at 210^{o}C. Helium was used as the carrier gas at a flow rate of 1 mL/min. The concentration of ethanol, butanol, acetic acid and butyric were estimated by obtained standard graphs for the respective metabolites based on its retention time and peak area. An FID detector was used to recorded data.

7.2.9 Hydrogen gas estimation

Gas samples were collected from the headspace sampling bottles using 20 mL capacity a gas tight syringes at different time intervals depending on the sample types. Each time, 5 mL of the sample was then injected in to a Varian CP-3800 gas chromatograph (Varian, Polo Alto, CA) equipped with a 500 µL sample loop. This volume was then directly injected *via* the Varian 1041 splitless on-column injector. Component separation was achieved using Haysep (c18-100 mesh, porous polymer column 2.0 m length and 0.32 cm inner diameter with 2 mm solid support) and A molecular sieve (13X, 60-80 mesh, packed column 1.5 m length 0.32 cm inner diameter with 2 mm solid support) with argon carrier at a flow rate of 3.6 mL/min. The instrument was equipped with Thermal Conductivity Detector (TCD).

The GC was controlled and automated by the Star GC workstation (Version 5.50) software package (Varian). The instrument was calibrated using standard hydrogen calibration gas supplied by BOC speciality gases. An overview of the methodology is shown in Figure 7-3.

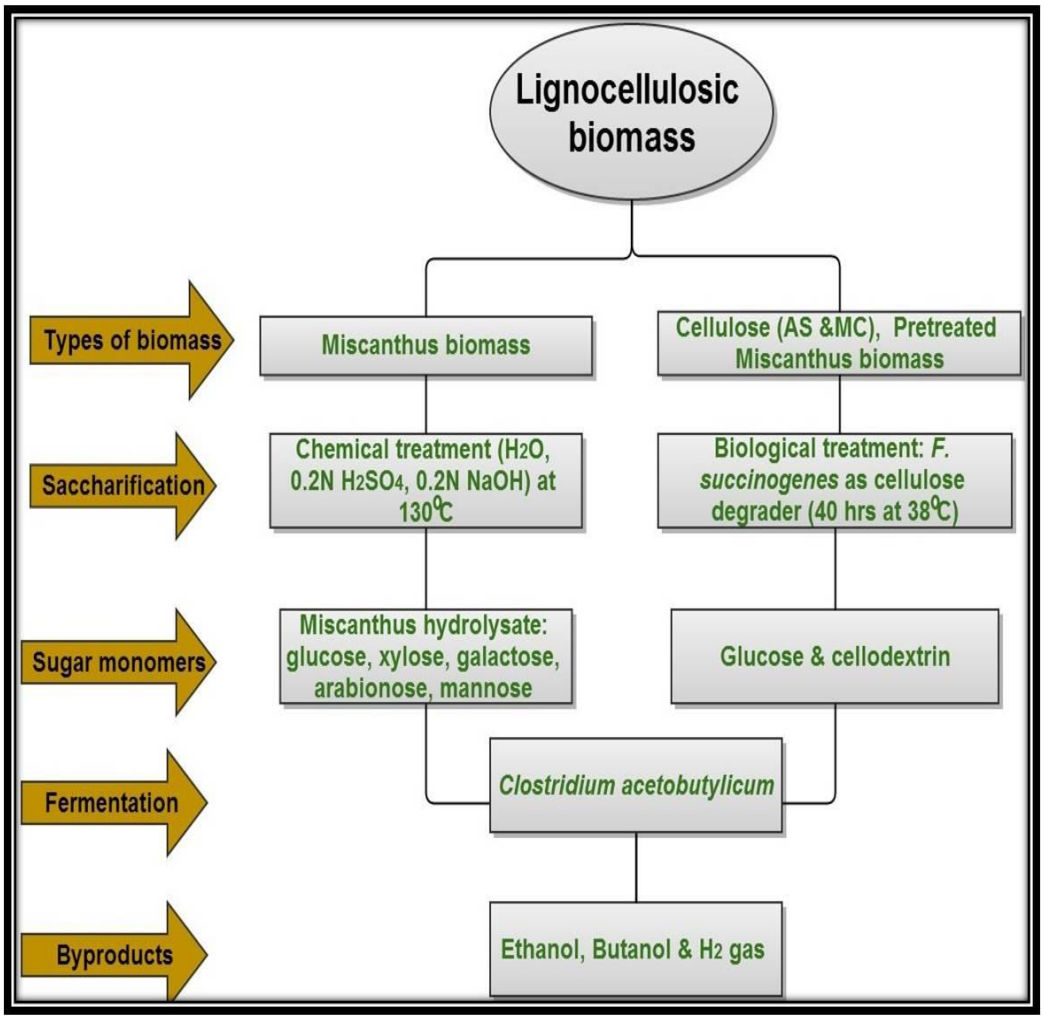

Figure 7-3. Overview of experimental design

7.3 Results and discussion

7.3.1 Sugar composition of miscanthus hydrolysates

The compositional analysis of the miscanthus hydrolysate is shown in Figure 7-4. In this study, we obtained miscanthus hydrolysate by treating with H_2O, 0.2N H_2SO_4 and 0.2N NaOH at 130^{o}C. To examine the ability of *C. acetobutylicum* ATCC 824 to utilise various sugars presented in the misnanthus hydrolysate, we analysed the concentrations of glucose, xylose, arabinose and mannose before and after fermentation. The concentrations of sugar components in the

hydrolysate varied from different treatments. The highest concentrations of sugars produced in hydrolysate derived from 0.2N H_2SO_4 treatment were 0.279 g/L glucose, 1.245/L xylose, 0.325 g/L arabinose and 0.022 g/L mannose. Whereas, the fermentable sugar productions were found to be lowest in H_2O treatment such as 0.031 g/L glucose, 0.033 g/L xylose, 0.022 g/L arabinose and 0.062 g/L mannose. Xylose was the most abundant sugar in the hydrolysates of the pre-treatment conditions, especially in H_2SO_4 treatment since acid pre-treatment is a robust treatment method that can degrade hemicelluloses leading to production of fermentable sugars such as xylose [474, 475]. Whereas, alkali treatment is less effective on hemicelluloses solubilisation [476, 477]. Therefore, concentrations of sugars are comparatively less than acid treatment. Glucose, xylose, arabinose and mannose are the major sugar components of biomass hydrolysates [478]. *C. acetobutylicum* can utilse a variety of carbohydrates including hexoses (eg. glucose) and pentoses (D-xylose and L-arabinose) [20, 280, 468]. After fermentation, concentrations of these sugars significantly reduced in all treatments (Figure 7-4), suggesting that *C. acetobutylicum* can utilise a variety of sugars as carbon source and can produce biofuel simultaneously.

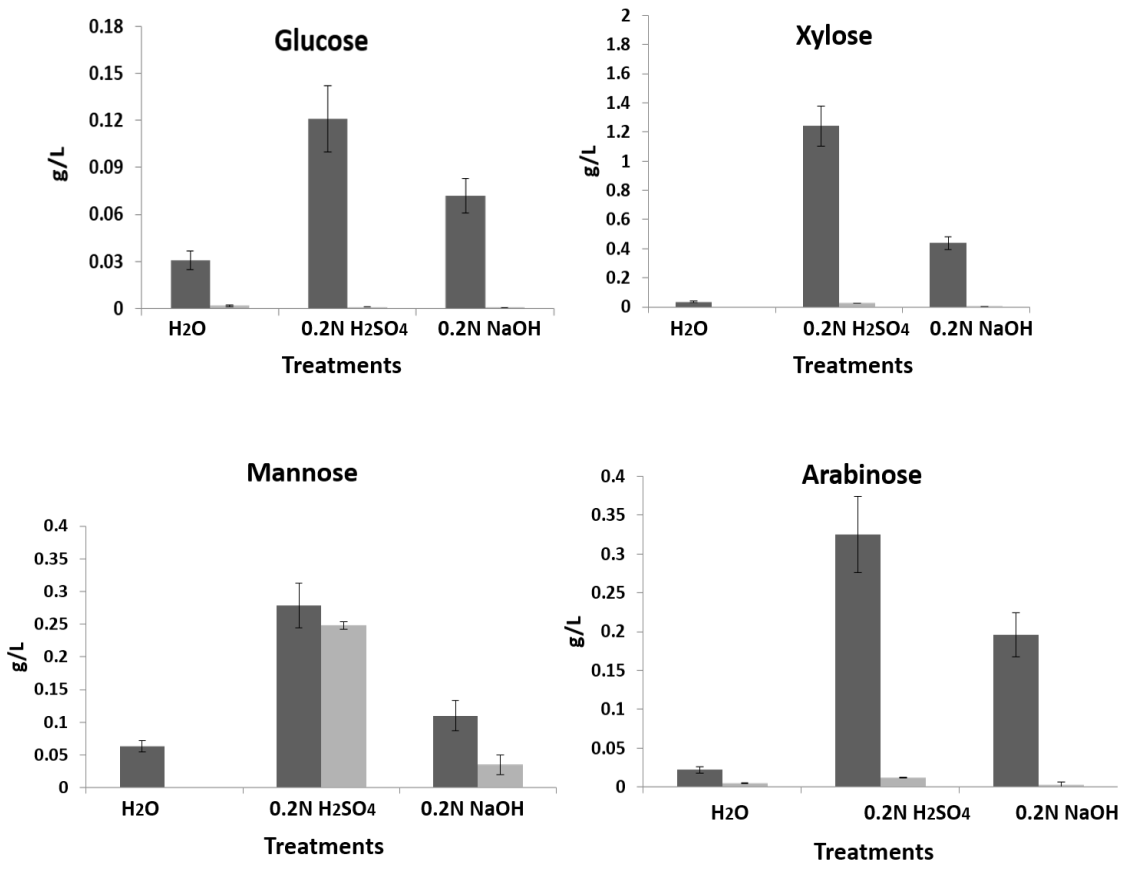

Figure 7-4 Fermentable sugars consumption by *C. acetobutylicum* during anaerobic fermentation of miscanthus hydrolysate. Figure represents initial (■) and final (□) concentration of sugars of fermentation. Data were taken from biological duplicates.

These sugars releasing into the hydrolysate solutions were used to produce ethanol, butanol and H_2 gas by *C. acetobutylicum*. The concentrations of by-products are compared among the treatment conditions.

7.3.2 Fermentation of Miscanthus biomass hydrolysate by Clostridium acetobutylicum

C. acetobutylicum possesses biphasic fermentation. During the exponential growth phase, acetic acid, butyric acid and H_2 gas are produced (acidogenesis) and during the late exponential phase followed by the stationary phase, acetone,

butanol and ethanol are produced [479]. A drop down in pH below 5 during acidogenesis triggers solvent formation [480].

Figure 7-5A to C shows production of ethanol, butanol and H_2 gas in different treatment conditions at 80 hrs and 120 hrs of incubation. Ethanol production (Figure 7-5A) shows little variation among the pre-treatments. However, ethanol production was relatively higher for 0.2N NaOH and 0.2N H_2SO_4 treatments compared to H_2O. The highest concentration of ethanol was approximately 0.042 g/L at 120 hrs of incubation in H_2SO_4 and NaOH treatment compared to 0.029 g/L for H_2O treatment. The highest butanol production was observed for the H_2SO_4 treatment (0.02 g/L). Butanol production was also observed in the NaOH treatment at 120 hrs of incubation (0.009 g/L), but the concentration was relatively lower than the butanol obtained from H_2SO_4 treatment (Figure 7-5B). Interestingly, there was no butanol production in the H_2O treatment. Absence of butanol production in H_2O treatment and less production of butanol in the 0.2N NaOH treatment might be because of lower concentration of sugars in the hydrolysates obtained by both treatments (Figure 7-4). The concentrations of sugars are comparatively higher in the 0.2N H_2SO_4 treatment that is reflected in the ethanol/butanol production yields. The effect of how sugar concentration in hydrolysates affects subsequent biofuel production was observed previously where an elevated level of glucose or sugars in the medium resulting in induced butanol production was observed [481, 482].

Hydrogen is a clean and efficient replacement to fossil fuels [483]. Hydrogen was also produced from hydrolysates in all treatments (Figure 7-5C). The highest production of H_2 was observed in H_2SO_4 treatment (16.25 mg/L of culture) while its concentrations were 11.46 mg/L of culture and 2.06 mg/L of culture for H_2O

and NaOH treatments respectively. The lowest production of H_2 gas was found in the NaOH treatment condition.The reason for reduced H_2 production in alkali treatment is unknown. However, it was reported that the NaOH treatment generates soluble lignin and other inhibitor by-products with sugars that may affect H_2 gas production [484, 485]. Our results suggested that the biomass treatment conditions significantly affected butanol, ethanol and H_2 productions. Overall, results showed that the H_2SO_4 treatment had a better yield of butanol, ethanol and H_2 gas over H_2O and NaOH treatment.

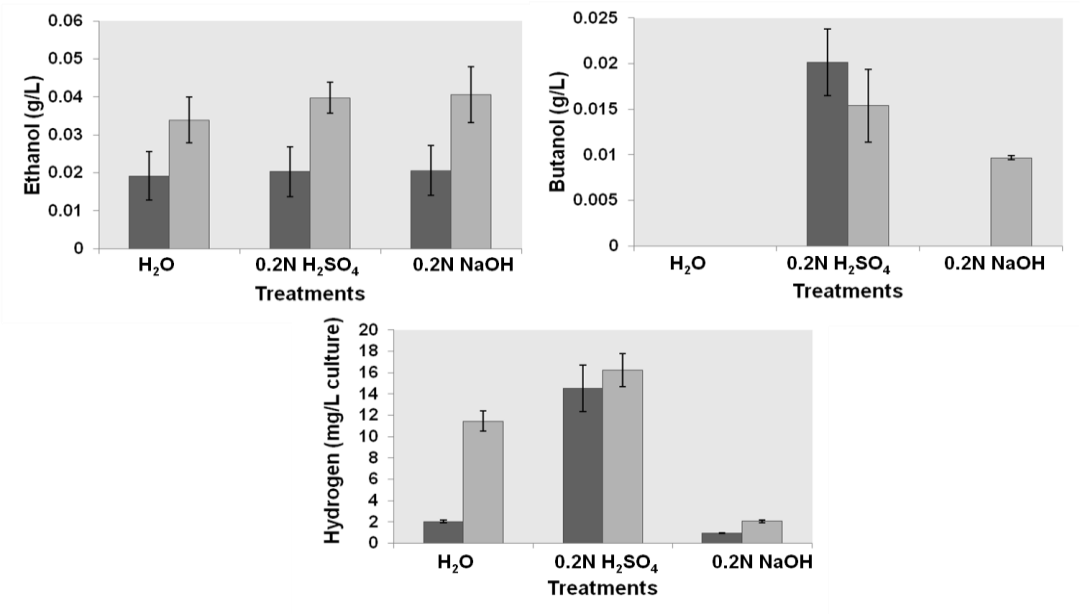

Figure 7-5. Biofuel generation from miscanthus hydrolysate using *C. acetobutylicum* ATCC 824.A) Ethanol, B) Butanol and C) H_2 gas.Samples taken at 80 hours (■) and 120 hours (□) of fermentation.

7.3.3 Fermentation of lignocellulosic biomass hydrolysed with F. succinogenes

7.3.3.1 Medium optimisation for co-culture conditions

In order to grow *F. succinogenes* S85 and *C. acetobutylicum* ATCC 824 as a co-culture, we modified media the so that it could allow both these two bacteria to grow in a single fermentation vessel. For that, we mixed the FS and CA media by 6 different concentrations and monitored the growth curves of cells. The optimum growth for both bacteria was observed at a combination of 40 % FS and 60 % CA media. Our results suggested that (Figure 7-6) *F. succinogenes* and *C. acetobutylicum* can grow simultaneously on 40 % FS and 60 % CA media with growth rates of 0.074 h^{-1} (doubling time 9.36) and 0.179 h^{-1} (doubling time 3.85) respectively. At this combination, the maximum $OD_{675nm,600nm}$ for *F. succinogenes* and *C. acetobutylicum* reached to 0.912 and 1.018 at 30 hours of incubation. The co-culture growth of both bacteria in the modified medium was observed by microscopy (Figure 7-7). This is the first study so far attempted to grow *F. succinogenes* and *C. acetobutylicum* together for simultaneous saccharification and fermentation. This modified medium (40 % FS plus 60 % CA) was used as a supplementary media with 5 g L^{-1} of each of AS cellulose, MC cellulose and miscanthus raw biomass (RB) as a sole carbon source. In this preliminary study, we observed hydrolysis of cellulosic material that was reflected in the final concentration of substrate left in the fermentation broth. The substrates concentrations were reduced from 5 g L^{-1} of AS cellulose, MC cellulose and MR biomass to 1.83 g L^{-1}, 2.29 g L^{-1} and 2.3 g L^{-1} respectively.

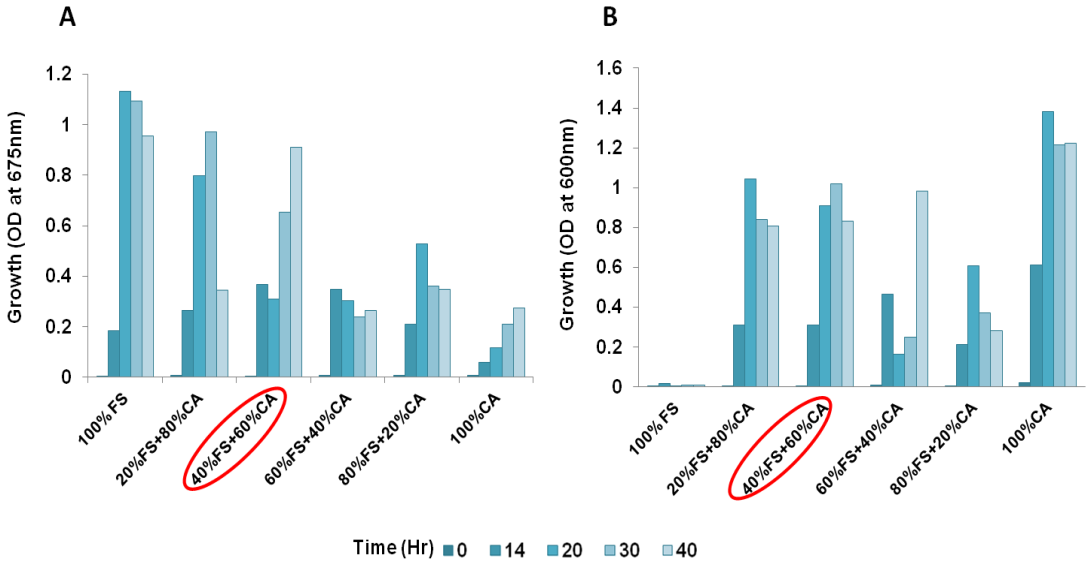

Figure 7-6. Growth profiles of *F. succinogenes* S85 (A) and *C. acetobutylicum* ATCC 824 (B) on modified medium used as co-culture.

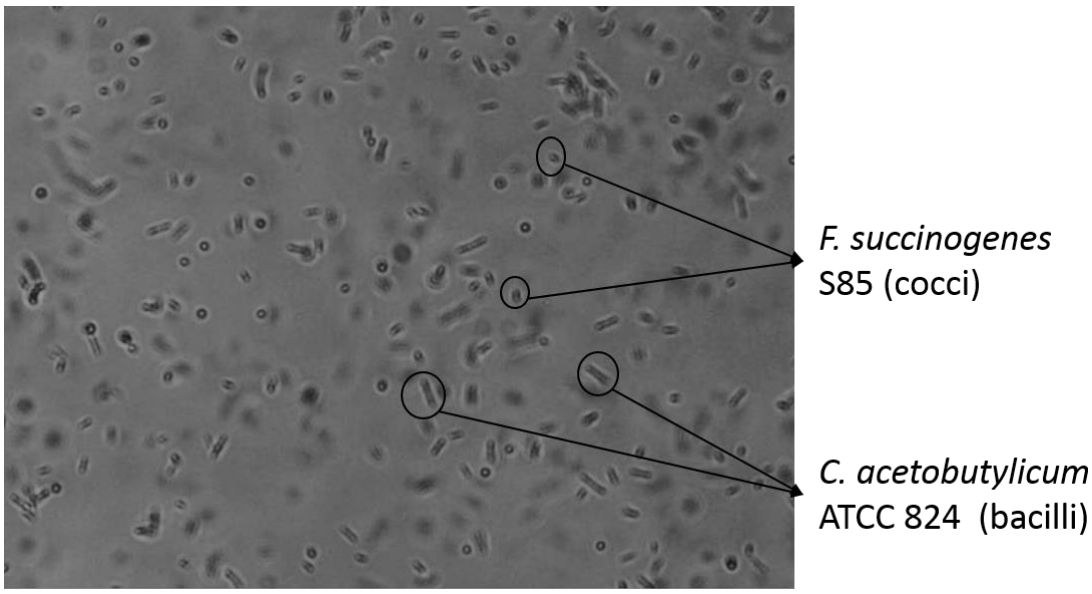

Figure 7-7.Mixed culture growth of *F. succinogenes* S85 and *C. acetobutylicum* ATCC 824 on modified media at 20 hours of incubation. Rod shaped cells represent *C. acetobutylicum* and coccoidal shaped cells represent *F. succinogenes*.

7.3.3.2 Fermentation by-products

The production of ethanol, butanol and H_2 gas were observed in all substrate conditions. However, depending on the types of substrates, the concentration of products varied (Figure 7-8).

Ethanol production was observed to be higher in AS cellulose supplemented medium and reached a maximum concentration of 0.27 g/L at 80 hours of incubation, followed by 0.21 g/L at 80 hours and 0.25 g/L at 120 hours in MC cellulose and MR biomass respectively. The ethanol concentration was comparatively higher in AS cellulose and MR biomass media. On the other hand, the maximum butanol production at 120 hours of incubation was 0.012 g/L, 0.014 g/L and 0.016 g/L for AS cellulose, MC cellulose and MR biomass supplemented media respectively. Butanol production showed little variation between the substrate conditions. However, slight differences in the concentrations were noted among the treatments. Hydrogen gas concentration was recorded as the highest in the RB (7.39 mg/L culture) followed by 6.8 mg/L culture in AS cellulose and 2.2 mg/L culture in MC cellulose media. The optimum productions of ethanol, butanol and H_2 was in the presence of AS cellulose. This might be because it was relatively easy substrate for hydrolysis by *F. succinogenes* S85 [486, 487]. Therefore, due to its greater susceptibility to release maximum sugars into the solution this might have then been used to produce ethanol, butanol and H_2 production by *C. acetobutylicum*. Our results suggest that the different substrates have significant effects on different forms of biofuel production by *C. acetobutylicum*. The ethanol, butanol and H_2 concentrations were found to be

comparatively higher than MC cellulose in MR biomass substrates. This indicates that alkali pre-treatment removed maximum lignin from the biomass and converted it into susceptible materials for enzymatic degradation [464]. This might allow *F. succinogenes* to boost hydrolysis and release not only glucose and cellodextrin but also the hemicellulose components that could be utilised by *C. acetobutylicum* for biofuels generation.

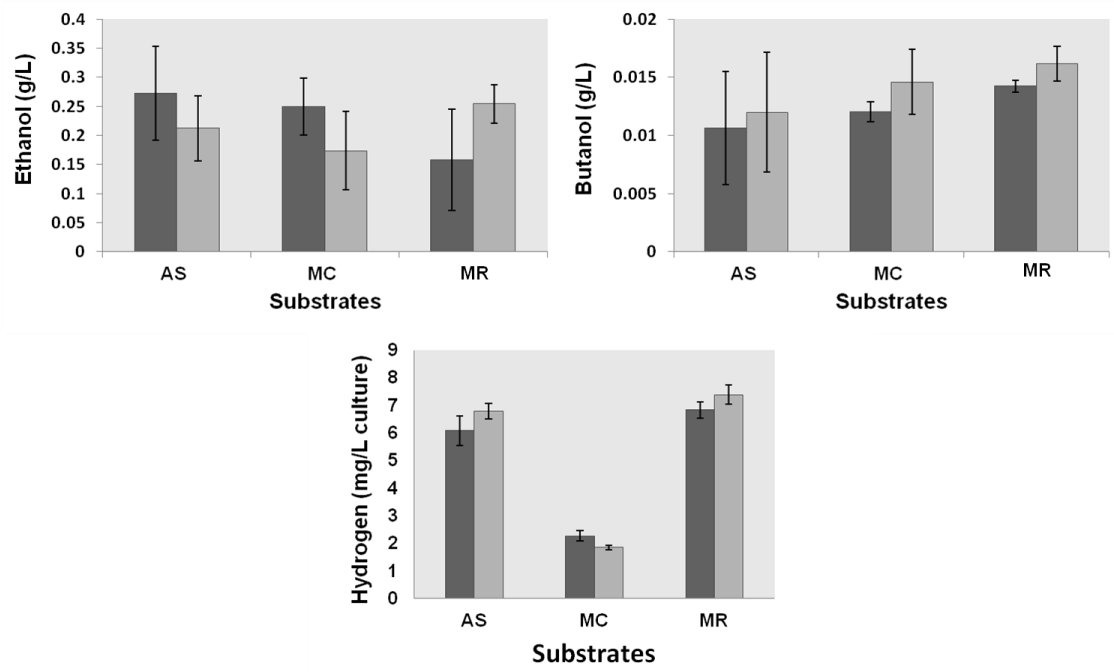

Figure 7-8 Biofuel generation from microbial co-culture system A) Ethanol, B) Butanol and C) H_2 gas. Sample were taken at 80 hours (■) and 120 hours (□) of fermentation.

In this preliminary study, we have shown for the first time that the two most efficient mesophilic lignocellulose degrading/fermenting co-cultures of *F. succinogenes* and *C. acetobutylicum* are able to grow together, producing C6 and C5 sugars and converting them to ethanol, butanol and H_2 gas as major fermentation by-products. No external enzymes or additives were required since cellulolytic/xylanolytic activity of *F. succinogenes* [87] generated sugars (C6 and

C5) that *C. acetobutylicum* could utilise and produce by-products via a fermentation process perfomed at 38^{o}C. Therefore, both strains represent a potentially useful starting point for the development of CBP at mesophilic temperature conditions.

7.4 Conclusions

The results from both chemical and biological hydrolysis and subsequent fermentation showed that both the techniques can be used for not only saccharification but also improving biofuel products using *C. acetobutylicum*. The results showed that among the chemical treatments, 0.2N H_2SO_4 treatment provided the best results since we can produce ethanol, butanol and H_2 in the same fermentation unit.

Similarly, in the preliminary study of a co-culture system, the results indicated that two anaerobes (mesophiles) with two different distinct abilities (saccharification & fermentation) can grow together and produce biofuel in the form of ethanol, butanol and H_2. It has also demonstrated the syntrophy between *F. succinogenes* and *C. acetobutylicum*. Furthermore, our results also showed that biological treatments showed maximum production of ethanol in all substrate conditions tested which was comparatively significantly higher than the chemical treatments tested. It can also be concluded that butanol can be obtained by either chemical or biological treatments, however, the production of butanol seems to depend on the sugar concentrations in the medium. H_2 gas production is higher in 0.2 H_2SO_4 treatments than any biological substrate condition. From the overall results, it can be concluded that both chemical and biological techniques showed promising results for future biofuels generation, since we could obtain alternative biofuels in different forms from a single fermentation unit. Nonetheless, biofuel production

using naturally capable bacteria still needs rigorous study for process optimisation. These two preliminary studies can be further improved to obtain higher product yields and rates.

Chapter 8

Conclusions and future work

8.1 General conclusion

I believe, the best solution for renewable energy possibly lies in nature's energy cycles: the plant biomass [488] and microbial systems [489]. This chapter discusses the aims, key findings and conclusions obtained from this PhD research study.

At present, the energy crisis and global warming are priority issues faced worldwide [6, 490, 491]. Lignocellulose is an abundant renewable resource and has many desirable features (such as its high energy content, significant compatibility, renewability, geographical availability, and eco friendliness) [5, 492] that can contribute to a future global energy alternative without competing with world food demand [493-495]. Biological conversion of lignocelluloses to biofuel is a promising solution, but still possesses major challenges for the researchers. For instance challenges include, the recalcitrant nature of lignocellulosic biomass [496], lack of genetic information about biofuel producing microbes and difficulties in the development of recombinant microorganisms [497]. Nonetheless, the innate capacity and potential of native microbes for lignocellulosic biofuel generation cannot be underestimated [6]. In that sense, anaerobic microbes have always been of particular interest to researchers [5] because of their significant contribution in pollution control and in production of bio-energy related value added products [498]. However, in order to achieve proper implementation for lignocellulosic biofuel generation and to develop microbial systems for biofuel generation rigorous research in terms of biological understanding of native microbes and discovery of new functions in industrially important model organisms is needed [6].

Therefore, based on their distinct but best capabilities, two anaerobic microbes were studied as model organisms in this thesis to demonstrate a possible route forwards: *Fibrobacter succinogenes* S85 (a specialised cellulose degrader) [43, 44] and *Clostridium acetobutylicum* (an efficient biofuel producer) [129, 131]. However, there are some questions that remain unanswered with these both microbes and they form the starting point of this thesis work. For instance these questions include, an unknown mechanism of cellulose degradation in *F. succinogenes* and low biofuel production titres in *C. acetobutylicum* respectively. The discovery of a cellulose degradation mechanism can provide knowledge on the novel enzymatic system involved in deconstruction (hydrolysis) of cellulosic material. At the same time, biological understanding of *C. acetobutylicum* can contribute to improved biofuel generation (fermentation). Using both microbes either separately or in combination can provide a (major) breakthrough in advanced biofuel generation from lignocellulosic biomass. To achieve aim of this PhD, the systematic application of proteomic tools (qualitative and quantitative) such as shotgun proteomics were applied. The data presented in Chapters 4-7 confirm the potential of these two microbes in future advanced biofuel generation. Moreover, the PhD also focused on the certain issues in proteomic techniques for improved protein coverage (especially separation techniques and quantitative membrane proteomics) to obtain better results.

Thus, respective studies from both microorganisms can be used for further implantation into a consolidated bioprocessing development (CBP) framework or relevant areas of exploitation. Conclusions from different studies are summarised in the subsections and key points and benefits obtained from the results were discussed.

8.1.1 Influence of lignin on proteome of C. acetobutylicum *ATCC 824*

The results obtained from the technical study (Chapter 3) on the proteome of *C. acetobutylicum* concluded that a better protein/peptide recovery can be obtained by Hydrophilic interaction liquid chromatography (HILIC) than with strong cation exchange (SCX) separation. From the results, it can be concluded that the major drawback of SCX separation is the clustering of peptides resulting in a compact distribution of all peptides in a narrrow window of the elution profile [263, 264] that affect subsequent LC-MS/MS analysis. HILIC separation is more uniform across the gradient, thus provides higher resolution and produced the maximum peptide/protein recovery of the methods tested (Please refer to Figure 3-4 and 3-6). This observation was considered to be valuable information and HILIC was used as the subsequent first dimension separation method in the quantitative proteomics studies of *C. acetobutylicum* and *F. succinogenes.*

The results obtained from the study of the influence of lignin on the metabolome and proteome of *C. acetobutylicum* (Chapter 4), revealed up/down regulation of several key proteins from different metabolic pathways including:carbohydrate metabolism, amino acid metabolism, TCA cycle and energy metabolism, cell cycle/cell division, signal transduction and chemotaxis, stress response, transcriptional and translation regulation. The results showed a deleterious effect of lignin on biofuel production (Please refer Figure 4-4, 4-5 and Table 4-1). Moreover, this is the first iTRAQ based proteomics study of *C. acetobutylicum* that sheds light on many adaptive, stress and metabolic strategies to respond under challenging environmental conditions by this bacterium.

8*.1.2 Membrane proteome study of* F. succinogenes *S85*

Since the mechanism of cellulose degradation lies within the membrane of *F. succinogenes*, a qualitative (Chapter 5) and quantitative (Chapter 6) membrane proteomics using biotinylation-iTRAQ based approach under different substrates conditions (glucose and cellulose) was carried out. From the preliminary results of a surface characterisation study, it is concluded that cellulose influences the surface topology of *F. succinogenes* S85. Colloidal surface characterisation using EPM, MATH assay and FTIR revealed a decrease in surface polysaccharides and increased in protein moieties on the cell surface upon exposure to cellulose. Thus, the study was further extended to surface-membrane biotinylation to identify proteins involved in cellulose degradation. In conclusion, a qualitative proteomic study identified several key proteins within membrane that are involved in cellulose degradation (Please refer to Table 5-1). The sub-membrane localisation (determined by PSORTb) of these proteins supported the previously proposed mechanism in this bacterium [57, 100]. However, several proteins, which are involved in cellulose degradation mechanism, were also found in glucose conditions and were the driving force behind extending the study towards the quantitative proteomic level.

An iTRAQ based quantitative study was performed on enriched membrane proteins obtained by ultracentrifugation and biotinylation. When glucose-grown cells were compared with cellulose grown cells, the maximum number of proteins were found with significant changes in biotinylation-iTRAQ than ultracentrifugation-iTRAQ quantitation (Please refer to Figure 6-4 and 6-6). It can

be concluded that biotinylation-iTRAQ analysis could be a better option for a quantitative membrane proteomics study and can be further implemented for membrane proteomics in other bacteria. The biological interpretation of data derived from biotinylation-iTRAQ quantitation (Figure 6-8) revealed several unknown proteins, including 10 proteins involved in carbohydrate degradation mechanism identified in cellulose grown cells with differential regulation. However, due to a lack of genetic information annotation about various proteins (unknown/hypothetical), it was difficult to propose possible firm mechanisms of cellulose degradation.

8.1.3 Bio-alcohol production from chemically and biologically treated biomass using C. acetobutylicum *ATCC 824*

This study demonstrated the great potential of *C. acetobutylicum* as a future biofuel generating candidate from lignocellulosic feedstock [171, 172]. The study aimed to produce biofuels from chemically and biologically hydrolysed biomass. In the first approach, our study concluded that *C. acetobutylicum* can utilise multiple sugar components as the carbon source from chemically derived miscanthus hydrolysate [468] and produced ethanol, butanol and H_2 gas (Figure 7-5). In another approach, the PhD study concluded that *F. succinogenes* and *C. acetobutylicum* can be utilised in co-culture for simultaneous saccharification and fermentation for CBP development. This was the first study so far where these two model organisms demonstrated syntrophy and produced ethanol, butanol and H_2 gas in single fermentation unit (Please refer to Figure 7-8). From the results it also can be concluded that these two mesophilic lignocellulose degrading/fermenting anaerobes can provide a promising breakthrough over

chemically achieved hydrolysis. The study can be further improved at an industrial scale as low cost technology.

8.2 Future work

The work in this study shed light on various useful aspects of technical development for proteomic studies such as peptide separation techniques to obtain improved protein recovery (Chapter 3) and development of strategies for membrane proteomics (Chapter 5 and 6). At the same time the study provided valuable insights in the cellulose degradation mechanism in *F. succinogenes* (such as membrane proteome involved in cellulose degradation and related membrane activities) and biofuel generation in *C. acetobutylicum* (possible effect of lignin on biofuel related pathways and adaptive responses). The proteomic study performed on both the bacteria provided information about influence of different substrate conditions on metabolic processes and also how we can utilise the individual capabilities of both these bacteria for improved biofuel generation. However, the study carried out in this project did not try to cover everything and there are various issues that need to be focused on in the future study.

The technical comparison between SCX and HILIC was particularly important, since the first dimension separation of peptide plays a crucial role in shotgun proteomics in terms of protein coverage by LC-MS/MS analysis. The technical comparison between SCX and HILIC (Chapter 3) shows that HILIC is a more suitable peptide separation technique compared to SCX. However, SCX has been most widely used first dimension separation in gel-free / shotgun proteomics [251, 252] and always been a priority for researchers. It means that in the future, HILIC can be an excellent alternative method in multidimensional

chromatography that can be used in analysis of protein modifications such as phosphoproteomics and glycoproteomics [197].

