METABOLIC ENGINEERING OF *Clostridium acetobutylicum* AND PRODUCT EXTENSION

Bunmi Busola Omorotionmwan, MSc.

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

June, 2020

Declaration

Unless otherwise acknowledged, the work presented in this thesis is my own. No part has been submitted for another degree in the University of Nottingham or any other institute of learning.

Bunmi Busola Omorotionmwan

ABSTRACT

The limited abundance of fossil fuels and its harmful effects on the environment has led to the increased global interest in more environmentally friendly sources of chemicals and fuels. The solventogenic *Clostridium acetobutylicum* represents a model organism for solvent production in this genus and in the so-called ABE fermentation produces acetone, butanol and ethanol in a 3:6:1 ratio.

The identification of those factors that influence solvent production through mutation of the encoding genes is key to the rational metabolic engineering and development of an industrial solvent producing *C. acetobutylicum* strain. Consequently, an attempt to knock-out genes previously impossible to knock-out was carried out, using a conditional variant of the ClosTron intron re-targeting system. While the knock-out of hydrogenase (*hydA*), butyryl CoA dehydrogenase (*bcd*) and crotonase (*crt*) genes were not possible in this thesis, the isolation of a null 3-hydroxybutyryl CoA dehydrogenase (*hbd)* mutant was demonstrated.

To further enhance metabolic engineering of this organism, it was also desirable to put in place a system for the introduction of genes encoding pivotal metabolic enzymes at discrete loci around the chromosome, similar to *pyrE*, that may be used to implement new metabolic pathways and to exemplify its utility by extending product formation in *C. acetobutylicum* to isopropanol. This will include the introduction of an inducible orthogonal expression system at a locus, which would control the expression of additional gene sets introduced at the other loci. A triple auxotrophic mutant of *C. acetobutylicum* was created and the utility of the *pheA* and *argH* loci in addition to the *pyrE* locus described. Integration at the *pyrE* and *pheA* loci were relatively easily accomplished compared to the *argH* locus which proved difficult despite several attempts. With the introduction of a synthetic acetone operon and a gene encoding a secondary dehydrogenase from *Clostridium beijeirinckii* NRRL B593 both under the control of a lactose inducible *tcdR* system, up to 18.47 g/l Isopropanol-Butanol-Ethanol (8.47 g/l isopropanol, 10.02 g/l butanol, 0.21 g/l ethanol) was produced. This represents the highest recorded production of isopropanol from *C. acetobutylicum* to date.

To my Mummy…

ACKNOWLEDGEMENTS

I want to appreciate my supervisors, Nigel Minton and Ying Zhang for their expert advice, guidance and support during my PhD. Thank you for the great opportunity to undertake my PhD research within the Synthetic Biology Research Center (SBRC) here at Nottingham. I thank the Petroleum Technology Development Fund (PTDF) of the Nigeria Government for funding my PhD.

I thank Hengzheng Wang for her bench support especially in my early days in the lab. Special thanks to Matt Abbott for the HPLC and GC analyses and his willingness every time to be of help with the sample preparation. I appreciate Evangelos Karvonis, the lab technician, who was always willing to go the extra mile for our media preparation. Special thanks to Klaus Winzer for his valuable contribution during the final stages of my thesis write-up. I want to thank my lab colleagues in B27 for the companionship and enlightening scientific discussions. To Cynthia, thank you for making lab work more fun! To my office mates in C02, thank you for your friendship and support. To all my friends and colleagues in the Clostridia Research Group (CRG), thank you for making my PhD period a very memorable one. I want to specially appreciate the Fadipes and Onuwes for the gift of friendship and support system without which I don't know how I would have survived these years.

To the loving memory of my late parents, Major T. A. Banire and D'ness B. Banire, I owe what I am today to your love and guidance through the years. This thesis is for you and I wish you both were alive to see this time. I want to appreciate my siblings, Seye, Seun and Toyin, your emotional support throughout my PhD gave me stability.

To my Son, Pneuma, thank you for understanding those times I couldn't be there, your smile was my strength after long lab days. To my husband, Craig, you know all of this could not have been possible without you, thank you for being my pillar and source of daily encouragement, these 4 years and for never letting me give up. Thank you!

My ultimate thanks is to the Lord, God Almighty to whom I owe my strength, joy and existence.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

2xYTG	Two-times Yeast-Tryptone-Glucose medium
5-FC	5-Fluorocytosine
5-FU	5-Fluorouracil
5- FOA	5-Fluorooratic acid
ACE	Allele-Coupled Exchange
Arg	Arginine
ATCC	American Type Culture Collection
BLAST	Basic Local Alignment Search Tool
bp	Base Pair
Cas	CRISPR associated
CAT	Chloramphenicol acetyl transferase
CBMS	Supplemented Clostridium Basal Broth
CGM	Clostridial Growth Medium
CRG	Clostridia Research Group
CRISPR	Clustered Regularly Interspaced Palindromic Repeats
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
EPB	Electroporation Buffer
Erm	Erythromycin
FOA	5-fluorooratic acid
g	Gram
InDel	Insertion/Deletion
kb	Kilobase Pair
KO	Knock-out
kV	kilovolt
L	Microlitre
LB	Luria-Bertani medium
LHA	Left Homology Arm
mg	Milligram

Ml	Millilitre
mM	Millimolar
NADH	nicotinamide adenine dinucleotide, reduced
NADPH	nicotinamide adenine dinucleotide phosphate, reduced
NEB	New England Biolabs
NGS	Next Generation Sequencing
OD	Optical Density
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
Phe	Phenylalanine
RBS	Ribosome Binding Site
RHA	Right Homology Arm
RNA	Ribonucleic Acid
RNAP	Ribonucleic Acid Polymerase
rpm	Revolutions Per Minute
sgRNA	Single guide RNA
SNP	Single Nucleotide Polymorphism
SOE	Splicing by Overlap Extension
TAE	Tris Base-Acetic Acid-EDTA Buffer
Th	Thiamphenicol
Ur	Uracil
UV	Ultraviolet
V	Volt
w/v	weight per volume
WT	Wild Type
µg	Microgram
%	Percentage
˚C	Degree Celsius

CHAPTER 1

General Introduction

CHAPTER 1

INTRODUCTION

Global energy demand and over-dependence on fossil fuels have contributed greatly to global warming and climate change. Climate change is an ever-growing environmental problem and arguably one of the greatest threats facing the planet. The signing of the agreement to lower greenhouse gas (GHG) emissions towards reducing the catastrophic impacts of climate change witnessed the largest number of countries ever (175) to sign an international agreement in a single day (Moon, 2016). Continued anthropogenic activities such as burning of fossil fuels lead to massive GHG emissions, especially CO_2, which cause excessive warming of the earth. Consequently, there has been an increased global interest in the development of renewable energy sources which are sustainable.

