Nitrogen metabolism and butanol production by South African *Clostridium beijerinckii* and *Clostridium saccharobutylicum* strains

Byron W.P. Reeve

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Molecular and Cell Biology

UNIVERSITY OF CAPE TOWN

March 2014

Acknowledgements

Putting all the thesis ideas in my head into words was a difficult task. This project has been my most important undertaking as a scientist to date. The skillset I learned in this project is invaluable. I hope to continue to grow as a scientist and be involved ABE fermentation and biofuels to a greater or lesser extent throughout my career as a scientist. I'd like to thank my supervisor, Associate Professor Shez Reid, for approaching me with the proposal for this project. I really could not ask for a better supervisor. Every Monday at 11:00 we sat down and discussed the progress of the week before. Every week there were triumphs and setbacks in the project which you took with equal grace, advice and encouragement. I'm particularly grateful to you for keen eye in helping me in all aspects of my scientific writing. I am grateful to Green Biologics Ltd., UK for funding myself and this project. A special thanks to Dr Edward Green for the positive feedback and support regarding the various progress reports. Thanks to, Dr Renia Gemmell and Dr Preben Krabben for their assistance with strain taxonomy and bioinformatics respectively, and Dr Nigel Minton for allowing me to learn how to use the ClosTron system at the University of Nottingham.

Personal thanks to my best friends, Paul Schwartz and Karl Thomson for the laughs and countless distractions. Greg Bowden, Kevin O'Connell, Bridget Price, you made life outside MCB possible. A special thanks to Dr Bruna Galvao for the jokes and the advice in all matters, Dr Samantha "Stats Queen" Nicholson for the many lame jokes and advice with the filthy statistics, Ailsa "There's a spelling mistake" Davidson for being willing to read a chapter. Thanks to, Dr Brian Kullin and Tuna Brock, for being such great bench buddies. Thanks to Associate Professor Valerie Abratt and all Anaerobe Unit members, past and present, who made my postgrad lab experience a pleasure. Thanks to my honours students, Tarryn-May Terry and Jun Mao, for their contributions to the MLST and acid survival sections of this project respectively. A special thank you to (soon to be) Dr Rosemary Meggersee, my good friend and sanity compass. We started undergrad together and it seems only fitting that we end our PhD journeys together.

I'd like to thank my family for all their love and support. Life was tough after the Reeve Clan moved away to Johannesburg in 2006. Fortunately you all made me feel like home was only a phone call or a flight away. To my younger brothers, Rowan and Lyndon thanks for always taking my mind off science and keeping me up to date in all things gaming. To my mother, Karen Reeve, I like to think I inherited my love for science from you. Thank you for always encouraging me to keep going, for all the advice over the many slow years of my bachelor life domestication and making me always feel excited to come home wherever I am in the world! You are the benchmark to which I compare all mothers. I have yet to meet your equal. To my father, Adrian Reeve, thank you for all the many conversations about work and science. When I grow up (although I wouldn't hold my breath), I can only hope to be a quarter the man you are. I am unashamed to say you are my role model and hero. A real life Batman.

Being a fulltime student is a costly affair. I am grateful to the UCT Research committee and the National Research Foundation for funding provided during my project.

Contents

Abstract

The acetone- butanol-ethanol (ABE) fermentation was one of the first fermentation processes to be industrialized on a large scale, and the dominant product, butanol is particularly significant due to its potential as a modern day fuel additive or fuel extender in the petrochemical industry.

A collection of 19 solventogenic *Clostridium beijerinckii* and 11 *Clostridium saccharobutylicum* strains isolated from the National Chemical Products (NCP) ABE fermentation plant in Germiston, South Africa, were classed according to species by a quick species-specific colony PCR and by rifampicin screening methods respectively. The species-specific PCR aims to provide a rapid means of assessing any contamination of an ABE batch fermentation by differentiating between *C. saccharobutylicum* and *C. beijerinckii* species. Random Amplification of Polymorphic DNA (RAPD) analysis generated four *C. beijerinckii* and two *C. saccharobutylicum* strain groups respectively. Multilocus Sequence Typing (MLST) was developed for a smaller selection of strains and showed a further two strain groups within the NCP *C. beijerinckii* strains and three groups within the *C. saccharobutylicum* strains. All taxonomic analysis performed in this study indicates that the NCP *C. beijerinckii* strains are genetically distinct from NCIMB 8052^{T}.

ABE solvent analysis for the strain collection on glucose, sucrose and xylose fermentation substrates revealed a wide diversity of solvent profiles for the NCP strain collection and a number of strains within each species which generate high yields of butanol for each of these substrate conditions.

The effect of supplementation of fermentation medium with the amino acids glutamate, glutamine, proline, lysine, histidine and asparagine was investigated. Glutamine, glutamate and histidine were shown to significantly increase butanol yields ($p < 0.05$, one-way ANOVA

Least Statistical Difference analysis). Acid shock studies determined that glutamate and histidine are involved in the Acid Tolerance Response (ATR), which may be crucial for the organisms to survive acidogenesis in order to produce solvents. These findings were validated through quantitative RT-PCR experiments where it was demonstrated that exposure to acidic conditions, like those found during acidogenesis, up-regulated the glutamate synthase (*gltA*) gene, which is responsible for producing endogenous glutamate, and its regulator (*nitR*), 2-fold compared to neutral conditions ($p < 0.05$ Pfaffl statistical analysis).

Bioinformatic analysis of the *C. beijerinckii* NCIMB 8052^T genome identified a number of putative nitrogen metabolism genes. Six of these genes were selected for gene inactivation using the ClostTron system. The putative gene targets consisted of two ammonium transporters, two PII –like global nitrogen metabolism regulators and a gene cluster containing a glutamine transporter and a histidine kinase which may interact with NitR. The plasmids for gene inactivation were all successfully conjugated into *C. beijerinckii* NCP260 and NCIMB 8052^T, however, integration of the ClosTron plasmids failed and no stable mutants were generated.

The findings of this research indicate that the UCT/NCP strains are phylogenetically closely related and consist of many robust strains able to produce high levels of solvents under a variety of fermentation conditions. A number of these South African strains show potential as specialist biobutanol producers in the current international ABE fermentation industry. In addition, supplementation of the fermentation substrates with glutamine, glutamate or histidine promotes better solvent yields, possibly because the amino acids offer acid shock protection.

Abbreviations

% Percentage

°C Degree(s) Celsius

aa Amino acid(s)

ABE Acetone-butanol-ethanol

Asn Asparagine

asRNA antisense RNA

ATP Adenosine triphosphate

atpD Gene encoding ATPase D

ATR Acid tolerance response

bp Base pair(s)

CBM Clostridial basal medium

cDNA Complementary DNA

CSC Commercial Solvents Corporation

DNA Deoxyribonucleic acid

dNTP Deoxyribonucleotide triphosphate

dsDNA Double stranded DNA

Erm Erythromycin

et al. Et alia

etfA Gene encoding electron transport flavoprotein alpha

g Gram(s)

g Relative centrifugal force

GC-MS Gas chromatography–mass spectrometry

gDNA Genomic DNA

Gln Glutamine

glnA Gene encoding glutamine synthase

gltAB Genes encoding glutamate synthetase subunits A and B

Glu Glutamate

GOGAT Glutamate synthase

GS Glutamine synthetase

GSMM Glucose salts mineral media

gyrB Gene encoding gyrase B

h Hour(s)

His Histidine

kb Kilobase pairs

KD Knock-down

KO Knock-out

L Litre(s)

LB Luria Bertani

Lys Lysine

M Molar

mg Milligram(s)

min Minute(s)

mL Millilitre(s)

MLST Multi locus sequence typing

mM Millimolar

n Number of samples

NAD(P)/NAD(P)H Nicotinamide adenine dinucleotide (phosphate)/reduced

NCBI National Centre for Biotechnology Information

NCP National Chemical Products

ng Nanogram(s)

nitR Gene encoding nitrogen regulator

ng Nanogram(s)

OD Optical density

OD_{600} Optical density at 600nm

oppB Gene encoding oligopeptide permease B

ori Origin of replication

p p-value; indicating probability

PCR Polymerase chain reaction

PFGE Pulse field gel elecrophoresis

pH Potential Hydrogen

Pro Proline

qRT-PCR Quantitative RT-PCR

RAPD Random amplification of polymorphic DNA

RCM Reinforced *Clostridium* medium

recA Gene encoding recombinase A

RFLP Restriction fragment length polymorphism

RNA ribonucleic acid

RNaseA Ribonuclease Alpha

rpoB Gene encoding RNA polymerase B

rRNA Ribosomal RNA

RT-PCR Reverse transcriptase PCR

s Second(s)

SNP Single nucleotide polymorphism

Spec Spectinomycin

ssDNA Single stranded DNA

Tet Tetracycline

U Enzymatic units

UCT University of Cape Town

UV Ultraviolet light

V Volts

w/ v Weight per volume

ʎ Lambda

µg Micro gram(s)

µL Micro litres(s)

µM Micromolar

Chapter One

Literature Review

Contents

1.1 Introduction to Acetone Butanol Ethanol (ABE) fermentation

The acetone- butanol-ethanol or ABE fermentation was one of the first fermentation processes to be industrialized on a large scale. The process involves the breakdown of a sugar rich substrate anaerobically to produce acetic and butyric acids by solventogenic members of the *Clostridium* genus in early log growth phase. These metabolic acids are converted to acetone, butanol and ethanol during stationary phase (Jones and Woods, 1986).

1.1.1 The history of ABE fermentation and the scientific contributions by the genus *Clostridium*

The production of butanol by microbial fermentation was first documented by Louis Pasteur in 1861 (McCutchan and Hickey, 1954). ABE fermentation was brought to the fore by Chaim Weizmann, who isolated and studied bacterial cultures capable of producing acetone and butanol. In 1910, Weizmann joined Strange and Graham Ltd. and began investigations into the production of butanol and isoamylalcohol by fermentative microbes in the hope of finding a way to create synthetic rubber (Morris, 1993). In 1912, Weizmann discovered a starch fermenting microbe which was designated strain BY. This strain was able to ferment starch in the form of maize mash to produce acetone. Strain BY was later named *Clostridium acetobutylicum* and became the organism under which the ABE fermentation process patent was formed. These early strains were all designated the genus and species name, *Clostridium acetobutylicum*, on the basis of being Gram-positive, anaerobic, rod-shaped spore-formers (Jones and Woods, 1986). The industry was most crucial during World War One (WWI) from 1914 when ABE fermentation was used as a means to produce acetone used in the synthesis of cordite which is the main component in the manufacture of munitions (Gabriel, 1928). After WWI, Weizmann's scientific contributions to the war effort resulted in the gain of

much influence with prominent members of state in the UK, culminating in the Balfour declaration. With the support of the British government, the State of Israel was founded in 1948 and Weizmann was made its first president.

After WWI, the patent for the ABE fermentation process was acquired by Commercial Solvents Corporation based in the USA (Gabriel, 1928). In 1922, the fermentation of a variety of carbohydrates by a strain isolated from American barley that exhibited similar solvent-producing properties to that of the Weismann strains was investigated (Robinson, 1922). Robinson concluded that this solventogenic strain was able to fully utilize glucose, fructose, mannose, sucrose, lactose, starch and dextrin in fermentations to produce acetone and butanol. Xylose, galactose, raffinose and arabinose were only partially utilized. Later, a comparison of 11 solventogenic strains able to utilize starch as a fermentation substrate was performed. It was found that there were only slight phenotypic variations. These 11 strains were likewise designated *C. acetobutylicum* (McCoy *et al.*, 1926). A similar study was conducted by Weyer and Rettger (1927) using phenotypically similar strains of *C. acetobutylicum* to ferment corn mash as a substrate.

In the 1930s it was demonstrated that ABE fermentation could be performed on sugar cane molasses as a substrate. A shift from starch based fermentation to sugar based fermentation occurred due to cheaper substrate costs and fewer difficulties processing sugar vs starch substrates. New strains were isolated for this purpose with higher yields reported, which saved on distillation costs (Hastings, 1978). In 1935, the patent on ABE fermentation expired and the process became available globally. Many new strains were isolated, developed and patented by companies; however, strain cataloguing was not systematic since there was no

knowledge of the taxonomic relationship between industrial solventogenic strains (Jones and Woods, 1986).

ABE fermentation notwithstanding, the research into biochemical aspects of the *Clostridium* genus persisted since the diversity and anaerobic metabolism of the Clostridia meant that there were many research inquiries into various aspects of classic anaerobic biochemistry. In 1934, the Stickland reaction was first discovered and documented among certain species of *Clostridium* (Stickland, 1934). The Stickland reaction is the fermentation of amino acids whereby one amino acid is oxidized and another amino acid is reduced (Nisman, 1954). Energy is conserved through adenosine triphosphate (ATP) formation during substrate-level phosphorylation (Fonknechten *et al.*, 2010). Co-factor work done using *C. kluyveri* by Stadtman in the 1960s led to our understanding of how the energy transfer molecule acetyl-CoA is formed (Fredericks and Stadtman, 1965). Nitrogen fixation, whereby atmospheric nitrogen is reduced to organic nitrogen compounds, was first demonstrated in a cell-free extract of *Clostridium pasteurianum* containing the nitrogenase enzyme and prompted further biochemical investigations into its mechanism of action (Carnahan *et al.*, 1960). Experiments using *C. kluyveri* revealed that the transhydrogenation from NADPH to NAD^+ is stimulated and inhibited by oxidized and reduced ferredoxin respectively (Thauer *et al.*, 1971). These discoveries, amongst others, have directed the course of study in understanding the biochemistry behind electron transport and anaerobic nitrogen metabolism within not only the genus, but other prokaryotes as well.

ABE fermentation was the major supplier of the world's solvents until the 1950s and 1960s. ABE fermentation facilities of industrial, semi-industrial and pilot scale plants were reported in China, Japan, USA, Canada, UK, France and the former USSR (Nimcevic and Gapes,

2000). When demand for acetone dropped after the world wars, ABE fermentation persisted. The predominant ABE product, 1-butanol (or n-butanol) was used in lacquers in the paint industry, aided by its use in the booming automobile paint industry that developed post-WWII. The ABE industry fell into decline after the 1960s when the oil industry expanded rapidly. Butanol is a by-product of the oil-refining process and as such, the ABE derived butanol prices could not compete with the petrochemical industry. In South Africa, the National Chemical Products (NCP) company operated one of the largest and most successful industrial ABE fermentation plants from the 1930s to the early 1980s (Keis *et al.*, 1995). This facility likewise had to be shut down due to a combination of the increase in the price of sugar cane substrate and competition from the petrochemical industry (Jones and Woods 1986). Since the 1980s, there has been only sporadic interest in ABE fermentation at an industrial scale.

In 1997, DNA-DNA hybridization experiments proved the existence of four genetically distinct strain species, including *C. acetobutylicum*. It was suggested that some strains from industrial strain collections designated as "*C. acetobutylicum*" may in fact belong to other species (Johnson *et al.*, 1997), and in 2001, it was confirmed that many strains previously designated *C. acetobutylicum* should be designated *C. acetobutylicum, Clostridium saccharoperbutylacetonicum, Clostridium beijerinckii* and *Clostridium saccharobutylicum.* Included in this study, were strains from the NCP strain collection which were also previously designated *C. acetobutylicum* (Keis, *et al.*, 2001). The major taxonomic revisal and current taxonomic advances related to industrial strain collections are the focus of Chapter Two.

Today, there is a need to be less reliant on limited fossil fuels and to lower greenhouse gas emissions. Renewable biofuels like bio-based butanol or biobutanol hold promise to supplement or replace finite fossil fuels. In 2006, China re-opened ten of its ABE plants with a reported total solvent production capacity of 210,000 tonnes from 2006 to the end of 2008 (Ni and Sun, 2009). In 2011, Brazil was confirmed to have an ABE plant that uses sugar cane juice to produce 8000 tonnes of solvents per annum (Green, 2011).

1.1.2 Biobutanol applications as a modern biofuel and the current limitations of ABE industry

The need to be less dependent on crude oil and other fossil fuels as well as the negative impact the processing and use of these fuels has on the environment is undeniable. The use of renewable biofuels to supplement existing fuel reserves has gained attraction particularly in the well-established bioethanol field with the USA and Brazil being world leaders in bioethanol production for blending into existing petroleum (Balat and Balat, 2009; Zhou and Thomson, 2009). The term "renewable" is particularly relevant since the biomass required for creation of these fuels originates most often from plant material that can be re-grown and regenerated. During the previous century, ABE fermentation by clostridia was the second largest fermentation process next to ethanol fermentation by yeast (Antoni *et al.*, 2007).The products of ABE fermentation are of commercial value as solvents used in both laboratory and industrial settings. The dominant product, butanol is particularly significant not only due to its potential as a modern day fuel additive or fuel extender in the petrochemical industry (Karakashev *et al.*, 2007), but also in the plastics industry (Parekh and Blaschek, 1999).

The generation of biobutanol by ABE fermentation has been extensively reviewed (Dürre, 2008; Dürre, 2011; Green, 2011; García *et al.*, 2011; Ezeji *et al.*, 2010). Butanol, compared to other biofuels like ethanol, is an ideal fuel additive with a higher energy output and mixes easily with petroleum based fuels due to its hydrophobic nature. In addition, butanol is less volatile with a higher flash point and lower vapour pressure making isolation, storage and transport of butanol safer than ethanol as a biofuel to support the ever increasing fuel demands of today (Dürre, 2011). Chemically modified butanol from ABE fermentation is proposed to be blended into or create existing diesel (Antoni *et al.*, 2007; Anbarasan *et al.*, 2012). Acetone, ethanol, H_2 and CO_2 gasses generated in the fermentation process can be separated and harvested for sustained commercial use as well (Qureshi and Blaschek, 2000). In 2005, the North American chemical company, Dupont, and the UK petroleum company, BP, declared joint interest in restarting industrial-scale ABE fermentation with the vision of bringing biobutanol into the petrochemical industry (Green, 2011).

There are currently a number of limitations that hamper the present-day commercial application of ABE fermentation which include:

1. The use of high starch or sugar crops as substrate
2. Cost of the substrate
3. Low yield of solvents and solvent toxicity
4. Expensive solvent recovery inputs

1: The use of high energy crops such as maize and wheat as a source of fermentation substrate is a point of contention, because the increasing global population necessitates adequate food supply. This food vs. fuel debate naturally favours food for domestic consumption by the world populous (Dürre, 2008). The use of high demand food crops for

first generation biofuel production is therefore not viable. There are at present, areas in the world where there is a surplus of readily available starch crops that can be used for ABE fermentation. A historical example of this is the Corn Belt in the USA that supplied the Commercial Solvent Corporation with substrate for ABE fermentation (Kumar and Gayen, 2011). Another example is the use of packing peanuts and palm sago containing starch as a substrate (Ezeji *et al.*, 2003; Liew *et al.*, 2006). However, there are some inherent issues with starch fermentation performance such as gelatinization during heat sterilization of starch which results in oligomers during cooling which are more difficult for amylases to break down into accessible simple sugars (Ezeji *et al.*, 2005). More importantly, there is the desire to move away from sugar based "food crop" substrates as is the case of first generation biofuels and transition to second generation biofuels in the form of non-food crop substrates, generally in the form of cheaper cellulosic plant materials.

2: The cost of the fermentation substrate can account for up to 60% of the cost of biofuel production (Demirbas, 2009). Naturally, the best course of action is to search for cost-effective alternative substrates that do not conflict with global food interests, are in consistent surplus or are deemed waste products. It has also been suggested that low cost waste materials such as domestic organic waste and plant material containing lignocellulose and hemicellulose be used as cost effective substrates (Papoutsakis, 2008). The use of agricultural waste material in industrial-scale ABE fermentation facilities in the former Soviet Union which operated until the late 1980s supports this concept (Zverlov *et al.*, 2006). Research into the ABE fermentation performances utilizing lignocellulosic substrates has been conducted. Cornstover and switchgrass hydrolysates have been used as ABE fermentation substrates. Cornstover consists of the remnant corn plant material after crop harvest and switch grass is a tall grass that grows predominantly in the USA that has been identified as a primary source of

biomass for biofuel production in the region. *C. beijerinckii* NCP260 was able to generate total solvent yields of 0.43 (grams solvent/ grams sugar) and 0.39 for cornstover and switchgrass hydrolysates pre-treated with lime respectively (Qureshi *et al.*, 2010b). The same research group used strain NCP260 to ferment barley straw hydrolysate pre-treated with both dilute sulphuric acid and lime to generate a total solvent yield of 0.43 (Qureshi *et al.*, 2010a). Rice bran pre-treated with commercial α-amylase and dilute hydrochloric acid was fermented by *C. beijerinckii* NCIMB 8052^{T} to generate a total solvent yield of 0.35 (Lee, 2009).

Alternative substrates can be found in already established industries. For instance, the use of the fibrous material known as bagasse which remains after sugar cane, cassava roots or sorghum cereal has been milled to extract the sugar-rich juices. All of these plants are grown in tropical and sub-tropical areas around the world. Sugar cane bagasse is considered waste material post-sugar refining processing and since it is linked in the sugar industry, there is a constant supply of sugar cane bagasse which can be used in ABE fermentation. Sorghum bagasse is primarily used as a low cost cattle fodder. The bagasse left over post-harvest is likewise an attractive low-cost substrate closely tied to an existing market. A number of researchers have investigated the use of bagasse as a fermentation substrate. Bagasse fermentation is often reliant on acid pre-treatment to liberate the sugars contained in the stalk material. Bagasse obtained from sorghum and cassava has been used as a fermentation substrate with solvent yields comparable to sugar-based fermentations (Yu *et al.*, 2012; Lu *et al.*, 2012). Wood pulp residue left over from paper pulping industry has been shown to be rich in the sugars necessary for ABE fermentation. A mutant *C. beijerinckii* strain was used in this process and a total solvent yield of 0.42 was achieved (Lu *et al.*, 2013a).

Alternative substrates such as seaweed have been considered for use in ABE fermentation. Algae are a fast-growing abundant source of biomass; however, costly commercial cellulases and high temperature substrate pre-treatments are required for the best fermentation yields although seaweed as a substrate does not require any additional nutrient supplementation to support ABE fermentation (van der Wal *et al.*, 2013). They are deemed attractive substrates, because they do not compete with terrestrial agricultural land space. Wastewater algae harvested from a waste water treatment facility was used as fermentation substrate that did not require additional nutrient supplementation (Efremenko *et al.*, 2012). With the aid of commercial xylanases and cellulases, *Clostridium saccharoperbutylacetonicum* was able to generate a total ABE solvent yield of 0.311 (Ellis *et al.*, 2012).

Inhibitors present in substrate hydrolysates contribute to low growth and solvent yields. Removal of fermentation inhibitors must be addressed in the context of process engineering. A study by Ezeji and colleagues on basic liquefied corn starch revealed that while it contained a multitude of sugars in the form of cellobiose, glucose, mannose, arabinose and xylose available for fermentation, there were also potent inhibitory phenolic compounds present like p-coumaric and ferulic acids. These inhibitory compounds were effective in concentrations as low as 0.3 g/L (Ezeji *et al.*, 2007). Lignocellulosic material is typically pre-treated with peroxidase in order to detoxify the substrate of phenolic substances before use in fermentation (Cho *et al.*, 2009).

The alternative substrates mentioned require expensive pre-treatments in the form of enzymatic catalysts to make fermentable sugars accessible, chemical hydrolysis to break down plant polymers into sugars or the addition of substances that remove fermentation inhibitors from the hydrolysate. Lignocellulose as a substrate is cheap, but the cost reduction

is offset by expensive substrate pre-treatment and detoxification. Strategies to improve second generation biofuel fermentation include: improve cellulose hydrolysing organisms and enzymes, improve the biomass pre-treatment process, breed/engineer biomass that is more compatible with hydrolysis and which produces less inhibitory compounds and finally, develop microbes that are more tolerant to inhibitory compounds. Substrate processing will remain a major impediment to the commercialization of ABE fermentation until the substrate treatment process is improved both functionally and economically.

3: The present low yield of solvents, particularly butanol, is a major obstacle which hinders commercial production of biobutanol via ABE fermentation. The low yield arises primarily due to the fact that the solvents produced by *Clostridium* become toxic to the organism at 1.5-2.0% concentration. Clostridia exhibit slow growth rates which results in low maximum cell densities and low solvents produced (Ezeji *et al.,* 2010). There are a number of ways of manipulating the metabolism and genetics of these organisms to increase solvent titres and solvent tolerance which will be discussed later in the review, however, the most obvious physical methods of improving solvent yield are to provide optimal reactor conditions to attain maximum cell densities. Linked to reactor conditions is the removal of the solvents before they have a chance to accumulate to toxic concentrations as is highlighted in the next section 4.

A high density continuous culture fermentation system is a more desirable form of reactor fermentation compared to traditional batch and fed batch fermentations. Continuous fermentation generally has improved solvent yields and productivities (Lee *et al.,* 2008). Various reactor designs have been implemented to achieve high cell densities including biofilm-based and membrane cell recycling bioreactors. The biofilm-based reactor uses

support structures, also known as absorbents, which form aggregates or granules of cells. These aggregated cells eventually form a protective biofilm network and generate high cell densities and increased solvent productivities (Qureshi *et al.*, 2005a). A subtype of the biofilm reactor is the plug flow reactor. The core reactor is a type of packed-bed reactor whereby the reactor is partitioned into different cylindrical compartments or plugs. The area to volume ratio is optimised so that there is maximum nutrient medium flow and therefore high cell densities result (Lienhardt *et al.*, 2002). Membrane cell recycling bioreactors present a means of concentrating cells before continuous culture. Cells are allowed to grow in broth and are collected and concentrated by membrane filtration. If the cell densities get too high, excess cells are removed by cell bleeding and cell densities replenished by cell recycling. When the cell filtrate reaches the desired density, continuous culturing with cell membrane filters takes place. Research conducted with this type of reactor reported a total solvent productivity of 7.55 g/L/h with a total solvent concentration of 8.58 g/L for a period of 200 h without cell degeneration (Tashiro *et al.*, 2005).

4: Expensive solvent recovery inputs tip the financial scales towards unfeasible when low solvent yields are combined with inefficient solvent recovery methods. The preferred method of solvent recovery in the last century was by distillation. This method is not economic since butanol has a higher boiling temperature than water and therefore the temperatures required to evaporate butanol result in high energy costs involved in the separation and distillation of butanol. Several promising techniques for in situ fermentation solvent recovery have been developed, namely pervaporation, vacuum recovery, gas-stripping and absorption. This results in greater cell viability and increases solvent yield. The removal of toxic butanol solvent products bypasses the destructive effects on the cell membrane and membrane proteins experienced at high butanol concentrations (Bowles and Ellefson, 1985).

Pervaporation keeps solvent levels low through the use of a selective membrane to collect the desired solvents that are partially vaporised and transported through the membrane under a vacuum (Izak *et al.*, 2008). A recent study reported the coupling of a polydimethylsiloxane (PDMS) membrane pervaporator to a continuous culture bioreactor. The pervaporation occurred intermittently in the first phase and then continuously in the later phase. *C. acetobutylicum* was able to generate a butanol titre and productivity of 28.03 g/L and 0.105 g/L/h respectively in the first phase and 61.43 g/L and 0.205 g/L/h respectively in the second phase (Chen *et al.*, 2013). Vacuum recovery is similar to regular solvent distillation methods employed in the previous century, but uses a vacuum pump to decrease the boiling temperature required for solvent evaporation. The solvent vapours are condensed and collected as usual (Jang *et al.*, 2012). A study reported ABE batch fermentation productivity and yield of 0.26 g/L/h and 0.35 g/g respectively using *C. beijerinckii* in a continuous and intermittent vacuum reactor system (Mariano *et al.*, 2011).

Gas-stripping is a solvent recovery process whereby H_2, CO_2 or N_2 carrier gasses are channelled through a bioreactor sparger which creates gas bubbles in the ABE reactor. The solvents vaporise and are liquefied when they pass through a condenser. The carrier gas can be recycled. This method is attractive since carbon dioxide and hydrogen gasses are natural products of ABE fermentation (Jang *et al.*, 2012). Genetically modified *C. acetobutylicum* was able to produce 35.6 g/L isopropanol-butanol-ethanol in 45 h with the aid of gas stripping to remove inhibitory solvent products (Lee *et al.*, 2011). A fed-batch study that incorporated gas-stripping reported 108.5 g/L total solvents produced by *C. acetobutylicum* with a total solvent yield and productivity of 0.22 g/g and 0.29 g/L/h respectively over a 263 h period (Lu *et al.*, 2012). Another study reported *C. acetobutylicum* butanol titres of 113.2

g/L and total solvent titres of 172 g/L with a solvent yield and productivity of 0.36 g/g and 0.53 g/L/h respectively over 326 h (Xue *et al.*, 2012). It is possible combine pervaporation and gas-stripping in the recovery process (Setlhaku *et al.*, 2013) however, it is thought that gas-stripping is more energy-efficient than pervaporation for commercial application (Ezeji *et al.*, 2004; Qureshi *et al.*, 2005b).

Absorption is a method of concentrating selected solvents from an aqueous solution. The absorbent is usually in the form of bead-like polymeric resins. The absorbent containing the desired super-concentrated solvent is heated with a much lower energy output and distilled, separated and condensed as normal. Silicalite absorption columns were used concentrate butanol in fermentation vessels in the range of 5–810 g/L of butanol. This technique is thought to be the most energy- and cost-effective method of butanol separation and collection (Qureshi *et al.*, 2005b). Dowex® Optipore SD-2 resin was used to absorb butanol in ABE batch fermentation vessels. *C. acetobutylicum* was able to produce high titres of 22.2 g/L butanol, above the typical inhibitory threshold of 2% and the solvent recovery was demonstrated to be cost-efficient and resins were reused without loss of butanol affinity (Nielsen and Prather, 2009).

In summary, the most urgent requirements for commercial application of ABE fermentation are decreasing pre-treatment costs and developing an energy-efficient method to continually remove toxic butanol as it is produced during ABE fermentation. Once these general criteria are met, strains can be selected based on the best butanol productivity for a particular substrate. Genetic and metabolic engineering can then be employed to further boost solvent yield and productivity.

1.2. ABE solvent metabolism

Solvent formation in solventogenic *Clostridium* is typically accomplished in two phases, namely acidogenesis and solventogenesis. Acidogenesis typically involves the breakdown of intracellular sugars like glucose into acetic and butyric acids which generates ATP essential for normal anaerobic cellular respiration and growth. Acidogenesis takes place during early- to mid-exponential phase and is characterised by a drop in intracellular and extracellular pH. Cell death becomes imminent when the pH begins to drop below pH 4.5. To counteract the drop in pH, the metabolic acids produced in the form of acetic and butyric acids are re-taken up from the extracellular environment and assimilated into the neutral pH solvents, acetone and butanol in late exponential and stationary phases. Parallel to the conversion of these acids to solvents, the cell prepares for sporulation which is a safety mechanism to escape the accumulating toxic effects of acetone and specifically butanol accumulation (Jones *et al.*, 2008).

1.2.1 Metabolic pathways involved during acidogenesis and solventogenesis

The metabolic pathways for acidogenesis and the subsequent solventogenesis are summarized in Figure 1.1 below. In this example, acidogenesis begins with the uptake of C-6 glucose which is broken down into two C-3 molecules of pyruvate via glycolysis to produce two molecules of ATP and NADH per molecule of glucose. Each molecule of pyruvate is then converted to acetyl-CoA by pyruvate-ferredoxinoxidoreductase (PFOR) which releases CO_2 and electrons are transferred via the ferredoxin pathway via hydrogenase (HYDA) with H_2 being the final electron acceptor. Acetyl-CoA can evolve along one of two metabolic pathways: the acetate or the butyrate pathway. In the acetate pathway, acetyl-CoA is converted to acetyl- phosphate by phosphotransacetylase (PTA). The acetyl- phosphate is

then dephosphorylated by acetate kinase (AK) which generates acetate and ATP. In the butyrate pathway, acetyl-CoA is converted to acetoacetyl CoA by thiolase (THL), the acetoacetyl-CoA is reduced to 3-hydroxybutyryl-CoA at the cost of NADH by β-hydroxybutyryl dehydrogenase (BHBD). 3-hydroxybutyryl-CoA is then hydrolysed to crotonyl-CoA by crotonase (CRO); the crotonyl-CoA is reduced to butyryl-CoA by butyryl-CoA dehydrogenase (BCD) which consumes NADH. Butyryl-CoA is converted to butyryl-phosphate by phosphotransbutyrylase (PTB) and converted to butyrate by butyrate kinase (BK) which generates ATP (Papoutsakis, 2008). Acetate and butyrate accumulate intracellularly and are secreted into the extracellular environment and dissociate to form acidic anions with the result that the intracellular and extracellular pH decreases proportionally to the amount of acids produced. Acidogenesis is thus the primary means of generating intracellular pools of ATP, NADH and NADPH. The redox equilibrium is maintained during fermentation with NAD^+ and NADP being regenerated through the formation of hydrogen gas.

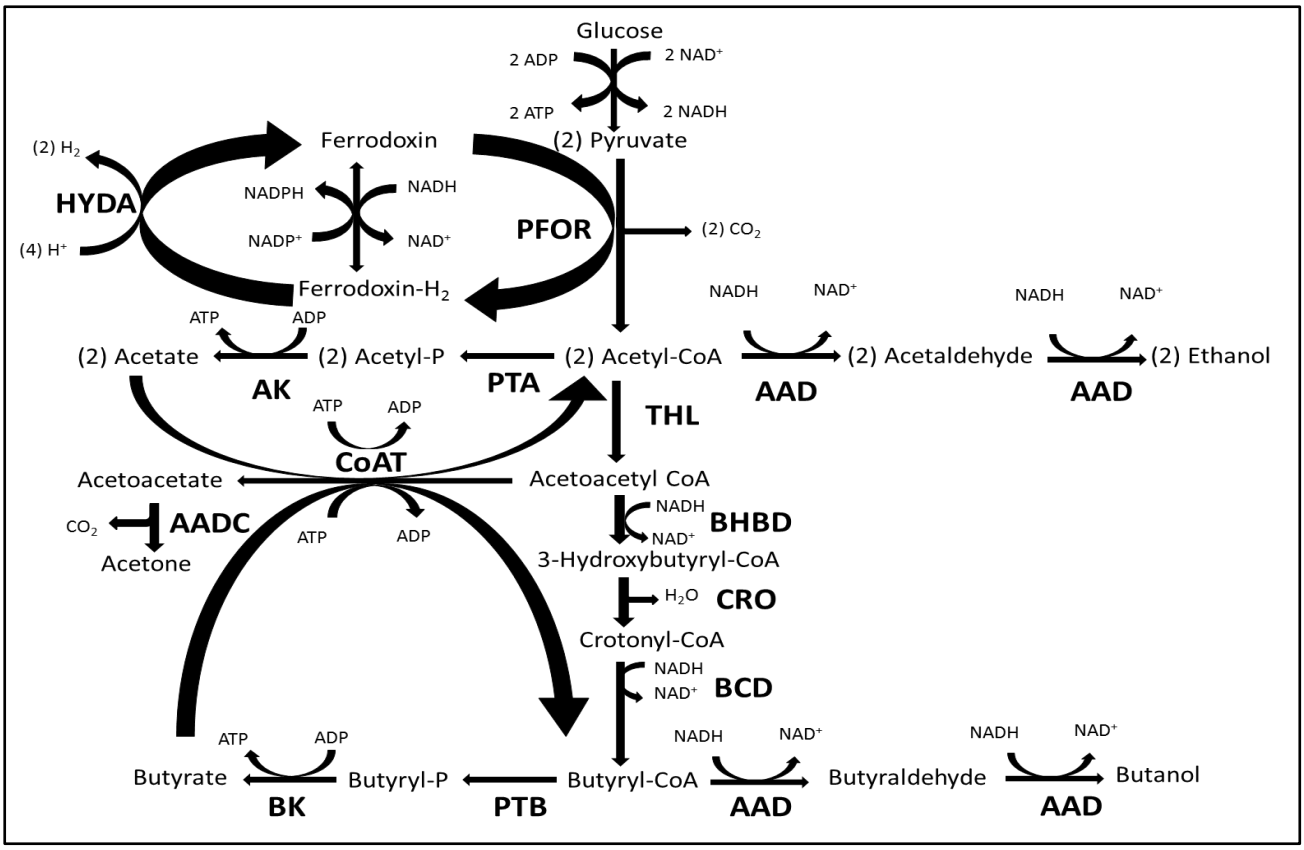


Figure 1.1: Metabolic pathways for the acidogenic and solventogenic phases. Enzymes are depicted in bold and abbreviated as follows: pyruvate-ferredoxinoxidoreductase (PFOR); hydrogenase (HYDA); phosphotransacetylase (PTA); acetate kinase (AK); thiolase (THL); CoA transferase (CoAT); acetoacetate decarboxylase (AADC); β-hydroxybutyryl dehydrogenase (BHBD); crotonase (CRO); butyryl-CoA dehydrogenase (BCD); phosphotransbutyrylase (PTB); butyrate kinase(BK); alcohol/aldehyde dehydrogenase (AAD). Adapted from Papoutsakis (2008) and Kumar and Gayen (2011).

During the onset of solventogenesis (Figure 1.1), acetic acid and butyric acid in the form of acetate and butyrate are assimilated by the acetone and butanol pathways respectively. Acetate and butyrate are reconverted back to acetyl-CoA and butyryl-CoA respectively at the cost of ATP by CoA Transferase (CoAT). Acetyl-CoA is converted to acetoacetyl-CoA by thiolase (THL). Acetoacetyl-CoA is converted to acetoacetate by CoA Transferase (CoAT) which is subsequently converted to acetone through the action of acetoacetate decarboxylase (AADC) which releases CO_2 as a by-product. Butyryl-CoA is converted to butyraldehyde by

aldehyde dehydrogenase (AAD) at the cost of NADH. Butyraldehyde is converted to butanol by alcohol dehydrogenase (AAD) at the cost of another NADH. Small amounts of ethanol are formed by the reduction of acetyl-CoA by alcohol dehydrogenase (AAD) to form acetaldehyde which is further reduced to ethanol by another alcohol dehydrogenase (AAD) (Kumar and Gayen, 2011). The reduction of metabolic intermediates to form butanol and ethanol consume NADH and increase the intracellular pools of NAD^+. The formation of these alcohols as electron acceptors during solventogenesis replaces the hydrogen gas electron acceptor as a means of recycling NADH.

The exact mechanism behind the trigger between acidogenesis and solventogenesis are unknown. It is proposed that intracellular butyryl-phosphate is a regulatory molecule that initiates solventogenesis (Zhao *et al.*, 2005). It has been observed that the shift from acidogenesis to solventogenesis coincided with increased intracellular concentrations of butyryl-CoA while concentrations of CoA and acetyl-CoA decreased (Boynton *et al.*, 1994). It was noted that intracellular concentrations of NADH and NADPH increase during acidogenesis and then declined during solventogenesis. At the peak of acidogenesis, the concentration of ATP is at its lowest, which is assumed to be due to ATP-dependent removal of H^+ from the cell. The shift is thought to be modulated by intracellular ADP: ATP and increasing concentrations of NADPH and NAD^+ (Grupe and Gottschalk, 1992).

In the case of butanol biochemical pathways as seen in Figure 1.1, the genes for solvent production in *C. acetobutylicum* are located on the pSOL1 mega plasmid. The genes are *adc*, *aad*, *ctfA* and *ctfB* which encode the acetoacetate decarboxylase, butyraldehyde dehydrogenase, α- and β-subunits of acetate/butyrate coenzyme A-transferase (AADC, AAD

and CoAT in Figure 1.1). If this plasmid is removed from *C. acetobutylicum*, solvent degeneration occurs (Cornillot *et al.*, 1997). These solvent producing enzymes were found arranged in what is known as the *sol* operon on the chromosome of *C. beijerinckii* (Wilkinson and Young, 1995). A similar *sol* operon configuration has likewise been identified in *C. saccharobutylicum* (Berezina *et al.*, 2009).

In the case of acidogenesis, if the extracellular environment is poorly buffered or the cells grow and divide too rapidly, uncontrolled acid production can lead to cell death before cells have an opportunity to assimilate the metabolic acids into solvents. This is known as acid crash. More specifically, Maddox *et al.*, (2000) noted that "acid crash" occurs in batch cultures without pH control or in continuous cultures with pH control problems when the concentration of undissociated acids in the media exceeds 57 – 60 mmol/L. This is in contrast to "acidogenic fermentations" which occur in batch or continuous cultures that have high starting pH values that decline due to rapid formation of acids followed by inhibition of solventogenesis when the total acid concentration reaches 240 – 250 mmol/L.

Solventogenesis presents its own obstacles to the growth of clostridia since the formation and accumulation of high concentrations of solvents is toxic to the cells that produce them (Nielsen and Prather, 2009). Solvents, particularly butanol have been shown to make the cell membrane more fluid and unselectively permeable. The ratio of cell membrane saturated: unsaturated fatty acids increases in response to solvent stress to compensate for this more fluid membrane. Eventually, however, the membrane ATPases cease functioning and there is a drastic drop in internal pH which leads to cell death (Borden and Papoutsakis, 2007). A study involving the over-expression of stress response genes encoded by the *groESL* operon

showed that these stress response proteins stabilize intracellular proteins during solvent stress, ensuring that cells remain viable long enough for sporulation (Tomas *et al.*, 2003). The mechanisms of acid and solvent survival remain mostly uncharacterized.

1.2.2 Metabolic and genetic engineering of solventogenic clostridia

The best way to understand the biochemical processes involved in acidogenesis and solventogenesis is to create mutants which remove selected enzymes from the biochemical pathway in question. Metabolic and genetic engineering of solventogenic *Clostridium* typically involves mutagenesis whereby genes pertaining to solvent metabolism are inactivated to increase solvent yields, improve solvent tolerance or favour the production of a specific solvent. Mutagenesis systems in solventogenic *Clostridium* remain limited. Initial mutants arose from the selection and improvement of industrial strains using indirect physical and direct chemical mutagens (Jones, 1993). Early work used direct mutagens such as ethyl methanesulfonate (EMS) or methylnitronitrosoguanidine (MNNG) increased the frequency of mutation. Indirect methods such as UV were ineffective (Bowring and Morris, 1985). Butanol-tolerant mutants of *C. acetobutylicum* were created by serial butanol enhancement (Lin and Blaschek, 1983). A combination of UV, MNNG and EMS exposure yielded an improved butanol producing *C. acetobutylicum* mutant (Syed *et al.*, 2008). Early targeted gene inactivation studies used non-replicating vectors. These vectors had very low efficiencies (Green *et al.*, 1996). Later, replicating vectors were incorporated into gene inactivation experiments (Harris *et al.*, 2002). However, these methods are slow and require considerable laboratory input.

Recently, gene inactivation by the Group II intron system, first isolated from *Lactobacillus lactis*, was developed in *Clostridium* to knock out the *buk* and *sol* operon genes (BK and AAD-AADC-CoAT enzymes in Figure 1.1) in *Clostridium acetobutylicum* using a system termed the TargeTron. The mutants generated produced increased butanol titres compared to wild type (Shao *et al.*, 2007). The same Group II intron system was used by Heap and colleagues to generate six different *C. acetobutylicum* mutants as well as five *Clostridium difficile* mutants with system termed the ClosTron. This was the first report of a mutagenesis system that could work in two different species within the *Clostridium* genus (Heap *et al.*, 2007). A system of generating knock-out (KO) mutants was generated by Sillers *et al.*, (2008) in *C. acetobutylicum* strain M5 which lacks the pSol1 mega plasmid (AAD, AADC and CoAT Figure 1.1). Strain M5 was complemented with *aad* expressed from the phosphotransbutyrylase (*ptb*) promoter, which restored butanol production to wild type levels. The *buk* and *ack* genes (BK and AK, Figure 1.1) were inactivated using custom developed pAKKO and pBKKO plasmids. These targets were inserted into the homologous region of the chromosome under selective pressure. The subsequent mutant strains exhibited up to 60% decreased acetate and butyrate production respectively (Sillers *et al.*, 2008). TargeTron- assisted disruption of the acetoacetate decarboxylase (*adc*) gene (AADC, Figure 1.1) decreased acetone formation and shifted the metabolism in favour of increased butanol formation (Jiang *et al.*, 2009). The TargeTron system was incorporated to delete phosphotransacetylase (*pta*) and butyrate kinase (*buk*) genes (PTA and BK, Figure 1.1) involved in acid formation to direct metabolism towards the formation of solvents in *C. beijerinckii* (Wang *et al.*, 2013).

Antisense RNA (asRNA) is a transcriptional repressor which inhibits the translation of gene products. The efficacy of asRNA has been reported in solventogenic clostridia (Fierro-Monti

et al., 1992; Liyanage *et al.*, 2000; Borden *et al.*, 2010). Gene expression manipulation by asRNA is known as knock-down (KD) mutation. KD mutagenesis using asRNA is useful if knocking the gene out is deleterious to the organism's viability. In these cases it is beneficial to rather down-regulate gene translation to low levels. KD was demonstrated in *C. acetobutylicum* to down-regulate the translation of acetoacetate decarboxylase and beta subunit coenzyme A transferase acetone formation enzymes (AADC and CoAT, Figure 1.1) and to decrease acetone formation in favour of butanol formation. The asRNAs were expressed on ampicillin-inducible plasmids. Down-regulation of these enzymes was confirmed by Western Blot and CoAT was determined to be primarily responsible for acetone formation (Tummala *et al.*, 2003b). The same group then used asRNA to down-regulate the translation of *ctfB* which led to the degradation of the whole *aad-ctfA-ctfB* transcript and drastically reduced acetone, but also butanol production, indicating that CtfB is responsible for this effect (Tummala *et al.*, 2003a). Butanol production was reduced in *C. saccharoperbutylacetonicum* when the *hupCBA* hydrogenase cluster was down regulated using targeted asRNAs (Nakayama *et al.*, 2008). asRNA was also used to KD sporulation sigma factor genes to increase butanol production by delaying the initiation of sporulation (Jones *et al.*, 2008).

A gene overexpression study demonstrated that a *C. acetobutylicum* fatty acid synthase overexpression strain did have increased solvent resistance, but did not produce more solvents (Zhao *et al.*, 2003). Overexpression of *groESL* operon genes resulted in increased expression of stress response genes which in turn increased solvent tolerance and titres (Tomas *et al.*, 2003). An overexpression plasmid construct was created utilizing the constitutively expressed *C. acetobutylicum* thiolase promoter coupled to genes which encode solvent enzymes. These enzymes included the acetoacetyl-CoA transferase subunits A and

B encoded by *ctfA* and *ctfB* genes, acetoacetate decarboxylase encoded by the *adc* gene (COAT and AADC, Figure 1.1) and a secondary-alcohol dehydrogenase enzyme from *C. beijerinckii* which catalyses the reduction of acetone to isopropanol. The transformed overexpression *C. acetobutylicum* strain was able to produce a significantly improved total solvent titre of 24.4 g/L Isopropanol-Butanol-Ethanol compared to the untransformed wild type (Collas *et al.*, 2012). Xylose uptake and utilization in *C. beijerinckii* was improved by the overexpression of a xylose-proton symporter encoded by the *xylT* gene cloned into a plasmid system linked to the *ptb* promoter. Solvent titres were increased in the over-expression transformant compared to wild type (Xiao *et al.*, 2012).

The pathways involved in acid and solvent production discussed only depict how the acids and later solvents are formed. Very little is known about the molecular mechanisms solventogenic clostridia use to protect themselves from the harmful low pH effects of acidogenesis and later solvent toxicity. If clostridial cells can mount an effective defence against the destructive effects of low pH environments, then more viable cells are able to enter solventogenesis to produce more solvents. Anaerobic bacteria including solventogenic clostridia are unable to maintain consistent cytoplasmic homeostasis independent of the extracellular environment. Intracellular pH decreases in parallel to the external environment. The internal environment is usually one unit higher than the external environment at any given time (Huang *et al.*, 1985). Since solventogenesis is preceded by acidogenesis, a detailed understanding of the mechanisms bacteria employ to protect themselves from acid shock is necessary.

1.3 The Acid Tolerance Response (ATR) in bacteria

Acid survival is a useful trait for all bacteria in their native environments due to the fact that fluctuations in environmental pH are commonplace. This is particularly relevant for Gram-positive bacteria that are found in food or as animal gastric/oral cavity commensals or that simply exist in acidic environments. Bacteria that produce acids are said to be acidogenic, while bacteria that resist acids are said to be aciduric. The most common form of acid stress is induced by organic acids which can travel through the cell membrane in their un-dissociated form. The cytoplasmic pH is higher than the growth medium, causing weak acids to dissociate, releasing protons thereby decreasing intracellular pH (Cotter and Hill, 2003). The ability to mount various defence mechanisms against the destructive effects of low intracellular pH is known as the Acid Tolerance Response (ATR) (Davis *et al.*, 1996). ATR has been defined as the exposure to a sub-lethal pH which increases the organism's tolerance to subsequent lethal pH exposure (Cotter *et al.*, 2000). The next sections of this review focus on the various mechanisms of ATR across different Gram-positive bacterial genera.

Proton pumps

Under normal cell conditions, the F_1F_0-ATPase combines H^+ influx across the cell membrane driven by the Proton Motive Force (PMF) with ATP generation. This ATPase can work in reverse too: it can secrete H^+ by using ATP generated during fermentative substrate-level phosphorylation. Gram-positive *Enterococcus hirae* (formerly *Streptococcus faecalis*) does not have a respiratory chain; therefore the F_1F_0 complex is not used for ATP synthesis, but solely for the secretion of protons out of the cell (Kobayashi *et al.*, 1986). The decrease in cytoplasmic pH increases ATPase activity (Arikado *et al.*, 1999). It was discovered that F_1F_0-ATPase-deficient mutants could not grow in pH 6 acidic conditions (Suzuki *et al.*,

1988). The permeability of a cell to protons is ultimately determined by active secretion facilitated by ATPases. The lowest pH environments encountered by different lactic acid bacterial species, for example, are directly related to the pH optima of the ATPase enzymes of the species in question (Sturr and Marquis, 1992). Membrane-bound ATPases in *Clostridium acetobutylicum* were shown contribute to the maintenance of an intracellular pH of 6.2 even after more than 100 mM of acetic and butyric acids were produced (Bowles and Ellefson, 1985).

Glutamate decarboxylase

In Gram-positive *Lactococcus*, the glutamate decarboxylase (GAD) system is involved in intracellular pH homeostasis. Extracellular glutamate is internalized and combined with a cytoplasmic H^+ to produce γ-aminobutyrate (GABA) via decarboxylation. The intracellular pH thus increases and the GABA can be exchanged for another extracellular glutamate through a glutamate-GABA antiporter exchange system (Small and Waterman, 1998). The GAD enzyme of *Lactococcus lactis* is active over a pH range of 4.0-5.5 with an optimal pH activity of pH 4.7 which implies that this system is uniquely tailored for coping with the acid stress generated by its own lactic acid metabolism (Nomura *et al.*, 1999). In *Lactobacillus*, it is theorized that three successive decarboxylation-antiporter cycles create a PMF large enough to facilitate the generation of ATP via the F_1F_0-ATPase system, coupling acid metabolism to ATP generation (Figure 1.2) (Higuchi *et al.*, 1997). The GAD system has also been reported in *Clostridium perfringens* (Cozzani *et al.*, 1970). To date, no such systems have been characterised in solventogenic *Clostridium*. A genome analysis to determine acid resistant mechanisms in *C. acetobutylicum* revealed that while no decarboxylase-antiporter system has been reported in the species, the *C. acetobutylicum* gene CAC3285 shared 44%

sequence identity with the *E. coli* glutamate/GABA antiporter gene, *gadC* (Borden *et al.*, 2010).

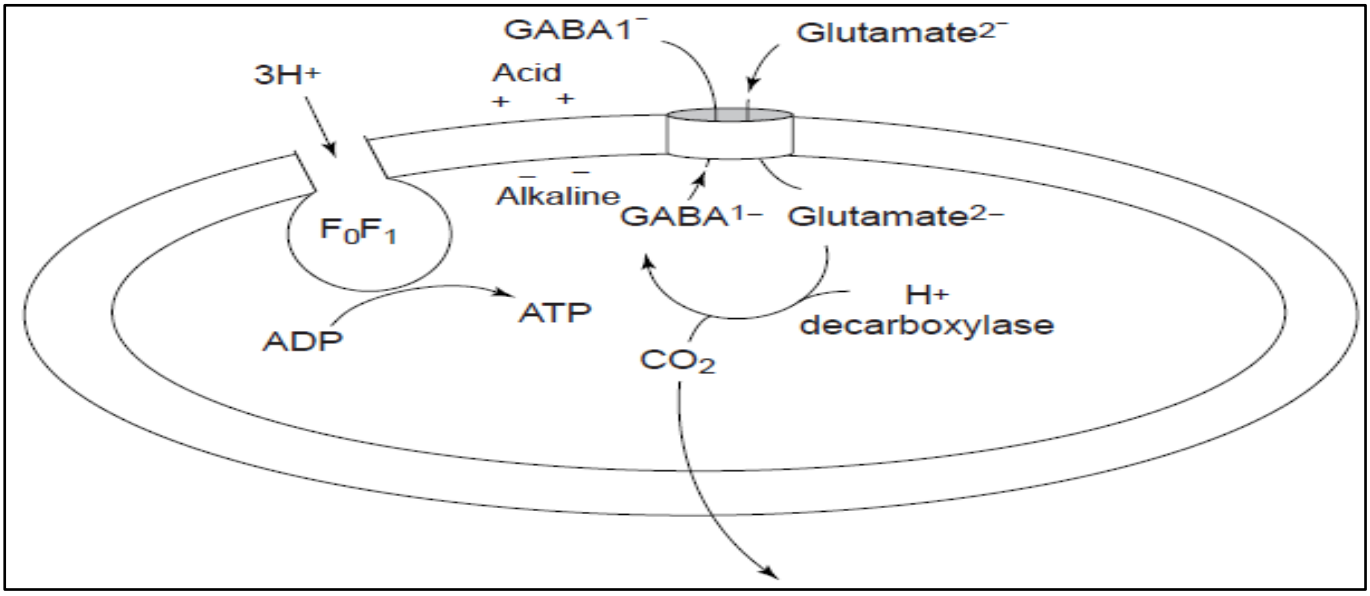


Figure 1.2: The glutamate decarboxylase and the GABA-glutamate antiporter system linked to F_0F_1 ATPase in *Lactobacillus*. Figure from Small and Waterman (1998).

Protection and repair macromolecules

Chaperonins maintain cellular protein integrity by facilitating protein folding, renaturation, protecting denatured proteins from degradation and the removal of damaged proteins (Cotter and Hill, 2003). The chaperone protein DnaK has been shown to increase the stability of the nucleotide excision repair protein, UvrA, in *E. coli* (Zou, 1998). DnaK is up-regulated in response to acid shock in *Streptococcus mutans* (Jayaraman *et al.*, 1997). The stress response chaperone protein GroEL is also expressed in response to acid shock. The increased expression of GroEL and DnaK induced by acid shock has been documented in *Lactobacillus delbrueckii* (Lim *et al.*, 2000). GroEL is likewise expressed under acidic conditions in *C. perfringens* and *S. mutans* (Villarreal *et al.*, 2000; Wilkins *et al.*, 2002). A study in *C.*

acetobutylicum indicated that the overexpression of the chaperonin operon genes *groESL* increased solvent production and tolerance (Tomas *et al.*, 2003). Overexpression of the stress genes, *groESL*, *grpE* and *htpG* in this species has been shown to improve butanol tolerance and long-term butanol adaptation (Mann *et al.*, 2012). There is evidence of the *dnaK-dnaJ* and *groEL-groES* gene operons being expressed at the peak of acid production of pH 5.5 at the transition point between acidogenesis and solventogenesis before any solvents are produced (Grimmler *et al.*, 2011). This implies that these stress response genes are involved in maintaining internal cell stability during peak acid stress as well as throughout solvent production.

Cell membrane changes

Membrane composition during acid stress is important as a primary barrier against the harmful effects of low pH. Membranes containing monosaturated- and long chain-fatty acids are more acid resistant (Quivey *et al.*, 2000). D-alanyl-lipoteichoic acid is a vital component of cell membranes. Inactivation of the *dltC* gene encoding the alanyl carrier protein, DCP, resulted in an acid-sensitive *S. mutans* mutant strain. Further proton permeability experiments confirmed that the *dltC* mutant's membrane was more permeable to H^+ than the wild type strain membrane (Boyd *et al.*, 2000).

Production of alkali

Bacteria use the production of alkali compounds to neutralize intracellular acids. Two of the most common forms of alkali generation involve urease and arginine deiminase (ADI) pathways which generate ammonia (NH_3). Ammonia combines with intracellular protons to produce ammonium (NH_4^+), thereby increasing the intracellular pH. Urease breaks down

intracellular urea into ammonia and CO_2. ADI coverts arginine into ornithine, ammonia and CO_2. These enzymes are particularly well characterised in oral bacteria (Cotter and Hill, 2003). In *Streptococcus salivarius*, the optimum pH for urease activity is at pH 7, and it completely loses activity at pH 4.3. This strongly suggests that the urease system in this organism is used as a preventative measure before acids accumulate at high intracellular concentrations (Sissons *et al.*, 1990). Below pH 5.5, the expression of urease genes can be increased by 600-fold compared to pH 7 conditions if sufficient sugars are present (Chen *et al.*, 1998). When arginine was added to growth media, enzymatic assays indicated that ADI improved cell survival at acid conditions as low as pH 3.1 for *Streptococcus rattus*, *Lactobacillus fermentum*, *Streptococcus sanguis* and *S. mutans* (Casiano-Colón and Marquis, 1988). A study of the ADI activity of *Lactobacillus sanfranciscensis* found in sour dough revealed that the enzyme had an optimal activity range of between pH 3.5-4.5. Furthermore, the organism could only survive lower pH conditions if arginine was supplemented into the media (De Angelis *et al.*, 2002). A study conducted in the pathogen, *Streptococcus pyogenes*, indicated that ADI-deficient mutant strains had severely reduced survival compared to wild type after exposure to pH 4 conditions for 6 h (Degnan *et al.*, 2000).

Gram-positive bacteria have many different endogenous methods of controlling intracellular pH in relation to the external environmental pH in order to survive. By contrast, the supplementation of certain exogenous amino acids into growth media has been shown to increase the ATR in certain species of bacteria. In Gram-negative *E. coli*, the supplementation of glutamine and glutamate into the growth media improves acid survival under pH 2.5 conditions. The conversion of glutamine to glutamate by the acid-activated glutaminase, YbaS, which releases ammonia, is a form of alkali production ATR mentioned previously: the ammonia binds to an intracellular proton to form ammonium which increases

the intracellular pH. It was further demonstrated that antiporter GadC, which exchanges extracellular glutamine with intracellular glutamate ensures a constant supply of intracellular glutamine is present for the glutaminase reaction (Lu *et al.*, 2013b). It was demonstrated that *E. coli* mutants deficient in the *gadB* gene, which encodes a glutamate decarboxylase, displayed an acid-sensitive phenotype. In addition, it was shown that glutamate supplementation improved survival of wild type *E. coli* in acidic conditions (Hersh *et al.*, 1996).

Amino supplementation in growth media has been shown to improve the ATR in Gram-positive bacteria as well. It was demonstrated in *Lactobacillus buchneri* that the histamine-histidine antiporter system was linked to the decarboxylation of histidine to form histamine and CO_2. The intracellular histamine was exchanged for an extracellular histidine by the histamine-histidine antiporter (Molenaar *et al.*, 1993). Media supplemented with glutamate, arginine and lysine enhanced the ATR of *Bacillus cereus* cells grown at pH 4.0. The survival of *B. cereus* cells was diminished when a competitive inhibitor of arginine decarboxylase, agmatine, was added to the growth medium, which implies that arginine decarboxylase plays a role in the ATR in this species (Senouci-Rezkallah *et al.*, 2011). The involvement of amino acids in the ATR in solventogenic *Clostridium* has not been widely studied despite the importance of intrinsic acidogenesis metabolism as a precursor for solventogenesis.

1.4 Nitrogen metabolism in bacteria

Both inorganic and organic nitrogen sources have been shown to be important in clostridial growth and solvent production (Long *et al.*, 1984; Stutz *et al.*, 2007). A comprehensive understanding of nitrogen metabolism in bacteria is important since nitrogen is incorporated into amino acids which are involved in every aspect of bacterial cell growth and development. The role of nitrogen assimilation enzymes in converting inorganic nitrogen, most often in the form of ammonia (NH_3), into organic nitrogen in the form of amino acids is vital to bacterial survival. Nitrogen assimilation is particularly important when environmental organic nitrogen availability is scarce. This section describes general nitrogen assimilation enzymes in Gram-negative bacteria with a focus on nitrogen assimilation across Gram-positive genera, detailing the mechanisms required for biological activity and the regulation of the components involved in nitrogen assimilation.