The influence of lignin on the *C. acetobutylicum* proteome (Chapter 4) shows a deleterious effect of lignin on various metabolic pathways and biofuel generation, but at the same time study also provided the adaptive response of *C. acetobutylicum* towards a lignin stress condition. The study was particularly important for metabolic engineering of *C. acetobutylicum* ATCC 824 for enhanced tolerance against lignocellulose derived microbial inhibitory compounds (e.g. furfural) or biotransformation of lignin compounds in this bacterium. This can be either achieved by improving *C. acetobutylicum* by recombinant technology or using other lignin degrading microbes that help to improve the biological conversion of lignocelluloses biomass hydrolysate to biofuel. It means that a future study must focus on these two aspects either remove lignin from biomass hydrolysate or biotransformed them to less inhibitory compounds. Influence of lignocellulose-derived compounds i.e. furfural and its biotransformation into less inhibitory compound were studied in other Clostridia species [282, 283]. Interestingly, our study also identified a significantly up-regulated protein, i.e. possible pectin degrading protein (CA_C3376) that possesses a 2 barrel cupin domain. The role of cupin (conserved beta barrel fold) with dioxygenases in aromatic ring degradation have been studied in soil bacteria [330]. Aromatic compound (2,4,6 trinitrotoluene (TNT)) degradation was extensively studied in *C. acetobutylicum*[291-293, 499-501]and that proved that it might have a unique metabolic pathway for the degradation of aromatic compounds, including lignin. Further characterization of this particular protein could provide new insights that may be useful for lignin to biofuels efforts.

In addition, metabolic engineering of *C. acetobutylicum* to improve biofuel production is another interesting area of research [129, 157]. The development of a recombinant strain of *C. acetobutylicum* ATCC 824, capable of cellulose degradation and fermentation [502] is an interesting area of research for CBP development.

Regarding the cellulose degradation capacity of *F. succinogenes*, the PhD study discusses importance of membrane activity. Data from the preliminary study of qualitative (Chapter 5) and quantitative membrane proteomics (Chapter 6) suggests further implementation of this approach at the proteomics as well as genomic level is needed, since most of the cellulose degradation related proteins were also observed in glucose grown cells. Moreover, a significant change in putative/uncharacterized membrane proteins suggests use of homology based searches in relevant organisms. The functions of these proteins need to be known in order to crack the mechanism of cellulose degradation. This will be the next step in this research. The analytical techniques that are modified in this study (particularly biotinylation enrichment with quantitative analysis) can be further implemented in order to overcome challenges with membrane proteomics that can be further applicable for other microbes.

Commercialization of microbial based biofuel production (CBP) needs development of low cost technology. The results from the preliminary study (Chapter 7) suggest that biofuel can be produced using chemical/biological saccharification and fermentation. However, future work should be focused on biological saccharification and fermentation (co-culture treatment for CBP) since it can provide economically feasible and eco-friendly technology for lignocellulosic biofuel generation.

References

1. Lynd, L.R., W.H.v. Zyl, J.E. McBRIDE, and M. Laser, *Consolidated bioprocessing of cellulosic biomass: an update.* Current Opinion in Biotechnology, 2005. **16**(5): p. 577-583.
2. Lettinga, G. and A.v. Haandel, *Anaerobic digestion for energy production and environmental protection.* Renewable energy: sources for fuels and electricity., 1993: p. 817-839.
3. Sialve, B., N. Bernet, and O. Bernard, *Anaerobic digestion of microalgae as a necessary step to make microalgal biodiesel sustainable.* Biotechnology Advances, 2009. **27**(4): p. 409-416.
4. Sabra, W., D. Dietz, D. Tjahjasari, and A.P. Zeng, *Biosystems analysis and engineering of microbial consortia for industrial biotechnology.* Engineering in Life Sciences, 2010. **10**(5): p. 407-421.
5. Lynd, L.R., P.J. Weimer, W.H. Van Zyl, and I.S. Pretorius, *Microbial cellulose utilization: fundamentals and biotechnology.* Microbiology and molecular biology reviews, 2002. **66**(3): p. 506-577.
6. Alper, H. and G. Stephanopoulos, *Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential?* Nature Reviews Microbiology, 2009. **7**(10): p. 715-723.
7. Collins, D., *International Energy Agency's "Bioenergy Task 43".* The Forestry Chronicle, 2013. **89**(3): p. 277-278.
8. Administration, E.I., *International Energy Outlook 2013 With Projections to 2040*2013: Government Printing Office.
9. Bauen, A., G. Berndes, M. Junginger, M. Londo, F. Vuille, R. Ball, T. Bole, C. Chudziak, A. Faaij, and H. Mozaffarian, *Bioenergy: a sustainable and reliable energy source. A review of status and prospects.* Bioenergy: a sustainable and reliable energy source. A review of status and prospects, 2009.
10. Duku, M.H., S. Gu, and E.B. Hagan, *A comprehensive review of biomass resources and biofuels potential in Ghana.* Renewable and Sustainable Energy Reviews, 2011. **15**(1): p. 404-415.
11. Havlik, P., U.A. Schneider, E. Schmid, H. Böttcher, S. Fritz, R. Skalský, K. Aoki, S.D. Cara, G. Kindermann, and F. Kraxner, *Global land-use implications of first and second generation biofuel targets.* Energy Policy, 2011. **39**(10): p. 5690-5702.
12. Rubin, E.M., *Genomics of cellulosic biofuels.* Nature, 2008. **454**(7206): p. 841-845.
13. Tilman, D., J. Hill, and C. Lehman, *Carbon-negative biofuels from low-input high-diversity grassland biomass.* Science, 2006. **314**(5805): p. 1598-1600.
14. Mukhopadhyay, A., A.M. Redding, B.J. Rutherford, and J.D. Keasling, *Importance of systems biology in engineering microbes for biofuel production.* Current Opinion in Biotechnology, 2008. **19**(3): p. 228-234.
15. de Souza, W.R., *Microbial Degradation of Lignocellulosic Biomass.* 2013.
16. Pienkos, P.T. and A. Darzins, *The promise and challenges of microalgal-derived biofuels.* Biofuels, Bioproducts and Biorefining, 2009. **3**(4): p. 431-440.
17. Scharf, M.E. and A. Tartar, *Termite digestomes as sources for novel lignocellulases.* Biofuels, Bioproducts and Biorefining, 2008. **2**(6): p. 540-552.
18. Ljungdahl, L. and K.-E. Eriksson, *Ecology of microbial cellulose degradation.* Advances in microbial ecology, 1985. **8**: p. 237-299.
19. Saha, B.C. *Lignocellulose biodegradation and applications in biotechnology*. in *ACS symposium series*. 2004. ACS Publications.

20. Aristidou, A. and M. Penttila, *Metabolic engineering applications to renewable resource utilization.* Current Opinion in Biotechnology, 2000. **11**(2): p. 187-198.
21. Vogel, J., *Unique aspects of the grass cell wall.* Current opinion in plant biology, 2008. **11**(3): p. 301-307.
22. Van Wyk, J.P., *Biotechnology and the utilization of biowaste as a resource for bioproduct development.* TRENDS in Biotechnology, 2001. **19**(5): p. 172-177.
23. de Vries, R.P. and J. Visser, *Aspergillus enzymes involved in degradation of plant cell wall polysaccharides.* Microbiology and molecular biology reviews, 2001. **65**(4): p. 497-522.
24. Saha, B.C., *Hemicellulose bioconversion.* Journal of Industrial Microbiology and Biotechnology, 2003. **30**(5): p. 279-291.
25. Wingren, A., M. Galbe, and G. Zacchi, *Techno-Economic Evaluation of Producing Ethanol from Softwood: Comparison of SSF and SHF and Identification of Bottlenecks.* Biotechnology progress, 2003. **19**(4): p. 1109-1117.
26. Sticklen, M.B., *Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol.* Nature Reviews Genetics, 2008. **9**(6): p. 433-443.
27. Del Rio, J.C., G. Marques, J. Rencoret, A.T. Martinez, and A. Gutierrez, *Occurrence of naturally acetylated lignin units.* Journal of agricultural and food chemistry, 2007. **55**(14): p. 5461-5468.
28. Aro, N., T. Pakula, and M. Penttila, *Transcriptional regulation of plant cell wall degradation by filamentous fungi.* FEMS microbiology reviews, 2005. **29**(4): p. 719-739.
29. Bayer, E.A., H. Chanzy, R. Lamed, and Y. Shoham, *Cellulose, cellulases and cellulosomes.* Current opinion in structural biology, 1998. **8**(5): p. 548-557.
30. Warren, R., *Microbial hydrolysis of polysaccharides.* Annual Reviews in Microbiology, 1996. **50**(1): p. 183-212.
31. Reinhold-Hurek, B., T. Hurek, M. Claeyssens, and M. Van Montagu, *Cloning, expression in Escherichia coli, and characterization of cellulolytic enzymes of Azoarcus sp., a root-invading diazotroph.* Journal of bacteriology, 1993. **175**(21): p. 7056-7065.
32. Thygesen, A., A.B. Thomsen, A.S. Schmidt, H. Jorgensen, B.K. Ahring, and L. Olsson, *Production of cellulose and hemicellulose-degrading enzymes by filamentous fungi cultivated on wet-oxidised wheat straw.* Enzyme and microbial technology, 2003. **32**(5): p. 606-615.
33. Kirby, R., *Actinomycetes and lignin degradation.* Advances in applied microbiology, 2005. **58**: p. 125-168.
34. Li, J., H. Yuan, and J. Yang, *Bacteria and lignin degradation.* Frontiers of Biology in China, 2009. **4**(1): p. 29-38.
35. Nie, G., N.S. Reading, and S.D. Aust, *Relative Stability of Recombinant Versus Native Peroxidases from Phanerochaete chrysosporium* Archives of biochemistry and biophysics, 1999. **365**(2): p. 328-334.
36. Schmeling, S., A. Narmandakh, O. Schmitt, K. Schuhle, and G. Fuchs, *Phenylphosphate synthase: a new phosphotransferase catalyzing the first step in anaerobic phenol metabolism in Thauera aromatica.* Journal of bacteriology, 2004. **186**(23): p. 8044-8057.
37. Lack, A., I. Tommasi, M. Aresta, and G. Fuchs, *Catalytic properties of phenol carboxylase In vitro study of CO2: 4-hydroxybenzoate isotope exchange reaction.* European Journal of Biochemistry, 1991. **197**(2): p. 473-479.
38. Fierobe, H.-P., A. Mechaly, C. Tardif, A. Belaich, R. Lamed, Y. Shoham, J.-P. Belaich, and E.A. Bayer, *Design and production of active cellulosome chimeras Selective incorporation of dockerin-containing enzymes into defined functional complexes.* Journal of Biological Chemistry, 2001. **276**(24): p. 21257-21261.
39. Jørgensen, H., J.B. Kristensen, and C. Felby, *Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities.* Biofuels, Bioproducts and Biorefining, 2007. **1**(2): p. 119-134.

40. Farrell, A.E., R.J. Plevin, B.T. Turner, A.D. Jones, M. O'hare, and D.M. Kammen, *Ethanol can contribute to energy and environmental goals.* Science, 2006. **311**(5760): p. 506-508.
41. Sun, Y. and J. Cheng, *Hydrolysis of lignocellulosic materials for ethanol production: a review.* Bioresource technology, 2002. **83**(1): p. 1-11.
42. Gilbert, H.J., *Cellulosomes: microbial nanomachines that display plasticity in quaternary structure.* Molecular microbiology, 2007. **63**(6): p. 1568-1576.
43. Suen, G., P.J. Weimer, D.M. Stevenson, F.O. Aylward, J. Boyum, J. Deneke, C. Drinkwater, N.N. Ivanova, N. Mikhailova, and O. Chertkov, *The complete genome sequence of Fibrobacter succinogenes S85 reveals a cellulolytic and metabolic specialist.* PloS one, 2011. **6**(4): p. e18814.
44. Jun, H.-S., M. Qi, J. Gong, E.E. Egbosimba, and C.W. Forsberg, *Outer membrane proteins of Fibrobacter succinogenes with potential roles in adhesion to cellulose and in cellulose digestion.* Journal of bacteriology, 2007. **189**(19): p. 6806-6815.
45. Kato, S., S. Haruta, Z.J. Cui, M. Ishii, and Y. Igarashi, *Effective cellulose degradation by a mixed-culture system composed of a cellulolytic Clostridium and aerobic non-cellulolytic bacteria.* FEMS Microbiology Ecology, 2004. **51**(1): p. 133-142.
46. Leschine, S.B., *Cellulose degradation in anaerobic environments.* Annual Reviews in Microbiology, 1995. **49**(1): p. 399-426.
47. Leschine, S. and E. Canale-Parola, *Mesophilic cellulolytic clostridia from freshwater environments.* Applied and environmental microbiology, 1983. **46**(3): p. 728-737.
48. Van Dyke, M. and A. McCarthy, *Molecular biological detection and characterization of Clostridium populations in municipal landfill sites.* Applied and environmental microbiology, 2002. **68**(4): p. 2049-2053.
49. Lee, S.K., H. Chou, T.S. Ham, T.S. Lee, and J.D. Keasling, *Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels.* Current Opinion in Biotechnology, 2008. **19**(6): p. 556-563.
50. Bayer, E.A. and R. Lamed, *The cellulose paradox: pollutant par excellence and/or a reclaimable natural resource?*, in *Microorganisms to Combat Pollution*1993, Springer. p. 47-64.
51. Mingardon, F., A. Chanal, A.M. Lopez-Contreras, C. Dray, E.A. Bayer, and H.-P. Fierobe, *Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes.* Applied and environmental microbiology, 2007. **73**(12): p. 3822-3832.
52. Beguin, P., *Molecular biology of cellulose degradation.* Annual Reviews in Microbiology, 1990. **44**(1): p. 219-248.
53. Schwarz, W., *The cellulosome and cellulose degradation by anaerobic bacteria.* Applied microbiology and biotechnology, 2001. **56**(5-6): p. 634-649.
54. Wilson, D.B., *Aerobic microbial cellulase systems.* Biomass recalcitrance: deconstructing the plant cell wall for bioenergy, 2008: p. 374-392.
55. Bayer, E.A., J.-P. Belaich, Y. Shoham, and R. Lamed, *The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides.* Annu. Rev. Microbiol., 2004. **58**: p. 521-554.
56. Martinez, D., J. Challacombe, I. Morgenstern, D. Hibbett, M. Schmoll, C.P. Kubicek, P. Ferreira, F.J. Ruiz-Duenas, A.T. Martinez, and P. Kersten, *Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion.* Proceedings of the National Academy of Sciences, 2009. **106**(6): p. 1954-1959.
57. Wilson, D.B., *Evidence for a novel mechanism of microbial cellulose degradation.* Cellulose, 2009. **16**(4): p. 723-727.

58. Ruminants, N.R.C.C.o.N.R.o.S., *Nutrient requirements of small ruminants: sheep, goats, cervids, and new world camelids*2007.
59. Hungate, R., *The anaerobic mesophilic cellulolytic bacteria.* Bacteriological reviews, 1950. **14**(1): p. 1.
60. Russell, J.B., *Rumen microbiology and its role in ruminant nutrition*2002: Cornell University.
61. Weimer, P.J., J.B. Russell, and R.E. Muck, *Lessons from the cow: what the ruminant animal can teach us about consolidated bioprocessing of cellulosic biomass.* Bioresource technology, 2009. **100**(21): p. 5323-5331.
62. Morrison, M., P.B. Pope, S.E. Denman, and C.S. McSweeney, *Plant biomass degradation by gut microbiomes: more of the same or something new?* Current Opinion in Biotechnology, 2009. **20**(3): p. 358-363.
63. Flint, H.J., *The rumen microbial ecosystem—some recent developments.* Trends in microbiology, 1997. **5**(12): p. 483-488.
64. Flint, H.J., E.A. Bayer, M.T. Rincon, R. Lamed, and B.A. White, *Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis.* Nature Reviews Microbiology, 2008. **6**(2): p. 121-131.
65. Orpin, C. and K. Joblin, *The rumen anaerobic fungi*, in *The rumen microbial ecosystem*1997, Springer. p. 140-195.
66. Stewart, C., H. Flint, and M. Bryant, *The rumen bacteria*, in *The rumen microbial ecosystem*1997, Springer. p. 10-72.
67. Williams, A.G. and G. Coleman, *The rumen protozoa*1997: Springer.
68. Krause, D.O., S.E. Denman, R.I. Mackie, M. Morrison, A.L. Rae, G.T. Attwood, and C.S. McSweeney, *Opportunities to improve fiber degradation in the rumen: microbiology, ecology, and genomics.* FEMS microbiology reviews, 2003. **27**(5): p. 663-693.
69. Toyoda, A., W. Iio, M. Mitsumori, and H. Minato, *Isolation and identification of cellulose-binding proteins from sheep rumen contents.* Applied and environmental microbiology, 2009. **75**(6): p. 1667-1673.
70. Forster, R.J., J. Gong, and R.M. Teather, *Group-specific 16S rRNA hybridization probes for determinative and community structure studies of Butyrivibrio fibrisolvens in the rumen.* Applied and environmental microbiology, 1997. **63**(4): p. 1256-1260.
71. Shinkai, T. and Y. Kobayashi, *Localization of ruminal cellulolytic bacteria on plant fibrous materials as determined by fluorescence in situ hybridization and real-time PCR.* Applied and environmental microbiology, 2007. **73**(5): p. 1646-1652.
72. Jungblut, P.R., E.-C. Müller, J. Mattow, and S.H. Kaufmann, *Proteomics Reveals Open Reading Frames inMycobacterium tuberculosis H37Rv Not Predicted by Genomics.* Infection and immunity, 2001. **69**(9): p. 5905-5907.
73. Michalet-Doreau, B., I. Fernandez, C. Peyron, L. Millet, and G. Fonty, *Fibrolytic activities and cellulolytic bacterial community structure in the solid and liquid phases of rumen contents.* Reproduction Nutrition Development, 2001. **41**(2): p. 187.
74. Tajima, K., R. Aminov, T. Nagamine, H. Matsui, M. Nakamura, and Y. Benno, *Diet-dependent shifts in the bacterial population of the rumen revealed with real-time PCR.* Applied and environmental microbiology, 2001. **67**(6): p. 2766-2774.
75. Ziemer, C.J., R. Sharp, M.D. Stern, M.A. Cotta, T.R. Whitehead, and D.A. Stahl, *Comparison of microbial populations in model and natural rumens using 16S ribosomal RNA-targeted probes.* Environmental microbiology, 2000. **2**(6): p. 632-643.
76. Dehority, B.A., *Rumen microbiology*. Vol. 372. 2003: Nottingham University Press Nottingham.
77. Koike, S. and Y. Kobayashi, *Development and use of competitive PCR assays for the rumen cellulolytic bacteria: Fibrobacter succinogenes, Ruminococcus albus and Ruminococcus flavefaciens.* FEMS microbiology letters, 2001. **204**(2): p. 361-366.

78. Hastie, P.M., K. Mitchell, and J.-A. Murray, *Semi-quantitative analysis of Ruminococcus flavefaciens, Fibrobacter succinogenes and Streptococcus bovis in the equine large intestine using real-time polymerase chain reaction.* British Journal of Nutrition, 2008. **100**(03): p. 561-568.
79. Hongoh, Y., P. Deevong, S. Hattori, T. Inoue, S. Noda, N. Noparatnaraporn, T. Kudo, and M. Ohkuma, *Phylogenetic diversity, localization, and cell morphologies of members of the candidate phylum TG3 and a subphylum in the phylum Fibrobacteres, recently discovered bacterial groups dominant in termite guts.* Applied and environmental microbiology, 2006. **72**(10): p. 6780-6788.
80. McDonald, J.E., R.J. Lockhart, M.J. Cox, H.E. Allison, and A.J. McCarthy, *Detection of novel Fibrobacter populations in landfill sites and determination of their relative abundance via quantitative PCR.* Environmental microbiology, 2008. **10**(5): p. 1310-1319.
81. McDonald, J.E., A.B. de Menezes, H.E. Allison, and A.J. McCarthy, *Molecular biological detection and quantification of novel Fibrobacter populations in freshwater lakes.* Applied and environmental microbiology, 2009. **75**(15): p. 5148-5152.
82. Chen, X., J. Wang, Y. Wu, and J. Liu, *Effects of chemical treatments of rice straw on rumen fermentation characteristics, fibrolytic enzyme activities and populations of liquid-and solid-associated ruminal microbes in vitro.* Animal Feed Science and Technology, 2008. **141**(1): p. 1-14.
83. Gupta, R.S., *The phylogeny and signature sequences characteristics of Fibrobacteres, Chlorobi, and Bacteroidetes.* Critical reviews in microbiology, 2004. **30**(2): p. 123-143.
84. Montogomery, L., B. Flesher, and D. Stahl, *Transfer of Bacteroides succinogenes (Hungate) to Fibrobacter succinogenes gen. nov. as Fibrobacter succinogenes comb. nov. and description of Fibrobacter intestinalis sp. nov.* Int. J. Syst. Bacteriol, 1988. **38**: p. 430-435.
85. Qi, M., K.E. Nelson, S.C. Daugherty, W.C. Nelson, I.R. Hance, M. Morrison, and C.W. Forsberg, *Novel molecular features of the fibrolytic intestinal bacterium Fibrobacter intestinalis not shared with Fibrobacter succinogenes as determined by suppressive subtractive hybridization.* Journal of bacteriology, 2005. **187**(11): p. 3739-3751.
86. Lin, C. and D.A. Stahl, *Comparative analyses reveal a highly conserved endoglucanase in the cellulolytic genus Fibrobacter.* Journal of bacteriology, 1995. **177**(9): p. 2543-2549.
87. Béra-Maillet, C., Y. Ribot, and E. Forano, *Fiber-degrading systems of different strains of the genus Fibrobacter.* Applied and environmental microbiology, 2004. **70**(4): p. 2172-2179.
88. Amann, R.I., C. Lin, R. Key, L. Montgomery, and D.A. Stahl, *Diversity Among Fibrobacter Isolates: Towards a Phylogenetic Classification.* Systematic and applied microbiology, 1992. **15**(1): p. 23-31.
89. Shinkai, T., R. Ohji, N. Matsumoto, and Y. Kobayashi, *Fibrolytic capabilities of ruminal bacterium Fibrobacter succinogenes in relation to its phylogenetic grouping.* FEMS microbiology letters, 2009. **294**(2): p. 183-190.
90. Russell, J.B., R.E. Muck, and P.J. Weimer, *Quantitative analysis of cellulose degradation and growth of cellulolytic bacteria in the rumen.* FEMS Microbiology Ecology, 2009. **67**(2): p. 183-197.
91. Cheng, K.-J., C. Stewart, D. Dinsdale, and J. Costerton, *Electron microscopy of bacteria involved in the digestion of plant cell walls.* Animal Feed Science and Technology, 1984. **10**(2): p. 93-120.
92. Miron, J. and D. Ben-Ghedalia, *The degradation and utilization of wheat-straw cell-wall monosaccharide components by defined ruminal cellulolytic bacteria.* Applied microbiology and biotechnology, 1992. **38**(3): p. 432-437.

93. Forano, E., A.-M. Delort, and M. Matulova, *Carbohydrate metabolism in Fibrobacter succinogenes: What NMR tells us.* Microbial Ecology in Health and Disease, 2008. **20**(2): p. 94-102.
94. Gaudet, G., E. Forano, G. Dauphin, and A.M. DELORT, *Futile cycling of glycogen in Fibrobacter succinogenes as shown by in situ1H-NMR and 13C-NMR investigation.* European Journal of Biochemistry, 1992. **207**(1): p. 155-162.
95. Matulova, M., R. Nouaille, P. Capek, M. Péan, E. Forano, and A.-M. Delort, *Degradation of wheat straw by Fibrobacter succinogenes S85: a liquid-and solid-state nuclear magnetic resonance study.* Applied and environmental microbiology, 2005. **71**(3): p. 1247-1253.
96. Doi, R.H. and A. Kosugi, *Cellulosomes: plant-cell-wall-degrading enzyme complexes.* Nature Reviews Microbiology, 2004. **2**(7): p. 541-551.
97. Miron, J., D. Ben-Ghedalia, and M. Morrison, *Invited Review: Adhesion Mechanisms of Rumen Cellulolytic Bacteria.* Journal of Dairy Science, 2001. **84**(6): p. 1294-1309.
98. Kudo, H., K.-J. Cheng, and J. Costerton, *Electron microscopic study of the methylcellulose-mediated detachment of cellulolytic rumen bacteria from cellulose fibers.* Canadian journal of microbiology, 1987. **33**(3): p. 267-272.
99. Weimer, P.J. and C.L. Odt. *Cellulose degradation by ruminal microbes: physiological and hydrolytic diversity among ruminal cellulolytic bacteria*. in *ACS Symposium Series*. 1995. ACS Publications.
100. Ransom-Jones, E., D.L. Jones, A.J. McCarthy, and J.E. McDonald, *The Fibrobacteres: an important phylum of cellulose-degrading bacteria.* Microbial ecology, 2012. **63**(2): p. 267-281.
101. Matheron, C., A.-M. Delort, G. Gaudet, and E. Forano, *Simultaneous but differential metabolism of glucose and cellobiose in Fibrobacter succinogenes cells, studied by in vivo 13C-NMR.* Canadian journal of microbiology, 1996. **42**(11): p. 1091-1099.
102. Wells, J.E., J.B. Russell, Y. Shi, and P.J. Weimer, *Cellodextrin efflux by the cellulolytic ruminal bacterium Fibrobacter succinogenes and its potential role in the growth of nonadherent bacteria.* Applied and environmental microbiology, 1995. **61**(5): p. 1757-1762.
103. Matheron, C., A.-M. Delort, G.v. Gaudet, and E. Forano, *Simultaneous but differential metabolism of glucose and cellobiose in Fibrobacter succinogenes cells, studied by in vivo 13C-NMR.* Canadian journal of microbiology, 1996. **42**(11): p. 1091-1099.
104. Jun, H., M. Qi, J. Ha, and C. Forsberg, *Fibrobacter succinogenes, a dominant fibrolytic ruminal bacterium: transition to the post genomic era.* Asian Australasian Journal of Animal Science, 2007. **20**(5): p. 802.
105. Wilson, D.B., *Three microbial strategies for plant cell wall degradation.* Annals of the New York Academy of Sciences, 2008. **1125**(1): p. 289-297.
106. Forsberg, C., E. Forano, and A. Chesson, *Microbial adherence to the plant cell wall and enzymatic hydrolysis.* Ruminant Physiology: Digestion, Metabolism, Growth and Reproduction. CABI Publishing, Wallingford, UK, 2000: p. 79-97.
107. Gong, J., E.E. Egbosimba, and C.W. Forsberg, *Cellulose-binding proteins of Fibrobacter succinogenes and the possible role of a 180-kDa cellulose-binding glycoprotein in adhesion to cellulose.* Canadian journal of microbiology, 1996. **42**(5): p. 453-460.
108. Arai, T., R. Araki, A. Tanaka, S. Karita, T. Kimura, K. Sakka, and K. Ohmiya, *Characterization of a cellulase containing a family 30 carbohydrate-binding module (CBM) derived from Clostridium thermocellum CelJ: importance of the CBM to cellulose hydrolysis.* Journal of bacteriology, 2003. **185**(2): p. 504-512.
109. Yoda, K., A. Toyoda, Y. Mukoyama, Y. Nakamura, and H. Minato, *Cloning, sequencing, and expression of a Eubacterium cellulosolvens 5 gene encoding an endoglucanase*

(Cel5A) with novel carbohydrate-binding modules, and properties of Cel5A. Applied and environmental microbiology, 2005. **71**(10): p. 5787-5793.
110. Gong, J. and C.W. Forsberg, *Factors affecting adhesion of Fibrobacter succinogenes subsp. succinogenes S85 and adherence-defective mutants to cellulose.* Applied and environmental microbiology, 1989. **55**(12): p. 3039-3044.
111. Elshahed, M.S., *Microbiological aspects of biofuel production: Current status and future directions.* Journal of Advanced Research, 2010. **1**(2): p. 103-111.
112. Bohutskyi, P. and E. Bouwer, *Biogas Production from Algae and Cyanobacteria Through Anaerobic Digestion: A Review, Analysis, and Research Needs*, in *Advanced Biofuels and Bioproducts*2013, Springer. p. 873-975.
113. Mata-Alvarez, J., S. Mace, and P. Llabres, *Anaerobic digestion of organic solid wastes. An overview of research achievements and perspectives.* Bioresource technology, 2000. **74**(1): p. 3-16.
114. Lawson, P.A., T.D. Allen, M.E. Caldwell, and R.S. Tanner, *Anaerobes: a piece in the puzzle for alternative biofuels.* Anaerobe, 2011. **17**(4): p. 206-210.
115. George, S., *Organic waste.* 2012.
116. George, S., *Change needed to encourage competition in organic waste treatment, OFT finds.* 2012.
117. Holm-Nielsen, J.B., T. Al Seadi, and P. Oleskowicz-Popiel, *The future of anaerobic digestion and biogas utilization.* Bioresource technology, 2009. **100**(22): p. 5478-5484.
118. Murphy, C.D., *The microbial cell factory.* Organic & biomolecular chemistry, 2012. **10**(10): p. 1949-1957.
119. Guillemont, J.E.G., P.J.-m. Raboisson, and N. Lounis, *Compounds with antibacterial activity against Clostridium*, 2013, US Patent 20,130,109,694.
120. Rao, R.P., N. Dufour, and J. Swana, *Using microorganisms to brew biofuels.* In Vitro Cellular & Developmental Biology-Plant, 2011. **47**(6): p. 637-649.
121. Cato, E., W.L. George, and S. Finegold, *Genus clostridium.* Bergey's manual of systematic bacteriology, 1986. **2**: p. 1141-1200.
122. Hyun, H. and J. Zeikus, *General biochemical characterization of thermostable pullulanase and glucoamylase from Clostridium thermohydrosulfuricum.* Applied and environmental microbiology, 1985. **49**(5): p. 1168-1173.
123. Hyun, H. and J. Zeikus, *Simultaneous and enhanced production of thermostable amylases and ethanol from starch by cocultures of Clostridium thermosulfurogenes and Clostridium thermohydrosulfuricum.* Applied and environmental microbiology, 1985. **49**(5): p. 1174-1181.
124. Mitchell, W.J., *Physiology of carbohydrate to solvent conversion by clostridia.* Advances in microbial physiology, 1997. **39**: p. 31-130.
125. Mousdale, D.M., *Biofuels: biotechnology, chemistry, and sustainable development*2008: CRC press.
126. Jones, D.T. and D.R. Woods, *Acetone-butanol fermentation revisited.* Microbiological reviews, 1986. **50**(4): p. 484.
127. Mishra, P. and A. Singh, *Microbial pentose utilization.* Advances in applied microbiology, 1993. **39**: p. 91-91.
128. Ni, Y. and Z. Sun, *Recent progress on industrial fermentative production of acetone–butanol–ethanol by Clostridium acetobutylicum in China.* Applied microbiology and biotechnology, 2009. **83**(3): p. 415-423.
129. Lütke-Eversloh, T. and H. Bahl, *Metabolic engineering of Clostridium acetobutylicum: recent advances to improve butanol production.* Current Opinion in Biotechnology, 2011. **22**(5): p. 634-647.
130. Jones, D.T. and S. Keis, *Origins and relationships of industrial solvent-producing clostridial strains.* FEMS microbiology reviews, 1995. **17**(3): p. 223-232.

131. Papoutsakis, E.T., *Engineering solventogenic clostridia.* Current Opinion in Biotechnology, 2008. **19**(5): p. 420-429.
132. Ezeji, T. and H.P. Blaschek, *Fermentation of dried distillers' grains and solubles (DDGS) hydrolysates to solvents and value-added products by solventogenic clostridia.* Bioresource technology, 2008. **99**(12): p. 5232-5242.
133. Qureshi, N., X.L. Li, S. Hughes, B.C. Saha, and M.A. Cotta, *Butanol production from corn fiber xylan using Clostridium acetobutylicum.* Biotechnology progress, 2006. **22**(3): p. 673-680.
134. Nölling, J., G. Breton, M.V. Omelchenko, K.S. Makarova, Q. Zeng, R. Gibson, H.M. Lee, J. Dubois, D. Qiu, and J. Hitti, *Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum.* Journal of bacteriology, 2001. **183**(16): p. 4823-4838.
135. Lee, S.F., C.W. Forsberg, and L. Gibbins, *Cellulolytic activity of Clostridium acetobutylicum.* Applied and environmental microbiology, 1985. **50**(2): p. 220-228.
136. Sabathe, F. and P. Soucaille, *Characterization of the cellulolytic complex (cellulosome) of Clostridium acetobutylicum.* FEMS microbiology letters, 2002. **217**(1): p. 15-22.
137. Behrens, S., W.J. Mitchell, and H. Bahl, *Molecular analysis of the mannitol operon of Clostridium acetobutylicum encoding a phosphotransferase system and a putative PTS-modulated regulator.* Microbiology, 2001. **147**(1): p. 75-86.
138. Dürre, P., *Formation of solvents in clostridia*, 2005, Boca Raton, FL, USA: CRC Press. p. 671-693.
139. Grossman, A.D., *Genetic networks controlling the initiation of sporulation and the development of genetic competence in Bacillus subtilis.* Annual review of genetics, 1995. **29**(1): p. 477-508.
140. Steiner, E., J. Scott, N.P. Minton, and K. Winzer, *An agr quorum sensing system that regulates granulose formation and sporulation in Clostridium acetobutylicum.* Applied and environmental microbiology, 2012. **78**(4): p. 1113-1122.
141. Harris, L.M., N.E. Welker, and E.T. Papoutsakis, *Northern, morphological, and fermentation analysis of spo0A inactivation and overexpression in Clostridium acetobutylicum ATCC 824.* Journal of bacteriology, 2002. **184**(13): p. 3586-3597.
142. Ravagnani, A., K.C. Jennert, E. Steiner, R. Grünberg, J.R. Jefferies, S.R. Wilkinson, D.I. Young, E.C. Tidswell, D.P. Brown, and P. Youngman, *Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia.* Molecular microbiology, 2000. **37**(5): p. 1172-1185.
143. Paredes, C.J., K.V. Alsaker, and E.T. Papoutsakis, *A comparative genomic view of clostridial sporulation and physiology.* Nature Reviews Microbiology, 2005. **3**(12): p. 969-978.
144. Jones, S.W., C.J. Paredes, B. Tracy, N. Cheng, R. Sillers, R.S. Senger, and E.T. Papoutsakis, *The transcriptional program underlying the physiology of clostridial sporulation.* Genome Biol, 2008. **9**(7): p. R114.
145. Grupe, H. and G. Gottschalk, *Physiological events in Clostridium acetobutylicum during the shift from acidogenesis to solventogenesis in continuous culture and presentation of a model for shift induction.* Applied and environmental microbiology, 1992. **58**(12): p. 3896-3902.
146. Rao, G., P. Ward, and R. Mutharasan, *Manipulation of End-Product Distribution in Strict Anaerobes.* Annals of the New York Academy of Sciences, 1987. **506**(1): p. 76-83.
147. Girbal, L., C. Croux, I. Vasconcelos, and P. Soucaille, *Regulation of metabolic shifts in Clostridium acetobutylicum ATCC 824.* FEMS microbiology reviews, 1995. **17**(3): p. 287-297.
148. Girbal, L. and P. Soucaille, *Regulation of solvent production in Clostridium acetobutylicum.* TRENDS in Biotechnology, 1998. **16**(1): p. 11-16.