1.1 Fossil Fuel

Fossil fuels are hydrocarbons derived from buried dead plants and animals. They include coal, oil and natural gas and are formed by natural processes such as anaerobic decomposition and crustal movement over millions of years (Hsu, 2017). Fossil fuels contain high concentration of carbon which release energy when burnt. They can be refined into petrochemical products such as gasoline, diesel, lubricants, wax, asphalt and jet fuel. However, they are non-renewable and their continued use is not sustainable. In 2018, there was an estimated 1.7%

increase in carbon dioxide emissions related to the rise in consumption of fossil fuels (REN, 2019).

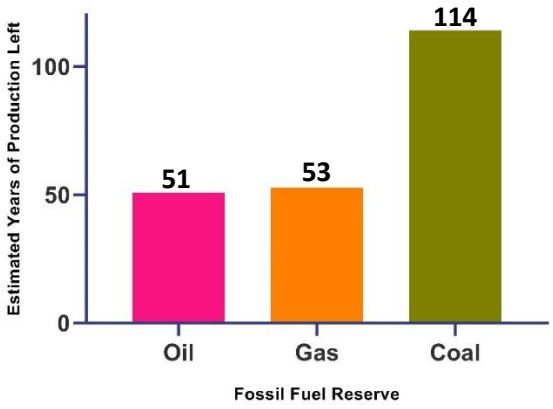


Figure 1. 1: Estimated Years of Fossil Fuel (Oil, Gas and Coal) Production Left.

Figure adapted from BP Statistical Review of World Energy, 2016 (BP, 2016)

Figure 1.1 gives the remaining years of global coal, oil and natural gas, estimated based on the reserves-to-product (R/P) ratio that measures the number of years of production left based on current reserves and annual production levels in 2015 (BP, 2016). It is worth noting that these values can change over time based on discovery of new fossil fuel reserves and alterations in yearly production.

1.2 Renewable Energy

Renewable energy sources are sustainable alternatives to fossil fuels and they include wind, solar, hydro, marine and biomass. The consumption of renewable energy has more than doubled between 2004 and 2017 in the EU; from about 8% in 2004 to about 18% of total final energy consumption (TFEC) in 2017 (Eurostat, 2019). Energy consumed in the transport sector make up about one-third of TFEC globally (REN, 2019) with the majority of the energy demand met by oil and petroleum products.

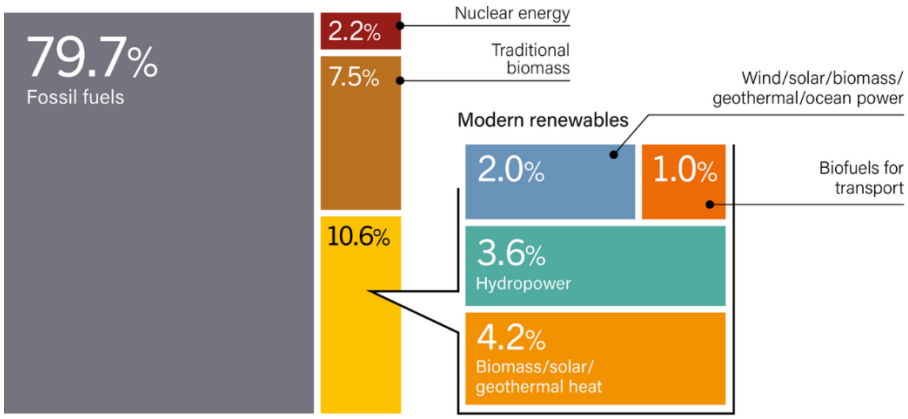


Figure 1. 2: Estimated Renewable Energy Share of Final Energy Consumption, 2017. Figure taken from the REN 21 Renewables 2019 Global Status Report (REN, 2019).

1.2.1 Biofuels

Biofuels derived from biomass provide an abundant and renewable source of energy with little or no emission of greenhouse gases (Durre, 2007) and include biodiesel, bioethanol and biobutanol. Typically, biofuels are blended in varying proportions with conventional transport fuels (Rodríguez-Fernández *et al.*, 2019). However, the development of biofuels has been hindered by a number of factors including feedstock uncertainties and the slow progress in development of biofuels for the aviation industry (REN, 2019) creating the need for new technologies aimed at developing biofuels to become more competitive in the market place.

Compared with ethanol, butanol has significantly improved properties such as higher energy content, less volatility and is safer to use, less corrosive, non-hygroscopic hence not prone to contamination, it can be used up to 100% in already existing automobile engines without prior modification and has potentials for use as aviation fuel (Lee *et al.*, 2008; Durre, 2007)

Table 1. 1 Comparison of Chemical and Physical Properties of Butanol and Other Fuels

Fuel	Energy Density (MJ/ L)	Air-to-Fuel Ratio	Research Octane Number	Water Solubility (%)
Gasoline	32	14.6	81-89	Negligible
Diesel	35.5	14.7	Nd	Negligible
Butanol-1	29.2	11.12	78	7
Ethanol	19.6	8.94	96	100

Source: Adapted from: Kolesinska *et al.*, 2019 *nd—no data

1.2.1.1 Butanol from Chemical Synthesis

There are three main ways of producing butanol chemically – Hydroformylation, Reppe synthesis and crotonyldehyde hydrogenation (figure 1.3). Hydroformylation is a two-step process that involves the addition of CO and H_2 to a carbon-carbon double bond producing aldehyde first and then hydrogenated to produce butanol (Kazemi Shariat Panahi *et al*., 2019). In the Reppe method, butanol is made directly by the reaction of propylene, CO and H_2O in the presence of a catalyst at low pressure and temperature. This process is however very expensive and not commercially successful (Albert Cotton, 1999). These first two processes rely totally on petroleum. The third route is the production of butanol from acetaldehyde by crotonyladehyde hydrogenation and it involves aldol condensation, dehydration and hydrogenation. (Kazemi Shariat Panahi *et al*., 2019). This method allows the option of using ethanol from biomass, which can be dehydrogenated to yield acetaldehyde.