The Gram-negative bacterium, *E. coli*, is a well characterised model organism both genetically and biochemically. As such, some of the most comprehensive knowledge of nitrogen metabolism in bacteria has been based on this organism, and this has been reviewed extensively by Merrick and Edwards (1995). *E. coli* grown in ammonia-rich environments uses glutamate dehydrogenase (GDH) to assimilate inorganic nitrogen sources into organic nitrogen. During intracellular ammonia limiting conditions below 1mM, the glutamine synthetase -glutamate synthase (GS-GOGAT) system is activated (Merrick and Edwards, 1995). The expression of the *glnA* gene, which encodes GS, is controlled by the NtrB-NtrC nitrogen sensor-regulator system. This two component system activates σ^{54}-dependent promoters in response to nitrogen limiting conditions (Weiss *et al.*, 2002). Cellular 2-

oxoglutarate and glutamine concentrations are sensed by the PII-type sensory protein which activates NtrB kinase to phosphorylate NtrC (Jiang and Ninfa, 2009).

Different mechanisms for transcriptional regulation are employed across the different bacterial genera, yet the ammonia assimilation enzymes, GS and GOGAT, as well as the nitrogen signal transduction proteins are widely conserved (Amon *et al.*, 2010). Examining the mechanisms of nitrogen assimilation in model systems such as *E. coli* forms the foundation of our understanding of nitrogen metabolism; however, the focus of this review will be nitrogen metabolism in Gram-positive bacteria, which demonstrates interesting differences. An understanding of metabolism in other Gram-positive bacteria is crucial, especially in the context of solventogenic species of *Clostridium* covered in later chapters. The regulation of nitrogen metabolism in the most well characterised Gram-positive bacteria has been well reviewed (Burkovski, 2003; Amon *et al.*, 2010). The following sections describe nitrogen regulation in the high GC content Actinobacteria (*Streptomyces*, *Mycobacterium* and *Corynebacterium*). Finally, the regulation of nitrogen metabolism will be described in the low GC content Firmicutes (*Bacillus*, *Lactobacillus* and *Clostridium*) with particular emphasis on *Clostridium*. Overall similarities across genera are summarised in Table 1.1.

Table 1.1: A comparison of the key components of ammonium assimilation and nitrogen control across different Gram-positive genera. Adapted from Amon *et al.* (2010). α and β designate the respective subgroups of the type I GS enzyme (Brown and Sonenshein, 1996).

		Firmicutes				Actinobacteria		Proteobacteria
Protein	**Function**	***Bacillus***	***Lactobacillus***	***Clostridium***	***Streptomyces***	***Mycobacterium***	***Corynebacterium***	
GS (type I)	Glutamine synthesis,	α	α	α	β	β	β	β
GlnA2-A4	ammonium assimilation	No	No	No	Yes	Yes	Some	No
GS (type II)		No	No	No	Yes	No	No	Some
GS (type III)		No	No	Some	No	Some	No	Some
GOGAT	Glutamate synthesis, ammonium assimilation	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GlnE	Adenylyl-transferase, regulation of GS activity	No	No	No	Yes	Yes	Yes	Yes
AmtB	Ammonium permease	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GlnK	PII-type signal transduction protein	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GlnD	Adenylyl-/uridylyltransferase, Regulation of GlnK	No	No	No	Yes	Yes	Yes	Yes
Transcriptional Regulator (s)		GlnR/TnrA/ CodY	GlnR	NitR	GlnR/GlnRII	GlnR	AmtR	NtrC

Nitrogen metabolism regulators in *Streptomyces*

The gene encoding the *Streptomyces coelicor* GS type I (*glnA*) is regulated by GlnR (Table 1.1) (Wray *et al.*, 1991). GlnR belongs to the OmpR-type family of transcription factors (Kenney, 2002). *Streptomyces* have evolved another OmpR-type regulator called GlnRII which recognises the promoter for the type II GS-encoding gene, *glnII* (Weisschuh *et al.*, 2000). The exact regulation of the putative ammonia assimilators GlnA2-A4 in the genus is unclear. Many nitrogen metabolic pathway genes share consensus sequence binding sites of GlnR. This means that multiple binding events with GlnR are possible which improves

regulatory sensitivity and control of promoter expression in response to nitrogen limitation. The *nirBD* operon encoding subunits for nitrite reductase has also been demonstrated to be under GlnR control (Tiffert *et al.*, 2008). The global regulatory proteins GlnK, GlnD and the ammonia transporter AmtB are encoded together in a gene cluster in *S. coelicor*. This operon is controlled by GlnR under limiting conditions (Fink *et al.*, 2002).

The post-translational regulation in this genus appears to emulate that of NtrBC described in *E. coli* earlier in this review; however, the mode of regulation differs. Actinomycetes have homologues of, uridylyl-transferase GlnD, and, adenylyltransferase, GlnE, regulator proteins. In *Streptomyces coelicolor*, GlnD can also adenylate the PII-like protein, GlnK, in response to nitrogen scarcity. In *E. coli*, the nitrogen levels of the cell are sensed by the uridylyl-transferase, GlnD, which signals GlnK. GlnK in turn, activates the adenylyltransferase, GlnE, which activates GS by adenylation (Forchhammer, 2008). Unlike the Ntr system in *E. coli*, the GlnD-GlnK interaction is not required for GlnE-dependent regulation of GS in *S. coelicor* (Hesketh *et al.*, 2002). In a recent study, the *glnA* promoter was used as a target for the phosphate system regulator, PhoP, and GlnR competitive binding to the same promoter region. GlnR demonstrated a higher affinity for binding, yet the nitrogen genes regulated by GlnR are controlled in part by PhoP. This ensures cross-talk between nitrogen and phosphate metabolic pathways during phosphate and nitrogen limitation (Sola-Landa *et al.*, 2013).

Nitrogen metabolism regulators in *Mycobacterium*

Genomic sequence screening and mutagenesis studies in *Mycobacterium smegmatis* indicated that the *amt1*, *amtB-glnK-glnD* and *glnA* genes were positively regulated by a GlnR-like protein (Table 1.1) (Amon *et al.*, 2008). In *Mycobacterium tuberculosis*, it was shown that the genes *nirBD* which encode the subunits of the nitrite reductase are up-regulated by GlnR in response to nitrogen starvation (Malm *et al.*, 2009). Two putative ammonium transporter genes, *amtA* and *amt1* have been reported in *M. tuberculosis*. The *amt1* gene is organized in a cluster together with the genes for a putative class II glutamine synthetase, a class II glutamine aminotransferase and glutamate synthase. No nitrogen-related operon structure was reported for *amtA* (Amon *et al.*, 2009).

Only the non-pathogenic *M. smegmatis* possesses a *gdhA* gene encoding an NADH-dependent GDH enzyme. This strongly implies that GS-GOGAT is the primary means of ammonia assimilation in other *Mycobacterium* species. The *glnA1* and *glnA2* genes together with *glnE*, which encodes for a GS-regulating adenylyl-transferase, are conserved in all mycobacterial genomes (Carroll *et al.*, 2008). It should be noted that despite the presence of *glnA1-4* genes in some species of *Mycobacterium*, only *glnA1* is required for homeostasis in *M. tuberculosis*. The functions of the other putative GS-encoding genes like those seen in *M. smegmatis* remain to be investigated (Harth *et al.*, 2005). The PII protein in *M. tuberculosis* is similar to orthologues seen in other species in terms of ATP and 2-oxoglutarate binding affinity (Bandyopadhyay *et al.*, 2010).

Nitrogen metabolism regulators in *Corynebacterium*

AmtR is the global nitrogen regulator in *Corynebacterium glutamicum* and other corynebacteria (Table 1.1) (Walter *et al.*, 2007). AmtR belongs to the TetR family of transcriptional regulators (Ramos *et al.,* 2005). Transcriptomic, proteomic and bioinformatic analyses have revealed a large network of 35 genes directly regulated by AmtR in the genus. These include the ammonia transport genes, *amtA* and *amtB*, the glutamate transporter gene *gluABC* and genes for the ammonia assimilation enzymes encoded by *gdh*, *glnA* and *gltAB*. AmtR also regulates alternative nitrogen source pathways such as creatine and urea utilization pathway genes encoded by *codA* and *ureABCEFGD* respectively (Buchinger *et al.*, 2009). The *amtA* and *amtB* ammonia transporter genes were not transcribed under nitrogen rich conditions. This is because AmtR binding sites are located upstream of *amtA* and *amtB* which allows AmtR to bind, thereby preventing transcription of these genes. In contrast, it was found that *glnA* expression was induced under low ammonia conditions (Jakoby *et al.*, 2000).

In corynebacteria, AmtR is a repressor that binds to nitrogen gene DNA, preventing transcription. During low ammonia conditions, the PII-like GlnK transduction protein activates and interacts with AmtR, which results in free transcription of nitrogen genes (Beckers *et al.,* 2005). GlnK must be adenylated by GlnD in order to interact with AmtR in the first place. GlnK is activated in response to nitrogen limitation. Additionally, the adenylation status of GlnK has been shown to be involved with localization of GlnK to the cytoplasm (Strösser *et al.*, 2004). Intracellular concentrations of 2-oxoglutarate and glutamine have been identified as candidate metabolites that are sensed by response regulator proteins which then modulate nitrogen metabolism (Forchhammer, 2004).

Nitrogen metabolism regulators in *Bacillus* and *Lactobacillus*

Nitrogen regulation in *Lactobacillus* appears to be similar to that of *Bacillus* (Amon *et al.*, 2010). The MerR-type regulatory proteins GlnR and TnrA have been shown to regulate nitrogen genes in *Bacillus subtilis* (Table 1.1) (Schreier and Rostkowski, 1995; Wray *et al.*, 1996). GlnR is a transcriptional repressor that down regulates the expression of the *glnRA* and *ureABC* operons (Brown and Sonenshein, 1996; Wray *et al.*, 1997). TnrA by contrast, is a transcriptional activator and up-regulates transcription of the ammonia transporter gene *amtB*, the PII-like protein encoded by *glnK*, the urease operon *ureABC* and the nitrate and nitrite reductase gene clusters comprising *nasBC* and *nasDEF* respectively (Wray *et al.*, 1996; Wray *et al.*, 1997; Nakano *et al.*, 1998).

The GOGAT encoding genes are under the control of the GltC regulator. GOGAT gene expression is down-regulated by GltC when intracellular concentrations of 2-oxoglutarate are high (Picossi *et al.*, 2007). Intracellular glutamine has also been shown to be indirectly sensed by GlnR and TnrA (Sonenshein, 2007). GlnR regulates its own expression through a transcriptional feed-back response loop. GlnR binds to the DNA sequence upstream of the *glnRA* operon. The resulting GlnR-feedback-inhibited GS inactivates TnrA by blocking the DNA binding capabilities of TnrA (Wray and Fisher, 2008). Therefore, GlnR represses nitrogen metabolism genes under nitrogen rich conditions and TnrA activates expression of nitrogen metabolism genes during nitrogen limitation conditions. Bacteria belonging to the *Bacillus* genus possess homologues of the global regulator protein CodY that has the ability to sense intracellular concentrations of branched-chain amino acids. CodY plays a key role as a mediator between nitrogen and carbon metabolism through direct repression of the urease

pathway genes. CodY also regulates the expression of genes involved in amino acid transport, catabolism and biosynthesis (Sonenshein, 2007).

Nitrogen metabolism regulators in *Clostridium*

Clostridium species have homologues of global regulators such as CodY (Varga *et al.*, 2004; Dineen *et al.*, 2007). No homologues of GlnR or TnrA have been found in *Clostridium* to date. Pathogenic species, such as *Clostridium difficile* do not have such high dependencies on nitrogen assimilation since their relationship within the human host drives reductive evolution of their genomes. As pathogens adapt and become dependent on their host environment, they begin to lose genetic information that is redundant for nutrient biosynthesis. In some cases this includes components of nitrogen assimilation (Gomez-Valero *et al.*, 2007; Titgemeyer *et al.*, 2007).

Soil anaerobes in the *Clostridium* genus are by contrast very much dependent on their nitrogen assimilation genes. It is interesting to note that *C. saccharobutylicum*, *C. beijerinckii*, *Clostridium tetani* and *Clostridium perfringens* possess type I GS enzymes while *C. acetobutylicum* has a type III GS (Reid and Stutz, 2005). The ABE-producing *C. saccharobutylicum* has a nitrogen assimilation gene cluster consisting of *glnA-nitR-gltAB* genes which encode GS, NitR and the two subunits of GOGAT respectively. NitR is an RNA-binding regulatory protein (Table 1.1) which belongs to the AmiR and NasR Transcription Antitermination Regulator (ANTAR) family (Shu and Zhulin, 2002). These proteins have domains which can be phosphorylated by a kinase, presumably under nitrogen limiting conditions. NitR has an RNA-binding domain that promotes transcriptional anti-termination and under nitrogen-limiting conditions, NitR applies its anti-termination activity,

which results in the bicistronic expression of *glnA-nitR* and *gltAB*. There is a second level of regulation by asRNA. During high nitrogen conditions, asRNA encoded in the *glnA* region binds to ribosomal binding sites of the two operons and down-regulates intracellular GS-GOGAT enzyme levels (Stutz *et al.*, 2007). This system seen in *C. saccharobutylicum* seems to be conserved in *Clostridium beijerinckii*, *Clostridium butyricum* and *Clostridium botulinum* (Amon *et al.*, 2010). To date, neither the ammonia transporter nor the PII-like regulator genes and proteins have been characterised in solventogenic *Clostridium* and no global regulators have been identified.

In summary, the ammonia transporter family Amt and their homologues are highly conserved across bacterial genera (Tremblay and Hallenbeck, 2009). GS and GOGAT are also largely conserved among prokaryotes and this reflects the importance of ammonia assimilation in the microbial world (Amon *et al.*, 2010). The genes encoding the PII-like GlnK regulatory proteins are found in most bacterial genome sequences to date, although these genes are controlled by many forms of post-translational interactions and regulation (Forchhammer, 2008).Global regulators are conserved within members of the same genus; however there is variability in their transcriptional regulation (Balleza *et al.*, 2009).

1.5 The importance of nitrogen metabolism in clostridial ABE fermentation

The conservation of the GS-GOGAT genes in prokaryotes is indicative of the importance of ammonia assimilation for survival. The role of these genes in solventogenic *Clostridium* is no less crucial. Our understanding of nitrogen metabolism and specifically the regulatory mechanisms that control nitrogen metabolism remains limited in this very diverse genus. The role of nitrogen metabolism in the ABE fermentation process specifically has not been investigated in any detail. There is evidence that the nitrogen source in the media affects solvent production. Experiments using ammonia-limited media demonstrated that *C. saccharobutylicum* did not accumulate enough acid end-products to trigger solventogenesis. Addition of ammonia to this media enhanced solvent production (Long *et al.*, 1984). The addition of organic nitrogen in the form of yeast extract and inorganic nitrogen in the form of ammonium nitrate to the media was found to be crucial to enhance solvent production (Abd-Alla and Elsadek El-Enany, 2012). Recently, a number of microarray and transcriptomic experiments have shown that genes involved in production or transport of certain amino acids are up-regulated during solvent stress. DNA microarray experiments that monitored the transcriptional events of *C. beijerinckii* during solvent stress identified glutamine, asparagine and histidine among other compounds as necessary for fermentation. Additionally, a lysine permease was shown to be upregulated during solvent stress (Heluane *et al.*, 2011). A microarray study on *C. acetobutylicum* cells exposed to a butanol pulse revealed that the genes involved in glutamine transport, glutamine synthesis, glutamate synthesis and proline transport were up-regulated in response to solvent stress (Janssen *et al.*, 2012). However, the role of individual amino acids during ABE fermentation has not been investigated.

1.6 Aims of this study

Preliminary work conducted on the NCP strain collection indicates that the strains are either *C. beijerinckii* or *C. saccharobutylicum* species (Keis *et al.*, 2001). The extent of the relatedness between strains of the same species of the UCT NCP strain collection remains unknown. The first aim of this thesis is to characterise the different strains in the two species phylogenetically at a strain level and to determine the solvent capabilities of the individual strains on glucose, sucrose and xylose carbohydrate substrates.

In addition, this project aims to elucidate the role of organic nitrogen supplementation, specifically, the effect of glutamine, glutamate, asparagine, lysine, histidine, and proline amino acid supplementation on the growth and solvent performances of the ABE-producing *C. beijerinckii* and *C. saccharobutylicum* NCP strains. Because amino acids can play a role in the ATR, acid shock experiments will determine if any of the amino acids investigated improve the ATR during acidogenesis and therefore improve the number of viable cells that transition to solventogenesis.

Bioinformatics will be used to identify possible nitrogen metabolism candidate genes in the genomic sequence of *C. beijerinckii* ATCC8052^{T}. Knock-out mutants of these candidate genes using ClosTron will be made to determine the effect these mutations have on nitrogen metabolism and solvent production. The role of the GS-GOGAT genes during acid formation is unclear. Quantitative PCR (qRT-PCR) experiments will be performed to determine if these genes are up-regulated in response to acid stress conditions. A better understanding of the fundamental biochemical pathways involved in nitrogen metabolism will create opportunities

to improve the growth, acid tolerance and ultimately solvent production of solventogenic *Clostridium* species.

Chapter Two
Taxonomy of the South African National Chemical Products (NCP) Strains

---

Contents

2.1 Abstract

A collection of 19 solventogenic *Clostridium beijerinckii* and 11 *Clostridium saccharobutylicum* strains isolated from the National Chemical Products (NCP) Acetone Butanol Ethanol (ABE) fermentation plant in Germiston, South Africa, were separated and classed according to species, by a quick species-specific colony PCR and rifampicin screening methods. Strain level grouping was attempted by 16s *rRNA* phylogenetic and electron transport chain flavoprotein alpha (*etfA*) gene Restriction Fragment Length Polymorphism (RFLP) analysis, but these methods did not have the resolving power to differentiate between closely related strains. Random Amplification of Polymorphic DNA (RAPD) analysis was performed on the strain collection and showed two strain groups for NCP *C. saccharobutylicum* strains and four groups within the *C. beijerinckii* species. Multilocus Sequence Typing (MLST) was employed on a smaller selection of strains based on the RAPD results. Nucleotide differences in the form of Single Nucleotide Polymorphisms (SNPs) for MLST analysis showed two strain groups within the NCP *C. beijerinckii* strains and three groups within the *C. saccharobutylicum* strains.

The species-specific PCR method aims to provide a rapid means of assessing any contamination of ABE batch fermentations by differentiating between *C. saccharobutylicum* and *C. beijerinckii* species, while the novel method of whole genome RAPD analysis indicates whether any strain diversity is present before high resolution techniques are employed. MLST analysis reveals a high degree of strain level discrimination and is able to detect even small differences between closely related NCP strains. All taxonomic analysis performed in this study indicates that the NCP *C. beijerinckii* strains are genetically distinct from NCIMB 8052^{T}.

2.2. Introduction

In the decades following Chaim Weizmann's historic discovery and industrialization of ABE forming microbes in the 1920s, commercial production of butanol peaked and a number of different bacterial cultures capable of producing acetone and butanol were isolated and used in various industrial fermentation plants. These strains were collectively designated the genus and species name, *Clostridium acetobutylicum*, on the basis of being Gram-positive, anaerobic, rod-shaped spore-formers (Jones and Woods, 1986).

Early phylogenetic work on the16s *rRNA* region done on a multitude of solvent and non-solvent producing clostridial species indicated that *C. acetobutylicum* and *C. beijerinckii* were closely related and clustered together in their own clade relative to other clostridia (Collins *et al.*, 1994). This was the last major analysis of the genus as a whole. As molecular techniques developed and tools such as Pulse Field Gel Electrophoresis (PFGE) were implemented in taxonomic characterisation studies, it was revealed that many of the earlier solvent producing strains that were initially broadly identified as *C. acetobutylicum*, in fact, consisted of many distinct clostridial species (Keis *et al.*, 1995). In the ABE fermentation industry, differences in solvent production profiles and substrate preference of the different isolates were becoming ever more apparent. A need to further classify solventogenic *Clostridium* strains arose. In the years before the revision of the major solventogenic species, it was noted that there were strains under the *C. acetobutylicum* species umbrella which, unlike the supposed ATCC 824^T type strain, were resistant to rifampicin antibiotic. Additionally, these strains did not produce riboflavin which typically reacts to form a bright yellow pigment in milk. These anomalies in phenotypic characterization studies prompted the first high resolution inter- and intra-species study of the solventogenic *Clostridium* group in

the form of DNA/DNA hybridization studies, which clearly showed that the group formerly known as *C. acetobutylicum* was comprised of at least four distinct species, including *C. acetobutylicum*, *Clostridium saccharoperbutylacetonicum*, *C. beijerinckii* and P262^{T} that would later be designated *C. saccharobutylicum* (Johnson *et al.*, 1997). South Africa's collection of commercial strains from the National Chemical Products (NCP) ABE plant in Germiston, which operated from 1936 to 1982 (Jones *et al.*, 2000), is no exception: the strain collection consists of *C. beijerinckii* and *C. saccharobutylicum* species and not *C. acetobutylicum* as was once thought (Keis *et al.*, 2001). Today there are a variety of solvent producing *Clostridium* species available. Some of the most prevalent players in the ABE fermentation process have been thoroughly screened and formally identified as *C. saccharoperbutylacetonicum* (ATCC 27021^{T}), *C. saccharobutylicum* (DSM 13864^{T} or NCP262^{T}), *C. beijerinckii* (NCIMB 8052^{T}) and *C. acetobutylicum* (ATCC 824^{T}) species (Keis *et al.*, 2001).

Commercial strains are particularly poorly classified, and strains with the same accession numbers are being found to have different genome sequences. Recently DNA sequence differences have been reported for the same strains of *Mesorhizobium loti* and *Vibrio pelagius* from different culture collections and it would appear that lyophilisation or other storage conditions, could lead to small strain differences in the genomes (Macian *et al.*, 2000; Willems *et al.*, 2001). With the advent of improved molecular and genetic manipulation of industrially relevant strains, correct identification of each strain, both physiologically and genetically, is becoming increasingly important. This is particularly crucial considering the non-model organism nature of clostridia and the limited metabolomic and genetic manipulation tools available for use within this group (Heap *et al.* 2007). With the advent of whole genome sequencing, further revisions of the taxonomic status of industrially relevant

Clostridium strains may well be necessary. However, this technology requires major input in terms of financial and analytical resources.

For practical purposes, 16s *rRNA* sequencing is most widely used to differentiate solvent-producing clostridia, with a divergence of greater than 3% in sequence identity being sufficient to separate species (Goebel and Stackebrandt, 1994). However, it is interesting that this region is so conserved in the clostridia. There are instances of species, considered to be separate, that show more than 99% sequence identity (Stackebrandt *et al.*, 1999). One of these examples is that of *C. acetobutylicum* and *Clostridium chromiireducens* (Inglett *et al.*, 2011). DNA/DNA hybridization remains the gold standard for strain separation, but for ease of use, other techniques based on differences in genomic DNA, such as PFGE and Restriction Fragment Length Polymorphism (RFLP) analysis, have been successfully used to identify species and type strains (Hou and Dutta, 2000 ; Keis *et al.* 2001).

To date, the University of Cape Town does not have any records of strain allocations in batch fermentations during the NCP facility's operation. The origins of the strains, the time period in which strains were used and the reasoning behind why certain strains were favoured above others, remain unknown. The strain level analysis is important academically, since it is unknown whether the NCP strains used from the early 1930s to the 1980s, are:

1) derived from a common ancestor and are clonal copies of the original parent strain (s), or

2) derived from a common parent strain and diverged from there, or

3) isolated separately for each strain.

From a practical aspect, species level classification of the strain collection is vital to maintain pure cultures in the fermentation process and be able to use the correct species for the correct substrate or feedstock. Strain level characterization is of paramount importance in an industrial context since solvent profiles and substrate preferences can be linked to closely related strains within a certain strain grouping within a species. These strains can be further analysed phenotypically or by more definitive means such as whole genome sequencing to identify genetic determinants of desirable phenotypes.

One approach to improving species-based taxonomy in solventogenic clostridia is to search for unique genes or gene configurations involved in growth, solvent production or substrate preference. These genes can act as a unique genetic barcode for the species in question and be used instead of laborious physiological tests or time intensive 16s *rRNA* sequencing. The nitrogen assimilation enzymes, encoded by the glutamine synthetase (*glnA*) and glutamate synthase (*gltAB*) genes, have been shown to have a unique gene configuration in both *C. saccharobutylicum* and *C. beijerinckii* (Stutz *et al.* 2007). To date, this gene operon configuration has only been seen in *C. saccharobutylicum*, *C. beijerinckii*, *Clostridium botulinum* and *Clostridium butyricum* species (Amon *et al.*, 2010). In contrast, analysis of the *C. acetobutylicum* ATCC 824^{T} genome shows it lacks this operon configuration and only possesses a *glnA* gene, encoding a type III glutamine synthetase far upstream from the *gltAB* genes. This makes the nitrogen assimilation genes a natural target for species level characterization since not only are the genes vital for the growth and solvent production of clostridial ABE fermenters, but the genes also act as a unique DNA signature for each of the three major species of interest in this study (Stutz *et al.* 2007). The availability of the full sequence of the type strains *C. acetobutylicum* ATCC 824^{T} and *C. beijerinckii* NCIMB 8052^{T} (closely related to the *C. beijerinckii* NCP strains), means that comparisons can be made to

NCP strains, particularly since the full genome sequence of *C. saccharobutylicum* has only recently become publicly available and is not fully annotated.

In an attempt to increase strain-level typing of the NCP collection, Restriction Fragment Length Polymorphism (RFLP) of variable nucleotide regions was investigated. The electron transfer flavoprotein alpha (*etfA*) gene was considered as a candidate region for RFLP analysis (Boynton *et al.*, 1996). EtfA protein is involved with the electron transport chain in the formation of acetoacetyl-CoA and butyryl-CoA precursors to acetone and butanol formation (Green *et al.* 1996). The region is variable at a nucleotide level and displays differences between the major solventogenic clostridial species (Personal correspondence, Green Biologics Ltd. UK). Although looking at variable genes or nucleotide regions rich in polymorphisms is a valid approach to gauge strain diversity, the method is constrained by the maximum amount of evolution and mutation that can occur within the region before function of the region is compromised. To this end, a genome-based approach that was analytically simple and cost effective for analysis of a large number of strains was investigated.

Random Amplification of Polymorphic DNA (RAPD) analysis fits these requirements. RAPD is a PCR-based technique that uses arbitrary primers of a short length and low annealing temperature conditions to promote non-stringent primer binding across many sites on the genome. The result is the generation of many PCR products of various sizes which produce multiple bands during agarose gel electrophoresis. It is assumed that related strains show similar banding patterns which are due to similar sequences around the primer binding sites across the genome (Williams *et al.*, 1990). Multiple differences in banding patterns

between two strains can therefore act as a proxy for strain polymorphisms which indicates strain diversity.

RAPD analysis has been used with the *Clostridium* genus since the 1990s, primarily in the clinical setting pertaining to outbreaks of *Clostridium difficile* infections in Europe (Van Dijck *et al.*, 1996). More recently, RAPD has been utilized in the food industry to track *Clostridium perfringens* and *C. difficile* infections in pig farming (Baker *et al.*, 2010) and monitoring outbreaks of environmental *C. botulinum* (Hannett *et al.*, 2011). In all of these examples, RAPD was used as a preliminary screening method to ascertain strain diversity of large sample populations in a cost-effective manner. Since there are presently no published studies using RAPD analysis on solventogenic clostridia, RAPD screening was used as an indicator of strain diversity of the NCP strains before considering any higher resolution strain typing methods. Most importantly, no prior sequencing information is required from the strains characterized in order for grouping to occur. Therefore, RAPD analysis may be applied to any solvent producing clostridial strain.

With improved sequencing technology and a drastic decrease in the cost of nucleotide sequencing, more sequence-based tools have become available. One approach is to isolate ubiquitous markers that contain highly variable sections across species or genera. Finding unique genetic markers is difficult since even unrelated species share many similarities in gene order and configuration. Sequence-based analysis normally targets metabolically important genes crucial for the organism's survival, and are termed "housekeeping" genes. Several housekeeping genes, such as *rpoB*, *gyrB* and *recA*, are commonly used in many molecular techniques due to their high degree of conservation between species. In clostridia,

however, there is often insufficient variation between these genes alone, therefore more housekeeping genes are required to improve strain phylogenetic resolution (Chalmers *et al.*, 2008).

Multilocus Sequence Typing (MLST) uses genes that are conserved across strains or species, but may contain regions of nucleotide variability. These areas of variability are referred to as Single Nucleotide Polymorphisms (SNPs) and form the basis for observing differences between strains in the downstream phylogenetic analysis. These genes are arranged in identical order to create a concatenated gene sequence for each strain. These sequences can then be aligned and compared to one another phylogenetically (Maiden, 2006). MLST is indeed a powerful tool for modern bacterial typing and has shown promise in related species of clostridia such as *Clostridium septicum* (Neumann and Rehberger, 2009).

In the ABE fermentation industry, it is important that commercial strains can be properly identified and classed. As with most batch fermentation processes, there is a risk of contamination by closely related species or environmental contaminants in the feedstock used in the industrial process. In this chapter, a rapid, colony-PCR based screening method is used as an efficient means of confirming the presence of the desired species in batch fermentations. The differences in the nitrogen assimilation gene configuration and nucleotide sequences between the three species under investigation enables the design of species-specific primers that can differentiate between the *C. saccharobutylicum, C. beijerinckii* and *C. acetobutylicum* species. Assignment of the correct species is validated through 16s *rRNA* sequencing, rifampicin screening and *etfA*-RFLP analysis to identify and elucidate the purity of all the strains present in the UCT NCP strain collection.

In addition, higher resolution RAPD analysis with a standardized set of primers is incorporated to generate a variety of banding patterns unique at a strain level. The strain level resolution is further increased by implementing MLST analysis of a subset of strains based on the RAPD data with the aim of acquiring reproducible strain level differences between strains within the same species, using a collection of sequenced standard housekeeping gene targets. Since there is presently no record of MLST being implemented in solventogenic *Clostridium*, a variety of known housekeeping genes were used with the aim of generating phylogenetic profiles that could adequately discriminate between strains within the *C. saccharobutylicum* and *C. beijerinckii* species respectively. These nucleotide data can be used to determine how closely related the NCP strains within the *C. saccharobutylicum* and *C. beijerinckii* strains are to one another and to other solventogenic species outside of the collection.

2.3 Materials and methods

2.3.1 Strains and media

The University of Cape Town has a collection of 30 commercial saccharolytic *Clostridium* strains recovered from the NCP ABE fermentation plant in Germiston South Africa, which consists of either *C. saccharobutylicum* or *C. beijerinckii* strains. These strains, together with the reference type strains, *C. beijerinckii* NCIMB 8052^T, *C. acetobutylicum* ATCC 824^T and *C. saccharobutylicum* $NCP262^T$ as controls, were used in the species identification process.

All strains were stored as spore stocks suspended in distilled water and kept at 4 °C. Spore stock suspensions of 100 µL were heat shocked at 80 °C for 10 min and plated on Clostridial Basal Medium (CBM) (1% glucose, 0.4% yeast extract (Difco), 0.0000002% biotin (Sigma), 0.05% cysteine HCL (Sigma), 0.1% $NaHCO_3$ (Merck), 0.2% casamino acids (Difco), 0.0001% thiamine-HCL, 0.0001% para-aminobenzoic acid (pABA), 0.02% $MgSO_4.7H_2O$, 0.0007% $MnSO_4.H_2O$ and 0.001% $FeSO_4.7H_2O$ (Merck) (O'Brien and Morris, 1971). Cultures were grown in an anaerobic chamber (model 1024 Forma Scientific) with a gas phase of 5% H_2, 10% CO_2 and 85% N_2, at 37 °C, for up to 48 h. Where no growth was observed, the process was repeated on the less robust spore stocks, using an intital treatment 60-70 °C for 5-10 min.

2.3.2 Rifampicin antibiotic screening

Rifampicin resistance screening was carried out as described by Keis and colleagues (Keis *et al.*, 1995). In short, each strain was revived in triplicate and plated. Three colonies per triplicate strain were picked at random and inoculated into 10 mL CBM broth and allowed to

grow anaerobically for 18 h at 37 °C. This resulted in nine cultures, representing each NCP strain for the initial rifampicin screening. A 10% inoculum was transferred to fresh 10 mL CBM broth and allowed to grow to an OD_{600} of 0.3-0.4. A volume of 100 µL of the cell suspension was spread plated onto a CBM plate. Two 6 mm 100 ng rifampicin discs, two 6 mm 10 ng rifampicin discs and a 6 mm negative water control disc (N.T Laboratory Supplies (Pty) Ltd.) were applied to the plate surface. Plates were incubated anaerobically at 37 °C for 48 h. A zone of clearance of approximately 25 mm in diameter around the 100 ng rifampicin disc indicates rifampicin-sensitive *C. saccharobutylicum,* while a smaller zone of approximately 7 mm indicates rifampicin-resistant *C. beijerinckii*.

2.3.3 Genomic DNA extraction

Strains were grown in CBM broth for 18 h and incubated at 37°C and cells were collected by centrifugation at room temperature at 16000 *g* for 5 min. Genomic DNA was extracted by initially exposing cells to a lysis buffer (20 mM Tris- HCl pH8, 2 mM EDTA, 1.2% Triton X-100 (Merck) and 20 mg/mL lysozyme (Fluka Biochemika, Sigma-Aldrich, Belgium). Proteinase K (Fermentas, Thermoscientific, Inqaba Biotechnical Industries (Pty) Ltd., South Africa) was added to cell lysis mixtures at a final concentration of 50 µg/mL and incubated for 60 min at 37 °C. Genomic DNA was extracted from the cell lysate using the Fermentas Genomic Purification Kit (Fermentas, #K0512) according to kit instructions with 2 µg/mL Ribonuclease A (Fermentas, #EN0531) added and incubated for 10 min at room temperature prior to the DNA precipitation step. The extracted genomic DNA was re-suspended in 100 µL nuclease free water, and stored at -20 °C.

2.3.4 Colony PCR preparation

Colonies were re-suspended in a lysis buffer (10 mM Tris- HCL pH 8 (Sigma), 1 mM EDTA (Sigma)), thoroughly vortexed and Proteinase K (Fermentas) was added to a final concentration of 50 µg/mL. The mixture was incubated at 55 °C for 15 min to lyse the cells and further incubated at 80 °C for 15 min to inactivate the Proteinase K. Lysed cells were centrifuged at 5000 *g* for 1 min and 1 µL of the supernatant containing DNA was used in downstream PCR reactions. This procedure was also used for cell suspensions from a fermentation batch, where 100-500 µL cell suspensions were centrifuged at 10000 *g* for 2 min and the cell pellet was then re-suspended in the afore mentioned lysis buffer.

2.3.5 PCR conditions

Primers were synthesised and purified using HPLC (University of Cape Town Oligo Synthesis Service) and their site specificity was tested using BLAST (Altschul *et al.*, 1990) as well as by PCR. All PCR reactions were conducted using KAPA Ready Mix (Kapa Biosystems (Pty) Ltd., South Africa) unless otherwise stated. The primer sequence and amplicon size for each PCR is listed in Table 2.1 below. The RAPD P3 primer as well as the 16s *rRNA* 27F and R5 primers were synthesized according to Alonso *et al.* (2001) and Weisburg *et al.* (1991), respectively. PCR amplification reactions were performed separately in a final reaction volume of 25 µL, which contained 1x Ready Mix buffer, 0.5 µM of forward and reverse primers, 100 ng of genomic DNA and made up to volume with nuclease free water. PCR reactions were carried out using a GeneAmp 9700 PCR system (Applied Biosystems).

Table 2.1: A summary of the primers used in this study, including their nucleotide sequences and PCR product sizes.

Primer Name	Application	Nucleotide Sequence 5'-3'	Product size (kb)
***C. beijerinckii* Nit-F**	Species-specific PCR	AGGTAAAGCAGCTGAGTGGG	1.0
***C. beijerinckii* Nit-R**		TCAGCGCCTGTACCACC	
***C. saccharobutylicum* Nit-F**	Species-specific PCR	TGCAGGAAAGCAGCAGAATGGG	1.1
***C. saccharobutylicum* Nit-R**		TTCATCAGCTCCCGTACCGCC	
***C. acetobutylicum* Nit-F**	Species-specific PCR	AGCTTATGCCTTCTGAATGGG	1.2
***C. acetobutylicum* Nit-R**		TGCCTGTGCTAGTGGCCTTCCGTC	
EtfA-F	*etfA* RFLP Strain typing	TGCTGGTATGCAAGATTC	1.1
EtfA-R		TAATCATAGAATCCTTTTCCTG	
P3	RAPD Analysis	GTAGACCCG	Variable
27F	16s *rRNA* Analysis	AGA GTT TGA TCI TGG CTC AG	1.5
5R		ACGGITACCTTGTTACGACTT	
AtpD-F	MLST Analysis	GGAGATGAAGATAAGGCTGTCG	0.179
AtpD-R		GCTGATTCTGGCAATTCATCG	
GyrB-F	MLST Analysis	GAGATTCGGCAGGTGGATCTGC	0.543
GyrB-R		CATCGGCTGCCATAGCATCTTC	
***C. saccharobutylicum* OppB-F**	MLST Analysis	GAGGCTTATGCCAGGAGGTCC	0.75
***C. saccharobutylicum* OppB-R**	MLST Analysis	GCAGTTACTCCCAATAAAGTAGTATAATC	
***C. beijerinckii* OppB-F**	MLST Analysis	TCTCTATTTGTAACTGACTGTACG	0.722
***C. beijerinckii* OppB-R**		TTCTAATGAGACTTATGCCAGG	
RecA-F	MLST Analysis	GAACATAAGTTCGATGGAAGGGAG	0.683
RecA-R		GCTCTTCCTCCAGTTGTTGTTTC	
RpoB-F	MLST Analysis	ATTGCAGATGGTCCATCTACAG	0.431
RpoB-R		TTCCAGCTTCTCCGTGAGG	

2.3.5.1 16s *rRNA* PCR

The 16s *rRNA* PCR cycling profile, using primers 27F and R5, consisted of an initial denaturing step at 94 °C for 5 min, followed by 30 cycles of denaturing at 94 °C for 1 min, annealing of the primers at 55 °C for 30 s, an elongation step at 72 °C for 1.5 min, followed by a final elongation step at 72 °C for 5 min.

2.3.5.2 Species-specific PCR

The vision behind the species-specific PCR was for it to be a quick means of strain screening. To that end, primers were designed to function under the same PCR conditions for all three species, to facilitate rapid analysis. The primers for the species-specific PCR reactions were designed to have the same annealing temperature with a product size of around 1 kb for all three species (Table 2.1). The PCR primers and thermal conditions were designed based on whole genome nucleotide data stored on the NCBI database for *C. beijerinckii* NCIMB 8052^T, *C. acetobutylicum* ATCC 824^T and GS-GOGAT operon nucleotide sequence of *C. saccharobutylicum* $NCP262^T$ (Stutz *et al.*, 2007).

The species-specific PCR cycling, using either primers *C. beijerinckii* Nit-F and Nit-R, *C. saccharobutylicum* Nit-F and Nit-R or *C. acetobutylicum* Nit-F and Nit-R, began with an initial denaturing step at 95 °C for 5 min, followed by 30 cycles of denaturing at 95 °C for 30 s, annealing of the primers at 50 °C for 30 s, an elongation step at 72 °C for 1 min, followed by a final elongation step at 72 °C for 5 min.

2.3.5.3 *etfA* PCR

The *etfA* PCR cycling profile, using primers EtfA-F and EtfA-R, consisted of an initial denaturing step at 94 °C for 5 min, followed by 35 cycles of denaturing at 94 °C for 1 min, annealing of the primers at 50 °C for 1 min, an elongation step at 72 °C for 1.5 min, followed by a final elongation step at 72 °C for 5 min.

2.3.5.4 RAPD PCR

The RAPD PCR cycling profile, using primer P3 only, consisted of an initial denaturing step at 95 °C for 5 min, followed by 45 cycles of denaturing step at 95 °C for 1 min, annealing of the primers at 34 °C for 1 min, an elongation step at 72 °C for 2 min, followed by a final elongation step at 72 °C for 5 min.

2.3.5.5 MLST PCR

The MLST PCR reactions were carried out using High Fidelity (Hi-Fi) Ready Mix (Kapa Biosystems). The *atpD, rpoB, C. saccharobutylicum oppB, C. beijerinckii oppB* and *recA* gene targets had a shared PCR profile which consisted of an initial denaturing step at 98 °C for 2 min, followed by 35 cycles of denaturing at 98 °C for 20 s, annealing of the primer at 58 °C for 15 s, an elongation step at 72 °C for 15 s, followed by a final elongation step at 72 °C for 5 min.

The *gyrB* PCR profile was the same as above, except the annealing temperature was set to 55 °C for the same duration. All MLST primers listed were based on primers designed to *C. botulinum* ATCC 13124 (Macdonald *et al.*, 2011) and then modified to match the sequence of

NCIMB 8052^{T}, to ensure amplification of MLST genes within the NCP strains. Primer binding sites were conserved between both NCP *C. beijerinckii* and *C. saccharobutylicum* species for all MLST genes, with the exception of the *oppB* gene which required primers to be designed separately for each species.

All DNA gels produced in species-specific and MLST PCRs were electrophoresed in 0.8% agarose (SeaKem® LE Agarose LONZA USA) at 100 V for 1-2 h. Gels of 1.5% agarose were used for RAPD analysis and were electrophoresed at 110 V for 1 h. *etfA* PCR products were electrophoresed in 1.5% agarose at 90 V for 30 min. All agarose gels contained 0.5 µg/mL ethidium bromide (Sigma) and were electrophoresed in 1X TAE buffer (40 mM Tris-acetate and 1mM EDTA) (Sigma). After electrophoresis, the gels were visualised under short wavelength UV light using the GelDoc (BioRad) and photographed. The Lambda-*Pst*I DNA ladder (λ) in all gels was phage λ DNA (Fermentas), digested with *Pst*I (Fermentas) which was routinely used as a molecular marker. A 100 or 1000 bp DNA ladder (Fermentas) was used when PCR products of different sizes were visualised on the same agarose gels.

2.3.6 *etfA* RFLP analysis

A region of the *etfA* gene of *C. beijerinckii* and *C. saccharobutylicum* strains was amplified by PCR. A total of 500 ng of the PCR product was digested with 10 U *Nla*III, 1x *Nla*III buffer and 100 ug/mL BSA (New England BioLabs) at 37 °C for 2 h, prior to electrophoresis in a 1.5% agarose gel. The resultant digest patterns were analysed visually under UV as described in 2.3.5.

2.3.7 RAPD analysis

Quantity One software (BioRad) was used for the RAPD analysis. The software uses a Dice algorithm and lanes/samples were grouped according to DNA band distance migrated (band size). A UPGMA (Unweighted Pair Group Method with Arithmetic Mean) phylogenetic tree was then constructed based on the patterning. The *C. saccharobutylicum* strains were analysed and grouped separately from the *C. beijerinckii* strains.

2.3.8 MLST analysis

The PCR products of the five genes to be used for MLST were purified using the BioFlux Biospin PCR Purification Kit (Catalogue # BSC03S1, Bioer Technology, Hangzou, P.R. China), according to the supplier's instructions. The purified product was then quantified using the NanoDrop (Nanodrop ND-330, Inquaba Biotechnical Industries (Pty) Ltd. South Africa) and ligated into PJET1.2vector (Thermo Scientific, #K1231) transformed into *E. coli* DH5α cells, according to the supplier's instructions. The recombinant plasmids were purified using a Bioflux Miniprep Kit (#BSC01S1, Bioer Technology) and sequenced (Macrogen Inc., Korea). Sequence data for strains NCP260, NCP262^{T} and NCP258 were supplied by Green

Biologics Ltd. (personal correspondence). MLST and 16s *rRNA* sequences were edited using Clustal V2.01 (Technelysium (Pty) Ltd., Australia) and aligned using Clustal W from the Mega software (Mega5.0 ©1993-2011) (Tamura *et al.*, 2011). Gene segments were arranged in the same orientation in the order *atpD*, *rpoB*, *gyrB*, *recA* followed by *oppB*, to create a concatenated gene sequence or pseudo gene of 2512 bp, as seen in Figure 2.1 below. Maximum Parsimony trees were generated with bootstrap values of 1000 (Takahashi and Nei, 2000).

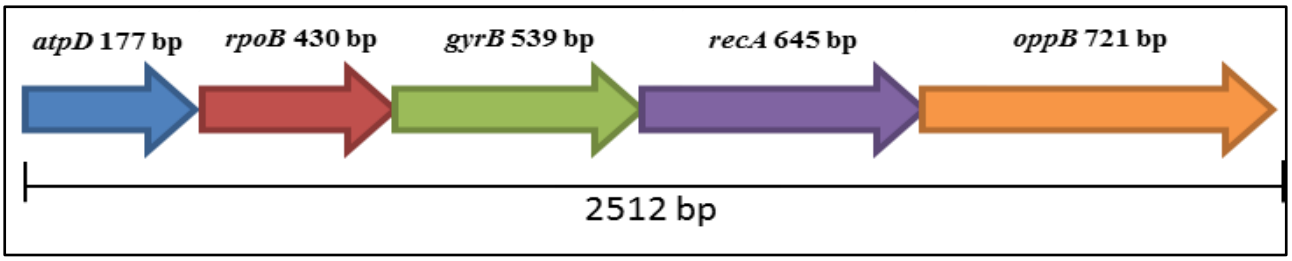


Figure 2.1: Combined sequence order and size of each MLST gene used in MLST analysis, assembled for individual strains to create a 2512 bp pseudogene.

2.4 Results and Discussion

2.4.1 Species level characterisation of the UCT NCP strain collection

2.4.1.1 16s *rRNA* gene sequencing

Five NCP strains, comprised of two *C. saccharobutylicum* and three *C. beijerinckii* strains were selected for 16s *rRNA* sequencing using the universal primers F27 and 5R (Weisburg *et al.*, 1991), to determine whether any strain-level taxonomic groupings were possible. DNA sequence from a number of solventogenic and non-solventogenic members of the genus were obtained from NCBI for comparison in the 16s *rRNA* analysis. While analysis of the 16s *rRNA* sequence could differentiate between the NCP *C. saccharobutylicum* and *C. beijerinckii* species, the 16s *rRNA* segment contained insufficient nucleotide variability to discern differences between NCP strains within the same species (Figure 2.2).

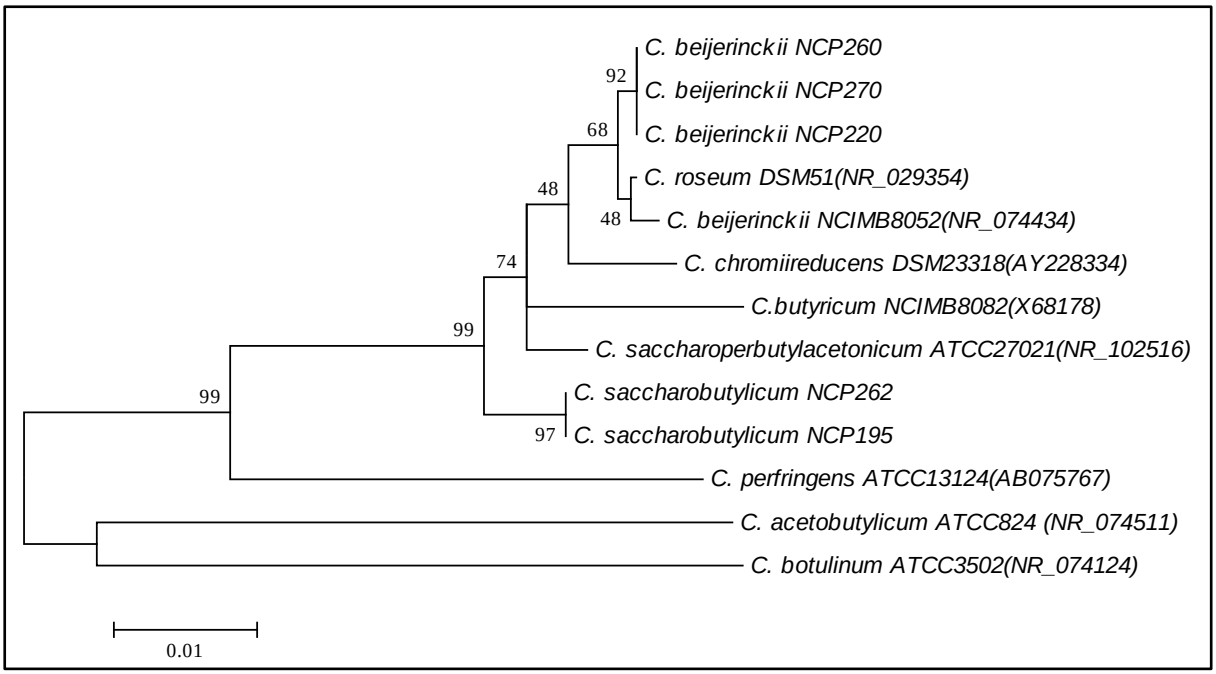


Figure 2.2: Unrooted 16s *rRNA* gene phylogenetic tree based on 1374 bp of amplified sequence for NCP strains compared to other members of the genus *Clostridium* using the Maximum Likelihood method. Bootstrap values calculated from 1000 resampled datasets. One scale bar =0.01 nucleotide substitutions per nucleotide position.

This preliminary analysis revealed several intriguing facts regarding the taxonomic nature of the NCP strain collection. Firstly, strains within a species are very closely related to one another, as is evident by strains within a species clustering together to form their own clade. Secondly, the *C. beijerinckii* strain NCIMB 8052^T clusters away from the NCP *C. beijerinckii* strains analysed. This division between strain NCIMB 8052^T and the NCP *C. beijerinckii* strains indicates that the NCP *C. beijerinckii* strains are evolutionarily distinct from NCIMB 8052^T.

Thirdly, *Clostridium roseum* DSM51^T appears to be more closely related to the NCIMB 8052^T than NCP *C. beijerinckii* strains, as depicted by both of the former species groups sharing a common tree node, yet this strain has been shown to be genetically distinct enough

to be designated its own species. Whilst very little is known about this species, it appears as if *C. roseum* does not produce ABE solvents, but does produce butyric and acetic acid. It can generally be classed as a hydrogen gas producer (Calusinska *et al.*, 2011). Despite the differences in metabolism, non-solventogenic *C. roseum* is still closely related to solventogenic *C. beijerinckii*. The bootstrap value within the *C. beijerinckii- C. roseum* clade of 68 indicates that there is nucleotide variability amongst the strains and the remaining 32% of the time the position of the non-NCP strains in the tree are probably subject to change. The short branch lengths within this clade indicate that all *C. beijerinckii* strains and the strains within that clade are closely related according to the 16s *rRNA* region analysed. This highlights the importance of both phenotypic and genetic characterization tests to assign new species status to new strains.

Clostridium chromiireducens DSM 23318^{T} is another non-solventogenic member of the genus which is involved in reduction of chromium and iron in contaminated soils and belongs to Group 1 of the genus *Clostridium* (Inglett *et al.*, 2011). It is interesting to note that the bootstrap value of 48 is not strong evidence for the position of *C. chromiireducens* in the tree. The solventogenic species *C. saccharoperbutylacetonicum* ATCC 27021^{T} forms its own distinct clade with a bootstrap value of 74, indicating that its position within the tree is likely. However, it does cluster by itself and has a much longer branch length, indicating it is not closely related to *C. beijerinckii* strains, as depicted by separate clades denoting a clear species barrier. The *C. saccharobutylicum* NCP strains cluster in their own clade with high identity between strains of the species observed once again, with a bootstrap value of 97 indicative of good support for clade position. This sequence homology of the 16s *rRNA* region is typified by the sequence identity values depicted in Table 2.2 below. BLAST results for the respective species are unsurprisingly not below 99% identity for the two species.

Table 2.2: BLAST results for selected NCP strains, containing BITS score and % identity for the 16s *rRNA* gene region.

Query Strain	Closest Homology	Nucleotide Identity	% Identity	BITS Score
NCP262^T	*C. saccharobutylicum* P262^T	1375/1375	100%	2540 bits (1375)
NCP195	*C. saccharobutylicum* P262^T	1375/1375	100%	2540 bits (1375)
NCP260	*C. beijerinckii* NCIMB 8052^T	1373/1374	99%	2532 bits (1371)
NCP220	*C. beijerinckii* NCIMB 8052^T	1374/1375	99%	2532 bits (1371)
NCP270	*C. beijerinckii* NCIMB 8052^T	1373/1374	99%	2532 bits (1371)

2.4.1.2 Colony-based species-specific PCR

Since the *glnA-nitR-gltAB* gene configuration in *C. saccharobutylicum* and *C. beijerinckii* is unusual amongst solventogenic clostridia, there is high species-level specificity (Amon *et al.*, 2010). In order to develop a rapid, highly-specific PCR-based identification procedure, species-specific primers for *C. saccharobutylicum*, *C. beijerinckii* and *C. acetobutylicum* strains were designed to these regions in the genome. The variable regions within the 3' end of *glnA* and the 5' end of *gltA* flanking the nitrogen regulator (*nitR*) of the nitrogen operon were used for primers for *C. saccharobutylicum* and *C. beijerinckii* species (Figures 2.3A and 2.3B). In the case of *C. acetobutylicum* ATCC 824^T, the operon configuration is not conserved, with no detectable *nitR*-like regulator gene present associated with the large glutamate synthase (*gltA*) gene. Primers were therefore designed against the intergenic region between the DNA helicase and the *gltA* gene and internal to the *gltA* gene (Figure 2.3C).

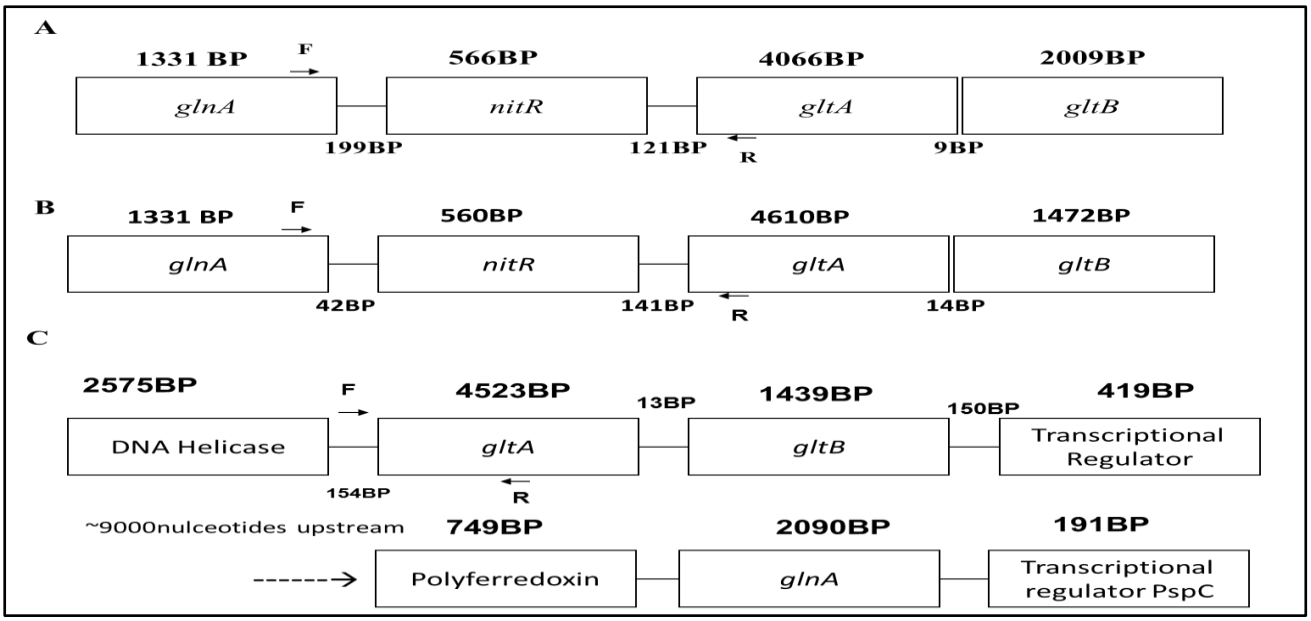


Figure 2.3: A schematic representation of the nitrogen assimilation gene configurations for **A:** *C. saccharobutylicum* NCP262^T, **B:** *C. beijerinckii* NCIMB 8052^T and **C:** *C. acetobutylicum* ATCC 824^T species. Arrows with an F and R denote forward and reverse primer binding sites respectively.

Colony PCR using the *Clostridium saccharobutylicum* species-specific primers was performed on selected strains. The product generated was specific only to *Clostridium saccharobutylicum* strains (Figure 2.4A). PCR was performed using species specific primers for *C. beijerinckii* (Figure 2.4B) and *C. acetobutylicum* strains (Figure 2.4C). We cannot rule out the possibility that other species outside of the three major species in this study may also have nitrogen assimilatory genes in this same configuration. However, for the purposes of this study, this colony PCR method has demonstrated adequate species level specificity within the NCP collection and no product was observed for the related *C. difficile* (data not shown) or Gram negative *E. coli* (Figures 2.4A-C).

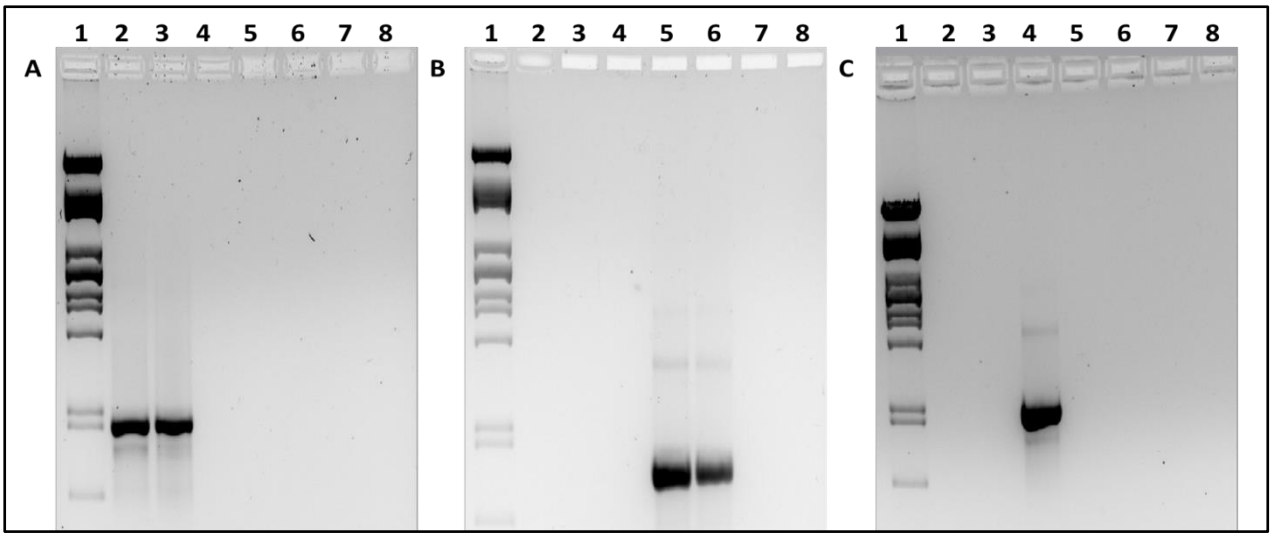


Figure 2.4: Species-specific colony PCR test for **A:** *C. saccharobutylicum*, **B:** *C. beijerinckii* and **C:** *C. acetobutylicum* strains. Lane 1, λ-*Pst*I DNA marker; 2, NCP262^T; 3, NCP108; 4, ATCC 824^T; 5, NCIMB 8052^T; 6, NCP260; 7, *E. coli* JM109 and 8, no template control.

Since Gram positive bacterial cells generally have a resistant cell wall (Niwa *et al.*, 2005), conventional colony PCR success of the NCP strains was inconsistent, possibly due to poor lysis of the cells. A cell detergent-lysozyme lysis treatment step was therefore included to more readily access the DNA for the subsequent PCR step. However, we discovered that the annealing temperature had to be kept lower than the theoretical optimum to ensure that positive test products were amplified for all NCP strains outside of NCIMB 8052^T, NCP262^T and ATCC 824^T. This lower thermal PCR profile ensured that all positive PCR products were correctly amplified. Either the low PCR annealing temperature or residue from the lysis cell pre-treatment may account for non-specific DNA bands seen in Figure 2.4B and 2.4C. If genomic DNA was used, then single amplification product was observed.