149. Nakayama, S.-i., T. Kosaka, H. Hirakawa, K. Matsuura, S. Yoshino, and K. Furukawa, *Metabolic engineering for solvent productivity by downregulation of the hydrogenase gene cluster hupCBA in Clostridium saccharoperbutylacetonicum strain N1-4.* Applied microbiology and biotechnology, 2008. **78**(3): p. 483-493.
150. Ezeji, T.C., N. Qureshi, and H.P. Blaschek, *Bioproduction of butanol from biomass: from genes to bioreactors.* Current Opinion in Biotechnology, 2007. **18**(3): p. 220-227.
151. Wiehn, M., K. Staggs, Y. Wang, and D.R. Nielsen, *In situ butanol recovery from Clostridium acetobutylicum fermentations by expanded bed adsorption.* Biotechnology progress, 2013.
152. Hillmann, F., C. Döring, O. Riebe, A. Ehrenreich, R.-J. Fischer, and H. Bahl, *The role of PerR in O2-affected gene expression of Clostridium acetobutylicum.* Journal of bacteriology, 2009. **191**(19): p. 6082-6093.
153. Schaffer, S., N. Isci, B. Zickner, and P. Dürre, *Changes in protein synthesis and identification of proteins specifically induced during solventogenesis in Clostridium acetobutylicum.* Electrophoresis, 2002. **23**(1): p. 110-121.
154. Mao, S., Y. Luo, T. Zhang, J. Li, G. Bao, Y. Zhu, Z. Chen, Y. Zhang, Y. Li, and Y. Ma, *Proteome reference map and comparative proteomic analysis between a wild type Clostridium acetobutylicum DSM 1731 and its mutant with enhanced butanol tolerance and butanol yield.* Journal of proteome research, 2010. **9**(6): p. 3046-3061.
155. Tomas, C.A., J. Beamish, and E.T. Papoutsakis, *Transcriptional analysis of butanol stress and tolerance in Clostridium acetobutylicum.* Journal of bacteriology, 2004. **186**(7): p. 2006-2018.
156. Janssen, H., C. Döring, A. Ehrenreich, B. Voigt, M. Hecker, H. Bahl, and R.-J. Fischer, *A proteomic and transcriptional view of acidogenic and solventogenic steady-state cells of Clostridium acetobutylicum in a chemostat culture.* Applied microbiology and biotechnology, 2010. **87**(6): p. 2209-2226.
157. Heap, J.T., O.J. Pennington, S.T. Cartman, G.P. Carter, and N.P. Minton, *The ClosTron: A universal gene knock-out system for the genus Clostridium.* Journal of Microbiological Methods, 2007. **70**(3): p. 452-464.
158. Qureshi, N. and H. Blaschek, *Recent advances in ABE fermentation: hyper-butanol producing Clostridium beijerinckii BA101.* Journal of Industrial Microbiology and Biotechnology, 2001. **27**(5): p. 287-291.
159. Inui, M., M. Suda, S. Kimura, K. Yasuda, H. Suzuki, H. Toda, S. Yamamoto, S. Okino, N. Suzuki, and H. Yukawa, *Expression of Clostridium acetobutylicum butanol synthetic genes in Escherichia coli.* Applied microbiology and biotechnology, 2008. **77**(6): p. 1305-1316.
160. Atsumi, S., A.F. Cann, M.R. Connor, C.R. Shen, K.M. Smith, M.P. Brynildsen, K.J. Chou, T. Hanai, and J.C. Liao, *Metabolic engineering of Escherichia coli for 1-butanol production.* Metabolic engineering, 2008. **10**(6): p. 305-311.
161. Steen, E.J., R. Chan, N. Prasad, S. Myers, C.J. Petzold, A. Redding, M. Ouellet, and J.D. Keasling, *Metabolic engineering of Saccharomyces cerevisiae for the production of n-butanol.* Microb Cell Fact, 2008. **7**(1): p. 36.
162. Nielsen, D.R., E. Leonard, S.-H. Yoon, H.-C. Tseng, C. Yuan, and K.L.J. Prather, *Engineering alternative butanol production platforms in heterologous bacteria.* Metabolic engineering, 2009. **11**(4): p. 262-273.
163. Berezina, O.V., N.V. Zakharova, A. Brandt, S.V. Yarotsky, W.H. Schwarz, and V.V. Zverlov, *Reconstructing the clostridial n-butanol metabolic pathway in Lactobacillus brevis.* Applied microbiology and biotechnology, 2010. **87**(2): p. 635-646.
164. Köpke, M., C. Held, S. Hujer, H. Liesegang, A. Wiezer, A. Wollherr, A. Ehrenreich, W. Liebl, G. Gottschalk, and P. Dürre, *Clostridium ljungdahlii represents a microbial production*

platform based on syngas. Proceedings of the National Academy of Sciences, 2010. **107**(29): p. 13087-13092.
165. Prosser, J.I., B.J. Bohannan, T.P. Curtis, R.J. Ellis, M.K. Firestone, R.P. Freckleton, J.L. Green, L.E. Green, K. Killham, and J.J. Lennon, *The role of ecological theory in microbial ecology.* Nature Reviews Microbiology, 2007. **5**(5): p. 384-392.
166. Bader, J., E. Mast-Gerlach, M. Popović, R. Bajpai, and U. Stahl, *Relevance of microbial coculture fermentations in biotechnology.* Journal of applied microbiology, 2010. **109**(2): p. 371-387.
167. Shong, J., M.R. Jimenez Diaz, and C.H. Collins, *Towards synthetic microbial consortia for bioprocessing.* Current Opinion in Biotechnology, 2012. **23**(5): p. 798-802.
168. Brenner, K., L. You, and F.H. Arnold, *Engineering microbial consortia: a new frontier in synthetic biology.* TRENDS in Biotechnology, 2008. **26**(9): p. 483-489.
169. Zuroff, T.R. and W.R. Curtis, *Developing symbiotic consortia for lignocellulosic biofuel production.* Applied microbiology and biotechnology, 2012. **93**(4): p. 1423-1435.
170. Zuroff, T.R., S.B. Xiques, and W.R. Curtis, *Consortia-mediated bioprocessing of cellulose to ethanol with a symbiotic Clostridium phytofermentans/yeast co-culture.* Biotechnology for biofuels, 2013. **6**(1): p. 59.
171. Yu, E., L. Deschatelets, and J. Saddler, *The bioconversion of wood hydrolyzates to butanol and butanediol.* Biotechnology Letters, 1984. **6**(5): p. 327-332.
172. Fond, O., J.M. Engasser, G. Matta-El-Amouri, and H. Petitdemange, *The acetone butanol fermentation on glucose and xylose. II. Regulation and kinetics in fed-batch cultures.* Biotechnology and bioengineering, 1986. **28**(2): p. 167-175.
173. Salimi, F., K. Zhuang, and R. Mahadevan, *Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing.* Biotechnology journal, 2010. **5**(7): p. 726-738.
174. Jiao, Y., A. Navid, B.J. Stewart, J.B. McKinlay, M.P. Thelen, and J. Pett-Ridge, *Syntrophic metabolism of a co-culture containing Clostridium cellulolyticum and Rhodopseudomonas palustris for hydrogen production.* international journal of hydrogen energy, 2012. **37**(16): p. 11719-11726.
175. Wang, A., N. Ren, Y. Shi, and D.-J. Lee, *Bioaugmented hydrogen production from microcrystalline cellulose using co-culture—Clostridium acetobutylicum X9 and Ethanoigenens harbinense B49.* international journal of hydrogen energy, 2008. **33**(2): p. 912-917.
176. Hanly, T.J. and M.A. Henson, *Dynamic metabolic modeling of a microaerobic yeast co-culture: predicting and optimizing ethanol production from glucose/xylose mixtures.* Biotechnology for biofuels, 2013. **6**(1): p. 44.
177. Wilkins, M.R., J.-C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser, and K.L. Williams, *Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it.* Biotechnology and genetic engineering reviews, 1996. **13**(1): p. 19-50.
178. Abdallah, C., E. Dumas-Gaudot, J. Renaut, and K. Sergeant, *Gel-based and gel-free quantitative proteomics approaches at a glance.* International journal of plant genomics, 2012. **2012**.
179. Profrock, D. and A. Prange, *Compensation of gradient related effects when using capillary liquid chromatography and inductively coupled plasma mass spectrometry for the absolute quantification of phosphorylated peptides.* Journal of Chromatography A, 2009. **1216**(39): p. 6706-6715.
180. Dycka, F., P. Bobal, K. Mazanec, and J. Bobalova, *Rapid and efficient protein enzymatic digestion: an experimental comparison.* Electrophoresis, 2012. **33**(2): p. 288-295.

181. Wolters, D.A., M.P. Washburn, and J.R. Yates, *An automated multidimensional protein identification technology for shotgun proteomics.* Analytical chemistry, 2001. **73**(23): p. 5683-5690.
182. Washburn, M.P., D. Wolters, and J.R. Yates, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nature biotechnology, 2001. **19**(3): p. 242-247.
183. Peng, J., J.E. Elias, C.C. Thoreen, L.J. Licklider, and S.P. Gygi, *Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome.* Journal of proteome research, 2003. **2**(1): p. 43-50.
184. Canas, B., D. López-Ferrer, A. Ramos-Fernández, E. Camafeita, and E. Calvo, *Mass spectrometry technologies for proteomics.* Briefings in functional genomics & proteomics, 2006. **4**(4): p. 295-320.
185. Lee, B., D. Lopez-Ferrer, B.C. Kim, H.B. Na, Y.I. Park, K.K. Weitz, M.G. Warner, T. Hyeon, S.-W. Lee, and R.D. Smith, *Rapid and efficient protein digestion using trypsin-coated magnetic nanoparticles under pressure cycles.* Proteomics, 2011. **11**(2): p. 309-318.
186. Zhang, X., A. Fang, C.P. Riley, M. Wang, F.E. Regnier, and C. Buck, *Multi-dimensional liquid chromatography in proteomics-A review.* Analytica chimica acta, 2010. **664**(2): p. 101-113.
187. Shevchenko, A., M. Wilm, O. Vorm, and M. Mann, *Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels.* Analytical chemistry, 1996. **68**(5): p. 850-858.
188. Shevchenko, A., J.H. Henrik Tomas, sbreve, J.V. Olsen, and M. Mann, *In-gel digestion for mass spectrometric characterization of proteins and proteomes.* Nature protocols, 2007. **1**(6): p. 2856-2860.
189. Hustoft, H.K., H. Malerod, S.R. Wilson, L. Reubsaet, E. Lundanes, and T. Greibrokk, *A Critical Review of Trypsin Digestion for LC-MS Based Proteomics.* 2012.
190. Yates III, J.R., *Mass spectral analysis in proteomics.* Annu. Rev. Biophys. Biomol. Struct., 2004. **33**: p. 297-316.
191. Chiu, C.W., C.L. Chang, and S.F. Chen, *Evaluation of peptide fractionation strategies used in proteome analysis.* Journal of separation science, 2012. **35**(23): p. 3293-3301.
192. Gygi, S.P., B. Rist, T.J. Griffin, J. Eng, and R. Aebersold, *Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags.* Journal of proteome research, 2002. **1**(1): p. 47-54.
193. Zhou, H., J. Dai, Q.H. Sheng, R.X. Li, C.H. Shieh, A. Guttman, and R. Zeng, *A fully automated 2-D LC-MS method utilizing online continuous pH and RP gradients for global proteome analysis.* Electrophoresis, 2007. **28**(23): p. 4311-4319.
194. Delmotte, N., M. Lasaosa, A. Tholey, E. Heinzle, and C.G. Huber, *Two-dimensional reversed-phase× ion-pair reversed-phase HPLC: an alternative approach to high-resolution peptide separation for shotgun proteome analysis.* Journal of proteome research, 2007. **6**(11): p. 4363-4373.
195. Gilar, M., P. Olivova, A.B. Chakraborty, A. Jaworski, S.J. Geromanos, and J.C. Gebler, *Comparison of 1-D and 2-D LC MS/MS methods for proteomic analysis of human serum.* Electrophoresis, 2009. **30**(7): p. 1157-1167.
196. Linden, J.C. and C.L. Lawhead, *Liquid chromatography of saccharides.* Journal of Chromatography A, 1975. **105**(1): p. 125-133.
197. Boersema, P.J., S. Mohammed, and A.J. Heck, *Hydrophilic interaction liquid chromatography (HILIC) in proteomics.* Analytical and bioanalytical chemistry, 2008. **391**(1): p. 151-159.
198. Hemstrom, P. and K. Irgum, *Hydrophilic interaction chromatography.* Journal of separation science, 2006. **29**(12): p. 1784-1821.

199. Fenn, J.B., M. Mann, C.K. Meng, S.F. Wong, and C.M. Whitehouse, *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.
200. Karas, M., D. Bachmann, U.e. Bahr, and F. Hillenkamp, *Matrix-assisted ultraviolet laser desorption of non-volatile compounds.* International journal of mass spectrometry and ion processes, 1987. **78**: p. 53-68.
201. Knochenmuss, R., *Ion formation mechanisms in UV-MALDI.* Analyst, 2006. **131**(9): p. 966-986.
202. Lazar, I.M. and J.L. Kabulski, *Microfluidic LC device with orthogonal sample extraction for on-chip MALDI-MS detection.* Lab Chip, 2013. **13**(11): p. 2055-2065.
203. Wysocki, V.H., K.A. Resing, Q. Zhang, and G. Cheng, *Mass spectrometry of peptides and proteins.* Methods, 2005. **35**(3): p. 211-222.
204. Lane, C., *Mass spectrometry-based proteomics in the life sciences.* Cellular and Molecular Life Sciences CMLS, 2005. **62**(7-8): p. 848-869.
205. Chernushevich, I.V., A.V. Loboda, and B.A. Thomson, *An introduction to quadrupole–time-of-flight mass spectrometry.* Journal of Mass Spectrometry, 2001. **36**(8): p. 849-865.
206. Rubakhin, S.S., J.C. Jurchen, E.B. Monroe, and J.V. Sweedler, *Imaging mass spectrometry: fundamentals and applications to drug discovery.* Drug discovery today, 2005. **10**(12): p. 823-837.
207. Ong, S.-E. and M. Mann, *Mass spectrometry–based proteomics turns quantitative.* Nature chemical biology, 2005. **1**(5): p. 252-262.
208. Bantscheff, M., M. Schirle, G. Sweetman, J. Rick, and B. Kuster, *Quantitative mass spectrometry in proteomics: a critical review.* Analytical and bioanalytical chemistry, 2007. **389**(4): p. 1017-1031.
209. Li, J., H. Steen, and S.P. Gygi, *Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents the yeast salinity stress response.* Molecular & Cellular Proteomics, 2003. **2**(11): p. 1198-1204.
210. Hansen, K.C., G. Schmitt-Ulms, R.J. Chalkley, J. Hirsch, M.A. Baldwin, and A. Burlingame, *Mass spectrometric analysis of protein mixtures at low levels using cleavable 13C-isotope-coded affinity tag and multidimensional chromatography.* Molecular & Cellular Proteomics, 2003. **2**(5): p. 299-314.
211. Thompson, A., J. Schäfer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, and C. Hamon, *Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS.* Analytical chemistry, 2003. **75**(8): p. 1895-1904.
212. Ow, S.Y., M. Salim, J. Noirel, C. Evans, I. Rehman, and P.C. Wright, *iTRAQ underestimation in simple and complex mixtures:"the good, the bad and the ugly".* Journal of proteome research, 2009. **8**(11): p. 5347-5355.
213. Ow, S.Y., T. Cardona, A. Taton, A. Magnuson, P. Lindblad, K. Stensjö, and P.C. Wright, *Quantitative shotgun proteomics of enriched heterocysts from Nostoc sp. PCC 7120 using 8-plex isobaric peptide tags.* Journal of proteome research, 2008. **7**(4): p. 1615-1628.
214. Callister, S.J., R.C. Barry, J.N. Adkins, E.T. Johnson, W.-j. Qian, B.-J.M. Webb-Robertson, R.D. Smith, and M.S. Lipton, *Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics.* Journal of proteome research, 2006. **5**(2): p. 277-286.
215. Mueller, L., T. Mairinger, G. Hermann, G. Koellensperger, and S. Hann, *Characterization of metal-tagged antibodies used in ICP-MS-based immunoassays.* Analytical and bioanalytical chemistry, 2013: p. 1-7.
216. Ong, S.-E., B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, and M. Mann, *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and*

accurate approach to expression proteomics. Molecular & Cellular Proteomics, 2002. **1**(5): p. 376-386.
217. Ross, P.L., Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, and S. Daniels, *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.* Molecular & Cellular Proteomics, 2004. **3**(12): p. 1154-1169.
218. Schwacke, R., U.-I. Flügge, and R. Kunze, *Plant membrane proteome databases.* Plant Physiology and Biochemistry, 2004. **42**(12): p. 1023-1034.
219. Hopkins, A.L. and C.R. Groom, *The druggable genome.* Nature reviews Drug discovery, 2002. **1**(9): p. 727-730.
220. Iwasaki, M., T. Masuda, M. Tomita, and Y. Ishihama, *Chemical cleavage-assisted tryptic digestion for membrane proteome analysis.* Journal of proteome research, 2009. **8**(6): p. 3169-3175.
221. Masuda, T., M. Tomita, and Y. Ishihama, *Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis.* Journal of proteome research, 2008. **7**(2): p. 731-740.
222. Wilchek, M. and E.A. Bayer, *[2] Introduction to avidin-biotin technology.* Methods in enzymology, 1990. **184**: p. 5-13.
223. Elia, G., *Biotinylation reagents for the study of cell surface proteins.* Proteomics, 2008. **8**(19): p. 4012-4024.
224. Abath, F., A. Almeida, and L. Ferreira, *Identification of surface-exposed Yersinia pestis proteins by radio-iodination and biotinylation.* Journal of medical microbiology, 1992. **37**(6): p. 420-424.
225. Bradburne, J., P. Godfrey, J.H. Choi, and J. Mathis, *In vivo labeling of Escherichia coli cell envelope proteins with N-hydroxysuccinimide esters of biotin.* Applied and environmental microbiology, 1993. **59**(3): p. 663-668.
226. Sabarth, N., S. Lamer, U. Zimny-Arndt, P.R. Jungblut, T.F. Meyer, and D. Bumann, *Identification of surface proteins of Helicobacter pylori by selective biotinylation, affinity purification, and two-dimensional gel electrophoresis.* Journal of Biological Chemistry, 2002. **277**(31): p. 27896-27902.
227. Smither, S.J., J. Hill, B.L. van Baar, A.G. Hulst, A.L. de Jong, and R.W. Titball, *Identification of outer membrane proteins of Yersinia pestis through biotinylation.* Journal of Microbiological Methods, 2007. **68**(1): p. 26-31.
228. Ge, Y. and Y. Rikihisa, *Surface-exposed proteins of Ehrlichia chaffeensis.* Infection and immunity, 2007. **75**(8): p. 3833-3841.
229. Swiatecka-Urban, A., S. Moreau-Marquis, D.P. MacEachran, J.P. Connolly, C.R. Stanton, J.R. Su, R. Barnaby, G.A. O'toole, and B.A. Stanton, *Pseudomonas aeruginosa inhibits endocytic recycling of CFTR in polarized human airway epithelial cells.* American Journal of Physiology-Cell Physiology, 2006. **290**(3): p. C862-C872.
230. Li, Y. and R. Sousa, *Expression and purification of E. coli BirA biotin ligase for in vitro biotinylation.* Protein expression and purification, 2012. **82**(1): p. 162-167.
231. Fliszar, K., J.M. Wiggins, C.M. Pignoli, G.P. Martin, and Z. Li, *Analysis of organic volatile impurities in pharmaceutical excipients by static headspace capillary gas chromatography.* Journal of Chromatography A, 2004. **1027**(1): p. 83-91.
232. Lewis, R.J., R.D. Johnson, and D.V. Canfield, *An Accurate Method for the Determination of Carbon Monoxide in Postmortem Blood Using GC/TCD*, 2002, DTIC Document.
233. Jonsson, P., J. Gullberg, A. Nordström, M. Kusano, M. Kowalczyk, M. Sjöström, and T. Moritz, *A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS.* Analytical chemistry, 2004. **76**(6): p. 1738-1745.

234. Pham, T.K., P.K. Chong, C.S. Gan, and P.C. Wright, *Proteomic analysis of Saccharomyces cerevisiae under high gravity fermentation conditions.* Journal of proteome research, 2006. **5**(12): p. 3411-3419.
235. Lee, D.Y. and O. Fiehn, *High quality metabolomic data for Chlamydomonas reinhardtii.* Plant methods, 2008. **4**(1): p. 7.
236. Gedela, S. and N.R. Medicherla, *Chromatographic techniques for the separation of peptides: Application to proteomics.* Chromatographia, 2007. **65**(9-10): p. 511-518.
237. Zhang, X., A. Fang, C.P. Riley, M. Wang, F.E. Regnier, and C. Buck, *Multi-dimensional liquid chromatography in proteomics—A review.* Analytica chimica acta, 2010. **664**(2): p. 101-113.
238. Angel, T.E., U.K. Aryal, S.M. Hengel, E.S. Baker, R.T. Kelly, E.W. Robinson, and R.D. Smith, *Mass spectrometry-based proteomics: existing capabilities and future directions.* Chemical Society Reviews, 2012. **41**(10): p. 3912-3928.
239. Beck, M., A. Schmidt, J. Malmstroem, M. Claassen, A. Ori, A. Szymborska, F. Herzog, O. Rinner, J. Ellenberg, and R. Aebersold, *The quantitative proteome of a human cell line.* Molecular systems biology, 2011. **7**(1).
240. Picotti, P., B. Bodenmiller, L.N. Mueller, B. Domon, and R. Aebersold, *Full Dynamic Range Proteome Analysis of S. cerevisiae by Targeted Proteomics.* Cell, 2009. **138**(4): p. 795-806.
241. Ghaemmaghami, S., W.-K. Huh, K. Bower, R.W. Howson, A. Belle, N. Dephoure, E.K. O'Shea, and J.S. Weissman, *Global analysis of protein expression in yeast.* Nature, 2003. **425**(6959): p. 737-741.
242. de Godoy, L.M., J.V. Olsen, J. Cox, M.L. Nielsen, N.C. Hubner, F. Fröhlich, T.C. Walther, and M. Mann, *Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast.* Nature, 2008. **455**(7217): p. 1251-1254.
243. O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins.* Journal of Biological Chemistry, 1975. **250**(10): p. 4007-4021.
244. Xue, H., B. Lu, and M. Lai, *The cancer secretome: a reservoir of biomarkers.* J Transl Med, 2008. **6**(1): p. 6-52.
245. Lopez-Ferrer, D., B. Canas, J. Vazquez, C. Lodeiro, R. Rial-Otero, I. Moura, and J. Capelo, *Sample treatment for protein identification by mass spectrometry-based techniques.* TrAC Trends in Analytical Chemistry, 2006. **25**(10): p. 996-1005.
246. Gygi, S.P., G.L. Corthals, Y. Zhang, Y. Rochon, and R. Aebersold, *Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology.* Proceedings of the National Academy of Sciences, 2000. **97**(17): p. 9390-9395.
247. Reinders, J., R.P. Zahedi, N. Pfanner, C. Meisinger, and A. Sickmann, *Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics.* Journal of proteome research, 2006. **5**(7): p. 1543-1554.
248. Lee, B., D. Lopez-Ferrer, B.C. Kim, H.B. Na, Y.I. Park, K.K. Weitz, M.G. Warner, T. Hyeon, S.-W. Lee, and R.D. Smith, *Rapid and efficient protein digestion using trypsin-coated magnetic nanoparticles under pressure cycles.* Proteomics, 2011. **11**(2): p. 309-318.
249. Ow, S.Y., M. Salim, J. Noirel, C. Evans, and P. Wright, *Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation.* Proteomics, 2011. **11**(11): p. 2341-2346.
250. Phillips, H.L., J.C. Williamson, K.A. van Elburg, A.P. Snijders, P.C. Wright, and M.J. Dickman, *Shotgun proteome analysis utilising mixed mode (reversed phase-anion exchange chromatography) in conjunction with reversed phase liquid chromatography mass spectrometry analysis.* Proteomics, 2010. **10**(16): p. 2950-2960.
251. Gaspari, M., *Peptide profiling by capillary separation techniques coupled to mass spectrometry*2006: Department of Analytical Biosciences, Faculty of Mathematics and Natural Sciences, Leiden University.

252. Sun, D., N. Wang, and L. Li, *Integrated SDS removal and peptide separation by strong-cation exchange liquid chromatography for SDS-assisted shotgun proteome analysis.* Journal of proteome research, 2012. **11**(2): p. 818-828.
253. López-Contreras, A.M., A.A. Martens, N. Szijarto, H. Mooibroek, P.A. Claassen, J. van der Oost, and W.M. de Vos, *Production by Clostridium acetobutylicum ATCC 824 of CelG, a cellulosomal glycoside hydrolase belonging to family 9.* Applied and environmental microbiology, 2003. **69**(2): p. 869-877.
254. Sambrook, J. and D.W. Russell, *Molecular cloning: a laboratory manual (3-volume set).* 2001.
255. Karunakaran, E. and C.A. Biggs, *Mechanisms of Bacillus cereus biofilm formation: an investigation of the physicochemical characteristics of cell surfaces and extracellular proteins.* Applied microbiology and biotechnology, 2011. **89**(4): p. 1161-1175.
256. Longworth, J., J. Noirel, J. Pandhal, P.C. Wright, and S. Vaidyanathan, *HILIC-and SCX-based quantitative proteomics of Chlamydomonas reinhardtii during nitrogen starvation induced lipid and carbohydrate accumulation.* Journal of proteome research, 2012. **11**(12): p. 5959-5971.
257. Zarei, M., A. Sprenger, F. Metzger, C. Gretzmeier, and J. Dengjel, *Comparison of ERLIC–TiO2, HILIC–TiO2, and SCX–TiO2 for global phosphoproteomics approaches.* Journal of proteome research, 2011. **10**(8): p. 3474-3483.
258. Zhu, B.-Y., C.T. Mant, and R.S. Hodges, *Hydrophilic-interaction chromatography of peptides on hydrophilic and strong cation-exchange columns.* Journal of Chromatography A, 1991. **548**: p. 13-24.
259. Betancourt, L.H., P.-J. De Bock, A. Staes, E. Timmerman, Y. Perez-Riverol, A. Sanchez, V. Besada, L.J. Gonzalez, J. Vandekerckhove, and K. Gevaert, *SCX charge state selective separation of tryptic peptides combined with 2D-RP-HPLC allows for detailed proteome mapping.* Journal of proteomics, 2013. **91**: p. 164-171.
260. Zhou, H., F. Wang, Y. Wang, Z. Ning, W. Hou, T.G. Wright, M. Sundaram, S. Zhong, Z. Yao, and D. Figeys, *Improved recovery and identification of membrane proteins from rat hepatic cells using a centrifugal proteomic reactor.* Molecular & Cellular Proteomics, 2011. **10**(10).
261. Brown, K.J., R. Marathi, A.A. Fiorillo, E.F. Ciccimaro, S. Sharma, D.S. Rowlands, S. Rayavarapu, K. Nagaraju, E.P. Hoffman, and Y. Hathout, *Accurate Quantitation of Dystrophin Protein in Human Skeletal Muscle Using Mass Spectrometry.* Journal of bioanalysis & biomedicine, 2012.
262. Gundry, R.L., M.Y. White, C.I. Murray, L.A. Kane, Q. Fu, B.A. Stanley, and J.E. Van Eyk, *Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow.* Current Protocols in Molecular Biology, 2009: p. 10.25. 1-10.25. 23.
263. Boersema, P.J., N. Divecha, A.J. Heck, and S. Mohammed, *Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy.* Journal of proteome research, 2007. **6**(3): p. 937-946.
264. Gilar, M., P. Olivova, A.E. Daly, and J.C. Gebler, *Orthogonality of separation in two-dimensional liquid chromatography.* Analytical chemistry, 2005. **77**(19): p. 6426-6434.
265. Wu, S. and Y. Zhu, *ProPAS: standalone software to analyze protein properties.* Bioinformation, 2012. **8**(3): p. 167.
266. Kuzyk, M.A., D. Smith, J. Yang, T.J. Cross, A.M. Jackson, D.B. Hardie, N.L. Anderson, and C.H. Borchers, *Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma.* Molecular & Cellular Proteomics, 2009. **8**(8): p. 1860-1877.

267. Le Bihan, T., H.S. Duewel, and D. Figeys, *On-line strong cation exchange μ-HPLC-ESI-MS/MS for protein identification and process optimization.* Journal of the American Society for Mass Spectrometry, 2003. **14**(7): p. 719-727.
268. McNulty, D.E. and R.S. Annan, *Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection.* Molecular & Cellular Proteomics, 2008. **7**(5): p. 971-980.
269. Shah, M., S. Song, Y. Lee, and R. Torget, *Effect of pretreatment on simultaneous saccharification and fermentation of hardwood into acetone/butanol.* Applied biochemistry and biotechnology, 1991. **28**(1): p. 99-109.
270. Lapez-Contreras, A.M., A.A. Martens, N. Szijarto, H. Mooibroek, P.A.M. Claassen, J. Van Der Oost, and W.M. De Vos, *Production by Clostridium acetobutylicum ATCC 824 of CelG, a cellulosomal glycoside hydrolase belonging to family 9.* Applied and environmental microbiology, 2003. **69**(2): p. 869-877.
271. Jesse, T., T. Ezeji, N. Qureshi, and H. Blaschek, *Production of butanol from starch-based waste packing peanuts and agricultural waste.* Journal of Industrial Microbiology and Biotechnology, 2002. **29**(3): p. 117-123.
272. Qureshi, N., X.L. Li, S. Hughes, B.C. Saha, and M.A. Cotta, *Butanol production from corn fiber xylan using Clostridium acetobutylicum.* Biotechnology progress, 2008. **22**(3): p. 673-680.
273. Tiam mun, L., A. Ishizaki, S. Yoshino, and K. Furukawa, *Production of acetone, butanol and ethanol from palm oil waste by Clostridium saccharoperbutylacetonicum N1-4.* Biotechnology Letters, 1995. **17**(6): p. 649-654.
274. Kobayashi, G., K. Eto, Y. Tashiro, K. Okubo, K. Sonomoto, and A. Ishizaki, *Utilization of excess sludge by acetone–butanol–ethanol fermentation employing Clostridium saccharoperbutylacetonicum N1-4 (ATCC 13564).* Journal of bioscience and bioengineering, 2005. **99**(5): p. 517-519.
275. Ennis, B. and I. Maddox, *The effect of pH and lactose concentration on solvent production from whey permeate using Clostridium acetobutylicum.* Biotechnology and bioengineering, 1987. **29**(3): p. 329-334.
276. Madihah, M., A. Ariff, K. Sahaid, A. Suraini, and M. Karim, *Direct fermentation of gelatinized sago starch to acetone–butanol–ethanol by Clostridium acetobutylicum.* World Journal of Microbiology and Biotechnology, 2001. **17**(6): p. 567-576.
277. Ezeji, T., N. Qureshi, and H.P. Blaschek, *Butanol production from agricultural residues: impact of degradation products on Clostridium beijerinckii growth and butanol fermentation.* Biotechnology and bioengineering, 2007. **97**(6): p. 1460-1469.
278. Cappelletti, B.M., V. Reginatto, E.R. Amante, and R.V. AntÃ´nio, *Fermentative production of hydrogen from cassava processing wastewater by Clostridium acetobutylicum.* Renewable Energy, 2011. **36**(12): p. 3367-3372.
279. Sun, Z. and S. Liu, *Production of n-butanol from concentrated sugar maple hemicellulosic hydrolysate by Clostridia acetobutylicum ATCC824.* biomass and bioenergy, 2012. **39**: p. 39-47.
280. Zverlov, V., O. Berezina, G. Velikodvorskaya, and W. Schwarz, *Bacterial acetone and butanol production by industrial fermentation in the Soviet Union: use of hydrolyzed agricultural waste for biorefinery.* Applied microbiology and biotechnology, 2006. **71**(5): p. 587-597.
281. Lee, S. and R.J. Mitchell, *Biological activities of lignin hydrolysate-related compounds.* Biochemistry and Molecurar Biology Reports, 2012. **45**(5): p. 265-274.
282. Zhang, Y. and T.C. Ezeji, *Transcriptional analysis of Clostridium beijerinckii NCIMB 8052 to elucidate role of furfural stress during acetone butanol ethanol fermentation.* Biotechnology for biofuels, 2013. **6**(1): p. 1-17.