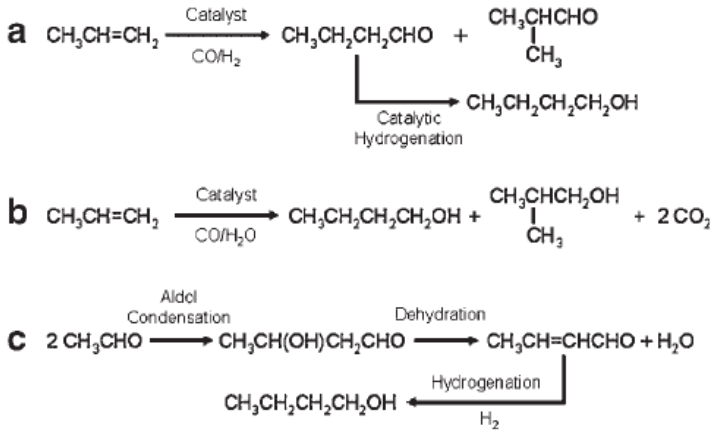


Figure 1. 3: Butanol from chemical synthesis

(a) Hydroformylation (b) Reppe process (c) crotonaldehyde hydrogenation. Figure taken from (Kazemi Shariat Panahi *et al*., 2019)

1.3 The Genus Clostridia

The genus *Clostridium* consists of Gram-positive, obligate anaerobic, endospore-forming rod-shaped bacteria and belong to the phylum Firmicutes (Durre, 2005). Although some species are notorious human and animal pathogens (such as *Clostridium difficile*, *Clostridium botulinum*, *Clostridium tetani* and *Clostridium perfringens*), the majority of them are innocuous and are useful in industrial biotransformations (include solventogenic species such as *Clostridium acetobutylicum*, *Clostridium saccharoperbutylacetonicum* and *Clostridium beijerinckii*) (Durre, 2008).

1.3.1 *Clostridium acetobutylicum*

C. acetobutylicum is a solventogenic Clostridia and represents the model organism for solvent production in this genus. It was the first sequenced solventogenic *Clostridium* (Nolling *et al*., 2001). The genome consists of a

3,940,880 bp chromosome and a 192,000 bp megaplasmid. There are 3,740 and 178 ORFs that have been identified in the chromosome and megaplasmid, respectively (Nolling *et al*., 2001). It is able to utilize a wide variety of substrates including monosacchraides and polysaccharides and produce the solvents, acetone, butanol and ethanol in the so called ABE fermentation (Jones and Woods, 1986)

1.3.2 The Acetone-Butanol-Ethanol (ABE) Fermentation by *C. acetobutylicum* - Brief History

The ABE fermentation discovered in 1861 by Louise Pasteur became popular in the 20th century and was the second largest industrial fermentation after ethanol. The increased demand for acetone for the production of cordite in England during the World War I (1914–18) necessitated the development of the ABE fermentation and in 1916 the first industrial scale ABE fermentation began based on the Weizmann *C. acetobutylicum* strain. 2016 marked the 100th anniversary of the first industrial operation of the Chaim Weizmann's ABE fermentation process. Within the first two years of establishment, 3,000 tons of acetone and 6,000 tons of butanol were produced (Moon *et al.*, 2016, Jones and Woods, 1986; Durre, 2007) However, by the 1960s the ABE production became less competitive due to the advancement of the petrochemical industry and increase in cost of feedstock (Bankar *et al*., 2013; Durre, 2007).

However, there has been a renewed interest in the production of biobutanol due to concerns about global warming and increase in oil prices. Butanol is not just a useful chemical used in the production of adhesives, detergents, surface cleaners, paint thinners and printing ink, it is importantly an alternative biofuel

and over the years several fermentative processes have evolved to meet the increasing demand for butanol (Sauer, 2016, Lee *et al*.; 2008, Durre, 2007; Moon *et al*., 2016).

In the ABE process, *C. acetobutylicum* breaks down sugar to produce the solvents acetone, butanol and ethanol in the ratio 3:6:1 (Jones and Woods, 1986). However, there are a number of limiting factors to the economic production of butanol, these include high substrate and product recovery costs, product toxicity and low product yield.

The cost of substrate including sourcing, availability and pre-treatment make up about 60 % of the overall production cost of the ABE fermentation (Papoutsakis, 2015; Jones and Woods, 1986). The costs of separation of these organic molecules from aqueous solution is quite high and it is estimated that the production economics of biofuels has to increase by 2 to 4-fold to be competitive in the market (Papoutsakis, 2015).

Longer chain alcohols are generally more toxic than short chain alcohols with butanol more toxic than ethanol. The antimicrobial effect of a solvent is directly related to its hydrophobicity. Alcohols increases the membrane permeability, impedes membrane protein function and ATP synthesis active nutrient transport, the membrane-bound ATPase activity, and glucose uptake (Dunlop, 2011). The butanol tolerance level for *C. acetobutylicum* is about 2% and produces an average of 13 g/l (Jones and Woods, 1986). Though butanol tolerance levels can be exceeded by production yield (Harris *et al*., 2000) as production continues even after growth ends, butanol toxicity is still a major economic problem with ABE fermentation.

1.3.3 The *C. acetobutylicum* ABE Metabolism

C. acetobutylicum is able to utilize a wide range of substrate including pentoses, hexoses, oligosaccharides and polysaccharides for metabolism (Jones and Woods, 1986; Lutke-Eversloh and Bahl, 2011)

The ABE fermentation can be typically divided into two phases – the acidogenic and the solventogenic phases. During the first phase, the cells are growing exponentially and carboxylic acids, acetate and butyrate are produced consequently, there is a drop in pH which could affect the proton gradient across membrane (Lütke-Eversloh and Bahl, 2011). These acids have been suggested to be inducers for the solventogenic phase and are taken up by cells as co-substrates for the production of the solvents, acetone, butanol and ethanol. The switch to solvent production is an adaptive mechanism to the low pH due to acid production.

Pyruvate is a key intermediary in the metabolism of Clostridia. Glucose is broken down to pyruvate via the Embden-Meyerhof-Panas pathway and the enzyme pyruvate:ferredoxin oxidoreductase, in the presence of coenzyme-A, cleaves it to acetyl-CoA, CO_2 and reduced ferredoxin. Acetyl-CoA can then be either oxidised to acetone, acetate and CO_2 or reduced to butanol, ethanol and butyrate (Gheshlaghi *et al*., 2009).