This screening method used colonies from an agar plate or samples from a fermentation batch culture and took a total of two hours to complete, which included lysis, PCR, gel electrophoresis and DNA imaging. This is a quick screening method suitable for laboratory or

industry and is highly applicable to the NCP strains. This screening method can also be used to confirm presence of desired species in batch without the need to stop the fermentation process for laborious taxonomic testing or worse yet discard the whole batch.

2.4.1.3 Rifampicin screening

The strains listed for 16s *rRNA* sequencing results in Table 2.2 above, were used as reference strains for the development of the species-specific PCR. The strain identity was confirmed by rifampicin antibiotic screening. As expected, nucleotide homology for the 16s *rRNA* region of the NCP *C. beijerinckii* strains matched NCIMB 8052^{T} closest and the NCP *C. saccharobutylicum* strains displayed closest identity to $NCP262^{T}$. The overall species level strain ID for each strain in the UCT NCP collection is listed below in Table 2.3. All strains were subjected to thorough rifampicin screening, which was validated by species-specific PCR performed on genomic DNA from a single colony. The UCT strain collection consists of exclusively *C. saccharobutylicum* and *C. beijerinckii* species.

Table 2.3: A summary of all University of Cape Town NCP strains and respective species allocation by species-specific PCR with a positive colony PCR product indicated as + and no PCR product as -. *C. acetobutylicum* ATCC 824^T is represented as a double negative. Rifampicin resistance is indicated by (R) and sensitivity by (S). Strain numbers consisting of mixed species are denoted as cb for *C. beijerinckii* and cs for *C. saccharobutylicum*.

NCP Name	**Strain Name**	***C. saccharobutylicum* Species-specific PCR**	***C. beijerinckii* Species-specific PCR**	**Rifampicin**
Type Strains:				
NCP262	*C. saccharobutylicum* P262^T	+	-	S
8052	*C. beijerinckii* NCIMB 8052^T	-	+	R
824	*C.acetobutylicum* ATCC 824^T	-	-	S
***C. beijerinckii* strains:**				
NCP106	UCT401	-	+	R
NCP172	UCT402	-	+	R
NCP193	UCT403	-	+	R
NCP195cb	UCT211	-	+	R
NCP200	UCT407	-	+	R
NCP202	UCT416	-	+	R
NCP220	UCT210	-	+	R
NCP254	UCT404	-	+	R
NCP258cb	UCT408	-	+	R
NCP259	UCT409	-	+	R
NCP260	UCT410	-	+	R
NCP261	UCT417	-	+	R
NCP263	UCT411	-	+	R
NCP264	UCT412	-	+	R
NCP270	UCT413	-	+	R
NCP271	UCT418	-	+	R
NCP272	UCT414	-	+	R
NCP280	UCT415	-	+	R

NCP J	UCT406	-	+	R
NRRL B593		-	+	R
***C. saccharobutylicum* strains:**				
NCP108	UCT202	+	-	S
NCP162	UCT203	+	-	S
NCP195cs	UCT211	+	-	S
NCP199	UCT204	+	-	S
NCP206	UCT206	+	-	S
NCP 200α206	UCT205	+	-	S
NCP249	UCT207	+	-	S
NCP258cs	UCT408	+	-	S
NCP265	UCT208	+	-	S
NCP268	UCT209	+	-	S

Most of the spore stocks for strains assigned an NCP strain number were shown to be of mixed species according to the last taxonomic studies conducted by Keis and colleagues (Keis *et al.,* 2001). The mixed strains were streaked out for single colonies and the strains listed in this study are pure insofar as the strain consists of one species. Strains NCP258 and NCP195 were found to be mixed, i.e. both *C. saccharobutylicum* and *C. beijerinckii* species were detected for a given spore stock, which differed from previous findings that stated these strains were *C. saccharobutylicum* strains. It was noticed in a previous paper by Keis and colleagues (Keis *et al.* 2001) that NCP258 and NC195 had a sensitive rifampicin phenotype resembling that of *C. acetobutylicum* (now *C. saccharobutylicum*). Both the *C. beijerinckii* and *C. saccharobutylicum* variants for NCP258 and NCP195 were isolated and streaked for single colonies. Successive generations were tested by species-specific PCR and 16s *rRNA* gene sequencing and separate strain spore stocks made for each species, in the case of these mixed stocks, with the result that the different species variants have cs and cb added to their

strain names to denote *C. saccharobutylicum* and *C. beijerinckii* species variants respectively. In the subsequent strain analyses, NCP258cs was used and is simply denoted NCP258.

The high occurrence of mixed NCP spore stocks is not all that surprising when one considers the historic nature of the original NCP strain collection used in the Germiston NCP facility three decades ago, and the ease with which spores can disperse in a laboratory. The taxonomy at the time was very limited and there were no physiological or genetic tools available to differentiate between the two major species within the NCP collection. To ensure spore stock purity, the entire spore stock suspension (not just single colonies) was subjected to rifampicin screening and the resultant rifampicin phenotype was used as confirmation of species identity. In any event, colonies were picked from all the original UCT NCP stocks and new spore stocks were made, based on the species identity listed in Table 2.3.

2.4.2 Strain level characterisation of the UCT NCP strain collection

2.4.2.1 *etfa* RFLP analysis

In order to develop an alternative strain screening method, any significant differences in known nucleotide regions were evaluated. The electron transfer flavoprotein alpha (*etfA*) gene was considered as a candidate region for restriction enzyme digest analysis. The EtfA protein is involved with the electron transport chain in the formation of acetoacetyl-CoA and butyryl-CoA, precursors to acetone and butanol formation. The region is variable at a nucleotide level, with high levels of variation between the major *C. beijerinckii* and *C. saccharobutylicum* species (Boynton *et al.*, 1996). The *etfA* region was digested using the restriction enzyme *Nla*III, since this has been found to be effective in discriminating between the banding patterns in the *eftA* region of solventogenic clostridial species (personal correspondence, Green Biologics Ltd., UK).

The amplified nucleotide regions were found to be similar within each species and the *Nla*III restriction enzyme digest did not offer any variation between the NCP *C. beijerinckii* strains (Figure 2.5A), as denoted by the conserved four band pattern generating 450, 356, 280 and 130 bp products respectively. However, they differed considerably from *C. beijerinckii* NCIMB 8052^{T} which generated 650, 450 and 130 bp products. The *C. saccharobutylicum* strains showed a conserved single band pattern with no visible difference between undigested control and the digested product of 1.1 kb due the presence of *Nla*III restriction enzyme sites too close to the edges of the amplicon (Figure 2.5B). The *etfA* RFLP screening method, whilst good for grouping a small collection of diverse strains into separate species or strain classes, provides too small a region to offer any insights into differences between closely related strains within the same species, such as the NCP strain collection.

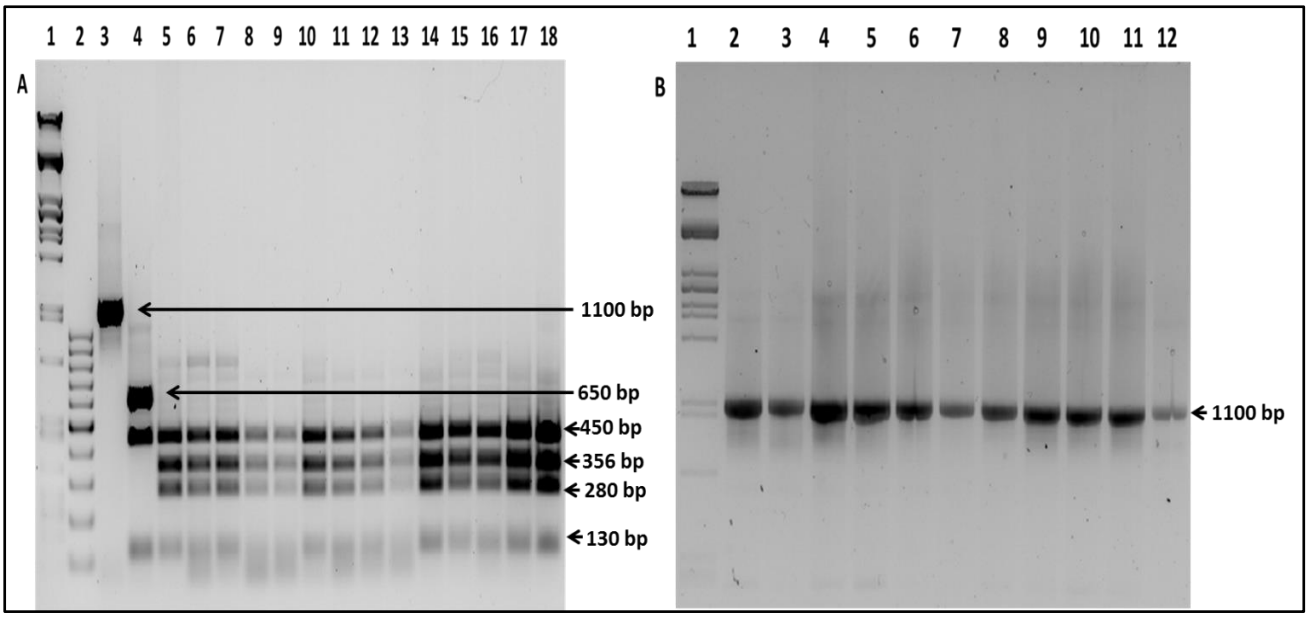


Figure 2.5A: *Nla*III RFLP of *etfA* region for *C. beijerinckii* NCP strains: lane 1, λ-*Pst*I marker; 2,100 bp marker; followed by RFLP of *etfa* regions for: 3, undigested NCIMB 8052^{T}; 4, NCIMB 8052^{T}; 5, NCP106; 6, NCP J; 7, NCP193; 8, NCP200; 9, NCP202; 10, NCP220; 11, NCP259; 12, NCP260; 13, NCP263; 14, NCP264; 15, NCP270; 16, NCP271; 17, NCP272 and 18, NCP172. **B:** *Nla*III RFLP of *etfA* region for *C. saccharobutylicum* NCP strains: lane 1, λ-*Pst*I marker; followed by RFLP of *etfA* regions for: 2, undigested NCP262^{T}; 3, NCP262; 4, NCP108; 5, NCP162; 6, NCP195; 7, NCP199; 8, NCP206; 9, NCP249; 10, NCP265; 11, NCP268 and 12, NCP258.

While it is true that *etfA* RFLP can distinguish between species by looking at the banding patterns generated, the species-specific colony PCR method is faster, since one can use colonies and not treated genomic DNA, as is a prerequisite for the 2 h *Nla*III digest. The grouping pattern observed in the *etfA* analysis is echoed by results seen in the 16s *rRNA* phylogenetic analysis mentioned previously (Figure 2.2). It is interesting to note that the last major review of the NCP strain collection conducted by Keis and colleagues, assigned NCP *C. beijerinckii* strains into "sub-group 3" and NCIMB 8052^{T} into "sub-group 2" for the species (Keis *et al.*, 2001). This is further proof that *C. beijerinckii* NCIMB 8052^{T} and the NCP *C. beijerinckii* strains are genetically diverse enough for the latter strains to be classified into a distinct sub-species grouping, when one considers solvent performance differences

between NCIMB 8052^{T} and the NCP *C. beijerinckii* strains, which will be covered in the next chapter.

2.4.2.2 RAPD analysis

Since the *etfA* gene region is too highly conserved in this group of bacteria to detect differences within NCP strains, a PCR technique based on the whole genome was considered in an attempt to sub-group the NCP strains. *C. beijerinckii* strains exhibited four strain groups, whilst the *C. saccharobutylicum* strains were less variable and consisted of two strain groups (Figures 2.6A and 2.6B).

It is interesting to note that two out of the ten *C. saccharobutylicum* strains fell into group 1 and the remaining eight strains fell into group 2 with respect to banding patterning for RAPD analysis (Figure 2.6A). This implies that most of the NCP *C. saccharobutylicum* strains are closely related and whilst strains may differ phenotypically even within the same species group, this conservation of banding patterns suggests a high level of genotypic relatedness within the species. The *C. beijerinckii* strains show four distinct groups, with NCP 195Cb showing an independent banding pattern. Nine out of the 16 *C. beijerinckii* strains belonged to strain group 2, two belonged to strain group 3 and four to strain group 4 (Figure 2.6B).

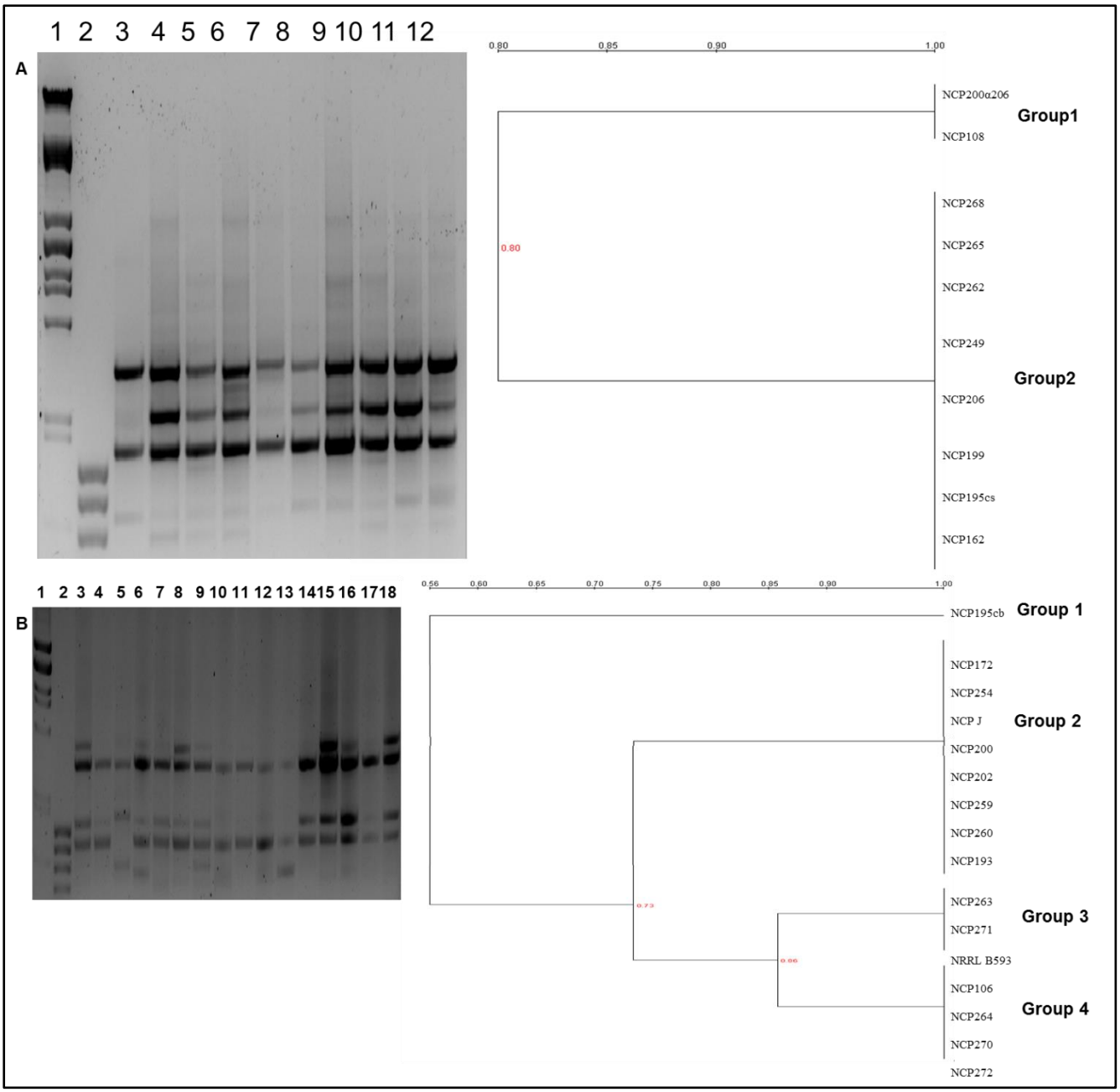


Figure 2.6: RAPD Analysis and clustering of NCP species using UPGMA analysis for **A:** *C. saccharobutylicum*: lane 1, λ-*Pst*I marker; 2, 100 bp marker; 3, NCP108; 4, NCP162; 5, NCP195cs; 6, NCP199; 7, NCP200α206; 8, NCP206; 9, NCP249; 10, NCP262; 11, NCP265 and 12, NCP268. **B:** *C. beijerinckii strains*: lane 1, λ-*Pst*I marker; 2, 100 bp marker; 3, NCP106; 4, NCP172; 5, NCP195cb; 6, NCP254; 7, NCP J; 8, NRRL B593; 9, NCP193; 10, NCP200; 11, NCP202; 12, NCP259; 13, NCP260; 14, NCP263; 15, NCP264; 16, NCP270; 17, NCP271 and 18, NCP272.

RAPD analysis indicates that there is greater genetic variability or strain diversity between the NCP *C. beijerinckii* strains compared to the NCP *C. saccharobutylicum* strains, as is

evident by the four *C. beijerinckii* strain groups compared to the two strain groups for *C. saccharobutylicum* strains. The differences in RAPD banding patterns may reflect phenotypic differences between strain groups. Any conserved, distinct bands shared between high performance solvent producing strains can be candidates for further sequencing and study. The DNA contained in these bands may contain a deletion or addition mutation to the genome that enables better solvent production.

Whilst RAPDs cannot offer definite NCP strain level grouping, an indication of strain clustering can be achieved with RAPD analysis, since it is probable that strains with similar or identical nucleotide sequences will exhibit similar banding migration patterns. However, RAPDs are very subjective since the grouping software relies on DNA bands on a gel image to present clearly. The low annealing temperature used in RAPDs also means that the genomic DNA template must be fresh and of high quality, since any impurities left over from the DNA extraction isolation could result in non-specific band amplification or no product yield at all (Williams *et al.*, 1990). There is also a certain amount of laboratory specificity, with different labs getting different RAPD patterns, either due to specificity of different Taq polymerases, electrophoresis or PCR conditions. It is advised that RAPD analysis be used as a preliminary screen to see if there are any obvious strain clusters for a given strain library.

A different range of clostridial RAPD primers can be used, which may have different specificities and generate entirely different DNA banding patterns. However, this RAPD analysis has confirmed diversity of the NCP strains within *C. saccharobutylicum* and *C. beijerinckii* species.

2.4.2.3 MLST Analysis

For best strain library results a combination of RAPD and MLST analysis should be performed (Vanhee *et al.*, 2009). MLST requires sequencing information and knowledge of target housekeeping genes of interest, which form the basis of nucleotide comparisons. Since MLST uses a suite of housekeeping genes with a PCR step in each, there is a chance that primer binding does not occur (Urwin and Maiden, 2003). MLST was implemented to confirm differences between strains at a nucleotide level, in the form of SNPs, so that the strains could be grouped more effectively. Primer sites and housekeeping genes were chosen as in Macdonald *et al.* (2011) for *C. botulinum* strains. The primer binding sites were modified to be specific to the NCP strains, using whole genome information based on *C. beijerinckii* NCIMB 8052^T and *C. saccharobutylicum* NCP262^T.

MLST was performed on a selection of NCP strains, based on the different groups observed in the RAPD analysis. *C. saccharobutylicum* strains NCP108 (RAPD strain group 1), NCP206, NCP262^T, NCP258 and NCP265 (RAPD group 2) along with *C. beijerinckii* strains NCP260 (RAPD group 2), NCP202 (RAPD group 2), NCP172 (RAPD group 2) and NCP J (RAPD group 2) were selected for MLST analysis. The genes used were *atpD, rpoB, gyrB, recA* and *oppB,* using the primers listed in Table 2.1. Phylogenetic analysis was performed on the DNA sequence of each gene in isolation (Appendix A, Figures A1-5), as well as on an assembly of a pseudogene arranged as described in Figure 2.1, to form the consensus phylogenetic tree shown in Figure 2.7 below.

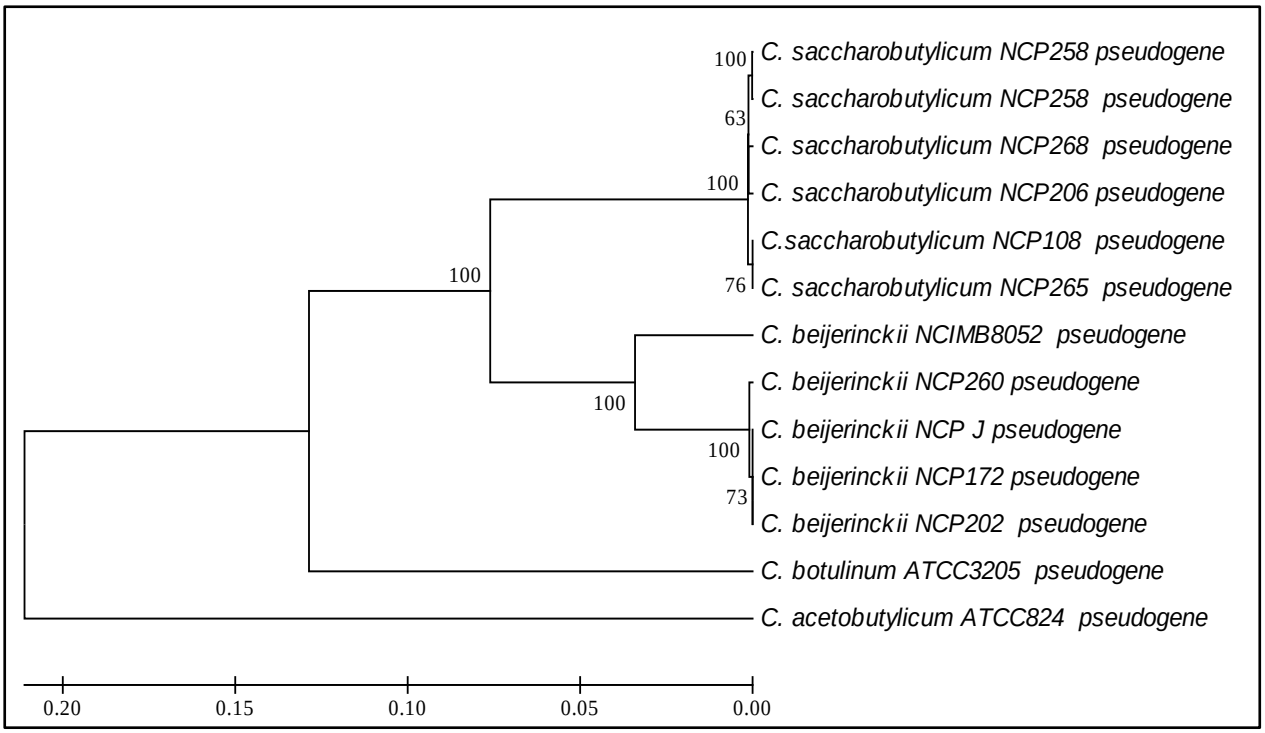


Figure 2.7: Assembled pseudogene phylogenetic tree based on 2512 bp of amplified sequence for NCP strains, compared to other members of the genus *Clostridium,* using the Maximum Likelihood method. Bootstrap values calculated from 1000 resampled datasets. One scale bar =0.05 nucleotide substitutions per nucleotide position.

The analysis of several strains within the NCP collection reveals some interesting insights into the strain grouping patterns at a nucleotide level between closely related NCP strains within the *C. beijerinckii* and *C. saccharobutylicum* NCP species. The most apparent observation is that the MLST analysis of the pseudogene as depicted in Figure 2.7 has much higher resolving power than that of the 16s *rRNA* analysis performed, as shown in Figure 2.2. This is evident by the fact that the MLST analysis resulted in four distinct strain clade groupings for *C. saccharobutylicum* species and in two clade groupings within the NCP *C. beijerinckii* species investigated. The node bootstrap values of 100 in all cases support the division of the species (Figure 2.7).

As depicted in the 16s *rRNA* phylogenetic analysis in Figure 2.2, it does appear as though the NCP strains within each of the *C. beijerinckii* and *C. saccharobutylicum* species analysed in Figure 2.7 are closely related to one another at a nucleotide level, since the branch lengths remain very short for the different groups within the species clades. This is the result of only a handful of SNPs evident within each housekeeping gene investigated. The nucleotide variability within each housekeeping gene differs substantially. This is a result of the combination of the length of the targeted housekeeping gene and mutations caused by selective pressures, through evolution of that particular gene (Maiden, 2006).

MLST analysis showed that the NCP *C. saccharobutylicum* strains NCP258 and NCP262^{T} cluster together with a bootstrap value of 100, indicating good placement. This clustering was retained in all five individual trees (Appendix A, Figures A1-5), confirming the close relationship between these strains. Strain NCP268 and NCP206 clustered together in the next group, with a bootstrap value of 63. NCP108 and NCP265 clustered together with a bootstrap value of 77. The second and third MLST *C. saccharobutylicum* groups with values of 63 and 77, indicate that there is nucleotide variation in these strains and that the groups assigned are supported 63% and 77% of the time throughout the 1000 bootstrap iterations of the dataset. These values make sense since only the *gyrB* gene region sequence (Appendix A, Figure A2) contained enough SNPs to differentiate between strains NCP268-NCP206 and NCP108-NCP265 to form their own respective clades.

In the case of NCP *C. beijerinckii* strains, there is once again a clear evolutionary distinction between NCIMB 8052^{T} and the NCP *C. beijerinckii* strains, as is evident by the bootstrap value of 100 at the intersecting node between the afore mentioned two strain groups and the

larger branch lengths. Interestingly, whilst NCP260 is closely related to the other NCP *C. beijerinckii* strains, it is has enough nucleotide variability within the pseudogene investigated to warrant segregation from its *C. beijerinckii* NCP relatives. A bootstrap value of 100 supports this division, although this variability is not retained in the *gyrB* and *recA* trees (Appendix A, Figures A2 and A5). The remaining *C. beijerinckii* strains NCP J, NCP172 and NCP202 are grouped together with a bootstrap value of 70 (Figure 2.7).

C. botulinum ATCC 3205 and *C. acetobutylicum* ATCC 824^T were used as outgroups in the MLST analysis, since these strains are evolutionarily distant relatives within the genus. It is interesting to note that, as in the 16s *rRNA* analysis in Figure 2.2, MLST indicates that ATCC 824^T is distantly related to *C. beijerinckii* and *C. saccharobutylicum* species. This highlights the importance of choice of genes used in MLST analysis, since targeting other genes related to solvent production may result in a shift in apparent relatedness between solventogenic *C. acetobutylicum* ATCC 824^T and its NCP *C. saccharobutylicum* and *C. beijerinckii* relatives.

The efficacy of housekeeping genes as a means of sequence comparison is effectively measured by the abundance of nucleotide variation or SNPs contained within a housekeeping gene or gene region. The 5' and 3' ends of the regions are typically conserved throughout a particular bacterial population. Table 2.4 below depicts the SNP frequencies found within the housekeeping gene segments used in this study. The precentage changes for *C. acetobutylicum* and *C. botulinum* species, whilst interesting, serve merely as an indication that there are SNP differences detected between the NCP strains and other members of the genus. Since 16s *rRNA* can adequately discriminate between species, using MLST at a genus level is not particularly viable or feasible, but it does give indications of relatedness at a sub-

species level. NCP262^T served as the reference strain for NCP *C. saccharobutylicum* strains in the SNP frequency analysis, whilst NCP260 was chosen as the reference strain for NCP *C. beijerinckii* strains rather than the type strain NCIMB 8052^T. This enabled the direct comparison of the relationship between the NCP *C. beijerinckii* strains, because it has already been established that the NCIMB 8052^T strain is genetically different from its NCP relatives.

Table 2.4: A list of SNP frequencies for the housekeeping genes used in the MLST analysis. All *C. beijerinckii*, *C. acetobutylicum* and *C. botulinum* strain SNP frequencies are relative to NCP260 and all *C. saccharobutylicum* strain SNP frequencies are relative to NCP262^T.

C. saccharobutylicum	**Strain**	**Number of SNPs per gene length (%polymorphisms)**					
		***atpD* (177bp)**	***gyrB* (539bp)**	***oppB* (721)**	***rpoB* (430bp)**	***recA* (645bp)**	**Pseudogene (2512bp)**
	NCP108	2 (1.2%)	4 (0.7%)	0	3 (0.7%)	1 (0.2%)	10 (0.4%)
	NCP206	2 (1.2%)	8 (1.5)	0	3 (0.7%)	1 (0.2%)	14 (0.6%)
	NCP265	2 (1.2%)	4 (0.7%)	0	3 (0.7%)	1 (0.2%)	10 (0.4%)
	NCP268	2 (1.2%)	3 (0.6%)	0	3 (0.7%)	1 (0.2%)	9 (0.4%)
	NC262^T	0	0	0	0	0	0
	NCP258	0	0	1 (0.1%)	0	0	1 (< 0.1%)
C. beijerinckii	**NCIMB 8052^T**	9 (5.1%)	3 (0.6%)	156 (21.6%)	7 (1.6%)	4 (0.6%)	179 (7.1%)
	NCP260	0	0	0	0	0	0
	NCP J	2 (1.2%)	0	2 (0.3%)	1 (0.2%)	0	5 (0.2%)
	NCP172	2 (1.2%)	0	2 (0.3%)	1 (0.2%)	0	5 (0.2%)
	NCP202	2 (1.2%)	0	2 (0.3%)	1 (0.2%)	0	5 (0.2%)
C. botulinum	**ATCC 3502**	36 (20.3%)	93 (17.3%)	191 (26.5%)	59 (13.7%)	142 (22.0%)	521 (20.7%)
C. acetobutylicum	**ATCC 824^T**	43 (24.3%)	138 (25.6%)	325 (45.1%)	77 (17.9%)	155 (24.0%)	738 (29.4%)

DNA sequence analysis of the *atpD* gene follows the same grouping trend as the overall pseudogene analysis. It should be noted that this region is 602 bp in *C. botulinum*. However, the corresponding gene in *C. beijerinckii* NCIMB 8052^T is only 480bp and the reverse primer site had to be modified to give a 179bp PCR product to facilitate PCR amplification. The *atpD* region presented the highest SNP frequency per gene length despite its short length (Table 2.4). This region could be expanded to provide greater resolving power amongst strains of the same species.

Closer inspection of the *gyrB* gene sequence analysis showed that the NCP *C. saccharobutylicum* strains NCP262^T-NCP258 and NCP 265-NCP108 group together and NCP268 and NCP206 form their own distinct groups within the larger cluster. There was no discriminatory resolving power within the NCP *C. beijerinckii* strains (Appendix A, Figure A2). This is emphasized by 0% SNP frequency values for NCP *C. beijerinckii* strains seen in Table 2.4. The inability of the *gyrB* gene sequence to separate the *C. beijerinckii* strains from one another at a nucleotide level may indicate that, based on this analysis, the NCP *C. beijerinckii* strains are relatively more closely related to each other than the NCP *C. saccharobutylicum* strains.

Investigation of the *oppB* gene revealed a reversal in the trend: the NCP *C. saccharobutylicum* strains, except NCP258, clustered together as one group within the species, whilst NCP *C. beijerinckii* strains have the typical split between NCP260 and the rest of the NCP strains (Appendix A, Figure A3). This observation is supported by 0% SNP frequencies for all NCP *C. saccharobutylicum* strains, except NCP258, which only has one SNP for the region, compared to NCP262^T and two SNPs separating NCP260 from the rest of

the *C. beijerinckii* strains (Table 2.4). Interestingly, the *oppB* gene shows the highest level of species level separation out of the candidate genes used in this analysis, with SNP frequencies of 26.5% and 45.1% for *C. botulinum* and *C. acetobutylicum* species respectively, compared to NCP260. There is SNP frequency of 21.6% for NCIMB 8052^T, compared to NCP260, which once again highlights how genetically distinct NCIMB 8052^T is compared to the NCP *C. beijerinckii* strains. This is accentuated by the larger branch lengths between NCIMB 8052^T and NCP *C. beijerinckii* strains.

Examination of the *rpoB* phylogenetic trees revealed *C. saccharobutylicum* NCP262 and NCP258 appear evolutionarily distinct from their counterparts, which cluster together separately within the species with high bootstrap values supporting this intra-species division (Appendix A, Figure A4). There is very little variation amongst the NCP *C. beijerinckii* strains with NCP260 clustering on its own by virtue of a single SNP in the *rpoB* region analysed. This is exemplified by the conservation of SNP frequencies between sub groups within the NCP species (Table 2.4). The *recA* gene region appears to be least effective in separating the NCP strains within both species. There is no resolving power within the *C. beijerinckii* species and only strains NCP 262^T and NCP258 are distinguished from the rest of the *C. saccharobutylicum* species. This poor resolving power is reflected in the SNP frequencies seen in Table 2.4, with no SNPs for *C. beijerinckii* and low frequencies within the *C. saccharobutylicum*.

In summary, the MLST technique was examined for efficacy in strain separation within the *C. beijerinckii* and *C. saccharobutylicum* species. The high bootstrap values obtained at the species level indicate that the five housekeeping genes selected are reliable genetic markers

for use in this technique. A review of the individual genes used in the MLST analysis and the SNP frequencies (Table 2.4) suggests that *atpD* and *gyrB* genes (Appendix A, Figures A1 and A2) contribute to the most resolving power within the *C. saccharobutylicum* strains, as is seen within the final concatenated sequence analysis shown in Figure 2.7. The *atpD* and *oppB* genes (Appendix A, Figures A1 and A3) provide the most SNPs per target gene length and contribute to overall grouping pattern within the *C. beijerinckii* species.

As mentioned previously, a housekeeping gene's ability to provide nucleotide variability to improve differentiating resolution can vary considerably. Whilst housekeeping genes vital to cell function and cell survival are useful candidates, due to ubiquitously conserved genes across genera, they are constrained by the fact that they can only undergo a limited number of nucleotide mutations, without adversely altering encoded protein function. It is interesting to note that at an amino acid level, the SNPs identified specify very few translated amino acid changes (Appendix A, Table A1). This is unsurprising, given that any changes in the amino acid composition of the proteins encoded by these housekeeping genes may alter the protein composition and overall functionality within the cell, resulting in poor cell growth or death. The complete lack of amino acid substitutions within the two NCP species for all the target translated MLST genes, except *C. saccharobutylicum* NCP 206 for the translated *gyrB* gene, gives the first indication that these strains are very closely related at a genetic level. This is especially apparent when one considers that the amino acid substitution level between NCP260 and *C. acetobutylicum* ATCC 824^{T}, in particular in contrast to the rest of the NCP strains, has a general frequency in excess of 20% (with the exception of the *rpoB* gene translated sequence).

A comparison between RAPD results and MLST results revealed that the two NCP *C. saccharobutylicum* strain groups assigned in the RAPD analysis were in fact comprised of three strain groups in the MLST analysis. The NCP *C. beijerinckii* strains consisted of four strain groups in the RAPD analysis, and the four strains chosen from the strain group 2 in the RAPD analysis were largely conserved in the MLST analysis and only NCP260 exhibited differences from the other NCP strains. MLST therefore confirms the grouping patterns observed in the RAPD analysis. The discrepancy between the two types of analyses lies within their nucleotide coverage. RAPD has entire genome coverage and is limited by the primer specificity. MLST is constrained by limited gene coverage, but can detect all differences in that region. A combination of both of these techniques must therefore be implemented to assess levels of strain diversity.

The MLST SNP frequencies are low for the individual genes as well as the assembled pseudogene, which indicates that the NCP strains within both species are very closely related to each other. The overall SNP frequency is too low to assign definitive sub-species strain grouping. The fact that there is some strain grouping, through the presence of only a few SNPs, indicates that these NCP strains are not simply clones of one another. The NCP strains could originate either through individual isolates from a similar geographic location and time, or more likely, a small group of individual parent ancestor strains were isolated and evolved over time. This nucleotide evolution could have been brought about by the selective pressure of many successive batch fermentations or substrate conditioning.

2.5 Conclusions

This study has revealed insights into the possible origins and relationships between the NCP *C. beijerinckii* and *C. saccharobutylicum* strains which were originally classified as *C. acetobutylicum*. Older characterization methods, such as phage typing and physiological tests like rifampicin screening, whilst valid, are very time intensive and have the risk of contamination or errors in the downstream experimental procedures. These restrictions make the species-specific PCR method a faster and lower risk option for screening solvent producing *Clostridium* strains. The PCR-based species-specific screening method should reduce the cost of screening ABE fermentation batch cultures and save time in the ABE fermentation industrial setting. The *etfA* RFLP and 16s *rRNA* analyses have shown that the UCT NCP strain collection is very closely related at a species level.

This is the first time RAPD and MLST analyses have been applied to solventogenic *Clostridium* species. Using RAPD analysis as a starting platform can establish ordered, grouped strain libraries with a better idea of the relatedness of the strains within the library, but has the disadvantage of variability in banding patterns obtained. For accurate strain level analysis, an additional multi-gene MLST approach should be used, which uses nucleotide data and physical SNPs between strains, as a basis of comparison. MLST resolving power in solventogenic *Clostridium* species can be improved by incorporating other genes showing more sequence variation or solvent-specific variable genes into the analysis. Alternately, the PCR products of candidate MLST genes selected in this study can be enlarged to produce a larger overall pseudogene, with the potential for more SNPs in the coverage area for potentially better strain resolution.

The strains within the UCT NCP collection fall into two distinct species, with a small number of nucleotide polymorphisms distinguishing the strain sub groups from each other, within the two species studied. The NCP *C. beijerinckii* strains are genetically distinct from NCIMB 8052^T, with less internal variation. These results are in agreement with the last major study of the NCP collection by Keis *et al.* (2001) where there were no visible differences between NCP strains within the *C. beijerinckii* subgroup 3. Within the *C. saccharobutylicum* species, MLST showed that $NCP262^T$ and NCP258 are closely related. This is in contrast to Keis *et al.* (2001) where all NCP *C. saccharobutylicum* strains were assigned to *C. saccharobutylicum* subgroup 1, except for NCP258, which was assigned to subgroup 2.

Overall, the strains within each species of the UCT NCP collection do differ from one another at a nucleotide level. It appears as though the NCP strains are genetically different, albeit closely related and are not identical strains with different designated strain numbers. However, it is not yet clear whether strains that group together are closely related due to geographic sampling differences or evolution from a common ancestor.

Chapter Three
Solvent Production Capabilities and Nitrogen Growth Studies in the National Chemical Products Strains

---

Contents

3.1 Abstract

The solvent production capabilities of the NCP strain collection were tested on fermentation media containing glucose, sucrose and xylose substrates as the sole carbon sources. ABE solvent analysis of these sugar fermentations revealed a wide diversity of solvent profiles for the NCP strain collection. There are a number of strains within each species which generate high yields of butanol for each of these substrate conditions. Strains NCP195, NCP265, NCP271, NCP J, NCP259 and NCP172 are able to produce high butanol yields under all three fermentation substrates tested.

The buffering capacity of phosphate was compared to that of calcium carbonate in NCP fermentation medium. Calcium carbonate buffered NCP media facilitated the best solvent production. There are many factors that contribute to good fermentation performance. One such factor is the role of the nitrogen source during fermentation. Supplementation with inorganic ammonium salts as the sole nitrogen source retarded growth onset and generated low solvent yields compared to organic nitrogen.

Supplementation of the amino acids glutamate, glutamine, proline, lysine, histidine or asparagine during fermentation was investigated in strains NCP271, NCP260 and NCP258. Glutamine, glutamate and histidine were shown to significantly increase butanol yields ($p < 0.05$, one-way ANOVA Least Statistical Difference analysis).

3.2 Introduction

The production of ABE solvents industrially is closely linked to the substrate used in the fermentation process and industrial strains are developed based on growth and solvent performance for a specific feedstock (Green, 2011). These substrates vary in carbon as well as nitrogen composition. Of chief concern for solvent production is the carbon content, or more specifically, the available sugars that make up the substrate selected for use in a particular fermentation process.

The monosaccharide glucose is the most accessible sugar in the ABE fermentation process due to its ubiquitous distribution amongst most fermentation substrates and its use as the primary source of energy by solventogenic *Clostridium*. The entry of glucose into the cell is controlled by the phosphotransferase system (PTS) whereby glucose is phosphorylated via a sequence of reactions initiated by phosphoenolpyruvate (PEP) in *C. beijerinckii* and *C. acetobutylicum* (Mitchell *et al.*, 1991; Tangney and Mitchell, 2007). This phosphorylated form of glucose can then be used in central cell metabolism, leading to solvent formation.

The disaccharide sucrose is another common fermentable sugar that is utilised in solvent production. Sugar cane molasses which contains sucrose has been used extensively as a substrate in ABE fermentation (Jones and Woods, 1986). Molasses is an inexpensive by-product of the sugar industry and it is considered an attractive substrate since it contains reduced sugars (glucose, sucrose and frcutose) and nitrogen compounds necessary to support fermentation (Najafpour and Shan, 2003). Sucrose transport into the cell is controlled by the sucrose PEP-dependent PTS in *C. beijerinckii* and *C. acetobutylicum* (Tangney *et al.*, 1998;

Tangney and Mitchell, 2000). The sucrose-6-phosphate is then hydrolysed by sucrose-6-phosphate hydrolases resulting in the formation of glucose-6-phosphate and fructose which are used in the glycoltic pathway (Reid *et al.*, 1999).

Lignocellulosic plant material that would otherwise be considered as waste material is an attractive fermentation substrate since it is inedible by humans and has a low cost (Jang *et al.*, 2012). Lignocellulosic biomass consists of three major components of variable proportions: cellulose, hemicellulose, and lignin which are often separated by the addition of acid or hydrolytic enzymes (Noguchi *et al.*, 2013). Hemicellulose is comprised of pentose sugars like xylose and mannose as well as hexose sugars such as glucose and galactose (Sun and Cheng, 2002). Many enzymatic steps are required to convert D-xylose to glyceraldehyde-3-P, which is a product which can feed into glycolysis (Wamelink *et al.*, 2008). The ability of solventogenic clostridia to utilise xylose as an energy source means that these strains are able to ferment treated lignocellulosic material as a substrate. This presents interesting alternative low cost substrates such as domestic organic waste that contains xylose (Claassen *et al.*, 2000). This waste substrate performs better when the lignocellulose is exposed to cellulase pre-treatment (López-Contreras *et al.*, 2000).

The other aspect of fermentation substrates often overlooked is the nitrogen composition of the fermentation media. Different nitrogen sources and overall nitrogen metabolism affect solvent production in saccharolytic *Clostridium* strains (Long *et al.*, 1984). Strains like *C. beijerinckii* NCIMB 8052^{T} cannot produce solvents under low ammonia conditions (Gottschal and Morris, 1981). Growth of solventogenic clostridia is severely hampered if the organic nitrogen source is limited (Stutz *et al.*, 2007). Recently, there is evidence that

indicates that substrates supplemented with both organic nitrogen in the form of yeast extract (YE) and inorganic nitrogen in the form of ammonium nitrate show increased solvent yields when compared to substrates supplemented with each of these nitrogen sources alone (Abd-Alla and Elsadek El-Enany, 2012).

Investigation into the amino acids comprising the organic nitrogen component of fermentation substrates and the role of specific amino acids in solventogenesis is limited. There is evidence to suggest some amino acids have a stimulatory effect on butanol production either directly through the properties of the particular amino acid or indirectly through the production or activation of alternative proteins that act as butanol metabolism enhancers. The next section outlines evidence in the literature for the biosynthesis of certain amino acids that were shown to be produced during solvent formation.

Transcriptomic and proteomic analysis of solventogenic clostridia during solvent production indicates that there is a clear difference in nitrogen metabolism between acidogenic and solventogenic states (Janssen *et al.*, 2010; Heluane *et al.*, 2011; Janssen *et al.*, 2012; Yang *et al.*, 2012). A number of amino acids including, histidine, lysine, proline, glutamine and glutamate were found to be synthesized or transported into the cell during solvent stresses. Transcriptomic analysis using microarray data tracked the regulation of genes in *C. beijerinckii* cultures that were exposed to a butanol pulse. Genes involved in histidine biosynthesis and lysine transport were significantly up-regulated during butanol stress (Table 3.1). Upon evaluation of the transcriptomic results obtained, the authors decided to incorporate glutamine, asparagine and histidine into the laboratory minimal medium as the

organic nitrogen sources in order to promote better growth and solvent production (Heluane *et al.*, 2011).

Table 3.1: A summary of up-regulated genes involved in amino acid transport and biosynthesis in solventogenic clostridia in response to different solvent stresses.

Amino Acid Involved	Gene product	Bacterium	Treatment	Reference
His	Histidinol dehydrogenase	*C. beijerinckii*	Butanol stress	(Heluane *et al.*, 2011)
His	Phosphoribosylformimino-5-aminoimidazole carboxamide ribonucleotide (ProFAR) isomerase	*C. beijerinckii*	Butanol stress	(Heluane *et al.*, 2011)
His	Phosphoribosyl-AMP cyclohydrolase	*C. beijerinckii*	Butanol stress	(Heluane *et al.*, 2011)
His	ATP phosphoribosyltransferase catalytic subunit	*C. beijerinckii*	Butanol stress	(Heluane *et al.*, 2011)
Lys	Lysine-specific permease	*C. beijerinckii*	Butanol stress	(Heluane *et al.*, 2011)
Lys	Dihydroxy-acid dehydratase	*C. beijerinckii*	Butanol stress	(Heluane *et al.*, 2011)
His	ProFar isomerase	*C. acetobutylicum*	Acidogenesis→solventogenesis	(Janssen *et al.*, 2010)
His	ATP phosphoribosyltransferase catalytic subunit	*C. acetobutylicum*	Acidogenesis→solventogenesis	(Janssen *et al.*, 2010)
His	Histidinol dehydrogenase	*C. acetobutylicum*	Acidogenesis→solventogenesis	(Janssen *et al.*, 2010)
Gln	*glnA* Glutamine synthetase type III	*C. acetobutylicum*	Butanol stress	(Janssen *et al.*, 2012)
Gln	*glnQ* glutamine ABC transporter	*C. acetobutylicum*	Butanol stress	(Janssen *et al.*, 2012)
Glu	*gltA* large subunit of NADH-dependent glutamate synthase	*C. acetobutylicum*	Acidogenesis→solventogenesis/ Butanol stress	(Janssen *et al.*, 2010, 2012)
Glu	*gltB* small subunit of NADH-dependent glutamate synthase	*C. acetobutylicum*	Butanol stress	(Janssen *et al.*, 2012)
Pro	Proline/glycine ABC Transporter	*C. acetobutylicum*	Butanol stress	(Janssen *et al.*, 2012)
Gln	Glutamine synthetase, catalytic region	*C. thermocellum*	Ethanol stress	(Yang *et al.*, 2012)
Gln	Glutamine	*C. thermocellum*	Ethanol stress	(Yang *et al.*,

	amidotransferase, class-II			2012)
Glu	Glutamate synthase (NADPH)	*C. thermocellum*	Ethanol stress	(Yang *et al.*, 2012)
Glu	Glutamate dehydrogenase	*C. thermocellum*	Ethanol stress	(Yang *et al.*, 2012)
His	Histidinol transferase	*C. thermocellum*	Ethanol stress	(Yang *et al.*, 2012)
His	Histidinol-phosphate aminotransferase	*C. thermocellum*	Ethanol stress	(Yang *et al.*, 2012)

A similar microarray approach was conducted by Janssen and colleagues using *C. acetobutylicum*. In this case, the expression of genes in acidogenesis compared to solventogenesis was monitored. It was shown that the genes involved in the biosynthesis of histidine and glutamate were up-regulated in the transition between acidogenesis and solventogenesis (Janssen *et al.*, 2010). Another series of experiments conducted by the same group compared acidogenic *C. acetobutylicum* cells maintained by chemostat to cells in the same stage exposed to a butanol pulse. The genes involved in the synthesis of glutamine and glutamate were upregulated in response to the butanol pulse. Interestingly, the genes for glutamine and proline transport were upregulated under these conditions as well (Janssen *et al.*, 2012) (Table 3.1).

Another study conducted with *Clostridium thermocellum* revealed that the genes involved in glutamine and glutamate biosynthesis as well as genes involved in the conversion of histidinol to glutamate (histidinol transferase) were up-regulated in response to an ethanol pulse compared to untreated cells (Yang *et al.*, 2012). Despite the fact that *Clostridium thermocellum* is not closely related to ABE-forming *Clostridium* species, the genes required for glutamine and glutamate biosynthesis recruited in response to solvent production are conserved across species. The exact effect these amino acids have on the alleviation of solvent stress and production remains unknown. These amino acids are either incorporated

into proteins linked to solvent stress or are themselves used in biochemical pathways involved in solvent stress. To date, no amino acid characterization studies have been performed that link any of these amino acids mentioned to increased solvent production.

In this study, the fermentation profiles for the entire NCP strain collection are determined for glucose-, sucrose- and xylose-based substrates. This gives an indication of the specificity and preference of certain strains to specific substrates. Strains that produce high yields of solvents under these conditions can be noted as strains with potential in current industrial fermentation processes. It is hypothesised that good nitrogen conditions underpin good growth and therefore solvent production. The impact of inorganic and organic nitrogen sources on growth and solvent production by these solventogenic clostridia is determined. Specifically, the supplementation of specific amino acids into fermentation media and the role this has on growth and solvent yields of representative NCP strains is investigated. Amino acids found to improve solvent yields can be used in fermentation media either as a direct supplement or by use of substrates rich in that particular amino acid.

3.3 Materials and Methods

3.3.1 Bacterial strains and media

All strains used in growth curve or fermentation studies (Table 3.2) were grown in 2% glucose reinforced clostridial medium (gRCM) (Oxoid) in an anaerobic chamber (model 1024 Forma Scientific) with a gas phase of 5% H_2, 10% CO_2 and 85% N_2 at 37 °C for 18 h. For nitrogen growth studies, cultures (18 h) were transferred to a 100 mL Schott bottle containing modified glucose-mineral salts-biotin minimal medium (GSMM) (2% glucose, 0.0008% $CaCl_2$, 0.004% K_2HPO_4, 0.004% KH_2PO_4, 0.008% NaCl, 0.02% $MgSO_4 \cdot 7H_2O$, 0.001% $MnSO_4 \cdot H_2O$, 0.001% $FeSO_4 \cdot 7H_2O$, 0.005% $ZnSO_4 \cdot 7H_2O$, 0.001% $Na_2MoO_4 \cdot 2H_2O$, 0.1% $NaHCO_3$, 0.0001% para-aminobenzoic acid (Merck), 0.05% cysteine, 0.0025% biotin, 0.0001% thiamine-HCl (Sigma), 0.000005% pyroxidine (DIFCO Labs), 0.000005% Ca-pantothenate (Merck), 0.000005% nicotinamide, 0.000005% riboflavin , 0.000000625% folic acid and 0.000000025% vitamin B_{12} (Sigma)) (Holdeman *et al.*, 1977). GSMM media was supplemented with high (0.2%) or low (0.0025%) organic nitrogen, in the form of Casamino Acids (DIFCO Labs), or 0.2% inorganic nitrogen in the form of either NH_4-acetate, $(NH_4)_2SO4$, NH_4Cl or NH_4NO_3 (Merck). Strains were diluted to a starting OD_{600} value of 0.1 in a total volume of 100 mL and grown anaerobically at 37 °C (Table 3.2). OD_{600} readings were measured every hour spectrophotometrically (Novaspec).

Table 3.2: A list of the strains used in this study.

Strain Name				
***C. beijerinckii Strains*:**			***C. saccharobutylicum* Strains:**	
NCIMB 8052^T	NCP202	NCP263	NCP108	NCP249
NCP J	NCP220	NCP270	NCP162	NCP258
NCP106	NCP254	NCP271	NCP195	$NCP262^T$
NCP172	NCP259	NCP272	NCP199	NCP265
NCP193	NCP260	NCP280	NCP200α206	NCP268
NCP200	NCP261	NRRL B593	NCP206	

In the case of the fermentation studies, a 10% inoculum of overnight cultures was transferred to fresh gRCM and incubated anaerobically at 37 °C until an OD_{600} of 0.4 was reached as a starter culture. A further 10% inoculum of this starter culture was then inoculated into a universal container containing NCP medium (5% glucose/sucrose/xylose (Sigma), 1% corn steep liquor (African Products (Pty) Ltd.), 0.4% yeast extract (DIFCO Labs), 0.5% $CaCO_3$ (Merck) and 0.2% $(NH_4)_2SO_4$) (Merck), pH6.5) to a total fermentation volume of 10 mL. All batch fermentations were carried out in triplicate and incubated anaerobically at 37 °C for 72 h. Amino acid fermentation trials and were carried out as outlined above, but the NCP medium was modified to incorporate only 0.1% yeast extract (YE) and the remaining organic nitrogen component was supplemented with 0.1-0.4% of either mono-sodium glutamate (Merck), proline, lysine, asparagine, histidine, or glutamine (Sigma) to a total fermentation volume of 100 mL. In the case of amino acid growth curves, amino acid fermentation medium was used, but the medium's $CaCO_3$ buffer was replaced with a phosphate buffer consisting of 1% K_2HPO_4 and 1% KH_2PO_4 (Merck) so that OD_{600} growth readings could be measured. All NCP media are assumed to be $CaCO_3$ buffered unless otherwise stated.

3.3.2 Solvent harvest and sample preparation

After 72 h fermentation, 2 mL fermentation samples were collected from triplicate fermentation vessels, centrifuged at 16000 *g* for 20 min and the supernatant harvested. The harvested solvents were kept on ice or frozen at -80 °C for up to a month prior to solvent analysis. Harvested solvent samples were mixed in glass GC vials (Chromacol, Thermo Fisher Scientific Inc.) 1:1 with an isobutanol (2-butanol) internal standard which was prepared as a w/v 10 g/L solution to a final volume of 1 mL. Standards for standard curve generation for acetone and butanol respectively were prepared w/v as 1, 5, 10, 15 and 20 g/L to a final volume of 1 mL with HPLC grade water. The same was performed for ethanol standards, but within the 0.1, 0.5, 1.0, 1.5 and 2.0 g/L range. Calibration standards were created by mixing acetone, butanol and ethanol for each concentration 1:1 with 10 g/L isobutanol internal standard to a total volume of 1 mL in glass GC vials.

3.3.3 Gas Chromatography-Mass Spectrophotometry (GC-MS) conditions

Calibrations for acetone, butanol and ethanol as well as fermentation samples quantifications were performed using the Agilent 7890A GC and QQQ 7000A MS fitted with an auto-sampler (Agilent Technologies Inc.). All samples were introduced into the polar polyethylene glycol Innowax column (19091N-233, Agilent). The GC program was operated with a simultaneous ramp temperature and ramp flow. The ramp oven temperature was initially set at 80 ℃ and held for 0.5 min. The temperature was then raised to 200 ℃ at rate of 10 ℃/min and held for 7.5 min with a total running temperature time of 20 min. The helium gas flow rate was set to 0.8 mL/min and was held for 8 min. The flow was then increased to 1.0 mL/min with an increasing rate of 1.0 mL/min^2 and held for 5 min. The flow rate was then

increased to 1.3 mL/min with the same increasing rate of 1.0 mL/min^2 with a total concurrent flow running time of 20 min.

Prior to sample injection, the 10 µL auto-sampler syringe was rinsed five times with HPLC grade 100% methanol (Sigma). Thereafter, a 0.1 µL sample was used as the injection volume into the column inlet. The syringe was then rinsed five times with a 100% methanol solution. The sample inlet temperature was set to 250 ℃ , the helium carrier gas pressure to 7.22 psi with a vent gas flow of 43 ml/min and a split ratio of 50:1.The amplitude and retention time of the isobutanol internal standard peaks in GC-MS runs was assumed to be constant for all standard curve and fermentation sample runs. Constant isobutanol peak amplitudes act as a reference control to determine whether the GC-MS sample run is behaving correctly. Only acetone, butanol and ethanol standard curves with a linear r^2 value of above 0.96 were used for quantitation of solvents in fermentation samples (Appendix B, Figure B1). ABE solvents were identified and quantified using Mass Hunter Workstation software B.03.01 (Agilent) (Appendix B, Figure B2).

3.3.4 Statistical analysis of solvents

Statistical significance of solvent concentration was determined using the one way ANOVA LSD test incorporating 12 degrees of freedom and an alpha threshold of 0.05 to the homogenous group using the software Statistica (StatSoft, Inc. (2012) STATISTICA (data analysis software system, version 11. www.statsoft.com).

3.4 Results and Discussion

3.4.1 Fermentation performances for common sugar substrates

The basic ABE fermentation capabilities of the UCT NCP strain collection that was characterised in Chapter Two is unknown. To this end, it was necessary to screen all the NCP strains for ABE solvent production under basic fermentation media conditions. Three sugar substrates were chosen as the carbon sources for these batch fermentations, namely, glucose, xylose and sucrose.

3.4.1.1 Fermentation performance for glucose-based medium

The monosaccharide glucose was selected, because this hexose sugar is ubiquitous in nature and as such, the pathways involved in general growth as well as solvent production rely heavily on glucose to function. Since glucose is a monomeric molecule, it does not require any enzymatic cleavage and it is easily transported into and around the cell (Jahreis *et al.*, 2008).

Of primary interest in this study are the butanol titres, however, acetone is also a valued solvent. Ethanol is also produced, but the concentrations are so small, that for the purposes of this study, ethanol is not a focus. The ABE solvent titres for the NCP *C. beijerinckii* strains that utilised glucose substrate as the primary carbon source are shown below (Figure 3.1). A total of eight of the NCP *C. beijerinckii* strains produced in excess of 10 g/L butanol, seven strains (including NCIMB 8052^{T}) produced between 8-10 g/L butanol and three strains produced butanol in concentrations of less than 8 g/L over a 72 h period. It should be noted

that strain NRRL B593 showed normal growth in the RCM starter culture, but did not adapt well to the fermentation conditions.

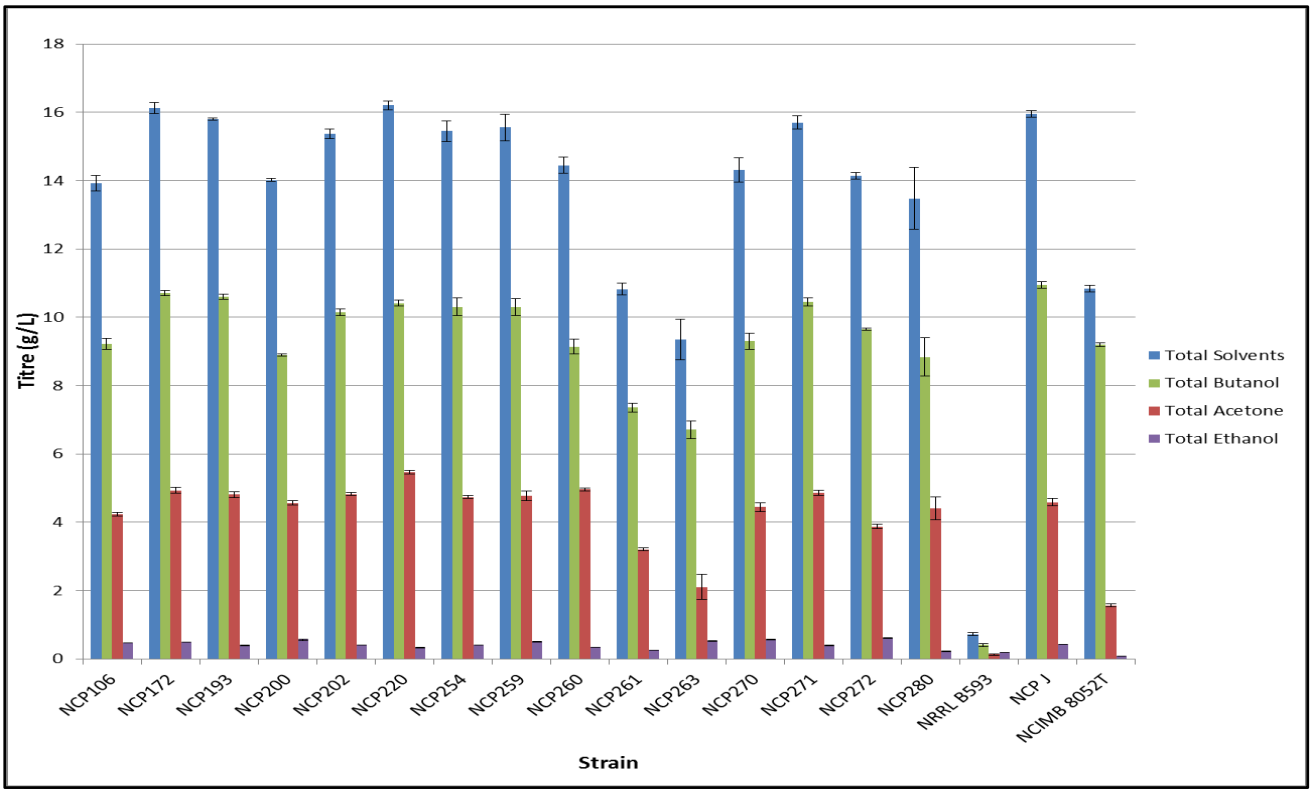


Figure 3.1: ABE solvent titres after 72 h for NCP *C. beijerinckii* strains utilising 5% glucose as the primary carbon source. Error bars represent the standard error of the mean, where n=3.