283. Zhang, Y., B. Han, and T.C. Ezeji, *Biotransformation of furfural and 5-hydroxymethyl furfural (HMF) by Clostridium acetobutylicum ATCC 824 during butanol fermentation.* New Biotechnology, 2012. **29**(3): p. 345-351.
284. Cadenas, A. and S. Cabezudo, *Biofuels as sustainable technologies: perspectives for less developed countries.* Technological Forecasting and Social Change, 1998. **58**(1): p. 83-103.
285. Fuchs, G., M. Boll, and J. Heider, *Microbial degradation of aromatic compounds from one strategy to four.* Nature Reviews Microbiology, 2011. **9**(11): p. 803-816.
286. Nakamura, Y., H. Miyafuji, H. Kawamoto, and S. Saka, *Acetic acid fermentability with Clostridium thermoaceticum and Clostridium thermocellum of standard compounds found in beech wood as produced in hot-compressed water.* Journal of wood science, 2011. **57**(4): p. 331-337.
287. Wang, S., H. Huang, J. Moll, and R.K. Thauer, *NADP+ reduction with reduced ferredoxin and NADP+ reduction with NADH are coupled via an electron-bifurcating enzyme complex in Clostridium kluyveri.* Journal of bacteriology, 2010. **192**(19): p. 5115-5123.
288. Li, F., J. Hinderberger, H. Seedorf, J. Zhang, W. Buckel, and R.K. Thauer, *Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from Clostridium kluyveri.* Journal of bacteriology, 2008. **190**(3): p. 843-850.
289. Huang, J. and J. Wiegel, *Cloning, characterization, and expression of a novel gene encoding a reversible 4-hydroxybenzoate decarboxylase from Clostridium hydroxybenzoicum.* Journal of bacteriology, 1999. **181**(16): p. 5119-5122.
290. Adrian, N. and E. Campbell, *TNT and RDX Degradation by Cell-Free Extracts of Clostridium Acetobutylicum*, 1999, DTIC Document.
291. Kutty, R. and G.N. Bennett, *Biochemical characterization of trinitrotoluene transforming oxygen-insensitive nitroreductases from Clostridium acetobutylicum ATCC 824.* Archives of microbiology, 2005. **184**(3): p. 158-167.
292. Hughes, J.B., C.Y. Wang, R. Bhadra, A. Richardson, G.N. Bennett, and F.B. Rudolph, *Reduction of 2, 4, 6-trinitrotoluene by Clostridium acetobutylicum through hydroxylamino-nitrotoluene intermediates.* Environmental toxicology and chemistry, 1998. **17**(3): p. 343-348.
293. Watrous, M.M., S. Clark, R. Kutty, S. Huang, F.B. Rudolph, J.B. Hughes, and G.N. Bennett, *2, 4, 6-Trinitrotoluene reduction by an Fe-only hydrogenase in Clostridium acetobutylicum.* Applied and environmental microbiology, 2003. **69**(3): p. 1542-1547.
294. Khan, T., R. Bhadra, and J. Hughes, *Anaerobic transformation of 2, 4, 6-TNT and related nitroaromatic compounds by Clostridium acetobutylicum.* Journal of Industrial Microbiology and Biotechnology, 1997. **18**(2-3): p. 198-203.
295. Cai, X., M. Servinsky, J. Kiel, C. Sund, and G.N. Bennett, *Analysis of redox responses during TNT transformation by Clostridium acetobutylicum ATCC 824 and mutants exhibiting altered metabolism.* Applied Microbiology and Biotechnology, 2012: p. 1-13.
296. Hedge, J.E. and B.T. Hofreiter, *Methods of estimating starch and carbohydrate.* Carbohydrate chemistry, 1962. **17**: p. 163-201.
297. Pham, T.K., S. Roy, J. Noirel, I. Douglas, P.C. Wright, and G.P. Stafford, *A quantitative proteomic analysis of biofilm adaptation by the periodontal pathogen Tannerella forsythia.* Proteomics, 2010. **10**(17): p. 3130-3141.
298. Mutschlechner, O., H. Swoboda, and J.R. Gapes, *Continuous two-stage ABE-fermentation using Clostridium beijerinckii NRRL B 592 operating with a growth rate in the first stage vessel close to its maximal value.* Journal of Molecular Microbiology and Biotechnology, 2000. **2**(1): p. 101-105.
299. Huang, W.-C., D.E. Ramey, and S.-T. Yang. *Continuous production of butanol by Clostridium acetobutylicum immobilized in a fibrous bed bioreactor*. in *Proceedings of*

the Twenty-Fifth Symposium on Biotechnology for Fuels and Chemicals Held May 4–7, 2003, in Breckenridge, CO. 2004. Springer.
300. Qureshi, N. and H. Blaschek, *Butanol production from agricultural biomass.* Food Biotechnology, K. Shetty, A. Pometto, G. Paliyath (eds), Taylor & Francis Group plc, Boca Raton, FL, 2005: p. 525-551.
301. Tashiro, Y., K. Takeda, G. Kobayashi, and K. Sonomoto, *High production of acetone–butanol–ethanol with high cell density culture by cell-recycling and bleeding.* Journal of biotechnology, 2005. **120**(2): p. 197-206.
302. Napoli, F., G. Olivieri, A. Marzocchella, and P. Salatino, *Bioenergy II: an assessment of the kinetics of butanol production by Clostridium acetobutylicum.* International Journal of Chemical Reactor Engineering, 2009. **7**(1).
303. Elias, J.E. and S.P. Gygi, *Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry.* Nature methods, 2007. **4**(3): p. 207-214.
304. Buswell, J.A., P. Ander, and K.-E. Eriksson, *Ligninolytic activity and levels of ammonia assimilating enzymes in Sporotrichum pulverulentum.* Archives of Microbiology, 1982. **133**(3): p. 165-171.
305. Ertani, A., O. Francioso, V. Tugnoli, V. Righi, and S. Nardi, *Effect of commercial lignosulfonate-humate on Zea mays L. metabolism.* Journal of agricultural and food chemistry, 2011. **59**(22): p. 11940-11948.
306. Manavalan, A., S.S. Adav, and S.K. Sze, *iTRAQ-based quantitative secretome analysis of Phanerochaete chrysosporium.* Journal of proteomics, 2011. **75**(2): p. 642-654.
307. Santos, P.M., D. Benndorf, and I. Sa-Correia, *Insights into Pseudomonas putida KT2440 response to phenol-induced stress by quantitative proteomics.* Proteomics, 2004. **4**(9): p. 2640-2652.
308. Lee, S.Y., Y.-H. Kim, and J. Min, *Conversion of phenol to glutamate and proline in Corynebacterium glutamicum is regulated by transcriptional regulator ArgR.* Applied microbiology and biotechnology, 2010. **85**(3): p. 713-720.
309. Shen, X.-H., N.-Y. Zhou, and S.-J. Liu, *Degradation and assimilation of aromatic compounds by Corynebacterium glutamicum: another potential for applications for this bacterium?* Applied microbiology and biotechnology, 2012. **95**(1): p. 77-89.
310. Mahler, H.R., *Studies on the fatty acid oxidizing system of animal tissues IV. The prosthetic group of butyryl coenzyme A dehydrogenase.* Journal of Biological Chemistry, 1954. **206**(1): p. 13-26.
311. Boynton, Z.L., G.N. Bennet, and F.B. Rudolph, *Cloning, sequencing, and expression of clustered genes encoding beta-hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and butyryl-CoA dehydrogenase from Clostridium acetobutylicum ATCC 824.* Journal of bacteriology, 1996. **178**(11): p. 3015-3024.
312. Leonardi, R. and P.L. Roach, *Thiamine Biosynthesis in Escherichia coli in vitro reconstitution of the thiazole synthase activity.* Journal of Biological Chemistry, 2004. **279**(17): p. 17054-17062.
313. Gonglez, B. and R. Vicuna, *Benzaldehyde lyase, a novel thiamine PPi-requiring enzyme, from Pseudomonas fluorescens biovar I.* Journal of bacteriology, 1989. **171**(5): p. 2401-2405.
314. Luders, S., C. Fallet, and E. Franco-Lara, *Proteome analysis of the Escherichia coli heat shock response under steady-state conditions.* Proteome Sci, 2009. **7**(1): p. 36-50.
315. Ali, G.M. and S. Komatsu, *Proteomic analysis of rice leaf sheath during drought stress.* Journal of proteome research, 2006. **5**(2): p. 396-403.
316. Zhang, Y., Y. Zhang, Y. Zhu, S. Mao, and Y. Li, *Proteomic analyses to reveal the protective role of glutathione in resistance of Lactococcus lactis to osmotic stress.* Applied and environmental microbiology, 2010. **76**(10): p. 3177-3186.

317. Moreno, J.I., R. Martin, and C. Castresana, *Arabidopsis SHMT1, a serine hydroxymethyltransferase that functions in the photorespiratory pathway influences resistance to biotic and abiotic stress.* The Plant Journal, 2005. **41**(3): p. 451-463.
318. Hanson, A.D., D.A. Gage, and Y. Shachar-Hill, *Plant one-carbon metabolism and its engineering.* Trends in plant science, 2000. **5**(5): p. 206-213.
319. Masai, E., Y. Katayama, and M. Fukuda, *Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds.* Bioscience, biotechnology, and biochemistry, 2007. **71**(1): p. 1-15.
320. Hanson, R.S. and T.E. Hanson, *Methanotrophic bacteria.* Microbiological reviews, 1996. **60**(2): p. 439-471.
321. Heluane, H., M.R. Evans, S.F. Dagher, and J.M. Bruno-BÃ¡rcena, *Meta-analysis and functional validation of nutritional requirements of solventogenic clostridia growing under butanol stress conditions and coutilization of d-glucose and d-xylose.* Applied and environmental microbiology, 2011. **77**(13): p. 4473-4485.
322. Gorsich, S.W., B.S. Dien, N.N. Nichols, P.J. Slininger, Z.L. Liu, and C.D. Skory, *Tolerance to furfural-induced stress is associated with pentose phosphate pathway genes ZWF1, GND1, RPE1, and TKL1 in Saccharomyces cerevisiae.* Applied microbiology and biotechnology, 2006. **71**(3): p. 339-349.
323. He, M.-x., B. Wu, Z.-x. Shui, Q.-c. Hu, W.-g. Wang, F.-r. Tan, X.-y. Tang, Q.-l. Zhu, K. Pan, and Q. Li, *Transcriptome profiling of Zymomonas mobilis under furfural stress.* Applied microbiology and biotechnology, 2012. **95**(1): p. 189-199.
324. Pryke, J.A. and T. ap Rees, *The pentose phosphate pathway as a source of NADPH for lignin synthesis.* Phytochemistry, 1977. **16**(5): p. 557-560.
325. Ralser, M., M.M. Wamelink, A. Kowald, B. Gerisch, G. Heeren, E.A. Struys, E. Klipp, C. Jakobs, M. Breitenbach, and H. Lehrach, *Dynamic rerouting of the carbohydrate flux is key to counteracting oxidative stress.* J Biol, 2007. **6**(10).
326. Kruger, N.J. and A. von Schaewen, *The oxidative pentose phosphate pathway: structure and organisation.* Current opinion in plant biology, 2003. **6**(3): p. 236-246.
327. Fuhrman, L.K., A. Wanken, K.W. Nickerson, and T. Conway, *Rapid accumulation of intracellular 2-keto 3-deoxy-6-phosphogluconate in an Entner Doudoroff aldolase mutant results in bacteriostasis.* FEMS microbiology letters, 1998. **159**(2): p. 261-266.
328. Yoon, S.-S., S.-H. Park, Y.-C. Kim, M. Shin, C.-K. Chong, and J.-D. Choi, *Cloning and characterization of phosphoglucomutase and phosphomannomutase derived from Sphingomonas chungbukensis DJ77.* Journal of biochemistry, 2008. **144**(4): p. 507-512.
329. Dunwell, J.M., *Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins.* Biotechnology and Genetic Engineering Reviews, 1998. **15**: p. 1.
330. Fetzner, S., *Ring-cleaving dioxygenases with a cupin fold.* Applied and environmental microbiology, 2012. **78**(8): p. 2505-2514.
331. Amador-Noguez, D., I.A. Brasg, X.-J. Feng, N. Roquet, and J.D. Rabinowitz, *Metabolome remodeling during the acidogenic-solventogenic transition in Clostridium acetobutylicum.* Applied and environmental microbiology, 2011. **77**(22): p. 7984-7997.
332. DeAngelis, K.M., D. Sharma, R. Varney, B. Simmons, N.G. Isern, L.M. Markilllie, C. Nicora, A.D. Norbeck, R.C. Taylor, and J.T. Aldrich, *Evidence supporting dissimilatory and assimilatory lignin degradation in Enterobacter lignolyticus SCF1.* Frontiers in microbiology, 2013. **4**.
333. Piper, P.W., C. Ortiz-Calderon, C. Holyoak, P. Coote, and M. Cole, *Hsp30, the integral plasma membrane heat shock protein of Saccharmyces cerevisiae, is a stress-inducible regulator of plasma membrane H+-ATPase.* Cell stress & chaperones, 1997. **2**(1): p. 12.

334. Giuliani, S.E., A.M. Frank, D.M. Corgliano, C. Seifert, L. Hauser, and F.R. Collart, *Environment sensing and response mediated by ABC transporters.* BMC genomics, 2011. **12**(Suppl 1): p. S8.
335. Janssen, H., C. Daring, A. Ehrenreich, B. Voigt, M. Hecker, H. Bahl, and R.-J. Fischer, *A proteomic and transcriptional view of acidogenic and solventogenic steady-state cells of Clostridium acetobutylicum in a chemostat culture.* Applied microbiology and biotechnology, 2010. **87**(6): p. 2209-2226.
336. Alsaker, K.V. and E.T. Papoutsakis, *Transcriptional program of early sporulation and stationary-phase events in Clostridium acetobutylicum.* Journal of bacteriology, 2005. **187**(20): p. 7103-7118.
337. Sullivan, L. and G.N. Bennett, *Proteome analysis and comparison of Clostridium acetobutylicum ATCC 824 and Spo0A strain variants.* Journal of Industrial Microbiology and Biotechnology, 2006. **33**(4): p. 298-308.
338. Tomas, C.A., K.V. Alsaker, H.P.J. Bonarius, W.T. Hendriksen, H. Yang, J.A. Beamish, C.J. Paredes, and E.T. Papoutsakis, *DNA array-based transcriptional analysis of asporogenous, nonsolventogenic Clostridium acetobutylicum strains SKO1 and M5.* Journal of bacteriology, 2003. **185**(15): p. 4539-4547.
339. Jean, D., V.r. Briolat, and G. Reysset, *Oxidative stress response in Clostridium perfringens.* Microbiology, 2004. **150**(6): p. 1649-1659.
340. Lim, K., A. Tempczyk, J.F. Parsons, N. Bonander, J. Toedt, Z. Kelman, A. Howard, E. Eisenstein, and O. Herzberg, *Crystal structure of YbaB from Haemophilus influenzae (HI0442), a protein of unknown function coexpressed with the recombinational DNA repair protein RecR.* Proteins: Structure, Function, and Bioinformatics, 2003. **50**(2): p. 375-379.
341. Schultz, J., F. Milpetz, P. Bork, and C.P. Ponting, *SMART, a simple modular architecture research tool: identification of signaling domains.* Proceedings of the National Academy of Sciences, 1998. **95**(11): p. 5857-5864.
342. Alsaker, K.V., T.R. Spitzer, and E.T. Papoutsakis, *Transcriptional analysis of spo0A overexpression in Clostridium acetobutylicum and its effect on the cell's response to butanol stress.* Journal of bacteriology, 2004. **186**(7): p. 1959-1971.
343. Mann, M.S., Z. Dragovic, G. Schirrmacher, and T. Lütke-Eversloh, *Over-expression of stress protein-encoding genes helps Clostridium acetobutylicum to rapidly adapt to butanol stress.* Biotechnology letters, 2012. **34**(9): p. 1643-1649.
344. Ito, K. and Y. Akiyama, *Cellular functions, mechanism of action, and regulation of FtsH protease.* Annu. Rev. Microbiol., 2005. **59**: p. 211-231.
345. Jayasekera, M.M.K., S.K. Foltin, E.R. Olson, and T.P. Holler, *Escherichia coli Requires the Protease Activity of FtsH for Growth.* Archives of Biochemistry and Biophysics, 2000. **380**(1): p. 103-107.
346. Tomoyasu, T., J. Gamer, B. Bukau, M. Kanemori, H. Mori, A.J. Rutman, A.B. Oppenheim, T. Yura, K. Yamanaka, and H. Niki, *Escherichia coli FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32.* The EMBO journal, 1995. **14**(11): p. 2551.
347. Cha, J.-H. and G.C. Stewart, *The divIVA minicell locus of Bacillus subtilis.* Journal of bacteriology, 1997. **179**(5): p. 1671-1683.
348. Garti-Levi, S., R. Hazan, J. Kain, M. Fujita, and S. Ben-Yehuda, *The FtsEX ABC transporter directs cellular differentiation in Bacillus subtilis.* Molecular microbiology, 2008. **69**(4): p. 1018-1028.
349. Mayer, F., *Cytoskeletons in prokaryotes.* Cell biology international, 2003. **27**(5): p. 429-438.
350. Kleinschnitz, E.M., A. Heichlinger, K. Schirner, J. Winkler, A. Latus, I. Maldener, W. Wohlleben, and G.n. Muth, *Proteins encoded by the mre gene cluster in Streptomyces*

coelicolor A3 (2) cooperate in spore wall synthesis. Molecular microbiology, 2011. **79**(5): p. 1367-1379.
351. Baker, M.D., P.M. Wolanin, and J.B. Stock, *Signal transduction in bacterial chemotaxis.* Bioessays, 2006. **28**(1): p. 9-22.
352. Perego, M. and J.A. Hoch, *Protein aspartate phosphatases control the output of two-component signal transduction systems.* Trends in Genetics, 1996. **12**(3): p. 97-101.
353. West, A.H. and A.M. Stock, *Histidine kinases and response regulator proteins in two-component signaling systems.* Trends in biochemical sciences, 2001. **26**(6): p. 369.
354. Skerker, J.M., M.S. Prasol, B.S. Perchuk, E.G. Biondi, and M.T. Laub, *Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis.* PLoS biology, 2005. **3**(10): p. e334.
355. Kawasaki, S., Y. Watamura, M. Ono, T. Watanabe, K. Takeda, and Y. Niimura, *Adaptive responses to oxygen stress in obligatory anaerobes Clostridium acetobutylicum and Clostridium aminovalericum.* Applied and environmental microbiology, 2005. **71**(12): p. 8442-8450.
356. Ballongue, J., J. Amine, E. Masion, H. Petitdemange, and R. Gay, *Induction of acetoacetate decarboxylase in Clostridium acetobutylicum.* FEMS microbiology letters, 1985. **29**(3): p. 273-277.
357. Reid, M.F. and C.A. Fewson, *Molecular characterization of microbial alcohol dehydrogenases.* Critical reviews in microbiology, 1994. **20**(1): p. 13-56.
358. Saxena, I.M. and R. Brown, *Cellulose biosynthesis: current views and evolving concepts.* Annals of Botany, 2005. **96**(1): p. 9.
359. Taherzadeh, M.J. and K. Karimi, *Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review.* International journal of molecular sciences, 2008. **9**(9): p. 1621-1651.
360. Jun, H.S., M. Qi, J. Gong, E.E. Egbosimba, and C.W. Forsberg, *Outer membrane proteins of Fibrobacter succinogenes with potential roles in adhesion to cellulose and in cellulose digestion.* Journal of bacteriology, 2007. **189**(19): p. 6806-6815.
361. Kudo, H., K.J. Cheng, and J.W. Costerton, *Electron microscopic study of the methylcellulose-mediated detachment of cellulolytic rumen bacteria from cellulose fibers.* Canadian journal of microbiology, 1987. **33**(3): p. 267-272.
362. Rijnaarts, H.H.M., W. Norde, J. Lyklema, and A.J.B. Zehnder, *The isoelectric point of bacteria as an indicator for the presence of cell surface polymers that inhibit adhesion.* Colloids and Surfaces B: Biointerfaces, 1995. **4**(4): p. 191-197.
363. Mukherjee, J., E. Karunakaran, and C.A. Biggs, *Using a multi-faceted approach to determine the changes in bacterial cell surface properties influenced by a biofilm lifestyle.* Biofouling, 2012. **28**(1): p. 1-14.
364. Weimer, P.J., Y. Shi, and C.L. Odt, *A segmented gas/liquid delivery system for continuous culture of microorganisms on insoluble substrates and its use for growth of Ruminococcus flavefaciens on cellulose.* Applied microbiology and biotechnology, 1991. **36**(2): p. 178-183.
365. Callaway, E. and S. Martin, *Effects of a Saccharomyces cerevisiae Culture on Ruminal Bacteria that Utilize Lactate and Digest Cellulose.* Journal of Dairy Science, 1997. **80**(9): p. 2035-2044.
366. Olsen, M.A. and S.D. Mathiesen, *The bacterial population adherent to plant particles in the rumen of reindeer fed lichen, timothy hay or silage.* Rangifer, 2011. **18**(2): p. 55-64.
367. Hsu, J.C. and M.H. Penner, *Influence of cellulose structure on its digestibility in the rat.* The Journal of nutrition, 1989. **119**(6): p. 872.
368. Smogyi, M., *Notes on sugar determination.* The Journal of biological chemistry, 1952. **195**(1): p. 19.

369. Updegraff, D.M., *Semimicro determination of cellulose inbiological materials.* Analytical Biochemistry, 1969. **32**(3): p. 420-424.
370. van Merode, A.E.J., H.C. van der Mei, H.J. Busscher, and B.P. Krom, *Influence of culture heterogeneity in cell surface charge on adhesion and biofilm formation by Enterococcus faecalis.* Journal of bacteriology, 2006. **188**(7): p. 2421.
371. Kim, S., B.L. Reuhs, and L.J. Mauer, *Use of Fourier transform infrared spectra of crude bacterial lipopolysaccharides and chemometrics for differentiation of Salmonella enterica serotypes.* Journal of applied microbiology, 2005. **99**(2): p. 411-417.
372. Marcotte, L., G. Kegelaer, C. Sandt, J. Barbeau, and M. Lafleur, *An alternative infrared spectroscopy assay for the quantification of polysaccharides in bacterial samples.* Analytical biochemistry, 2007. **361**(1): p. 7-14.
373. Kamnev, A.A., *FTIR spectroscopic studies of bacterial cellular responses to environmental factors, plant-bacterial interactions and signalling.* Spectroscopy, 2008. **22**(2): p. 83-95.
374. Wilson, W.W., M.M. Wade, S.C. Holman, and F.R. Champlin, *Status of methods for assessing bacterial cell surface charge properties based on zeta potential measurements.* Journal of Microbiological Methods, 2001. **43**(3): p. 153-164.
375. Rosenberg, M., D. Gutnick, and E. Rosenberg, *Adherence of bacteria to hydrocarbons: a simple method for measuring cell-surface hydrophobicity.* FEMS Microbiology Letters, 2006. **9**(1): p. 29-33.
376. Rosenberg, M., *Microbial adhesion to hydrocarbons: twenty five years of doing MATH.* FEMS microbiology letters, 2006. **262**(2): p. 129-134.
377. Tahmourespour, A., R. Kasra Kermanshahi, and N.A.A.R. Salehi Rasoul, *The relationship between cell surface hydrophobicity and antibiotic resistance of streptococcal strains isolated from dental plaque and caries.* Iranian Journal of Basic Medical Sciences, 2008.
378. Naumann, D., *Infrared spectroscopy in microbiology.* Encyclopedia of analytical chemistry, 2000.
379. Bosch, A., D. Serra, C. Prieto, J. Schmitt, D. Naumann, and O. Yantorno, *Characterization of Bordetella pertussis growing as biofilm by chemical analysis and FT-IR spectroscopy.* Applied microbiology and biotechnology, 2006. **71**(5): p. 736-747.
380. Ge, Y. and Y. Rikihisa, *Anaplasma phagocytophilum delays spontaneous human neutrophil apoptosis by modulation of multiple apoptotic pathways.* Cellular microbiology, 2006. **8**(9): p. 1406-1416.
381. Wells, J.E. and J.B. Russell, *Why do many ruminal bacteria die and lyse so quickly?* Journal of dairy science, 1996. **79**(8): p. 1487-1495.
382. Kwaszewska, A.K., A. Brewczynska, and E.M. Szewczyk, *Hydrophobicity and biofilm formation of lipophilic skin corynebacteria.* Polish Journal of Microbiology, 2006. **55**(3): p. 189-193.
383. Slama, R., K. Bekir, H. Miladi, A. Noumi, and A. Bakhrouf, *Adhesive ability and biofilm metabolic activity of Listeria monocytogenes strains before and after cold stress.* African Journal of Biotechnology, 2012. **11**(61): p. 12475-12482.
384. Marshall, K.C., *Interfaces in Microbial Ecology. 1976.* Soil Science, 1977. **123**(5): p. 344.
385. Eboigbodin, K., J. Newton, A. Routh, and C. Biggs, *Bacterial quorum sensing and cell surface electrokinetic properties.* Applied microbiology and biotechnology, 2006. **73**(3): p. 669-675.
386. Kim, C., H. Jung, J.H. Kim, and C.S. Shin, *Effect of monascus pigment derivatives on the electrophoretic mobility of bacteria, and the cell adsorption and antibacterial activities of pigments.* Colloids and Surfaces B: Biointerfaces, 2006. **47**(2): p. 153-159.
387. Hong, Y. and D.G. Brown, *Electrostatic behavior of the charge-regulated bacterial cell surface.* Langmuir, 2008. **24**(9): p. 5003-5009.

388. Sonohara, R., N. Muramatsu, H. Ohshima, and T. Kondo, *Difference in surface properties between Escherichia coli and Staphylococcus aureus as revealed by electrophoretic mobility measurements.* Biophysical chemistry, 1995. **55**(3): p. 273-277.
389. Hayashi, H., H. Seiki, S. Tsuneda, A. Hirata, and H. Sasaki, *Influence of growth phase on bacterial cell electrokinetic characteristics examined by soft particle electrophoresis theory.* Journal of colloid and interface science, 2003. **264**(2): p. 565-568.
390. Nancy, Y.Y., J.R. Wagner, M.R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, S.C. Sahinalp, M. Ester, and L.J. Foster, *PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes.* Bioinformatics, 2010. **26**(13): p. 1608-1615.
391. Cantarel, B.L., P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, and B. Henrissat, *The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics.* Nucleic acids research, 2009. **37**(suppl 1): p. D233-D238.
392. Malburg Jr, L.M., A.H. Iyo, and C.W. Forsberg, *A novel family 9 endoglucanase gene (celD), whose product cleaves substrates mainly to glucose, and its adjacent upstream homolog (celE) from Fibrobacter succinogenes S85.* Applied and environmental microbiology, 1996. **62**(3): p. 898-906.
393. Qi, M., H.S. Jun, and C.W. Forsberg, *Characterization and synergistic interactions of Fibrobacter succinogenes glycoside hydrolases.* Applied and environmental microbiology, 2007. **73**(19): p. 6098-6105.
394. McGavin, M.J., C.W. Forsberg, B. Crosby, A.W. Bell, D. Dignard, and D.Y. Thomas, *Structure of the cel-3 gene from Fibrobacter succinogenes S85 and characteristics of the encoded gene product, endoglucanase 3.* Journal of bacteriology, 1989. **171**(10): p. 5587-5595.
395. Brumm, P., D. Mead, J. Boyum, C. Drinkwater, K. Gowda, D. Stevenson, and P. Weimer, *Functional annotation of Fibrobacter succinogenes S85 carbohydrate active enzymes.* Appl Biochem Biotechnol, 2010. **10**: p. s12010-12010.
396. Malburg, S.R., L.M. Malburg Jr, T. Liu, A.H. Iyo, and C.W. Forsberg, *Catalytic properties of the cellulose-binding endoglucanase F from Fibrobacter succinogenes S85.* Applied and environmental microbiology, 1997. **63**(6): p. 2449-2453.
397. Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, and A. Bairoch, *Protein identification and analysis tools on the ExPASy server.* The proteomics protocols handbook, 2005: p. 571-607.
398. Dyrlov Bendtsen, J., H. Nielsen, G. von Heijne, and S. Brunak, *Improved prediction of signal peptides: SignalP 3.0.* Journal of molecular biology, 2004. **340**(4): p. 783-795.
399. Morris, E.J. and O.J. Cole, *Relationship between cellulolytic activity and adhesion to cellulose in Ruminococcus albus.* Journal of general microbiology, 1987. **133**(4): p. 1023-1032.
400. Morrison, M. and J. Miron, *Adhesion to cellulose by Ruminococcus albus: a combination of cellulosomes and Pil-proteins? 1.* FEMS microbiology letters, 2000. **185**(2): p. 109-115.
401. Rakotoarivonina, H., G. Jubelin, M. Hebraud, B. Gaillard-Martinie, E. Forano, and P. Mosoni, *Adhesion to cellulose of the Gram-positive bacterium Ruminococcus albus involves type IV pili.* Microbiology, 2002. **148**(6): p. 1871-1880.
402. Huang, L. and C.W. Forsberg, *Cellulose digestion and cellulase regulation and distribution in Fibrobacter succinogenes subsp. succinogenes S85.* Applied and environmental microbiology, 1990. **56**(5): p. 1221-1228.
403. McDermid, K.P., C.W. Forsberg, and C.R. MacKenzie, *Purification and properties of an acetylxylan esterase from Fibrobacter succinogenes S85.* Applied and environmental microbiology, 1990. **56**(12): p. 3805-3810.
404. Kam, D.K., H.-S. Jun, J.K. Ha, G.D. Inglis, and C.W. Forsberg, *Characteristics of adjacent family 6 acetylxylan esterases from Fibrobacter succinogenes and the interaction with*

the Xyn10E xylanase in hydrolysis of acetylated xylan. Canadian journal of microbiology, 2005. **51**(10): p. 821-832.
405. Yoshida, S., R.I. Mackie, and I.K.O. Cann, *Biochemical and domain analyses of FSUAxe6B, a modular acetyl xylan esterase, identify a unique carbohydrate binding module in Fibrobacter succinogenes S85.* Journal of bacteriology, 2009. **192**(2): p. 483-493.
406. Doi, R., *Cellulases of mesophilic microorganisms.* Annals of the New York Academy of Sciences, 2008. **1125**(1): p. 267.
407. Di Guilmi, A.M., A.a. Dessen, O. Dideberg, and T. Vernet, *The glycosyltransferase domain of penicillin-binding protein 2a from Streptococcus pneumoniae catalyzes the polymerization of murein glycan chains.* Journal of bacteriology, 2003. **185**(15): p. 4418-4423.
408. Henrissat, B., M. Claeyssens, P. Tomme, L. Lemesle, and J.P. Mornon, *Cellulase families revealed by hydrophobic cluster analysi.* Gene, 1989. **81**(1): p. 83-95.
409. Gold, N.D. and V.J.J. Martin, *Global view of the Clostridium thermocellum cellulosome revealed by quantitative proteomic analysis.* Journal of bacteriology, 2007. **189**(19): p. 6787-6795.
410. Wu, H., K.P. Mintz, M. Ladha, and P.M. Fivesâ€Taylor, *Isolation and characterization of Fap1, a fimbriae-associated adhesin of Streptococcus parasanguis FW213.* Molecular microbiology, 1998. **28**(3): p. 487-500.
411. Branda, S.S., Ã.s. Vik, L. Friedman, and R. Kolter, *Biofilms: the matrix revisited.* Trends in microbiology, 2005. **13**(1): p. 20-26.
412. Pell, A.N. and P. Schofield, *Microbial adhesion and degradation of plant cell walls.* Forage Cell Wall Structures and Digestibility: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Madison, 1993: p. 397-424.
413. Skorstengaard, K., M.S. Jensen, P. Sahl, T.E. Petersen, and S. Magnusson, *Complete primary structure of bovine plasma fibronectin.* European Journal of Biochemistry, 1986. **161**(2): p. 441-453.
414. Dean, D.C., C.L. Bowlus, and S. Bourgeois, *Cloning and analysis of the promotor region of the human fibronectin gene.* Proceedings of the National Academy of Sciences, 1987. **84**(7): p. 1876.
415. Py, B., B.G. Isabelle, J. Haiech, M. Chippaux, and F. Barras, *Cellulase EGZ of Erwinia chrysanthemi: structural organization and importance of His98 and Glu133 residues for catalysis.* Protein engineering, 1991. **4**(3): p. 325-333.
416. Badieyan, S., D.R. Bevan, and C. Zhang, *Study and design of stability in GH5 cellulases.* Biotechnology and Bioengineering, 2012.
417. Weimann, A., Y. Trukhina, P.B. Pope, S.G.A. Konietzny, and A.C. McHardy, *De novo prediction of the genomic components and capabilities for microbial plant biomass degradation from (meta-) genomes.* Biotechnology for biofuels, 2013. **6**(1): p. 1-13.
418. Chen, Y., A.J. Stipanovic, W.T. Winter, D.B. Wilson, and Y.J. Kim, *Effect of digestion by pure cellulases on crystallinity and average chain length for bacterial and microcrystalline celluloses.* Cellulose, 2007. **14**(4): p. 283-293.
419. Braun, T.F., S. Poulson, J.B. Gully, J.C. Empey, S. Van Way, A. Putnam, and D.F. Blair, *Function of proline residues of MotA in torque generation by the flagellar motor of Escherichia coli.* Journal of bacteriology, 1999. **181**(11): p. 3542-3551.
420. D'Andrea, L.D. and L. Regan, *TPR proteins: the versatile helix.* Trends in biochemical sciences, 2003. **28**(12): p. 655-662.
421. Adindla, S., K.K. Inampudi, K. Guruprasad, and L. Guruprasad, *Identification and analysis of novel tandem repeats in the cell surface proteins of archaeal and bacterial genomes using computational tools.* Comparative and functional genomics, 2004. **5**(1): p. 2-16.

422. Schulein, R., I. Gentschev, H.J. Mollenkopf, and W. Goebel, *A topological model for the haemolysin translocator protein HlyD.* Molecular and General Genetics MGG, 1992. **234**(1): p. 155-163.
423. Stewart, C.S., C. Paniagua, D. Dinsdale, K.J. Cheng, and S.H. Garrow, *Selective isolation and characteristics of Bacteriodes succinogenes from the rumen of a cow.* Applied and Environmental Microbiology, 1981. **41**(2): p. 504-510.
424. Russell, J., *Fermentation of cellodextrins by cellulolytic and noncellulolytic rumen bacteria.* Applied and environmental microbiology, 1985. **49**(3): p. 572-576.
425. Shi, Y. and P. Weimer, *Utilization of individual cellodextrins by three predominant ruminal cellulolytic bacteria.* Applied and environmental microbiology, 1996. **62**(3): p. 1084-1088.
426. Kota, U. and M.B. Goshe, *Advances in qualitative and quantitative plant membrane proteomics.* Phytochemistry, 2011. **72**(10): p. 1040-1060.
427. Kline, K.G., G.L. Finney, and C.C. Wu, *Quantitative strategies to fuel the merger of discovery and hypothesis-driven shotgun proteomics.* Briefings in functional genomics & proteomics, 2009. **8**(2): p. 114-125.
428. Barrera, N.P. and C.V. Robinson, *Advances in the mass spectrometry of membrane proteins: from individual proteins to intact complexes.* Annual review of biochemistry, 2011. **80**: p. 247-271.
429. Huber, T., D. Steiner, D. Röthlisberger, and A. Plückthun, *In vitro selection and characterization of DARPins and Fab fragments for the co-crystallization of membrane proteins: The Na-citrate symporter CitS as an example.* Journal of structural biology, 2007. **159**(2): p. 206-221.
430. Gilmore, J.M. and M.P. Washburn, *Advances in shotgun proteomics and the analysis of membrane proteomes.* Journal of proteomics, 2010. **73**(11): p. 2078-2091.
431. Wessel, D. and U.-I. Flugge, *A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids.* Analytical biochemistry, 1984. **138**(1): p. 141-143.
432. Sambrook, J., *Russel (2001). Molecular cloning: a laboratory manual*, 2001, Cold spring harbor laboratory predd, cold spring harbour, new york.
433. Ow, S.Y., M. Salim, J. Noirel, C. Evans, I. Rehman, and P.C. Wright, *iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly"*. Journal of proteome research, 2009. **8**(11): p. 5347-5355.
434. Helbig, A.O., A.J. Heck, and M. Slijper, *Exploring the membrane proteome—challenges and analytical strategies.* Journal of proteomics, 2010. **73**(5): p. 868-878.
435. Seiler, S. and S. Fleischer, *Isolation of plasma membrane vesicles from rabbit skeletal muscle and their use in ion transport studies.* Journal of Biological Chemistry, 1982. **257**(22): p. 13862-13871.
436. Elschenbroich, S., Y. Kim, J.A. Medin, and T. Kislinger, *Isolation of cell surface proteins for mass spectrometry-based proteomics.* Expert review of proteomics, 2010. **7**(1): p. 141-154.
437. Jarrold, B., J. DeMuth, K. Greis, T. Burt, and F. Wang, *An effective skeletal muscle prefractionation method to remove abundant structural proteins for optimized two-dimensional gel electrophoresis.* Electrophoresis, 2005. **26**(11): p. 2269-2278.
438. Huber, L.A., K. Pfaller, and I. Vietor, *Organelle Proteomics Implications for Subcellular Fractionation in Proteomics.* Circulation research, 2003. **92**(9): p. 962-968.
439. Zhao, Y., W. Zhang, Y. Kho, and Y. Zhao, *Proteomic analysis of integral plasma membrane proteins.* Analytical chemistry, 2004. **76**(7): p. 1817-1823.
440. Nunomura, K., K. Nagano, C. Itagaki, M. Taoka, N. Okamura, Y. Yamauchi, S. Sugano, N. Takahashi, T. Izumi, and T. Isobe, *Cell surface labeling and mass spectrometry reveal diversity of cell surface markers and signaling molecules expressed in undifferentiated*

mouse embryonic stem cells. Molecular & Cellular Proteomics, 2005. **4**(12): p. 1968-1976.
441. Tang, X., W. Yi, G.R. Munske, D.P. Adhikari, N.L. Zakharova, and J.E. Bruce, *Profiling the membrane proteome of Shewanella oneidensis MR-1 with new affinity labeling probes.* Journal of proteome research, 2007. **6**(2): p. 724-734.
442. Myers-Morales, T., C. Cowan, M.E. Gray, C.R. Wulff, C.E. Parker, C.H. Borchers, and S.C. Straley, *A surface-focused biotinylation procedure identifies the Yersinia pestis catalase KatY as a membrane-associated but non-surface-located protein.* Applied and environmental microbiology, 2007. **73**(18): p. 5750-5759.
443. Rigden, D.J., *Analysis of glycoside hydrolase family 98: catalytic machinery, mechanism and a novel putative carbohydrate binding module.* FEBS letters, 2005. **579**(25): p. 5466-5472.
444. Morimoto, K., S. Karita, T. Kimura, K. Sakka, and K. Ohmiya, *Cloning, sequencing, and expression of the gene encoding Clostridium paraputrificum chitinase ChiB and analysis of the functions of novel cadherin-like domains and a chitin-binding domain.* Journal of bacteriology, 1997. **179**(23): p. 7306-7314.
445. Joris, B., S. Englebert, C.-P. Chu, R. Kariyama, L. Daneo-Moore, G.D. Shockman, and J.-M. Ghuysen, *Modular design of the Enterococcus hirae muramidase-2 and Streptococcus faecalis autolysin.* FEMS microbiology letters, 1992. **91**(3): p. 257-264.
446. Reichenbecher, M., F. Lottspeich, and K. Bronnenmeier, *Purification and properties of a cellobiose phosphorylase (CepA) and a cellodextrin phosphorylase (CepB) from the cellulolytic thermophile Clostridium stercorarium.* European Journal of Biochemistry, 1997. **247**(1): p. 262-267.
447. Hosking, E.R., C. Vogt, E.P. Bakker, and M.D. Manson, *The Escherichia coli MotAB Proton Channel Unplugged.* Journal of molecular biology, 2006. **364**(5): p. 921-937.
448. Krishnan, S. and N.V. Prasadarao, *Outer membrane protein A and OprF: versatile roles in Gram-negative bacterial infections.* FEBS Journal, 2012. **279**(6): p. 919-931.
449. Zhang, W., H. Wang, J. Liu, Y. Zhao, K. Gao, and J. Zhang, *Adhesive ability means inhibition activities for lactobacillus against pathogens and S-layer protein plays an important role in adhesion.* Anaerobe, 2013.
450. Green, R., C.C. Hanfrey, K.A. Elliott, D.E. McCloskey, X. Wang, S. Kanugula, A.E. Pegg, and A.J. Michael, *Independent evolutionary origins of functional polyamine biosynthetic enzyme fusions catalysing de novo diamine to triamine formation.* Molecular microbiology, 2011. **81**(4): p. 1109-1124.
451. Worthington, R.J., J.J. Richards, and C. Melander, *Small molecule control of bacterial biofilms.* Organic & Biomolecular Chemistry, 2012. **10**(37): p. 7457-7474.
452. Fischer, G. and F.X. Schmid, *The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell.* Biochemistry, 1990. **29**(9): p. 2205-2212.
453. Kloda, A. and B. Martinac, *Molecular identification of a mechanosensitive channel in archaea.* Biophysical journal, 2001. **80**(1): p. 229-240.
454. Kloda, A. and B. Martinac, *Structural and functional differences between two homologous mechanosensitive channels of Methanococcus jannaschii.* The EMBO journal, 2001. **20**(8): p. 1888-1896.
455. Schülein, R., I. Gentschev, H.-J. Mollenkopf, and W. Goebel, *A topological model for the haemolysin translocator protein HlyD.* Molecular and General Genetics MGG, 1992. **234**(1): p. 155-163.
456. Kumar, P., D.M. Barrett, M.J. Delwiche, and P. Stroeve, *Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production.* Industrial & Engineering Chemistry Research, 2009. **48**(8): p. 3713-3729.