Electron transfer is controlled by NADH-ferredoxin oxidoreductase, NADPH-ferredoxin oxidoreductase and hydrogenase and these determine the nature of the fermentation. Reduced ferredoxin can either transfer electrons to protons via hydrogenase to form hydrogen or via NAD(P)H-ferredoxin oxidoreductase to form the reduced NAD(P)H. During acidogenesis, hydrogen is produced

however during solventogenesis, there is shift to NAD(P)H production towards the production of solvents.

1.3.3.1 Acid Forming Pathways

During acidogenesis, acetate and butyrate are produced from acetyl-coA and butyryl-CoA respectively. Phosphotransacetylase (*pta*) and acetate kinase (*ack*) are the enzymes responsible for the formation of acetate with the production of 2 mols of ATP. For the synthesis of butyrate, 2 molecules of acetyl-CoA are condensed to acetoacteyl-CoA which is further reduced to butyryl-CoA and converted to butyrate through phosphotransbutyrylase (*ptb*) and butyrate kinase (*buk*) with the generation of 1 mole of ATP (Jones and Woods, 1986).

The lactic acid pathway sometimes becomes active as a less efficient alternative to regenerate NADH under certain growth conditions such as iron limitation (Bahl *et al*., 1986) or inhibition of hydrogenase by carbon monoxide (Kim *et al.*, 1984). Figure 1.4 shows the metabolic pathway in *C. acetobutylicum.* The products from acidogenesis are given in grey boxes while the products from solventogenesis are given in black boxes.

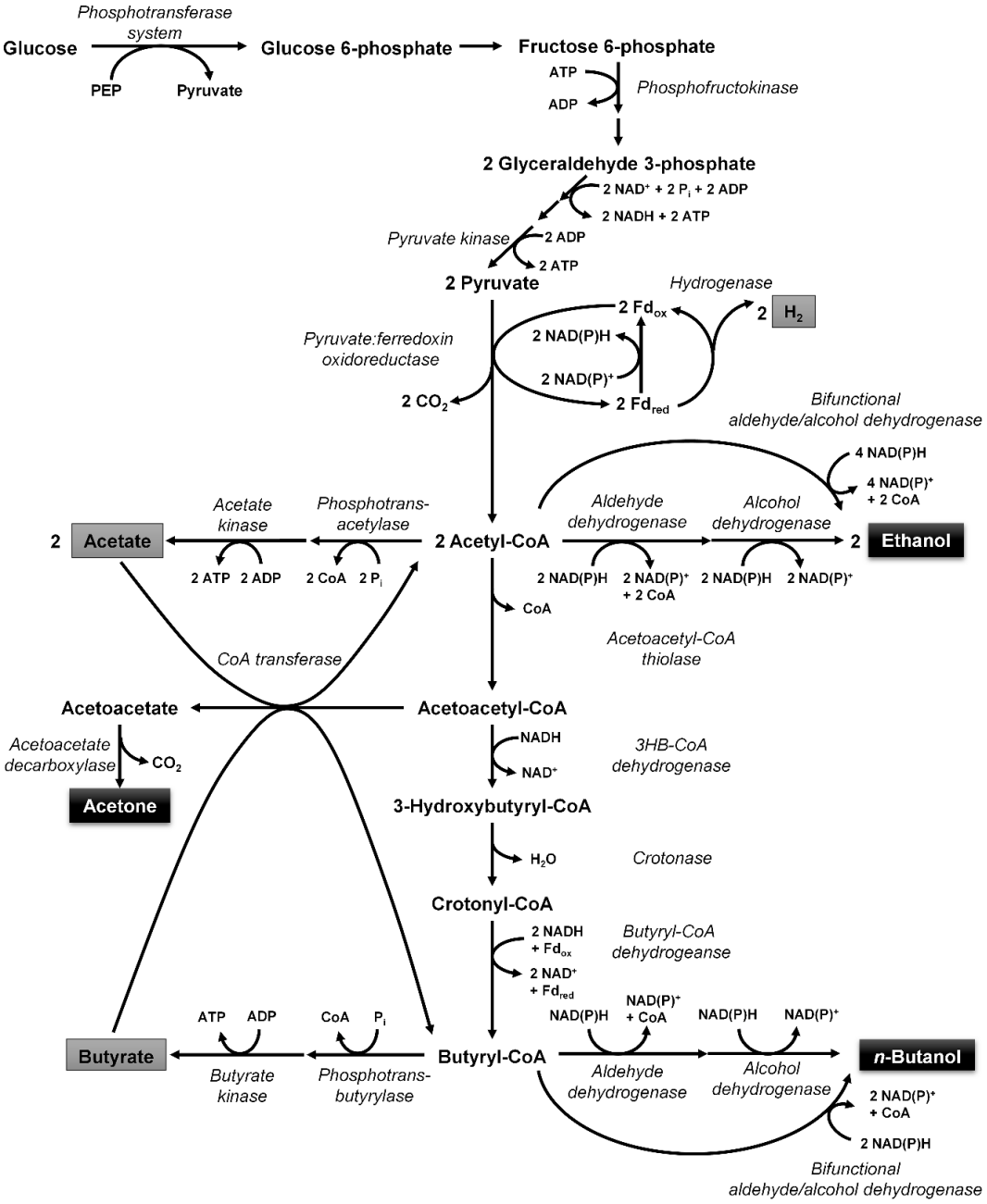


Figure 1. 4: The metabolic pathways in *C. acetobutylicum* ABE fermentation.

Products in grey boxes (acetate, butyrate and H_2) are made during acidogenesis while those in black boxes (acetone, ethanol and butanol) are produced during solventogenesis. Figure taken from (Moon *et al.*, 2016)

1.3.3.2 Solvent Forming Pathways

With the onset of solventogenesis, there is the re-assimilation of acetate and butyrate and they are metabolized concomitantly with glucose. The acid uptake is via the acetoacetyl-CoA:acyl-CoA transferaseA/B (*ctfA/B*) with acetoacetyl-CoA serving as the CoA donor (Lütke-Eversloh and Bahl, 2011) and this re-assimilation has been directly coupled with the formation of acetone (Jones and Woods, 1986). Acetone is formed in two steps from acetoacetyl-CoA; first is the removal of CoA by *ctfA/B* yielding acetoacetate and the subsequent decarboxylation by acetoacetate decarboxylase (*adc*) to form acetone.