The NCP *C. beijerinckii* solvent titres under glucose conditions were arranged according to best butanol production titres and yields (Appendix B, Table B1). Total solvent yield in this particular case was expressed as the amount (in grams) of ABE solvents produced per 50g of sugar used as a fermentation substrate. In other words, the solvent yield is a ratio expressed as grams solvent/50g sugar added. The same is true for butanol yield. In the case of NCP fermentation medium, strains NCP J and NCP172 had the highest butanol yields of 0.219 and 0.214 respectively. This equates to 21.9% and 21.4% of the sugar added to the fermentation system converted to butanol. Of particular interest is the ratio of butanol: acetone. In classic ABE fermentation, the ratio of butanol: acetone should equal two. Strains that exhibit high

butanol yields as well as a ratio of butanol: acetone great than 2.0 are considered desirable from a biobutanol biofuel standpoint, since the organism intrinsically favours the production of butanol over acetone, resulting in the best butanol yields. The top eight NCP *C. beijerinckii* strains, except NCP 220, all had butanol: acetone ratio of greater than 2.0 (Appendix B Table B1). NCIMB 8052^T exhibited a high butanol: acetone ratio of 5.88, favouring butanol formation. Strains NCP J and NCP172 had butanol: acetone ratios of 2.39 and 2.17 respectively along with total solvent yields of 0.319 and 0.323 respectively which indicates that these strains intrinsically produce and tolerate high concentrations of solvents under glucose conditions.

The solvent yields and titres for the NCP *C. saccharobutylicum* strains are shown below (Figure 3.2). A total of six of the *C. saccharobutylicum* strains produced more than 10 g/L of butanol, four strains produced 8-10 g/L of butanol and one strain produced less than 8 g/L butanol over the course of the fermentation. NCP162 did not respond well to the NCP fermentation medium, despite normal gRCM medium starter culture readings. Of interest is strain NCP249 which presents the best butanol yield of 0.218 and has an unusually high butanol: acetone ratio of 4.62 and a total solvent yield of 0.272 (Appendix B Table B2). However, a strain like NCP195 has a butanol yield of 0.214 with a lower butanol: acetone ratio of 1.83, indicating more acetone is produced represented by a larger total solvent yield of 0.309. Depending on the final desired solvent output, either strain could be used, however, NCP249 could be deemed more desirable in the context of biobutanol formation and up-scaling butanol production since a greater proportion of the cells' energy is directed towards butanol formation.

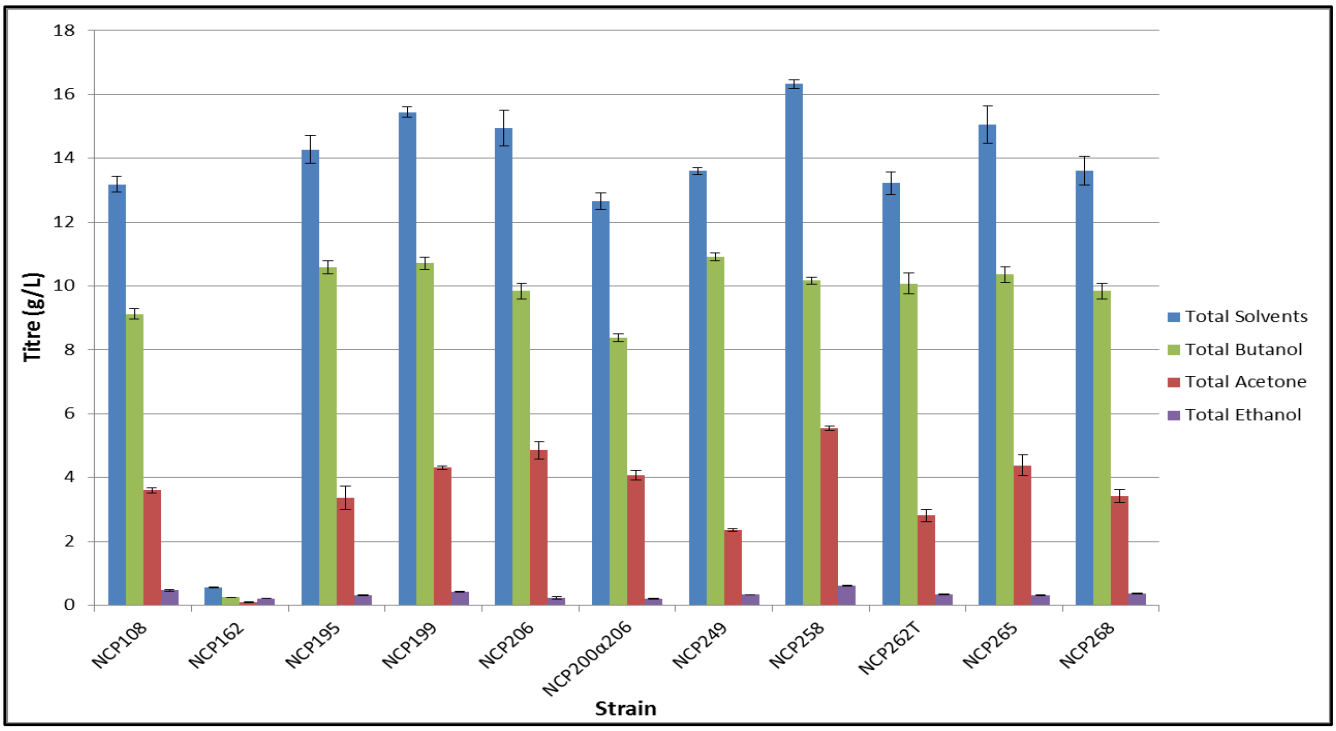


Figure 3.2: ABE solvent titres after 72 h for NCP *C. saccharobutylicum* strains utilising 5% glucose as the primary carbon source. Error bars represent the standard error of the mean, where n=3.

In a recent study utilizing NCP260, a total solvent yield of 0.41 was recorded using glucose as a substrate (Qureshi *et al.*, 2010b). A comparative study using the major solventogenic industrial strains reported solvent yields on glucose of 0.282 for *C. saccharobutylicum* 262^{T} (Ezeji and Blaschek, 2008).Overall, the butanol performances between NCP *C. saccharobutylicum* and *C. beijerinckii* strains under glucose conditions are similar. However, the range of the ratios of butanol: acetone in the *C. saccharobutylicum* strains (Appendix B, Table B2) is more varied than that of the *C. beijerinckii* strains (Appendix B, Table B1). The total solvent yield data between the species utilizing glucose as the sugar source is in agreement with the last comparison of the NCP collection to other major solventogenic species that reported a total solvent yield range of 0.2-0.3 for *C. saccharobutylicum* and 0.25-30 for *C. beijerinckii* strains (Shaheen *et al.*, 2000). The observation that *C. saccharobutylicum* strain solvent yields are more variable is supported by the aforementioned study as well as this study.

3.4.1.2 Fermentation performance for sucrose-based medium

The disaccharide sucrose was selected as a substrate based on its significant historical use as the preferred fermentation substrate in the form of sugar cane molasses during the latter part of the NCP facility's operation in South Africa. Sucrose is comprised of the monosaccharides, glucose and fructose. As such, this sugar is versatile since the sugar can be transported as either a disaccharide or hydrolysed monosaccharides; however, sucrose metabolism means that the organism must produce its own sucrases to liberate these monomers for metabolic consumption. These monosaccharides can be metabolically utilised in either fructose- or glucose-related metabolic systems.

The solvent titres for *C. beijerinckii* strains under sucrose fermentation conditions are depicted below (Figure 3.3). Four strains (including NCIMB 8052^T) produced more than 10 g/L of butanol, seven strains produced between 8-10 g/L of butanol and the remaining seven strains produced less than 8 g/L of butanol over a 72 h period. The butanol: acetone ratios for the *C. beijerinckii* strains were generally higher than 2.2. This indicated that a higher proportion of butanol than acetone is formed (Appendix B, Table B3). NCP254 performed the best out of the *C. beijerinckii* strains with a butanol yield of 0.216, a butanol: acetone ratio of 2.38 and a solvent yield of 0.314. NCIMB 8052^T and NCP172 had higher butanol yields of 0.208 and 0.200, butanol: acetone ratios of 3.13 and 3.10 and solvent yields of 0.28 and 0.272 respectively, under these conditions compared to the other strains in the species in this study. NCP263 and NRRL B593 did not produce solvents under sucrose-based fermentation conditions.

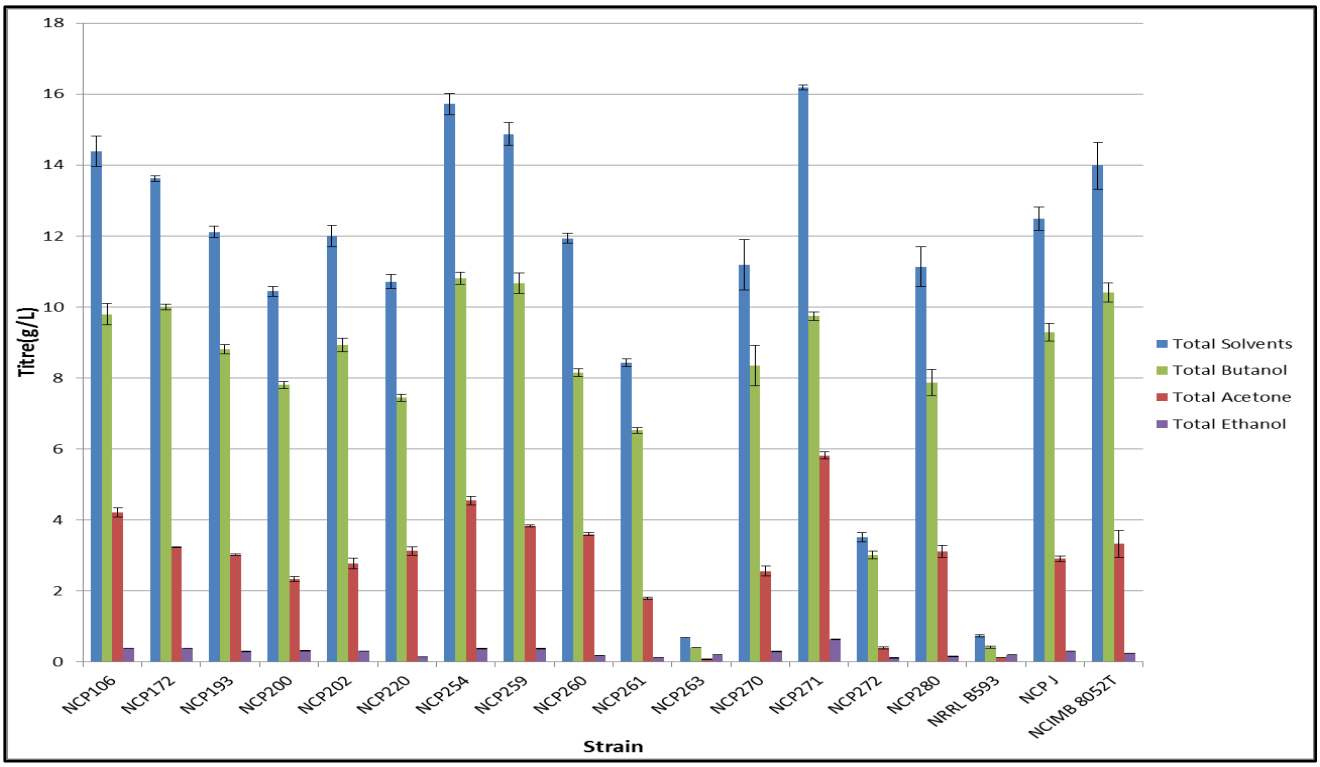


Figure 3.3: ABE solvent titres after 72 h for NCP *C. beijerinckii* strains utilising 5% sucrose as the primary carbon source. Error bars represent the standard error of the mean, where n=3.

The solvent titres for *C. saccharobutylicum* strains under sucrose conditions are illustrated below (Figure 3.4). Four strains achieved butanol titres of more than 10 g/L, three strains between 8-10 g/L and four strains less than 8 g/L. NCP162 did not produce solvents in the sucrose-based NCP fermentation medium. The butanol: acetone ratio for these strains was generally in excess of 2.2 for sucrose substrate (Appendix B, Table B4). NCP268 had the highest butanol yield of 0.227, a butanol: acetone ratio of 2.5 and an overall solvent yield of 0.327. NCP268 displays a high level of intrinsic solvent producing capabilities and importantly, can tolerate higher amounts of total solvent concentrations under sucrose conditions compared to the other strains examined in this study.

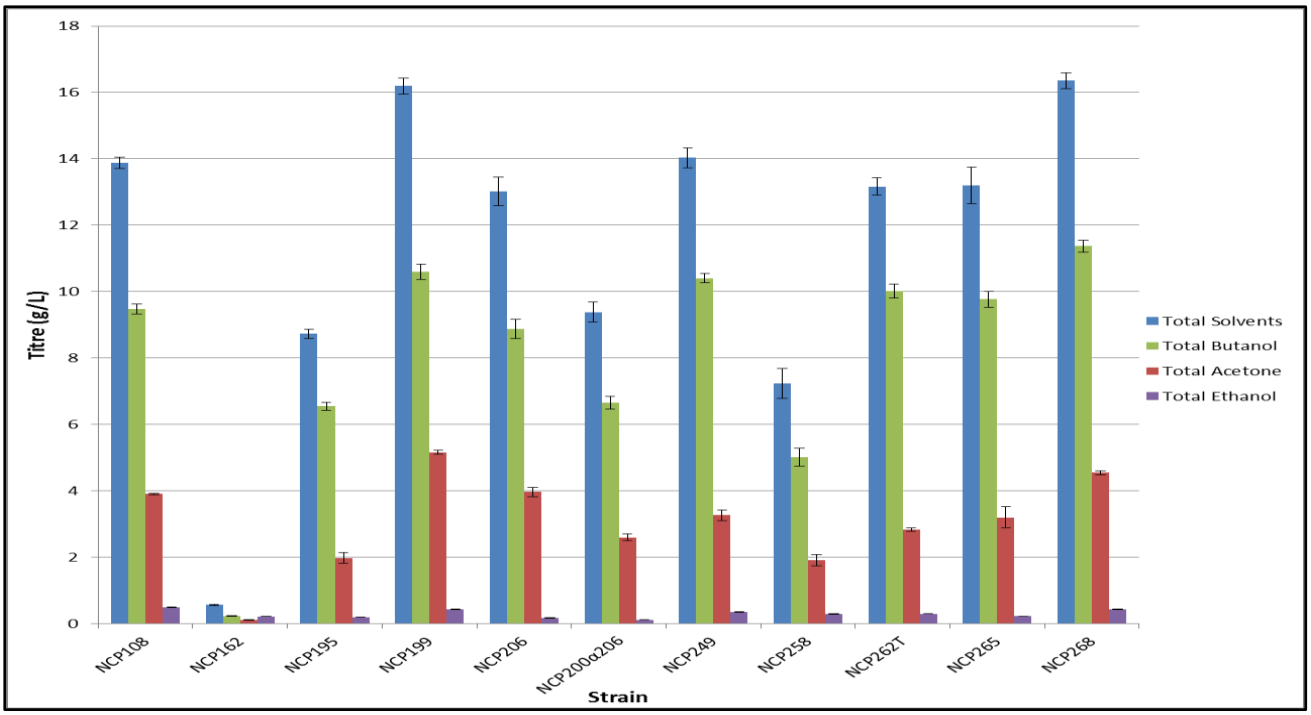


Figure 3.4: ABE solvent titres after 72 h for NCP *C. saccharobutylicum* trains utilising 5 % sucrose as the primary carbon source. Error bars represent the standard error of the mean, where n=3.

Overall, the NCP strains produced lower levels of the different solvents under sucrose conditions compared to glucose conditions. This is surprising when one considers that the NCP strains were historically utilised in molasses-based fermentations. However, it seems likely that only a small subset of these strains were actually used for sugar cane fermentations. In a study comparing solvent performance of a selection of NCP strains grown on 6% molasses media, solvent yields of 0.298 and 0.305 were recorded for *C. saccharobutylicum* strains NCP262^{T} and NCP258 respectively. Solvent yields of 0.315 and 0.183 were recorded for *C. beijerinckii* NCP260 and NCIMB 8052^{T} respectively (Shaheen *et al.*, 2000). Butanol yields of 0.27 were observed for NCIMB 8052^{T} utilizing tropical maize stalk juice as the sugar substrate (Wang and Blaschek, 2011). A recent study conducted using NCP262^{T} in a batch fermentation system based on molasses media reported a solvent yield of 0.33 (Ni *et al.*, 2012). The same group reproduced the conditions under continuous

fermentation and reported a total solvent titre of 13.75 g/L with a productivity of 0.439 g/L/h over a period of 220 h (Ni *et al.*, 2013).

It is interesting to note that under sucrose conditions, there is a general departure from the classic 6:3:1 butanol: acetone: ethanol ratio. The ratio of butanol: acetone in particular exceeds 2.2 under sucrose conditions. Utilizing sucrose as a carbon source for cell metabolism has benefits since both glucose and fructose metabolic pathways facilitate energy production and the subsequent generation of solvents. However, this biochemical pathway versatility comes at the price of enzymatic hydrolysis of sucrose and the requirement for ATP to phosphorylate glucose and fructose (Reid *et al.*, 1999). The different strains may have varying native sucrase activity levels and it is possible that a portion of the sucrose is never hydrolysed to liberate fructose and glucose for metabolism.

3.4.1.3 Fermentation performance for xylose-based medium

The pentose sugar xylose was selected as a substrate, because of its abundance in nature in the form of plant matter. Since xylose is the principal component of hemicellulose, it is an attractive sugar substrate for fermentation, because a large proportion of all plant biomass is comprised of hemicellulose and therefore xylose. If fermenting bacteria can utilise the xylose component of hemicellulose, then most forms of cellulosic plant material become accessible as fermentation substrates.

The solvent titres for *C. beijerinckii* strains under xylose-based conditions are illustrated below (Figure 3.5). A number of the *C. beijerinckii* strains did not grow well or produce

solvents under xylose-based fermentation conditions. Four strains achieved butanol titres between 8-10 g/L and 14 strains (including NCIMB 8052^{T}) produced less than 8 g/L of butanol over a 72 h period. The butanol: acetone ratios varied; however, the *C. beijerinckii* strains generally had butanol: acetone ratios of generally more than 2.36 and up to 19.30 (Appendix B, Table B5). NCP259 had the highest butanol yield of 0.195, a butanol: acetone ratio of 2.66 and total solvent yield of 0.277.

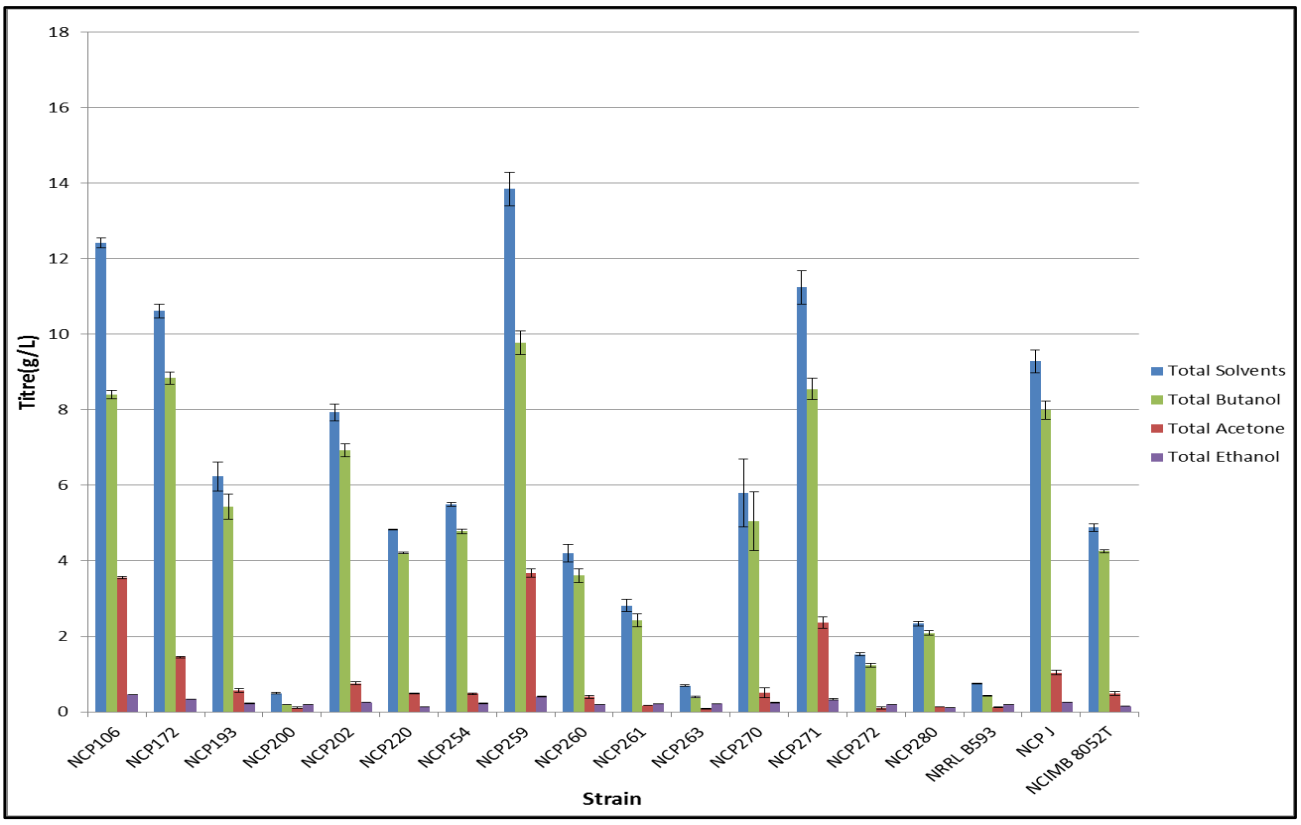


Figure 3.5: ABE solvent titres after 72 h for NCP *C. beijerinckii* strains utilising 5% xylose as the primary carbon source. Error bars represent the standard error of the mean, where n=3.

The solvent titres under xylose-based fermentation conditions for *C. saccharobutylicum* strains are depicted below (Figure 3.6). One strain produced in excess of 10 g/L of butanol, two between 8-10 g/L and the remaining eight strains produced less than 8 g/L of butanol over 72 h. NCP195 had the highest butanol yield of 0.214, a butanol: acetone ratio of 2.54

and a total solvent yield of 0.306 (Appendix B, Table B6). This is the most promising candidate for xylose based fermentations since it can produce and tolerate higher levels of solvents. It is possible that NCP195 could degrade cellulosic material like hemicellulose since there is evidence that *C. saccharobutylicum* can produce solvents in hemicellulose-based media (Berezina *et al.*, 2009). Total solvent yields of 0.28 have been reported in a *glcG* deficient mutant of *C. acetobutylicum*. This did not prevent D-glucose from being metabolized and D-xylose metabolism was favoured by the co-overexpression of a D-xylose proton-symporter (cac1345), D-xylose isomerase (cac2610), and xylulokinase (cac2612 in the mixed glucose-xylose medium (Xiao *et al.*, 2011). *C. saccharoperbutylacetonicum* displayed a solvent yield of 0.311 on media containing cellobiose and xylose with 35 g/L glucose and 39 g/L xylose respectively (Noguchi *et al.*, 2013).

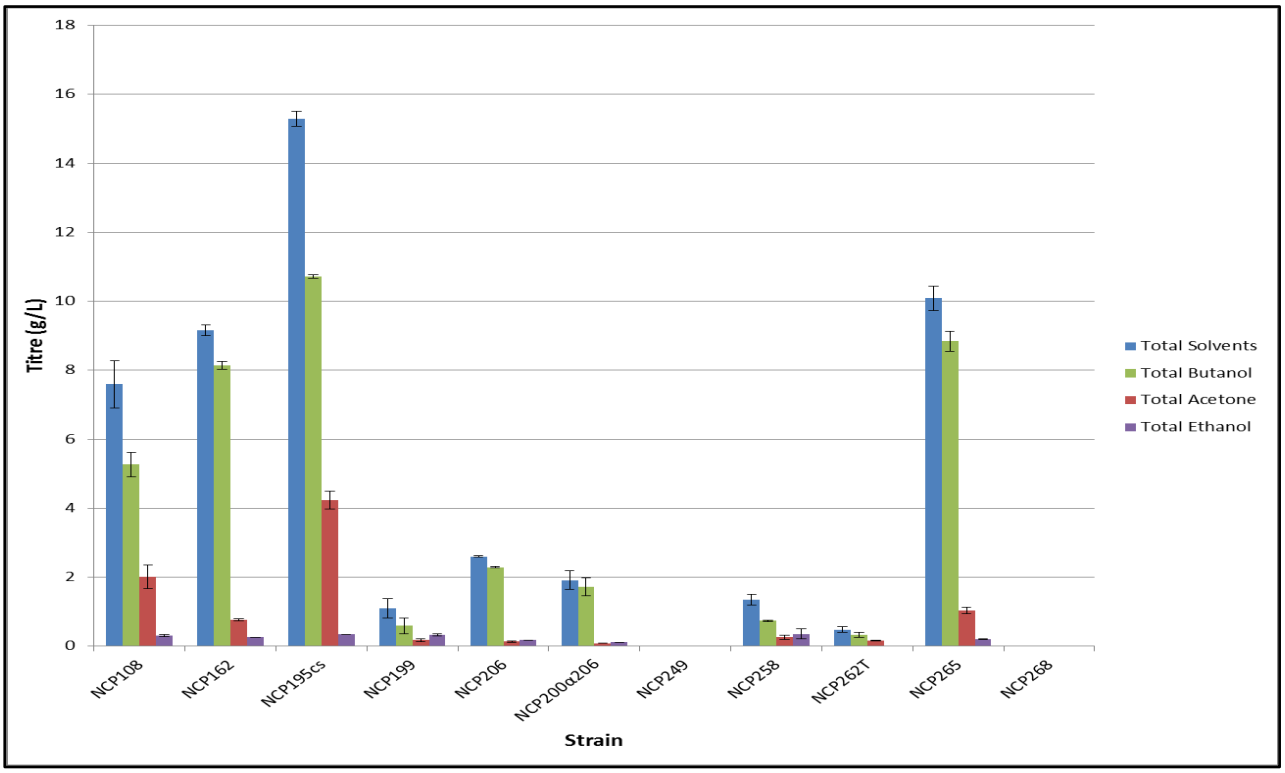


Figure 3.6: ABE solvent titres after 72 h for NCP *C. saccharobutylicum strains* utilising 5% xylose as the primary carbon source. Error bars represent the standard error of the mean, where n=3.

There are a variety of solvent profiles and performances under glucose, sucrose and xylose fermentation conditions shown by different NCP strains within the same species. Solvent production variation arising from different starting cell biomass was ruled out due to the fact that starter cultures were all grown up to OD_{600} of 0.4 and a 10% starter culture inoculum introduced into the fermentation vessels all set to the same final volume. The different solvent performances for the same 72 h fermentation period seem indicative of the intrinsic strain performance capabilities of the strains; however, further optimization of a particular strain may require individual growth conditions. This diversity in the phenotypes of solvent yields and ratios is further evidence that the NCP strain collections differ at a genetic level.

Unfortunately, there does not appear to be any correlation between the strain-level groups found in the RAPD or MLST analysis from Chapter Two, and the solvent profiles presented in this chapter. This could be because the gene segments used for the strain level comparisons were not chosen because they were directly involved in solvent production. The variety in solvent profiles for the different strains warrants some form of solvent-related gene analysis as the basis of any MLST work performed on solventogenic *Clostridium.*

A number of strains have butanol and total solvent yields that exceed 0.2 and 0.3 respectively for the different sugar substrates tested (Appendix B, Tables B1-6). Different NCP strains could be tailored to certain types of feedstock depending on what sugar substrate is used by a particular fermentation plant. A number of strains, namely, NCP195, NCP265, NCP271, NCP J, NCP259 and NCP172 show consistent solvent production for all substrates and can be considered to be robust strains with respect to ability to use different substrates to produce

solvents. These strains could be incorporated into fermentation facilities where the substrate availability varies or there is a seasonal supply of different crops or substrates made available for use in fermentations.

3.4.2 Investigation into the role of nitrogen in growth

The effects of nitrogen components in fermentation media are insufficiently characterised. While atomic nitrogen is not incorporated into solvents, it is very much a part of all amino acids which form the proteins and enzymes that facilitate growth of the organism and ultimately solvent production.

3.4.2.1 The effects of inorganic nitrogen on growth

The impact of inorganic nitrogen on the growth of the NCP strains was assessed. Inorganic nitrogen in the form of ammonium salts was compared to organic nitrogen in the form of casamino acids as done previously (Stutz *et al.*, 2007). Growth conditions were compared between media supplemented with casamino acids and ammonium acetate for *C. beijerinckii* (Figure 3.7) and *C. saccharobutylicum* (Figure 3.8) strains.

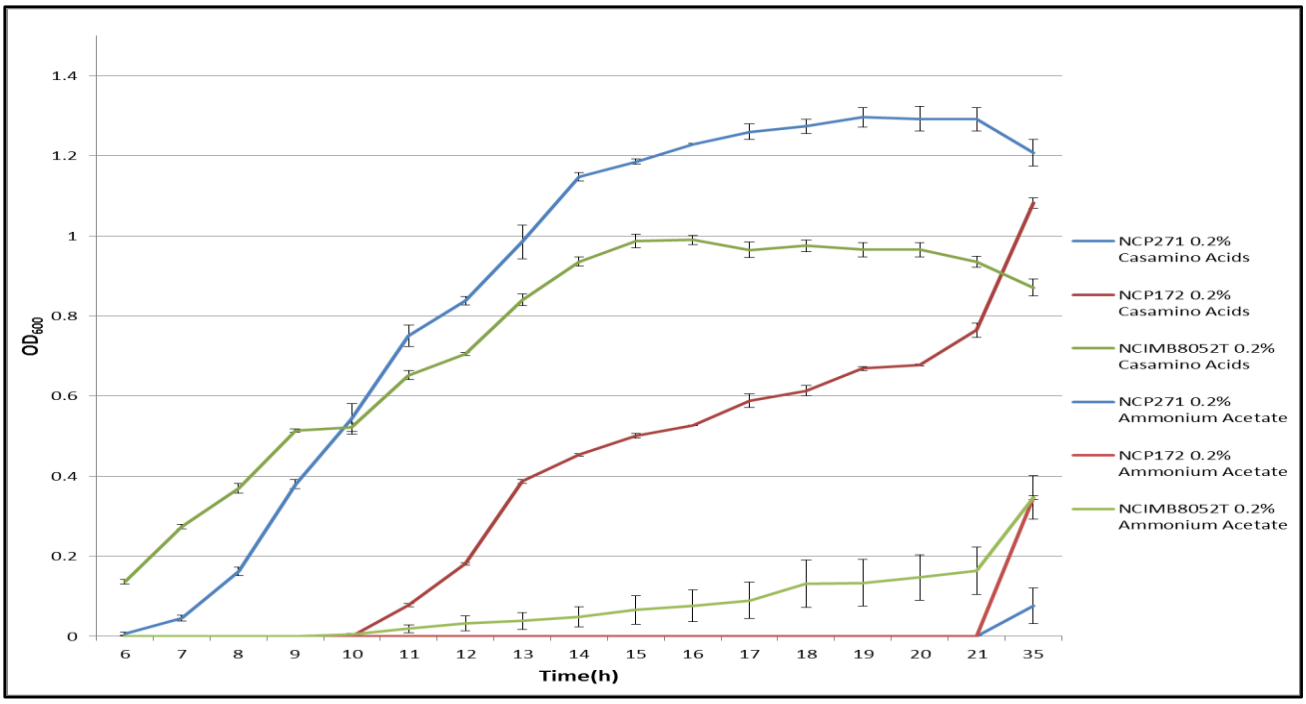


Figure 3.7: Growth of NCP *C. beijerinckii* strains in minimal media supplemented with 0.2% casamino acids compared to 0.2% ammonium acetate as the inorganic nitrogen source. Error bars represent the standard error of the mean, where n=3.

The *C. beijerinckii* strains did not respond well to ammonium acetate as the primary nitrogen source in the media. Growth rates were severely reduced and maximum growth levels were retarded compared to media supplemented with casamino acids (Figure 3.7). The *C. saccharobutylicum* strains by contrast, reacted better to ammonium acetate as the sole nitrogen source in the medium (Figure 3.8). All the *C. saccharobutylicum* strains under ammonium acetate nitrogen conditions reached a maximum OD_{600} value of 0.7 with the exception of NCP262^{T}, which reached 1.0.

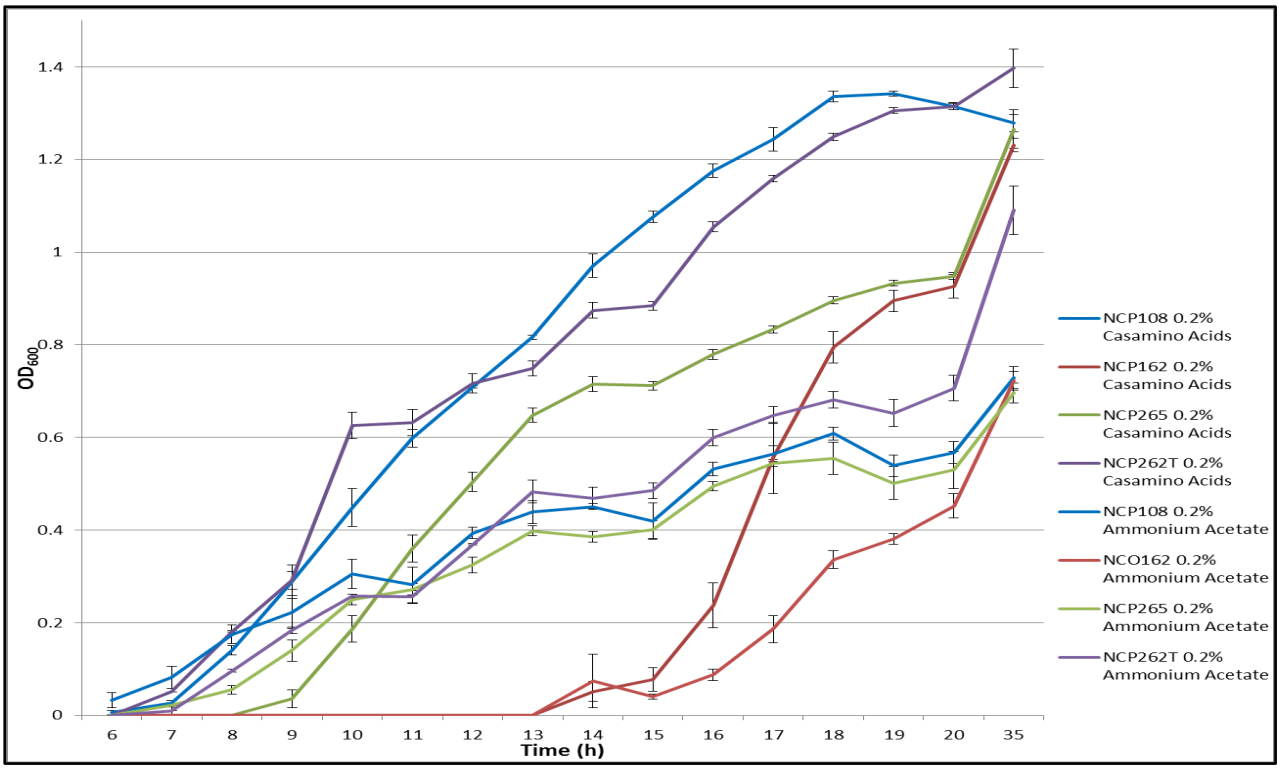


Figure 3.8: Growth of NCP *C. saccharobutylicum* strains in minimal media supplemented with 0.2% casamino acids compared to 0.2% ammonium acetate as the inorganic nitrogen source. Error bars represent the standard error of the mean, where n=3.

The same growth conditions were repeated for the same strains supplemented with ammonium sulphate, ammonium chloride and ammonium nitrate. All of these ammonium salts inhibited the start of growth and gave a reduced growth rate showing only limited growth after 35 h (data not shown). It was concluded that inorganic nitrogen alone is insufficient to support adequate growth and therefore solvent production. The delay in growth under inorganic nitrogen conditions is likely due to the fact that the organism must assimilate the inorganic ammonium ions (NH_4^+) into the amino acids necessary to maintain cell growth and division instead of transporting and using the amino acids found in the casamino acid supplemented media directly. GC analysis was performed on minimal media after 96 h of growth and it was determined that this defined minimal medium was not suitable for solvent production or any meaningful solvent comparisons (data not shown).

3.4.2.2 The effects of organic nitrogen on growth

Organic nitrogen in laboratory fermentation media is usually in the form of amino acids contained in casamino acids or yeast extract (YE). The importance of organic nitrogen in maintaining growth is depicted in Figure 3.9 for *C. beijerinckii* and Figure 3.10 for *C. saccharobutylicum* strains.

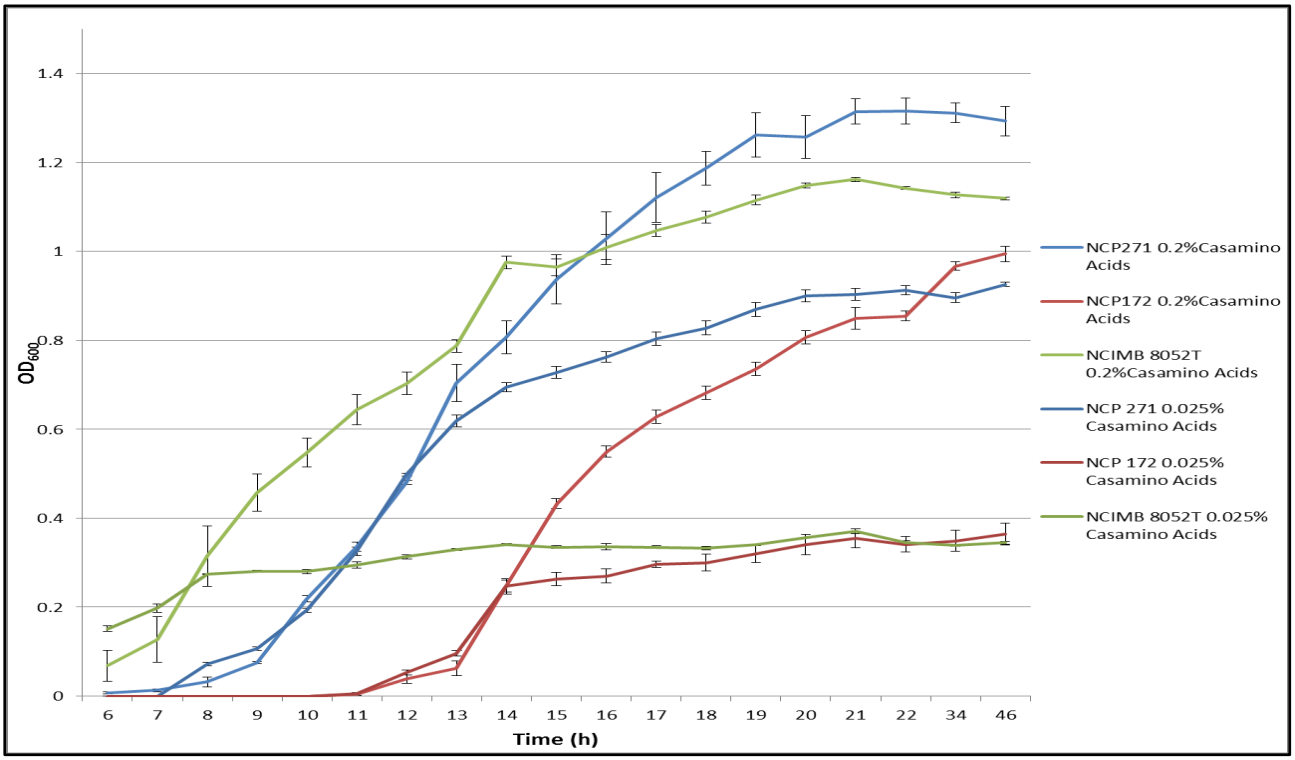


Figure 3.9: Growth of NCP *C. beijerinckii* strains in minimal media supplemented with 0.2% compared to 0.025% casamino acids as the organic nitrogen source. Error bars represent the standard error of the mean, where n=3.

The *C. beijerinckii* strains reacted differently to one another when grown in nitrogen limited media in terms of their maximum growth (Figure 3.9). NCP271 grew to a maximum OD_{600} value of 1.3 under high organic nitrogen conditions (0.2% casamino acids) and 0.9 under low organic nitrogen conditions (0.025% casamino acids). NCIMB 8052^T grew to a maximum OD_{600} value of 1.18 under high organic nitrogen conditions and 0.3 under low organic

nitrogen conditions. NCP172 grew to a maximum OD_{600} value of 1.0 under high organic nitrogen conditions and 0.3 under low nitrogen conditions. While different growth rates were observed for all the *C. saccharobutylicum* strains, they all grew to a similar maximum OD_{600} value of 1.255-1.4 under high nitrogen conditions and they all grew to a maximum of 0.64-0.77 under low nitrogen conditions (Figure 3.10).

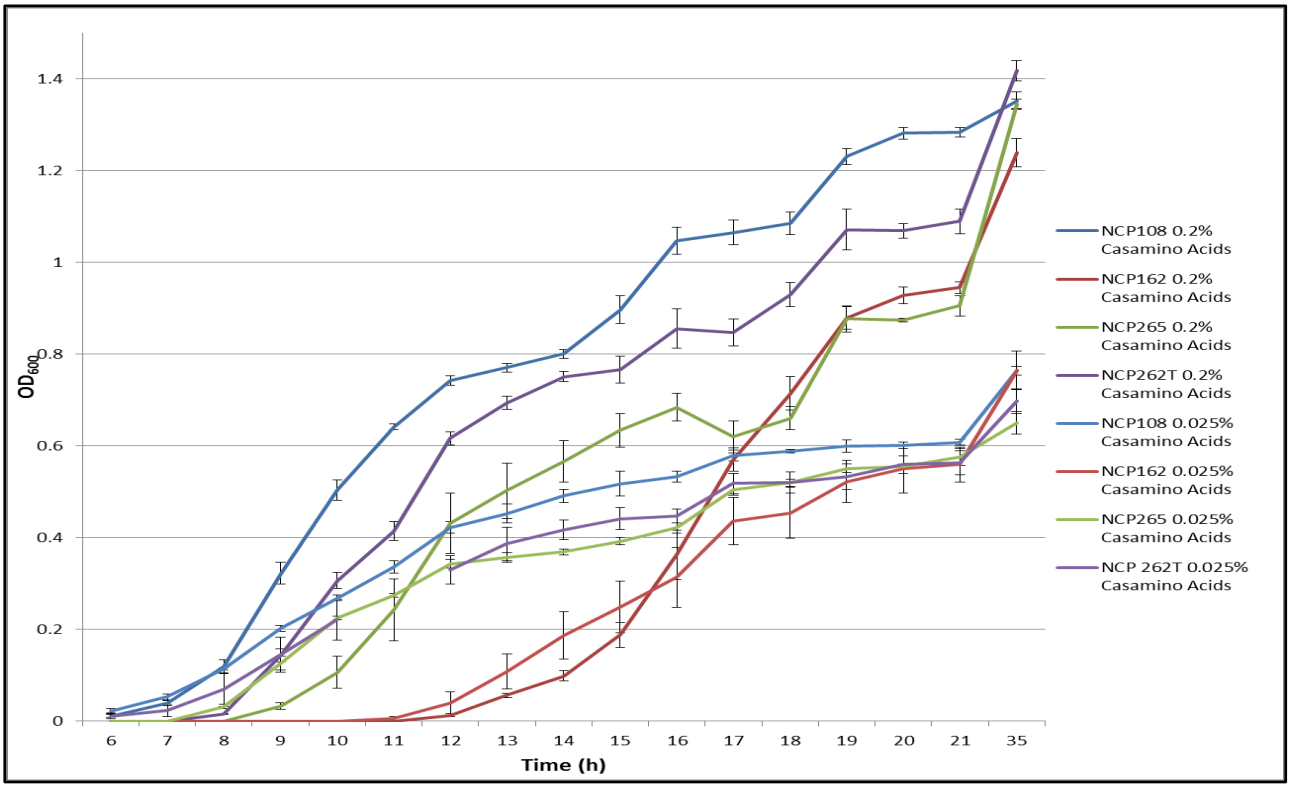


Figure 3.10: Growth of NCP *C. saccharobutylicum* strains in minimal media supplemented with 0.2% compared to 0.025% casamino acids as the organic nitrogen source. Error bars represent the standard error of the mean, where n=3.

Unsurprisingly, when the percentage casamino acids present in growth media was changed from a high organic nitrogen content of 0.2% to a low organic nitrogen content of 0.025%, there was a decrease in maximum growth attained by each strain for both species. This is in agreement with observations from Stutz *et al.* (2007), where similar growth studies were

performed on NCP262^{T}. The *C. beijerinckii* strains grew faster and achieved maximum growth by 21 h under these media conditions compared to the *C. saccharobutylicum* strains that only achieved maximum growth after 24 h. The *C. beijerinckii* strains did not cope as well with nitrogen stress compared to the *C. saccharobutylicum* strains with the exception of *C. beijerinckii* NCP271 that grew better than the *C. saccharobutylicum* strains. This simple experiment demonstrates that high organic nitrogen content in the media is required to maintain the optimal growth of clostridial cultures.

3.4.3 Investigation into the effects of amino acids during fermentation and growth

The increased transcription products involved in amino acid biosynthesis during solventogenesis (Table 3.1), prompted investigations into the effect of supplementing fermentation media with individual amino acids. Conventional NCP medium was used as a starting point for fermentation studies. Determining the composition of complex YE is impossible. The molecular mass of the various amino acids ranges from 132-155 g/mol. Therefore equal weight/volume concentrations were used to compare YE supplemented media to amino acid supplemented media. The NCP medium recipe was modified so that the organic nitrogen YE component was reduced from 4 g/L to a basal level of 1 g/L and then supplemented with various amino acids. This basal level of 1 g/L of YE is referred to as the unsupplemented control in the course of this chapter. Normal fermentation experiments were conducted on the modified NCP media and solvents assayed.

3.4.3.1 Amino acid trials in phosphate buffered fermentation media

Normal NCP medium contains calcium carbonate ($CaCO_3$) as a buffer, which makes it unsuitable for spectrophotometric- based growth studies. The $CaCO_3$ buffer in the medium was replaced with a phosphate buffer to facilitate spectrophotometric readings of NCP strain growth under various amino acid supplementations. Fermentation trials were performed to compare the buffering capacities between phosphate and $CaCO_3$ buffered NCP media using 4 g/L of YE. NCP260 produced less butanol under phosphate buffered compared to calcium carbonate buffered conditions for 4 g/L YE, however, this was used in initial studies so that the growth of the bacterial strains could be monitored.

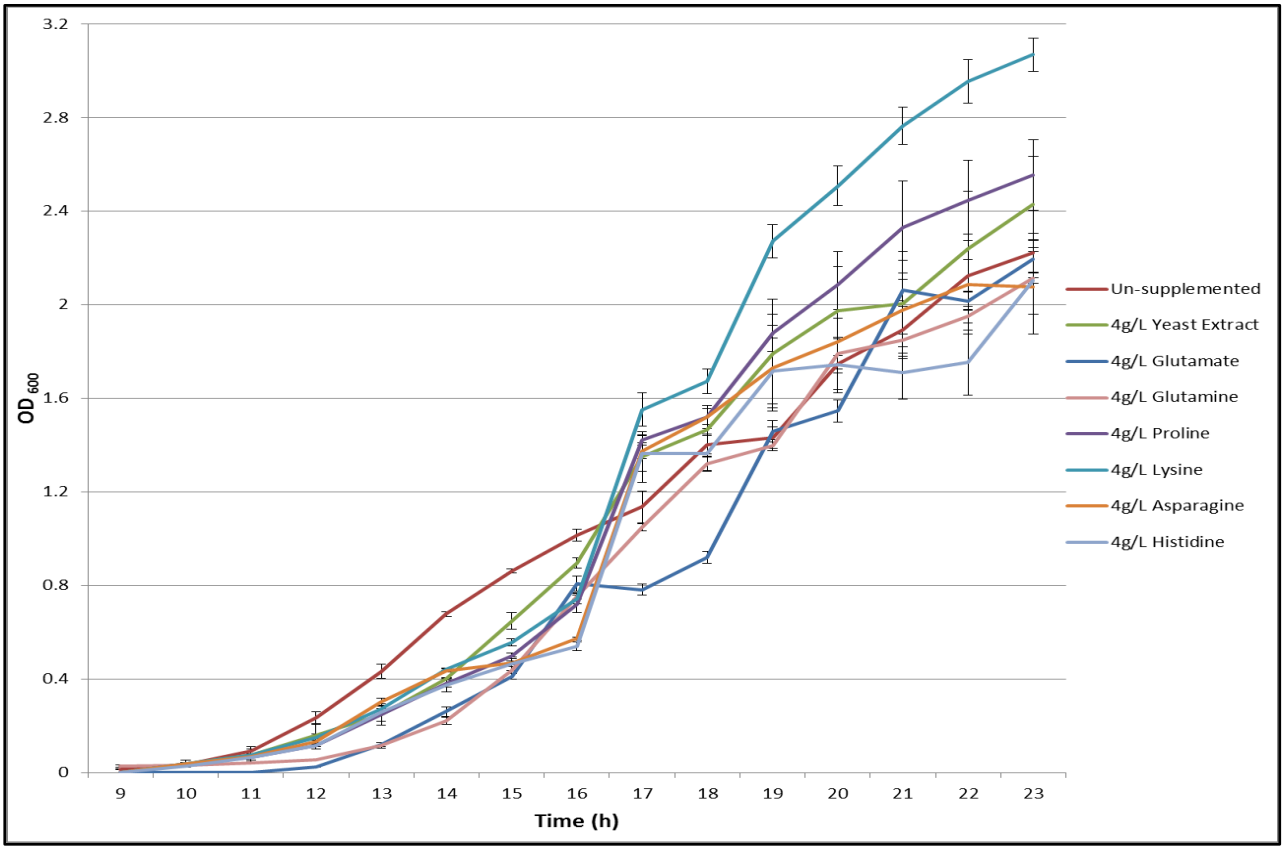


Figure 3.11: Growth of NCP260 in phosphate buffered 5% glucose NCP media supplemented with 4 g/L of various amino acids compared to 4 g/L YE- and un-supplemented controls. Error bars represent the standard error of the mean, where n=3.

The growth of NCP260 in phosphate buffered NCP media supplemented with various amino acids is shown above (Figure 3.11). Interestingly, NCP260 had a more pronounced lag phase in media supplemented with 4 g/L of glutamine, glutamate, histidine, asparagine and un-supplemented medium compared to 4 g/L YE supplemented medium after 23 h (Figure 3.11). The lag phase was shorter for NCP260 grown in media supplemented with 4 g/L lysine and proline compared to 4 g/L YE supplemented media ending at 12 h. However, the growth rates in the presence of different amino acids did not differ substantially, and the final OD_{600} values after 23 h were very similar.

Growth studies were performed on phosphate buffered media containing increasing concentrations of glutamine and glutamate. The initial lag phase between 9-14 h increased with increasing concentrations of glutamate present in the media (Figure 3.12). However, the final maximum level of growth after 48 h was similar at an OD_{600} of 2.2. The same trend was observed for glutamine: there was a greater lag phase with increasing concentrations of glutamine of between 12-14 h (Figure 3.13).

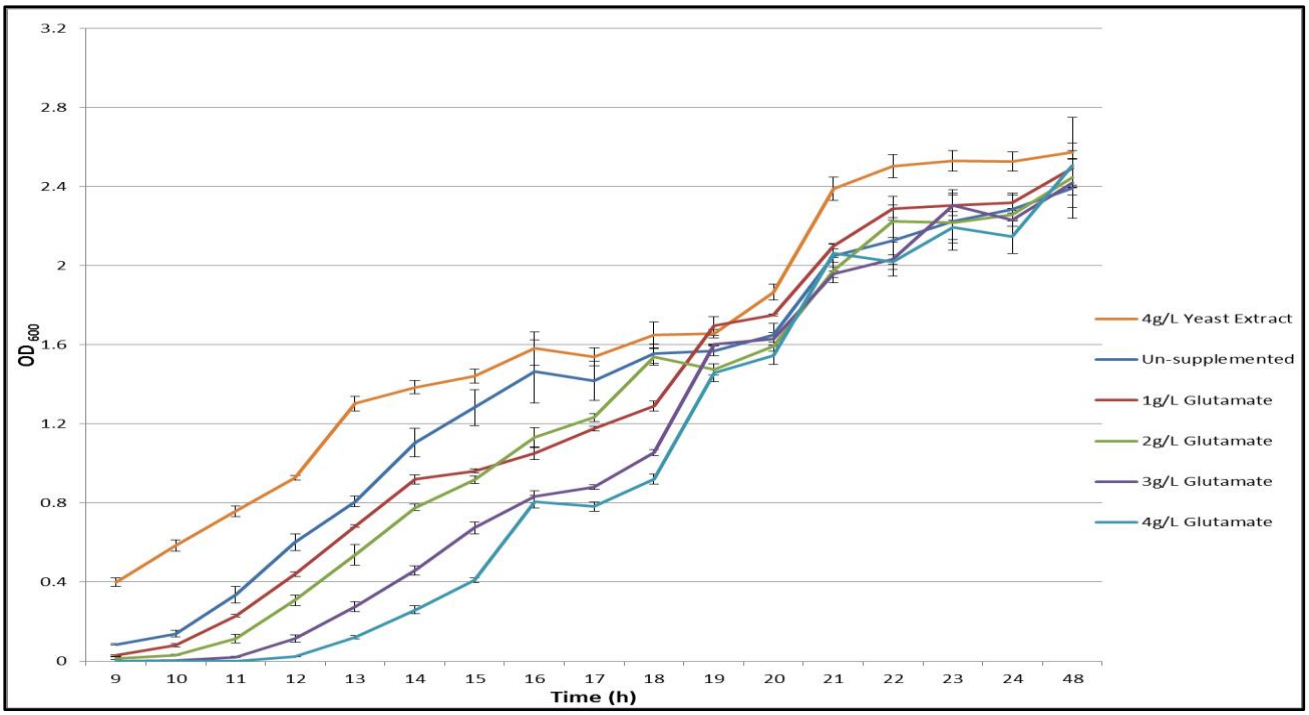


Figure 3.12: Growth of NCP260 on phosphate buffered 5% glucose NCP media supplemented with 1, 2, 3 and 4 g/L glutamate compared to 4 g/L YE- and un-supplemented controls. Error bars represent the standard error of the mean, where n=3.

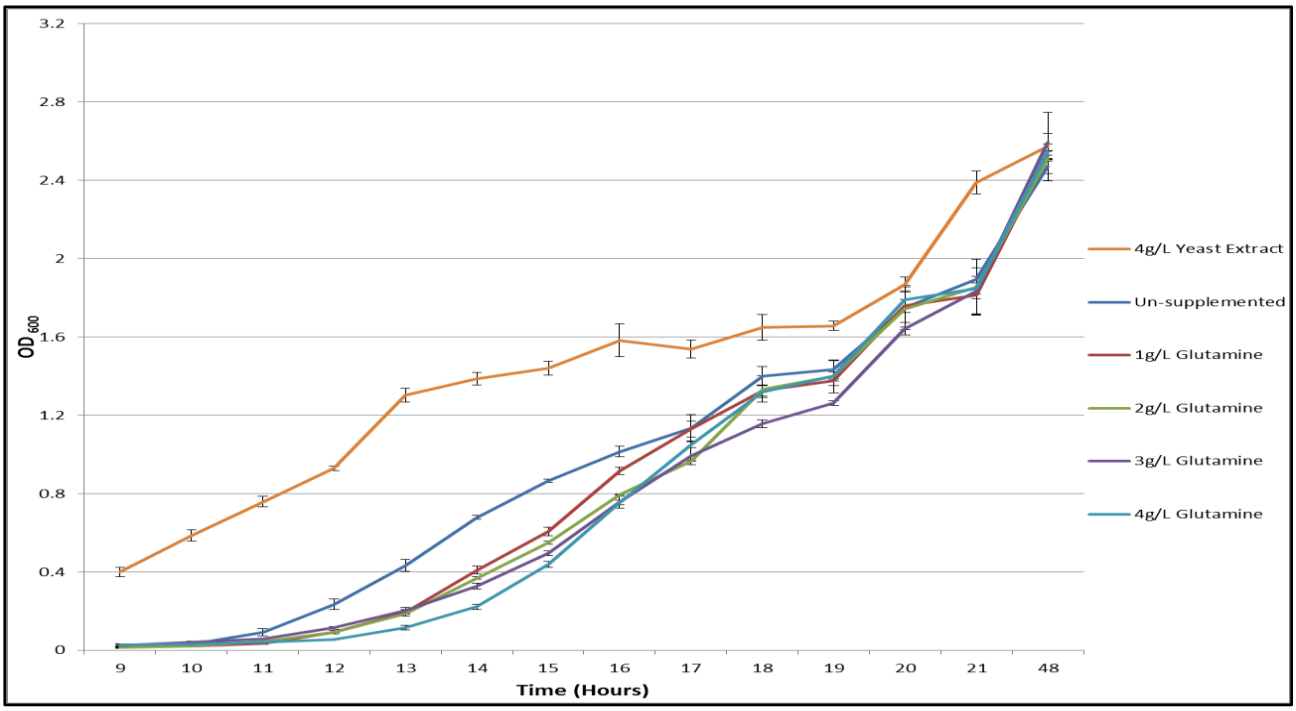


Figure 3.13: Growth of NCP260 on phosphate buffered 5% glucose NCP media supplemented with 1, 2, 3 and 4 g/L glutamine compared to 4 g/L YE- and un-supplemented controls. Error bars represent the standard error of the mean, where n=3.

The increasing lag phase times as a result of increasing concentrations of glutamine and glutamate is counterintuitive since one would expect cells to start growing faster with increasing concentrations of amino acids available for anabolic processes. YE supplemented medium contains many amino acids all of which can enter the cell through a variety of methods and transporters, thus there is no bottle neck for amino acids entering the cell and the cell can use the amino acids in metabolism immediately. In the case of a low basal level of YE supplemented with only a single amino acid, the cell must constantly take up that amino acid through specific amino acid transporters and then break the amino acid down and re-assimilate it into other amino acids using the α-ketoglutarate transaminase system. The metabolic shift combined with the constant energy required to take up one particular amino acid may account for the lag in growth observed with increasing concentrations of glutamine and glutamate.

Glutamine, glutamate and histidine supplementation under phosphate buffered media conditions significantly improved total solvent, butanol and acetone titres (Figure 3.14). The growth rate when supplementing with these amino acids is initially slower than other amino acids and YE supplemented conditions (Figure 3.11), however, all amino acid supplementation conditions resulted in grow to a similar maximum OD_{600} of 2.2 or more after 23 h. This strongly suggests that glutamine, glutamate and histidine confer beneficial qualitative properties to the solvent making process since it was demonstrated that there was no relationship between maximum growth and solvent production.