457. Egüés, I., C. Sanchez, I. Mondragon, and J. Labidi, *Effect of alkaline and autohydrolysis processes on the purity of obtained hemicelluloses from corn stalks.* Bioresource technology, 2012. **103**(1): p. 239-248.
458. Hill, J., E. Nelson, D. Tilman, S. Polasky, and D. Tiffany, *Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels.* Proceedings of the National Academy of Sciences, 2006. **103**(30): p. 11206-11210.
459. Balan, V., B. Bals, S.P. Chundawat, D. Marshall, and B.E. Dale, *Lignocellulosic biomass pretreatment using AFEX*, in *Biofuels*2009, Springer. p. 61-77.
460. Adsul, M., M. Singhvi, S. Gaikaiwari, and D. Gokhale, *Development of biocatalysts for production of commodity chemicals from lignocellulosic biomass.* Bioresource technology, 2011. **102**(6): p. 4304-4312.
461. Lee, H.-J., W.-S. Lim, and J.-W. Lee, *Improvement of ethanol fermentation from lignocellulosic hydrolysates by the removal of inhibitors.* Journal of Industrial and Engineering Chemistry, 2013.
462. Lynd, L.R., *Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy.* Annual review of energy and the environment, 1996. **21**(1): p. 403-465.
463. Zhu, J. and X. Pan, *Woody biomass pretreatment for cellulosic ethanol production: technology and energy consumption evaluation.* Bioresource technology, 2010. **101**(13): p. 4992-5002.
464. Chaturvedi, V. and P. Verma, *An overview of key pretreatment processes employed for bioconversion of lignocellulosic biomass into biofuels and value added products.* 3 Biotech, 2013. **3**(5): p. 415-431.
465. Hsu, T., R. Ladisch, and G. Tsao, *Alcohol from cellulose.* chemical intermediates, 1980. **1203**(3): p. 3.
466. LaForge, F. and C. Hudson, *The Preparation of Several Useful Substances from Corn Cobs.* Industrial & Engineering Chemistry, 1918. **10**(11): p. 925-927.
467. Chang, V.S. and M.T. Holtzapple. *Fundamental factors affecting biomass enzymatic reactivity*. in *Twenty-First Symposium on Biotechnology for Fuels and Chemicals*. 2000. Springer.
468. Qureshi, N. and T.C. Ezeji, *Butanol,'a superior biofuel'production from agricultural residues (renewable biomass): recent progress in technology.* Biofuels, Bioproducts and Biorefining, 2008. **2**(4): p. 319-330.
469. Guedon, E., M. Desvaux, and H. Petitdemange, *Improvement of cellulolytic properties of Clostridium cellulolyticum by metabolic engineering.* Applied and environmental microbiology, 2002. **68**(1): p. 53-58.
470. Wyman, C.E., B.E. Dale, R.T. Elander, M. Holtzapple, M.R. Ladisch, and Y. Lee, *Comparative sugar recovery data from laboratory scale application of leading pretreatment technologies to corn stover.* Bioresource technology, 2005. **96**(18): p. 2026-2032.
471. Wang, L. and H. Chen, *Increased fermentability of enzymatically hydrolyzed steam-exploded corn stover for butanol production by removal of fermentation inhibitors.* Process Biochemistry, 2011. **46**(2): p. 604-607.
472. George, H., J. Johnson, W. Moore, L. Holdeman, and J. Chen, *Acetone, isopropanol, and butanol production by Clostridium beijerinckii (syn. Clostridium butylicum) and Clostridium aurantibutyricum.* Applied and environmental microbiology, 1983. **45**(3): p. 1160-1163.
473. Ramirez, R.S., M. Holtzapple, and N. Piamonte, *Fundamentals of Biomass Pretreatment at High pH.* Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals, 2013: p. 145-167.

474. Kaparaju, P., M. Serrano, A.B. Thomsen, P. Kongjan, and I. Angelidaki, *Bioethanol, biohydrogen and biogas production from wheat straw in a biorefinery concept.* Bioresource technology, 2009. **100**(9): p. 2562-2568.
475. Saha, B.C., L.B. Iten, M.A. Cotta, and Y.V. Wu, *Dilute acid pretreatment, enzymatic saccharification and fermentation of wheat straw to ethanol.* Process Biochemistry, 2005. **40**(12): p. 3693-3700.
476. Talebnia, F., D. Karakashev, and I. Angelidaki, *Production of bioethanol from wheat straw: an overview on pretreatment, hydrolysis and fermentation.* Bioresource technology, 2010. **101**(13): p. 4744-4753.
477. McIntosh, S. and T. Vancov, *Optimisation of dilute alkaline pretreatment for enzymatic saccharification of wheat straw.* biomass and bioenergy, 2011. **35**(7): p. 3094-3103.
478. Xiao, H., Y. Gu, Y. Ning, Y. Yang, W.J. Mitchell, W. Jiang, and S. Yang, *Confirmation and elimination of xylose metabolism bottlenecks in glucose phosphoenolpyruvate-dependent phosphotransferase system-deficient Clostridium acetobutylicum for simultaneous utilization of glucose, xylose, and arabinose.* Applied and environmental microbiology, 2011. **77**(22): p. 7886-7895.
479. Wietzke, M. and H. Bahl, *The redox-sensing protein Rex, a transcriptional regulator of solventogenesis in Clostridium acetobutylicum.* Applied Microbiology and Biotechnology, 2012. **96**(3): p. 749-761.
480. Hüsemann, M.H. and E.T. Papoutsakis, *Solventogenesis in Clostridium acetobutylicum fermentations related to carboxylic acid and proton concentrations.* Biotechnology and bioengineering, 1988. **32**(7): p. 843-852.
481. He, Q. and H. Chen, *Improved efficiency of butanol production by absorbent fermentation with a renewable carrier.* Biotechnology for biofuels, 2013. **6**(1): p. 121.
482. Chen, L., C. Xin, P. Deng, J. Ren, H. Liang, and F. Bai, *Butanol production from hydrolysate of Jerusalem artichoke juice by Clostridium acetobutylicum L7].* Sheng wu gong cheng xue bao= Chinese journal of biotechnology, 2010. **26**(7): p. 991.
483. Kyazze, G., R. Dinsdale, F.R. Hawkes, A.J. Guwy, G.C. Premier, and I. Donnison, *Direct fermentation of fodder maize, chicory fructans and perennial ryegrass to hydrogen using mixed microflora.* Bioresource Technology, 2008. **99**(18): p. 8833-8839.
484. Cao, G., N. Ren, A. Wang, D.-J. Lee, W. Guo, B. Liu, Y. Feng, and Q. Zhao, *Acid hydrolysis of corn stover for biohydrogen production using Thermoanaerobacterium thermosaccharolyticum W16.* international journal of hydrogen energy, 2009. **34**(17): p. 7182-7188.
485. Klinke, H.B., A. Thomsen, and B.K. Ahring, *Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass.* Applied Microbiology and Biotechnology, 2004. **66**(1): p. 10-26.
486. Stalbrand, H., S.D. Mansfield, J.N. Saddler, D.G. Kilburn, R.A.J. Warren, and N.R. Gilkes, *Analysis of Molecular Size Distributions of Cellulose Molecules during Hydrolysis of Cellulose by Recombinant Cellulomonas fimiβ-1, 4-Glucanases.* Applied and environmental microbiology, 1998. **64**(7): p. 2374-2379.
487. Malburg, S., L. Malburg, T. Liu, A.H. Iyo, and C.W. Forsberg, *Catalytic properties of the cellulose-binding endoglucanase F from Fibrobacter succinogenes S85.* Applied and environmental microbiology, 1997. **63**(6): p. 2449-2453.
488. Kraan, S., *Mass-cultivation of carbohydrate rich macroalgae, a possible solution for sustainable biofuel production.* Mitigation and Adaptation Strategies for Global Change, 2013. **18**(1): p. 27-46.
489. Himmel, M.E., Q. Xu, Y. Luo, S.-Y. Ding, R. Lamed, and E.A. Bayer, *Microbial enzyme systems for biomass conversion: emerging paradigms.* Biofuels, 2010. **1**(2): p. 323-341.
490. Lacis, A.A., G.A. Schmidt, D. Rind, and R.A. Ruedy, *Atmospheric CO2: principal control knob governing Earth's temperature.* science, 2010. **330**(6002): p. 356-359.

491. Shafiee, S. and E. Topal, *When will fossil fuel reserves be diminished?* Energy Policy, 2009. **37**(1): p. 181-189.
492. Demain, A.L., *Biosolutions to the energy problem.* Journal of industrial microbiology & biotechnology, 2009. **36**(3): p. 319-332.
493. Ragauskas, A.J., C.K. Williams, B.H. Davison, G. Britovsek, J. Cairney, C.A. Eckert, W.J. Frederick, J.P. Hallett, D.J. Leak, and C.L. Liotta, *The path forward for biofuels and biomaterials.* science, 2006. **311**(5760): p. 484-489.
494. Jonsson, L.J., B. Alriksson, and N.-O. Nilvebrant, *Bioconversion of lignocellulose: inhibitors and detoxification.* Biotechnology for biofuels, 2013. **6**(1): p. 16.
495. Metzger, J.O. and A. Hüttermann, *Sustainable global energy supply based on lignocellulosic biomass from afforestation of degraded areas.* Naturwissenschaften, 2009. **96**(2): p. 279-288.
496. Olson, D.G., J.E. McBride, A. Joe Shaw, and L.R. Lynd, *Recent progress in consolidated bioprocessing.* Current opinion in biotechnology, 2012. **23**(3): p. 396-405.
497. Stephanopoulos, G., *Challenges in engineering microbes for biofuels production.* science, 2007. **315**(5813): p. 801-804.
498. Khanal, S., *Anaerobic biotechnology for bioenergy production: principles and applications*2009: John Wiley & Sons.
499. Zhang, C. and J.B. Hughes, *Biodegradation pathways of hexahydro-1, 3, 5-trinitro-1, 3, 5-triazine (RDX) by Clostridium acetobutylicum cell-free extract.* Chemosphere, 2003. **50**(5): p. 665-671.
500. Khan, T.A., *Anaerobic transformation of 2, 4, 6-trinitrotoluene by Clostridium acetobutylicum.* ETD http://hdl. handle. net/1911/14073, 1996.
501. Cai, X., M. Servinsky, J. Kiel, C. Sund, and G.N. Bennett, *Analysis of redox responses during TNT transformation by Clostridium acetobutylicum ATCC 824 and mutants exhibiting altered metabolism.* Applied microbiology and biotechnology, 2013. **97**(10): p. 4651-4663.
502. Kovács, K., B.J. Willson, K. Schwarz, J.T. Heap, A. Jackson, D.N. Bolam, K. Winzer, and N.P. Minton, *Secretion and assembly of functional mini-cellulosomes from synthetic chromosomal operons in Clostridium acetobutylicum ATCC 824.* Biotechnology for biofuels, 2013. **6**(1): p. 117.

Appendix 4.1 Proteins identiefied with significant changes in comparision of ExpC/ExpCL

Locus ID	Protein	Fold change	% Coverage	Peptides	Locus ID	Protein	Fold change	% Coverage	Peptides
	Alcohol metabolism					Signal transduction/chemotaxis/secretion/trafficing			
CA_P0165	Acetoacetate decarboxylase	1.36	45	26	**CA_C3734**	tRNA modification GTPase	-1.21	2	8
CA_C3392	NADH-dependent butanol dehydrogenase	-3.05	9	11	**CA_C2942**	S-ribosylhomocysteine lyase	1.20	17	4
	Amino acid biosynthesis				**CA_C2433**	HtrA-like serine protease	1.35	21	30
CA_C3254	Gamma-glutamyl phosphate reductase	1.44	5	9	**CA_C2220**	Chemotaxis histidine kinase, CheA (Contains CheW-like adaptor domain)	-1.30	10	20
CA_C3600	4-hydroxy-tetrahydrodipicolinate synthase 2	-1.17	7	3	**CA_C2161**	Flagellar motor switch protein FliG	1.35	4	15
CA_C2973	2-keto-3-deoxy-6-phosphogluconate aldolase	-1.12	35	15	**CA_C0585**	N-terminal CheY reciever domain fused to C-terminal uncharacterized CheX-like domain	-1.40	3	5
CA_C2658	Glutamine synthetase type III	1.20	12	46		Stress responce			
CA_C2235	Cysteine synthase	1.69	30	14	**CA_C3598**	Reverse rubrerythrin-1 (revRbr 1) (NADH peroxidase) (NPXase) (Npx)	-1.20	20	19
CA_C1684	TYPA/BIPA type GTPase	-1.22	8	12	**CA_C2637**	Lon protease	-1.39	4	34
CA_C0568	Aspartate semialdehyde dehydrogenase	-1.25	11	14	**CA_C1549**	Glutathione peroxidase	-1.23	10	3
	Carbohydrate metabolism					Transcription			
CA_C0517	6-phosphofructokinase	1.42	10	19	**CA_C3472**	Protein containing transcriptional regulator domain	-1.29	2	15
CA_C0518	Pyruvate kinase	1.34	11	43	**CA_C3198**	Transcription elongation factor GreA (Transcript cleavage factor GreA)	1.20	18	11
CA_C3376	Possible pectin degradation protein	-4.46	10	3	**CA_C3142**	DNA-directed RNA polymerase subunit beta' (RNAP subunit beta')	1.29	9	50
CA_C2680	Glucose-6-phosphate isomerase	1.93	24	18	**CA_C2889**	Transcription termination factor Rho	1.24	5	19
CA_C2613	Transcriptional regulators of NagC/XylR family	1.43	9	8	**CA_C2295**	Probable transcriptional regulatory protein	2.09	6	10
CA_C2337	Phosphomannomutase	-1.58	14	47	**CA_C1719**	DNA-directed RNA polymerase subunit omega (RNAP omega subunit)	1.38	43	5
CA_C1408	Phospho-beta-glucosidase	1.54	2	9	**CA_C0807**	Cold shock protein	1.78	46	20
CA_C1287	HIT family hydrolase	2.77	16	3	**CA_C0461**	Mercuric resistance operon regulatory protein, MerR family	1.52	5	5
CA_C0713	Enolase	1.26	32	24		Translation			
CA_C0712	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	1.52	15	22	**CA_C1722**	Peptide deformylase 1 (PDF 1)	1.46	9	4
	Cell cycle/cell division				**CA_C1723**	Methionyl-tRNA formyltransferase	1.33	7	19
CA_C3202	ATP-dependent zinc metalloprotease FtsH	-3.23	7	20	**CA_C3147**	50S ribosomal protein L1	1.21	35	37
CA_C2120	Cell division protein SepF	9.76	4	5	**CA_C3145**	50S ribosomal protein L7/L12	-1.08	46	294
CA_C1013	FTSA related protein, predicted ATPases of the HSP70 family	1.43	1	7	**CA_C3140**	30S ribosomal protein S12	-1.17	10	20
CA_C2118	Cell division protein DivIVA	1.98	8	13	**CA_C3138**	Elongation factor G (EF-G)	1.37	24	64
CA_C0498	Cell division protein	-1.38	6	7	**CA_C3133**	50S ribosomal protein L3	1.20	23	31
	Cell envalope/cell wall biogenesis				**CA_C3132**	50S ribosomal protein L4	1.44	9	26
CA_C3194	UDP-N-acetylmuramoylalanine--D-glutamate ligase	-1.29	4	3	**CA_C3124**	30S ribosomal protein S17	1.16	19	76
CA_C2335	UTP-glucose-1-phosphate uridylyltransferase	-1.44	22	17	**CA_C3123**	50S ribosomal protein L14	1.18	67	17
CA_C2333	Glucose-1-phosphate thymidylyltransferase	-1.47	16	10	**CA_C3106**	30S ribosomal protein S11	1.40	21	7
CA_C2332	dTDP-glucose 4,6-dehydratase	-1.33	7	13	**CA_C3105**	30S ribosomal protein S4 A	1.55	33	26

	Cofactor biosynthesis				**CA_C3146**	50S ribosomal protein L10	1.25	48	43
CA_C0594	Pyridoxal biosynthesis lyase PdxS	1.27	31	22	**CA_C2888**	50S ribosomal protein L31	1.49	19	8
	DNA metabolism				**CA_C2669**	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B 1	1.67	7	10
CA_C0006	DNA gyrase subunit B	-1.43	4	9	**CA_C2361**	Translation initiation factor IF-3	-1.70	7	12
CA_C0007	DNA gyrase subunit A	-1.21	11	57	**CA_C2360**	50S ribosomal protein L35	-1.63	19	7
CA_C3729	Stage 0 sporulation J, ParB family of DNA-binding proteins	-1.23	6	15	**CA_C2359**	50S ribosomal protein L20	-1.55	32	46
CA_C3036	Superfamily I DNA helicase	-1.17	0	25	**CA_C1802**	Translation initiation factor IF-2	1.33	3	14
CA_C2382	Single-strand DNA-binding protein	1.59	35	21	**CA_C1790**	Ribosome-recycling factor (RRF)	-1.42	14	21
CA_C0127	Recombination protein	-1.56	4	9	**CA_C1788**	Elongation factor Ts (EF-Ts)	-1.31	38	60
	Energy metabolism/electron transport				**CA_C1787**	30S ribosomal protein S2	1.36	40	32
CA_C2709	Electron transfer flavoprotein subunit alpha (Alpha-ETF)	-1.08	41	171	**CA_C1259**	50S ribosomal protein L27	1.20	42	65
	Fatty acid biosynthesis/lipid metabolism					Transportation and binding			
thlB	Thiolase B	1.28	12	92	**CA_C2869**	ATP synthase subunit b (ATP synthase F(0) sector subunit b)	1.39	25	18
CA_C2708	3-hydroxybutyryl-CoA dehydrogenase	-1.11	43	340	**CA_C3632**	Oligopeptide ABC transporter, periplasmic substrate-binding component	-1.20	8	20
CA_C2711	Acyl-CoA dehydrogenase, short-chain specific	-1.15	29	20	**CA_C3354**	Probable cation efflux pump	-1.10	1	37
CA_C2712	3-hydroxybutyryl-CoA dehydratase	-1.16	53	55	**CA_C3282**	ABC-type multidrug/protein/lipid transport system, ATPase component	-1.20	4	7
CA_C3076	Phosphate butyryltransferase	-1.25	47	47	**CA_C3012**	ATPase component of ABC transporter, with duplicated ATPase domains	1.79	1	12
CA_C3575	Malonyl CoA-acyl carrier protein transacylase	-1.19	15	9	**CA_C1816**	Ribonuclease Y	-1.18	3	20
CA_C3572	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	-1.18	4	9	**CA_C1590**	2-oxoglutarate/malate translocator	1.50	2	3
	Intermediary metabolism/multibiosynthetic pathway				**CA_C1353**	Phosphotransferase system IIC component	1.35	2	5
CA_C2873	Acetyl-CoA acetyltransferase	-1.09	40	149	**CA_C0570**	PTS enzyme II, ABC component	-1.23	1	8
CA_C1743	Acetate kinase	-1.40	10	5	**CA_C0147**	ABC transporter, ATP-binding protein	-1.55	24	31
hydA	Hydrogenase I	1.13	2	13	**CA_C0146**	Related to ABC transporter permease component	-1.61	2	5
CA_C3371	2-enoate reductase	-3.02	2	10	**CA_C0139**	Predicted permease	-1.44	2	10
CA_C3090	Fumarate hydratase	-1.95	5	5	**CA_C2255**	Predicted permease	1.16	1	24
CA_C2935	Predicted acetyltransferase	1.12	7	4	**CA_C1705**	Periplasmic phosphate-binding protein	2.99	2	6
CA_C2584	Protein containing ChW-repeats	1.46	2	5	**CA_C0880**	Periplasmic amino acid binding protein	1.45	5	5
CA_C2572	Possible aminoglycoside phosphotransferase	-4.48	2	16	**CA_C0111**	Glutamine-binding periplasmic protein fused to glutamine permease	1.74	2	7
CA_C2283	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	1.30	13	15		Unknown function			
CA_C1819	Aspartate Aminotransferase	-4.63	5	5	**CA_C3537**	Fragment of SECA	-1.26	11	8
CA_C0944	Transketolase	1.94	5	5	**CA_C3341**	Multimeric flavodoxin WrbA family protein	1.36	26	19
CA_C0097	Porphobilinogen deaminase	2.18	4	4	**CA_C3264**	Uncharacterized conserved protein, YTFJ B.subtilis ortholog	1.40	23	9
	Nucleoside/nucleotide metabolism				**CA_C3248**	Uncharacterized protein	1.34	13	11
CA_C3627	7-cyano-7-deazaguanine synthase	-1.22	5	15	**CA_C3094**	Uncharacterized consrved protein	-1.96	9	4
CA_C3224	PUR operon repressor, Adenine/guanine phosphoribosyltransferase family	-1.18	2	6	**CA_C3008**	CBS domain containing protein	1.41	8	34
CA_C2879	Uracil phosphoribosyltransferase	1.28	24	22	**CA_C2853**	Uncharacterized protein	1.40	2	6
CA_C1395	Bifunctional purine biosynthesis protein PurH(Inosinicase)	1.26	16	34	**CA_C2817**	Predicted membrane protein	1.73	2	11
CA_C0480	Oxygen-sensitive ribonucleoside-triphosphate reductase nrdD	-1.38	6	15	**CA_C2643**	Uncharacterized protein	-1.29	3	3
	Protein biosynthesis/modification				**CA_C2564**	Uncharacterized protein	-1.27	26	44
CA_C3260	Asparagine--tRNA ligase	1.52	11	28	**CA_C2387**	Uncharacterized protein	1.34	39	12

CA_C3195	Glycine--tRNA ligase	-1.21	4	13	**CA_C2364**	Uncharacterized protein	-1.73	3	11
CA_C3189	ATPases with chaperone activity clpC, two ATP-binding domai	-1.23	6	11	**CA_C2251**	Uncharacterized conserved membrane protein	1.38	1	5
CA_C3038	Isoleucine--tRNA ligase	1.35	3	15	**CA_C2193**	Ucharacterized protein, CGEB homolog	1.79	2	27
CA_C2847	Ribosome-associated protein Y	1.28	12	39	**CA_C2085**	Uncharacterized protein from alkaline shock protein family	-1.69	30	10
CA_C2740	Histidine--tRNA ligase	-1.59	4	8	**CA_C1880**	Uncharacterized protein	1.56	13	3
CA_C2370	33 kDa chaperonin	-1.62	3	4	**CA_C1828**	TldD protein	1.30	2	7
CA_C2362	Threonine--tRNA ligase	-1.57	4	26	**CA_C1817**	Stage V sporulation protein, spoVS	-1.29	52	26
CA_C2357	Phenylalanine--tRNA ligase alpha subunit	-1.59	6	10	**CA_C1717**	UPF0296 protein	1.19	37	6
CA_C2356	Phenylalanine--tRNA ligase beta subunit	-1.48	3	12	**CA_C1679**	UPF0297 protein	-1.79	20	9
CA_C2269	Aspartate--tRNA ligase	-1.29	10	17	**CA_C1306**	Uncharacterized protein	1.21	5	3
CA_C2264	Serine hydroxymethyltransferase	1.23	10	20	**CA_C1304**	Uncharacterized conserved protein, predicted metal-binding	1.70	4	4
CA_C1297	N-terminal of elongation factor Ts	1.59	4	8	**CA_C1242**	MreB	1.37	10	13
CA_C0990	Glutamate-tRNA ligase	1.36	17	24	**CA_C1171**	Uncharacterized protein	-1.81	6	3
CA_C0646	Leucine-tRNA ligase	1.15	2	12	**CA_C0504**	FHA-domain containing secreted protein	1.32	4	36
	Protein fate								
CA_C2846	Protein translocase subunit SecA	-1.44	9	48					
CA_C2278	Protein translocase subunit SecD	1.25	2	4					
CA_C1052	Membrane protease subunit, stomatin/prohibitin homolog	1.63	19	16					

Appendix 4.2 Proteins identiefied with significant changes in comparision of ExpCL/StaCL

Locus ID	Protein	Fold change	% Coverage	Peptides	Locus ID	Protein	Fold change	% Coverage	Peptides
	Alcohol metabolism (solventogenesis)				**CA_C3674**	Two CBS domain containing protein	-4.41	20	11
CA_P0165	Acetoacetate decarboxylase	1.40	45	26	**CA_C3150**	Preprotein translocase subunit SecE	-2.30	9	6
CA_C3299	NADH-dependent butanol dehydrogenase A	-1.60	5	11	**CA_C3006**	Zn-dependent peptidase, insulinase family	-1.30	3	19
CA_C3298	NADH-dependent butanol dehydrogenase B	-1.68	14	31	**CA_C2846**	Protein translocase subunit SecA	1.68	9	48
CA_C3392	NADH-dependent butanol dehydrogenase	8.86	9	11	**CA_C2769**	Peptidyl-prolyl cis-trans isomerase	1.28	10	7
	Amino acid biosynthesis				**CA_C2646**	Signal peptidase I	-1.93	7	3
CA_C3600	4-hydroxy-tetrahydrodipicolinate synthase 2	-1.91	7	3	**CA_C2278**	Protein translocase subunit SecD	1.94	2	4
CA_C3254	Gamma-glutamyl phosphate reductase	1.36	5	9	**CA_C1760**	Signal peptidase I	-2.94	3	6
CA_C2856	S-adenosylmethionine synthase	-1.34	9	50		Ribosome metabolism			
CA_C2378	4-hydroxy-tetrahydrodipicolinate synthase 1	-2.98	7	9	**CA_C3146**	50S ribosomal protein L10	-1.38	48	43
CA_C0091	Ketol-acid reductoisomerase	-1.58	10	7	**CA_C1964**	Ribosomal protein S1	-1.77	2	3
CA_C0021	Serine--tRNA ligase	-2.15	2	7	**CA_C1798**	Ribosome maturation factor RimP	-1.23	6	4
CA_C2973	2-keto-3-deoxy-6-phosphogluconate aldolase	1.60	35	15		signal transduction/chemotaxis/ secretion/trafficing			
CA_C2783	O-acetylhomoserine sulfhydrylase	2.28	8	12	**CA_C3734**	tRNA modification GTPase	-1.38	2	8
CA_C2644	Carbamoyl-phosphate synthase large chain	-1.92	2	14	**CA_C3430**	Membrane associated, signal transduction histidine kinase-like ATPase	-1.54	2	4
CA_C2095	Aminopeptidase P AMPP/PEPQ family enzyme	-2.35	2	5	**CA_C2942**	S-ribosylhomocysteine lyase	-1.35	17	4
CA_C1684	TYPA/BIPA type GTPase	-1.58	8	12	**CA_C2433**	HtrA-like serine protease	1.36	21	30

CA_C0568	Aspartate semialdehyde dehydrogenase	-1.46	11	14	**CA_C2220**	Chemotaxis histidine kinase, CheA	2.95	10	20
CA_C1001	Aspartate aminotransferase	-3.46	3	3	**CA_C2218**	Chemotaxis signal receiving protein CheY	-1.90	15	8
	Carbohydrate metabolism(sugar, polysachharide)				**CA_C1820**	Phosphocarrier Protein	1.23	42	103
CA_C0517	6-phosphofructokinase (Phosphofructokinase)	-1.35	10	19	**CA_C0585**	N-terminal CheY reciever domain fused to C-terminal uncharacterized CheX-like domain	-1.41	3	5
CA_C0518	Pyruvate kinase	-1.21	11	43		Sporulation/genrmination			
CA_C0709	Glyceraldehyde-3-phosphate dehydrogenase	1.30	38	171	**CA_C3223**	Putative septation protein SpoVG	-1.84	45	25
CA_C0710	Phosphoglycerate kinase	1.58	48	97		Stress responce			
CA_C0711	Triosephosphate isomerase	1.56	48	35	**CA_C3714**	18 kDa heat shock protein	-1.72	29	22
CA_C3661	Glycosyltransferase	-2.75	5	4	**CA_C3598**	Reverse rubrerythrin-1 (revRbr 1) (NADH peroxidase)	-1.97	20	19
CA_C3376	Possible pectin degradation protein	16.27	10	3	**CA_C3315**	Chaperone protein HtpG	1.35	4	8
CA_C3032	Galactose mutarotase related enzyme	2.21	4	4	**CA_C1549**	Glutathione peroxidase	1.94	10	3
CA_C2660	Pyruvate carboxylase	-1.28	2	19		Transcription			
CA_C2337	Phosphomannomutase	2.64	14	47	**CA_C1300**	RNA polymerase sigma factor RpoD	1.44	16	19
CA_C1408	Phospho-beta-glucosidase	1.79	2	9	**CA_C3472**	Protein containing transcriptional regulator domain	-1.70	2	15
CA_C1287	HIT family hydrolase	-1.50	16	3	**CA_C3283**	Transcriptional regulator	-1.97	15	9
CA_C1036	Pyruvate kinase	-1.30	13	20	**CA_C3198**	Transcription elongation factor GreA	1.35	18	11
CA_C0980	Pyruvate-formate lyase	-1.49	5	9	**CA_C3149**	Transcription antitermination protein nusG	-2.78	5	5
CA_C0743	6-phospho-beta-glucosidase	-1.48	15	18	**CA_C3143**	DNA-directed RNA polymerase subunit beta	-2.49	7	30
	Cell cycle/cell division				**CA_C3142**	DNA-directed RNA polymerase subunit beta	-1.18	9	50
CA_C3459	Homolog of cell division GTPase FtsZ, diverged	-1.50	2	4	**CA_C2990**	Cold shock protein	-1.75	25	15
CA_C3202	ATP-dependent zinc metalloprotease FtsH	5.90	7	20	**CA_C2939**	Response regulator	-1.23	3	6
CA_C2641	Trigger factor	-1.29	34	89	**CA_C2889**	Transcription termination factor Rho	-1.21	5	19
CA_C2120	Cell division protein SepF	-6.15	4	5	**CA_C2842**	Transcription accessory protein TEX	1.30	6	13
CA_C2118	Cell division protein DivIVA	-2.11	8	13	**CA_C2084**	N utilization substance protein B homolog	-1.59	7	6
CA_C1812	DNA translocase FtsK	-2.03	1	8	**CA_C1838**	Adenosine tRNA methylthiotransferase	-2.24	6	12
CA_C0497	Cell division ATP-binding protein	1.68	14	9	**CA_C1808**	Polyribonucleotide nucleotidyltransferase	-1.45	10	22
	Cell envalope/cell wall biogenesis				**CA_C1799**	Transcription terminator NusA	1.68	16	16
CA_C2874	Putative UDP-N-acetylglucosamine 2-epimerase	-2.01	8	15	**CA_C0461**	Mercuric resistance operon regulatory protein	-1.87	5	5
CA_C3222	Bifunctional protein GlmU	-2.04	5	11		Translation			
CA_C3194	UDP-N-acetylmuramoylalanine-D-glutamate ligase	1.13	4	3	**CA_C1722**	Peptide deformylase 1 (PDF 1)	-1.87	9	4
CA_C2895	D-alanine-D-alanine ligase	1.84	14	6	**CA_C1723**	Methionyl-tRNA formyltransferase	-1.43	7	19
CA_C2335	UTP-glucose-1-phosphate uridylyltransferase	2.54	22	17	**CA_C3722**	30S ribosomal protein S18	-3.12	27	8
CA_C2333	Glucose-1-phosphate thymidylyltransferase	2.36	16	10	**CA_C3717**	50S ribosomal protein L9	1.87	42	34
CA_C2332	dTDP-glucose 4,6-dehydratase	1.97	7	13	**CA_C3148**	50S ribosomal protein L11	-2.05	46	17
	Cofactor biosynthesis				**CA_C3145**	50S ribosomal protein L7/L12	1.20	46	294
CA_C3626	GTP cyclohydrolase 1	-2.00	6	10	**CA_C3140**	30S ribosomal protein S12	-2.40	10	20
CA_C3586	Putative competence-damage inducible protein	-1.81	2	7	**CA_C3139**	30S ribosomal protein S7	-1.28	40	32
CA_C3292	NifU homolog involved in Fe-S cluster formation	1.61	7	6	**CA_C3138**	Elongation factor G (EF-G)	-1.72	24	64
CA_C3291	Selenocysteine lyase, NifS family	1.82	3	9	**CA_C3136**	Elongation factor Tu (EF-Tu)	-1.09	46	272