The solvent formation genes of *C. acetobutylicum* are present on the pSOL1 mega plasmid. The *adhE1, ctfA* and *ctfB* make up the *sol*operon (figure 1.5) and are transcribed as a tricistron while the *adc* gene is contiguously located and transcribed in the opposite direction to the *sol* operon (Cornillot *et al.*, 1997).

Figure 1. 5: The *sol* operon gene locus.

adhE1 is the bifunctional aldehyde-alcohol dehydrogenase; *ctfA* and *ctfB* are two coenzyme A transferase subunits and *adc* codes for the acetoacetate decarboxylase.

The *adhE2* also resides in the pSOL1 plasmid but is not clustered with *adhE1* as observed in *C. ljungdahlii* (Leang *et al.*, 2013) and their expressions also differ (Yoo *et al.*, 2015; Fontaine *et al.*, 2002)

Three metabolic states have been identified in *C acetobutylicum;* acidogenesis (at neutral pH, the production of acetate and butyrate), solventogenesis (at low pH, the production of acetone, ethanol and butanol) and alcohologenesis (at neutral pH and high NAD(P)H, the production of ethanol and butanol) (Yoo *et al*., 2015, Girbal *et al*., 1995). Under solventogenesis, the *adhE1* is primarily responsible for the butyraldehyde dehydrogenase activity while *bdhA, bdhB* and *bdhC* are responsible for butanol dehydrogenase activity. However. during alcohologenesis, *adhE2* is largely responsible for both butyraldehyde dehydrogenase and butanol dehydrogenase activities (Yoo *et al.*, 2015).

Both solventogenesis and sporulation are regulated by the master regulator, Spo0A (Alsaker *et al*., 2004). Typically, sporulation occurs as cells enter into stationary phase. Deletion of the encoding gene, *spo0A*, results in a mutant that is deficient in solvent production and unable to septate. When *spo0A* was overexpressed, solvent genes were overexpressed but increased solvent production was not observed due to expedited and enhanced sporulation (Harris *et al*., 2002). The transition from acidogenesis to solventogenesis is a complex process involving genetic regulation, metabolic and physiological shift, cellular signal integration and sporulation (Liao *et al.*, 2015, Jones *et al*., 2008). During this switch, there is the up-regulation of the *sol* operon genes, *adhE1, ctfA, ctfB* and *adc* and stress genes (Alsaker *et al*., 2010, Alsaker and Papoutsakis, 2005) and it could be influenced by redox balance and environmental factors (Wietzke and Bahl, 2012, Ezeji *et al*., 2010). The transition to solventogenesis from acidogenesis is complicated and at least 245 genes have been shown by microarray to be differentially expressed during this switch (Grimmler *et al*., 2011) but the detailed mechanism is yet to be well understood.

1.3.3.3 Rex transcriptional regulation

In *C. acetobutylicum*, the redox sensing transcriptional repressor (Rex) has been found to modulate its DNA-binding activity in response to the NADH/NAD+ ratio and is involved in the redox-dependent solventogenic shift in this organism (Zhang et al., 2014, Wietzke and Bahl, 2012).

This Rex ORF is located upstream of the crt-bcd-etfAB-hbd operon encoding the enzymes responsible for the conversion of acetoacetyl-CoA to butyryl-CoA. There appears to be a general relation of Rex to the C2/C4 clostridial metabolic pathway as similar gene arrangements have been identified in others solventogenic clostridia (Wietzke and Bahl, 2012). Rex boxes are found in the 5′UTR of the crt-bcd-etfAB-hbd operon and the thiolase gene as well as in the promoter region of the bifunctional aldehyde/alcohol dehydrogenase 2 (*adhE2*) gene and lactic acid dehydrogenase (*ldh*).

A de-repression of the adhE2 gene with increase in the expression of the *adhE2* gene and NADH-dependent alcohol dehydrogenase activities were demonstrated in a *C. acetobutylicum* rex negative mutant (Wietzke and Bahl, 2012). The Rex-regulated expression of the NADH-consuming enzymes in response to increased NAD(P)H availability help maintain redox homeostasis in the cell (Zhang et al., 2014). Additionally, a de-repression of Rex and prevention of the formation of Rex-DNA complex was observed with the addition of an increasing concentration of NADH of up to 5 mM. In contrast, the Rex binding activity was restored with the addition of 5-fold excess of oxidised NAD+ indicative of Rex metabolic gene expression by Rex based on the NADH/NAD+ ratio (Wietzke and Bahl, 2012).

In more recent studies, when the *rexA* gene was deleted in the *C. acetobutylicum* mutant, CAB1057, a significant increase in the yield and flux of alcohol formation was noted in addition to a four-fold decrease of the hydrogen yield and fluxes, associated with formation of NADH from reduced ferredoxin (Nguyen *et al*., 2018).

1.3.3.4 The Central Metabolic Pathway

The central metabolic ABE pathway from acetyl-CoA consists of thiolase (Thl, encoded by *thl*) which catalyses the condensation of two molecules of acetyl-CoA to acetoacetyl-CoA, the precursor of 4-carbon products. It also catalyses the reverse exergonic thiolytic cleavage (Mann and Lütke-Eversloh, 2013). In the next three steps butyryl-CoA is formed. The next enzyme after Thl is the NADH dependent 3–hydroxybutyryl-CoA dehydrogenase (Hbd, encoded by *hbd*), which catalyses the reduction of acetoacetyl-CoA to 3-hydroxybutyryl-CoA (Gheshlaghi *et al*., 2009). Crotonase (Crt, encoded by *crt*), otherwise known as enoyl-CoA hydratase, catalyses the dehydration of 3-hydroxybutyryl-CoA to crotonyl-CoA (Seedorf *et al*., 2008). Butyryl-CoA dehydrogenase (Bcd, encoded by *bcd*) is responsible for the reduction of crotonyl-CoA to butyryl-CoA in the presence of NADH. Bcd is similar to the eukaryotic acyl-CoA dehydrogenase which catalyses the oxidation of acyl-CoA into enoyl-CoA but functions in the reverse direction.