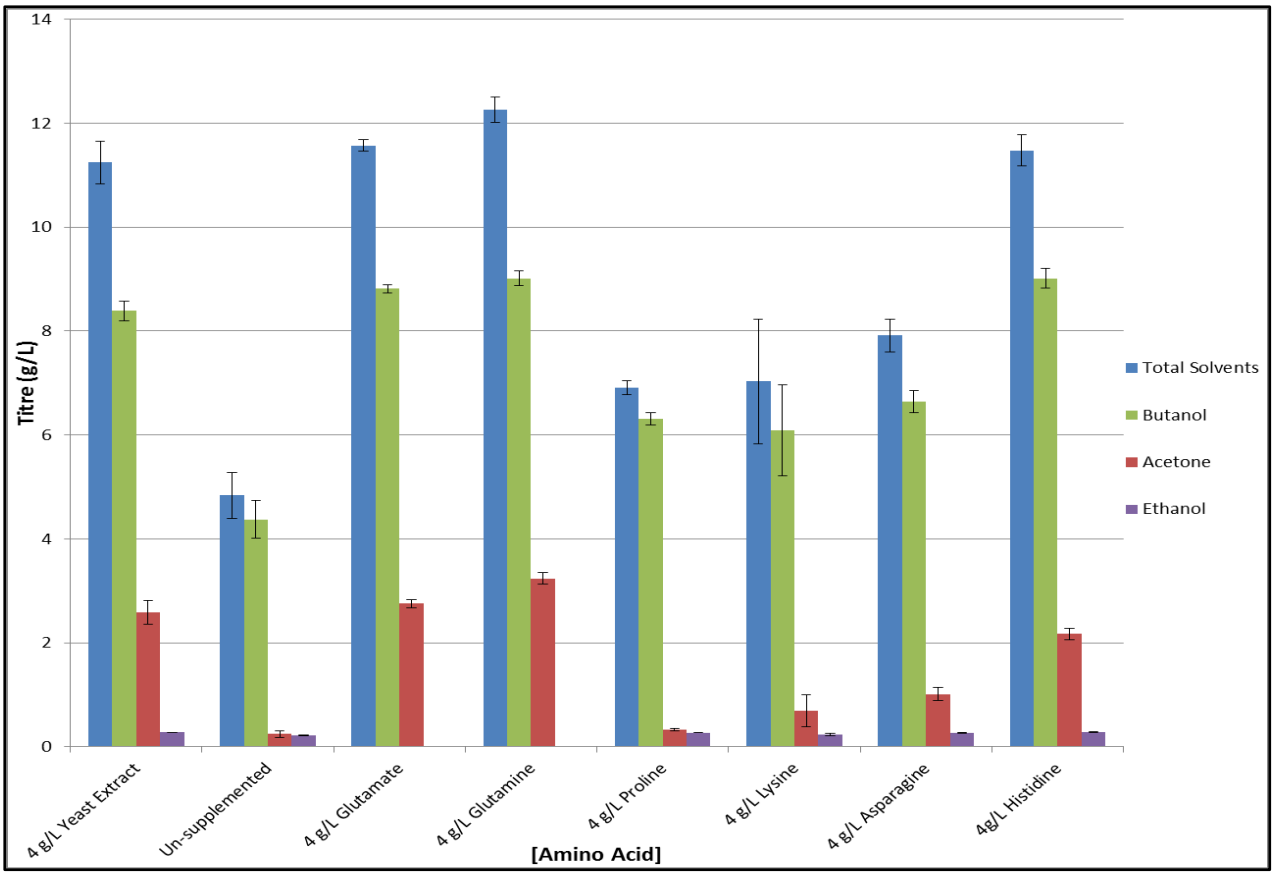


Figure 3.14: ABE solvent titres for NCP260 utilising various amino acids as the organic nitrogen source in phosphate buffered 5% glucose NCP media conditions. Error bars represent the standard error of the mean, where n=3.

Solvent production was measured after these single amino acids were added to regular NCP media which already contained 4 g/L of YE. Phosphate buffered NCP media containing 4 g/L of YE was supplemented with an additional 4 g/L of glutamate, glutamine, proline, lysine, histidine and asparagine and the butanol titres measured after 72 h of fermentation (Figure 3.15). There was no statistically significant difference between phosphate buffered NCP medium which contained the standard 4 g/L of YE and the same medium supplemented with additional 4 g/L of the various amino acids. This is unsurprising since the YE likely contains adequate quantities of all the amino acids supplemented, and indicates that the 4% YE contains sufficient quantities of all other growth factors such as vitamins to maintain optimal growth.

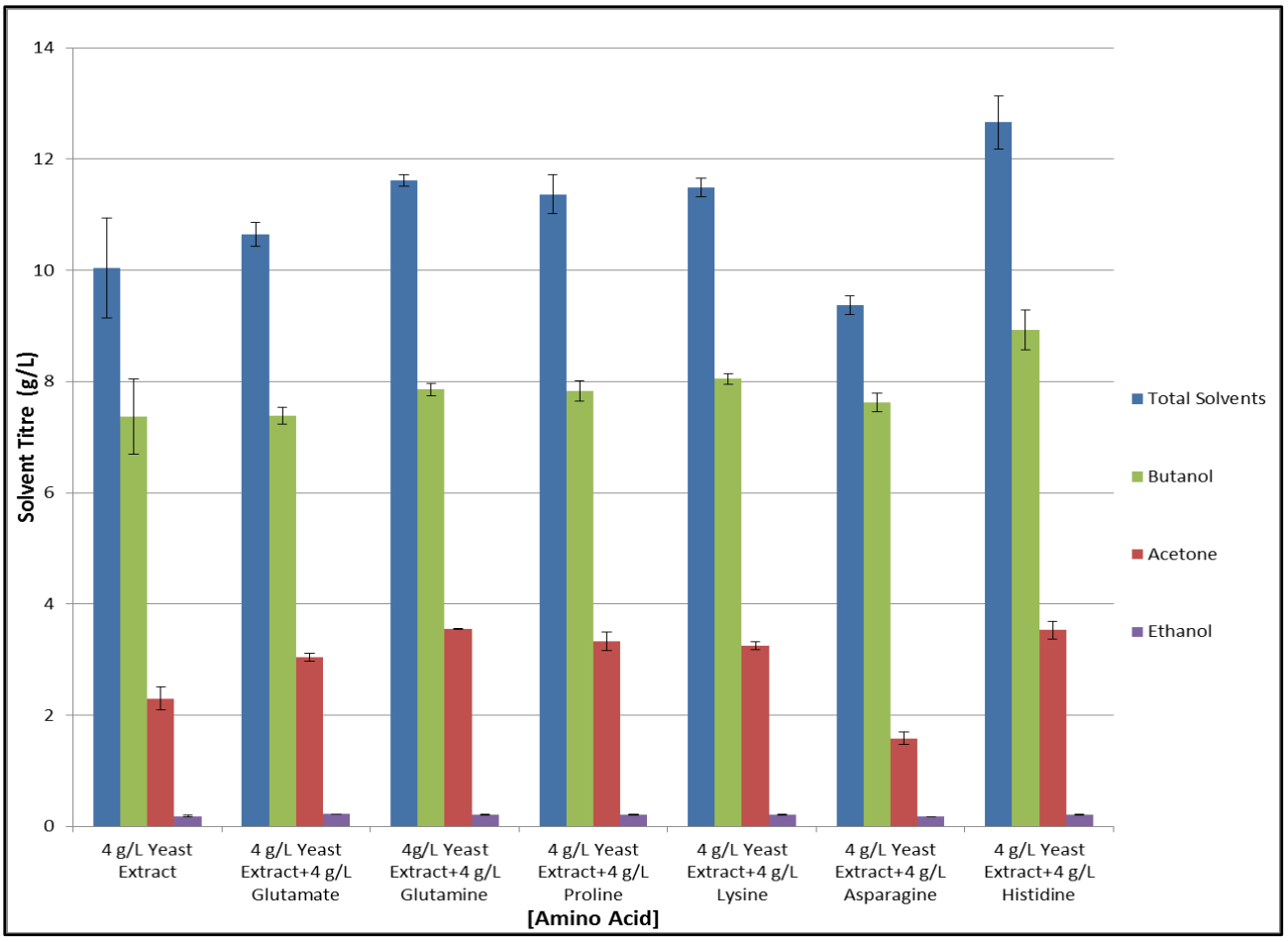


Figure 3.15: ABE solvent titres for NCP260 grown in phosphate buffered 5% glucose NCP media supplemented with 4 g/L YE as well as additional amino acids. Error bars represent the standard error of the mean, where n=3.

3.4.3.2 Amino acid trials in calcium carbonate buffered fermentation media

It was thought that the particularly low solvent titres of the un-supplemented media may be due to the phosphate buffer not being as effective as the $CaCO_3$ buffer for this particular concentration combined with the stress of low organic nitrogen availability (Figure 3.14). To that end, the effect of amino acid supplementation on a number of NCP strains was further investigated in calcium carbonate buffered NCP media. Fermentation trials were performed to compare butanol titres as a proxy for the buffering capacities between phosphate and $CaCO_3$ buffered NCP media using 4 g/L of YE or 4 g/L each amino acid.

NCP258 was selected as a representative of the NCP *C. saccharobutylicum* species since it was demonstrated in the MLST studies discussed in Chapter Two that it is closely related to NCP262^{T}. NCP271 and NCP260 were selected as representatives of the NCP *C. beijerinckii* species as NCP271 was shown to be able to adequately ferment glucose, sucrose and xylose sugar substrates in section 3.3.1 and NCP260 is closely related to NCIMB 8052^{T}, which has the whole genome sequence available. This section details the effects of glutamate, glutamine, lysine, proline, histidine and asparagine supplementation in fermentation media on butanol production.

The butanol titres for *C. saccharobutylicum* NCP258 under various organic nitrogen conditions are shown below (Table 3.3). NCP258 produced significantly more solvents under YE, glutamine, glutamate and histidine supplementation conditions compared to the un-supplemented control. In all cases, with the exception of asparagine supplementation, there is a trend of increasing butanol titres with increasing amino acid concentrations. High concentrations of 4 g/L of glutamate, glutamine and histidine were able to restore the butanol titres to that of normal 4 g/L YE cultures butanol titres. Increasing the concentration of lysine, asparagine and proline did not restore butanol titres to normal 4 g/L YE cultures butanol levels. Interestingly, the addition of 1 g/L of glutamate immediately increases the butanol titres to that of normal fermentation conditions.

Table 3.3: Butanol titres (g/L) for NCP258 on 5% glucose NCP media supplemented with various amino acids. ± represent the standard deviation of the mean, where n=3. Green shading indicates a statistically significant increase in butanol titres ($p < 0.05$) compared to the un-supplemented control.

[Butanol] for Glutamate Supplementation					
Un-supplemented	**1 g/L Glutamate**	**2 g/L Glutamate**	**3 g/L Glutamate**	**4 g/L Glutamate**	**4 g/L YE**
7.79±0.51	10.51±0.85	9.79±0.85	9.63±0.38	10.63±0.34	10.17±0.18
[Butanol] for Glutamine Supplementation					
Un-supplemented	**1 g/L Glutamine**	**2 g/L Glutamine**	**3 g/L Glutamine**	**4 g/L Glutamine**	**4 g/L YE**
7.79±0.51	7.92±0.68	8.12±0.50	9.29±0.11	10.09±0.08	10.17±0.18
[Butanol] for Lysine Supplementation					
Un-supplemented	**1 g/L Lysine**	**2 g/L Lysine**	**3 g/L Lysine**	**4 g/L Lysine**	**4 g/L YE**
7.79±0.51	7.49±1.63	7.88±1.77	9.47±0.87	8.86±1.50	10.17±0.18
[Butanol] for Proline Supplementation					
Un-supplemented	**1 g/L Proline**	**2 g/L Proline**	**3 g/L Proline**	**4 g/L Proline**	**4 g/L YE**
7.79±0.51	8.24±0.61	8.00±0.70	7.96±0.93	9.19±0.63	10.17±0.18
[Butanol] for Histidine Supplementation					
Un-supplemented	**1 g/L Histidine**	**2 g/L Histidine**	**3 g/L Histidine**	**4 g/L Histidine**	**4 g/L YE**
7.79±0.51	8.93±0.35	9.30±1.21	10.74±1.96	9.82±0.98	10.17±0.18
[Butanol] for Asparagine Supplementation					
Un-supplemented	**1 g/L Asparagine**	**2 g/L Asparagine**	**3 g/L Asparagine**	**4 g/L Asparagine**	**4 g/L YE**
7.79±0.51	8.38±0.33	8.46±0.17	7.95±0.24	8.44±0.35	10.17±0.18

Butanol titres produced under different amino acid conditions for *C. beijerinckii* NCP271 are shown below (Table 3.4). Media supplementation with high concentrations of all amino acids, except lysine resulted in significantly more solvents produced by NCP271 compared to un-supplemented conditions. The addition of proline immediately increased the butanol titres of NCP271 to normal fermentation butanol titres. There is a trend of increasing butanol titres with increasing concentrations of amino acids, with the exception of lysine. High

concentrations of glutamate, proline and histidine restored butanol titres to that of 4 g/L YE normal NCP medium. Increasing the concentrations of glutamine, lysine and asparagine did not restore butanol titres to normal.

Table 3.4: Butanol titres (g/L) for NCP271 on 5% glucose NCP media supplemented with various amino acids. ± represent the standard deviation of the mean, where n=3. Green shading indicates a statistically significant increase in butanol titres ($p < 0.05$) compared to the un-supplemented control.

[Butanol] for Glutamate Supplementation					
Un-supplemented	**1 g/L Glutamate**	**2 g/L Glutamate**	**3 g/L Glutamate**	**4 g/L Glutamate**	**4 g/L YE**
7.19±0.35	7.92±0.68	8.12±0.50	9.29±0.11	10.09±0.08	10.45±0.21
[Butanol] for Glutamine Supplementation					
Un-supplemented	**1 g/L Glutamine**	**2 g/L Glutamine**	**3 g/L Glutamine**	**4 g/L Glutamine**	**4 g/L YE**
7.19±0.35	7.39±0.29	7.39±0.31	7.60±0.10	8.50±0.37	10.45±0.21
[Butanol] for Lysine Supplementation					
Un-supplemented	**1 g/L Lysine**	**2 g/L Lysine**	**3 g/L Lysine**	**4 g/L Lysine**	**4 g/L YE**
7.19±0.35	7.49±1.63	7.88±1.77	9.47±0.87	8.86±1.50	10.45±0.21
[Butanol] for Proline Supplementation					
Un-supplemented	**1 g/L Proline**	**2 g/L Proline**	**3 g/L Proline**	**4 g/L Proline**	**4 g/L YE**
7.19±0.35	9.19±0.63	9.96±0.93	10.00±0.70	10.24±0.61	10.45±0.21
[Butanol] for Histidine Supplementation					
Un-supplemented	**1 g/L Histidine**	**2 g/L Histidine**	**3 g/L Histidine**	**4 g/L Histidine**	**4 g/L YE**
7.19±0.35	8.93±0.35	9.30±1.21	10.74±1.96	9.82±0.98	10.45±0.21
[Butanol] for Asparagine Supplementation					
Un-supplemented	**1 g/L Asparagine**	**2 g/L Asparagine**	**3 g/L Asparagine**	**4 g/L Asparagine**	**4 g/L YE**
7.19±0.35	8.71±0.26	8.46±0.17	8.48±0.29	8.44±0.35	10.45±0.21

The butanol titres for *C. beijerinckii* NCP260 under various amino acid conditions are shown below (Table 3.5). Significant increases in NCP260 butanol titre were observed for high concentrations of glutamate, glutamine, proline, lysine and histidine compared to the un-supplemented control media conditions. The addition of asparagine had no statistically significant effect on butanol titres for NCP260, except for 4 g/L of asaparagine supplementation. Low concentrations of glutamate, glutamine and histidine increased butanol titres compared to the YE supplemented control. There is a trend of increasing butanol titres with increasing amino acids for all amino acids used with the exception of asparagine. This is in contrast to the other *C. beijerinckii* strain, NCP271 where even low concentrations of proline and asparagine significantly increased butanol titres.

Table 3.5: Butanol titres (g/L) for NCP260 on 5% glucose NCP media supplemented with various amino acids. ± represent the standard deviation of the mean where n=3. Green shading indicates a statistically significant increase in butanol titres ($p < 0.05$) compared to the un-supplemented control.

[Butanol] for Glutamate Supplementation					
Un-supplemented	**1 g/L Glutamate**	**2 g/L Glutamate**	**3 g/L Glutamate**	**4 g/L Glutamate**	**4 g/L YE**
7.22±0.77	8.86±0.62	8.85±0.79	9.15±0.17	9.50±0.12	9.15±0.36
[Butanol] for Glutamine Supplementation					
Un-supplemented	**1 g/L Glutamine**	**2 g/L Glutamine**	**3 g/L Glutamine**	**4 g/L Glutamine**	**4 g/L YE**
7.22±0.77	9.27±0.75	8.50±0.82	9.54±0.21	10.08±0.73	9.15±0.36
[Butanol] for Lysine Supplementation					
Un-supplemented	**1 g/L Lysine**	**2 g/L Lysine**	**3 g/L Lysine**	**4 g/L Lysine**	**4 g/L YE**
7.22±0.77	6.84±0.34	7.41±0.35	8.12±0.49	8.89±0.43	9.15±0.36
[Butanol] for Proline Supplementation					
Un-supplemented	**1 g/L Proline**	**2 g/L Proline**	**3 g/L Proline**	**4 g/L Proline**	**4 g/L YE**
7.22±0.77	5.07±0.93	8.13±1.01	7.84±0.23	9.81±0.39	9.15±0.36
[Butanol] for Histidine Supplementation					
Un-supplemented	**1 g/L Histidine**	**2 g/L Histidine**	**3 g/L Histidine**	**4 g/L Histidine**	**4 g/L YE**
7.22±0.77	8.28±0.54	8.96±0.13	9.00±0.16	8.97±0.25	9.15±0.36
[Butanol] for Asparagine Supplementation					
Un-supplemented	**1 g/L Asparagine**	**2 g/L Asparagine**	**3 g/L Asparagine**	**4 g/L Asparagine**	**4 g/L YE**
7.22±0.77	7.61±0.07	7.63±0.13	7.63±0.20	8.00±0.11	9.15±0.36

Table 3.6 shows a comparison of the butanol titres obtained for NCP260 in phosphate compared to calcium carbonate buffered NCP media. This strain produced significantly more butanol in phosphate medium containing any of the amino acids in comparison to un-supplemented media. However, only glutamate, glutamine and histidine were comparable to that of the YE supplemented media control. Calcium carbonate buffering resulted in

significant increased butanol production levels in NCP media containing 4 g/L proline, lysine, glutamate, glutamine and histidine respectively, in comparison to un-supplemented media. Clearly, carbonate buffering is more beneficial to producing solvents since the butanol titres were comparable to that of 4 g/L YE-supplemented media.

Table 3.6: Comparative butanol titres (g/L) for NCP260 on phosphate versus calcium carbonate buffered NCP media containing 5% glucose and supplemented with various amino acids. ± represent the standard deviation of the mean, where n=3. Green shading indicates a statistically significant increase in butanol titres ($p < 0.05$) compared to the un-supplemented control.

[Butanol] (g/L) in phosphate buffered NCP media							
Un-supplemented	**4 g/L Asparagine**	**4 g/L Proline**	**4 g/L Lysine**	**4 g/L Glutamate**	**4 g/L Glutamine**	**4 g/L Histidine**	**4 g/L YE**
4.38±0.64	6.64±0.37	6.31±0.20	6.09±1.52	8.82±0.14	9.02±0.24	9.02±0.33	8.39±0.33
[Butanol] (g/L) in calcium carbonate buffered NCP media							
Un-supplemented	**4 g/L Asparagine**	**4 g/L Proline**	**4 g/L Lysine**	**4 g/L Glutamate**	**4 g/L Glutamine**	**4 g/L Histidine**	**4 g/L YE**
7.22±0.77	8.00±0.11	9.81±0.39	8.89±0.43	9.50±0.12	10.08±0.73	8.97±0.25	9.15±0.36

The results from the amino acid trials are variable and differ slightly between the two species investigated. *C. saccharobutylicum* NCP258 only shows a significant butanol response at higher concentrations of glutamate, glutamine and histidine. The *C. beijerinckii* strains exhibit a far more acute response to the limited nitrogen conditions for glutamate, glutamine and histidine. *C. beijerinckii* NCP271 responded to all amino acid supplementation except lysine. Interestingly, even low concentrations of proline and asparagine were sufficient to supplement NCP271.

Many aminoacyl-tRNA synthetases and systems involved in amino acid biosynthesis in Gram positive bacteria are regulated through interactions of uncharged tRNA effector molecules

with a ribosomal T-box domain. This domain is typically found upstream of the coding sequence and forms an antiterminator structure. Regulation of biosynthesis genes is therefore controlled by these effector molecules (Grundy *et al.*, 1997). There is evidence that asparagine synthesis which is regulated by T-box-tRNA^{Asn} interactions may be connected to the acidogenesis to solventogenesis switch (Saad *et al.*, 2012). Since asparagine is supplemented directly in the fermentation, there is no need for the synthesis of this amino acid, which may affect the regulation of the pathway and explain why asparagine supplementation was largely ineffective in increasing butanol titres. Intercellular accumulation of proline has been shown to increase cell viability of yeast cells during ethanol fermentation (Takagi *et al.*, 2005). As yet, there is no evidence to link proline supplementation to ABE solvent production in *Clostridium.*

Glutamate is central to the biosynthesis of other amino acids whereby transaminase enzymes convert glutamate to other free α-amino acids: Glutamate + α-keto acid → α-ketoglutarate + α-amino acids (Reid and Stutz, 2005). The role of glutamine may involve glutamine synthetase which catalyses a transamination reaction of the amide group from glutamine to form glutamate: Glutamine + 2-ketoglutarate + NADPH + H^+ → 2-glutamate + $NADP^+$. This newly formed glutamate can then be used in transamination reactions as mentioned above. Intracellular glutamate pools are clearly central to metabolic homeostasis and therefore solventogenesis. The role of lysine in fermentation remains unknown at this point.

Interestingly, amino acid supplementation conditions reduced the amount of acetone and ethanol produced for all strains investigated. The butanol: acetone ratios under amino acid supplementation conditions were higher compared to the normal YE conditions. Much less

acetone was produced under amino acid supplementation conditions resulting in a butanol: acetone ratio of greater than 2.0, unlike the normal 4 g/L YE NCP media conditions for all three strains investigated (Appendix B, Figures B3-B5).

The difference in solvent titres between un-supplemented phosphate buffered media and media supplemented with 4 g/L YE indicates that amino acids are involved in regulating pH as well as biosynthesis. Supplementation of glutamate, glutamine and histidine specifically increased butanol titres phosphate buffering conditions. This warrants further investigation into the role of these amino acids in acid stress since surviving acidogenesis is a prerequisite to solventogenesis resulting in increased solvent formation.

3.5 Conclusions

The UCT NCP strain collection has a diverse sugar substrate preference. There are a number of candidate strains for glucose, sucrose and xylose substrate fermentation that favour butanol metabolism to produce high yields and titres of biobutanol over a 72 h fermentation period. Strains NCP J, NCP249 NCP172 and NCP199 ferment glucose and generate high butanol yields of 0.219, 0.218, 0.214 and 0.214 respectively. Strains NCP254, NCP259, NCP268 and NCP199 utilise sucrose well in fermentations, resulting in butanol yields of 0.216, 0.213, 0.227 and 0.212 respectively. Strains NCP195, NCP259, NCP172 and NCP265 metabolise xylose to produce butanol yields of 0.214, 0.195, 0.177 and 0.177 respectively. Strains that performed on xylose substrates could be tested on crude hemicellulose substrates for potential second generation biobutanol industrial applications. Strains NCP195, NCP265, NCP271, NCP J, NCP259 and NCP172 are able to ferment all three substrates to produce high titres and yields of butanol. These strains are robust and can be used in many industrial applications.

Laboratory media supplemented with inorganic ammonia salts resulted in a delayed growth onset and a decrease in maximum growth, resulting in poor solvent production. Organic nitrogen is essential for normal growth and solvent production. The same maximum cell densities are observed regardless of which amino acid is supplemented into the fermentation media. *C. beijerinckii* strains NCP271 and NCP260 and *C. saccharobutylicum* strain NCP258 demonstrated significantly improved ($p < 0.05$) solvent production when grown in media supplemented with glutamate, glutamine and histidine. These amino acids improve the quality of cultures, improving butanol yields, without drastically changing the biomass quantity. These amino acids may play a role in acid survival during acidogenesis since they

consistently contributed to improved butanol yields under both calcium carbonate and phosphate buffered media conditions.

Chapter Four
The Role of Nitrogen Metabolism in Maintaining Intracellular Amino Acid Pools during Acidogenesis

Contents

4.1 Abstract

The nitrogen metabolism system in solventogenic *Clostridium* is not well characterised. The role of selected nitrogen metabolism genes was investigated using bioinformatics based on the genome sequence of *C. beijerinckii* ATCC8052^T. Two putative ammonia transporter genes, a putative glutamine transporter gene, two putative PII-like nitrogen regulator genes and a putative signal transduction histidine kinase gene were identified. Strategies for construction of knock out (KO) mutants were designed to investigate the loss of function of these six genes using the ClosTron site-directed mutagenesis system. Mutants of the ammonia and glutamine transporter genes should show disruption of glutamine and ammonia traffic into the cell, which may affect cell growth and solvent production. PII-like nitrogen regulator mutants were attempted to assess whether these putative regulators have any global control over nitrogen metabolism in *C. beijerinckii*. The histidine kinase mutant was attempted to assess whether it interacts with the NitR regulator of the GS-GOGAT system. Successful transfer of targeted ClosTron plasmids was confirmed in NCIMB 8052^T and NCP260 strains, however, integration under erythromycin selective pressure was not observed and no mutants were obtained.

The role of the amino acids, asparagine, proline, lysine, glutamine, glutamate and histidine in generating an Acid Tolerance Response (ATR) was investigated. Only histidine and glutamate supplementation significantly increased cell survival relative to ammonium acetate supplementation during low pH 5.0 conditions. Transcription of the *glnA*, *gltA* and *nitR* genes encoding GS, GOGAT and their nitrogen regulator, NitR, was monitored during acid shock conditions by quantitative Real Time PCR (qRT-PCR). *gltA* and *nitR* gene transcription was upregulated two-fold during pH 5.8 acid induction compared to uninduced pH 6.5 conditions.

Interestingly, *glnA* mRNA levels remained unchanged for both conditions. The transcription events of the GS-GOGAT operon during acid shock correlate with the observation that the biosynthesis of glutamate and not glutamine is responsible for the ATR during acidogenesis. Glutamate and histidine increase cell survival during acidogenesis so that more viable cells enter solventogenesis, ultimately improving solvent yields.

4.2 Introduction

Nitrogen metabolism and the maintenance of intracellular nitrogen balance are of critical importance to support life in all organisms. Without the appropriate nitrogen distribution, the biosynthesis of components necessary for cell growth and division cannot occur since enzymes and cell structures are primarily comprised of nitrogen-containing amino acids. The transport and biosynthesis of these amino acids are therefore of paramount importance to normal growth and survival of bacteria. This is particularly evident in the natural habitats of bacteria where the environment is continuously changing and nutrient starvation is commonplace. The reaction to nutrient starvation has been described as the stringent response (Chatterji and Ojha, 2001). Because access to constant nutrient supply in the form of organic nitrogen is rare in their natural environment, bacteria have evolved various mechanisms of assimilating organic and inorganic nitrogen from the environment. This can be achieved either through nitrogen fixation of atmospheric nitrogen (Chen *et al.*, 2001) or through the assimilation of inorganic nitrogen compounds such as ammonia (NH_4^+) (Usdin *et al.*, 1986). In times of nitrogen limitation, solventogenic *C. saccharobutylicum* and *C. beijerinckii* have been shown to use a system of assimilating inorganic ammonia from the soil environment into organic nitrogen in the form of the amino acids glutamate and glutamine (Stutz *et al.*, 2007).

This ammonia assimilation system is comprised of the enzymes glutamine synthetase (GS) and glutamate synthase (GOGAT) (Figure 4.1). Glutamine can be formed from ammonia using GS which catalyses the ATP-dependent amidation of glutamate. Glutamate can be synthesized from glutamine, 2-ketoglutarate and an intracellular proton using GOGAT in an NADPH dependent manner to produce glutamate and NAD^+. These two enzymes typically

operate in a cyclic manner so that overall the net reaction is as follows: NH_4^+ + 2-ketoglutarate + NADPH + H^+ → glutamate + $NADP^+$. A final transaminase reaction can convert the resulting glutamate into various amino acids (Reid and Stutz, 2005).

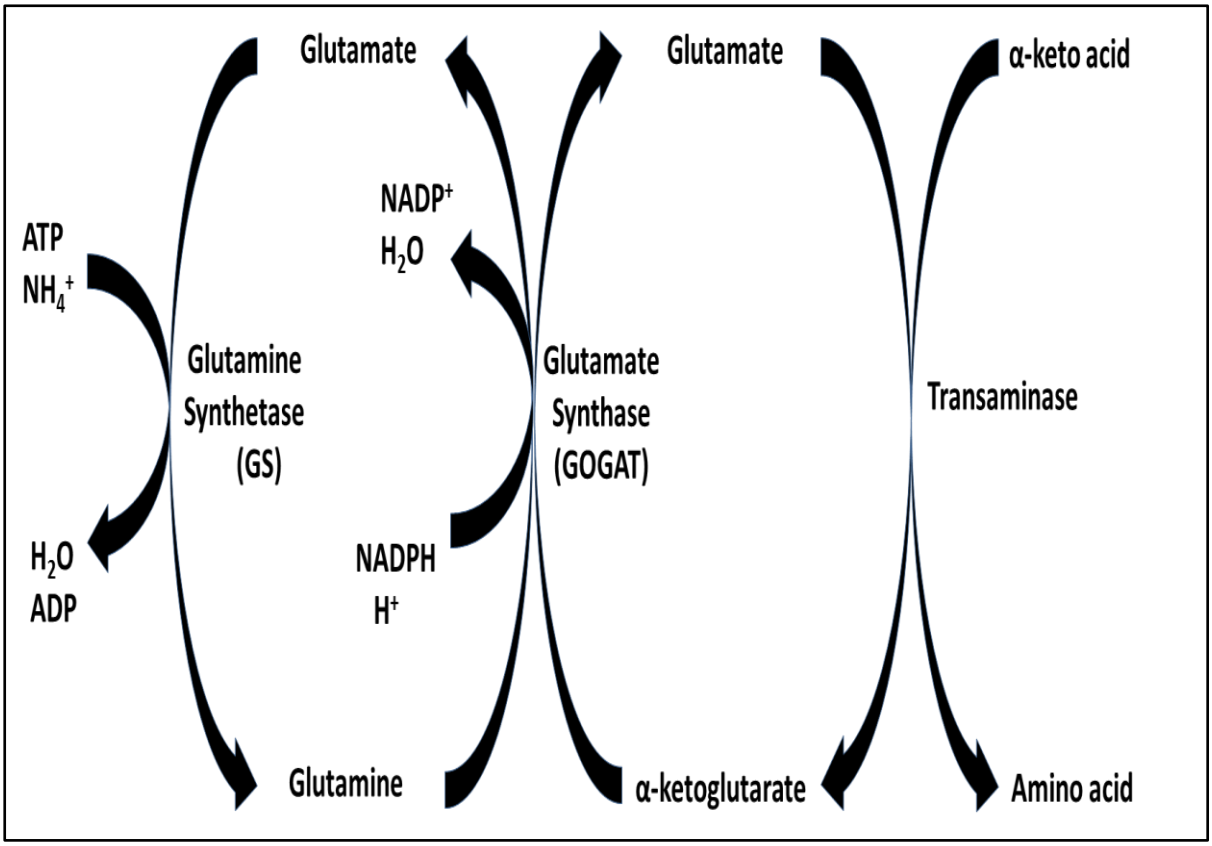


Figure 4.1: The glutamine synthetase (GS), glutamate synthase (GOGAT) and transaminase reactions including the components used and products formed in the reactions. Diagram adapted from Prescott, Harley and Klein's Microbiology (Willey *et al.*, 2008).

The GS-GOGAT system is therefore a core component of nitrogen metabolism and acts as a safety mechanism that ultimately ensures amino acid production when cells are under nitrogen or nutrient stress. The ammonia required in this reaction originates in the external environment and enters the cell through ammonia channels or transporters. These channels have been described in Gram negative *E. coli* and Gram positive *Corynebacterium*

glutamicum (Meier-Wagner *et al.*, 2001; Thornton *et al.*, 2006). Glutamine transporters control the traffic of glutamine into the cell which in turn can affect the GS-GOGAT enzyme activity. A Homologue of the glutamine transporter, *glnQ*, has been identified in *Streptococcus* (Tamura *et al.*, 2002). These ammonia and glutamine transporters have not been characterised in solventogenic *Clostridium* yet.

In *C. saccharobutylicum* and *C. beijerinckii*, the genes encoding the GS and GOGAT enzymes are clustered on the chromosome, with a gene encoding the regulator NitR to form a *glnA-nitR-gltAB* gene configuration. The signal mechanism that governs this system is unclear; however, the GS-GOGAT system may be linked to a global nitrogen regulation system. During nitrogen limitation in *E. coli*, the PII protein senses cellular 2-oxoglutarate and glutamine concentrations and activates the histidine kinase, NtrB, which in turn phosphorylates the nitrogen regulator, NtrC (Jiang *et al.*, 2000). NtrC controls the transcription of nitrogen-regulated genes including GS (Weiss *et al.*, 2002). A paralogue of the PII protein was found to regulate nitrogen metabolism in *Synechococcus elongates* (Laichoubi *et al.*, 2011). In *C. saccharobutylicum*, the NitR is an ANTAR (AmiR and NasR Transcription Antitermination Regulators) regulator, and as such, would interact with a histidine kinase (Shu and Zhulin, 2002). To date, no global regulator such as PII, or a NitR histidine kinase has been identified or characterised at molecular level in the solventogenic *Clostridium* strains.

In addition to maintaining nitrogen balance, solventogenic *Clostridium* species must survive the low pH conditions of acidogenesis before entering solventogenesis (Scotcher *et al.*, 2005). The accumulation of metabolic acids in solventogenic clostridia is far greater in

industrial fermentations than in their natural environment. This is because culturing in industry provides optimal growth conditions and an excess of nutrients. When metabolic acids reach sudden high concentrations, the destructive effect of low pH causes cell death before these acids can be re-assimilated into neutral solvents. In industry this phenomenon is commonly referred to as "acid crash" and results in major financial losses (Maddox *et al.*, 2000). Different buffering systems and pH control systems are employed in both batch and continuous culture ABE fermentation systems in order to regulate the rate of decrease in pH during acidogenesis (Zhang *et al.*, 2009; Ni *et al.*, 2012). Cells enter solventogenesis when the concentrations of acid end-products, among other factors, trigger the shift to solventogenesis (Richter *et al.*, 2012; Yang *et al.*, 2013).

The adaptations that solventogenic clostridia evolved in response to surviving the stresses of acidogenesis imposed by their anaerobic fermentative metabolism is a form of Acid Tolerance Response (ATR). Amino acids have been implicated in generating ATR in different species of bacteria either directly through media supplementation or indirectly through biosynthesis of a specific amino acid (Molenaar *et al.*, 1993; Hersh *et al.*, 1996; Lu *et al.*, 2013b). The role of amino acids in the ATR during acidogenesis in solventogenic clostridia has not been investigated.

In the previous chapter, it was demonstrated that supplementation of fermentations with glutamine, glutamate or histidine resulted in higher levels of butanol. In order to determine whether this effect was linked to the ATR in *Clostridium*, the contribution of asparagine, lysine, proline, glutamate, glutamine and histidine to the ATR and to improving cell survival were determined in this chapter. The expression levels of the GS-GOGAT operon genes

during acid shock were determined by qRT-PCR. A model is proposed to explain the increase in solvent yields observed and the role of these amino acids in the ATR.

Bioinformatics was used to identify and characterise a number of nitrogen metabolism genes with putative functions in *C. beijerinckii* NCIMB 8052^{T}. These include two ammonia transporter genes and a glutamine transporter gene. In addition, two genes encoding PII-like global regulators and a gene encoding histidine kinase clustered with the glutamine transporter gene were identified as genes with a potential regulatory function. In order to confirm functionality of these genes, the ClosTron mutagenesis system was implemented in an attempt to generate mutants in these six nitrogen metabolism gene targets.

4.3 Materials and Methods

4.3.1 Acid survival assay

Overnight (18 h) cultures of NCP260 cells were grown anaerobically in Clostridial Basal Medium (CBM) broth (O'Brien and Morris, 1971). A 10% 18 h inoculum was introduced into Glucose Salts Mineral Medium (GSMM) (Holdeman *et al.*, 1977) supplemented with 1mM of asparagine, glutamine, glutamate, lysine, proline, histidine or 0.2% ammonium acetate (in the case of no amino acid supplementation) as the sole nitrogen source. Exposure to the sub-lethal pH 5 in media supplemented with ammonium acetate was used as a non-amino acid control. The medium was adjusted to pH 6.5 for 100% survival pH control experiments and pH 5 for acid shock experiments and incubated anaerobically at 37 ℃ for 4 h. The cultures were serially diluted from 10^{-1} to 10^{-3} and spread-plated onto CBM agar. After 24 h, colony counts were performed and the number of pH 5 colonies was compared to that of pH 6.5 and percentage survival calculated. The percentage survivors from pH 5 relative to pH 6.5 conditions were calculated as an average of three biological repeats and each biological repeat consisted of three technical repeats. Statistically significant differences in percentage cell survival for different amino acid supplementations in the acid resistance assay was determined by an ANOVA LSD test with an alpha threshold of 0.05 to the homogenous group (degrees of freedom = 12) using the software Statistica (StatSoft, Inc. (2012). STATISTICA (data analysis software system), version 11. www.statsoft.com).

4.3.2 RNA extraction

NCP260 was inoculated into 10 mL 5% glucose Reinforced Clostridial Medium (RCM) (Oxoid) broth pH 6.5 and grown anaerobically at 37 °C for 18 h. The inoculum was transferred to fresh 20 mL 5% glucose RCM broth at either pH 6.5 for uninduced conditions or pH 5.8 for induced conditions and diluted to an OD_{600} value of 0.1. Cells were grown to early exponential phase corresponding to an OD_{600} value of 0.4. The experiment consisted of three biological repeats for the uninduced state and three biological repeats for the induced condition. Cells were stabilised prior to RNA extraction by the addition of RNA Protect (Qiagen) in a 2: 1 RNA Protect to cell suspension volume ratio and mixed by thorough vortexing. Cells were pelleted by centrifugation at 5000 *g* for 10 min. Total mRNA was extracted from the cell pellets using the RNeasy Mini Kit (Qiagen), with the following modification: an off-column DNase step was incorporated before the RNA ethanol precipitation step. The solution was exposed to10 U DNase (Thermo Scientific) and incubated at 37 ℃ for 30 min.

The concentration of the RNA was determined using NanoDrop spectrophotometry (Nanodrop ND-330, Inqaba Biotechnical Industries (PTY) Ltd. South Africa). In order to check for the presence of any contaminating DNA in the RNA extracts, a 16s *rRNA* PCR was performed on extracted RNA and electrophoresed on a 1% agarose gel at 60 V. The integrity of the RNA was determined by 3-(N-morpholino) propanesulfonic acid (MOPS) gel (4.185% MOPS, 0.68% CH_3COONa, 0.5 M EDTA, 1% agarose, pH 7) electrophoresis. The RNA sample was added to RNA loading buffer (1 µg RNA, 50% formamide; 1.7% formaldehyde, 5% 10x MOPS buffer, 10% tracking dye and 0.01% ethidium bromide) and electrophoresed for 3 h at 50 V. A 1x MOPS buffer was used as the running buffer. The agarose gels were

visualised under short wavelength UV light using the GelDoc (BioRad) and photographed. The RNA was stored at -80 °C.

4.3.3 cDNA synthesis

The Bioline Tetro cDNA Synthesis Kit (BIO-65043, Bioline) was used to generate cDNA from 500 ng of the purified RNA template in a 40 µL total reaction volume for each biological repeat. This was performed in duplicate according to manufacturer's instructions using random hexamer primers. Successful conversion of RNA to cDNA was assessed by formation of 16s *rRNA* PCR product for the converted cDNA. Each biological cDNA conversion was confirmed in duplicate. Once successful conversion was confirmed, the cDNA duplicates were pooled for each biological repeat. Pooled cDNA stocks comprising of equal volumes of both induced and uninduced biological samples were used to generate the qRT-PCR standard curves. Smaller aliquots of individual biological and pooled cDNA were made to ensure cDNA stocks were only freeze-thawed once for downstream reactions. The cDNA aliquots were stored at -20 ℃.

4.3.4 qRT-PCR primers, reagents and cycling conditions

Primers were designed using Beacon Designer program (Premier Biosoft). The amplicon size and nucleotide sequence of all primers used in this experiment are shown below (Table 4.1). The arrangement and relative positions of qRT-PCR primers are depicted below (Figure 4.2). Primers were synthesised and purified using High Pressure Liquid Chromatography (HPLC) (University of Cape Town Oligo Synthesis Service) and their site specificity was tested using BLAST (Altschul *et al.*, 1990) as well as by PCR.

Table 4.1: A list of the qRT-PCR primers used in this study.

Primer Name	PCR Product	Nucleotide Sequence 5'-3'	Product size (BP)
qRT-PCR 16s F	16s *rRNA*	AGCTAGTTGGTGAGGTAACGGC	125
qRT-PCR 16s R		TGCAATATTCCCCACTGCTGCC	
qRT-PCR *glnA* F	*glnA*	TCTTGCGGCTGTTACTAACCCG	123
qRT-PCR *glnA* R		AGCTGCTGGTACTCTTATTAATGC	
qRT-PCR *nitR* F	*nitR*	AGTTATGCAACATTCCCCAGACC	123
qRT-PCR *nitR* R		ATAAGGCTCATTTACGATAGC	
qRT-PCR *gltA* F	*gltA*	TCGGAAACGAAGGGGATTACGC	125
qRT-PCR *gltA* R		TCCACCCAATTAGTTCAAGTCC	

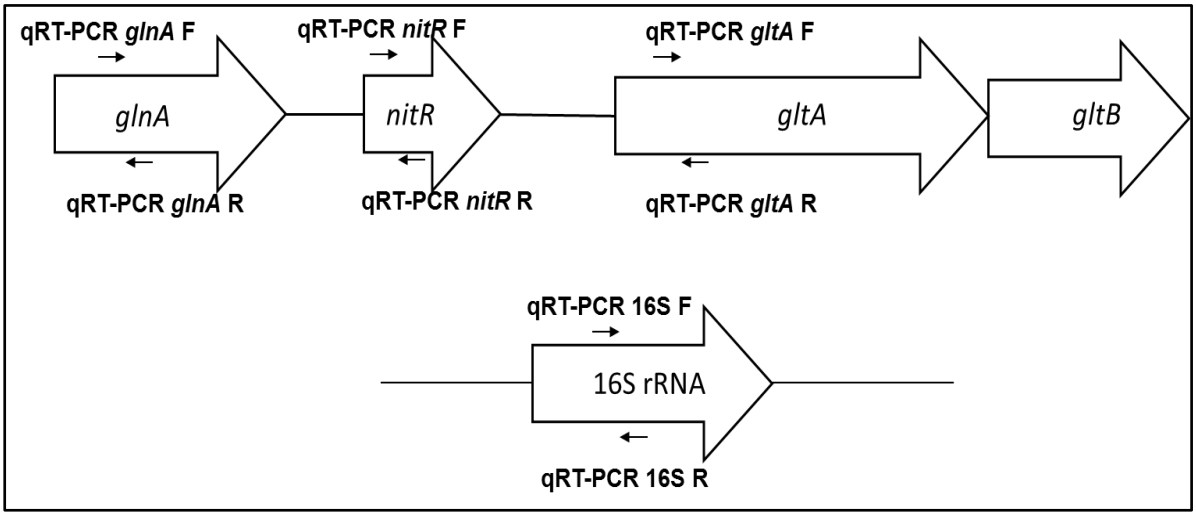


Figure 4.2: Schematic representation of the GS-GOGAT nitrogen gene cluster, 16s *rRNA* gene and the relative positions of the primer pairs used in qRT-PCR.

All qRT-PCR reactions were performed using the Celtic Quantace SensiMix SYBR No-ROX Kit (QT650-02, Celtic Molecular Diagnostics (Pty) Ltd.) at a total reaction volume of 25 µL which contained a final concentration of 1x SensiMix, 0.5 µM of forward and reverse primers, 2 µL of cDNA and made up to volume with nuclease-free water. qRT-PCR reactions were carried out using the Rotor-Gene 6000 (Corbett) and all runs were performed using a 96 well rotor. Cycling conditions consisted of an initial denaturation of 95 °C for 10 min and 40

cycles of 95 °C for 15 s, 52 °C for 20 s and 72 °C for 15 s. Analysis of data was performed using the Rotor-Gene 6000 Real Time Rotary Analyser Software (V1.7, Corbett). The design and implementation of the qRT-PCR experiment was done according to the relevant Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin *et al.*, 2009).

4.3.5 qRT-PCR experimental design and QIAgility automated robot sample preparation

Four qRT-PCR runs were performed for the 16s *rRNA*, *glnA*, *nitR* and *gltA* genes with each gene having its own standard curve, positive and no template controls per run. The controls for each run included an RNA no template control (to determine the baseline of SYBR green fluorescence), and a genomic DNA (gDNA) positive control (to ensure that the PCR was successful). Three biological samples of each set for both induced (pH 5.8) and uninduced (pH 6.5) conditions were tested, with each biological sample being analysed in technical duplicate. Each biological RNA sample was run with each primer pair in three separate qRT-PCR runs. The standard curves were created by pooling cDNA for each biological repeat for each condition and then establishing a dilution series from 10^{0}, $2x10^{-1}$, 10^{-1}, $2x10^{-2}$, 10^{-2}, $2x10^{-3}$, 10^{-3} to 10^{-4} using sterile, nuclease-free water in technical duplicate for each primer pair (Appendix C, Figure C1). Each run had its own standard curve for each primer pair, including the reference gene, to allow for differences in cDNA and PCR efficiency. All r^2 values above the 0.95 threshold and all efficiency values above 0.7 for each primer pair were used in the data analysis and the validation of the qRT-PCR experiment (Bustin *et al.*, 2009). The specificity of the primers for their gene targets was demonstrated through melt curves for standard curve and biological sample cDNA template (Appendix C, Figure C2).

The QIAgility automated sampler robot (Qiagen, Hilden, Germany) was used to ensure accurate pipetting of reagents and templates resulting in high levels of experimental technical reproducibility for all qRT-PCR experiments. The QIAgility software (QIAgility version 4.15.1, Qiagen) was used to create four runs which mirrored the experimental layout of the qRT-PCR runs. The automated device was programmed to create and dispense master mixes for the various reference and target gene primer pairs, serial dilutions of the pooled cDNA used in standard curves for said genes and to aliquot the cDNA and control template into master mix in the Rotor-Gene 6000 96 sample carousel tubes. Primers stocks, 10^0 pooled NCP260 cDNA, individual biological cDNA, gDNA, Sensimix, nuclease-free water and empty eppendorf tubes were provided to create and aliquot master mixes, serially diluted pooled cDNA and sample template using sterile 200 µL and 50 µL conductive filter tips (Qiagen). Samples were contained inside Rotor Disc tubes using Rotor Heat Sealing Film (Qiagen) and heat sealed at high temperature using a Rotor Disc Heat Sealer (Qiagen).

4.3.6 qRT-PCR data and statistical analysis

The Corbet Rotogene machine internal specialised software, Rotor-Gene 6000 Real Time Rotary Analyser Software (V1.7, Corbett), was used for the standard curves, melt curves and cycle threshold (CT) analyses. The Pfaffl method was then applied for statistical data analysis (Pfaffl, 2001). The relative cDNA concentration for both induced and uninduced conditions was determined for each sample and for each primer pair, utilising standard curves generated for the respective primers. Each of the technical repeats was used to obtain an average value for the biological run under each condition. CT values that had a discrepancy of more than four cycles were discarded as outliers. The 16s *rRNA* gene was used as the reference gene.

Statistical significance and reproducibility was assessed by comparing the efficiency and CT values for each primer pair under each condition. Any experimentally statistically significant divergent values were excluded from the data analysis (Bustin *et al.*, 2009). A mean value for each biological repeat was established for each condition for each gene. These values were then calibrated as per the Pfaffl equation using their primer and run-specific efficiency values. All values were then normalised against the calibrated 16s *rRNA* value. The values were averaged by using all three biological repeats in induced and uninduced conditions. The relative abundance of each gene during the induced state (pH 5.8) was calculated and compared to the uninduced (pH 6.5) value. The relative increases in these values were then evaluated for statistical significance (Pfaffl, 2001).

4.3.7 Bioinformatics of nitrogen metabolism genes

The genome of *C. beijerinckii* NCIMB 8052^T was screened for nitrogen metabolism genes. The nucleotide sequence of target genes, the genes flanking target genes and the intergenic regions between genes were collected from the NCBI database. The nucleotide sequences of the target genes were shown to be 100% conserved between NCP260 and NCIMB 8052^T (Personal correspondence, Green Biologics Ltd.) The ClosTron targeted mutagenesis algorithm, was used to generate six *C. beijerinckii* NCIMB 8052^T / NCP260 targeted gene insertion regions. The translated amino acids sequences of the target genes were queried by BLAST to confirm identity and putative function. The gene targets selected for ClosTron mutagenesis are shown below along with their putative gene functions and gene accession numbers (Table 4.2). These target genes encode two putative ammonia transporters, Cbei_2061 and Cbei_5034 as well as an amino acid transporter, Cbei_4173. Genes encoding

two putative PII-like nitrogen regulators Cbei_4942 and Cbei_5033 along with a histidine kinase, Cbei_4175, were also targeted.

Table 4.2: List of mutagenesis gene targets and plasmids used in this study.

Plasmid Name	Putative Gene Target	Gene Accession Number	Gene Size (bp)
pMLT007S-E2_cbei2061Amt	Ammonia transporter	Cbei_2061	1245
pMLT007S-E2_cbei5034Amt		Cbei_5034	1245
pMTL007S-E2_cbei4173GlnQ	Glutamine transporter	Cbei_4173	723
pMLT007S-E2_cbei4942PII	PII-like Nitrogen Regulator	Cbei_4942	348
pMLT007S-E2_cbei5033PII		Cbei_5033	339
pMLT007S-E2_ cbei4175HK	Histidine kinase	Cbei_4175	1263
pMTL007S-E2	Untargeted Control Plasmid	N/A	N/A

4.3.8 ClosTron insertion targets and plasmid design

The target gene regions were inputted into the ClosTron Intron Targeting and Design Tool at www.clostron.com (Heap *et al.*, 2009). The intron design tool is based on the Perutka algorithm that allows for efficient computation of the best nucleotide target sites for insertion by the Ll.LtrB- transposon (Perutka *et al.*, 2004). The intron targeted ClosTron plasmids were synthesised by DNA 2.0 Inc. (www.DNA2.0.com) using the pMTL007S-E2 plasmid series specific for *C. beijerinckii* (Heap *et al.*, 2010) with a targeted region of 309 bp designed to recognise the gene targets mentioned. The insertion sites were cloned and sequenced by DNA2.0 to confirm correct targeting. An example of one of the targeted pMTL007S-E2 plasmids and key components can be seen in Figure 4.3. The other five targeted plasmids were constructed in the same way with the exception that the targeted insert

region, as depicted in red in Figure 4.3, was specific for regions Cbei_2061, Cbei_5034 and Cbei_4942, Cbei_4173 and Cbei_4175 respectively (Table 4.2).

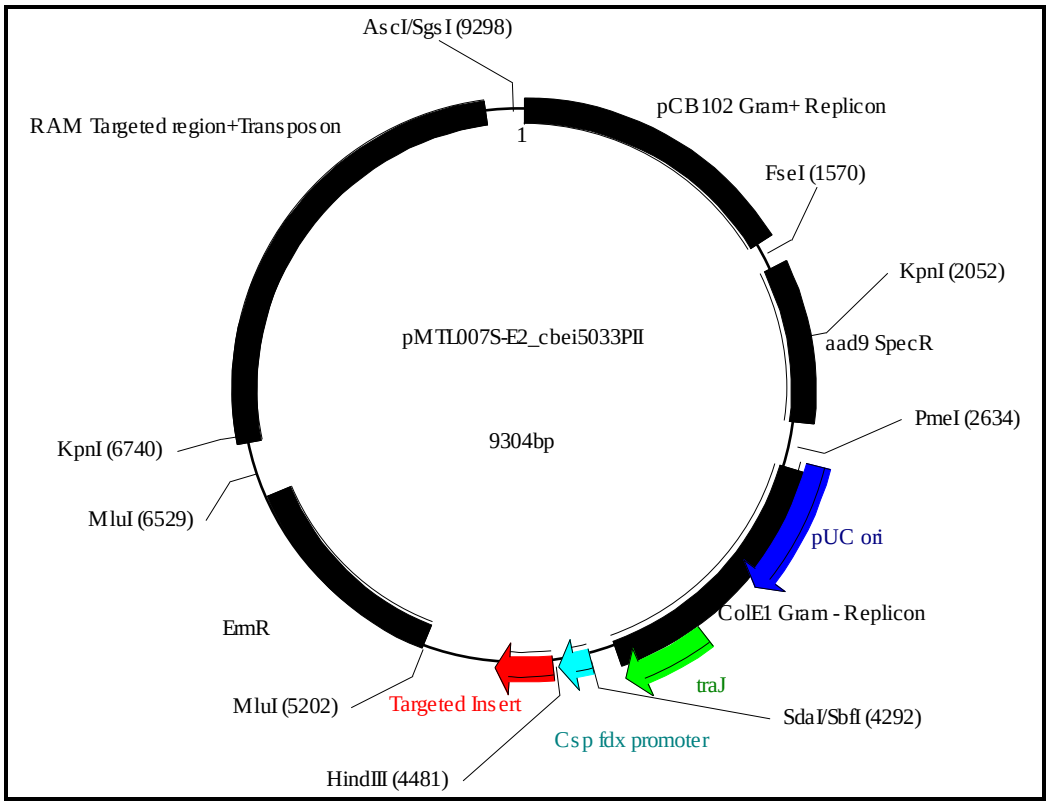


Figure 4.3: Vector map of the Clostron plasmid pMTL006S-E2_cbei5033PII. Key restriction enzyme sites are shown and key elements are indicated in colour.

The pMTL007S-E2 plasmid series carries an *aad9* gene which encodes a spectinomycin resistance plasmid selective marker. The plasmid contains an erythromycin activated Retrotransposition-Activated selectable Marker (*ermB*-RAM) which is responsible for conferring erythromycin resistance after the transposon integration event under erythromycin selective conditions. Plasmid replication inside *E. coli* donor cells is enabled by pUC origin of replication. The *traJ* element encodes a 5' UTR cis-acting RNA which promotes the plasmid transfer. A pCB102 replicon is present to facilitate replication of the plasmid once

inside the *C. beijerinckii* host. The Csp-fdx promoter from *Clostridium sporogenes* directs expression of the group II intron. The targeted insert region is tightly coupled to the *ermB*-RAM which integrates into the host *C. beijerinckii* chromosome under erythromycin selective conditions (Heap *et al.*, 2010).

4.3.9 Preparation of ClosTron donor *E. coli* electrocompetent cells

E. coli CA434 cells were grown in Luria Broth (LB) medium at 37 °C for 18h with agitation (Heap *et al.*, 2009). A 1% 18 h inoculum was transferred to fresh 50 mL LB and incubated at 37 °C with agitation until an OD_{600} of 0.6. Cultures were then placed on ice for 15 min. Cells were pelleted by centrifugation at 5000 *g* for 15 min at 4 °C. Cell pellets were washed with 10 mL ice cold distilled water and the cells collected as before. Cell pellets were resuspended in 2 mL 10% glycerol and centrifuged as before. The final cell pellet was resuspended in 300 µL 10% glycerol and stored as 40 µL aliquots at -80 °C.

4.3.10 Electrotransformation of ClosTron donor *E. coli* electrocompetent cells

Plasmid donor *E. coli* 40 µL electrocompetent cell aliquots were transferred to 2 mm electroporation cuvettes (BioRad), which were pre-frozen at -80 °C, and mixed with 50 ng of the corresponding ClosTron plasmid. The Micropulser (Biorad) was used to deliver a single, time-constant pulse at 2.5 kV, 200 Ω and 25 µF. An untargeted ClosTron plasmid control (Table 4.2) and a viability control using distilled water were also included for electroporation. Electroporation was conducted in triplicate for each construct with the appropriate controls. Electroporated cells were immediately diluted with 500 µL recovery LB media and kept on

ice for 1 min. The recovered cells were transferred to a 1.5 µL eppendorf tube and incubated at 37 °C with agitation for 1 h. The 500 µL cell suspension was then split in half with one volume spread plated onto Luria Agar (LA) plates containing 250 µg/mL spectinomycin (Sigma) and 10 µg/mL tetracycline (Sigma), and incubated at 37 °C for 18 h. Viability controls were performed on plates containing 10 µg/mL tetracycline only. Tetracycline is used to precent non-*E. coli* contamination, because this strain of *E. coli* is resistant to this anitibiotic and spectinomycin is selective for ClosTron plasmids. The other half of the electroporated *E. coli* volume was transferred to LB containing 250 µg/mL spectinomycin and 10 µg/mL tetracycline and incubated at 37 °C for 18 h with agitation.

Plasmids were extracted from *E. coli* using Miniprep Kit (Bioflux #BSC01S1, Bioer Technology) as per manufacturer's instructions. Successful plasmid uptake was confirmed by two restriction enzyme digests of the plasmid. One restriction enzyme digest used *Kpn*I and the other used *Asc*I and *Sda*I. The *Kpn*I digest was performed using 1X unique *Kpn*I buffer, 1 µg plasmid and 2 U *Kpn*I enzyme (Fermentas). The *Sda*I-*Asc*I digest was performed using 2X Tango buffer, 1 µg plasmid and 2 U *Sda*I and *Asc*I enzymes (Fermentas). The digests were performed at 37 ℃ for 2 h. Successful enzymatic digests were confirmed using 1% agarose gel electrophoresis at 80 V and visualised as in 4.3.2.

4.3.11 ClosTron plasmid conjugation of donor *E. coli* with recipient *C. beijerinckii* cells

Plasmid recipient *C. beijerinckii* NCP260 and NCIMB 8052^T conjugation cells were prepared in parallel with plasmid donor *E. coli* electrotransformed cells in an anaerobic cabinet (model 1024 Forma Scientific) with a gas composition of 5% H_2, 10% CO_2 and 85% N_2. Single

colonies of NCP260 and NCIMB 8052^T recipient cells were inoculated into 2% glucose RCM broth and incubated anaerobically at 37 °C for 18 h. After 18 h, electrotransformed *E. coli* cells were transferred to 1.5 mL eppendorf tubes and pelleted by centrifugation at 5000 *g* for 2 min. The pellet was washed twice in 1 mL 1x Phosphate Buffered Saline (PBS) and centrifuged as before. The pH of the PBS was adjusted to pH 7 for all experiments. The pellet was transferred to the anaerobic cabinet and resuspended in 200 µL of an 18 h *C. beijerinckii* cell culture.

Cell suspensions were plated as 25 µL spots onto RCM plates containing no selection and incubated agar-side-down anaerobically for conjugation to occur for 6-7 h at 37 °C. RCM agar plates containing conjugants were then flooded with 500 µL 1xPBS. Cells were scraped from the surface and 100 µL of the resulting cell slurry was spread plated in triplicate onto RCM agar containing 750 µg/mL spectinomycin and 250 µg/mL D-cycloserine (Sigma) and incubated at 37 °C until clostridial growth was visible. Clostridial colonies were transferred to fresh RCM agar containing 750 µg/mL spectinomycin and 250 µg/mL D-cycloserine. D-cycloserine was used to eliminate any undesirable background donor *E. coli* cells. Resultant *C. beijerinckii* cells which harboured the ClosTron plasmid were grown in RCM broth containing 750 µg/mL spectinomycin and 250 µg/mL D-cycloserine. The conjugant ClosTron plasmids were extracted and confirmed as for *E. coli* plasmids in 4.3.10. Single colonies that grew on 750 µg/mL spectinomycin and 250 µg/mL D-cycloserine RCM plates were then transferred to RCM plates which contained 250 µg/mL D-cycloserine and 10 µg/mL erythromycin. Integration of the RAM into the chromosome of *C. beijerinckii* cells would be indicated by the growth of mutant colonies on media which contained 10 µg/mL erythromycin.