CA_C3290	Iron-regulated ABC-type transporter membrane component	2.12	8	8	**CA_C3133**	50S ribosomal protein L3	1.36	23	31
CA_C3289	Iron-regulated ABC-type transporter membrane component	1.44	6	22	**CA_C3132**	50S ribosomal protein L4	1.42	9	26
CA_C2475	Possible 5-Nitroimidazole antibiotics resistance protein	-2.51	5	6	**CA_C3131**	50S ribosomal protein L23	-1.19	78	100
	DNA metabolism (replication, recombination, repair)				**CA_C3129**	30S ribosomal protein S19	1.55	42	48
CA_C1283	Chaperone protein DnaJ	1.53	2	9	**CA_C3127**	30S ribosomal protein S3	1.40	10	19
CA_C0007	DNA gyrase subunit A	-1.18	11	57	**CA_C3126**	50S ribosomal protein L16	1.80	10	8
CA_C3735	Predicted RNA-binding protein Jag, SpoIIIJ-associated	-1.56	7	7	**CA_C3125**	50S ribosomal protein L29	1.78	39	22
CA_C3729	Stage 0 sporulation J, ParB family of DNA-binding proteins	-1.48	6	15	**CA_C3123**	50S ribosomal protein L14	1.23	67	17
CA_C3723	Single-stranded DNA-binding protein 3	-1.62	9	11	**CA_C3122**	50S ribosomal protein L24	1.44	39	38
CA_C3587	DNA replication protein DnaD	-1.73	2	7	**CA_C3121**	50S ribosomal protein L5	1.58	60	94
CA_C3211	DNA binding protein HU	-1.46	50	197	**CA_C3119**	30S ribosomal protein S8	1.43	50	30
CA_C3036	Superfamily I DNA helicase	1.64	0	25	**CA_C3118**	50S ribosomal protein L6	1.61	32	22
CA_C2707	8-oxoguanine-DNA-glycosylase	1.72	8	8	**CA_C3117**	50S ribosomal protein L18	1.59	39	34
CA_C0127	Recombination protein RecR	2.39	4	9	**CA_C3116**	30S ribosomal protein S5	1.76	47	20
CA_C0126	Nucleoid-associated protein	1.74	15	13	**CA_C3114**	50S ribosomal protein L15	2.11	27	17
	Energy metabolism/electron transport				**CA_C3105**	30S ribosomal protein S4 A	1.33	33	26
CA_C2709	Electron transfer flavoprotein subunit alpha	-1.31	41	171	**CA_C2888**	50S ribosomal protein L31	1.28	19	8
CA_C2229	Pyruvate-flavodoxin oxidoreductase	-1.58	24	157	**CA_C2361**	Translation initiation factor IF-3	-1.69	7	12
	Fatty acid biosynthesis/lipid metabolism				**CA_C2360**	50S ribosomal protein L35	-1.76	19	7
thlB	Thiolase B	1.29	12	92	**CA_C2359**	50S ribosomal protein L20	-1.60	32	46
CA_C2708	3-hydroxybutyryl-CoA dehydrogenase	-1.27	43	340	**CA_C1807**	30S ribosomal protein S15	-1.80	37	22
CA_C2711	Acyl-CoA dehydrogenase, short-chain specific	1.16	29	20	**CA_C1787**	30S ribosomal protein S2	-1.23	40	32
CA_C3075	Butyrate kinase 1 (BK 1) (BKI)	-1.31	36	81	**CA_C1274**	30S ribosomal protein S20	1.34	18	21
CA_C3575	Malonyl CoA-acyl carrier protein transacylase	-1.93	15	9	**CA_C1259**	50S ribosomal protein L27	-1.45	42	65
CA_C3573	3-oxoacyl-[acyl-carrier-protein] synthase 2	-2.01	4	6		Transport and binding			
CA_C3572	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	-1.87	4	9	**CA_C2869**	ATP synthase subunit b	-1.54	25	18
CA_C3571	3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	-3.38	22	6	**CA_C3632**	Oligopeptide ABC transporter	-1.58	8	20
CA_C3570	Biotin carboxylase	-4.30	5	8	**CA_C3288**	Iron-regulated ABC transporter ATPase subunit	1.68	13	8
CA_C3569	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	-3.81	9	10	**CA_C3282**	ABC-type multidrug/protein/lipid transport system	-2.38	4	7
CA_C0489	Holo-[acyl-carrier-protein] synthase (Holo-ACP synthase)	1.42	6	4	**CA_C3087**	Phosphoenolpyruvate-protein phosphotransferase	-1.60	14	39
	Intermediary metabolism/other metabolic pathways/multibiosynthetic pathways				**CA_C3012**	ATPase component of ABC transporter	2.40	1	12
CA_C2873	Acetyl-CoA acetyltransferase	-1.51	40	149	**CA_C2982**	MinD family ATPase	1.65	7	4
CA_C3601	Nudix (MutT) family hydrolase	-2.07	5	5	**CA_C2734**	ABC-type multidrug transport system	-1.19	2	44
CA_C3576	Dioxygenase related to 2-nitropropane dioxygenase	-1.89	11	20	**CA_C1816**	Ribonuclease Y (RNase Y)	-1.45	3	20
CA_C3371	2-enoate reductase	12.00	2	10	**CA_C1590**	2-oxoglutarate/malate translocator	1.74	2	3
CA_C3314	Nitroreductase family protein	1.46	7	13	**CA_C0570**	PTS enzyme II, ABC component	-1.42	1	8
CA_C3221	Ribose-phosphate pyrophosphokinase	-1.94	11	12	**CA_C0268**	ABC transporter ATP-binding protein	-1.63	3	5
CA_C3090	Fumarate hydratase, subunit B	-1.91	5	5	**CA_C0147**	ABC transporter, ATP-binding protein	2.45	24	31
CA_C2935	Predicted acetyltransferase	-1.27	7	4	**CA_C0146**	Related to ABC transporter permease component	2.58	2	5
CA_C2775	Phosphohydrolase from calcineurin family	1.88	2	5	**CA_C0139**	Predicted permease	2.19	2	10

CA_C2572	Possible aminoglycoside phosphotransferase	5.75	2	16	**CA_P0068**	Mannose-specific phosphotransferase system component IID	1.33	10	7
CA_C2542	FAD/FMN-containing dehydrogenase	-6.71	4	14	**CA_C2864**	ATP synthase epsilon chain	1.45	18	10
CA_C2283	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	1.82	13	15	**CA_C0984**	Methionine import ATP-binding protein MetN	-1.75	2	3
CA_C1778	Amidase from nicotinamidase family	-2.13	5	3	**CA_C2255**	Predicted permease	-2.17	1	24
CA_C1356	Thiamine biosynthesis enzyme ThiH	1.27	4	12		Unknown			
CA_C0896	Chorismate synthase	-1.86	6	8	**CA_C3725**	Uncharacterized protein	-1.80	9	6
CA_C0797	Phosphoenolpyruvate synthase	-2.07	1	5	**CA_C3721**	Hypothetical secreted protein	1.68	8	31
CA_C0094	Ferredoxin-nitrite reductase	-1.33	21	17	**CA_C3720**	Uncharacterized protein	2.02	11	3
	Nucleoside/nucleotide metabolism				**CA_C3708**	Uncharacterized protein	-2.68	5	4
CA_C3627	7-cyano-7-deazaguanine synthase	-1.97	5	15	**CA_C3707**	Uncharacterized protein	-2.66	14	5
CA_C3602	HD superfamily hydrolase	-1.87	10	8	**CA_C3703**	Uncharacterized protein	-3.43	3	6
CA_C3593	Adenylosuccinate synthetase (AMPSase) (AdSS)	-1.91	11	11	**CA_C3592**	Uncharacterized protein	-2.16	15	8
CA_C3224	PUR operon repressor, Adenine/guanine phosphoribosyltransferase family	-1.38	2	6	**CA_C3284**	DegV domain-containing protein	-1.75	4	12
CA_C3112	Adenylate kinase	1.38	43	40	**CA_C3248**	Uncharacterized protein	1.34	13	11
CA_C2892	CTP synthase	-1.77	4	3	**CA_C3212**	Fusion of Uroporphyrinogen-III methylase	1.47	4	9
CA_C2879	Uracil phosphoribosyltransferase	-1.19	24	22	**CA_C3094**	Uncharacterized consrved protein, associated with phosphate permease	-2.07	9	4
CA_C2701	Inosine-5'-monophosphate dehydrogenase	1.51	26	44	**CA_C2758**	Uncharacterized protein, YPUA B.subtilis ortholog	1.43	5	6
CA_C2275	Adenine phosphoribosyltransferase	1.98	12	4	**CA_C2723**	Deacethylase/dipeptidase/desuccinylase family of Zn-dependent hydrolases	1.93	14	35
CA_C2064	Purine nucleoside phosphorylase	-1.32	3	5	**CA_C2655**	Uncharacterized membrane-associated protein, DedA family	1.60	6	7
CA_C1963	5'-nucleotidase/2',3'-cyclic phosphodiesterase related enzyme	-1.86	1	14	**CA_C2643**	Uncharacterized protein	-1.26	3	3
CA_C1821	Adenylosuccinate lyase	-1.63	1	7	**CA_C2629**	Hypothetical secreted protein	-1.90	3	4
CA_C1789	Uridylate kinase (UK)	-2.05	12	3	**CA_C2611**	Uncharacterized protein	1.40	3	6
CA_C1395	Bifunctional purine biosynthesis protein PurH	-1.14	16	34	**CA_C2564**	Uncharacterized protein	1.52	26	44
CA_C1390	N5-carboxyaminoimidazole ribonucleotide mutase	-2.40	9	8	**CA_C2387**	Uncharacterized protein	1.41	39	12
CA_C0480	Oxygen-sensitive ribonucleoside-triphosphate reductase nrdD	2.06	6	15	**CA_C2364**	Uncharacterized protein	-1.87	3	11
CA_C0027	Orotate phosphoribosyltransferase	1.71	3	3	**CA_C2251**	Uncharacterized conserved membrane protein	-1.69	1	5
	Protein biosinthesis/modification				**CA_C1945**	Phage related anti-repressor protein	-1.57	9	8
CA_C2703	60 kDa chaperonin	-2.06	54	316	**CA_C1892**	Uncharacterized protein	-2.77	11	16
CA_C2704	10 kDa chaperonin	-2.41	56	24	**CA_C1886**	Uncharacterized phage related protein	-3.09	7	3
CA_C3260	Asparagine--tRNA ligase	1.35	11	28	**CA_C1868**	Uncharacterized secreted protein	-2.55	4	4
CA_C3201	Formate--tetrahydrofolate ligase	1.36	12	23	**CA_C1849**	Predicted flavoprotein	-1.28	2	3
CA_C3197	Lysine--tRNA ligase	1.60	11	21	**CA_C1817**	Stage V sporulation protein	-1.22	52	26
CA_C2991	Methionine--tRNA ligase	-1.43	3	3	**CA_C1629**	Putative intracellular protease/amidase	-1.75	20	9
CA_C2830	Acylphosphatase	-2.24	14	10	**CA_C1306**	Uncharacterized protein	1.49	5	3
CA_C2399	Valine--tRNA ligase	-1.43	3	19	**CA_C1249**	Site-determining protein	-1.99	13	8
CA_C2370	33 kDa chaperonin	-1.90	3	4	**CA_C0660**	Uncharacterized protein	1.63	10	4
CA_C2362	Threonine--tRNA ligase	-1.72	4	26	**CA_C0556**	Uncharacterised conserved protein	-1.39	6	9
CA_C2356	Phenylalanine--tRNA ligase beta subunit	1.75	3	12	**CA_C0334**	Uncharacterized protein	2.67	29	3
CA_C2269	Aspartate--tRNA ligase	-1.28	10	17	**CA_C0034**	Uncharacterized protein	-2.00	38	12

CA_C2264	Serine hydroxymethyltransferase	1.71	10	20	**CA_P0012**	Pediocin immunity protein	-1.39	12	11
CA_C0990	Glutamate--tRNA ligase	1.25	17	24	**CA_C2385**	Uncharacterized protein	1.77	5	3
CA_C0646	Leucine--tRNA ligase	1.34	2	12	**CA_C2134**	Predicted GTPase	1.40	10	17
	Protein fate								
CA_C3716	Lon-like ATP-dependent protease	-2.25	1	9					

Appendix 4.3 Proteins identiefied with significant changes in comparision of ExpC/StaC

Locus ID	Protein	Fold change	% Coverage	Peptides	Locus ID	Protein	Fold change	% Coverage	Peptides
	Alcohol metabolism					Signal transduction/chemotaxis/secretion/ trafficing			
CA_P0165	Acetoacetate decarboxylase	1.36	45	26	**CA_C3734**	tRNA modification GTPase	-1.21	2	8
CA_C3392	NADH-dependent butanol dehydrogenase	-3.05	9	11	**CA_C2942**	S-ribosylhomocysteine lyase	1.20	17	4
	Amino acid biosynthesis				**CA_C2433**	HtrA-like serine protease	1.35	21	30
CA_C3254	Gamma-glutamyl phosphate reductase	1.44	5	9	**CA_C2220**	Chemotaxis histidine kinase, CheA (Contains CheW-like adaptor domain)	-1.30	10	20
CA_C3600	4-hydroxy-tetrahydrodipicolinate synthase 2	-1.17	7	3	**CA_C2161**	Flagellar motor switch protein FliG	1.35	4	15
CA_C2973	2-keto-3-deoxy-6-phosphogluconate aldolase	-1.12	35	15	**CA_C0585**	N-terminal CheY reciever domain fused to C-terminal uncharacterized CheX-like domain	-1.40	3	5
CA_C2658	Glutamine synthetase type III	1.20	12	46		Stress responce			
CA_C2235	Cysteine synthase	1.69	30	14	**CA_C3598**	Reverse rubrerythrin-1 (revRbr 1) (NADH peroxidase) (NPXase) (Npx)	-1.20	20	19
CA_C1684	TYPA/BIPA type GTPase	-1.22	8	12	**CA_C2637**	Lon protease	-1.39	4	34
CA_C0568	Aspartate semialdehyde dehydrogenase	-1.25	11	14	**CA_C1549**	Glutathione peroxidase	-1.23	10	3
	Carbohydrate metabolism					Transcription			
CA_C0517	6-phosphofructokinase	1.42	10	19	**CA_C3472**	Protein containing transcriptional regulator domain	-1.29	2	15
CA_C0518	Pyruvate kinase	1.34	11	43	**CA_C3198**	Transcription elongation factor GreA (Transcript cleavage factor GreA)	1.20	18	11
CA_C3376	Possible pectin degradation protein	-4.46	10	3	**CA_C3142**	DNA-directed RNA polymerase subunit beta' (RNAP subunit beta')	1.29	9	50
CA_C2680	Glucose-6-phosphate isomerase	1.93	24	18	**CA_C2889**	Transcription termination factor Rho	1.24	5	19
CA_C2613	Transcriptional regulators of NagC/XylR family	1.43	9	8	**CA_C2295**	Probable transcriptional regulatory protein	2.09	6	10
CA_C2337	Phosphomannomutase	-1.58	14	47	**CA_C1719**	DNA-directed RNA polymerase subunit omega (RNAP omega subunit)	1.38	43	5
CA_C1408	Phospho-beta-glucosidase	1.54	2	9	**CA_C0807**	Cold shock protein	1.78	46	20
CA_C1287	HIT family hydrolase	2.77	16	3	**CA_C0461**	Mercuric resistance operon regulatory protein, MerR family	1.52	5	5
CA_C0713	Enolase	1.26	32	24		Translation			
CA_C0712	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	1.52	15	22	**CA_C1722**	Peptide deformylase 1 (PDF 1)	1.46	9	4
	Cell cycle/cell division				**CA_C1723**	Methionyl-tRNA formyltransferase	1.33	7	19
CA_C3202	ATP-dependent zinc metalloprotease FtsH	-3.23	7	20	**CA_C3147**	50S ribosomal protein L1	1.21	35	37
CA_C2120	Cell division protein SepF	9.76	4	5	**CA_C3145**	50S ribosomal protein L7/L12	-1.08	46	294

CA_C1013	FTSA related protein, predicted ATPases of the HSP70 family	1.43	1	7	**CA_C3140**	30S ribosomal protein S12	-1.17	10	20
CA_C2118	Cell division protein DivIVA	1.98	8	13	**CA_C3138**	Elongation factor G (EF-G)	1.37	24	64
CA_C0498	Cell division protein	-1.38	6	7	**CA_C3133**	50S ribosomal protein L3	1.20	23	31
	Cell envalope/cell wall biogenesis				**CA_C3132**	50S ribosomal protein L4	1.44	9	26
CA_C3194	UDP-N-acetylmuramoylalanine--D-glutamate ligase	-1.29	4	3	**CA_C3124**	30S ribosomal protein S17	1.16	19	76
CA_C2335	UTP-glucose-1-phosphate uridylyltransferase	-1.44	22	17	**CA_C3123**	50S ribosomal protein L14	1.18	67	17
CA_C2333	Glucose-1-phosphate thymidylyltransferase	-1.47	16	10	**CA_C3106**	30S ribosomal protein S11	1.40	21	7
CA_C2332	dTDP-glucose 4,6-dehydratase	-1.33	7	13	**CA_C3105**	30S ribosomal protein S4 A	1.55	33	26
	Cofactor biosynthesis				**CA_C3146**	50S ribosomal protein L10	1.25	48	43
CA_C0594	Pyridoxal biosynthesis lyase PdxS	1.27	31	22	**CA_C2888**	50S ribosomal protein L31	1.49	19	8
	DNA metabolism				**CA_C2669**	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B 1	1.67	7	10
CA_C0006	DNA gyrase subunit B	-1.43	4	9	**CA_C2361**	Translation initiation factor IF-3	-1.70	7	12
CA_C0007	DNA gyrase subunit A	-1.21	11	57	**CA_C2360**	50S ribosomal protein L35	-1.63	19	7
CA_C3729	Stage 0 sporulation J, ParB family of DNA-binding proteins	-1.23	6	15	**CA_C2359**	50S ribosomal protein L20	-1.55	32	46
CA_C3036	Superfamily I DNA helicase	-1.17	0	25	**CA_C1802**	Translation initiation factor IF-2	1.33	3	14
CA_C2382	Single-strand DNA-binding protein	1.59	35	21	**CA_C1790**	Ribosome-recycling factor (RRF)	-1.42	14	21
CA_C0127	Recombination protein	-1.56	4	9	**CA_C1788**	Elongation factor Ts (EF-Ts)	-1.31	38	60
	Energy metabolism/electron transport				**CA_C1787**	30S ribosomal protein S2	1.36	40	32
CA_C2709	Electron transfer flavoprotein subunit alpha (Alpha-ETF)	-1.08	41	171	**CA_C1259**	50S ribosomal protein L27	1.20	42	65
	Fatty acid biosynthesis/lipid metabolism					Transportation and binding			
thlB	Thiolase B	1.28	12	92	**CA_C2869**	ATP synthase subunit b (ATP synthase F(0) sector subunit b)	1.39	25	18
CA_C2708	3-hydroxybutyryl-CoA dehydrogenase	-1.11	43	340	**CA_C3632**	Oligopeptide ABC transporter, periplasmic substrate-binding component	-1.20	8	20
CA_C2711	Acyl-CoA dehydrogenase, short-chain specific	-1.15	29	20	**CA_C3354**	Probable cation efflux pump	-1.10	1	37
CA_C2712	3-hydroxybutyryl-CoA dehydratase	-1.16	53	55	**CA_C3282**	ABC-type multidrug/protein/lipid transport system, ATPase component	-1.20	4	7
CA_C3076	Phosphate butyryltransferase	-1.25	47	47	**CA_C3012**	ATPase component of ABC transporter, with duplicated ATPase domains	1.79	1	12
CA_C3575	Malonyl CoA-acyl carrier protein transacylase	-1.19	15	9	**CA_C1816**	Ribonuclease Y	-1.18	3	20
CA_C3572	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	-1.18	4	9	**CA_C1590**	2-oxoglutarate/malate translocator	1.50	2	3
	Intermediary metabolism/multibiosynthetic pathway				**CA_C1353**	Phosphotransferase system IIC component	1.35	2	5
CA_C2873	Acetyl-CoA acetyltransferase	-1.09	40	149	**CA_C0570**	PTS enzyme II, ABC component	-1.23	1	8
CA_C1743	Acetate kinase	-1.40	10	5	**CA_C0147**	ABC transporter, ATP-binding protein	-1.55	24	31
hydA	Hydrogenase I	1.13	2	13	**CA_C0146**	Related to ABC transporter permease component	-1.61	2	5
CA_C3371	2-enoate reductase	-3.02	2	10	**CA_C0139**	Predicted permease	-1.44	2	10
CA_C3090	Fumarate hydratase	-1.95	5	5	**CA_C2255**	Predicted permease	1.16	1	24
CA_C2935	Predicted acetyltransferase	1.12	7	4	**CA_C1705**	Periplasmic phosphate-binding protein	2.99	2	6
CA_C2584	Protein containing ChW-repeats	1.46	2	5	**CA_C0880**	Periplasmic amino acid binding protein	1.45	5	5
CA_C2572	Possible aminoglycoside phosphotransferase	-4.48	2	16	**CA_C0111**	Glutamine-binding periplasmic protein fused to glutamine permease	1.74	2	7
CA_C2283	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	1.30	13	15		Unknown function			
CA_C1819	Aspartate Aminotransferase	-4.63	5	5	**CA_C3537**	Fragment of SECA	-1.26	11	8
CA_C0944	Transketolase	1.94	5	5	**CA_C3341**	Multimeric flavodoxin WrbA family protein	1.36	26	19

CA_C0097	Porphobilinogen deaminase	2.18	4	4	**CA_C3264**	Uncharacterized conserved protein, YTFJ B.subtilis ortholog	1.40	23	9
	Nucleoside/nucleotide metabolism				**CA_C3248**	Uncharacterized protein	1.34	13	11
CA_C3627	7-cyano-7-deazaguanine synthase	-1.22	5	15	**CA_C3094**	Uncharacterized consrved protein	-1.96	9	4
CA_C3224	PUR operon repressor, Adenine/guanine phosphoribosyltransferase family	-1.18	2	6	**CA_C3008**	CBS domain containing protein	1.41	8	34
CA_C2879	Uracil phosphoribosyltransferase	1.28	24	22	**CA_C2853**	Uncharacterized protein	1.40	2	6
CA_C1395	Bifunctional purine biosynthesis protein PurH(Inosinicase)	1.26	16	34	**CA_C2817**	Predicted membrane protein	1.73	2	11
CA_C0480	Oxygen-sensitive ribonucleoside-triphosphate reductase nrdD	-1.38	6	15	**CA_C2643**	Uncharacterized protein	-1.29	3	3
	Protein biosynthesis/modification				**CA_C2564**	Uncharacterized protein	-1.27	26	44
CA_C3260	Asparagine--tRNA ligase	1.52	11	28	**CA_C2387**	Uncharacterized protein	1.34	39	12
CA_C3195	Glycine--tRNA ligase	-1.21	4	13	**CA_C2364**	Uncharacterized protein	-1.73	3	11
CA_C3189	ATPases with chaperone activity clpC, two ATP-binding domai	-1.23	6	11	**CA_C2251**	Uncharacterized conserved membrane protein	1.38	1	5
CA_C3038	Isoleucine--tRNA ligase	1.35	3	15	**CA_C2193**	Ucharacterized protein, CGEB homolog	1.79	2	27
CA_C2847	Ribosome-associated protein Y	1.28	12	39	**CA_C2085**	Uncharacterized protein from alkaline shock protein family	-1.69	30	10
CA_C2740	Histidine--tRNA ligase	-1.59	4	8	**CA_C1880**	Uncharacterized protein	1.56	13	3
CA_C2370	33 kDa chaperonin	-1.62	3	4	**CA_C1828**	TldD protein	1.30	2	7
CA_C2362	Threonine--tRNA ligase	-1.57	4	26	**CA_C1817**	Stage V sporulation protein, spoVS	-1.29	52	26
CA_C2357	Phenylalanine--tRNA ligase alpha subunit	-1.59	6	10	**CA_C1717**	UPF0296 protein	1.19	37	6
CA_C2356	Phenylalanine--tRNA ligase beta subunit	-1.48	3	12	**CA_C1679**	UPF0297 protein	-1.79	20	9
CA_C2269	Aspartate--tRNA ligase	-1.29	10	17	**CA_C1306**	Uncharacterized protein	1.21	5	3
CA_C2264	Serine hydroxymethyltransferase	1.23	10	20	**CA_C1304**	Uncharacterized conserved protein, predicted metal-binding	1.70	4	4
CA_C1297	N-terminal of elongation factor Ts	1.59	4	8	**CA_C1242**	MreB	1.37	10	13
CA_C0990	Glutamate-tRNA ligase	1.36	17	24	**CA_C1171**	Uncharacterized protein	-1.81	6	3
CA_C0646	Leucine-tRNA ligase	1.15	2	12	**CA_C0504**	FHA-domain containing secreted protein	1.32	4	36
	Protein fate								
CA_C2846	Protein translocase subunit SecA	-1.44	9	48					
CA_C2278	Protein translocase subunit SecD	1.25	2	4					
CA_C1052	Membrane protease subunit, stomatin/prohibitin homolog	1.63	19	16					

Appendix 4.4 Proteins identiefied with significant changes in comparision of StaC/StaCL

Locus ID	**Protein name**	**Fold change**	**% coverage**	**Peptides**	**Locus ID**	**Protein name**	**Fold change**	**% coverage**	**Peptides**
	Alcohol metabolism (solventogenesis)					Signal transduction/chemotaxis/secretion/trafficing			
CA_P0165	Acetoacetate decarboxylase	-2.18	45	26	CA_C3734	tRNA modification GTPase MnmE	-1.37	2	8
CA_C3392	NADH-dependent butanol dehydrogenase	3.79	9	11	CA_C2942	S-ribosylhomocysteine lyase	1.50	17	4

	Amino acid biosynthesis				CA_C2220	Chemotaxis histidine kinase, CheA	2.60	10	20
CA_C3254	Gamma-glutamyl phosphate reductase	-1.10	5	9		Sporulation/genrmination			
CA_C3171	3-isopropylmalate dehydrogenase	-2.89	4	3	CA_C3223	Putative septation protein SpoVG	-1.27	45	25
CA_C1684	TYPA/BIPA type GTPase	-2.09	8	12		Stress responce			
CA_C1001	Aspartate aminotransferase	-1.66	3	3	CA_C3714	18 kDa heat shock protein	-1.31	29	22
	Carbohydrate metabolism(sugar, polysachharide)				CA_C2637	Lon protease	-1.22	4	34
CA_C0517	6-phosphofructokinase	-1.68	10	19		Transcription			
CA_C0518	Pyruvate kinase	-1.57	11	43	CA_C1300	RNA polymerase sigma factor RpoD	-1.87	16	19
CA_C3376	Possible pectin degradation protein	5.34	10	3	CA_C3283	Transcriptional regulator, MarR/EmrR family	-1.36	15	9
CA_C3032	Galactose mutarotase related enzyme	1.50	4	4	CA_C2990	Cold shock protein	-1.54	25	15
CA_C2523	Glycosyltransferase	-1.62	1	4	CA_C2939	Response regulator	1.61	3	6
CA_C2337	Phosphomannomutase	2.59	14	47	CA_C2842	Transcription accessory protein TEX	-3.86	6	13
CA_C1036	Pyruvate kinase	-1.70	13	20	CA_C2430	Transcription elongation factor	1.54	10	6
CA_C0743	6-phospho-beta-glucosidase	-1.34	15	18	CA_C2084	N utilization substance protein B homolog	-1.26	7	6
CA_C0484	Phosphoglucosamine mutase	1.56	8	12	CA_C1843	Predicted transcriptional regulator, YDCN B.subtilis ortholog	-1.59	5	3
	Cell cycle/cell division				CA_C1838	(Dimethylallyl)adenosine tRNA methylthiotransferase MiaB	-1.55	6	12
CA_C2641	Trigger factor	1.26	34	89		Translation			
CA_C1875	Uncharacterized protein	-1.64	12	6	CA_C1722	Peptide deformylase 1	-2.08	9	4
	Cell envalope/cell wall biogenesis				CA_C1723	Methionyl-tRNA formyltransferase	-1.62	7	19
CA_C2335	UTP-glucose-1-phosphate uridylyltransferase	2.56	22	17	CA_C3722	30S ribosomal protein S18	-2.26	27	8
CA_C2333	Glucose-1-phosphate thymidylyltransferase	2.45	16	10	CA_C3717	50S ribosomal protein L9	2.55	42	34
CA_C2332	dTDP-glucose 4,6-dehydratase	1.88	7	13	CA_C3140	30S ribosomal protein S12	-1.46	10	20
	DNA metabolism (replication, recombination, repair)				CA_C3139	30S ribosomal protein S7	-1.18	40	32
CA_C1283	Chaperone protein DnaJ	-1.52	2	9	CA_C3138	Elongation factor G (EF-G)	-1.30	24	64
CA_C3735	Predicted RNA-binding protein Jag, SpoIIIJ-associated	-1.39	7	7	CA_C3132	50S ribosomal protein L4	1.24	9	26
CA_C3729	Stage 0 sporulation J, ParB family of DNA-binding proteins	-1.41	6	15	CA_C3131	50S ribosomal protein L23	-1.32	78	100
CA_C3723	Single-stranded DNA-binding protein 3	-1.46	9	11	CA_C3106	30S ribosomal protein S11	-1.54	21	7
CA_C3211	DNA binding protein HU	-1.28	50	197	CA_C3105	30S ribosomal protein S4 A	-1.18	33	26
CA_C3036	Superfamily I DNA helicase	1.37	0	25	CA_C2888	50S ribosomal protein L31	1.60	19	8
CA_C2707	8-oxoguanine-DNA-glycosylase	1.72	8	8	CA_C1790	Ribosome-recycling factor	-1.79	14	21
CA_C1785	DNA topoisomerase	-1.48	1	12	CA_C1788	Elongation factor Ts	-1.46	38	60
CA_C1092	Predicted metal-dependent phosphoesterase (PHP family), YciV ortholog	-1.92	8	3	CA_C3146	50S ribosomal protein L10	-1.33	48	43

CA_C0127	Recombination protein RecR	2.39	4	9	CA_C1259	50S ribosomal protein L27	-1.27	42	65
CA_C0126	Nucleoid-associated protein	1.73	15	13		Transport and binding			
	Energy metabolism/electron transport				CA_C2869	ATP synthase subunit b	-1.70	25	18
CA_C2710	Electron transfer flavoprotein subunit beta	-1.67	54	212	CA_C2870	ATP synthase subunit c	-1.84	9	8
	Fatty acid biosynthesis/lipid metabolism				CA_C3354	Probable cation efflux pump	-1.20	1	37
thlB	Thiolase B	-1.81	12	92	CA_C3282	ABC-type multidrug/protein/lipid transport system, ATPase component	-1.78	4	7
CA_C2708	3-hydroxybutyryl-CoA dehydrogenase	-1.17	43	340	CA_C3012	ATPase component of ABC transporter, with duplicated ATPase domains	1.48	1	12
CA_C3571	3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	-1.38	22	6	CA_C1353	Phosphotransferase system IIC component, possibly N-acetylglucosamine-specific	1.91	2	5
CA_C3570	Biotin carboxylase	-1.55	5	8	CA_C0386	PTS cellobiose-specific component IIC	-1.30	6	11
CA_C3569	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	-1.34	9	10	CA_C0268	ABC transporter ATP-binding protein	-1.47	3	5
CA_C0489	Holo-[acyl-carrier-protein] synthase	1.78	6	4	CA_C0147	ABC transporter, ATP-binding protein	2.35	24	31
	intermediary metabolism/other metabolic pathways/multibiosynthetic pathways				CA_C0146	Related to ABC transporter permease component	2.88	2	5
CA_C2873	Acetyl-CoA acetyltransferase	-1.22	40	149	CA_C0139	Predicted permease	2.05	2	10
CA_C3371	2-enoate reductase (Two distinct NAD(FD)-dependent dehydrogenase domains)	5.07	2	10	CA_C1705	Periplasmic phosphate-binding protein	2.96	2	6
CA_C2935	Predicted acetyltransferase	1.60	7	4		Unknown			
CA_C2572	Possible aminoglycoside phosphotransferase (Protein kinase related), diverged	2.53	2	16	CA_C3721	Hypothetical secreted protein	2.62	8	31
CA_C2542	FAD/FMN-containing dehydrogenase	-2.16	4	14	CA_C3708	Uncharacterized protein	-1.66	5	4
CA_C1847	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	-1.21	7	28	CA_C3707	Uncharacterized protein, homolog of Bacillus firmus	-1.63	14	5
CA_C1356	Thiamine biosynthesis enzyme ThiH	1.47	4	12	CA_C3703	Uncharacterized protein	-1.74	3	6
CA_C0797	Phosphoenolpyruvate synthase	-1.18	1	5	CA_C3537	Fragment of SECA	-1.37	11	8
	Nucleoside/nucleotide metabolism				CA_C3341	Multimeric flavodoxin WrbA family protein	-1.34	26	19
CA_C3627	7-cyano-7-deazaguanine synthase	-1.20	5	15	CA_C3284	DegV domain-containing protein	-1.46	4	12
CA_C2064	Purine nucleoside phosphorylase	1.73	3	5	CA_C3008	CBS domain containing protein	-1.32	8	34
CA_C1546	Pyrimidine-nucleoside phosphorylase	1.60	6	10	CA_C2655	Uncharacterized membrane-associated protein, DedA family	2.92	6	7

CA_C0480	Oxygen-sensitive ribonucleoside-triphosphate reductase nrdD	2.04	6	15	CA_C2643	Uncharacterized protein	2.14	3	3
	Protein biosinthesis/modification				CA_C1945	Phage related anti-repressor protein	-1.61	9	8
CA_C2703	60 kDa chaperonin	-1.50	54	316	CA_C1892	Uncharacterized protein	-1.40	11	16
CA_C3178	Proline--tRNA ligase	-1.71	5	16	CA_C1868	Uncharacterized secreted protein, homolog YXKC Bacillus subtilis	-1.96	4	4
CA_C3177	Cysteine--tRNA ligase	-2.72	3	4	CA_C1828	TldD protein	2.24	2	7
CA_C3154	RRNA methylase, YACO B.subtilis ortholog	-2.53	3	4	CA_C1735	Predicted kinase related to hydroxyacetone kinase, YLOV ortholog	2.02	6	18
CA_C2830	Acylphosphatase	-1.49	14	10	CA_C1679	UPF0297 protein	-2.04	20	9
CA_C2356	Phenylalanine--tRNA ligase beta subunit	1.77	3	12	CA_C1635	Predicted nucleic acid binding protein, containing 2 S1 domains, YITL B.subtilis ortholog	-2.28	7	6
CA_C2264	Serine hydroxymethyltransferase	1.71	10	20	CA_C1629	Putative intracellular protease/amidase, ThiJ family	-2.14	20	9
CA_C2094	Elongation factor P	-2.26	20	10	CA_C1510	Uncharacterized protein	-1.98	2	5
CA_C1678	Alanine--tRNA ligase	-1.57	2	11	CA_C1306	Uncharacterized protein	1.96	5	3
	Protein fate				CA_C1249	Site-determining protein	-3.18	13	8
CA_C3716	Lon-like ATP-dependent protease	-1.50	1	9	CA_C1182	Phage related protein, YorG B.subtilis homolog	1.55	6	8
CA_C3674	Two CBS domain containing protein	-1.72	20	11	CA_C0660	Uncharacterized protein	2.02	10	4
CA_C3006	Zn-dependent peptidase, insulinase family	-1.34	3	19	CA_C0635	Zinc finger domain	-1.36	38	27
CA_C1760	Signal peptidase I	-1.43	3	6					

Appendix 5.1 List of membrane associated proteins involved in various membrane activities in *F. succinogenes* S85

Locus ID	Protein description	Glucose	MC cellulose	AS cellulose	Location[a]	Gravy index[b]	Molecular mass (kDa)[b]	pI[b]	Presence of signal peptide (amino acid position)[c]	Reference
Fisuc_0344 FSU_0758	TonB family protein	√	-	-	Non cytoplasmic	-0.447	53.371	8.84	Yes (20-21)	-
Fisuc_0552 FSU_0976	Periplasmic solute binding protein	√	-	-	Unknown	-0.271	32.87	9.39	No	-