1.3.3.5 Regulation of Electron Flow in *C. acetobutylicum*

During glycolysis, there is the generation of less ATP and more of the NAD(P)H required for growth. Additional ATP is generated during acidogenesis, however, only a portion of the NAD(P)H is used up and these excess electrons are disposed of through the enzyme hydrogenase (HydA, encoded by *hydA*) using protons as terminal electron acceptors (Jones and Woods, 1986). Although the formation of acetate produces more ATP than the production of butyrate, there is a net generation of NADH as no NADH is consumed in acetate production, however, the production of butyrate uses up NADH, hence is redox neutral. The NADH ferredoxin oxidoreductase catalyses the transfer of electrons from NADH to reduced ferredoxin which then produces hydrogen through hydrogenase (Jones and Woods, 1986). Hydrogen production has been a target in the metabolic engineering of *C. acetobutylicum* for enhanced butanol production (Nguyen, 2016; Lütke-Eversloh and Bahl, 2011). The acetate/butyrate ratio can be regulated by the NADH ferredoxin oxidoreductase and hydrogenase complex where in the case of enzyme inhibition, there is the excess of NADH which for redox balance must be utilized in the production of butyrate. During solventogenesis, there is a reduction in hydrogenase activity and carbon and electron flow is directed to the production of solvents which are more reduced products. The NADPH ferredoxin oxidoreductase mediates the formation of NADPH from reduced ferredoxin (Jones and Woods, 1986).

More recently, the electron bifurcation has been described as a third type of energy conservation (Seedorf *et al.*, 2008). Here the endergonic reduction of the low potential ferredoxin (-500 mV) by the high potential NADH (-280 mV) is coupled with the exergonic reaction from crotonyl-CoA (- 10mV) to butyryl-

CoA. This coupling of reactions is achieved by the enzyme complex, butyryl-CoA dehydrogenase/electron transfer flavoprotein (Bcd/Etf) (Li *et al.*, 2008; Buckel and Thauer, 2013). The RNF (*Rhodobacter* nitrogen fixation) complex plays a major part in energy conservation by coupling ferredoxin oxidation to reduced NADH with the pumping of H+/Na+ across the cytoplasmic membrane (Buckel and Thauer, 2013; Buckel and Thauer, 2018). This was first discovered in *C. kluyveri* (Seedorf *et al.*, 2008). Although the Bcd/Etf complex is present in *C. acetobutylicum*, Rnf-homologous genes have not been identified in this organism nor in *C. pasteurianum* (Lütke-Eversloh and Bahl, 2011), consequently, it could be inferred that the reaction is not coupled with electron transport across membrane.

1.4 Fermentation and Downstream Purification

A major advantage of solventogenic clostridia is the ability to utilize a wide range of substrates. Various alternative substrates have been used for solvent production by solventogenic Clostridia and include Jerusalem artichokes (Vandecasteele, 1985), maize mash (Shaheen *et al.*, 2000), cassava (Lu *et al.*, 2012), microalgae (Wang *et al.*, 2016), cane molasses (Shaheen *et al.*, 2000), barley straw (Qureshi *et al.*, 2010). Fermentation can be carried out via batch, fed-batch or continuous processes. While the batch fermentation is easy to set up, and does not require much set up costs and control, continuous fermentation may offer superior productivity, reduction in butanol inhibition and less time for re-inoculation and sterilization (Vandecasteele, 1985; Ezeji *et al.*, 2007).

The *in situ* product recovery (ISPR) has been useful to improve product recovery and reduce the toxic effect of butanol (Cheng *et al.*, 2019). ISPR can be done using gas stripping, pervaporation, adsorption or liquid-liquid extraction (Xue *et*

al., 2017). For the production of ABE from lignocellulose, various processes could be implemented such as separate hydrolysis and fermentation (SHF), simultaneous saccharification and fermentation (SSF), simultaneous saccharification and co-fermentation (SSCF), and consolidated bioprocessing (CBP). The CBP seems to be the most economic and preferred one because it combines enzyme production, hydrolysis of cellulose and fermentation in one in just one process (Cheng *et al*., 2019). Figure 1.6 illustrates these various production processes from lignocellulose.

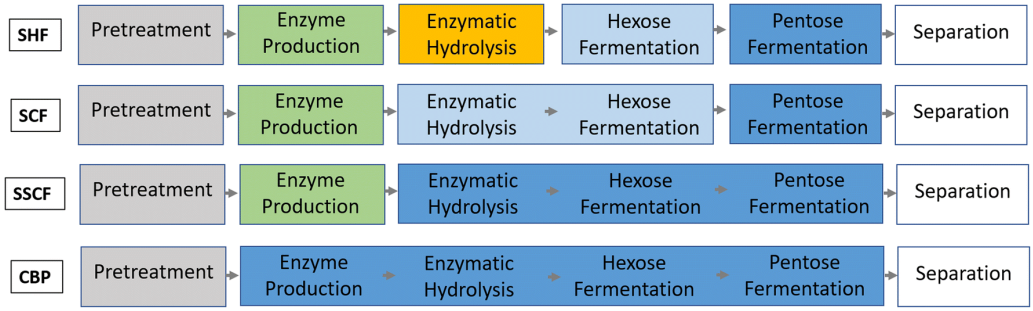


Figure 1. 6: Types of bioprocessing for solvent production from lignocellulose.

Separate Hydrolysis and Fermentation (SHF), Simultaneous Saccharification and Fermentation (SSF), Simultaneous Saccharification and Co-fermentation (SSCF), Consolidated Bioprocessing (CBP). Figure taken from Cheng *et al*., 2019.

1.5 Development of Genetic Tools in clostridia Engineering

In the past, research in clostridia and manipulation of its metabolism has been limited by lack of efficient genetic tools. As a result, most directed mutagenesis has been by homologous recombination resulting in single cross-over integrants or double cross-over mutants. The former is highly unstable while the latter is usually very difficult to isolate. Over the years however there has been the

development of more systems towards improving metabolic engineering of clostridia and some of these are discussed hereafter.

1.5.1 Shuttle Vectors

A shuttle vector is a plasmid that can replicate and be manipulated in two different hosts species for example, *Clostridium* and *Escherichia coli*. It basically consists of a selectable marker and origins of replication functional in both hosts. The advantage of the shuttle vector is that a plasmid can be manipulated in an *E. coli* host and the recombinant plasmid transformed into the *C. acetobutylicum* which is more difficult to use. A number of Clostridia vectors have been described (Minton *et al*., 1993; Purdy *et al*., 2002; Heap *et al*., 2009).