4.3.12 Chromosomal integration of mutants and PCR screening

The nucleotide sequences for all ClosTron gene targets were conserved between NCP260 and NCIMB8052^{T} (NCP260 sequence data supplied by Green Biologics Ltd., personal correspondence). Primers were designed to these gene targets flanking the RAM insertion site (Table 4.3). Successful integration of the RAM to target region would be confirmed by PCR of the gene region. The PCR conditions consisted of 96 °C for 5 min, followed by 30 cycles of 96 °C for 2 min, 55 °C for 1 min, 72 °C for 2.5 min and a final 72 °C for 5 min. Wild type cells generate a normal PCR amplicon size whereas mutant cells generate an amplicon size of 1.8 kb larger than wild type which indicates RAM integration into the target gene.

Table 4.3: A list of PCR primers used in this study and their expected PCR product sizes.

Primer Name	ClosTron Gene Target and Accession Number	Nucleotide Sequence 5'-3'	Wild Type Product Size (kb)	Expected Mutant Product Size (kb)
2061Amt F	Ammonium Transporter Cbei_2061	ATGCAGGTAGTGCGCTCGGAGC	0.771	2.571
2061AmtR		ACATATTTAGTGCGACAGCCCC		
5034Amt F	Ammonium Transporter Cbei_5034	TTGCAATAACACCGGCGGC	0.842	2.642
5034Amt R		TCCAACTTCACCAGTCCTAGCG		
4942PII F	PII Nitrogen Regulator Cbei_4942	ACTTGTGGATTAACCAAAAGGG	0.987	2.787
4942PII R		TCAGGTAGATATTCGGACCAGC		
5033PII F	PII Nitrogen Regulator Cbei_5033	TGGGATGCGGACAACAACGTGG	0.212	2.012
5033PII R		AGCTGAATTTCCTCTCTCACC		
4173GlnQ F	Glutamine Transporter Cbei_4173	TTGAGATAAAGCAGGTTAAACAGG	0.564	2.564
4173GlnQ R		TACTTTCCCAGATGTCGGCTCC		
4175HK F	Histidine Kinase Cbei_4175	AATGACCGGGTTCTTAGAGG	0.459	2.459
4175HK R		TTCAACTTCTATACTTCCTCCC		
ClosInsert F	pMTL007S-E2 Insert Region	AGAAAGGAGGAAAAAGGCTATAGC	1.764	N/A
ClosInsert R		TGGTTTGCACCACCCTCTTCGG		
ErmRAM F	pMTL007S-E2 *ermB*-RAM region	TGATTGCCAAGCACGTCCCC	0.740	N/A
ErmRAM R		AGGCATTCTTGTTTAGGGTATCCC		

4.4 Results and Discussion

4.4.1 The role of amino acids during acidogenesis

As mentioned in Chapter Three, certain amino acids improved butanol yields independent of maximum growth conditions for the same amino acid in the same media. Specifically, when a low basal level of yeast extract (1 g/L) was incorporated into the media, additional supplementation of the amino acids glutamine, glutamate and histidine had a significant ($p<0.05$) effect on improving the butanol yield when these amino acids were the primary organic nitrogen sources in the medium. Another study on *C. beijerinckii* also concluded that media supplemented with additional glutamine and histidine improved growth and solvent production (Heluane *et al.*, 2011). This trend was observed for both *C. beijerinckii* (NCP260 and NCP271) and *C. saccharobutylicum* (NCP258) species. To that end, *C. beijerinckii* NCP260 was the focus of this study as this species should be amenable to genetic modification, and the whole genome sequence data of the closely related NCIMB 8052^T is available on NCBI .

4.4.1.1 The effect of amino acid supplementation on acid survival

The hypothesis is that glutamine, glutamate and histidine improve solvent production by improving the Acid Tolerance Response, thereby ensuring more mature cells survive acidogenesis to produce more solvents during solventogenesis. To that end, NCP260 was subjected to a series of acid survival experiments conducted in minimal medium which was supplemented with asparagine, proline, lysine, glutamine, glutamate and histidine. The results of the acid survival experiment performed on NCP260 are depicted below (Figure 4.4). Glutamate and histidine supplementation at pH 5 resulted in significant ($p<0.05$) acid survival of cells of 40% and 68% respectively relative to pH 6.5 conditions compared to 7%

survival in ammonium acetate supplemented medium under the same relative pH conditions. This is an increase in survival at pH 5 of 4- and 7-fold, respectively, relative to ammonium acetate supplemented medium.The media supplemented with other amino acids, including glutamine did not significantly improve the ATR at pH 5.

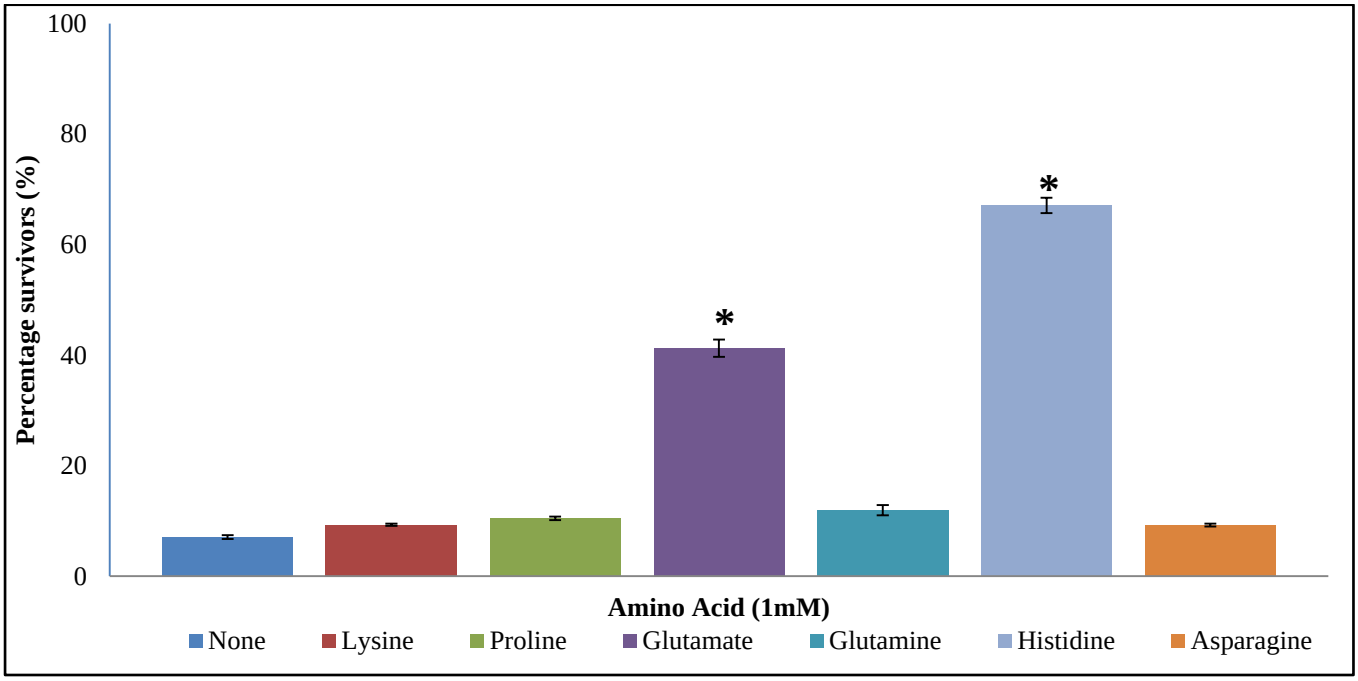


Figure 4.4: Percentage survival of NCP260 expressed as colony survivors of acid shock at pH 5 for 4 h relative to pH 6.5 for various amino acid supplementations. Error bars represent the standard error of the mean, where n=3.* indicates statistically significant % survival ($p<0.05$).

These observations are consistent with other ATR amino acid supplementation studies. Supplementing the growth media of *Lactobacillus casei* with histidine caused an accumulation of intracellular histidine resulting in significantly improved acid survival (Broadbent *et al.*, 2010). Glutamate has been shown to increase acid survival in *Bacillus cereus* (Hersh *et al.*, 1996) as well as *E. coli* (Kilimann *et al.*, 2005; Senouci-Rezkallah *et al.*, 2011).

The acid protective mechanism for histidine and glutamate in *C. beijerinckii* is unclear. A glutamate decarboxylase (GAD) system, encoded by *gadB*, was identified in *Lactococcus lactis* (Nomura *et al.*, 1999). In this system, an intracellular proton and glutamate is decarboxylated to produce gamma-aminobutyric acid (GABA) and CO_2. The decarboxylation process decreases intracellular acidity through the consumption of intracellular protons. Interestingly, the GAD system is linked to a glutamate-GABA anti-porter, encoded by the *gadC* gene in *L. lactis* whereby extracellular glutamate is transported into the cell and intracellular GABA is excreted (Sanders *et al.*, 1998). A GAD enzyme was also found in *Clostridium perfringens* (Cozzani *et al.*, 1970). However, bioinformatic analysis conducted in this study reveals there is no evidence of such GAD enzymes or a glutamate-GABA anti-porter present in the genomes of solventogenic clostridia. One mechanism which may be active in ATR in *C. beijerinckii* in the case of glutamate, is the reduction of intracellular protons in the form of acid during the conversion of glutamate and α-ketoacids to α-amino acid and α-ketoglutarate by transaminases (Figure 4.1). Glutamate is a core amino acid that can be converted to any amino acid. Glutamate in the form of D-glutamic acid is a key component of the Gram positive bacterial cell wall peptidoglycan (Bui *et al.*, 2012). Glutamate incorporated into cell wall peptidoglycan therefore contributes by strengthening the primary physical barrier between the intracellular and extracellular environment which may allow the organism to adapt to changes in the environment, including pH shifts experienced in the ATR (Lam *et al.*, 2009).

Glutamine was shown not to be involved in acid survival, yet it does increase solvent yield. Metabolism of pyruvate from glucose generates α-ketoglutarate (Cvitkovitch *et al.*, 1997). Glutamine, α-ketoglutarate and an intracellular proton are consumed in the GOGAT reaction to generate glutamate. It is possible that glutamine supplementation does confer acid

resistance indirectly through the conversion of glutamine to glutamate via GOGAT and then the resulting glutamate is incorporated into the transaminase reaction pathway. However, it may be that glutamine is not taken up efficiently under acid conditions.

A histidine moelcule can be decarboxylated by histidine decarboxylase to form histamine and CO_2. This decarboxylation reaction also requires the consumption of an intracellular proton. A Histidine decarboxylase, encoded by the *hdcA* gene, was shown to be linked to a histidine-histamine antiport system encoded by the *hdcC* gene in *Lactobacillus buchneri* (Martin *et al.*, 2005). A histidine decarboxylase enzyme has likewise been identified in *C. perfringens* (Van Poelje and Snell, 1990). Bioinformatic analysis of the genomes of solventogenic clostridia conducted in this study indicates that there is no sequence homology to *hdcA* or *hdcC* genes. It is possible for bacteria such as members of the *Bacillus* genus to degrade intracellular histidine into glutamate. This reaction is initiated by the histidase enzyme, typically encoded by the *hutH* gene: histidine → urocanate + NH_3. Urocanate is ultimately transformed into glutamate through multiple enzymes encoded by the *hutU*, *hutI* and either *hutF* or *hutG* genes (Bender, 2012). It has been shown that *C. saccharobutylicum* NCP262^T does not possess histidase activity to degrade histidine to glutamate (Usdin *et al.*, 1986). Given the closely related genetic nature between *C. beijerinckii* and *C. saccharobutylicum*, it seems likely that *C. beijerinckii* does not possess histidase activity either. Bioinformatic analysis of the NCIMB 8052^T genome conducted in this study revealed no histidine degrading genes. However, the mechanism of acid survival conferred by histidine may be explained when one considers the role of histidinol dehydrogenase. Histidinol dehydrogenase has not been reported as being involved in ATR in other bacteria, but it could play a role when one considers: L-histidine + 2 NADH + 2 H^+ ⇌ L-histidinol + 2 NAD^+. This reaction likewise utilises intracellular protons thereby raising intracellular pH. The histidine imidazole

functional group has a pK_a value of close to 6.0 and as such, histidine and histidine-containing peptides have been shown to have intracellular buffering capacity in eukaryotes (Abe, 2000).

4.4.1.2 The role of GS-GOGAT during acid shock

Glutamate and histidine were shown to be involved in the ATR. Because nothing is known of the genes involved in histidine degradation in solventogenic *Clostridium*, the role of the GS-GOGAT enzymes during acidic conditions was assessed using qRT-PCR in *C. beijerinckii* NCP260. The transcription levels of the *glnA*, *nitR* and *gltA* genes were monitored during acid shock conditions (Figure 4.5). It was found that prolonged exposure to pH 5.8 conditions stimulated a statistically significant two fold up-regulation in the *nitR* and *gltA* genes. Interestingly, the *glnA* gene expression level was found to be unchanged during pH 5.8 acid shock conditions compared to pH 6.5.

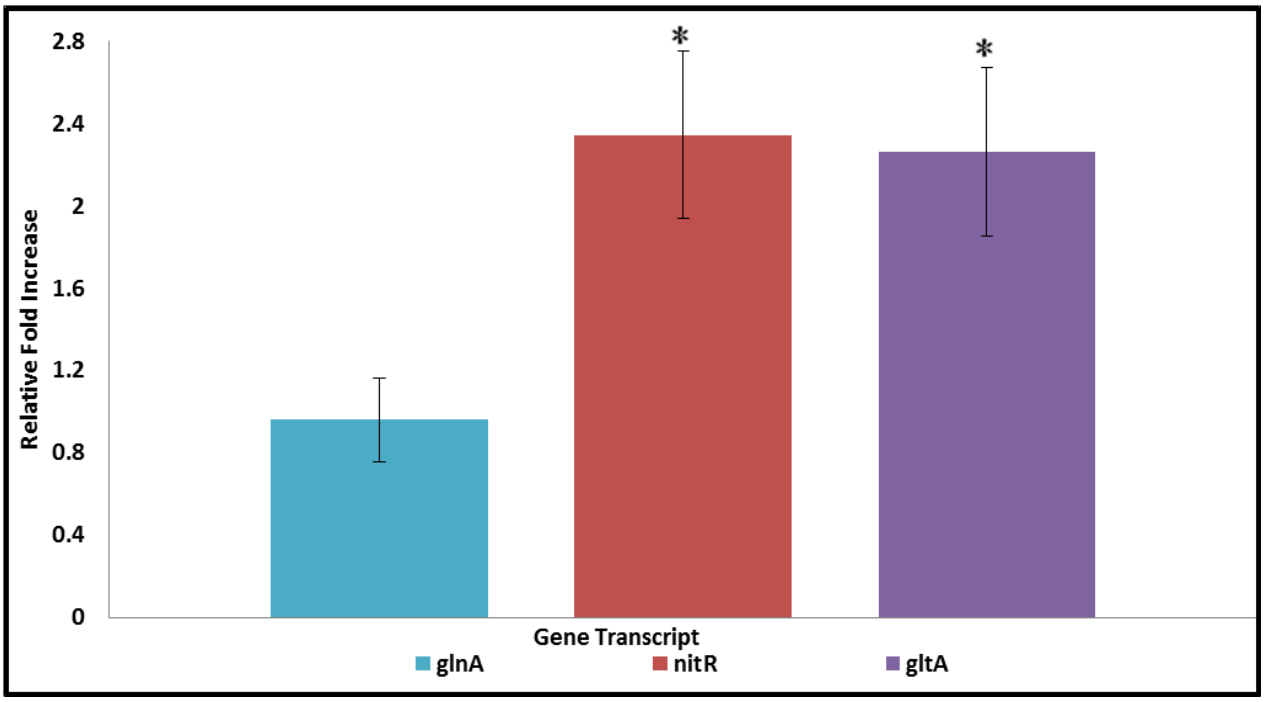


Figure 4.5: Relative fold increase in transcription of NCP260 GS-GOGAT operon genes *glnA*, *nitR* and *gltA* during pH 5.8 acid shock relative to pH 6.5. Error bars represent the standard error of the mean, where n= 9. * indicate statistically significant fold increase ($p<0.05$).

The up-regulation of only the *gltA* gene during acid stress is in agreement with the results from the acid shock assay: glutamate and not glutamine is required for acid shock survival (Figure 4.4). These findings correlate with a study conducted on *C. acetobutylicum* which monitored transcriptional expression in the switch from acidogenesis to solventogenesis whereby the *gltA* gene was up-regulated while *glnA* gene expression remained unchanged (Janssen *et al.*, 2010).

The GS-GOGAT system is the primary means of assimilating inorganic nitrogen in the form of ammonia in solventogenic *C. beijerinckii* and *C. saccharobutylicum* strains. During nitrogen limiting conditions, the GS-GOGAT operon displays bicistronic expression of the *glnA-nitR* genes and the *gltA-gltB* genes (Figure 4.6) (Stutz *et al.*, 2007). The nitrogen regulator NitR, is an ANTAR anti-terminator protein (Shu and Zhulin, 2002). NitR acts as a

positive transcriptional regulator, allowing the co-ordinated expression of *glnA-nitR* genes as well as *gltAB* genes by binding to transcription termination secondary structures, facilitating continuous read-through of DNA by RNA polymerase (Quixley, 1999). There is evidence that an antisense RNA (asRNA) of 43bp in length regulates the expression of the *glnA* gene by binding onto *glnA* mRNA transcripts (Fierro-Monti *et al.*, 1992). It was noted that *gltA* mRNA transcripts were in excess of asRNA under conditions which induced transcription of *glnA* which implies that only *glnA* mRNA is subject to asRNA repression under low nitrogen conditions (Stutz, 2000). During acidic conditions, expression of the *glnA* gene is unchanged despite higher levels of *nitR* expression. This implies that there is some form of post-transcriptional regulation during acid conditions that prevents transcription of *glnA* mRNA. Under low pH conditions, a possible explanation is that asRNA binds to *glnA* and represses translation of the GS enzyme product. This may be reflected in the qRT-PCR results if mRNA transcripts cannot be successfully converted to cDNA when asRNA is bound to the RNA.

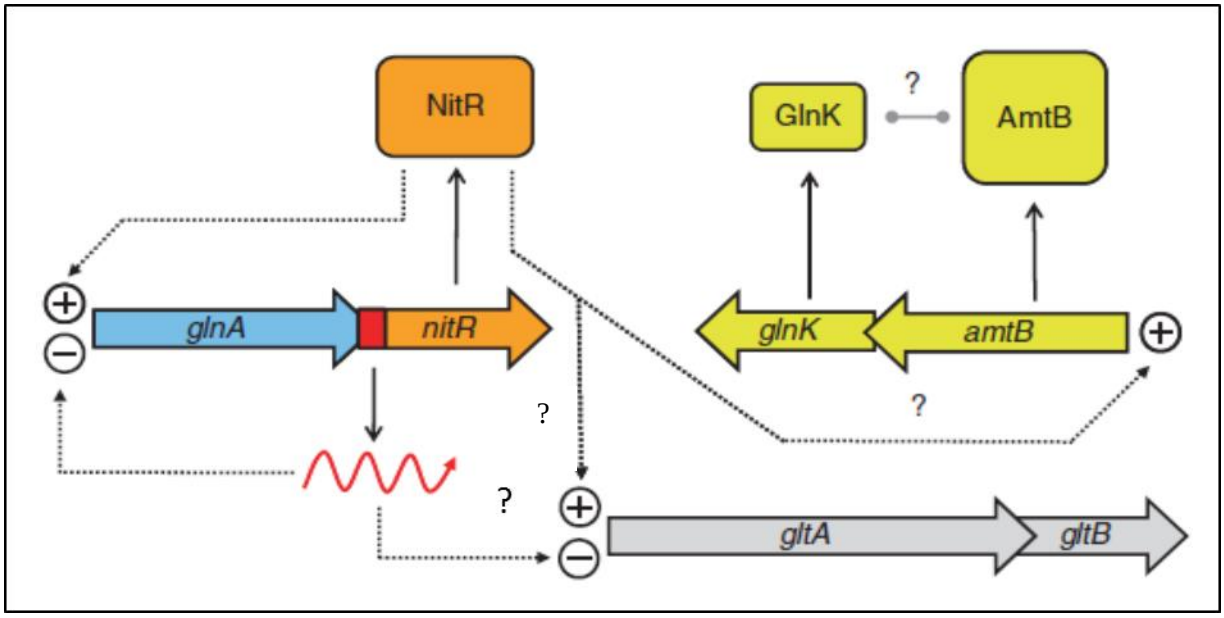


Figure 4.6: The regulatory system of nitrogen metabolism in *C. saccharobutylicum*. Question marks indicate putative interactions and the red coiled line indicates asRNA. The *glnK, amtB, glnA, nitR, gltA* and *gltB* genes encode a global regulator, ammonia transporter, glutamine synthetase, nitrogen regulator and subunits A and B of the glutamate synthetase (Figure from Amon *et al.*, 2010).

The need for *gltA* transcription during acid shock conditions could be due to the fact that accumulation of cytoplasmic pools of glutamate is advantageous in increasing intracellular pH through GOGAT activity and also transaminase reactions to form other amino acids. Glutamate appears central as a flexible amino acid for pH homeostasis and also acts as a precursor amino acid for biosynthesis of other amino acids which form proteins and enzymes required to mount an ATR. Identification of the pathways involved in histidine degradation and their role in the ATR, would further elucidate this phenomenon in solventogenic clostridia.

4.4.2 Mutagenesis of putative nitrogen metabolism genes

The GS-GOGAT system is integral to central amino acid biosynthesis in clostridial cells; therefore creating mutants in these genes would be highly deleterious to the organism. In addition, nothing is known about the global regulation of nitrogen assimilation in these bacteria, therefore other gene candidates which may play a role, but would not interfere with the GS-GOGAT enzymes directly were chosen as targets for KO mutations.

4.4.2.1 Bioinformatics of metabolism gene targets

The genome of *C. beijerinckii* NCIMB 8052^T was analysed using NCBI nucleotide BLAST for candidate nitrogen metabolism genes for use in mutagenesis studies. Two putative ammonia transporter genes, Cbei_ 2061 (Figure 4.7A) and Cbei_5034 (Figure 4.7B) were selected for mutagenesis, since ammonia is essential for solventogenesis and it is established that GS-GOGAT requires ammonia for assimilation into glutamate and glutamine. In addition, two putative PII-like nitrogen regulators, Cbei_5033 (Figure 4.7B) and Cbei_4942 (Figure 4.7C), were targeted in order to determine whether they are involved in global regulation of nitrogen metabolism. The PII protein, encoded by Cbei_5033, is particularly interesting as it lies adjacent to the ammonium transporter, Cbei_5034. A putative amino acid transporter, Cbei_4173 (Figure 4.7D), which shares homology to the *glnQ* glutamine transporter gene in the Gram-positive *S. mutans*, was targeted in order to determine how much of a role external glutamine has on normal cell function. The *glnQHMP* operon configuration in *S. mutans* closely resembles that seen in Cbei_4173-4170 (Chen *et al.*, 2010; Krastel *et al.*, 2010). Lastly, a putative histidine kinase which lies within the same gene cluster as the glutamine transporter Cbei_4175 (Figure 4.7D), was targeted to determine whether this kinase interacts with the ANTAR type NitR regulator of the GS-GOGAT

operon. To elucidate the roles of these putative nitrogen genes, the ClosTron system was used to inactivate these gene targets by homologous insertion.

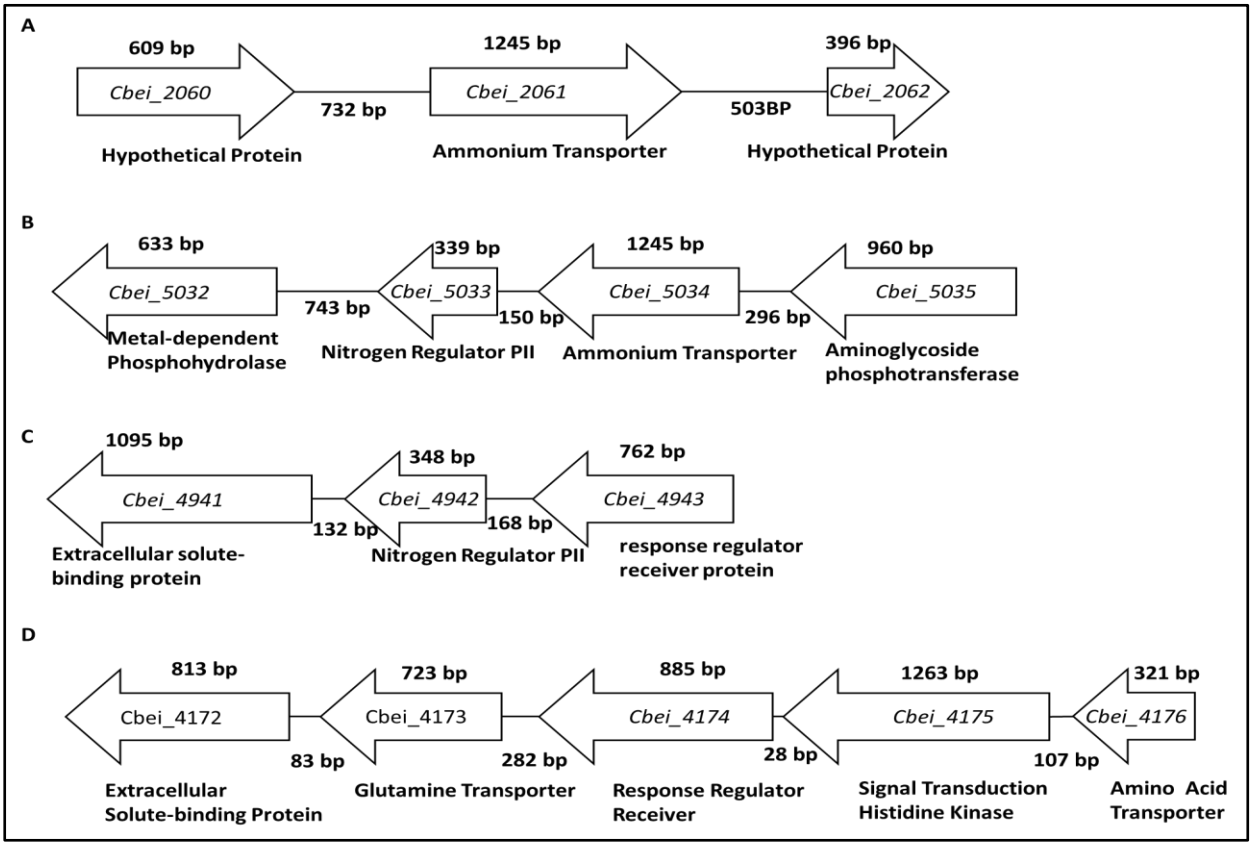


Figure 4.7: Schematic representation of the *C. beijerinckii* NCIMB8052^T gene clusters surrounding the putative genes encoding **A:** Cbei_2061 ammonium transporter, **B:** Cbei_5033 nitrogen regulator PII and Cbei_5034 ammonium transporter, **C:** Cbei_4942 nitrogen regulator PII and **D:** glutamine transporter Cbei_4173 and histidine kinase Cbei_4175.

A BLAST search of the translated *C. beijerinckii* NCIMB 8052^T nitrogen metabolism genes was conducted to confirm the annotated putative function of the genes. The translated ammonium transporters Cbei_2061 and Cbei_5034 belong to the ammonia transporter superfamily and contain conserved ammonia permease functional domains. BLAST analysis of Cbei_2061 and Cbei_5034 revealed an amino acid sequence identity of 91% to two

different *C. saccharobutylicum* ammonia transporter proteins respectively (Table 4.4). However, there have been no documented functional studies on this putative ammonia transporter homologue. The translated *C. beijerinckii* ammonium transporter genes (Cbei_2061 and Cbei_5034) displayed 99% and 83% identity respectively to that of *Clostridium pasteurianum*, which is supported by functional studies (Kleiner & Fitzke, 1994). Interestingly, the translated *C. beijerinckii* ammonia transporters (Cbei_2061 and Cbei_5034) share 46% and 45% identity respectively to that of *E. coli* (Jayakumar *et al.*, 1986). *Saccharomyces cerevisiae* ammonia transporter proteins displayed 37% and 39% sequence identity compared to the Cbei_2061 and Cbei_5034 proteins respectively (Marini *et al.*, 1994). The functional conservation of amino acid identity in these model organisms in a Gram-negative bacterium and eukaryotic yeast support the annotated function assigned to Cbei_2061 and Cbei_5034.

Table 4.4: BLAST results detailing the species and amino acid identities of the proteins encoded by candidate nitrogen metabolism genes.

Query Gene Accession Number	Protein Functional Super Family and Domains	Species Identity	Amino Acid Identity (%)	Reference
Cbei_2061	Ammonia transporter superfamily	*C. pasteurianum* NRRL B598	99%	Kleiner & Fitzke, 1994
		C. saccharobutylicum NCP262^{T}	91%	Annotation only
		E. coli ATCC 9637	46%	Jayakumar *et al.*, 1986
		S. cerevisiae S288c	37%	Marini *et al.*, 1994
Cbei_5034	Ammonia permease domains	*C. pasteurianum* NRRL B598	83%	Kleiner & Fitzke, 1994
		C. saccharobutylicum NCP262^{T}	91%	Annotation only
		E. coli ATCC 9637	45%	Jayakumar *et al.*, 1986
		S. cerevisiae S288c	39%	Marini *et al.*, 1994
Cbei_4173	ABC ATPase transporter superfamily Glutamine and histidine transporter domains	*C. saccharoperbutylacetonicum* N1-4	90%	Annotation only
		E. coli ATCC 9637	54%	Nohno *et al.*, 1986
		S. mutans GS-5	58%	Chen *et al.*, 2010
Cbei_4942	PII superfamily	*C. saccharoperbutylacetonicum* N1-4	91%	Annotation only
		E. coli ATCC 9637	40%	Adler *et al.*, 1975
Cbei_5033		*C. saccharoperbutylacetonicum* N1-4	88%	Annotation only
		E. coli ATCC 9637	46%	Adler *et al.*, 1975
Cbei_4175	Histidine kinase-like ATPase superfamily	*C. pasteurianum* NRRL B598	99%	Annotation Only
		S. mutans GS-5	7.8%	Senadheera *et al.*, 2009

The glutamine transporter, Cbei_4173 belongs to the ABC ATPase transporter superfamily and contains glutamine and histidine transporter domains (Table 4.4). Cbei_4173 shared translated sequence identity of 90% to the translated sequence of *Clostridium saccharoperbutylacetonicum* N1-4, but there are no functional studies in this organism to validate the annotation. Cbei_4173 displayed 54% identity to a known glutamine transporter in *E. coli* (Nohno *et al.*, 1986). The putative *C. beijerinckii* glutamine transporter showed 58% identity to that of *Streptococcus mutans*. The corresponding gene has been designated *glnQ* and occurs in an operon that mirrors that of Cbei_4172-4176 (Figure 4.7D) (Chen *et al.*, 2010). The fact that the Cbei_4172-4176 genes are arranged close to one another and have putative nitrogen metabolism functions is a good indicator that Cbei_4173 and Cbei_4175 are involved in nitrogen metabolism.

Cbei_4175 belongs to the histidine kinase-like ATPase superfamily (Table 4.4). This kinase shared 99% translated amino acid identity to *C. pasteurianum* NRRL B598 by annotation only. Interestingly, Cbei_4175 shared only 7.8% identity with that of *S. mutans*. The role of this kinase, VicK, in *S. mutans* has been shown to be linked to acid tolerance (Senadheera *et al.*, 2009). While both the kinase encoded by Cbei_4175 and VicK possess conserved functional sensor kinase domains, the amino acid sequence differences between species indicates that the function of the kinase in *Clostridium* may differ from that of *S. mutans*. However, the role of Cbei_4175 in *C. beijerinckii* is unclear. The kinase may be a promising candidate for an activator of the ANTAR type NitR regulator in the GS-GOGAT operon (Shu and Zhulin, 2002), although it may regulate the *glnQ* operon, particularly as it lies adjacent to a putativer response regulator, Cbei_4174. Since this histidine kinase gene is arranged in the same operon structure as that seen in *S. mutans*, it is likely to be involved in nitrogen regulation (Chen *et al.*, 2010).

The PII-like nitrogen regulator gene Cbei_4942, showed a translated sequence identity of 91% to *Clostridium saccharoperbutylacetonicum*. No research has been conducted into the PII proteins of solventogenic *Clostridium* and as such, the protein identities for all clostridial PII proteins are based on annotation only. Cbei_4942 had 40% amino acid identity to that of the PII protein in *E. coli*, which has been reported to have regulatory activity (Adler *et al.*, 1975). The other putative PII-like regulatory gene Cbei_5033 indicated an 88% translated sequence identity to *Clostridium saccharoperbutylacetonicum*. This protein had 46% identity to that of the PII protein in *E. coli* (Adler *et al.*, 1975).

4.4.2.2 Electroporation and confirmation of targeted ClosTron plasmids

The six targeted ClosTron plasmids were electroporated into donor *E. coli* CA434 as described in section 4.3.10. The presence of the spectinomycin resistant plasmids in *E. coli* cells was confirmed by restriction enzyme digest of the plasmids. The plasmids (Table 4.2) were digested with *Kpn*I which has two equidistant recognition sites on the pMLT plasmids (Figure 4.8), generating two plasmid fragments of 4616 bp and 4688 bp that appear as a single DNA band of approximately 4650 bp in the *Kpn*I digest gel image (Figure 4.8, lanes 3 and 6). The plasmids were also digested with *Sda*I and *Asc*I to generate two plasmid fragments of 4957 bp and 4347 bp (Figure 4.8, lanes 4 and 7). The same set of restriction digests were performed for electrotransformed *E. coli* pMTL007S-E2_4942PII and pMTL007S-E2_2061Amt, pMTL007S-E2_4173GlnQ and pMTL007S-E2_4175HK plasmids (data not shown).

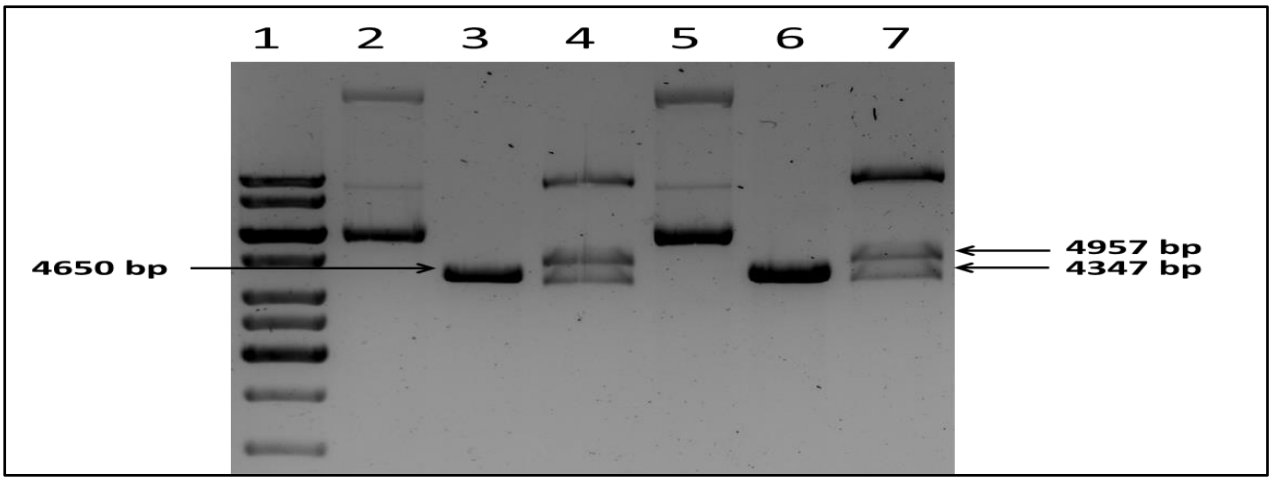


Figure 4.8: Restriction enzyme digests of spectinomycin resistant *E. coli* pMTL007S-E2_5033PII and pMTL007S-E2_5034Amt plasmids. Lane 1, molecular marker; lane 2, undigested pMTL007S-E2_5033PII; lane 3, pMTL007S-E2_5033PII digested with *Kpn*I; lane 4, pMTL007S-E2_5033PII digested with *Sda*I –*Asc*I; lane 5, undigested pMTL007S-E2_5034Amt; lane 6, pMTL007S-E2_5034Amt digested with *Kpn*I and lane 7, pMTL007S-E2_5034Amt digested with *Sda*I –*Asc*I.

4.4.2.3 Conjugation and confirmation of ClosTron plasmids

Once the the six ClosTron plasmids were confirmed in the *E. coli* donor cells by restriction digests, conjugation proceeded between *E. coli* and *C. beijerinckii* NCP260 or NCIMB 8052^{T} cells. Plasmid extraction and restriction enzyme digest of spectinomycin resistant NCP260 cells confirmed the pMTL007S-E2_5033PII, pMTL007S-E2_5034Amt, pMTL007S-E2_4942PII and pMTL007S-E2_2061Amt plasmids had been conjugated successfully into NCP260 (Figure 4.9A and B). The same restriction digests were performed for conjugant pMTL007S-E2_4173GlnQ and pMTL007S-E2_4175HK plasmids (data not shown). The identity of the spectinomycin resistant NCP260 cells was confirmed by the species-specific PCR mentioned in Chapter Two as well as 16s *rRNA* sequencing. NCIMB 8052^{T} ex-conjugants were confirmed to contain the relevant plasmids in a similar manner (data not shown).

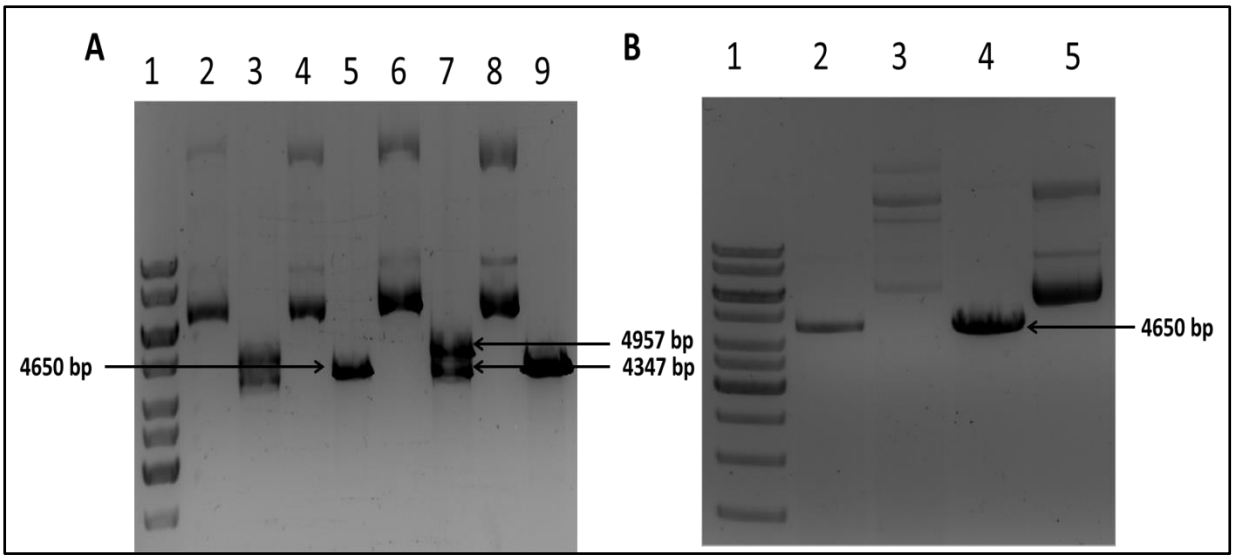


Figure 4.9: Restriction enzyme digests of spectinomycin resistant *C. beijerinckii* NCP260 ex-conjugants containing **A:** pMTL007S-E2_5033PII and pMTL007S-E2_5034Amt plasmids. Lane 1, molecular marker; lane 2, undigested pMTL007S-E2_5033PII; lane 3, pMTL007S-E2_5033PII digested with *Sda*I-*Asc*I; lane 4, undigested pMTL007S-E2_5033PII; lane 5, pMTL007S-E2_5033PII digested with *Kpn*I; lane 6, undigested pMTL007S-E2_5034Amt; lane 7, pMTL007S-E2_5034Amt digested with *Sda*I- *Asc*I; lane 8, undigested pMTL007S-E2_5034Amt and lane 9, pMTL007S-E2_5034Amt digested with *Kpn*I. **B:** pMTL007S-E2_4942PII and pMTL007S-E2_2061Amt plasmids. Lane 1, molecular marker; lane 2, pMTL007S-E2_4942PII digested with *Kpn*I; lane 3, undigested pMTL007S-E2_4942PII; lane 4, pMTL007S-E2_2061Amt digested with *Kpn*I and lane 5, undigested pMTL007S-E2_2061Amt.

4.4.2.4 ClosTron insertional inactivation of target genes in the chromosome of *C. beijerinckii*

Once the ClosTron plasmids were conjugated into the *C. beijerinckii* recipient cells, integration of the *ermB-RAM* (conferring erythromycin resistance) into the targeted region of the host *C. beijerinckii* chromosome is selected for by the presence of erythromycin in the media. Integration optimisation was attempted by using media which contained a range of 1, 2, 4, 6, 8, 25 and 50 µg/mL erythromycin in addition to media containing the recommended 10 µg/mL erythromycin. Fresh erythromycin antibiotic stocks were made specifically for each round of integration attempted. The screening volume was changed from three ex-

conjugant colonies transferred onto erythromycin media as described in the literature to all visible conjugation colonies (>50) transferred to erythromycin media. No cell growth on erythromycin selective media was observed, indicating that no stable integration was achieved for any of the constructs.

The insertion and the *ermB*-RAM regions for all targeted ClosTron plasmids were sequenced using primer pairs ClosInsert F- ClosInsert R and ErmRAM F- ErmRAM R, respectively (Table 4.3). The targeted ClosTron pMLT007S-E2 plasmids were confirmed to possess the correct targeting sequence and an *ermB*-RAM for each construct (data not shown). The lack of integration could therefore be due to the fact that inactivation of the chosen target gene is lethal to the host organism; however, it is unlikely that six different target genes would all be essential in this manner. Other explanations may be that the target sites are not specific enough for integration for the selected gene targets, or the recommended system for integration in *C. beijerinckii* is poorly optimised. The ammonia and glutamine transporter gene mutants should still be able to grow on enriched RCM which contains other organic nitrogen without having to depend on ammonia or glutamine transport to generate amino acids. The disruption of the PII nitrogen regulator genes and histidine kinase, however, may be deleterious to cell viability. Since six different target genes were selected for single-round insertion inactivation, it is unlikely that the inactivation of all six of the gene targets would be deleterious.

The target region specificity was determined by the log-odds scores generated by the Perutka algorithm utilised in generating the integration sites at www.clostron.com (Perutka *et al.*, 2004). Log-odds scores were 9-10 for all targets, which is regarded as normal for site-specific

integration. It is likely that the ClosTron system in *C. beijerinckii* requires further refinement. There are very few reports in the literature of successful knock-out integration in *C. beijerinckii*. To date, the only published system utilising group II mobile introns for mutagenesis in *C. beijerinckii* was conducted using the TargeTron system devised by Sigma Aldrich (Wang *et al.*, 2013). The insertion site determination and plasmid construction methods behind TargeTron are proprietary information and not publically available in the literature. The final erythromycin-induced RAM integration used in that particular study was achieved using 25 µg/mL and not the typical 10 µg/mL concentration as is recommended for ClosTron pMTL007S-E2 plasmids used in this study (Heap *et al.*, 2010). There is evidence that erythromycin is unsuitable in fermentative bacteria due to denaturation of the antibiotic under low pH conditions (Mermelstein *et al.*, 1994).

Some economic factors that limit the use of the pMTL007S-E2 plasmids in *C. beijerinckii* are the use of spectinomycin and D-cycloserine as selective markers for plasmid retention and the elimination of *E. coli* donor cell background respectively. Such high concentrations of antibiotic and the media requirements for attaining *C. beijerinckii* conjugants make this system costly. Other antibiotics or antibiotic derivatives that have a better efficacy should be used as an alternative so that lower concentrations of the antibiotic are used to achieve the same result in a more cost-effective manner.

4.4.3 A model for amino acid biosynthesis and acid tolerance

Supplementing the media with high concentrations of amino acids is an artificial situation. The importance of the intracellular concentrations of amino acids formed during various stages in clostridial growth is best understood when cells are required to synthesise their own

amino acids from inorganic ammonia as a sole nitrogen source. In a study conducted with *C. acetobutylicum*, researchers investigated the transition between the exponential acidogenic growth phase and the early-mid stationary solventogenic phase (Amador-Noguez *et al.*, 2011). They tracked intracellular metabolites by introducing radio-labelled glucose as the sole source of carbon for biosynthesis of metabolites. These metabolites include biosynthesised amino acids. Since the sole source of nitrogen incorporated into the defined minimal media was in the form of ammonia, all metabolites detected were derived through the biosynthesis of amino acids from glucose and ammonia.

Table 4.5 summarises the molecular concentrations of the metabolites synthesised in that study. The intracellular accumulation of these metabolites wase tracked over time and showed that NADPH- and ATP-dependent biosynthesis of amino acids occurs predominantly during acidogenesis when glucose was being metabolised. Unsurprisingly, metabolites produced in glycolysis of glucose such as acetyl-CoA, butyryl-CoA, ATP, NADH and NADPH were elevated during active growth of cells in acidogenesis and declined in solventogenesis. Lysine concentrations were consistent throughout all phases of growth and solvent production. Proline accumulated during exponential phase, but was scarce during stationary phase. It is interesting to note that both proline and lysine were found to be exported from the cell and accumulated in the extracellular environment, which implies that these amino acids are potentially used as precursors for the synthesis of other amino acids or maintaining redox potentials and are not actively incorporated into cellular components. Asparagine biosynthesis increased from acidogenesis to solventogenesis. Interestingly, aspartate, a derivative of asparagine is present in the cell in large quantities during active growth and then decreases during solventogenesis. To note, histidine levels were consistently low during both phases. Intracellular glutamate and glutamine were found in the highest

concentrations during initial growth and decreased to lower concentrations during solventogenesis, but were still present at high concentrations compared to other metabolites. These high levels of intracellular glutamate have also been noted in metabolomics studies in *E. coli* (Bennet *et al.*, 2009).

Table 4.5: Intracellular metabolite concentrations during acidogenesis and solventogenesis of *C. acetobutylicum* with glucose as the sole carbon source and ammonia as the sole nitrogen source (adapted from Amador-Noguez *et al.*, 2011).

Metabolite	Concentration (mM)	
	Acidogenesis	Solventogenesis
Acetyl-CoA	1.6736	0.4575
Asparagine	0.0353	0.4925
Aspartate	7.0482	1.1698
ATP	3.2656	0.9530
Butyryl-CoA	1.8968	0.6180
Glutamate	26.6743	3.2572
Glutamine	8.7740	2.0087
Histidine	0.0802	0.0427
Lysine	1.3066	1.0608
NAD^+	1.6046	1.2327
NADH	0.3393	0.1228
$NADP^+$	0.0042	0.0024
NADPH	0.2415	0.0177
Proline	1.2303	0.1908

The abundance of certain amino acids in the solventogenic clostridial cells can be explained by the ATP and cell resources required in producing these amino acids. The energetic cost of the biosynthesis of amino acids varies between amino acids (Table 4.6). The energy cost is primarily measured in ATP, the energy currency of the cell, but also in NADH and NADPH which are vital to maintaining redox potentials for metabolic reactions. The ATP cost of biosynthesis of an amino acid is related to the ATP cost of acquiring the starting metabolite and then the number of enzymatic steps involved has a proportional relationship to ATP cost (Craig and Weber, 1998).

Table 4.6: The bioenergetic costs (in ATP) for the biosynthesis of amino acids in *E. coli.* Adapted from Craig and Weber, (1998).

Amino Acid	Metabolite Required	Metabolite ATP Cost	NADH Cost	NADPH Cost	Total Biosynthesis ATP Cost	Number of Enzymatic Steps
Asn	Oxaloacetate	3	0	1	4	1
Lys	Oxaloacetate, pyruvate	2	0	4	18.5	10
Pro	α-ketogluterate	1	0	3	12.5	4
Glu	α-ketogluterate	0	0	1	8.5	1
Gln	α-ketogluterate	1	0	1	9.5	2
His	Ribose-5-phosphate	6	-3	1	33	7

Only three ATPs are required to gain access to oxaloacetate as a metabolic template and only one enzymatic reaction to transform the metabolite oxaloacetate into asparagine with no enzymatic ATP costs required. Lysine and proline are quite energy inefficient amino acids to produce since they require 18.5 and 12.5 ATPs and cost four and three molecules of NADPH respectively for their biosynthesis. Glutamate and glutamine have intermediate energy costs of 8.5 and 9.5 molecules of ATP respectively per amino acid molecule synthesised. Interestingly, these molecules only consume one molecule of NADPH per amino acid synthesised. Glutamine and glutamate are needed in high levels throughout cell growth and are inexpensive to produce. However, further media supplementation of glutamate in particular decreases the energy consumed in producing large amounts of intracellular glutamate that are central to amino acid biosynthesis of other amino acids and are involved in regulating intracellular pH.

Histidine is an unusual amino acid since it costs 33 ATPs for its biosynthesis. Despite its expensive energy costs, it only consumes one molecule of NADPH and it generates 3 molecules of NADH during its biosynthesis. The high energy cost of histidine (Table 4.6) may explain why this amino acid was not found in high concentrations during acidogenesis and solventogenesis under minimal media conditions (Table 4.5). Histidine can act as a high-energy carrier molecule that can be broken down to generate ATP or used to regulate pH. The expensive bioenergetic cost for biosynthesis of histidine can be bypassed through histidine supplementation in fermentation media.

To summarise, the biosynthesis of amino acids is tightly regulated to balance the ATP and reducing NADH/NADPH costs of production with the costs of converting metabolic acids to solvents. The biosynthesis of amino acids like proline and lysine is expensive both in terms of energy cost and in terms of redox equivalents. Supplementing media with asparagine, proline and lysine does not appear to increase the ATR overall. This implies that despite bypassing the ATP cost, these amino acids are not biologically flexible and not crucial for the acid response. Histidine is energetically expensive to synthesise, however it has a low NADPH cost and contributes to reducing NADH. These observations offer some interesting insights as to why certain amino acids may be favoured in the cellular pools during fermentation. ABE fermentation requires reduced NADH for multiple steps in the metabolism of pyruvate to final ABE solvent products. Supplementing media with histidine circumvents the expensive biosynthesis ATP costs, so that ATP can then be used in cell metabolism leading to butanol formation, however this comes at the cost of losing NADH produced during histidine biosynthesis. Asparagine, glutamate and glutamine are cheaper to synthesise both energetically and in terms of redox equivalents. However, only glutamate increased the ATR. Cellular pools of glutamine are assumed to be converted into glutamate

by GOGAT activity. Glutamate is ultimately used directly or indirectly in the ATR as glutamate itself or in the form of the necessary amino acids formed through the conversion of glutamate into α-amino acids through transaminase reactions.

4.5 Conclusions

The ClosTron pMTL007S-E2 plasmid shuttle conjugation transfer system has been demonstrated in *C. beijerinckii* NCIMB 8052^{T} as well as NCP260. In this study, six different target sequences failed to yield stable gene inactivated integrants. The host chromosomal integration procedure for the *C. beijerinckii* requires further optimisation and a standard published method for the species containing detailed protocols is needed. Acid shock studies showed that glutamate and histidine enhance the survival of *C. beijerinckii* NCP260 in low pH, but glutamine does not. The low pH conditions experienced by solventogenic *Clostridium* during acidogenesis induces the expression of the *gltA* and *nitR* genes encoding GOGAT and the NitR regulator respectively, while expression of *glnA* encoding GS remains constant. Increasing the cellular pools of glutamate by amino acid supplementation increases the ATR in *C. beijerinckii*. The improvement in ATR could be due to the GOGAT-mediated conversion of glutamine which consumes protons to form glutamate, and in addition, glutamate participates in transaminase reactions to form α-amino acids which increase intracellular pH. High concentrations of intracellular histidine are thought to increase the intracellular pH through the production of histidinol which consumes protons as a means of mounting an ATR. Supplying histidine in media would also decrease the energetic burden of the cell and energy that would be used for costly histidine biosynthesis is focused on ATR and butanol production instead. Although glutamine was not shown to be directly involved in the ATR, it is also required for the activity of GOGAT and therefore would be required for generating an ATR during acidic conditions; however, the intracellular pools of glutamine have been shown to be high during acidogenesis (Amador-Noguez *et al.*, 2011) and may be sufficient to mount an ATR during acid conditions. Supplementing fermentation media with glutamate and histidine improves the ATR which allows more cells to survive acidogenesis and enter solventogenesis, ultimately improving solvent yields.

Chapter Five
General Conclusions

Contents

5.1 Aim and hypothesis

The aim of this thesis was to characterise the UCT strain collection in terms of taxonomy and optimal substrates and to investigate the nitrogen metabolism system in ABE producing *Clostridium beijerinckii* and *Clostridium saccharobutylicum*. The taxonomic characterisation included species- and strain-level typing. The nitrogen metabolism characterisation included bioinformatic assessment of potential nitrogen metabolism genes, growth and fermentation studies under various nitrogen conditions and finally, the role of amino acids in mounting an ATR to survive acidogenesis to produce ABE solvents.

The findings of this study have generated the following hypothesis: glutamate, glutamine and histidine present in the fermentation media of solventogenic *C. beijerinckii* and *C. saccharobutylicum* may be used in the GOGAT and histidol dehydrogenase reactions to decrease intracellular proton concentration which increases acid survival and solvent production.

5.2 Species taxonomy of the NCP strain collection

All solventogenic *Clostridium* strains were originally classified as *C. acetobutylicum* based on morphology. Later studies showed that this group consisted of at least four species, *C. acetobutylicum*, *C. beijerinckii*, *C. saccharobutylicum* and *C. saccharoperbutylacetonicum* (Johnson *et al.*, 1997; Keis *et al.*, 2001). The UCT NCP strain collection required species level characterisation to catalogue the collection. A species-specific PCR was developed which utilises the unique *glnA-nitR-gltAB* gene arrangement (Stutz *et al.*, 2007), and this was optimised as a rapid, colony-PCR screen. The species-specific PCR, rifampicin screening and

16s *rRNA* sequencing revealed that the NCP strain collection consisted of 19 *Clostridium beijerinckii* and 11 *Clostridium saccharobutylicum* strains.

5.3 Strain level characterisation of the NCP strain collection

Research into strain typing of solventogenic *Clostridium* is limited. Early strain level studies used PFGE to generate genomic DNA fingerprints of the species within the major solventogenic groups. Previous analysis of the NCP strain collection indicated one strain type for *C. beijerinckii* and two strain types for *C. saccharobutylicum* (Keis *et al.*, 2001). There are no records of the origins of particular strains or of when each strain was used during the NCP facility's operation. Modern strain typing methods were applied to determine the genetic relationship between the strains of the two NCP species (Alonso *et al.*, 2001; Macdonald *et al.*, 2011).

RAPD analysis generated two strain groups for *C. saccharobutylicum* and four groups for *C. beijerinckii*. MLST analysis revealed a subset of three *C. saccharobutylicum* strain groups from strains within the two RAPD strain groups, and indicated that the *C. beijerinckii* RAPD strain group 2 could be further subdivided into two strain groups. With genome sequencing becoming more accessible to research laboratories, further consolidation of these strain groups will become evident

These findings demonstrate that the NCP strains are closely related to each other and may have originated from a similar geographic location or evolved from a selection of common strain ancestors over many successive batch fermentations.

5.4 Sugar substrate fermentation performance

ABE solvent performance is primarily indicated by the ability to convert carbon in the form of sugars into solvents (Jones and Woods, 1986). *Clostridium* strains are able to ferment a wide variety of sugar rich substrates (Shaheen *et al.*, 2000; Wang and Blaschek, 2011) and lignocellulosic plant material (Yu *et al.*, 2012; Lu *et al.*, 2013a). The solvent producing capabilities of the UCT NCP collection were assessed in glucose-, sucrose- and xylose-based media.

Strains NCP J, NCP249, NCP172 and NCP199 ferment glucose and generate high butanol yields of 0.219, 0.218, 0.214 and 0.214 respectively. Strains NCP254, NCP259, NCP268 and NCP199 utilise sucrose well in fermentations, resulting in butanol yields of 0.216, 0.213, 0.227 and 0.212 respectively. Strains NCP195, NCP259, NCP172 and NCP265 metabolise xylose to produce butanol yields of 0.214, 0.195, 0.177 and 0.177 respectively. Strains NCP195, NCP265, NCP271, NCP J, NCP259 and NCP172 are able to ferment all three substrates to produce high titres and yields of butanol. Solvent analysis of the NCP strain collection indicates a broad diversity in substrate preference with many strains able to produce high yields of butanol from the different substrates under laboratory conditions. These strains can be used in the existing ABE fermentation industry.

5.5 The effect of nitrogen sources on growth

Nitrogen metabolism and the assimilation of nitrogen from the environment are of paramount importance to all organisms. Gram-positive bacteria have evolved an efficient network of nitrogen metabolism genes, sensor proteins and regulators to ensure optimal growth conditions are maintained even in times of nutrient limitation (Amon *et al.*, 2010). The GS-GOGAT nitrogen assimilation system in *C. saccharobutylicum* converts extracellular ammonia into the amino acids, glutamine and glutamate. These amino acids can be converted into other α-amino acids by transaminases in order to maintain the intracellular amino acid pools required for optimal growth (Stutz *et al.*, 2007). Inorganic nitrogen in the form of ammonium salts have been shown to improve the growth and solvent production of *Clostridium* (Long *et al.*, 1984). Organic nitrogen has likewise been shown to be of vital importance to maintaining optimal growth (Stutz *et al.*, 2007). A balanced combination of both organic and inorganic nitrogen has been shown to improve butanol yields (Abd-Alla and Elsadek El-Enany, 2012).

Strains grown in media supplemented with ammonium salts experienced an increased lag-phase and decreased maximum growth compared to media supplemented with organic nitrogen. Media supplemented with 2% casamino acids supported optimal growth compared to 0.02% casamino acid-supplemented media.

5.6 The effect of amino acids on solvent production

DNA microarray experiments that monitored the transcriptional events of *C. beijerinckii* during solvent stress identified genes involved in glutamine and histidine biosynthesis as being up-regulated. A lysine permease was also shown to be upregulated during solvent stress (Heluane *et al.*, 2011). A microarray study on *C. acetobutylicum* cells exposed to a butanol pulse revealed that the genes involved in glutamine transport, glutamine synthesis, glutamate synthesis and proline transport were up-regulated in response to solvent stress (Janssen *et al.*, 2012). However, the role of individual amino acids during ABE fermentation has not been investigated.

A selection of NCP strains were selected for growth and fermentation studies in NCP media modified to contain a reduced level of YE and supplemented with asparagine, glutamate, glutamine, lysine or proline. Amino acid growth studies indicated that strain growth reached a similar maximum growth value and the growth rates in the presence of the different amino acids supplementations did not differ substantially. It was established that carbonate buffered fermentation media generated much higher solvent titres than phosphate buffered media. Solvent analysis in calcium carbonate buffered media supplemented with the various amino acids indicated that glutamate, glutamine and histidine significantly ($p < 0.05$) increased the butanol titres of NCP258, NCP271 and NCP260 compared to the un-supplemented controls.

5.7 Effect of amino acids on acid survival

The ability to mount various defence mechanisms against the destructive effects of low intracellular pH is known as the Acid Tolerance Response (ATR) (Davis *et al.*, 1996). This defence response primes bacterial cells at a sub-lethal pH which increases the organism's tolerance to subsequent lethal pH exposure (Cotter *et al.*, 2000). Gram-positive bacteria employ numerous acid response mechanisms in response to acid stress (Cotter and Hill, 2003). Amino acids have been shown to improve the ATR of bacteria (Broadbent *et al.*, 2010; Lu *et al.*, 2013b).

Acid survival experiments were conducted in media supplemented with asparagine, glutamate, glutamine, lysine or proline as the sole nitrogen source. Only glutamate and histidine were shown to significantly ($p < 0.05$) improve cell survival at low pH. qRT-PCR experiments monitored the expression of the important nitrogen assimilation genes, *glnA-nitR-gltA*, during acid shock. The *nitR* and *gltA* genes were significantly ($p < 0.05$) up-regulated during acidic conditions. Glutamate and the GOGAT enzyme, responsible for glutamate production, are likely to be involved in the ATR during acidogenesis. The molecular nature of the role played by histidine in the ATR has not yet been identified, however, the gene encoding histidinol dehydrogenase has been shown to be up-regulated during the transition between acidogenesis and solventogenesis (Janssen *et al.*, 2010), so the transcription of this gene could be quantitated during acid shock conditions to elucidate its role in ATR.