Fisuc_0663 FSU_1094	Toluene tolerance family protein	√	-	-	Non cytoplasmic	-0.449	21.87	9.22	Yes (18-19)	-
Fisuc_2918 FSU_0181	Capsular exopolysaccharide family	√	-	-	Cytoplasmic membrane	-0.165	80.26	7.97	No	-
Fisuc_0891 FSU_1339	Cell wall/surface repeat protein	√	-	-	Non cytoplasmic	-0.229	94.84	5.25	Yes (24-25)	-
Fisuc_2182 FSU_2721	Lauroyl/myristoyl acyltransferase-like protein	√	-	-	Unknown	-0.225	33.849	9.79	No	-
dtd Fisuc_0587 FSU_1014	D-tyrosyl-tRNA(Tyr) deacylase	√	-	-	Unknown	-0.054	15.934	5.53	No	-
surA Fisuc_0518 FSU_0941	Peptidyl-prolyl cis-trans isomerase SurA	√	-	-	Periplasm	-0.31	48.24	6.63	Yes (19-20)	-
Fisuc_0663 FSU_1094	Toluene tolerance family protein	√	-	-	Non cytoplasmic	-0.449	21.87	9.22	Yes (18-19)	-
Fisuc_2767 FSU_0019	ADP-ribosylation/Crystallin J1	√	-	-	Unknown	-0.228	29.82	4.99	No	-
FSU_1456	Diguanylate cyclase (GGDEF) domain protein	√	-	-	Cytoplasmic membrane	0.325	46.51	8.72	No	-
FSU_1566	Putative 1-acyl-sn-glycerol-3-phosphate acetyltransferase	√	-	-	Cytoplasmic membrane	0.026	28.6	9.3	No	-
Fisuc_1756 FSU_2256	Peptidyl-prolyl cis-trans isomerase, FKBP-type	√	-	-	Periplasm	-0.115	42.84	6.36	Yes (18-19)	-
Fisuc_1795 FSU_2295	Histidine kinase	√	-	-	Cytoplasmic membrane	-0.312	139.27	5.41	No	-
Fisuc_1518 FSU_2000	Antigen-like protein	√	-	-	Non cytoplasmic	-0.192	86.54	5.43	Yes (19-20)	-
Fisuc_1108	Phospholipid/glycerol acyltransferase	√	-	-	Cytoplasmic membrane	0.047	28.62	9.3	No	-
FSU_2398	TPR domain protein	√	√	√	Non cytpplasmic	-0.617	146.67	8.5	Yes (20-21)	[360]
FSU_2396	OmpA family protein	√	√	√	Outer membrane	-0.408	55.97	4.86	Yes (28-29)	[360]
Fisuc_1892 FSU_2397	TPR domain protein	√	√	√	Non cytpplasmic	-0.341	83.84	5.51	Yes (23-24)	[360]
Fisuc_1891	OmpA/MotB domain protein	√	√	√	Unknown	-0.458	53.24	4.74	No	-
Fisuc_1893	TPR repeat-containing protein	√	√	√	Unknown	-0.619	146.7	8.5	Yes (20-21)	-
Fisuc_1894 FSU_2400	MotA/TolQ/ExbB proton channel	√	√	√	Cytoplasmic membrane	0.461	23.06	9.14	No	-
Fisuc_1980	Sporulation domain protein	√	√	√	Unknown	-0.476	13.35	9.56	No	-
mscL Fisuc_2074 FSU_2602	Large-conductance mechanosensitive channel	√	√	√	Cytoplasmic membrane	0.462	15.94	9.21	No	-
Fisuc_0201	MotA/TolQ/ExbB proton channel	√	√	√	Cytoplasmic membrane	0.319	24.61	7.78	No	-
Fisuc_2503 FSU_3071	Extracellular solute-binding protein family 3	√	√	√	periplasmic	-0.132	28.75	5.47	Yes (21-22)	-
Fisuc_2249 FSU_2794	Fibrobacter succinogenes major paralogous domain protein	√	√	√	Non cytoplasmic	-0.387	70.12	4.57	No	-
Fisuc_0289 FSU_0701	Efflux transporter, RND family, MFP subunit	√	√	√	Non cytoplasmic	-0.115	37.256	9.45	No	-
Fisuc_0299 FSU_0711	Tetratricopeptide repeat protein	√	√	√	Periplasm	-0.656	49.015	9.14	Yes (29-30)	-
Fisuc_0331	Pentapeptide repeat protein	√	√	√	Extracellular	-0.401	47.89	9	Yes (23-24)	-

secG Fisuc_2367 FSU_2921	Preprotein translocase, SecG subunit	√	√	√	Cytoplasmic membrane	0.449	16.955	8.73	No	-
Fisuc_0369	WD40 domain protein beta Propeller	√	√	√	Non cytoplasmic	-0.373	44.076	7.74	Yes (21-22)	-
Fisuc_2509 FSU_3077	OmpA family protein	√	√	√	Outer membrane	-0.381	83.917	4.75	Yes (19-20)	-
Fisuc_2762	PpiC-type peptidyl-prolyl cis-trans isomerase	√	√	√	Outer membrane	-0.291	71.067	5.11	No	-
Fisuc_0884 FSU_1330	Secretion protein HlyD family protein	√	√	√	Cytoplasmic membrane	-0.189	36.175	8.8	No	-
Fisuc_0775 FSU_1216	Polysaccharide biosynthesis/export protein	√	√	√	Out ermembrane	-0.181	41.49	5.74	Yes (21-22)	-
secDF Fisuc_0858 FSU_1302	Protein-export membrane protein SecD	√	√	√	Cytoplasmic membrane	0.285	93.186	8.96	No	-
Fisuc_2892	OmpA/MotB domain protein	√	√	√	Outer membrane	-0.36	32.21	5.53	Yes (19-20)	-
Fisuc_2917 FSU_0180	OmpA family protein	√	√	√	Outer membrane	-0.272	70.79	5.27	Yes (17-18)	-
Fisuc_0978	Capsular exopolysaccharide family	√	√	√	Cytoplasmic membrane	-0.242	77.96	8.94	No	-
Fisuc_1192 FSU_1653	Periplasmic solute binding protein	√	√	√	Cytoplasmic membrane	-0.125	36.391	5.03	Yes (20-21)	-
Fisuc_1226 FSU_1687	ABC transporter related protein	√	√	√	Cytoplasmic membrane	-0.323	30.503	8.6	No	-
Fisuc_1230 FSU_1691	Extracellular solute-binding protein family 5	√	√	√	Unknown	-0.336	67.554	5.69	Yes (21-22)	-
Fisuc_1591 FSU_2077	Capsular exopolysaccharide family	√	√	√	Cytoplasmic membrane	-0.092	78.322	8.14	No	-
Fisuc_1592	OmpA/MotB domain protein	√	√	√	Outer membrane	-0.386	73.841	4.72	Yes (17-18)	-
FSU_0151	OmpA family protein	√	√	√	Outer membrane	-0.566	23.937	6.65	No	-
Fisuc_2987 FSU_0252	Ankyrin	√	√	√	Non cytoplasmic	-0.252	26.75	8.97	Yes (19-20)	-
Fisuc_1980	Sporulation domain protein	√	√	√	Unknown	-0.476	13.35	9.56	No	-
groEL groL Fisuc_0061 FSU_0456	60 kDa chaperonin	√	√	√	Unknown	-0.128	57.52	5.41	No	-
mrcA Fisuc_0062 FSU_0457	Penicillin-binding protein 1A	√	√	√	Unknown	-0.394	90.22	9.33	No	-
gapA Fisuc_0102 FSU_0503	Glyceraldehyde 3-phosphate dehydrogenase A	√	√	√	Unknown	-0.03	36.81	6.14	No	-
nuoB Fisuc_2126 FSU_2661	NADH-quinone oxidoreductase subunit B	√	√	√	Cytoplasmic membrane	-283	23.97	6.62	No	-
Fisuc_2131	NADH dehydrogenase (Quinone)	√	√	√	Unknown	-0.255	57.9	6.83	No	-
frdC Fisuc_2492 FSU_3060	Succinate dehydrogenase (Or fumarate reductase) cytochrome b subunit, b558 family	√	√	√	Cytoplasmic membrane	0.404	31.59	9.04	No	-
Fisuc_2502 FSU_3070	4Fe-4S ferredoxin iron-sulfur binding domain protein	√	√	√	Cytoplasmic membrane	-0.115	27.52	6.52	No	-
fdrB Fisuc_2494 FSU_3062	4Fe-4S ferredoxin iron-sulfur binding domain protein	√	√	√	Cytoplasmic membrane	-0.102	28	6.52	No	-

Fisuc_2732 FSU_3303	Rhodanese domain protein	√	√	√	Non cytoplasmic	0.033	15.785	8.81	Yes (20-21)	-
Fisuc_0872 FSU_1318	Peptidyl-prolyl cis-trans isomerase	√	√	√	Outer membrane	-0.451	29.67	7.65	No	-
gdhA Fisuc_2811 FSU_0066	Glu/Leu/Phe/Val dehydrogenase	√	√	√	Unknown (multiple location)	-0.235	48.622	6.86	No	-
rplM Fisuc_0975 FSU_1421	50S ribosomal protein L13	√	√	√	Periplasm	-0.113	15.6	9.8	No	-
Fisuc_1010 FSU_1457	BatA protein	√	√	√	Cytoplasmic membrane	0.013	40.98	9.18	No	-
nifJ Fisuc_2881 FSU_0139	Pyruvate-flavodoxin oxidoreductase	√	√	√	Unknown	-0.138	129.294	7.23	No	-
Fisuc_2905 FSU_0167	FKBP-type peptidyl-prolyl cis-trans isomerase domain protein/thiol-disulfide oxidoreductase domain protein	√	√	√	Periplasm	-0.496	48.48	8.83	Yes (21-22)	-
Fisuc_3109 FSU_0380	P3 protein	√	√	√	Unknown	-0.226	60.975	5.93	No	-
Fisuc_3112 FSU_0383	DSBA oxidoreductase	√	√	√	Non cytoplasmic	-0.413	27.626	8.58	Yes (24-25)	-
Fisuc_2974 FSU_0239	PPIC-type PPIASE domain protein	√	√	√	Non cytoplasmic	-0.422	36.986	9.5	Yes (22-23)	-
rplL Fisuc_1272	50S ribosomal protein L7/L12	√	√	√	Non cytoplasmic	0.206	12.792	5.27	No	-
Fisuc_1259 FSU_1722	Histidine kinase	√	√	√	Cytoplasmic membrane	-0.321	143.171	5.06	No	-
FSU_2120	FG-GAP repeat protein	√	√	√	Unknown	-0.216	121.23	5.79	No	-
Fisuc_1632	FG-GAP repeat protein	√	√	√	Non cytoplasmic	-0.193	122.994	5.93	Yes (23-24)	-
FSU_0746	Pentapeptide repeat domain protein	√	√	√	Extracellular	-0.396	48.123	9	Yes (25-26)	-
FSU_0013	PPIC-type PPIASE domain protein	√	√	√	Outer membrane	-0.267	74.36	5.35	No	-
Fisuc_3033 FSU_0298	Mechanosensitive ion channel family protein	√	√	-	Cytoplasmic membrane	0.55	29.387	6.32	No	-
Fisuc_1658 FSU_2147	TPR repeat-containing protein	√	√	-	Non cytoplasmic	-0.676	27.923	7.61	Yes (22-23)	-
yidC Fisuc_1750 FSU_2248	Inner membrane protein oxaA	√	√	-	Cytoplasmic membrane	0.023	67.957	9.09	No	-
Fisuc_1429 FSU_1897	LemA family protein	√	√	-	Unknown	0.068	20.494	7.81	No	-
Fisuc_1008 FSU_1454	BatB protein	√	√	-	Cytoplasmic membrane	-0.006	38.346	9.54	No	-
Fisuc_1456 FSU_1929	LemA family protein	√	√	-	Unknown	-0.109	22.275	8.77	No	-
Fisuc_1229	Binding-protein-dependent transport systems inner membrane component	√	-	√	Cytoplasmic membrane	0.227	51.588	9.32	No	-
FSU_1690	Putative oligopeptide/dipeptide ABC transporter, permease protein	√	-	√	Cytoplasmic membrane	0.24	50.827	9.28	No	-
Fisuc_0392	Mucin-associated surface protein (MASP)	√	-	√	Cytoplasmic membrane	0.085	22.813	4.75	Yes (17-18)	-

Fisuc_0002 FSU_0395	Diaminopimelate dehydrogenase	√	-	√	Unknown	-0.17	35.752	6.76	No	-
Fisuc_2432 FSU_2995	Adenylate/guanylate cyclase domain protein	√	-	√	Cytoplasmic membrane	-0.341	115.294	6.03	No	-
hom Fisuc_2253 FSU_2798	Homoserine dehydrogenase	√	-	√	Unknown	0.142	45.884	5.9	No	-
fabG Fisuc_2016 FSU_2539	3-oxoacyl-(Acyl-carrier-protein) reductase	√	-	√	Unknown	0.094	25.131	6.35	No	[360]
Fisuc_1897 FSU_2403	TonB family protein	-	√	√	Unknown	-0.393	32.452	9.84	No	-
typA Fisuc_0016 FSU_0409	GTP-binding protein TypA	-	√	√	Cytoplasmic membrane	-0.291	67.979	5.16	No	-
Fisuc_0042 FSU_0435	MotA/TolQ/ExbB proton channel	-	√	√	Non cytoplasmic	-0.079	57.895	9.54	Yes (49-50)	-
Fisuc_0457 FSU_0874	Band 7 protein	-	√	√	Cytoplasmic membrane	-0.322	55.171	5.4	No	-
Fisuc_0617	Extracellular solute-binding protein family 1	-	√	√	Unknown	-0.213	57.545	7.77	No	-
Fisuc_1151 FSU_1609	OmpA family protein	-	√	√	Non cytoplasmic	-0.274	22.491	8.74	No	-
Fisuc_1465 FSU_1938	Extracellular ligand-binding receptor	-	√	√	Non cytoplasmic	-0.212	67.212	9.25	Yes (19-20)	-
FSU_1047	Extracellular solute-binding protein	-	√	√	Unknown	-0.218	57.575	7.77	No	-
Fisuc_0288 FSU_0700	Outer membrane efflux protein	-	√	√	Outer membrane	-0.378	52.346	5.31	Yes (21-22)	[360]
nuoE Fisuc_2129 FSU_2664	NADH dehydrogenase (Ubiquinone) 24 kDa subunit	-	√	√	Unknown	-0.212	37.941	6.75	No	-
nuoF Fisuc_2130 FSU_2665	NADH dehydrogenase (Quinone)	-	√	√	Unknown	-0.065	46.835	6.02	No	-
murG Fisuc_0566 FSU_0991	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	-	√	√	Cytoplasmic membrane	0.028	38.473	8.72	No	-
Fisuc_0242 FSU_0652	MORN variant repeat protein	-	√	√	Non cytoplasmic	-0.658	34.161	6.96	Yes (17-18)	-
Fisuc_0851 FSU_1295	Carboxyl-terminal protease	-	√	√	Cytoplasmic membrane	-0.359	66.051	9.19	Yes (24-25)	-
Fisuc_0871 FSU_1317	Peptidase M23	-	√	√	Non cytoplasmic	-0.201	48.044	9.71	No	-
Fisuc_1443 FSU_1914	Peptidase M23	-	√	√	Outer membrane	-0.327	30.171	9.65	No	-
FSU_1305	Putative BatD protein	-	√	√	Unknown	-0.305	70.698	9.43	No	-
Fisuc_0038 FSU_0431	TPR repeat-containing protein	-	√	-	Non cytoplasmic	-0.384	39.988	6.22	No	-
Fisuc_2028 FSU_2553	Efflux transporter, RND family, MFP subunit -		√	-	Cytoplasmic membrane	-0.24	46.108	9.5	No	-

Fisuc_0120	Tetratricopeptide TPR_2 repeat protein	-	√	-	Unknown	-0.413	143.314	5.69	Yes (27-28)	-
Fisuc_0480 FSU_0898	Efflux transporter, RND family, MFP subunit	-	√	-	Cytoplasmic membrane	-0.036	38.448	9.52	No	-
Fisuc_0539	Type II and III secretion system protein	-	√	-	Outermembrane	-0.022	65.248	5.9	Yes (22-23)	-
Fisuc_0743 FSU_1181	ABC transporter related protein	-	√	-	Cytoplasmic membrane	-0.071	29.041	5.61	No	-
Fisuc_0885 FSU_1331	Outer membrane efflux protein	-	√	-	Outer membrane	-0.345	61.621	5.22	Yes (19-20)	-
atpD Fisuc_2839 FSU_0095	V-type ATPase, D subunit	-	√	-	Unknown	-0.448	23.981	9.87	No	-
Fisuc_1227 FSU_1688	Oligopeptide/dipeptide ABC transporter, ATP-binding protein	-	√	-	Cytoplasmic membrane	-0.138	36.622	8.33	No	-
Fisuc_1395 FSU_1863	Capsular polysaccharide biosynthesis domain protein	-	√	-	Unknown	-0.002	43.793	5.5	Yes (17-18)	-
secY Fisuc_1402 FSU_1870	Preprotein translocase subunit secY	-	√	-	Cytoplasmic membrane	0.57	49.184	9.58	No	-
Fisuc_1571 FSU_2056	Outer membrane efflux protein	-	√	-	Outer membrane	-0.387	47.008	5.45	Yes (20-21)	-
Fisuc_1621 FSU_2110	Extracellular solute-binding protein	-	√	-	Non cytoplasmic	-0.057	58.667	5.89	Yes (28-29)	-
Fisuc_0105	Dihydroorotate oxidase	-	√	-	Unknown	-0.316	43.717	8.7	No	-
nuoI1 Fisuc_2133	NADH-quinone oxidoreductase subunit I 1	-	√	-	Unknown	-0.423	21.693	7.57	No	-
Fisuc_0012 FSU_0405	Putative peptide chain release factor	-	√	-	Unknown	-0.529	17.922	9.66	No	-
FSU_0506	Dihydroorotate oxidase	-	√	-	Cytoplasmic membrane	-0.228	42.32	8.89	No	-
psd Fisuc_0585 FSU_1012	Phosphatidylserine decarboxylase	-	√	-	Cytoplasmic membrane	-0.207	31.835	9.5	No	-
ftsH1 Fisuc_2778	ATP-dependent zinc metalloprotease FtsH 1	-	√	-	Cytoplasmic membrane	-0.37	77.197	6.04	No	-
Fisuc_2788 FSU_0040	Peptidase M23	-	√	-	Unknown	-0.195	34.358	9.54	No	-
Fisuc_1007	Peptidase M16 domain protein	-	√	-	Unknown	-0.091	55.733	6.44	Yes (22-23)	-
Fisuc_3072 FSU_0338	Penicillin binding transpeptidase domain protein	-	√	-	Unknown	-0.385	49.85	9.07	No	-
Fisuc_1133 FSU_1591	Putative transporter	-	√	-	Cytoplasmic membrane	0.726	58.663	9.18	No	-
Fisuc_1652	Mammalian cell entry related domain protein	-	√	-	Unknown	-0.029	37.004	5.38	No	-
Fisuc_1744 FSU_2242	S Diguanylate cyclase	-	√	-	Unknown	-0.101	59.019	7.66	No	-
ftsH FSU_0030	Cell division protein FtsH	--	√	-	Cytoplasmic membrane	-0.376	74.874	5.72	No	-
nuoBC Fisuc_2344 FSU_2894	NADH-quinone oxidoreductase, B subunit	-	√	-	Cytoplasmic membrane	-0.291	48.044	5.88	No	-

FSU_1453	Peptidase M16 domain protein	-	√	-	Unknown	-0.116	54.068	6.07	No	-
FSU_2141	Mce-like protein	-	√	-	Unknown	-0.013	35.369	5.23	No	-
Fisuc_0149 FSU_0552	Sulfate ABC transporter, periplasmic sulfate-binding protein	-	-	√	Periplasm	-0.401	37.761	5.68	Yes (22-23)	-
Fisuc_0197 FSU_0604	OmpA/MotB domain protein	-	-	√	Unknown	-0.398	21.493	4.95	No	-
Fisuc_0555 FSU_0979	Cell wall/surface repeat protein	-	-	√	Outer membrane	-0.084	135.996	5.23	Yes (15-16)	-
leuC Fisuc_0067 FSU_0466	3-isopropylmalate dehydratase large subunit	-	-	√	Unknown	-0.232	50.913	6.07	No	-
leuD Fisuc_0068 FSU_0467	3-isopropylmalate dehydratase, small subunit	-	-	√	Unknown	-0.163	22.009	6.21	No	-
Fisuc_0093	von Willebrand factor type A	-	-	√	Unknown	-0.315	25.004	4.61	No	-
plsX Fisuc_2015 FSU_2538	Phosphate acyltransferase	-	-	√	Unknown	0.133	34.28	5.38	No	-
acpP Fisuc_2017 FSU_2540	Acyl carrier protein	-	-	√	Unknown	-0.252	8.9	4.28	No	-
Fisuc_2059 FSU_2587	Oxidoreductase domain protein	-	-	√	Periplasm	-0.306	45.143	8.41	No	-
Fisuc_0149 FSU_0552	Sulfate ABC transporter, periplasmic sulfate-binding protein	-	-	√	Periplasm	-0.401	37.761	5.68	Yes (22-23)	-
Fisuc_0163 FSU_0566	SirA family protein	-	-	√	Unknown	-0.417	9.127	4.46	No	-
Fisuc_0186 FSU_0593	Endoribonuclease L-PSP	-	-	√	Unknown	0.27	16.319	6.08	No	-
Fisuc_2806	Beta-ketoacyl synthase	-	-	√	Cytoplasmic membrane	-0.037	83.389	5.25	No	-
cysK_1 Fisuc_0494 FSU_0912	Cysteine synthase A	-	-	√	Cytoplasmic membrane	-0.106	32.556	6.77	No	-
ndk Fisuc_2491 FSU_3059	Nucleoside diphosphate kinase	-	-	√	Extracellular	-0.171	16.563	5.94	No	-
purT Fisuc_0556 FSU_0980	Phosphoribosylglycinamide formyltransferase 2	-	-	√	Cytoplasmic membrane	-0.082	42.916	5.68	No	-
Fisuc_0618 FSU_1048	Diguanylate cyclase	-	-	√	Unknown	-0.216	64.355	7.69	No	-
Fisuc_0655	Aminotransferase class I and II	-	-	√	Unknown	-0.102	47.04	6.1	No	-
Fisuc_0656 FSU_1087	HD domain protein	-	-	√	Cytoplasmic membrane	0.207	69.545	6.05	No	-
Fisuc_0221	50S ribosomal protein L31	-	-	√	Unknown	-0.824	11.311	9.4	No	-
FSU_1314	Leucine-rich repeat domain protein	-	-	√	Extracellular	-0.122	27.822	4.68	No	-
Fisuc_0868	Leucine-rich repeat protein	-	-	√	Extracellular	-0.093	27.807	4.68	No	-
Fisuc_2819 FSU_0075	Peptidyl-prolyl cis-trans isomerase	-	-	√	Non cytoplasmic	-0.07	18.707	6.13	Yes (19-20)	-
Fisuc_0989 FSU_1435	LicC domain protein	-	-	√	Unknown	-0.369	28.65	5.27	No	-
fabF_3 Fisuc_307	3-oxoacyl-[acyl-carrier-protein] synthase 2	-	-	√	Cytoplasmic membrane	-0.049	43.761	5.67	No	-

0 FSU_0336										
Fisuc_1598 FSU_2085	Biotin/lipoic acid binding domain protein	-	-	√	Non cytoplasmic	0.368	11.998	4.85	No	-
FSU_0061	Beta-ketoacyl synthase family protein	-	-	√	Cytoplasmic membrane	-0.026	83.924	5.25	No	-
rpmE FSU_0628	50S ribosomal protein L31	-	-	√	Unknown	-0.827	9.949	9.47	No	-
rfbC Fisuc_0777 FSU_1218	dTDP-4-dehydrorhamnose 3,5-epimerase	-	-	√	Unknown	-0.327	21.178	5.33	No	-

Appendix 5.2 List of membrane associated proteins with unknown activities in *F. succinogenes* S85

Locus ID	Protein description	Glucose	MC cellulose	AS cellulose	Location[a]	Gravy Index[b]	Molecular mass (kDa)[b]	pI[b]	Presence of signal peptide (amino acid position)[c]	Reference
Fisuc_0463	Putative uncharacterized protein	√	-	-	Noncytoplasmic	-0.587	42.59	6.62	Yes (24-25)	-
Fisuc_2413 FSU_2974	Putative uncharacterized protein	√	-	-	Cytoplasmic membrane	0.619	23.33	9.68	No	-
Fisuc_0380	Putative uncharacterized protein	√	-	-	Unknown	-1.293	6.3	4.33	No	-
Fisuc_2763 FSU_0015	Putative uncharacterized protein	√	-	-	Non cytoplasmic	-0.339	39.8	7.62	Yes (21-22)	-
Fisuc_0719 FSU_1156	Putative uncharacterized protein	√	-	-	Outer membrane	-0.252	45.96	5.49	Yes (17-18)	-
Fisuc_0720 FSU_1157	Putative lipoprotein	√	-	-	Non cytoplasmic	-0.42	54.63	5.62	Yes (22-23)	-
Fisuc_1385 FSU_1851	Putative uncharacterized protein	√	-	-	Unknown	-0.289	81.68	6.27	Yes (19-20)	-
Fisuc_0783 FSU_1224	Putative uncharacterized protein	√	-	-	Outer membrane	-0.121	30.12	4.46	Yes (20-21)	-
Fisuc_0202 FSU_0609	Lipoprotein	√	-	-	Non cytoplasmic	-0.501	19.76	5.01	Yes (19-20)	([360]
Fisuc_1316 FSU_1783	Membrane protein	√	-	-	Non cytoplasmic	-0.451	20.24	8.83	Yes (20-21)	[360]
FSU_1374	Putative uncharacterized protein	√	-	-	Unknown	0.129	10.48	10.89	No	-
FSU_1403	Conserved domain protein	√	-	-	Unknown	-0.425	9.27	9.13	No	-
FSU_0881	Putative uncharacterized protein	√	-	-	Non cytoplasmic	-0.585	41.95	6.35	Yes (19-20)	-
FSU_0795	Putative uncharacterized protein	√	-	-	Unknown	-0.783	7.95	4.24	No	-
Fisuc_1854	Putative uncharacterized protein	√	-	-	Unknown	-0.29	8.28	10.22	No	-

Fisuc_1476	Putative uncharacterized protein	√	-	-	Non cytoplasmic	-0.486	43.31	8.83	Yes (20-21)	-
Fisuc_2868	Putative uncharacterized protein	√	-	-	Unknown	-0.555	27.73	6.13	No	-
Fisuc_2965 FSU_0230	Membrane protein	√	√	√	Non cytoplasmic	-0.207	44.92	4.64	Yes (19-20)	[360]
Fisuc_1898 FSU_2404	Membrane protein	√	√	√	Non cytoplasmic	-0.268	48.26	4.79	No	[360]
Fisuc_1527 FSU_2009	Membrane protein	√	√	√	Outer membrane	-0.263	78.75	5.09	Yes (25-26)	[360]
Fisuc_1528 FSU_2010	Membrane protein	√	√	√	Non cytoplasmic	-0.195	78.195	5.64	Yes (	([360]
Fisuc_0033 FSU_0426	Putative uncharacterized protein	√	√	√	Unknown	-0.711	22.895	10.11	No	-
Fisuc_0043	Putative uncharacterized protein	√	√	√	Non cytoplasmic	-0.278	29.93	9.3	Yes (24-25)	-
Fisuc_0220 FSU_0627	Putative lipoprotein	√	√	√	Unknown	-0.373	32.7	9.54	No	-
Fisuc_0222	Putative uncharacterized protein	√	√	√	Unknown	-0.296	21.09	9.76	No	[360]
Fisuc_2147	Putative uncharacterized protein	√	√	√	Unknown	-0.111	24.82	6.34	No	-
Fisuc_2269	Putative uncharacterized protein	√	√	√	Unknown	-0.227	26.82	10.02	No	-
Fisuc_0328 FSU_0743	Putative uncharacterized protein	√	√	√	Periplasm	-0.314	25.13	6.63	No	-
Fisuc_2370 FSU_2924	Putative lipoprotein	√	√	√	Unknown	-0.59	16.756	5.37	No	-
Fisuc_0382 FSU_0797	Putative uncharacterized protein	√	√	√	Cytoplasmic membrane	-0.117	37.419	5.02	Yes (18-19)	[360]
Fisuc_2555 FSU_3125	Putative uncharacterized protein	√	√	√	Non cytoplasmic	-1.191	18.92	9.06	Yes (21-22)	-
Fisuc_2572 FSU_3142	Putative lipoprotein	√	√	√	Periplasm	-0.205	71.611	5.67	Yes (21-22)	-
Fisuc_2588 FSU_3158	Conserved domain protein	√	√	√	Cytoplasmic membrane	-0.201	64.29	8.22	No	-
Fisuc_0578	Putative uncharacterized protein	√	√	√	Unknown	-0.608	129.335	4.65	No	-
Fisuc_0741	Putative uncharacterized protein	√	√	√	Extracellular	-0.093	44.11	4.74	Yes (20-21)	-
Fisuc_0752 FSU_1190	Putative lipoprotein	√	√	√	Unknown	-0.669	22.029	4.67	Yes (21-22)	-
Fisuc_0767 FSU_1207	Putative lipoprotein	√	√	√	Non cytoplasmic	-0.105	37.411	4.38	Yes (24-25)	-
Fisuc_2795 FSU_0049	Putative lipoprotein	√	√	√	Non cytpplasmic	-0.79	12.637	5.48	Yes (23-24)	-
Fisuc_0888 FSU_1335	Putative uncharacterized protein	√	√	√	Non cytoplasmic	-0.487	51.75	5.28	Yes (17-18)	-
Fisuc_2863	Putative uncharacterized protein	√	√	√	Non cytoplasmic	-0.228	56.8	5.06	Yes (17-18)	-
Fisuc_1013 FSU_1460	Putative uncharacterized protein	√	√	√	Non cytoplasmic	-0.303	16.368	4.8	Yes (21-22)	-
Fisuc_1021 FSU_1468	Putative lipoprotein	√	√	√	Non cytoplasmic	-0.41	19.71	6.72	Yes (19-20)	-

Fisuc_2982	Putative uncharacterized protein	√	√	√	Outer membrane	-0.152	36.539	4.87	Yes (19-20)	-
Fisuc_3015 FSU_0280	Putative uncharacterized protein	√	√	√	Non cytoplasmic	-0.229	37.819	5.41	Yes (20-21)	-
Fisuc_3024 FSU_0289	Putative lipoprotein	√	√	√	Unknown	-0.16	31.238	4.86	Yes (5-6)	-
Fisuc_1203 FSU_1664	Putative uncharacterized protein	√	√	√	Unknown	-0.34	26.03	7.75	No	-
Fisuc_1526 FSU_2008	Putative uncharacterized protein	√	√	√	Non cytoplasmic	-0.338	85.239	5.5	Yes (19-20)	-
Fisuc_1529 FSU_2011	Putative lipoprotein	√	√	√	Non cytoplasmic	-0.279	24.182	4.8	Yes (23-24)	-
Fisuc_1597	Putative uncharacterized protein	√	√	√	Non cytoplasmic	0.144	29.528	9.01	Yes (22-23)	-
FSU_0247	Putative uncharacterized protein	√	√	√	Outer membrane	-0.23	34.983	4.8	No	-
FSU_2503	Putative uncharacterized protein	√	√	√	Unknown	-0.289	14.66	9.67	Yes (16-17)	-
FSU_2684	Putative uncharacterized protein	√	√	√	Cytoplasmic membrane	-0.1	26.421	6.93	No	-
FSU_2814	Putative uncharacterized protein	√	√	√	Unknown	-0.227	26.822	10.07	No	-
FSU_0629	Putative lipoprotein	√	√	√	Non cytoplasmic	-0.145	23.248	9.74	No	-
FSU_0784	Conserved domain protein	√	√	√	Non cytoplasmic	-0.281	45.949	7.03	Yes (36-37)	-
FSU_0792	Putative uncharacterized protein	√	√	√	Non cytoplasmic	-0.412	102.455	5.29	No	-
FSU_3096	Putative uncharacterized protein	√	√	√	Unknown	-0.316	34.113	5.24	No	-
FSU_1004	Putative uncharacterized protein	√	√	√	Unknown	-0.618	127.845	4.62	No	-
FSU_2084	Putative uncharacterized protein	√	√	√	Non cytoplasmic	0.154	29.558	9.01	Yes (22-23)	-
FSU_1179	Conserved domain protein	√	√	√	Extracellular	-0.071	46.207	4.82	No	-
Fisuc_2069	Putative uncharacterized protein	√	√	-	Outer membrane	-0.207	42.41	4.73	Yes (18-19)	-
Fisuc_0345 FSU_0759	Conserved domain protein	√	√	-	Periplasm	-0.603	39.887	5.86	No	-
Fisuc_2544 FSU_3113	Putative uncharacterized protein	√	√	-	Unknown	-0.026	36.422	9.55	No	-
Fisuc_2890	Putative uncharacterized protein	√	√	-	Unknown	-0.16	28.093	5.57	No	-
Fisuc_1435 FSU_1905	Putative uncharacterized protein	√	√	-	Non cytoplasmic	-0.15	35.581	5.52	Yes (26-27)	-
Fisuc_3021	Putative uncharacterized protein	√	√	-	Cytoplasmic membrane	-0.005	67.29	4.51	No	-
FSU_2597	Putative uncharacterized protein	√	√	-	Outer membrane	-0.22	41.786	4.64	No	-
Fisuc_0841 FSU_1285	Conserved domain protein	√	-	√	Unknown	-0.197	27.296	6.91	No	-
Fisuc_0100 FSU_0501	Putative lipoprotein	√	-	√	Unknown	-0.225	37.536	4.78	No	-
Fisuc_2020 FSU_2544	Putative uncharacterized protein	√	-	√	Non cytoplasmic	-0.055	28.17	5.1	Yes (18-19)	-