C. acetobutylicum ATCC 824 possesses Cac8241, a powerful restriction system which recognises the sequence 5'-GCNGC-3'. Shuttle vectors can be methylated to prevent degradation by this system. Plasmids can be methylated in *E coli* containing pAN2 prior to transformation into *Clostridium* (Heap *et al*., 2009).

1.5.2 Antisense RNA

Antisense RNA (asRNA) are small molecules which can regulate gene expression by hybridizing with complementary mRNA transcripts (Desai and Papoutsakis, 1999). asRNA forms a duplex with mRNA and prevents access to ribosome binding site, duplex RNA-specific RNases could also degrade the duplex mRNA formed. Some asRNAs could be naturally occurring (Saberi *et al*., 2016), the synthesis of glutamine synthetase (*glnA*) of *Clostridium* sp. strain NCP262 (formerly *C. acetobutylicum* P262) enzyme can be regulated by naturally occurring asRNA molecules (Ellison *et al*., 1985). An advantage of the

asRNA is that it usually does not involve complete inhibition of protein synthesis so may not result in lethal mutations. (Desai and Papoutsakis, 1999).

The use of asRNA to regulate the *buk* and *ptb* genes (Desai and Papoutsakis, 1999), as well as downregulation of the acetone formation pathway (Tummala *et al*., 2003a; Tummala *et al*., 2003b), has been described.

1.5.3 ClosTron

The ClosTron is an insertional mutagenesis system based on the mobile group II intron from *Lactococcus lactis* (Ll.ltrb). This system was first described by Zhong *et al.* (2003) and the retargeted group II intron was called TargeTron. Heap and co-workers adapted this system for *Clostridium* and called it the ClosTron (Heap *et al*., 2007). Here, desired changes are made to the group II intron which makes it insert preferentially to the DNA region of interest. There are four basic steps in the ClosTron mutagenesis: design of intron, construction of plasmid, transfer of plasmid and isolation of mutant (Heap *et al*., 2010). A 350bp fragment containing the retargeted intron is inserted into the ClosTron plasmid, pMTL007, between the *Bsr*GI and *Hin*dIII restriction sites. This fragment contains an antibiotic marker (Retrotransposition Marker, RAM) which is non-functional as it is interrupted by a group I intron from phage *td* (Cousineau *et al*., 1998). The retargeted ClosTron plasmid is introduced into *C. acetobutylicum* by transformation from an *E. coli* donor. Isolation of integrants is made possible by the positive selection of erythromycin resistance which becomes active once the group I intron self-splices out of the RAM (figure 1.7). PCR is carried out to confirm if intron insertion occurred at the desired DNA location.

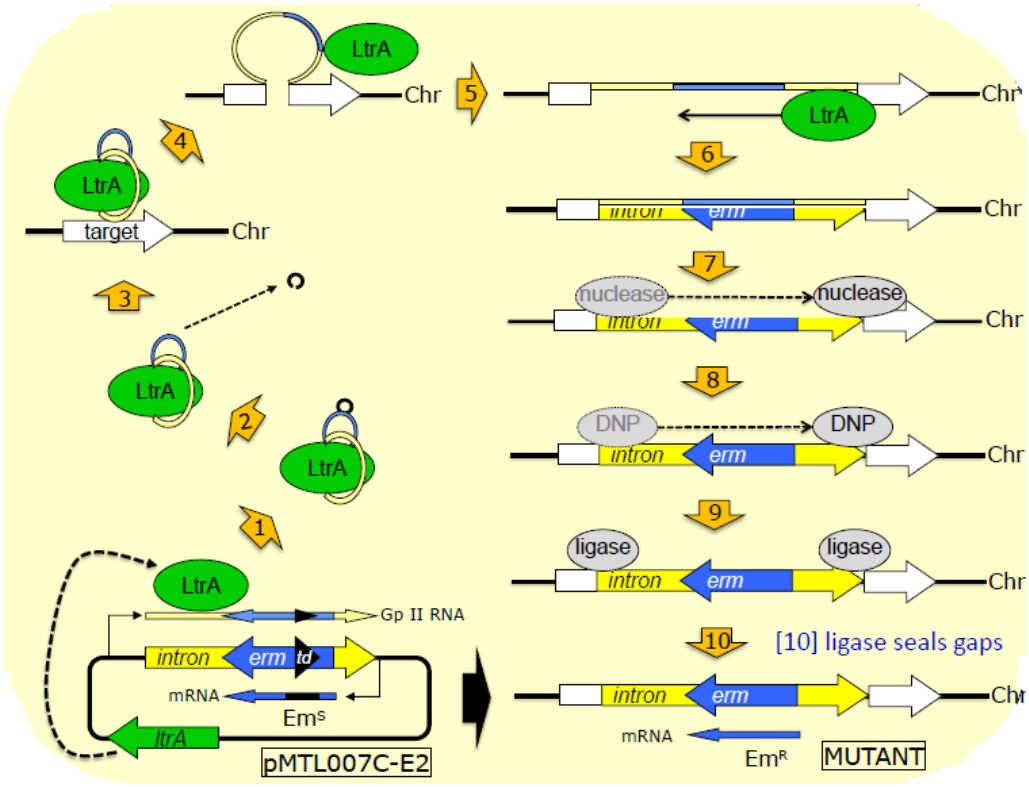


Figure 1. 7: Schematic representation of ClosTron mutant generation.

The ClosTron plasmid, pMTL007C-E2 encompasses a group II intron encoding region (yellow) which has the ermB gene (blue) inserted but inactivated by the insertion of a phage td group I intron (black). The *td* intron is self-catalytic and mediates orientation specific splicing from RNA transcripts. Within the *ermB* gene, the *td* is in the wrong direction. pMTL007C-E2 does not confer Em^R. When the opposite DNA strand is transcribed, *td* becomes in the right orientation and forms a ribonuclear protein, RNP with the LtrA protein, *td* is also spliced out (2). The RNP recognises and binds (because of the re-targeted sequence incorporated into the group II intron encoding region) to specific sequences in the target gene within the chromosome (3). DNA target is nicked by the LtrA (4) and RNA is inserted (5). LtrA posseses a reverse transcriptase activity which synthesises the complementary DNA strand (6) and host nucleases degrade the inserted RNA (7). DNA polymerase (DNP) then synthesises the opposite DNA strand (8) while ligase seals up the gap (9). With the complete integration of intron (10), host is now Em^R and mutants can be selected for by antibiotic resistance (Kuehne and Minton, 2012).