5.8 A model for acid survival and solvent production

Quantitative RT-PCR experiments established that the *gltA* gene, encoding GOGAT, is up-regulated during acid response, however, the expression of the *glnA* gene was not increased at the same time. Although this was not tested in our acid shock study, histidinol dehydrogenase has been shown by Janssen and colleagues (2010) to be up-regulated in during the transition between acidogenesis and solventogenesis. A theoretical model for nitrogen metabolism and acid stress in *Clostridium* (Figure 5.1) is presented, based on experiments conducted in this study as well as evidence in the literature. The metabolic pathways that provide the precursor molecules necessary for biosynthesis of 16 amino acids are indicated (Amador-Noguez *et al.*, 2011). The total ATP energy and NADPH/NADH reducing costs required for their biosynthesis are shown (Craig and Weber, 1998). The amino acids involved in improving butanol yields are indicated in blue, sources of acid are indicated in red and extracellular compounds are indicated.

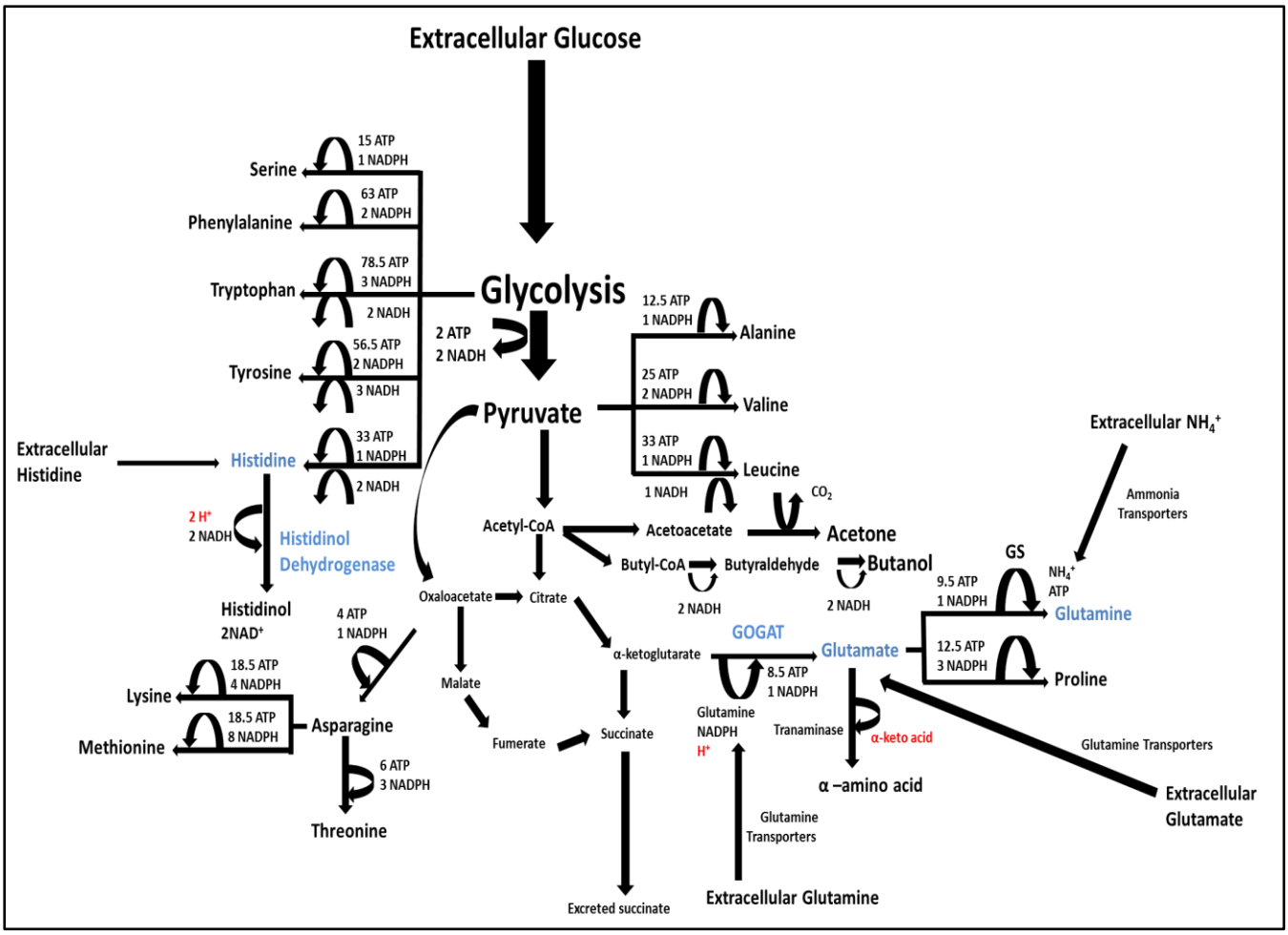


Figure 5.1: A hypothetical model of the amino acid biosynthesis, ATR and solvent production in *C. beijerinckii* and *C. saccharobutylicum*. ATP and NADH/NADPH amino acid biosynthesis costs are indicated. Amino acids and enzymes involved in the ATR are indicated in blue and sources of intracellular protons are indicated in red. The amino acid biosynthesis pathways are from Amador-Noguez *et al.* (2011) and bioenergetics costs of amino acids from Craig and Weber (1998).

Histidine can either be transported into the cell or biosynthesised. The histidine biosynthetic and metabolic pathways in solventogenic *Clostridium* have not yet been characterised. The biosynthesis of histidine requires the precursor molecule, ribose 5-phosphate, which is derived from the glycolysis pathway. Overall, histidine costs a net 33 ATPs and one NADPH molecule to synthesise, with two molecules of NADH being produced in the process. During acid stress conditions, an ATR would be mounted when histidine is converted to histidinol by histidinol dehydrogenase which consumes two molecules of NADH and two intracellular

protons, thereby increasing intracellular pH. One molecule of glutamine costs net 9.5 ATP and one NADPH and requires extracellular ammonia and glutamate as precursors in the glutamine synthetase (GS) reaction. The biosynthesis of glutamate requires α-ketoglutarate and glutamine as precursors and costs 8.5 ATP and one NADPH molecule per molecule of glutamate synthesised. In the process of glutamate formation, glutamate synthase (GOGAT) consumes a molecule of NADPH and an intracellular proton, thereby increasing intracellular pH. Glutamate can be modified by transaminase enzymes to produce α-amino acids which consume acidic α-keto acids, further increasing intracellular pH as a form of ATR. The glutamine required in the GOGAT reaction may be derived from the extracellular environment via the glutamine transporter rather than being synthesised in the GS reaction, which may explain why *glnA* transcription is not up-regulated during acid shock. In this study, both glutamine and glutamate have been shown to increase solvent production, presumably because they are both required for the action of GOGAT which contributes to the ATR as outlined above.

5.9 Future work

The RAPD primers used were designed for *C. difficile* as this technique has not been used in the solventogenic *Clostridium* strains (Alonso *et al.*, 2001). Screening of the genome of solventogenic *Clostridium* for novel RAPD primer sites and testing a selection of RAPD primers may yield better strain typing patterns. Similarly, the MLST primers used in this study were originally designed to regions within traditional housekeeping genes of *C. botulinum* (Macdonald *et al.*, 2011). Better strain resolution may be accomplished if the targeted MLST genes contain more SNPs and the PCR product sizes of the genes are

increased to give more specificity. Using solvent genes, which are under fermentation selective pressure, as MLST targets may show interesting strain differences.

In order to validate the proposed mechanisms in the model, a number of experiments should be performed. The ClosTron plasmid integration efficiencies inside the *C. beijerinckii* hosts must be optimised and higher efficiency plasmids used in order to generate stable integrant mutants for functional studies. KO mutants of candidate nitrogen genes, particularly the glutamine transporter gene, should be generated and functionality experiments for the gene targets performed. The role of histidinol dehydrogenase in the ATR should be investigated by quantitative-RT-PCR experiments following acid shock and gene knock-out experiments can be performed to determine if this enzyme is involved in the ATR. Furthermore, acid shock experiments with either glutamine or glutamate as the sole nitrogen source should be performed in conjunction with qRT-PCR and enzymatic activity assay experiments to determine whether GS is indeed involved in the ATR.

5.10 Conclusion

The NCP strain collection consists of genetically similar strains able to ferment a diverse array of substrates to produce high yields of biobutanol. The amino acids glutamine, glutamate and histidine are vital to maintaining high butanol yields. The glutamate synthase enzyme appears to be central to maintaining intracellular pools of glutamate during acid stress conditions. Glutamate, glutamine and histidine may all contribute towards the ATR of solventogenic *Clostridium,* which would increase the viability of cells during acidogenesis and, hence, solvent production.

Appendix A

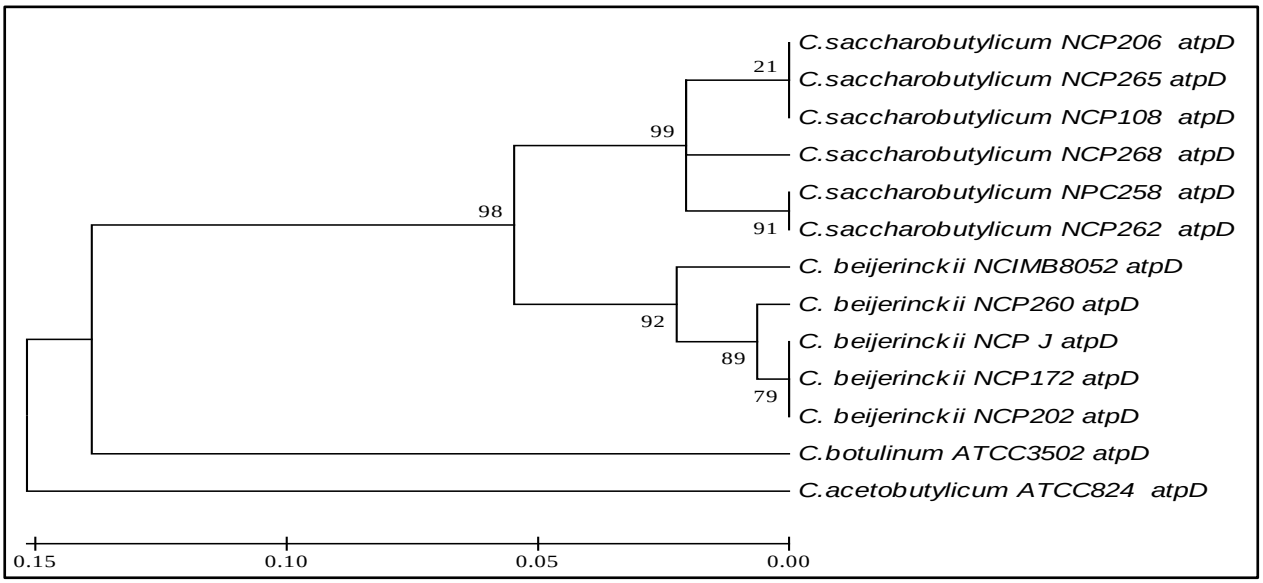


Figure A1: *atpD* gene phylogenetic tree based on 179 bp of amplified sequence for NCP strains, compared to other members of the genus *Clostridium* using the Maximum Likelihood method. Bootstrap values calculated from 1000 resampled datasets. One scale bar =0.05 nucleotide substitutions per nucleotide position.

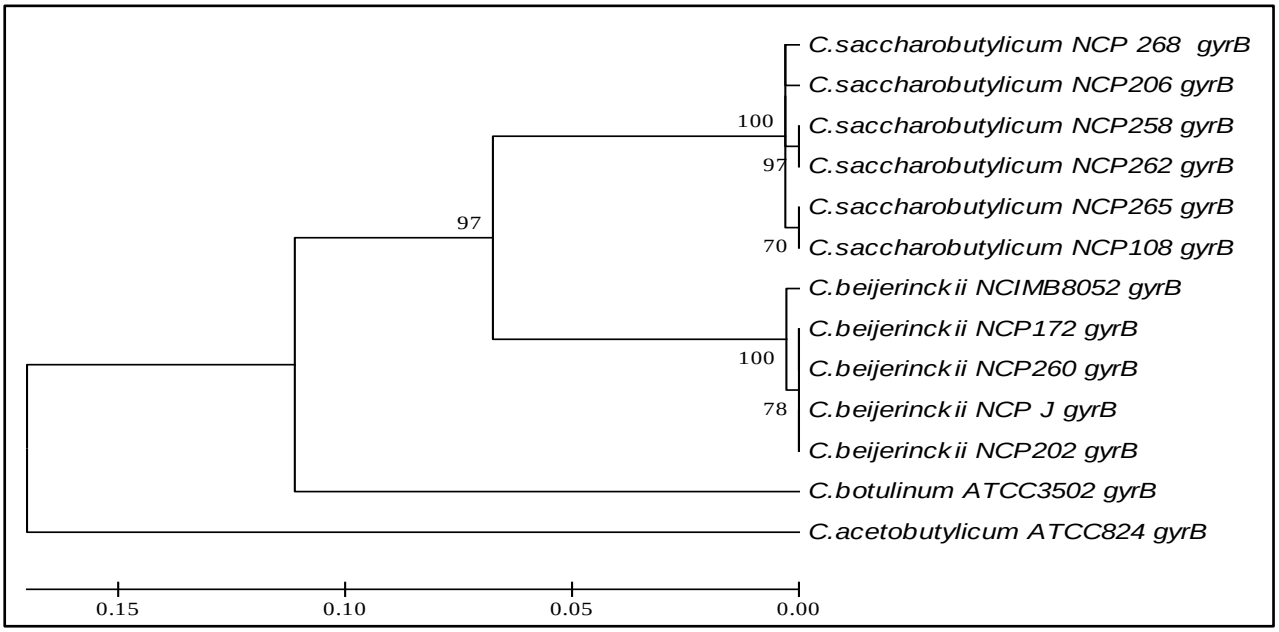


Figure A2: *gyrB* gene phylogenetic tree based on 543 bp of amplified sequence for NCP strains compared to other members of the genus *Clostridium* using the Maximum Likelihood method. Bootstrap values calculated from 1000 resampled datasets. One scale bar =0.05 nucleotide substitutions per nucleotide position.

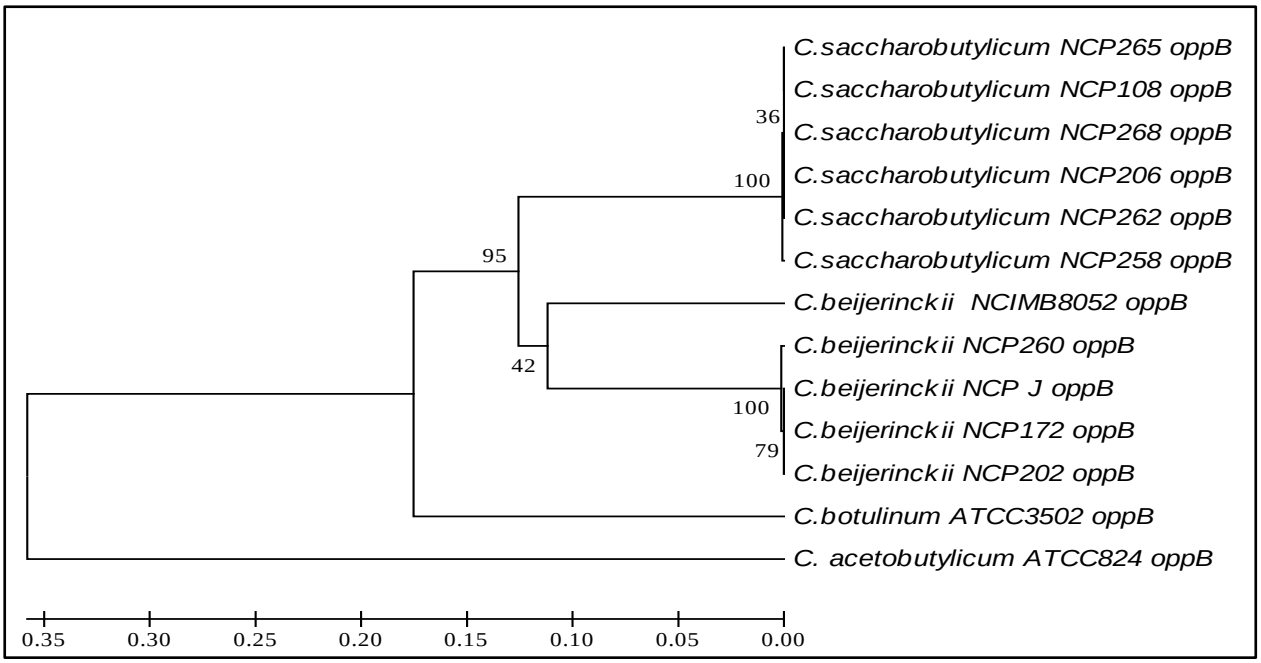


Figure A3: *oppB* gene phylogenetic tree based on 722 bp of amplified sequence for NCP strains compared to other members of the genus *Clostridium* using the Maximum Likelihood method. Bootstrap values calculated from 1000 resampled datasets. One scale bar =0.05 nucleotide substitutions per nucleotide position.

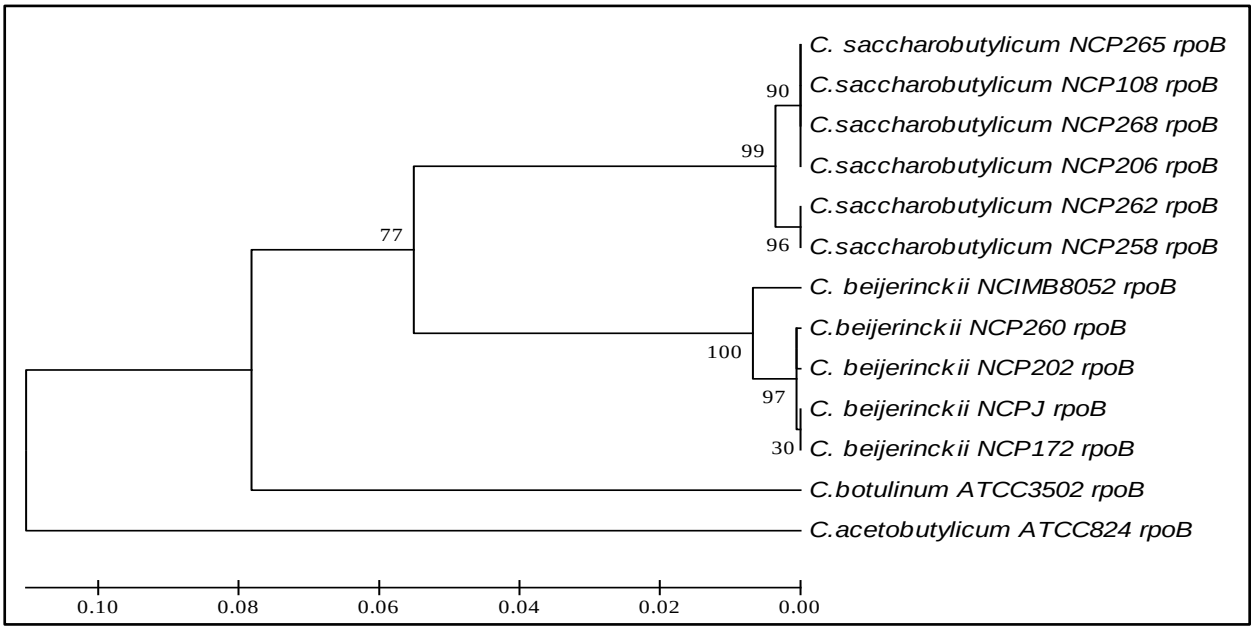


Figure A4: *rpoB* gene phylogenetic tree based on 431 bp of amplified sequence for NCP strains compared to other members of the genus *Clostridium* using the Maximum Likelihood method. Bootstrap values calculated from 1000 resampled datasets. One scale bar =0.02 nucleotide substitutions per nucleotide position.

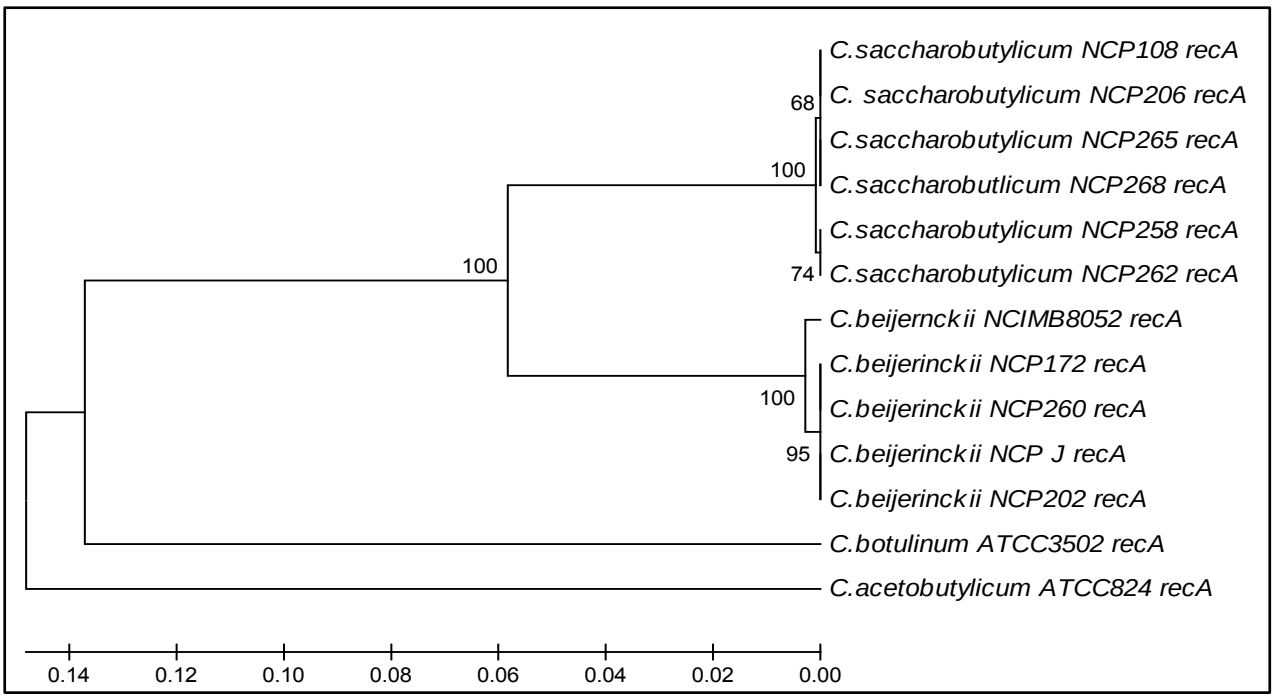


Figure A5: *recA* gene phylogenetic tree based on 683 bp of amplified sequence for NCP strains compared to other members of the genus *Clostridium* using the Maximum Likelihood method. Bootstrap values calculated from 1000 resampled datasets. One scale bar =0.02 nucleotide substitutions per nucleotide position.

Table A1: A summary of the amino acid substitution frequency for the housekeeping genes used in MLST analysis. All *C. beijerinckii*, *C. acetobutylicum* and *C. botulinum* strain amino acid substitutions relative to NCP260 and all *C. saccharobutylicum* strain substitutions relative to NCP262^T.

C. saccharobutylicum	**Strain**	**Number of amino acid substitutions per polypeptide length (% aa polymorphisms)**					
		atpD **(58aa)**	***gyrB*** **(179aa)**	***oppB*** **(240aa)**	***rpoB*** **(143aa)**	***recA*** **(214aa)**	**Pseudogene (834aa)**
	NCP108	0	0	0	0	0	0
	NCP206	0	2 (0.3%)	0	0	0	2 (0.2%)
	NCP265	0	0	0	0	0	0
	NCP268	0	0	0	0	0	0
	NC262^T	0	0	0	0	0	0
	NCP258	0	0	0	0	0	0
C. beijerinckii	NCIMB 8052^T	1 (1.7%)	0	46 (19.2%)	0	0	47 (5.6%)
	NCP260	0	0	0	0	0	0
	NCP J	0	0	0	0	0	0
	NCP172	0	0	0	0	0	0
	NCP202	0	0	0	0	0	0
C. botulinum	ATCC 3502	13 (22.4%)	29 (16.2%)	66 (27.5%)	2 (1.4%)	40 (18.7%)	150 (18%)
C. acetobutylicum	ATCC 824	19 (32.8%)	39 (21.8%)	128 (53.3%)	10 (7.%)	43 (20.1%)	239 (28.7%)

Appendix B

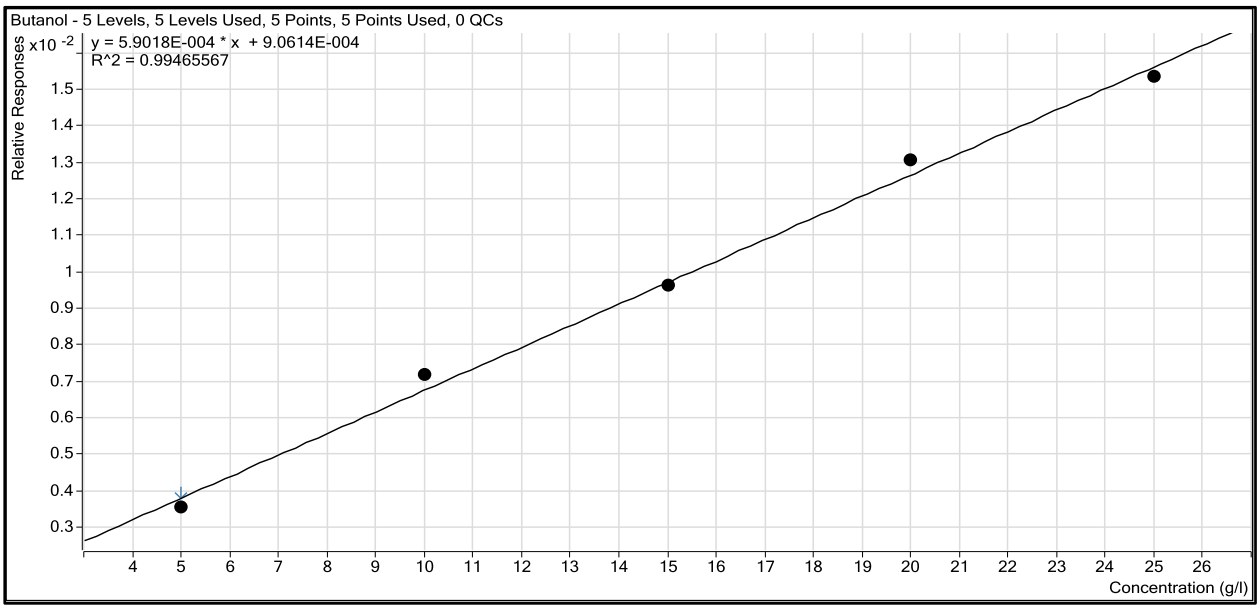
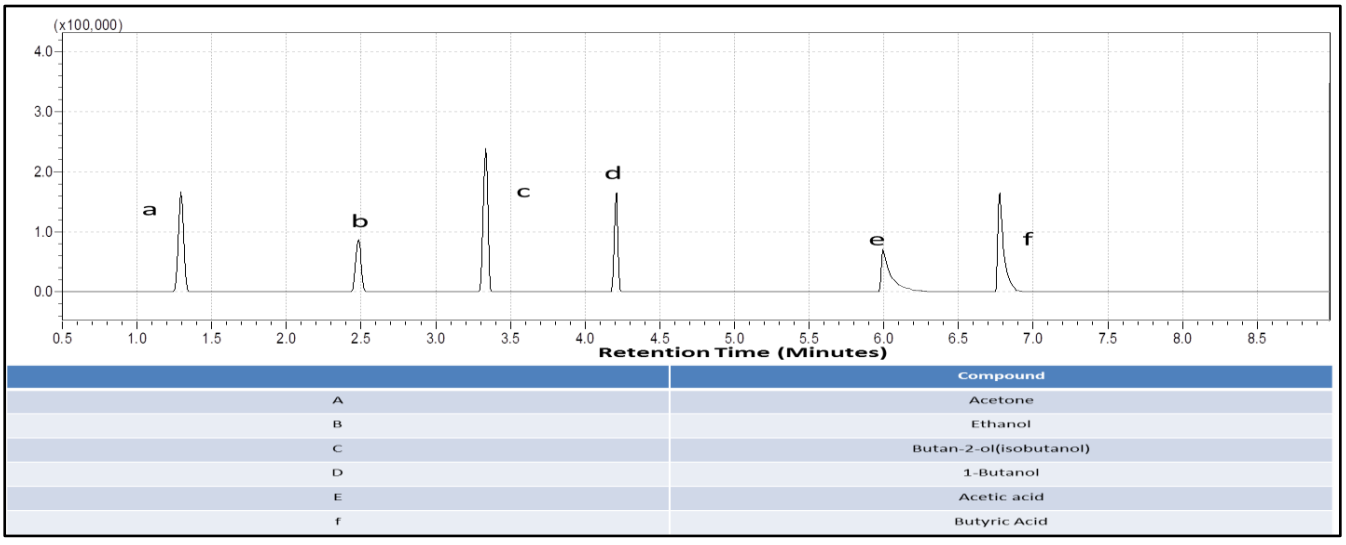


Figure B1: A typical butanol standard curve generated by Mass Hunter Workstation software for 10 g/L isobutanol standard combined with equal volumes of 1,5,10 and 15 g/L butanol standards.



	Compound
A	Acetone
B	Ethanol
C	Butan-2-ol(isobutanol)
D	1-Butanol
E	Acetic acid
f	Butyric Acid

Figure B2: A typical GC-MS run profile for ABE solvents, acids and isobutanol internal standard.

Table B1: Catalogue of ABE solvent production for the UCT NCP *C. beijerinckii* strains in glucose-based NCP media. Standard error of the mean is indicated by ±, where n=3.

Strain	Butanol Titre (g/L)	Butanol Yield (g/50g sugar)	Acetone Titre (g/L)	Ethanol Titre (g/L)	Total Solvent Titre (g/L)	Total Solvent Yield(g/50g sugar)	Butanol: Acetone
NCP J	10.95±0.10	0.219	4.59±0.11	0.42±0.00	15.96±0.09	0.319	2.39
NCP172	10.71±0.08	0.214	4.94±0.09	0.48±0.00	16.13±0.16	0.323	2.17
NCP193	10.60±0.07	0.212	4.81±0.08	0.39±0.00	15.80±0.02	0.316	2.20
NCP271	10.45±0.12	0.209	4.86±0.08	0.39±0.00	15.70±0.19	0.314	2.15
NCP220	10.42±0.08	0.208	5.46±0.05	0.32±0.01	16.20±0.13	0.324	1.91
NCP254	10.31±0.25	0.206	4.74±0.04	0.40±0.00	15.45±0.29	0.309	2.18
NCP259	10.30±0.25	0.206	4.77±0.14	0.49±0.01	15.56±0.39	0.311	2.16
NCP202	10.14±0.10	0.203	4.82±0.05	0.41±0.00	15.38±0.14	0.308	2.10
NCP272	9.65±0.04	0.193	3.88±0.06	0.62±0.01	14.14±0.09	0.283	2.49
NCP270	9.30±0.24	0.186	4.45±0.13	0.57±0.01	14.32±0.36	0.286	2.09
NCP106	9.21±0.16	0.184	4.24±0.06	0.47±0.01	13.92±0.23	0.278	2.17
NCIMB 8052^{T}	9.20±0.06	0.184	1.56±0.04	0.08±0.00	10.85±0.10	0.217	5.88
NCP260	9.15±0.36	0.183	4.96±0.08	0.34±0.01	14.45±0.42	0.289	1.84
NCP200	8.90±0.04	0.178	4.57±0.06	0.56±0.01	14.02±0.04	0.280	1.95
NCP280	8.84±0.57	0.177	4.41±0.33	0.22±0.01	13.48±0.90	0.270	2.00
NCP261	7.36±0.13	0.147	3.21±0.05	0.26±0.00	10.82±0.17	0.216	2.29
NCP263	6.71±0.26	0.134	2.11±0.37	0.53±0.01	9.35±0.59	0.187	3.19
NRRL B593	0.41±0.04	0.008	0.12±0.01	0.19±0.00	0.73±	0.015	3.28

Table B2: Catalogue of ABE solvent production for the UCT NCP *C. saccharobutylicum* strains in glucose-based NCP media. Standard error of the mean is indicated by ±, where n=3.

Strain	**Butanol Titre (g/L)**	**Butanol Yield (g/50g sugar)**	**Acetone Titre (g/L)**	**Ethanol Titre (g/L)**	**Total Solvent Titre (g/L)**	**Total Solvent Yield (g/50g sugar)**	**Butanol: Acetone**
NCP249	10.92±0.12	0.218	2.36±0.04	0.33±0.00	13.61±0.11	0.272	4.62
NCP199	10.71±0.20	0.214	4.31±0.06	0.42±0.00	15.44±0.16	0.309	2.49
NCP195	10.59±0.19	0.212	3.37±0.37	0.32±0.01	14.27±0.44	0.285	3.14
NCP265	10.35±0.25	0.207	4.38±0.33	0.32±0.01	15.05±0.58	0.301	2.37
NCP258	10.17±0.10	0.203	5.55±0.06	0.61±0.02	16.33±0.14	0.327	1.83
NCP262^{T}	10.08±0.32	0.202	2.81±0.19	0.33±0.01	13.22±0.35	0.264	3.58
NCP206	9.84±0.24	0.197	4.85±0.28	0.24±0.04	14.94±0.56	0.299	2.03
NCP268	9.84±0.25	0.197	3.41±0.20	0.36±0.01	13.61±0.44	0.272	2.88
NCP108	9.12±0.16	0.182	3.60±0.08	0.46±0.02	13.18±0.24	0.264	2.53
NCP200 α206	8.38±0.13	0.168	4.07±0.14	0.20±0.01	12.65±0.26	0.253	2.06
NCP162	0.25±0.00	0.005	0.09±0.00	0.21±0.00	0.55±0.01	0.011	2.70

Table B3: Catalogue of ABE solvent production for the UCT NCP *C. beijerinckii* strains in sucrose-based NCP media. Standard error of the mean is indicated by ±, where n=3.

Strain	Butanol Titre (g/L)	Butanol Yield (g/50g sugar)	Acetone Titre (g/L)	Ethanol Titre (g/L)	Total Solvent Titre (g/L)	Total Solvent Yield (g/50g sugar)	Butanol: Acetone
NCP254	10.81±0.18	0.216	4.54±0.12	0.36±0.01	15.72±0.29	0.314	2.38
NCP259	10.66±0.29	0.213	3.84±0.03	0.37±0.01	14.87±0.32	0.297	2.78
NCIMB 8052^{T}	10.41±0.28	0.208	3.32±0.38	0.24±0.01	13.98±0.66	0.280	3.13
NCP172	10.01±0.08	0.200	3.22±0.01	0.39±0.00	13.62±0.08	0.272	3.10
NCP106	9.80±0.29	0.196	4.21±0.14	0.38±0.00	14.38±0.43	0.288	2.33
NCP271	9.74±0.11	0.195	5.82±0.10	0.63±0.01	16.19±0.07	0.324	1.67
NCP J	9.28±0.25	0.186	2.90±0.08	0.30±0.00	12.49±0.33	0.250	3.20
NCP202	8.93±0.19	0.179	2.77±0.15	0.30±0.00	12.00±0.30	0.240	3.22
NCP193	8.81±0.14	0.176	3.02±0.02	0.29±0.00	12.12±0.16	0.242	2.92
NCP270	8.34±0.57	0.167	2.56±0.14	0.29±0.01	11.19±0.71	0.224	3.26
NCP260	8.15±0.11	0.163	3.60±0.03	0.18±0.00	11.93±0.14	0.239	2.27
NCP280	7.87±0.37	0.157	3.11±0.18	0.15±0.01	11.13±0.56	0.223	2.53
NCP200	7.80±0.10	0.156	2.33±0.07	0.31±0.00	10.44±0.13	0.209	3.35
NCP220	7.44±0.10	0.149	3.12±0.12	0.14±0.00	10.71±0.20	0.214	2.38
NCP261	6.52±0.08	0.130	1.79±0.03	0.12±0.00	8.43±0.11	0.169	3.64
NCP272	3.01±0.10	0.060	0.39±0.03	0.11±0.00	3.51±0.13	0.070	7.67
NRRL B593	0.42±0.04	0.008	0.12±0.01	0.20±0.00	0.74±0.03	0.015	3.37
NCP263	0.40±0.01	0.008	0.07±0.01	0.21±0.00	0.68±0.01	0.014	5.60

Table B4: Catalogue of ABE solvent production for the UCT NCP *C. saccharobutylicum* strains in sucrose-based NCP media. Standard error of the mean is indicated by ±, where n=3.

Strain	Butanol Titre (g/L)	Butanol Yield (g/50g sugar)	Acetone Titre (g/L)	Ethanol Titre (g/L)	Total Solvent Titre (g/L)	Total Solvent Yield (g/50g sugar)	Butanol: Acetone
NCP268	11.37±0.18	0.227	4.55±0.06	0.43±0.01	16.35±0.25	0.327	2.50
NCP199	10.59±0.23	0.212	5.16±0.06	0.44±0.01	16.19±0.24	0.324	2.05
NCP249	10.41±0.14	0.208	3.26±0.17	0.36±0.01	14.03±0.30	0.281	3.19
NCP262^{T}	10.02±0.21	0.200	2.84±0.05	0.30±0.00	13.17±0.26	0.263	3.53
NCP265	9.77±0.24	0.195	3.20±0.32	0.22±0.00	13.19±0.55	0.264	3.05
NCP108	9.48±0.14	0.190	3.91±0.02	0.49±0.01	13.87±0.17	0.277	2.43
NCP206	8.88±0.29	0.178	3.97±0.14	0.17±0.01	13.02±0.43	0.260	2.24
NCP200 α206	6.66±0.19	0.133	2.60±0.10	0.12±0.01	9.39±0.29	0.188	2.56
NCP195	6.55±0.12	0.131	1.98±0.16	0.20±0.01	8.73±0.14	0.175	3.31
NCP258	5.02±0.27	0.100	1.92±0.17	0.29±0.01	7.23±0.45	0.145	2.61
NCP162	0.23±0.00	0.005	0.12±0.01	0.21±0.00	0.57±0.01	0.011	1.99

Table B5: Catalogue of ABE solvent production for the UCT NCP *C. beijerinckii* strains in xylose-based NCP media. . Standard error of the mean is indicated by ±, where n=3.

Strain	Butanol Titre (g/L)	Butanol yield (g/50g sugar)	Acetone Titre (g/L)	Ethanol Titre (g/L)	Total Solvents Titre (g/L)	Total solvent yield (g/50g sugar)	Butanol: Acetone
NCP259	9.77±0.31	0.195	3.68±0.12	0.40±0.01	13.84±0.44	0.277	2.66
NCP172	8.84±0.16	0.177	1.45±0.02	0.33±0.00	10.61±0.18	0.212	6.10
NCP271	8.55±0.28	0.171	2.36±0.15	0.33±0.01	11.24±0.44	0.225	3.63
NCP106	8.40±0.12	0.168	3.55±0.02	0.46±0.00	12.42±0.14	0.248	2.36
NCP J	7.99±0.24	0.160	1.04±0.06	0.25±0.00	9.28±0.30	0.186	7.68
NCP202	6.92±0.17	0.138	0.76±0.05	0.25±0.00	7.93±0.22	0.159	9.16
NCP193	5.44±0.33	0.109	0.57±0.05	0.23±0.01	6.23±0.39	0.125	9.58
NCP270	5.05±0.77	0.101	0.50±0.13	0.24±0.01	5.79±0.90	0.116	10.01
NCP254	4.78±0.05	0.096	0.49±0.02	0.22±0.00	5.49±0.05	0.110	9.85
NCIMB 8052^{T}	4.25±0.05	0.085	0.48±0.05	0.15±0.00	4.88±0.10	0.098	8.84
NCP220	4.21±0.02	0.084	0.49±0.01	0.14±0.00	4.84±0.01	0.097	8.61
NCP260	3.61±0.18	0.072	0.40±0.05	0.20±0.00	4.21±0.23	0.084	9.07
NCP261	2.42±0.16	0.048	0.17±0.00	0.21±0.01	2.81±0.16	0.056	14.08
NCP280	2.09±0.06	0.042	0.13±0.00	0.11±0.00	2.33±0.06	0.047	15.54
NCP272	1.23±0.05	0.025	0.10±0.03	0.20±0.00	1.52±0.04	0.030	12.49
NRRL B593	0.42±0.01	0.008	0.12±0.00	0.20±0.00	0.75±0.01	0.015	3.46
NCP263	0.39±0.02	0.008	0.09±0.01	0.22±0.00	0.69±0.01	0.014	4.43
NCP200	0.19±0.01	0.004	0.11±0.02	0.19±0.00	0.49±0.02	0.010	1.78

Table B6: Catalogue of ABE solvent production for the UCT NCP *C. saccharobutylicum* strains in xylose-based NCP media. . Standard error of the mean is indicated by ±, where n=3.

Strain	Butanol Titre (g/L)	Butanol yield (g/50g sugar)	Acetone Titre (g/L)	Ethanol Titre (g/L)	Total Solvents Titre (g/L)	Total solvent yield (g/50g sugar)	Butanol: Acetone
NCP195	10.72±0.05	0.214	4.23±0.26	0.34±0.01	15.29±0.22	0.306	2.54
NCP265	8.84±0.29	0.177	1.04±0.09	0.21±0.01	10.09±0.36	0.202	8.52
NCP162	8.14±0.12	0.163	0.76±0.03	0.26±0.00	9.16±0.15	0.183	10.71
NCP108	5.27±0.35	0.105	2.01±0.34	0.31±0.03	7.59±0.69	0.152	2.62
NCP206	2.29±0.02	0.046	0.14±0.02	0.18±0.01	2.60±0.02	0.052	16.92
NCP200 α206	1.72±0.26	0.034	0.09±0.00	0.11±0.00	1.92±0.26	0.038	19.30
NCP258	0.74±0.02	0.015	0.26±0.05	0.35±0.15	1.35±0.16	0.027	2.88
NCP199	0.59±0.23	0.012	0.18±0.03	0.33±0.03	1.09±0.27	0.022	3.35
NCP262^{T}	0.32±0.07	0.006	0.16±0.02	0.00	0.48±0.09	0.010	2.03
NCP249	0.00	0.000	0.00	0.00	0.00	0.000	0.00
NCP268	0.00	0.000	0.00	0.00	0.00	0.000	0.00

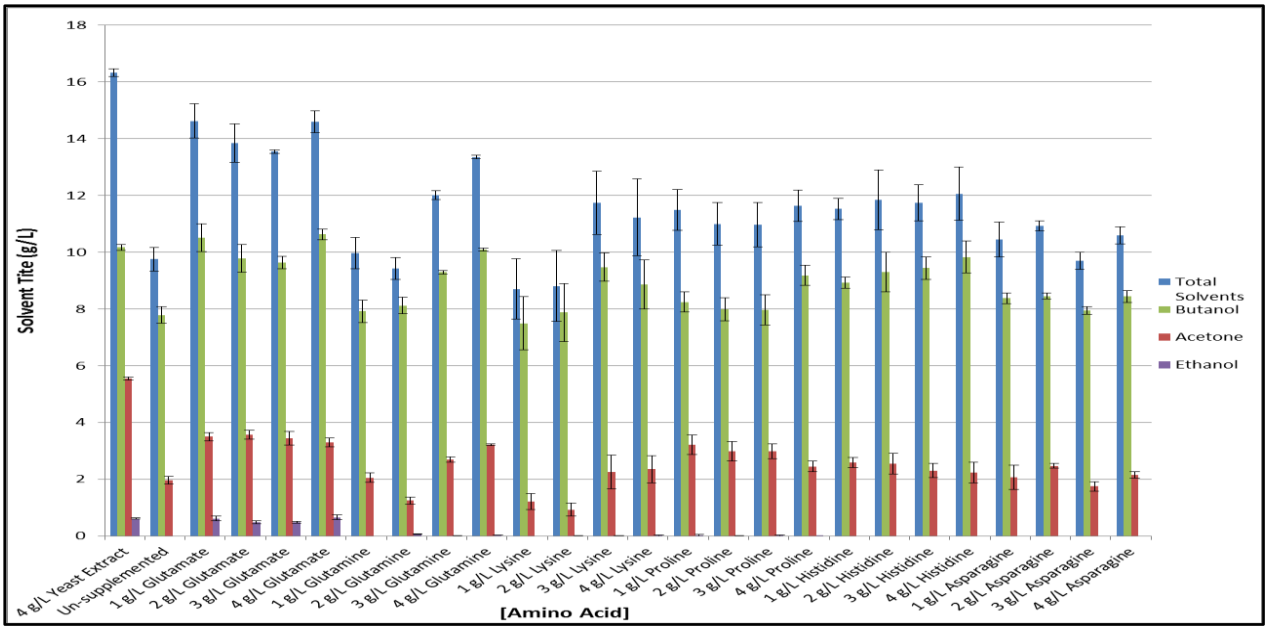


Figure B3: ABE solvent titres for NCP258 utilising various amino acids as the organic nitrogen source in 5% glucose NCP media. Error bars represent the standard error of the mean, where n=3.

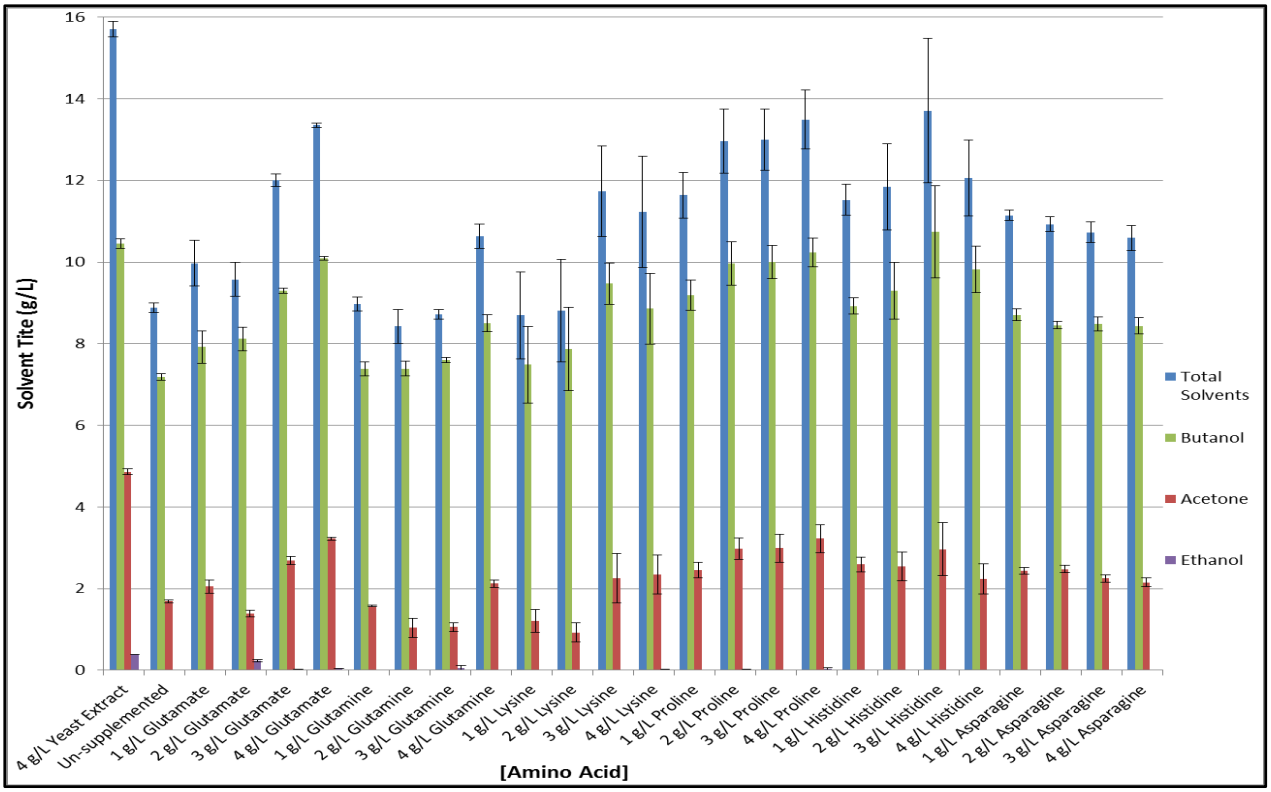


Figure B4: ABE solvent titres for NCP271 utilising various amino acids as the organic nitrogen source in 5% glucose NCP media. Error bars represent the standard error of the mean, where n=3.

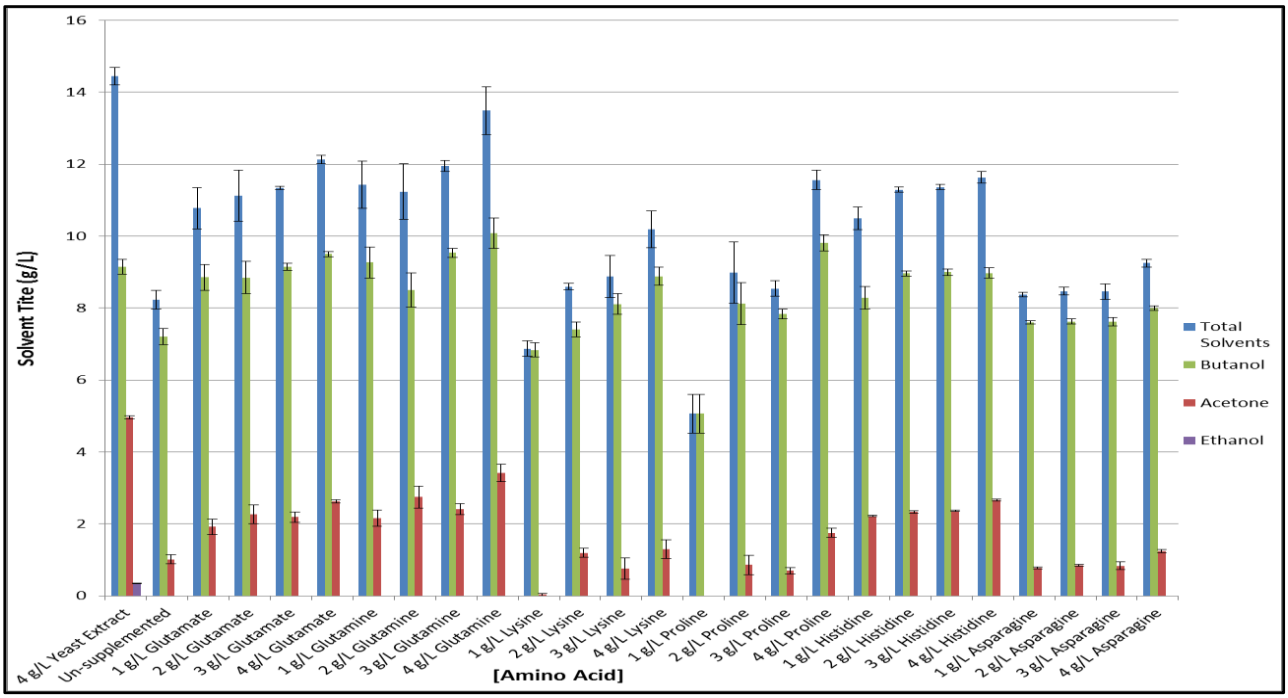


Figure B5: ABE solvent titres for NCP260 utilising various amino acids as the organic nitrogen source in 5% glucose NCP media. Error bars represent the standard error of the mean, where n=3.

Appendix C

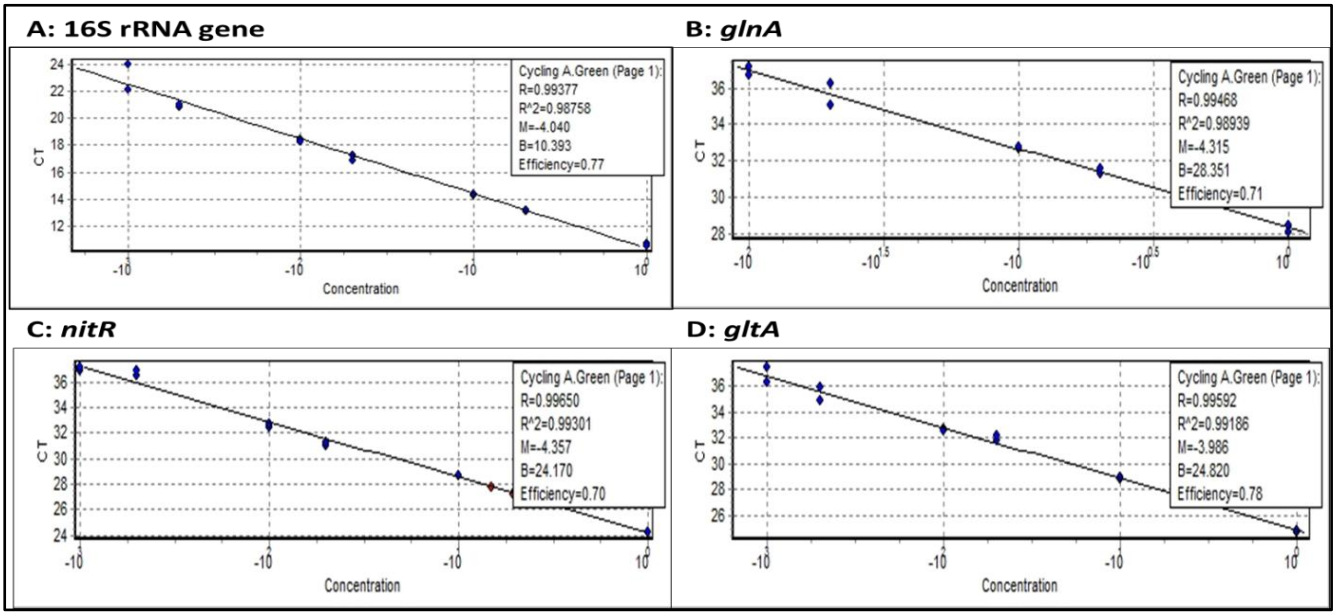


Figure C1: Representative standard curves for each primer pair showing R^2 and efficiency values as well as the M gradient value. Standard curves for **A**, 16s *rRNA* gene primers; **B**, *glnA* gene primers; **C**, *nitR* gene primers and **D**, *gltA* gene primers.

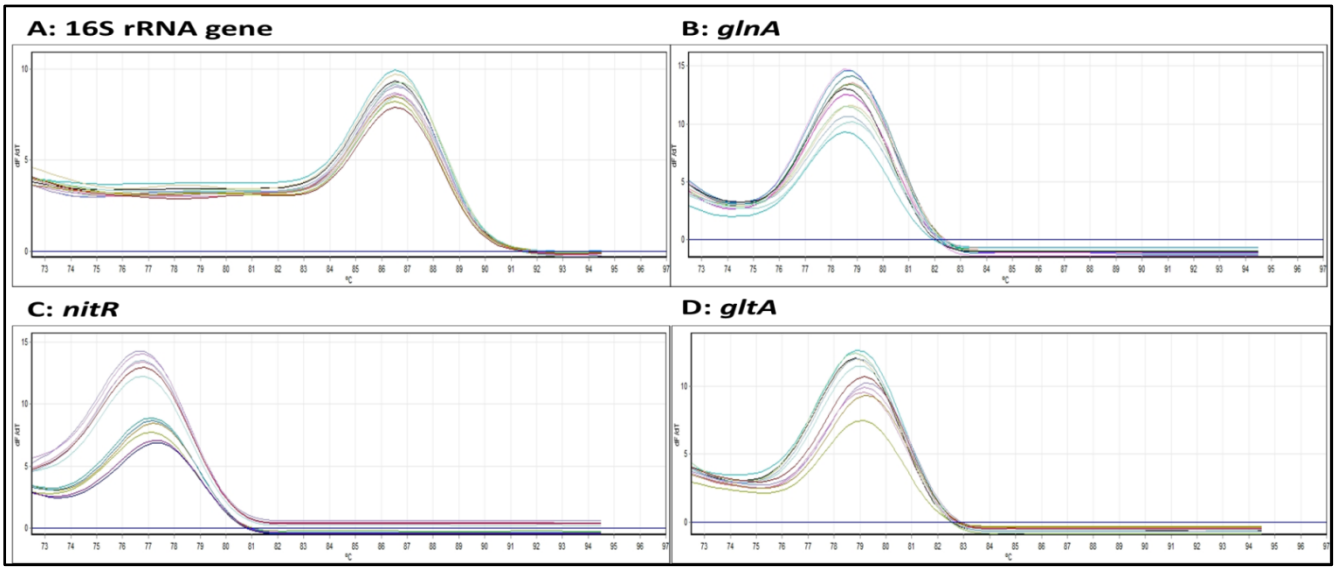


Figure C2: Representative melt curves for each primer pair and all biologicals in a single run. Melt curves for A: 16s *rRNA*; B: *glnA*; C: *nitR* and D: *gltA* gene primers.

Literature cited

Abd-Alla, M. H. & Elsadek El-Enany, A.-W. 2012. Production of acetone-butanol-ethanol from spoilage date palm (Phoenix dactylifera L.) fruits by mixed culture of *Clostridium acetobutylicum* and *Bacillus subtilis*. *Biomass and Bioenergy,* 42**,** 172-178.

Abe, H. 2000. Role of histidine-related compounds as intracellular proton buffering constituents in vertebrate muscle. *Biochemistry (Mosc),* 65**,** 757-765.

Adler, S. P., Purich, D. & Stadtman, E. R. 1975. Cascade control of *Escherichia coli* glutamine synthetase. Properties of the PII regulatory protein and the uridylyltransferase-uridylyl-removing enzyme. *J Biol Chem,* 250**,** 6264-6272.

Alonso, R., Gros, S., Peláez, T., García-de-Viedma, D., Rodríguez-Créixems, M. & Bouza, E. 2001. Molecular analysis of relapse vs re-infection in HIV-positive patients suffering from recurrent *Clostridium difficile* associated diarrhoea. *J Hosp Infect,* 48**,** 86-92.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-410.

Amador-Noguez, D., Brasg, I. A., Feng, X. J., Roquet, N. & Rabinowitz, J. D. 2011. Metabolome remodeling during the acidogenic-solventogenic transition in *Clostridium acetobutylicum*. *Appl Environ Microbiol,* 77**,** 7984-7997.

Amon, J., Bräu, T., Grimrath, A., Hänssler, E., Hasselt, K., Höller, M., Jessberger, N., Ott, L., Szököl, J., Titgemeyer, F. & Burkovski, A. 2008. Nitrogen control in *Mycobacterium smegmatis*: nitrogen-dependent expression of ammonium transport and assimilation proteins depends on the OmpR-type regulator GlnR. *J Bacteriol,* 190**,** 7108-7116.

Amon, J., Titgemeyer, F. & Burkovski, A. 2009. A genomic view on nitrogen metabolism and nitrogen control in mycobacteria. *J Mol Microbiol Biotechnol,* 17**,** 20-29.

Amon, J., Titgemeyer, F. & Burkovski, A. 2010. Common patterns - unique features: nitrogen metabolism and regulation in Gram-positive bacteria. *FEMS Microbiol Rev,* 34**,** 588-605.

Anbarasan, P., Baer, Z. C., Sreekumar, S., Gross, E., Binder, J. B., Blanch, H. W., Clark, D. S. & Toste, F. D. 2012. Integration of chemical catalysis with extractive fermentation to produce fuels. *Nature,* 491**,** 235-239.

Antoni, D., Zverlov, V. V. & Schwarz, W. H. 2007. Biofuels from microbes. *Appl Microbiol Biotechnol,* 77**,** 23-35.

Arikado, E., Ishihara, H., Ehara, T., Shibata, C., Saito, H., Kakegawa, T., Igarashi, K. & Kobayashi, H. 1999. Enzyme level of enterococcal F_1F_o-ATPase is regulated by pH at the step of assembly. *Eur J Biochem,* 259**,** 262-268.

Baker, A. A., Davis, E., Rehberger, T. & Rosener, D. 2010. Prevalence and diversity of toxigenic *Clostridium perfringens* and *Clostridium difficile* among swine herds in the midwest. *Appl Environ Microbiol,* 76**,** 2961-2967.