Fisuc_2072 FSU_2600	Putative uncharacterized protein	√	-	√	Outer membrane	-0.117	38.859	6.05	Yes (23-24)	-
Fisuc_0411 FSU_0825	Putative membrane protein	√	-	√	Cytoplasmic membrane	0.785	25.496	8.83	No	-
Fisuc_0283	Putative uncharacterized protein	√	-	√	Non cytoplasmic	-0.268	23.748	5.32	Yes (21-22)	-
Fisuc_2475 FSU_3041	Putative uncharacterized protein	√	-	√	Non cytoplasmic	-0.037	21.899	6.08	Yes (21-22)	-
Fisuc_2506 FSU_3074	Putative uncharacterized protein	√	-	√	Non cytoplasmic	-0.391	34.551	4.51	Yes (21-22)	-
Fisuc_0633	Putative uncharacterized protein	√	-	√	Non cytoplasmic	-0.347	33.313	4.92	Yes (21-22)	-
Fisuc_0742	Putative uncharacterized protein	√	-	√	Extracellular	-0.711	45.103	4.55	No	-
Fisuc_2897 FSU_0158	Putative lipoprotein	√	-	√	Non cytoplasmic	-0.017	25.293	5.3	Yes (19-20)	-
Fisuc_2739 FSU_3310	Putative membrane protein	√	-	√	Cytoplasmic membrane	0.119	115.48	6.13	No	-
Fisuc_1319 FSU_1786	Putative uncharacterized protein	√	-	√	Non cytoplasmic	0.183	27.32	5.58	Yes (19-20)	-
Fisuc_1485 FSU_1966	Putative uncharacterized protein	√	-	√	Non cytoplasmic	-0.184	21.636	8.98	Yes (21-22)	-
Fisuc_1490 FSU_1971	Putative lipoprotein	√	-	√	Unknown	-0.541	32.012	5.03	Yes (21-22)	-
FSU_0694	Putative uncharacterized protein	√	-	√	Unknown	-0.406	17.201	5.19	No	-
FSU_1180	Putative lipoprotein	√	-	√	Extracellular	-0.732	46.485	4.69	No	-
FSU_1064	Putative uncharacterized protein	√	-	√	Non cytoplasmic	-0.422	35.657	5.4	Yes (42-43)	-
Fisuc_1875 FSU_2377	Putative uncharacterized protein	-	√	√	Non cytoplasmic	-0.5	37.593	5.53	Yes (18-19)	-
Fisuc_0081 FSU_0479	Putative uncharacterized protein	-	√	√	Unknown	-0.151	20.617	9.01	No	-
Fisuc_2158	Putative uncharacterized protein	--	√	√	Unknown	0.032	28.263	8.7	No	-
Fisuc_2226	Putative uncharacterized protein		√	√	Unknown	-0.092	16.115	9.12	Yes (20-21)	-
Fisuc_0482 FSU_0900	Putative uncharacterized protein	-	√	√	Outer membrane	-0.259	54.535	5.68	Yes (21-22)	-
Fisuc_2326	Putative uncharacterized protein	-	√	√	Non cytoplasmic	-0.554	23.324	8.7	Yes (34-35)	-
Fisuc_2612 FSU_3182	Conserved domain protein	-	√	√	Non cytoplasmic	-0.293	72.606	8.63	Yes (23-24)	-
Fisuc_2704 FSU_3272	Conserved domain protein	-	√	√	Unknown	-0.216	116.053	6.23	No	-
Fisuc_2775 FSU_0027	Putative uncharacterized protein	-	√	√	Unknown	-0.704	45.781	9.24	No	-
Fisuc_0756 FSU_1194	Putative lipoprotein	-	√	√	Periplasm	-0.128	32.275	8.54	Yes (22-23)	-
Fisuc_0861	Putative uncharacterized protein	-	√	√	Unknown	-0.354	67.25	9.38	Yes (17-18)	-
Fisuc_0866	Putative uncharacterized protein	-	√	√	Unknown	-0.534	264.045	5.11	No	-

Fisuc_1223 FSU_1684	Putative uncharacterized protein	-	√	√	Non cytoplasmic	0.167	31.235	5.21	Yes (19-20)	-
Fisuc_1317 FSU_1784	Putative lipoprotein	-	√	√	Non cytoplasmic	-0.371	22.441	5.18	Yes (20-21)	-
FSU_2695	Putative uncharacterized protein	-	√	√	Unknown	0.018	27.822	8.39	No	-
FSU_2768	Putative uncharacterized protein	-	√	√	Unknown	-0.101	18.525	9.23	No	-
FSU_2876	Putative uncharacterized protein	-	√	√	Non cytoplasmic	-0.559	25.727	9.18	No	-
FSU_1310	Putative uncharacterized protein	-	√	√	Unknown	-0.511	273.261	5.11	Yes (27-28)	-
Fisuc_1899	Putative uncharacterized protein	-	√	-	Non cytoplasmic	-0.448	14.624	6.74	No	-
Fisuc_1907 FSU_2415	Putative lipoprotein	-	√	-	Non cytoplasmic	-0.463	92.291	5.3	Yes (27-28)	-
Fisuc_1926 FSU_2435	Putative uncharacterized protein	-	√	-	Cytoplasmic membrane	0.122	28.113	6.33	No	-
Fisuc_1954 FSU_2474	Putative uncharacterized protein	-	√	-	Extracellular	-0.26	64.462	6.63	No	-
Fisuc_0009	Putative uncharacterized protein	-	√	-	Unknown	-0.286	146.239	7.21	No	-
Fisuc_0058 FSU_0452	Conserved domain protein	-	√	-	Unknown	-0.234	87.056	8.49	No	-
Fisuc_0224 FSU_0631	Putative uncharacterized protein	-	√	-	Outer membrane	-0.463	107.908	4.89	Yes (17-18)	-
Fisuc_2227	Putative uncharacterized protein	-	√	-	Non cytoplasmic	-0.46	35.518	6.52	No	-
Fisuc_2263	Putative uncharacterized protein	-	√	-	Unknown	-0.366	24.371	4.61	No	-
Fisuc_0474 FSU_0892	Putative uncharacterized protein	-	√	-	Unknown	0.11	24.669	5.58	No	[360]
Fisuc_2552	Putative uncharacterized protein	-	√	-	Unknown	-0.076	17.607	10.09	No	-
Fisuc_2573	Putative uncharacterized protein	-	√	-	Unknown	-0.105	29.498	6	No	-
Fisuc_0657 FSU_1088	Putative lipoprotein	-	√	-	Non cytoplasmic	-0.103	34.721	4.56	Yes (22-23)	-
Fisuc_0670 FSU_1106	Conserved domain protein	-	√	-	Unknown	-0.267	41.663	6.57	No	-
Fisuc_3030	Putative uncharacterized protein	-	√	-	Unknown	-0.221	28.986	9.39	No	-
Fisuc_1171 FSU_1633	Putative lipoprotein	-	√	-	Non cytoplasmic	-0.789	25.395	8.83	Yes (29-30)	-
Fisuc_1651	Putative uncharacterized protein	-	√	-	Non cytoplasmic	-0.437	28.289	9.11	Yes (21-22)	-
Fisuc_1755	Putative uncharacterized protein	-	√	-	Unknown	-0.272	21.012	6.83	No	-
Fisuc_1844 FSU_2348	Putative lipoprotein	-	√	-	Non cytoplasmic	-0.324	41.623	5.73	No	-
FSU_2405	Putative lipoprotein	-	√	-	Cytoplasmic membrane	-0.442	14.378	7.76	No	-
FSU_0402	Conserved domain protein	-	√	-	Outer membrane	-0.27	151.901	6.61	Yes (22-23)	-
FSU_0522	Putative lipoprotein	-	√	-	Non cytoplasmic	-0.408	35.975	6.52	Yes (17-18)	-

FSU_2808	Putative uncharacterized protein	-	√	-	Unknown	-0.37	24.617	4.61	No	-
FSU_3143	Putative uncharacterized protein	-	√	-	Unknown	-0.051	28.285	5.71	No	-
Fisuc_0064	Putative uncharacterized protein	-	-	√	Unknown	-0.237	15.357	6.58	No	-
Fisuc_0213 FSU_0620	Putative uncharacterized protein	-	-	√	Unknown	-0.021	25.063	4.39	No	-
Fisuc_2190 FSU_2729	Conserved domain protein	-	-	√	Unknown	-0.585	120.809	7.73	No	[360]
Fisuc_0357 FSU_0772	Putative uncharacterized protein	-	-	√	Non cytoplasmic	-0.431	26.7	9.4	Yes (17-18)	-
Fisuc_2314	Putative uncharacterized protein	-	-	√	Non cytoplasmic	0.021	20.342	4.74	Yes (20-21)	-
Fisuc_0417 FSU_0831	Putative uncharacterized protein	-	-	√	Outer membrane	-0.284	87.192	5.33	Yes (19-20)	-
Fisuc_0455 FSU_0872	Conserved domain protein	-	-	√	Non cytoplasmic	-0.342	27.372	4.63	Yes (20-21)	-
Fisuc_0558 FSU_0982	Putative lipoprotein	-	-	√	Extracellular	-0.497	72.266	4.99	Yes (21-22)	-
Fisuc_0770	Putative uncharacterized protein	-	-	√	Outer membrane	-0.439	237.304	5.1	No	-
Fisuc_0820 FSU_1263	Putative uncharacterized protein	-	-	√	Unknown	-0.295	26.781	5.36	No	-
Fisuc_0886 FSU_1333	Putative lipoprotein	-	-	√	Periplasm	-0.188	60.005	4.68	Yes (23-24)	-
Fisuc_2895	Putative uncharacterized protein	-	-	√	Unknown	-0.186	11.273	4.92	No	-
Fisuc_1141 FSU_1599	Putative lipoprotein	-	-	√	Unknown	-0.391	45.816	5.28	Yes (21-22)	-
Fisuc_1474 FSU_1953	Conserved domain protein	-	-	√	Non cytoplasmic	-0.338	163.156	4.86	Yes (20-21)	-
Fisuc_1536 FSU_2018	Conserved domain protein	-	-	√	Unknown	-0.277	15.993	9.91	No	-
Fisuc_1660	Putative uncharacterized protein	--	-	√	Unknown	-0.347	32.623	4.83	Yes (21-22)	-
FSU_0103	Putative uncharacterized protein	-	-	√	Unknown	0.316	7.586	5.03	No	-
FSU_1086	Putative aspartate aminotransferase	-	-	√	Unknown	-0.107	47.327	6.41	No	-
FSU_1405	Putative uncharacterized protein	-	-	√	Unknown	-0.508	34.935	5.87	No	-
FSU_2149	Putative uncharacterized protein	-	-	√	Unknown	-0.274	35.374	4.91	No	-

Appendix 6.1 Proteins identied with significant changes by ultracentrifugationc method

Locus ID	Protein description	Location	Fold change	Peptides	pI	Molecular mass (kDa)	Cazy family
Fisuc_1894 FSU_2400	MotA/TolQ/ExbB proton channel (MotA/TolQ/ExbB proton channel family protein)	Cytoplasmic membrane	-2.58724	87	9.14	23055.47	
FSU_1029	Membrane protein	Unknown (non-cytoplasmic)	-2.34746	25	4.84	44313.23	
Fisuc_0600	Putative uncharacterized protein	Unknown (non-cytoplasmic)	-2.34746	25	4.92	43210.07	
gdhA Fisuc_2811 FSU_0066	Glutamate dehydrogenase	Unknown (multiple localisation)	-1.96646	179	6.86	48622.33	
Fisuc_2041 FSU_2567	Putative type IV pilin (Type IV pilin)	Unknown	-1.9435	56	5.43	17516.89	
Fisuc_1151 FSU_1609	OmpA family protein (OmpA/MotB domain protein)	Unknown	-1.89984	39	8.74	22491.64	
Fisuc_2021 FSU_2545	Putative uncharacterized protein	Unknown	-1.76776	31	4.81	30990.91	
Fisuc_2370 FSU_2924	Putative lipoprotein	Unknown	-1.69647	37	5.37	16756.85	
ndk	Nucleoside diphosphate kinase (NDK)	Extracellular	-1.68239	19	5.94	16563.08	
Fisuc_1897 FSU_2403	TonB family protein	Unknown	-1.65723	66	9.84	32452.41	
Fisuc_2308	Putative uncharacterized protein	Unknown (non-cytoplasmic)	-1.59106	22	4.62	15281.1	
Fisuc_1523 FSU_2005	Cellulase (Glycoside hydrolase family 5)	Unknown	-1.58382	29	4.94	42062.02	GH5
Fisuc_2987 FSU_0252	Ankyrin (Ankyrin repeat family protein)	Unknown (non-cytoplasmic)	-1.54579	45	8.97	26752.83	
Fisuc_0718 FSU_1155	Putative uncharacterized protein	Unknown	-1.53088	92	9.71	42666.34	
Fisuc_0718 FSU_1155	Putative uncharacterized protein	Unkown	-1.53088	92	9.71	42666.34	
Fisuc_2249 FSU_2794	Fibrobacter succinogenes major paralogous domain protein	Unknown (non-cytoplasmic)	-1.51936	163	4.57	70117.73	
Fisuc_1525 FSU_2007	Cellulose-binding domain protein	Unknown (non-cytoplasmic)	-1.51902	62	6.46	29200.5	CBM30
Fisuc_0008 FSU_0401	Putative membrane protein	Cytoplasmic membrane	-1.47998	21	10.08	31517.2	
Fisuc_1898 FSU_2404	Membrane protein	Unknown	-1.40206	76	4.79	48258.76	
Fisuc_0767 FSU_1207	Putative lipoprotein	Unknown	-1.35703	78	4.38	37411.72	
Fisuc_2672 FSU_3241	Polysaccharide biosynthesis protein	Cytoplasmic membrane	-1.34682	55	9.61	53880.45	
Fisuc_1802	Glycoside hydrolase family 8	Unknown	-1.32508	96	5.63	79810.17	GH8
FSU_2303	Glycoside hydrolase family 8	Unknown	-1.32508	96	5.55	81395.08	GH8
Fisuc_2897 FSU_0158	Putative lipoprotein	Unknown	-1.29479	49	5.3	25293.25	
Fisuc_1224 FSU_1685	Cellulase	Extracellular	-1.28745	119	7.55	80027.53	GH5
murG	pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	Cytoplasmic membrane	-1.22431	49	8.72	38473.62	
Fisuc_2544 FSU_3113	Putative uncharacterized protein	Uknown	-1.22126	94	9.55	36422.18	
Fisuc_3015 FSU_0280	Putative uncharacterized protein	Unknown	-1.1863	51	5.41	37819.7	
Fisuc_3112 FSU_0383	DSBA oxidoreductase (Putative outer membrane protein)	Unknown	-1.16959	142	8.58	27626.66	
Fisuc_1892 FSU_2397	TPR domain protein (TPR repeat-containing protein)	Unknown	-1.16089	146	5.51	83838.55	
Fisuc_0824 FSU_1267	Insecticidal toxin-like protein	Unknown (multiple localisation)	1.106848	383	4.8	377705.1	
Fisuc_0775 FSU_1216	Polysaccharide biosynthesis/export protein (Polysaccharide export protein)	Outer membrane	1.191759	101	5.74	41486.58	

surA Fisuc_0518 FSU_0941	Peptidyl-prolyl cis-trans isomerase SurA	Periplasm	1.195907	82	6.63	48242.47	
purT	Phosphoribosylglycinamide formyltransferase 2	Cytoplasmic membrane	1.202272	77	5.68	42916.42	
Fisuc_2386	Integral membrane sensor hybrid histidine kinase	Cytoplasmic membrane	1.216033	140	5.19	158187.8	
Fisuc_3033 FSU_0298	Mechanosensitive ion channel family protein (MscS Mechanosensitive ion channel)	Cytoplasmic membrane	1.216997	58	6.32	29387.64	
Fisuc_0851 FSU_1295	Carboxyl-terminal protease	Cytoplasmic membrane	1.2197	111	9.19	66051.97	
nadB Fisuc_2761 FSU_0012	L-aspartate oxidase	Unknown (multiple localisation)	1.223465	61	8.9	57977.04	
Fisuc_0801 FSU_1244	Putative lipoprotein, TIGR02171	Unknown	1.224646	43	5.2	101218.9	
mrcA Fisuc_0062 FSU_0457	Penicillin-binding protein 1A	Cytoplasmic membrane	1.246277	70	9.33	90219.42	
Fisuc_1592	OmpA/MotB domain protein	Outer membrane	1.25916	364	4.72	73841.2	
Fisuc_1518 FSU_2000	Antigen-like protein	Unknown	1.260768	42	5.43	86541.08	
FSU_0013	PPIC-type PPIASE domain protein	Outer membrane	1.27923	272	5.35	74360.87	
lepA	Elongation factor 4 (EF-4)	Cytoplasmic membrane	1.283092	91	5.11	67876.94	
Fisuc_1021 FSU_1468	Putative lipoprotein	unknown	1.291164	109	6.72	19714.23	
Fisuc_3023 FSU_0288	Putative uncharacterized protein	unknown	1.292245	63	6.32	26091.54	
Fisuc_2995 FSU_0260	Putative lipoprotein	Outer membrane	1.296667	95	5.23	98102.75	
Fisuc_1247 FSU_1709	Putative lipoprotein	Unknown	1.300218	40	4.75	66401.18	
Fisuc_2917 FSU_0180	OmpA family protein (OmpA/MotB domain protein)	Outer membrane	1.302168	89	5.27	70794.57	
Fisuc_2965 FSU_0230	Membrane protein	Unknown	1.302248	71	4.64	44916.19	
Fisuc_2503 FSU_3071	Extracellular solute-binding protein family 3	Periplasm	1.321313	62	5.47	28753.02	
yidC	Membrane protein insertase YidC	Cytoplasmic membrane	1.323256	75	9.09	67957.4	
fdrA Fisuc_2493 FSU_3061	Succinate dehydrogenase or fumarate reductase, flavoprotein subunit	Cytoplasmic membrane	1.325557	201	6.93	70572.9	
Fisuc_1141 FSU_1599	Putative lipoprotein	Unknown	1.329258	47	5.28	45816.09	
Fisuc_2974 FSU_0239	PPIC-type PPIASE domain protein	Unknown (multiple localisation)	1.338031	105	9.5	36986.91	
fdrB Fisuc_2494 FSU_3062	4Fe-4S ferredoxin iron-sulfur binding domain protein	Cytoplasmic membrane	1.341483	75	6.52	28006.3	
Fisuc_0741	Putative uncharacterized protein	Extracellular	1.364853	52	4.74	44119.89	
FSU_1179	Conserved domain protein	Extracellular	1.364853	52	4.82	46207.38	
Fisuc_0289 FSU_0701	Efflux transporter, RND family, MFP subunit	Unknown (non-cytoplasmic)	1.372731	76	9.45	37256.16	
Fisuc_1476	Putative uncharacterized protein	Unknown	1.440277	57	8.83	43314.81	
Fisuc_3111 FSU_0382	Carbohydrate binding family 11	Unknown	1.4502	353	7.81	118616.7	CBM30,CBM30,CBM11,GH51
Fisuc_1877	Putative uncharacterized protein	Unknown	1.454864	11	4.4	33132.57	
secG Fisuc_2367 FSU_2921	Preprotein translocase, SecG subunit	Cytoplasmic membrane	1.514361	12	8.73	16955.93	
Fisuc_1885	Putative uncharacterized protein	Cytoplasmic membrane	1.516182	14	4.64	21989.72	
FSU_2390	Putative uncharacterized protein	Cytoplasmic membrane	1.516182	14	5.1	25007.25	
Fisuc_2704 FSU_3272	Conserved domain protein (	Unknown	1.550747	225	6.23	116053.6	GH116
Fisuc_1595 FSU_2082	Phosphoglycerate mutase	Unknown (multiple localisation)	1.552123	65	4.9	73212.87	
Fisuc_0283	Putative uncharacterized protein	Unknown	1.862633	13	5.32	23748.93	

Fisuc_0331	Pentapeptide repeat protein	Extracellular	2.208716	244	9	47890.88	
FSU_0746	Pentapeptide repeat domain protein	extracellular	2.208716	244	9	48123.18	

Appendix 6.2 Proteins identied with significant changes by biotitinylation method

Locus ID	Protein Description	Location	Fold change	Peptides	pI	Molecular mass (kDa)	Cazy family
gdhA Fisuc_2811 FSU_0066	Glutamate dehydrogenase	Unknown (multiple location)	-3.0159	179	6.86	48622.33	
FSU_1029	Membrane protein	Unknown (non-cytoplasmic)	-2.7668	25	4.84	44313.23	
Fisuc_0600	Putative uncharacterized protein	Unknown (non-cytoplasm)	-2.7668	25	4.92	43210.07	
Fisuc_2704 FSU_3272	Conserved domain protein	Unknown	-2.5693	225	6.23	116053.63	GH116
plsY Fisuc_2247	Glycerol-3-phosphate acyltransferase	cytoplasmic membrane	-2.2812	13	9.58	23437.00	
Fisuc_0242 FSU_0652	MORN variant repeat protein	Unknown	-2.1069	48	6.96	34161.37	
Fisuc_2249 FSU_2794	Fibrobacter succinogenes major paralogous domain protein	Unknown	-1.9215	163	4.57	70117.73	
Fisuc_0209	Putative uncharacterized protein	Unknown (non-cytoplasm)	-1.8560	16	9.19	23256.38	
Fisuc_0820 FSU_1263	Putative uncharacterized protein	Unknown	-1.8256	50	5.36	26781.48	
Fisuc_1425 FSU_1893	Cellulase	Unknown (multiple location)	-1.7971	18	6.15	58365.87	GH45
nifJ Fisuc_2881 FSU_0139	Pyruvate-flavodoxin oxidoreductase	Unknown	-1.7897	326	7.23	129294.35	
Fisuc_2775 FSU_0027	Putative uncharacterized protein	Unknown	-1.7777	52	9.24	45781.82	
Fisuc_1010 FSU_1457	BatA protein	cytoplasmic membrane	-1.7436	73	9.18	40980.70	
Fisuc_1151 FSU_1609	OmpA family protein (OmpA/MotB domain protein)	Unknown (non-cytoplasmic)	-1.6272	39	8.74	22491.64	
Fisuc_3111 FSU_0382	Carbohydrate binding family 11	Unnkown (non-cytoplasmic)	-1.5937	353	7.81	118616.73	CBM30,CBM30,CBM11,GH51
Fisuc_2868	Putative uncharacterized protein	Unknown	-1.5687	27	6.13	27737.28	
Fisuc_2900 FSU_0162	Cellodextrin-phosphorylase	cytoplasmic membrane	-1.5566	119	8.16	93662.58	GH94
Fisuc_0818	Putative uncharacterized protein	Unnkown	-1.5505	33	6.07	19547.20	
Fisuc_0718 FSU_1155	Putative uncharacterized protein	Unnkown	-1.5050	92	9.71	42666.34	
Fisuc_0288 FSU_0700	Outer membrane efflux protein	Unknown	-1.4644	95	5.31	52346.45	
Fisuc_0557 FSU_0981	Conserved domain protein	Unknown	-1.4605	34	4.63	60784.58	

Fisuc_0165 FSU_0568	HesA/MoeB/ThiF family protein	Outer membrane	-1.4243	42	7.6	29259.20	
Fisuc_1897 FSU_2403	TonB family protein	Unknown	-1.4031	66	9.84	32452.41	
Fisuc_0033 FSU_0426	Putative uncharacterized protein	Unknown	-1.4018	37	10.11	22895.15	
Fisuc_0457 FSU_0874	Band 7 protein	cytoplasmic membrane	-1.3886	166	5.4	55171.27	
speA	Biosynthetic arginine decarboxylase (ADC)	Unknown (multiple location)	-1.3559	66	5.06	71331.96	
lepA	Elongation factor 4	cytoplasmic membrane	-1.3354	91	5.11	67876.94	
Fisuc_0767 FSU_1207	Putative lipoprotein	Unknown	-1.3019	78	4.38	37411.72	
Fisuc_1527 FSU_2009	Membrane protein	Outer membrane	1.1721	163	5.09	78756.34	
Fisuc_1893	TPR repeat-containing protein	Unknown (multiple location)	1.1974	334	8.5	146703.86	
FSU_2396	OmpA family protein	Outer membrane	1.2168	143	4.86	55974.95	
Fisuc_1891	OmpA/MotB domain protein	Unknown (multiple location)	1.2168	143	4.74	53241.56	
Fisuc_2380 FSU_2934	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.2565	69	7.64	49323.02	
Fisuc_0417 FSU_0831	Putative uncharacterized protein	Outer membrane	1.2705	65	5.33	87192.47	
Fisuc_1465 FSU_1938	Extracellular ligand-binding receptor	Unknown (non-cytoplasmic)	1.2826	96	9.25	67212.86	
Fisuc_3112 FSU_0383	DSBA oxidoreductase (Putative outer membrane protein)	Unknown (non-cytoplasmic)	1.2977	142	8.58	27626.66	
Fisuc_0220 FSU_0627	Putative lipoprotein	Unknown	1.2983	152	9.54	32699.79	
Fisuc_0369	WD40 domain protein beta Propeller	Unknown (non-cytoplasmic)	1.3124	76	7.74	44076.06	
FSU_0784	Conserved domain protein	Unknown (non-cytoplasmic)	1.3124	76	7.03	45949.36	
Fisuc_0401 FSU_0816	Glycosyl hydrolase family 98	cytoplasmic membrane	1.3135	55	8.45	97071.66	CBM51
Fisuc_1571 FSU_2056	Outer membrane efflux protein	Outer membrane	1.3602	61	5.45	47008.86	
Fisuc_1597	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.3616	43	9.01	29528.42	
FSU_2084	Putative uncharacterized protein	Unknown	1.3616	43	9.01	29558.51	
Fisuc_1518 FSU_2000	Antigen-like protein	Unknown (non Cytoplasmic)	1.3724	42	5.43	86541.08	
Fisuc_2386	Integral membrane sensor hybrid histidine kinase	cytoplasmic membrane	1.3728	140	5.19	158187.80	
FSU_2943	Sensor histidine kinase/response regulator	cytoplasmic membrane	1.3728	140	5.25	159080.95	
Fisuc_0228 FSU_0636	30S ribosomal protein S20	Unknown	1.3846	59	10.86	9326.06	

Fisuc_0886 FSU_1333	Putative lipoprotein	Periplasm	1.3989	31	4.68	60005.03	
Fisuc_2613 FSU_3183	Conserved domain protein	Unknown (non-cytoplasmic)	1.4033	57	8.76	60589.08	
Fisuc_1230 FSU_1691	Extracellular solute-binding protein family 5	Unknown (multiple location)	1.4059	147	5.69	67554.14	
Fisuc_1203 FSU_1664	Putative uncharacterized protein	Unkonwn	1.4072	32	7.75	26030.60	
Fisuc_1530 FSU_2012	Chitinase	Unknown (non-cytoplasmic)	1.4096	50	4.42	37482.07	CBP9, GH18
Fisuc_0289 FSU_0701	Efflux transporter, RND family, MFP subunit	Unknown (non-cytoplasmic)	1.4168	76	9.45	37256.16	
Fisuc_2071 FSU_2599	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.4276	49	6.3	59200.25	
Fisuc_1525 FSU_2007	Cellulose-binding domain protein	Unknown (non-cytoplasmic)	1.4308	62	6.46	29200.50	CBM30
Fisuc_1894 FSU_2400	MotA/TolQ/ExbB proton channel	Cytoplasmic membrane	1.4323	87	9.14	23055.47	
Fisuc_2370 FSU_2924	Putative lipoprotein	Unknown	1.4402	37	5.37	16756.85	
surA Fisuc_0518 FSU_0941	Peptidyl-prolyl cis-trans isomerase SurA	Periplasm	1.4421	82	6.63	48242.47	
Fisuc_1898 FSU_2404	Membrane protein	Unknown (non-cytoplamsic)	1.4467	76	4.79	48258.76	
Fisuc_2965 FSU_0230	Membrane protein	Unknown (non-cytoplamsic)	1.4494	71	4.64	44916.19	
Fisuc_2509 FSU_3077	OmpA family protein	Outer membrane	1.4560	168	4.75	83917.26	
Fisuc_2227	Putative uncharacterized protein	Unknown (non Cytoplasmic)	1.4660	59	6.52	35518.63	
FSU_2769	Putative lipoprotein	Unknown (multiple location)	1.4660	59	6.52	35975.27	
Fisuc_0841 FSU_1285	Conserved domain protein	Unknown	1.4911	22	6.91	27296.18	
Fisuc_0025 FSU_0418	Alpha/beta hydrolase fold-3 domain protein	Unkonwn	1.4952	16	5.15	31189.18	
Fisuc_0249 FSU_0659	Putative uncharacterized protein	Unknown (multiple location)	1.5062	58	9.23	65289.10	
Fisuc_2069	Putative uncharacterized protein	Outer membrane	1.5093	48	4.73	42410.23	
FSU_2597	Putative uncharacterized protein	Outer membrane	1.5093	48	4.64	41786.39	
Fisuc_2732 FSU_3303	Rhodanese domain protein (Rhodanese-like protein)	Unknown (non Cytoplasmic)	1.5136	53	8.81	15785.50	
Fisuc_3015 FSU_0280	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.5142	51	5.41	37819.70	
nuoD Fisuc_2343	NADH-quinone oxidoreductase subunit D	Unknown	1.5306	37	6.68	43020.42	

Fisuc_3067 FSU_0333	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.5353	47	6.34	28052.26	
Fisuc_0433 FSU_0847	LysM domain protein	Unknown (non-cytoplasmic)	1.5393	37	8.84	43059.42	CBM50
Fisuc_0382 FSU_0797	Putative uncharacterized protein	cytoplasmic membrane	1.5431	53	5.02	37419.11	
Fisuc_1954 FSU_2474	Putative uncharacterized protein	Extracellular	1.5515	53	6.63	64462.39	
Fisuc_2041 FSU_2567	Putative type IV pilin	Unknown	1.5522	56	5.43	17516.89	
Fisuc_2147	Putative uncharacterized protein	Unknown	1.5587	77	6.34	24824.41	
FSU_2684	Putative uncharacterized protein	cytoplasmic membrane	1.5587	77	6.93	26421.32	
Fisuc_1247 FSU_1709	Putative lipoprotein	Unknown (non-cytoplasmic)	1.5647	40	4.75	66401.18	
Fisuc_1863	RDD domain containing protein	cytoplasmic membrane	1.5701	65	8.8	66169.31	
Fisuc_2917 FSU_0180	OmpA family protein	Outer membrane	1.5824	89	5.27	70794.57	
Fisuc_0081 FSU_0479	Putative uncharacterized protein	Unknown	1.5833	25	9.01	20617.89	
Fisuc_1651	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.6040	18	9.11	28289.21	
secG Fisuc_2367 FSU_2921	Preprotein translocase, SecG subunit	cytoplasmic membrane	1.6067	12	8.73	16955.93	
Fisuc_1642 FSU_2131	Diguanylate cyclase	cytoplasmic membrane	1.6145	52	7.66	50265.73	
Fisuc_0741	Putative uncharacterized protein	Extracellular	1.6215	52	4.74	44119.89	
FSU_1179	Conserved domain protein	Extracellular	1.6215	52	4.82	46207.38	
nuoI_1 nuoI nuoI2 FSU_2668	NADH-quinone oxidoreductase subunit I	Unknown	1.6466	15	6.3	20711.82	
Fisuc_3030	Putative uncharacterized protein	Unknown	1.6677	66	9.39	28986.61	Probably GH38
Fisuc_1756 FSU_2256	Peptidyl-prolyl cis-trans isomerase	Periplasm	1.6680	48	6.36	42840.21	
Fisuc_2475 FSU_3041	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.6852	17	6.08	21899.85	
Fisuc_2795 FSU_0049	Putative lipoprotein	Unknown (non-cytoplasmic)	1.6883	26	5.48	12637.95	
Fisuc_0482 FSU_0900	Putative uncharacterized protein	Outer membrane	1.6890	14	5.68	54535.86	
Fisuc_2300	Putative uncharacterized protein	Unknown	1.6923	6	4.67	26089.98	
Fisuc_0357 FSU_0772	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.6980	40	9.4	26700.92	
Fisuc_1021 FSU_1468	Putative lipoprotein	Unnkown (non-cytoplasmic)	1.7026	109	6.72	19714.23	
Fisuc_2905 FSU_0167	FKBP-type peptidyl-prolyl cis-trans isomerase domain protein	Periplasm	1.7209	153	8.83	48480.20	

Fisuc_3023 FSU_0288	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.7223	63	6.32	26091.54
Fisuc_2293 FSU_2840	Fibro-slime family protein	Unnkown (non-cytoplasmic)	1.7333	16	4.68	74180.59
Fisuc_1284 FSU_1752	Putative lipoprotein	Unknown (non-cytoplasmic)	1.7460	19	4.72	59406.80
Fisuc_1218	Putative uncharacterized protein	Unknown	1.7529	47	6.05	37566.90
FSU_1679	Putative uncharacterized protein	Unknown	1.7529	47	6.37	39672.47
FSU_0013	PPIC-type PPIASE domain protein	Outermembrane	1.7645	272	5.35	74360.87
Fisuc_1456 FSU_1929	LemA family protein	Unknown	1.7938	46	8.77	22275.67
Fisuc_3024 FSU_0289	Putative lipoprotein	Unkown	1.8072	48	4.86	31238.58
Fisuc_0331	Pentapeptide repeat protein	Extracellular	1.8397	244	9	47890.88
FSU_0746	Pentapeptide repeat domain protein	Extracellular	1.8397	244	9	48123.18
Fisuc_1592	OmpA/MotB domain protein	Outer membrane	1.8504	364	4.72	73841.20
FSU_2876	Putative uncharacterized protein	Unknown (multiple location)	1.8763	19	9.18	25727.55
Fisuc_0463	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.9281	33	6.62	42589.83
FSU_0881	Putative uncharacterized protein	Unknown (non-cytoplasmic)	1.9281	33	6.35	41950.03
fdrB Fisuc_2494 FSU_3062	4Fe-4S ferredoxin iron-sulfur binding domain protein	Cytoplasmic membrane	1.9358	75	6.52	28006.30
fdrA Fisuc_2493 FSU_3061	Succinate dehydrogenase or fumarate reductase	Cytoplasmic membrane	1.9413	201	6.93	70572.90
Fisuc_0178	Putative uncharacterized protein	Unnkown (non-cytoplasmic)	1.9705	21	5.16	31792.34
FSU_0584	Putative uncharacterized protein	Unknown	1.9705	21	5.12	34401.34
Fisuc_2752 FSU_0003	Putative lipoprotein	unknown (non-cytoplasmic)	1.9771	22	4.96	50091.38
FSU_3017	Putative uncharacterized protein	unknown	2.0101	25	4.9	39328.79
Fisuc_0455 FSU_0872	Conserved domain protein	Unknown (non-cytoplasmic)	2.0249	19	4.63	27372.49
Fisuc_0872 FSU_1318	Peptidyl-prolyl cis-trans isomerase	Outer membrane	2.0274	48	7.65	29667.95
Fisuc_1485 FSU_1966	Putative uncharacterized protein	Unknown (non-cytoplasmic)	2.0441	30	8.98	21636.80
Fisuc_1878 FSU_2382	Putative lipoprotein	Unknown	2.0734	18	4.7	19189.52
Fisuc_1877	Putative uncharacterized protein	Unknown (non-cytoplasmic)	2.0895	11	4.4	33132.57

Fisuc_2892	OmpA/MotB domain protein	Outer membrane	2.0916	73	5.53	32218.71
Fisuc_3033 FSU_0298	Mechanosensitive ion channel family protein	cytoplasmic membrane	2.1914	58	6.32	29387.64
Fisuc_1008 FSU_1454	BatB protein	cytoplasmic membrane	2.1986	29	9.54	38346.69
Fisuc_0752 FSU_1190	Putative lipoprotein	Unknown	2.3211	49	4.67	22029.48
Fisuc_0722	Peptidase M16 domain protein	Unnkown (non-cytoplasmic)	2.3314	37	8.94	55299.51
FSU_1159	Peptidase, M16 family	Unknown (non-cytoplasmic)	2.3314	37	8.94	55604.86
Fisuc_0657 FSU_1088	Putative lipoprotein	Unknown (non-cytoplasmic)	2.3861	33	4.56	34721.24
Fisuc_2262 FSU_2807	Membrane protein	Unknown	2.4590	70	6.92	38496.51
Fisuc_0775 FSU_1216	Polysaccharide biosynthesis/export protein	Outer membrane	2.5195	101	5.74	41486.58
Fisuc_2308	Putative uncharacterized protein	Unknown (non-cytoplasmic)	2.6323	22	4.62	15281.10
Fisuc_1660	Putative uncharacterized protein	Unkown	2.8948	11	4.83	32623.58
FSU_2149	Putative uncharacterized protein	Unknown	2.8948	11	4.91	35374.94
Fisuc_0202 FSU_0609	Lipoprotein (Putative lipoprotein)	Unknown (non-cytoplasmic)	3.1529	66	5.01	19759.43
Fisuc_1316 FSU_1783	Membrane protein (PEGA domain protein)	Unknown (non-cytoplasmic)	3.5892	29	8.83	20237.14
Fisuc_1317 FSU_1784	Putative lipoprotein	Unknown (non-cytoplasmic)	4.0268	43	5.18	22441.30
Fisuc_2974 FSU_0239	PPIC-type PPIASE domain protein	Unknown (non-cytoplasmic)	4.4198	105	9.5	36986.91