Balat, M. & Balat, H. 2009. Recent trends in global production and utilization of bio-ethanol fuel. *Applied Energy,* 86**,** 2273-2282.

Balleza, E., López-Bojorquez, L. N., Martínez-Antonio, A., Resendis-Antonio, O., Lozada-Chávez, I., Balderas-Martínez, Y. I., Encarnación, S. & Collado-Vides, J. 2009. Regulation by transcription factors in bacteria: beyond description. *FEMS Microbiol Rev,* 33**,** 133-151.

Bandyopadhyay, A., Arora, A., Jain, S., Laskar, A., Mandal, C., Ivanisenko, V. A., Fomin, E. S., Pintus, S. S., Kolchanov, N. A., Maiti, S. & Ramachandran, S. 2010. Expression and molecular characterization of the *Mycobacterium tuberculosis* PII protein. *J Biochem,* 147**,** 279-289.

Beckers, G., Strösser, J., Hildebrandt, U., Kalinowski, J., Farwick, M., Krämer, R. & Burkovski, A. 2005. Regulation of AmtR-controlled gene expression in *Corynebacterium glutamicum*: mechanism and characterization of the AmtR regulon. *Mol Microbiol,* 58**,** 580-595.

Bender, R. A. 2012. Regulation of the histidine utilization (hut) system in bacteria. *Microbiol Mol Biol Rev,* 76**,** 565-584.

Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J. & Rabinowitz, J.D. 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol,* 5, 593-599.

Berezina, O. V., Brandt, A., Yarotsky, S., Schwarz, W. H. & Zverlov, V. V. 2009. Isolation of a new butanol-producing *Clostridium* strain: high level of hemicellulosic activity and structure of solventogenesis genes of a new *Clostridium saccharobutylicum* isolate. *Syst Appl Microbiol,* 32**,** 449-459.

Borden, J. R., Jones, S. W., Indurthi, D., Chen, Y. & Papoutsakis, E. T. 2010. A genomic-library based discovery of a novel, possibly synthetic, acid-tolerance mechanism in *Clostridium acetobutylicum* involving non-coding RNAs and ribosomal RNA processing. *Metab Eng,* 12**,** 268-281.

Borden, J. R. & Papoutsakis, E. T. 2007. Dynamics of Genomic-Library Enrichment and Identification of Solvent Tolerance Genes for *Clostridium* acetobutylicum. *Applied and Environmental Microbiology,* 73**,** 3061-3068.

Bowles, L. K. & Ellefson, W. L. 1985. Effects of butanol on *Clostridium acetobutylicum*. *Appl Environ Microbiol,* 50**,** 1165–1170.

Bowring, S.N. & Morris, J.G. 1985. Mutagenesis of *Clostridium acetobutylicum*. *J Appl Bacteriol,* 58, 577-584.

Boyd, D. A., Cvitkovitch, D. G., Bleiweis, A. S., Kiriukhin, M. Y., Debabov, D. V., Neuhaus, F. C. & Hamilton, I. R. 2000. Defects in D-alanyl-lipoteichoic acid synthesis in *Streptococcus mutans* results in acid sensitivity. *J Bacteriol,* 182**,** 6055-6065.

Boynton, Z. L., Bennett, G. N. & Rudolph, F. B. 1994. Intracellular concentrations of coenzyme A and Its derivatives from *Clostridium acetobutylicum* ATCC 824 and their roles in enzyme regulation. *Appl Environ Microbiol,* 60**,** 39-44.

Boynton, Z. L., Bennett, G. N. & Rudolph, F. B. 1996. Cloning, sequencing, and expression of clustered genes encoding β-hydroxybutyryl-Coenzyme A (CoA) dehydrogenase crotonase, and butyryl-CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824. *J Bacteriol,* 178**,** 3015–3024.

Broadbent, J. R., Larsen, R. L., Deibel, V. & Steele, J. L. 2010. Physiological and transcriptional response of *Lactobacillus casei* ATCC 334 to acid stress. *J Bacteriol,* 192**,** 2445-2458.

Brown, S. W. & Sonenshein, A. L. 1996. Autogenous regulation of the *Bacillus subtilis glnRA* operon. *J Bacteriol,* 178**,** 2450-2454.

Buchinger, S., Strösser, J., Rehm, N., Hänssler, E., Hans, S., Bathe, B., Schomburg, D., Krämer, R. & Burkovski, A. 2009. A combination of metabolome and transcriptome analyses reveals new targets of the *Corynebacterium glutamicum* nitrogen regulator AmtR. *J Biotechnol,* 140**,** 68-74.

Bui, N. K., Eberhardt, A., Vollmer, D., Kern, T., Bougault, C., Tomasz, A., Simorre, J. P. & Vollmer, W. 2012. Isolation and analysis of cell wall components from *Streptococcus pneumoniae*. *Anal Biochem,* 421**,** 657-666.

Burkovski, A. 2003. Ammonium assimilation and nitrogen control in *Corynebacterium glutamicumand* its relatives: an example for new regulatory mechanisms in actinomycetes. *FEMS Microbiol Rev,* 27**,** 617-628.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J. & Wittwer, C. T. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem,* 55**,** 611-622.

Calusinska, M., Joris, B. & Wilmotte, A. 2011. Genetic diversity and amplification of different clostridial [FeFe] hydrogenases by group-specific degenerate primers. *Lett Appl Microbiol,* 53**,** 473-480.

Carnahan, J. E., Mortenson, L. E. & Castle, J. E. 1960. Selective inhibition of nitrogen fixation in *Clostridium pasteurianum*. *J Bacteriol,* 80**,** 311-315.

Carroll, P., Pashley, C. A. & Parish, T. 2008. Functional analysis of GlnE, an essential adenylyl transferase in *Mycobacterium tuberculosis*. *J Bacteriol,* 190**,** 4894-4902.

Casiano-Colón, A. & Marquis, R. E. 1988. Role of the arginine deiminase system in protecting oral bacteria and an enzymatic basis for acid tolerance. *Appl Environ Microbiol,* 54**,** 1318-1324.

Chalmers, G., Bruce, H. L., Hunter, D. B., Parreira, V. R., Kulkarni, R. R., Jiang, Y. F., Prescott, J. F. & Boerlin, P. 2008. Multilocus sequence typing analysis of *Clostridium perfringens* isolates from necrotic enteritis outbreaks in broiler chicken populations. *J Clin Microbiol,* 46**,** 3957-3964.

Chatterji, D. & Ojha, A. K. 2001. Revisiting the stringent response, ppGpp and starvation signaling. *Curr Opin Microbiol,* 4**,** 160-165.

Chen, C., Xiao, Z., Tang, X., Cui, H., Zhang, J., Li, W. & Ying, C. 2013. Acetone-butanol-ethanol fermentation in a continuous and closed-circulating fermentation system with PDMS membrane bioreactor. *Bioresour Technol,* 128**,** 246-251.

Chen, J., Toth, J. & Kasap, M. 2001. Nitrogen-fixation genes and nitrogenase activity in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. *J Ind Microbiol Biotechnol,* 27 281-286.

Chen, P. M., Chen, Y. Y., Yu, S. L., Sher, S., Lai, C. H. & Chia, J. S. 2010. Role of GlnR in acid-mediated repression of genes encoding proteins involved in glutamine and glutamate metabolism in Streptococcus mutans. *Appl Environ Microbiol,* 76**,** 2478-2486.

Chen, Y. Y., Hall, T. H. & Burne, R. A. 1998. *Streptococcus salivarius* urease expression: involvement of the phosphoenolpyruvate: sugar phosphotransferase system. *FEMS Microbiol Lett,* 165**,** 117-122.

Cho, D. H., Lee, Y. J., Um, Y., Sang, B.-I. & Kim, Y. H. 2009. Detoxification of model phenolic compounds in lignocellulosic hydrolysates with peroxidase for butanol production from *Clostridium beijerinckii*. *Appl Microbiol Biotechnol,* 83**,** 1035-1043.

Claassen, P. A., Budde, M. A. & López-Contreras, A. M. 2000. Acetone, butanol and ethanol production from domestic organic waste by solventogenic clostridia. *J Mol Microbiol Biotechnol,* 2**,** 39-44.

Collas, F., Kuit, W., Clément, B., Marchal, R., López-Contreras, A. M. & Monot, F. 2012. Simultaneous production of isopropanol, butanol, ethanol and 2,3-butanediol by *Clostridium acetobutylicum* ATCC 824 engineered strains. *AMB Express,* 2**,** 1-10.

Collins, M. D., Lawson, P. A., Willems, A., Cordoba, J. J., Fernandez-Garayzabal, J., Garcia, P., Cai, J., Hippe, H. & Farrow, J. A. E. 1994. The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol,* 44**,** 812-826.

Cornillot, E., Nair, R. V., Papoutsakis, E. T. & Soucaille, P. 1997. The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. *J Bacteriol,* 179**,** 5442-5447.

Cotter, P. D., Gahan, C. G. & Hill, C. 2000. Analysis of the role of the *Listeria monocytogenes* F_0F_1-ATPase operon in the acid tolerance response. *Int J Food Microbiol,* 60**,** 137-146.

Cotter, P. D. & Hill, C. 2003. Surviving the acid test: responses of Gram-positive bacteria to low pH. *Microbiology and Molecular Biology Reviews,* 67**,** 429-453.

Cozzani, I., Misuri, A. & Santoni, C. 1970. Purification and general properties of glutamate decarboxylase from *Clostridium perfringens*. *Biochem J,* 118**,** 135-141.

Craig, C. L. & Weber, R. S. 1998. Selection costs of amino acid substitutions in ColE1 and ColIa gene clusters harbored by *Escherichia coli*. *Mol Biol Evol,* 15**,** 774-776.

Cvitkovitch, D. G., Gutierrez, J. A. & Bleiweis, A. S. 1997. Role of the citrate pathway in glutamate biosynthesis by *Streptococcus mutans*. *J Bacteriol,* 179**,** 650-655.

Davis, M. J., Coote, P. J. & O'Byrne, C. P. 1996. Acid tolerance in *Listeria monocytogenes*: the adaptive acid tolerance response (ATR) and growth-phase-dependent acid resistance. *Microbiology,* 142**,** 2975-2982.

De Angelis, M., Mariotti, L., Rossi, J., Servili, M., Fox, P. F., Rollan, G. & Gobbetti, M. 2002. Arginine catabolism by sourdough lactic acid bacteria: purification and characterization of the arginine deiminase pathway enzymes from *Lactobacillus sanfranciscensis* CB1. *Applied and Environmental Microbiology,* 68**,** 6193-6201.

Degnan, B. A., Fontaine, M. C., Doebereiner, A. H., Lee, J. J., Mastroeni, P., Dougan, G., Goodacre, J. A. & Kehoe, M. A. 2000. Characterization of an isogenic mutant of *Streptococcus pyogenes* Manfredo lacking the ability to make streptococcal acid glycoprotein. *Infect Immun,* 68**,** 2441–2448.

Demirbas, A. 2009. Political, economic and environmental impacts of biofuels: A review. *Applied Energy,* 86**,** S108-S117.

Dineen, S. S., Villapakkam, A. C., Nordman, J. T. & Sonenshein, A. L. 2007. Repression of *Clostridium difficile* toxin gene expression by CodY. *Mol Microbiol,* 66**,** 206-219.

Dürre, P. 2008. Fermentative butanol production: bulk chemical and biofuel. *Ann N Y Acad Sci,* 1125**,** 353-362.

Dürre, P. 2011. Fermentative production of butanol--the academic perspective. *Curr Opin Biotechnol,* 22**,** 331-336.

Efremenko, E. N., Nikolskaya, A. B., Lyagin, I. V., Senko, O. V., Makhlis, T. A., Stepanov, N. A., Maslova, O. V., Mamedova, F. & Varfolomeev, S. D. 2012. Production of biofuels from pretreated microalgae biomass by anaerobic fermentation with immobilized *Clostridium acetobutylicum* cells. *Bioresource Technology,* 114**,** 342-348.

Ellis, J. T., Hengge, N. N., Sims, R. C. & Miller, C. D. 2012. Acetone, butanol, and ethanol production from wastewater algae. *Bioresource Technology,* 111**,** 491-495.

Ezeji, T. & Blaschek, H. P. 2008. Fermentation of dried distillers' grains and solubles (DDGS) hydrolysates to solvents and value-added products by solventogenic clostridia. *Bioresour Technol,* 99**,** 5232-5242.

Ezeji, T., Milne, C., Price, N. D. & Blaschek, H. P. 2010. Achievements and perspectives to overcome the poor solvent resistance in acetone and butanol-producing microorganisms. *Appl Microbiol Biotechnol,* 85**,** 1697-1712.

Ezeji, T. C., Groberg, M., Qureshi, N. & Blaschek, H. P. 2003. Continuous production of butanol from starch-based packing peanuts. *Appl Biochem Biotechnol,* 105, 375-382.

Ezeji, T. C., Qureshi, N. & Blaschek, H. P. 2004. Butanol fermentation research: upstream and downstream manipulations. *Chem Rec,* 4**,** 305-314.

Ezeji, T. C., Qureshi, N. & Blaschek, H. P. 2005. Continuous butanol fermentation and feed starch retrogradation: butanol fermentation sustainability using *Clostridium beijerinckii* BA101. *J Biotechnol,* 115**,** 179-187.

Ezeji, T. C., Qureshi, N. & Blaschek, H. P. 2007. Production of acetone butanol (AB) from liquefied corn starch, a commercial substrate, using *Clostridium beijerinckii* coupled with product recovery by gas stripping. *J Ind Microbiol Biotechnol,* 34**,** 771-777.

Fierro-Monti, I. P., Reid, S. J. & Woods, D. R. 1992. Differential expression of a *Clostridium acetobutylicum* antisense RNA: implications for regulation of glutamine synthetase. *J Bacteriol,* 174**,** 7642-7647.

Fink, D., Weissschuh, N., Reuther, J., Wohlleben, W. & Engels, A. 2002. Two transcriptional regulators GlnR and GlnRII are involved in regulation of nitrogen metabolism in *Streptomyces coelicolor* A3(2). *Mol Microbiol,* 46**,** 331-347.

Fonknechten, N., Chaussonnerie, S., Tricot, S., Lajus, A., Andreesen, J. R., Perchat, N., Pelletier, E., Gouyvenoux, M., Barbe, V., Salanoubat, M., Le Paslier, D., Weissenbach, J., Cohen, G. N. & Kreimeyer, A. 2010. *Clostridium sticklandii*, a specialist in amino acid degradation:revisiting its metabolism through its genome sequence. *BMC Genomics,* 11**,** 555.

Forchhammer, K. 2004. Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. *FEMS Microbiol Rev,* 28**,** 319-333.

Forchhammer, K. 2008. P(II) signal transducers: novel functional and structural insights. *Trends Microbiol,* 16**,** 65-72.

Fredericks, W. W. & Stadtman, E. R. 1965. The nature of the cofactor requirements of the hydrogenase system from *Clostridium kluyveri*. *J Biol Chem,* 240**,** 4809-4815.

Gabriel, C. L. 1928. Butanol fermentation process. *Ind Eng Chem,* 20**,** 1063-1067.

García, V., Päkkilä, J., Ojamo, H., Muurinen, E. & Keiski, R. L. 2011. Challenges in biobutanol production: How to improve the efficiency? *Renewable and Sustainable Energy Reviews,* 15**,** 964-980.

Goebel, B. M. & Stackebrandt, E. 1994. Cultural and phylogenetic analysis of mixed microbial populations found in natural and commercial bioleaching environments. *Appl Environ Microbiol,* 60**,** 1614-1621.

Gómez-Valero, L., Rocha, E. P., Latorre, A. & Silva, F. J. 2007. Reconstructing the ancestor of *Mycobacterium leprae*: the dynamics of gene loss and genome reduction. *Genome Res,* 17**,** 1178-1185.

Green, E. M. 2011. Fermentative production of butanol- the industrial perspective. *Curr Opin Biotechnol,* 22**,** 337-343.

Green, E. M., Bennett, G. N., Boynton, Z. L., Harris, L. M., Rudolph, F. B. & Papoutsakis, E. T. 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. *Microbiology,* 142**,** 2079-2086.

Gottschal, J. & Morris, J. G. 1981. Non-production of acetone and butanol by *Clostridium acetobutylicum* during glucose- and ammonium-limitation in continuous culture. *Biotechnol Lett*, 3, 525–530.

Grimmler, C., Janssen, H., Krausse, D., Fischer, R. J., Bahl, H., Dürre, P., Liebl, W. & Ehrenreich, A. 2011. Genome-wide gene expression analysis of the switch between acidogenesis and solventogenesis in continuous cultures of *Clostridium acetobutylicum*. *J Mol Microbiol Biotechnol,* 20**,** 1-15.

Grundy, F. J., Hodil, S. E., Rollins, S. M. & Henkin, T. M. 1997. Specificity of tRNA-mRNA interactions in *Bacillus subtilis tyrS* antitermination. *J Bacteriol,* 179**,** 2587-2594.

Grupe, H. & Gottschalk, G. 1992. Physiological events in *Clostridium acetobutylicum* during the shift from acidogenesis to solventogenesis in continuous culture and presentation of a model for shift induction. *Appl Environ Microbiol,* 58**,** 3896-3902.

Hannett, G. E., Stone, W. B., Davis, S. W. & Wroblewski, D. 2011. Biodiversity of *Clostridium botulinum* type E associated with a large outbreak of botulism in wildlife from Lake Erie and Lake Ontario. *Appl Environ Microbiol,* 77**,** 1061-1068.

Harris, L. M., Welker, N. E. & Papoutsakis, E. T. 2002. Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *Journal of Bacteriology,* 184**,** 3586-3597.

Harth, G., Maslesa-Galić, S., Tullius, M. V. & Horwitz, M. A. 2005. All four *Mycobacterium tuberculosis glnA* genes encode glutamine synthetase activities but only GlnA1 is abundantly expressed and essential for bacterial homeostasis. *Mol Microbiol,* 58**,** 1157-1172.

Hastings, J. H. J. 1978. Acetone-butyl alcohol fermentation. *Economic microbiology, vol. 2. Primary products of metabolism*. A. H. Rose (ed.).31-45 Academic Press, Inc., New York.

Heap, J. T., Kuehne, S. A., Ehsaan, M., Cartman, S. T., Cooksley, C. M., Scott, J. C. & Minton, N. P. 2010. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *Journal of Microbiological Methods,* 80**,** 49-55.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P. & Minton, N. P. 2007. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J Microbiol Methods,* 70**,** 452-464.

Heap, J. T., Pennington, O. J., Cartman, S. T. & Minton, N. P. 2009. A modular system for *Clostridium* shuttle plasmids. *Journal of Microbiological Methods,* 78**,** 79-85.

Heluane, H., Evans, M. R., Dagher, S. F. & Bruno-Bárcena, J. M. 2011. Meta-analysis and functional validation of nutritional requirements of solventogenic clostridia growing under butanol stress conditions and coutilization of D-glucose and D-xylose. *Appl Environ Microbiol,* 77**,** 4473-4485.
Hersh, B. M., Farooq, F. T., Barstad, D. N., Blankenhorn, D. L. & Slonczewski, J. L. 1996. A glutamate-dependent acid resistance gene in *Escherichia coli*. *J Bacteriol,* 178**,** 3978-3981.
Hesketh, A., Fink, D., Gust, B., Rexer, H. U., Scheel, B., Chater, K., Wohlleben, W. & Engels, A. 2002. The GlnD and GlnK homologues of *Streptomyces coelicolor* A3(2) are functionally dissimilar to their nitrogen regulatory system counterparts from enteric bacteria. *Mol Microbiol,* 46**,** 319-330.
Higuchi, T., Hayashi, H. & Abe, K. 1997. Exchange of glutamate and gamma-aminobutyrate in a *Lactobacillus* strain. *J Bacteriol,* 179**,** 3362-3364.
Holdeman, L. V. E., Cato, E. P. & Moore, W. E. C. (1977). *Anaerobe Laboratory Manual*. 15-16. Blacksburg.
Hou, L. H. & Dutta, S. K. 2000. Phylogenetic characterization of several para- and meta-PCB dechlorinating *Clostridium* species: 16s rDNA analysis. *Lett Appl Microbiol,* 30**,** 238-243.
Huang, L., Gibbins, L. N. & Forsberg, C. W. 1985. Transmembrane pH gradient and membrane potential in *Clostridium acetobutylicum* during growth under acetogenic and solventogenic conditions. *Appl Environ Microbiol,* 50**,** 1043-1047.
Inglett, K. S., Bae, H. S., Aldrich, H. C., Hatfield, K. & Ogram, A. V. 2011. *Clostridium chromiireducens* sp. nov., isolated from Cr(VI)-contaminated soil. *Int J Syst Evol Microbiol,* 61**,** 2626-2631.
Izák, P., Schwarz, K., Ruth, W., Bahl, H. & Kragl, U. 2008. Increased productivity of *Clostridium acetobutylicum* fermentation of acetone, butanol, and ethanol by pervaporation through supported ionic liquid membrane. *Appl Microbiol Biotechnol,* 78**,** 597-602.
Jahreis, K., Pimentel-Schmitt, E. F., Brückner, R. & Titgemeyer, F. 2008. Ins and outs of glucose transport systems in eubacteria. *FEMS Microbiol Rev,* 32**,** 891-907.
Jakoby, M., Nolden, L., Meier-Wagner, J., Krämer, R. & Burkovski, A. 2000. AmtR, a global repressor in the nitrogen regulation system of *Corynebacterium glutamicum*. *Mol Microbiol,* 37**,** 964-977.
Jang, Y. S., Malaviya, A., Cho, C., Lee, J. & Lee, S. Y. 2012. Butanol production from renewable biomass by clostridia. *Bioresour Technol,* 123**,** 653-663.
Janssen, H., Döring, C., Ehrenreich, A., Voigt, B., Hecker, M., Bahl, H. & Fischer, R. J. 2010. A proteomic and transcriptional view of acidogenic and solventogenic steady-state cells of *Clostridium acetobutylicum* in a chemostat culture. *Appl Microbiol Biotechnol,* 87**,** 2209-2226.
Janssen, H., Grimmler, C., Ehrenreich, A., Bahl, H. & Fischer, R. J. 2012. A transcriptional study of acidogenic chemostat cells of *Clostridium acetobutylicum*- solvent stress caused by a transient n-butanol pulse. *J Biotechnol,* 161**,** 354-365.
Jayakumar, A., Schulman, I., Macneil, D. & Barnes, E. M. J. 1986. Role of the *Escherichia coli glnALG* operon in regulation of ammonium transport. *J Bacteriol,* 166**,** 281-284.
Jayaraman, G. C., Penders, J. E. & Burne, R. A. 1997. Transcriptional analysis of the *Streptococcus mutans hrcA, grpE* and *dnaK* genes and regulation of expression in response to heat shock and environmental acidification. *Mol Microbiol,* 25**,** 329-341.
Jiang, P., Atkinson, M. R., Srisawat, C., Sun, Q. & Ninfa, A. J. 2000. Functional dissection of the dimerization and enzymatic activities of *Escherichia coli* nitrogen regulator II and their regulation by the PII protein. *Biochemistry,* 39**,** 13433-13449.
Jiang, P. & Ninfa, A. J. 2009. Alpha-ketoglutarate controls the ability of the *Escherichia coli* PII signal transduction protein to regulate the activities of NRII (NtrB but does not control the binding of PII to NRII). *Biochemistry,* 48**,** 11514-11521.
Jiang, Y., Xu, C., Dong, F., Yang, Y., Jiang, W. & Yang, S. 2009. Disruption of the acetoacetate decarboxylase gene in solvent-producing *Clostridium acetobutylicum* increases the butanol ratio. *Metab Eng,* 11**,** 284-291.
Johnson, J. L., Toth, J., Santiwatanakul, S. & Chen, J. S. 1997. Cultures of "*Clostridium acetobutylicum*" from various collections comprise *Clostridium acetobutylicum*, *Clostridium beijerinckii*, and two other distinct types based on DNA-DNA reassociation. *Int J Syst Bacteriol,* 47**,** 420-424.
Jones, D.T. 1993. Mutagenesis and its application in biotechnology. *The clostridia and biotechnology*. 77-97. Butterworth-Heineman. London.
Jones, D. T., Shirley, M., Wu, X. & Keis, S. 2000. Bacteriophage infections in the industrial Acetone Butanol (AB) fermentation process. *J Mol Microbiol Biotechnol,* 2**,** 21-26.
Jones, D. T. & Woods, D. R. 1986. Acetone-Butanol Fermentation Revisited. *Microbiol Rev,* 50**,** 484-524.
Jones, S. W., Paredes, C. J., Tracy, B., Cheng, N., Sillers, R., Senger, R. S. & Papoutsakis, E. T. 2008. The transcriptional program underlying the physiology of clostridial sporulation. *Genome Biol,* 9**,** R114.
Karakashev, D., Thomsen, A. B. & Angelidaki, I. 2007. Anaerobic biotechnological approaches for production of liquid energy carriers from biomass. *Biotechnol Lett,* 29**,** 1005-1012.

Keis, S., Bennett, C. F., Ward, V. K. & Jones, D. T. 1995. Taxonomy and phylogeny of industrial solvent-producing clostridia. *Int J Syst Bacteriol,* 45**,** 693-705.
Keis, S., Shaheen, R. & Jones, D. T. 2001. Emended descriptions of *Clostridium acetobutylicum* and *Clostridium beijerinckii,* and descriptions of *Clostridium saccharoperbutylacetonicum* sp. nov. and *Clostridium saccharobutylicum* sp. nov. *Int J Syst Evol Microbiol,* 51**,** 2095–2103.

Kenney, L. J. 2002. Structure/function relationships in OmpR and other winged-helix transcription factors. *Curr Opin Microbiol,* 5**,** 135-141.
Kilimann, K. V., Hartmann, C., Vogel, R. F. & Gänzle, M. G. 2005. Differential inactivation of glucose- and glutamate dependent acid resistance of *Escherichia coli* TMW 2.497 by high-pressure treatments. *Syst Appl Microbiol,* 28**,** 663-671.
Kleiner, D. & Fitzke, E. 1994. Evidence for ammonia translocation by *Clostridium pasteurianum*. *Biochem Biophys Res Commun*, 86, 211-217.

Kobayashi, H., Suzuki, T. & Unemoto, T. 1986. Streptococcal cytoplasmic pH is regulated by changes in amount and activity of a proton-translocating ATPase. *J Biol Chem,* 261**,** 627-630.
Krastel, K., Senadheera, D. B., Mair, R., Downey, J. S., Goodman, S. D. & Cvitkovitch, D. G. 2010. Characterization of a glutamate transporter operon, *glnQHMP*, in *Streptococcus mutans* and its role in acid tolerance. *J Bacteriol,* 192**,** 984-993.
Kumar, M. & Gayen, K. 2011. Developments in biobutanol production: New insights. *Applied Energy,* 88**,** 1999-2012.
Laichoubi, K. B., Beez, S., Espinosa, J., Forchhammer, K. & Contreras, A. 2011. The nitrogen interaction network in *Synechococcus* WH5701, a cyanobacterium with two PipX and two P(II)-like proteins. *Microbiology,* 157**,** 1220-1228.
Lam, H., Oh, D. C., Cava, F., Takacs, C. N., Clardy, J., De Pedro, M. A. & Waldor, M. K. 2009. D-amino acids govern stationary phase cell wall remodeling in bacteria. *Science,* 325**,** 1552-1555.
Lee, J. 2009. Fermentation of rice bran and defatted rice bran for butanol production using *Clostridium beijerinckii* NCIMB 8052. *Journal of Microbiology and Biotechnology,* 19**,** 482-490.
Lee, J., Jang, Y. S., Choi, S. J., Im, J. A., Song, H., Cho, J. H., Seung, D. Y., Papoutsakis, E. T., Bennett, G. N. & Lee, S. Y. 2011. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for isopropanol-butanol-ethanol fermentation. *Applied and Environmental Microbiology,* 78**,** 1416-1423.
Lee, S. Y., Park, J. H., Jang, S. H., Nielsen, L. K., Kim, J. & Jung, K. S. 2008. Fermentative butanol production by Clostridia. *Biotechnol Bioeng,* 101**,** 209-228.
Lienhardt, J., Schripsema, J., Qureshi, N. & Blaschek, H. P. 2002. Butanol production by *Clostridium beijerinckii* BA101 in an immobilized cell biofilm reactor: increase in sugar utilization. *Appl Biochem Biotechnol,* 98, 591-598.
Liew, S. T., Arbakariya, A., Rosfarizan, M. & Raha, A. R. 2006. Production of solvent (acetone– butanol–ethanol) in continuous fermentation by *Clostridium saccharobutylicum* DSM 13864 using gelatinised sago starch as a carbon source. *Malays J Microbiol,* 2**,** 42-50.
Lim, E. M., Ehrlich, S. D. & Maguin, E. 2000. Identification of stress-inducible proteins in *Lactobacillus delbrueckii* subsp. *bulgaricus*. *Electrophoresis,* 21**,** 2557-2561.
Lin, Y. L. & Blaschek, H. P. 1983. Butanol production by a butanol-tolerant strain of *Clostridium acetobutylicum* in extruded corn broth. *Appl Environ Microbiol,* 45**,** 966-973.
Liyanage, H., Young, M. and Kashket, E. R. 2000. Butanol Tolerance of *Clostridium beijerinckii* NCIMB 8052 associated with down regulation of *gldA* by antisense RNA. *J Mol Microbiol Biotechnol*, 2, 87-93.
Long, S., Jones, D. T. & Woods, D. R. 1984. Initiation of solvent production, clostridial stage and endospore formation in *Clostridium acetobutylicum* P262. *Appl Environ Microbiol*, 20, 256–261.
López-Contreras, A. M., Claassen, P. A., H., M. & De Vos, W. M. 2000. Utilisation of saccharides in extruded domestic organic waste by *Clostridium acetobutylicum* ATCC 824 for production of acetone, butanol and ethanol. *Appl Microbiol Biotechnol,* 54**,** 162-167.
Lu, C., Dong, J. & Yang, S. T. 2013a. Butanol production from wood pulping hydrolysate in an integrated fermentation-gas stripping process. *Bioresour Technol,* 143**,** 467-475.
Lu, C., Zhao, J., Yang, S.-T. & Wei, D. 2012. Fed-batch fermentation for n-butanol production from cassava bagasse hydrolysate in a fibrous bed bioreactor with continuous gas stripping. *Bioresource Technology,* 104**,** 380-387.
Lu, P., Ma, D., Chen, Y., Guo, Y., Chen, G.-Q., Deng, H. & Shi, Y. 2013b. L-glutamine provides acid resistance for *Escherichia coli* through enzymatic release of ammonia. *Cell Research,* 23**,** 635-644.
Macdonald, T. E., Helma, C. H., Shou, Y., Valdez, Y. E., Ticknor, L. O., Foley, B. T., Davis, S. W., Hannett, G. E., Kelly-Cirino, C. D., Barash, J. R., Arnon, S. S., Lindström, M., Korkeala, H., Smith, L. A., Smith, T. J. & Hill, K. K. 2011. Analysis of *Clostridium botulinum* serotype E strains by using multilocus

sequence typing, amplified fragment length polymorphism, variable-number tandem-repeat analysis, and botulinum neurotoxin gene sequencing. *Appl Environ Microbiol,* 77**,** 8625-8634.
Macián, M. C., Wolfgang, L., Chleifer, K., Garay, E. & Pujalte, M. J. 2000. *Vibrio pelagius*: Differences of the type strain deposited at various culture collections. *Syst Appl Microbiol,* 23**,** 373-375.
Maddox, I. S., Steiner, E., Hirsch, S., Wessner, S., Gutierrez, N. A., Gapes, J. R. & Schuster, K. C. 2000. The cause of "acid-crash" and "acidogenic fermentations" during the batch acetone-butanol-ethanol (ABE-) fermentation process. *J Mol Microbiol Biotechnol,* 2**,** 95-100.
Maiden, M. C. 2006. Multilocus sequence typing of bacteria. *Annu Rev Microbiol,* 60**,** 561-588.
Malm, S., Tiffert, Y., Micklinghoff, J., Schultze, S., Joost, I., Weber, I., Horst, S., Ackermann, B., Schmidt, M., Wohlleben, W., Ehlers, S., Geffers, R., Reuther, J. & Bange, F. C. 2009. The roles of the nitrate reductase NarGHJI, the nitrite reductase NirBD and the response regulator GlnR in nitrate assimilation of *Mycobacterium tuberculosis*. *Microbiology,* 155**,** 1332-1339.
Mann, M. S., Dragovic, Z., Schirrmacher, G. & Lütke-Eversloh, T. 2012. Over-expression of stress protein-encoding genes helps *Clostridium acetobutylicum* to rapidly adapt to butanol stress. *Biotechnol Lett,* 34**,** 1643-1649.
Mariano, A. P., Qureshi, N., Filho, R. M. & Ezeji, T. C. 2011. Bioproduction of butanol in bioreactors: New insights from simultaneous in situ butanol recovery to eliminate product toxicity. *Biotechnology and Bioengineering,* 108**,** 1757-1765.
Marini, A. M., Vissers, S., Urrestarazu, A. & André, B. 1994. Cloning and expression of the MEP1 gene encoding an ammonium transporter in *Saccharomyces cerevisiae*. *EMBO J,* 13**,** 3456-3463.
Martín, M. C., Fernández, M., Linares, D. M. & Alvarez, M. A. 2005. Sequencing, characterization and transcriptional analysis of the histidine decarboxylase operon of *Lactobacillus buchneri*. *Microbiology,* 151**,** 1219-1228.
McCoy, E., Fred, E. B., Peterson, W. H. & Hastings, E. G. 1926. A cultural study of the acetone butyl alcohol organism: One plate. *J Infect Dis,* 39**,** 457-483.
McCutchan, W. N., & R. J. Hickey. 1954. The butanol acetone fermentations. *Ind Ferment,* 1, 347-388.
Meier-Wagner, J., Nolden, L., Jakoby, M., Siewe, R., Krämer, R. & Burkovski, A. 2001. Multiplicity of ammonium uptake systems in *Corynebacterium glutamicum* role of Amt and AmtB. *Microbiology,* 147**,** 135-143.
Mermelstein, L. D., Welker, N. E., Petersen, D. J., Bennett, G. N. & Papoutsakis, E. T. 1994. Genetic and metabolic engineering of *Clostridium acetobutylicum* ATCC 824. *Ann N Y Acad Sci,* 721**,** 54-68.
Merrick, M. J. & Edwards, R. A. 1995. Nitrogen Control in Bacteria. *Microbiol Rev,* 59**,** 604–622.
Mitchell, W. J., Shaw, J. E. & Lesley Andrews, L. 1991. Properties of the glucose phosphotransferase system of *Clostridium acetobutylicum* NCIB 8052. *Appl Environ Microbiol,* 57**,** 2534-2539.
Molenaar, D., Bosscher, J. S., Ten Brink, B., Driessen, A. J. & Konings, W. N. 1993. Generation of a proton motive force by histidine decarboxylation and electrogenic histidine histamine antiport in *Lactobacillus buchneri*. *J Bacteriol,* 175**,** 2864-2870.
Morris, J. G. 1993. History and future potential of the clostridia in biotechnology. *The clostridia and biotechnology*. 1-23. Butterworth-Heineman. London.

Najafpour, G. D. & Shan, C. P. 2003. Enzymatic hydrolysis of molasses. *Bioresour Technol,* 86**,** 91-94.
Nakano, M. M., Hoffmann, T., Zhu, Y. & Jahn, D. 1998. Nitrogen and oxygen regulation of *Bacillus subtilis nasDEF* encoding NADH-dependent nitrite reductase by TnrA and ResDE. *J Bacteriol,* 180**,** 5344-5350.
Nakayama, S., Kosaka, T., Hirakawa, H., Matsuura, K., Yoshino, S. & Furukawa, K. 2008. Metabolic engineering for solvent productivity by downregulation of the hydrogenase gene cluster *hupCBA* in *Clostridium saccharoperbutylacetonicum* strain N1-4. *Appl Microbiol Biotechnol,* 78**,** 483-493.
Neumann, A. P. & Rehberger, T. G. 2009. MLST analysis reveals a highly conserved core genome among poultry isolates of *Clostridium septicum*. *Anaerobe,* 15**,** 99-106.
Ni, Y. & Sun, Z. 2009. Recent progress on industrial fermentative production of acetone-butanol-ethanol by *Clostridium acetobutylicum* in China. *Appl Microbiol Biotechnol,* 83**,** 415-423.
Ni, Y., Wang, Y. & Sun, Z. 2012. Butanol production from cane molasses by *Clostridium saccharobutylicum* DSM 13864: batch and semicontinuous fermentation. *Appl Biochem Biotechnol,* 166**,** 1896-1907.
Ni, Y., Xia, Z., Wang, Y. & Sun, Z. 2013. Continuous butanol fermentation from inexpensive sugar-based feedstocks by *Clostridium saccharobutylicum* DSM 13864. *Bioresour Technol,* 129**,** 680-685.
Nielsen, D. R. & Prather, K. J. 2009. In situ product recovery ofn-butanol using polymeric resins. *Biotechnology and Bioengineering,* 102**,** 811-821.
Nimcevic, D. & Gapes, J. R. 2000. The acetone-butanol fermentation in pilot plant and pre-industrial scale. *J Mol Microbiol Biotechnol,* 2**,** 15-20.
Nisman, B. 1954. The Stickland reaction. *Bacteriol Rev,* 18**,** 16-42.

Niwa, T., Kawamura, Y., Katagiri, Y. & Ezaki, T. 2005. Lytic enzyme, labiase for a broad range of Gram-positive bacteria and its application to analyze functional DNA/RNA. *J Microbiol Methods,* 61**,** 251-260.

Noguchi, T., Tashiro, Y., Yoshida, T., Zheng, J., Sakai, K. & Sonomoto, K. 2013. Efficient butanol production without carbon catabolite repression from mixed sugars with *Clostridium saccharoperbutylacetonicum* N1-4. *J Biosci Bioeng,* 116**,** 716-721.

Nohno, T., Saito, T. & Hong, J. S.1986. Cloning and complete nucleotide sequence of the *Escherichia coli* glutamine permease operon (*glnHPQ*). *Mol Gen Genet,* 205, 260-269.

Nomura, M., Nakajima, I., Fujita, Y., Kobayashi, M., Kimoto, H., Suzuki, I. & Aso, H. 1999. *Lactococcus lactis* contains only one glutamate decarboxylase gene. *Microbiology,* 145**,** 1375-1380.

O'Brien, R. W. & Morris, J. G. 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *J Gen Microbiol,* 68**,** 307-318.

Papoutsakis, E. T. 2008. Engineering solventogenic clostridia. *Curr Opin Biotechnol,* 19**,** 420-429.

Parekh, M. & Blaschek, H. P. 1999. Butanol production by hypersolvent-producing mutant *Clostridium beijerinckii* BA101 in corn steep water medium containing maltodextrin. *Biotechnol Lett,* 21**,** 45–48.

Perutka, J., Wang, W., Goerlitz, D. & Lambowitz, A. M. 2004. Use of computer-designed Group II Introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. *Journal of Molecular Biology,* 336**,** 421-439.

Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res,* 29**,** 2002-2007.

Picossi, S., Belitsky, B. R. & Sonenshein, A. L. 2007. Molecular mechanism of the regulation of *Bacillus subtilis gltAB* expression by GltC. *J Mol Biol,* 365**,** 1298-1313.

Quivey, R. G., Faustoferri, R., Monahan, K. & Marquis, R. 2000. Shifts in membrane fatty acid profiles associated with acid adaptation of *Streptococcus mutans*. *FEMS Microbiol Lett,* 189**,** 89-92.

Quixley, K. W. M. 1999. Studies on the regulation of nitrogen metabolism in *Clostridium acetobutylicum* NCP262 and *Clostridium beijerinckii* NCIMB 8052. MSc thesis, University of Cape Town, South Africa.

Qureshi, N., Annous, B. A., Ezeji, T. C., Karcher, P. & Maddox, I. S. 2005a. Biofilm reactors for industrial bioconversion processes: employing potential of enhanced reaction rates. *Microbial Cell Factories,* 4**,** 24.

Qureshi, N. & Blaschek, H. P. 2000. Economics of butanol fermentation using hyper-butanol producing *Clostridium beijerinckii* BA101. *Food and Bioproducts Processing,* 78**,** 139-144.

Qureshi, N., Hughes, S., Maddox, I. S. & Cotta, M. A. 2005b. Energy-efficient recovery of butanol from model solutions and fermentation broth by adsorption. *Bioprocess Biosyst Eng,* 27**,** 215-222.

Qureshi, N., Saha, B. C., Dien, B., Hector, R. E. & Cotta, M. A. 2010a. Production of butanol (a biofuel) from agricultural residues: Part I – Use of barley straw hydrolysate. *Biomass and Bioenergy,* 34**,** 559-565.

Qureshi, N., Saha, B. C., Hector, R. E., Dien, B., Hughes, S., Liu, S., Iten, L., Bowman, M. J., Sarath, G. & Cotta, M. A. 2010b. Production of butanol (a biofuel) from agricultural residues: Part II – Use of corn stover and switchgrass hydrolysates. *Biomass and Bioenergy,* 34**,** 566-571.

Ramos, J. L., Martínez-Bueno, M., Molina-Henares, A. J., Terán, W., Watanabe, K., Zhang, X., Gallegos, M. T., Brennan, R. & Tobes, R. 2005. The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev,* 69**,** 326-356.

Reid, S. J., Rafudeen, M. S. & Leat, N. G. 1999. The genes controlling sucrose utilization in *Clostridium beijerinckii* NCIMB 8052 constitute an operon. *Microbiol,* 145**,** 1461-1472.

Reid, S. J. & Stutz, H. E. 2005.Chapter 11: Nitrogen assimilation in clostridia. *Handbook on Clostridia,* 240–250. Taylor and Francis. Edited by Dürre, P.

Richter, H., Qureshi, N., Heger, S., Dien, B., Cotta, M. A. & Angenent, L. T. 2012. Prolonged conversion of n-butyrate to n-butanol with *Clostridium saccharoperbutylacetonicum* in a two-stage continuous culture with in-situ product removal. *Biotechnol Bioeng,* 109**,** 913-921.

Robinson, G. C. 1922. A study of the acetone and butyl alcohol fermentation of various carbohydrates. *J Biol Chem,* 53**,** 125-154.

Saad, N. Y., Schiel, B., Brayé, M., Heap, J. T., Minton, N. P., Dürre, P. & Becker, H. D. 2012. Riboswitch (T-box)-mediated control of tRNA-dependent amidation in *Clostridium acetobutylicum* rationalizes gene and pathway redundancy for asparagine and asparaginyl-tRNA^{asn} synthesis. *J Biol Chem,* 287**,** 20382-20394.

Sanders, J. W., Leenhouts, K., Burghoorn, J., Brands, J. R., Venema, G. & Kok, J. 1998. A chloride-inducible acid resistance mechanism in *Lactococcus lactis* and its regulation. *Mol Microbiol,* 27**,** 299-310.

Schreier, H. J. & Rostkowski, C. A. 1995. *Bacillus subtilis glnR* mutants defective in regulation. *Gene,* 161**,** 51-56.

Scotcher, M. C., Rudolph, F. B. & Bennett, G. N. 2005. Expression of *abrB310* and SinR, and effects of decreased *abrB310* expression on the transition from acidogenesis to solventogenesis, in *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol,* 71**,** 1987-1995.

Senadheera, D., Krastel, K., Mair, R., Persadmehr, A., Abranches, J., Burne, R. A. & Cvitkovitch, D. G. 2009. Inactivation of VicK affects acid production and acid survival of *Streptococcus mutans*. *J Bacteriol,* 191**,** 6415-6424.

Senouci-Rezkallah, K., Schmitt, P. & Jobin, M. P. 2011. Amino acids improve acid tolerance and internal pH maintenance in *Bacillus cereus* ATCC14579 strain. *Food Microbiol,* 28**,** 364-372.

Setlhaku, M., Heitmann, S., Górak, A. & Wichmann, R. 2013. Investigation of gas stripping and pervaporation for improved feasibility of two-stage butanol production process. *Bioresource Technology,* 136**,** 102-108.

Shaheen, R., Shirley, M. & Jones, D. T. 2000. Comparative fermentation studies of industrial strains belonging to four species of solvent-producing clostridia. *J Mol Microbiol Biotechnol,* 2**,** 115-124.

Shao, L., Hu, S., Yang, Y., Gu, Y., Chen, J., Jiang, W. & Yang, S. 2007. Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. *Cell Res,* 17**,** 963-965.

Shu, C. J. & Zhulin, I. B. 2002. ANTAR: an RNA-binding domain in transcription antitermination regulatory proteins. *Trends Biochem Sci,* 27**,** 3-5.

Sillers, R., Chow, A., Tracy, B. & Papoutsakis, E. T. 2008. Metabolic engineering of the non-sporulating, non-solventogenic *Clostridium acetobutylicum* strain M5 to produce butanol without acetone demonstrate the robustness of the acid-formation pathways and the importance of the electron balance. *Metab Eng,* 10**,** 321-332.

Sissons, C. H., Perinpanayagam, H. E. R., Hancock, E. M. & Cutress, T. W. 1990. pH Regulation of urease levels in *Streptococcus salivarius*. *Journal of Dental Research,* 69**,** 1131-1137.

Small, P. L. C. & Waterman, S. R. 1998. Acid stress, anaerobiosis and gadCB: lessons from *Lactococcus lactis* and *Escherichia coli*. *Trends Microbiol,* 6**,** 214-216.

Sola-Landa, A., Rodríguez-García, A., Amin, R., Wohlleben, W. & Martín, J. F. 2013. Competition between the GlnR and PhoP regulators for the *glnA* and *amtB* promoters in *Streptomyces coelicolor*. *Nucleic Acids Res,* 41**,** 1767-1782.

Sonenshein, A. L. 2007. Control of key metabolic intersections in *Bacillus subtilis*. *Nat Rev Microbiol,* 5**,** 917-927.

Stackebrandt, E., Kramer, I., Swiderski, J. & Hippe, H. 1999. Phylogenetic basis for a taxonomic dissection of the genus *Clostridium*. *FEMS Immunol Med Microbiol,* 24**,** 253-258.

Stickland, L. H. 1934. Studies in the metabolism of the strict anaerobes (genus *Clostridium*): The chemical reactions by which *Cl. sporogenes* obtains its energy. *Biochem J,* 28**,** 1746-1759.

Strösser, J., Lüdke, A., Schaffer, S., Krämer, R. & Burkovski, A. 2004. Regulation of GlnK activity: modification, membrane sequestration and proteolysis as regulatory principles in the network of nitrogen control in *Corynebacterium glutamicum*. *Mol Microbiol,* 54**,** 132-147.

Sturr, M. G. & Marquis, R. E. 1992. Comparative acid tolerances and inhibitor sensitivities of isolated F-ATPases of oral lactic acid bacteria. *Appl Environ Microbiol,* 58**,** 2287-2291.

Stutz, H. 2000. Regulation and characterization of the nitrogen assimilatory gene cluster in *Clostridium acetobutylicum* P262. *PhD thesis*. University of Cape Town, South Africa.

Stutz, H. E., Quixley, K. W., McMaster, L. D. & Reid, S. J. 2007. Co-regulation of the nitrogen-assimilatory gene cluster in *Clostridium saccharobutylicum*. *Microbiology,* 153**,** 3081-3090.

Sun, Y. & Cheng, J. 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresour Technol,* 83**,** 1-11.

Suzuki, T., Unemoto, T. & Kobayashi, H. 1988. Novel streptococcal mutants defective in the regulation of H^+-ATPase biosynthesis and in F_0 complex. *J Biol Chem,* 263**,** 11840-11843.

Syed, Q., Nadeem, M. & Nelofer, R. 2008. Enhanced butanol production by mutant strains of *Clostridium acetobutylicum* in molasses medium. *Turk J Biochem,* 33**,** 25-30.

Takagi, H., Takaoka, M., Kawaguchi, A. & Kubo, Y. 2005. Effect of L-proline on sake brewing and ethanol stress in *Saccharomyces cerevisiae*. *Appl Environ Microbiol,* 71**,** 8656-8662.

Takahashi, K. & Nei, M. 2000. Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. *Mol Biol Evol,* 17**,** 1251–1258.

Tamura, G. S., Nittayajarn, A. & Schoentag, D. L. 2002. A glutamine transport gene, *glnQ*, is required for fibronectin adherence and virulence of group B streptococci. *Infection and Immunity,* 70**,** 2877-2885.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol,* 28**,** 2731-2739.

Tangney, M. & Mitchell, W. J. 2000. Analysis of a catabolic operon for sucrose transport and metabolism in *Clostridium acetobutylicum* ATCC 824. *J Mol Microbiol Biotechnol,* 2**,** 71-80.
Tangney, M. & Mitchell, W. J. 2007. Characterisation of a glucose phosphotransferase system in *Clostridium acetobutylicum* ATCC 824. *Appl Microbiol Biotechnol,* 74**,** 398-405.
Tangney, M., Rousse, C., Yazdanian, M. & Mitchell, W. J. 1998. Note: Sucrose transport and metabolism in *Clostridium*. *J Appl Microbiol,* 84**,** 914-919.
Tashiro, Y., Takeda, K., Kobayashi, G. & Sonomoto, K. 2005. High production of acetone–butanol–ethanol with high cell density culture by cell-recycling and bleeding. *J Biotechnol,* 120**,** 197-206.
Thauer, R. K., Rupprecht, E., Ohrloff, C., Jungermann, K. & Decker, K. 1971. Regulation of the reduced nicotinamide adenine dinucleotide phosphate-ferredoxin reductase system in *Clostridium kluyveri*. *J Biol Chem,* 246**,** 954-959.
Thornton, J., Blakey, D., Scanlon, E. & Merrick, M. 2006. The ammonia channel protein AmtB from *Escherichia coli* is a polytopic membrane protein with a cleavable signal peptide. *FEMS Microbiol Lett,* 258**,** 114-120.
Tiffert, Y., Supra, P., Wurm, R., Wohlleben, W., Wagner, R. & Reuther, J. 2008. The *Streptomyces coelicolor* GlnR regulon: identification of new GlnR targets and evidence for a central role of GlnR in nitrogen metabolism in actinomycetes. *Mol Microbiol,* 67**,** 861-880.
Titgemeyer, F., Amon, J., Parche, S., Mahfoud, M., Bail, J., Schlicht, M., Rehm, N., Hillmann, D., Stephan, J., Walter, B., Burkovski, A. & Niederweis, M. 2007. A genomic view of sugar transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *J Bacteriol,* 189**,** 5903-5915.
Tomas, C. A., Welker, N. E. & Papoutsakis, E. T. 2003. Overexpression of *groESL* in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. *Applied and Environmental Microbiology,* 69**,** 4951-4965.
Tremblay, P. L. & Hallenbeck, P. C. 2009. Of blood, brains and bacteria, the Amt/Rh transporter family: emerging role of Amt as a unique microbial sensor. *Mol Microbiol,* 71**,** 12-22.
Tummala, S. B., Junne, S. G. & Papoutsakis, E. T. 2003a. Antisense RNA downregulation of coenzyme A transferase combined with alcohol-aldehyde dehydrogenase overexpression leads to predominantly alcohologenic *Clostridium acetobutylicum* fermentations. *Journal of Bacteriology,* 185**,** 3644-3653.
Tummala, S. B., Welker, N. E. & Papoutsakis, E. T. 2003b. Design of antisense RNA constructs for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*. *Journal of Bacteriology,* 185**,** 1923-1934.
Urwin, R. & Maiden, M. C. J. 2003. Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol,* 11**,** 479-487.
Usdin, K. P., Zappe, H., Jones, D. T. & Woods, D. R. 1986. Cloning, expression, and purification of glutamine synthetase from *Clostridium acetobutylicum*. *Appl Environ Microbiol,* 52**,** 413-419.
van Der Wal, H., Sperber, B. L., Houweling-Tan, B., Bakker, R. R., Brandenburg, W. & López-Contreras, A. M. 2013. Production of acetone, butanol, and ethanol from biomass of the green seaweed *Ulva lactuca*. *Bioresour Technol,* 128**,** 431-437.
van Dijck, P., Avesani, V. & Delmée, M. 1996. Genotyping of outbreak-related and sporadic isolates of *Clostridium difficile* belonging to serogroup C. *J Clin Microbiol,* 34**,** 3049–3055.
van Poelje, P. D. & Snell, E. E. 1990. Cloning, sequencing, expression, and site-directed mutagenesis of the gene from *Clostridium perfringens* encoding pyruvoyl-dependent histidine decarboxylase. *Biochemistry,* 29**,** 132-139.
Vanhee, L. M. E., Symoens, F., Jacobsen, M. D., Nelis, H. J. & Coenye, T. 2009. Comparison of multiple typing methods for *Aspergillus fumigatus*. *Clin Microbiol Infect,* 15**,** 643–650.
Varga, J., Stirewalt, V. L. & Melville, S. B. 2004. The CcpA protein is necessary for efficient sporulation and enterotoxin gene (*cpe*) regulation in *Clostridium perfringens*. *J Bacteriol,* 186**,** 5221-5229.
Villarreal, L., Heredia, N. L. & García, S. 2000. Changes in protein synthesis and acid tolerance in *Clostridium perfringens* type A in response to acid shock. *Int Microbiol,* 3, 113-116.

Walter, B., Hänssler, E., Kalinowski, J. & Burkovski, A. 2007. Nitrogen metabolism and nitrogen control in corynebacteria: variations of a common theme. *J Mol Microbiol Biotechnol,* 12**,** 131-138.
Wamelink, M. M., Struys, E. A. & Jakobs, C. 2008. The biochemistry, metabolism and inherited defects of the pentose phosphate pathway: a review. *J Inherit Metab Dis,* 31**,** 703-717.
Wang, Y. & Blaschek, H. P. 2011. Optimization of butanol production from tropical maize stalk juice by fermentation with *Clostridium beijerinckii* NCIMB 8052. *Bioresour Technol,* 102**,** 9985-9990.
Wang, Y., Li, X., Milne, C. B., Janssen, H., Lin, W., Phan, G., Hu, H., Jin, Y. S., Price, N. D. & Blaschek, H. P. 2013. Development of a gene knockout system using mobile group II introns (Targetron) and genetic disruption of acid production pathways in *Clostridium beijerinckii*. *Appl Environ Microbiol,* 79**,** 5853-5863.

Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. 1991. 16s *rRNA* ribosomal DNA amplification for phylogenetic study. *J Bacteriol,* 173**,** 697-703.
Weiss, V., Kramer, G., Dünnebier, T. & Flotho, A. 2002. Mechanism of regulation of the bifunctional histidine kinase NtrB in *Escherichia coli*. *J Mol Microbiol Biotechnol,* 4**,** 229-233.
Weisschuh, N., Fink, D., Vierling, S., Bibb, M. J., Wohlleben, W. & Engels, A. 2000. Transcriptional analysis of the gene for glutamine synthetase II and two upstream genes in *Streptomyces coelicolor* A3(2). *Mol Gen Genet,* 264**,** 461-469.
Weyer, E. R. & Rettger, L. F. 1927. A comparative study of six different strains of the organism commonly concerned in large-scale production of butyl alcohol and acetone by the biological process. *J Bacteriol,* 6**,** 399-424.
Wilkins, J. C., Homer, K. A. & Beighton, D. 2002. Analysis of *Streptococcus mutans* proteins modulated by clture under acidic conditions. *Applied and Environmental Microbiology,* 68**,** 2382-2390.
Wilkinson, S. R. & Young, M. 1995. Physical map of the *Clostridium beijerinckii* (formerly *Clostridium acetobutylicum*) NCIMB 8052 chromosome. *J Bacteriol,* 177**,** 439-448.
Willems, A., Hoste, B., Tang, J., Janssens, D. & Gillis, M. 2001. Differences between subcultures of the *Mesorhizobium loti* type strain from different culture collections. *Syst Appl Microbiol,* 24**,** 549-553.
Willey, J., Sherwood, L. & Woolverton, C. 2008. *Prescott, Harley and Klein's Microbiology*, 200–201. McGraw-Hill.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, A. & S. V., T. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res,* 18**,** 6531-6535.
Wray, L. V., Atkinson, M. R. & Fisher, S. H. 1991. Identification and cloning of the *glnR* locus, which is required for transcription of the glnA gene in *Streptomyces coelicolor* A3(2). *J Bacteriol,* 173**,** 7351-7360.
Wray, L. V., Ferson, A. E. & Fisher, S. H. 1997. Expression of the *Bacillus subtilis ureABC* operon is controlled by multiple regulatory factors including CodY, GlnR, TnrA, and Spo0H. *J Bacteriol,* 179**,** 5494-5501.
Wray, L. V., Ferson, A. E., Rohrer, K. & Fisher, S. H. 1996. TnrA, a transcription factor required for global nitrogen regulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A,* 93**,** 8841-8845.
Wray, L. V., Jr. & Fisher, S. H. 2008. *Bacillus subtilis* GlnR contains an autoinhibitory C-terminal domain required for the interaction with glutamine synthetase. *Mol Microbiol,* 68**,** 277-285.
Xiao, H., Gu, Y., Ning, Y., Yang, Y., Mitchell, W. J., Jiang, W. & Yang, S. 2011. Confirmation and elimination of xylose metabolism bottlenecks in glucose phosphoenolpyruvate-dependent phosphotransferase system-deficient *Clostridium acetobutylicum* for simultaneous utilization of glucose, xylose, and arabinose. *Appl Environ Microbiol,* 77**,** 7886-7895.
Xiao, H., Li, Z., Jiang, Y., Yang, Y., Jiang, W., Gu, Y. & Yang, S. 2012. Metabolic engineering of D-xylose pathway in *Clostridium beijerinckii* to optimize solvent production from xylose mother liquid. *Metab Eng,* 14**,** 569-578.
Xue, C., Zhao, J., Lu, C., Yang, S. T., Bai, F. & Tang, I. C. 2012. High-titer n-butanol production by *Clostridium acetobutylicum* JB200 in fed-batch fermentation with intermittent gas stripping. *Biotechnol Bioeng,* 109**,** 2746-2756.
Yang, S., Giannone, R. J., Dice, L., Yang, Z. K., Engle, N. L., Tschaplinski, T. J., Hettich, R. L. & Brown, S. D. 2012. *Clostridium thermocellum* ATCC27405 transcriptomic, metabolomic and proteomic profiles after ethanol stress. *BMC Genomics,* 13**,** 336.
Yang, X., Tu, M., Xie, R., Adhikari, S. & Tong, Z. 2013. A comparison of three pH control methods for revealing effects of undissociated butyric acid on specific butanol production rate in batch fermentation of *Clostridium acetobutylicum*. *AMB Express,* 3**,** 1-8.
Yu, J., Zhang, T., Zhong, J., Zhang, X. & Tan, T. 2012. Biorefinery of sweet sorghum stem. *Biotechnol Adv,* 30**,** 811-816.
Zhang, Y., Ma, Y., Yang, F. & Zhang, C. 2009. Continuous acetone-butanol-ethanol production by corn stalk immobilized cells. *J Ind Microbiol Biotechnol,* 36**,** 1117-1121.
Zhao, Y., Hindorff, L. A., Chuang, A., Monroe-Augustus, M., Lyristis, M., Harrison, M. L., Rudolph, F. B. & Bennett, G. N. 2003. Expression of a cloned cyclopropane fatty acid synthase gene reduces solvent formation in *Clostridium acetobutylicum* ATCC 824. *Applied and Environmental Microbiology,* 69**,** 2831-2841.
Zhao, Y., Tomas, C. A., Rudolph, F. B., Papoutsakis, E. T. & Bennett, G. N. 2005. Intracellular butyryl phosphate and acetyl phosphate concentrations in *Clostridium acetobutylicum* and their implications for solvent formation. *Appl Environ Microbiol,* 71**,** 530-537.
Zhou, A. & Thomson, E. 2009. The development of biofuels in Asia. *Applied Energy,* 86**,** S11-S20.

Zou, Y. 1998. Involvement of molecular chaperonins in nucleotide excision repair. Dnak leads to increased thermal stability of UvrA, catalytic UvrB loading, enhanced repair, and increased UV resistance. *Journal of Biological Chemistry,* 273**,** 12887-12892.
Zverlov, V. V., Berezina, O., Velikodvorskaya, G. A. & Schwarz, W. H. 2006. Bacterial acetone and butanol production by industrial fermentation in the Soviet Union: use of hydrolyzed agricultural waste for biorefinery. *Appl Microbiol Biotechnol,* 71**,** 587-597.