The ClosTron has been refined and streamlined (Heap *et al*., 2010). The second generation of ClosTron plasmid was derived from the pMTL80000 modular

plasmid system. This new plasmid vector, pMTL007-E2 has flippase recognition target (FRT) sites flanking the ErmBRAM allowing recombination and causing the removal of the *ermB* gene. This allows marker re-cycling and re-use of ClosTron to make multiple mutants in a particular strain (Heap *et al*., 2010). This has, however, only been possible in *Clostridium acetobutylicum*. Intron design and plasmid construction can now be performed using the free to use online algorithm at http://ClosTron.com. Figure 1.8 shows the ClosTron plasmids.

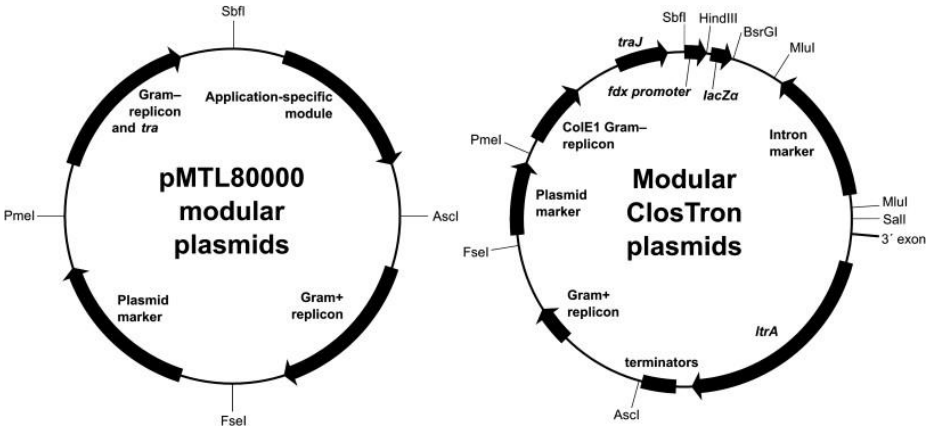


Figure 1. 8: Modular ClosTron plasmids

Derived from the pMTL80000 modular plasmid system. Figure taken from Heap *et al.*, 2010.

1.5.4 Allele Coupled Exchange (ACE)

ClosTron mutagenesis has allowed the efficient and reproducible creation of mutants in Clostridia, however, it has some drawbacks. Firstly, it results in polar effects upstream or downstream of target gene where insertion has taken place. Secondly, using ClosTron, cargo DNA exceeding 1.0kb cannot be inserted.

ACE (Allele Coupled Exchange) technology allows the integration of DNA into *Clostridium* using a pseudo-suicide plasmid with an antibiotic resistance (*catP*) This plasmid consists of a Long Homology Arm, LHA (1200bp) and a Short

Homology Arm, SHA (300bp) both derived from the host (Heap *et al*., 2012b). The initial recombination event takes place through the LHA resulting in a duplication of both homology arms. Subsequent recombination between the two SHA results in the creation of a new selectable allele. It is so called 'Allele Coupled Exchange' because this new allele consists of a chromosomal allele coupled with a plasmid- borne allele. This provides a system to 'Knock in' the desired DNA fragment which has been inserted between the SHA and the LHA.

Figure 1.9 shows ACE based on the targeting of the *pyrE* gene.

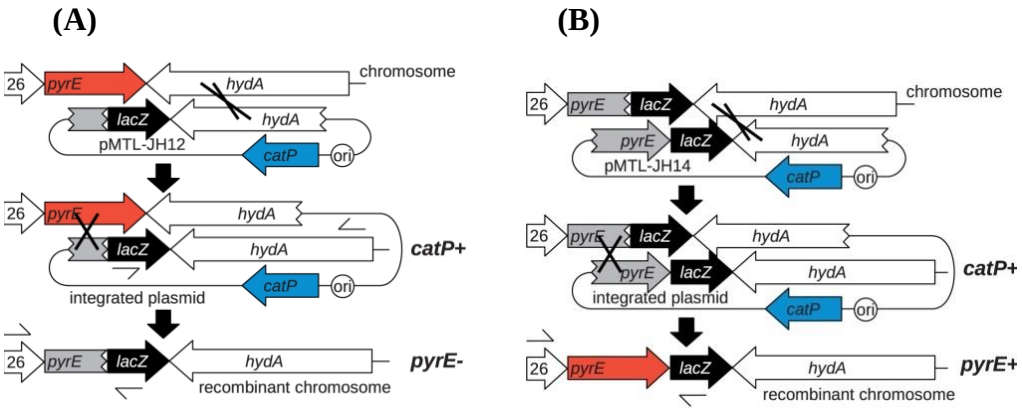


Figure 1. 9: ACE based on *pyrE* gene

(A) Creation of the *pyrE* mutant: Initial recombination between chromosome and pMTL-JH12 takes place by the LHA resulting in plasmid integration which can be selected for on Tm plates. A second recombination occurs via the SHA leading to plasmid excision and double crossover can be selected using FOA. (B) *pyrE* repair: this is based on similar principles as the *pyrE* deletion, but here, in pMTL-JH14 the 300 bp internal portion of *pyrE* that comprises the SHA is followed immediately by the remainder of the *pyrE* coding sequence. The first recombination is via LHA and with the second recombination event, double cross over clones become selectable as they now become *pyrE* positive and uracil prototrophic. Figure taken from (Heap *et al*., 2012b)

The *pyrE* gene codes for orotate phosphoribosyl transferase required for the *de novo* synthesis of pyrimidine (Yamagishi *et al*., 1996). When a *pyrE* ACE vector is introduced into the cell, in a small percentage of the population the ACE vector integrates by homologous recombination between one of the homology arms and the equivalent region in the chromosome. As the LHA is four times longer than the SHA, these single crossover integrants are far more likely due to recombination at the LHA. They can be detected as larger colonies compared to bacteria cells carrying autonomous plasmids on media containing the antibiotic thiamphenicol. This is because every daughter cell of integrants carry a copy of the *catP* gene, which confers thiamphenicol resistance. All progeny are, therefore, resistant to thiamphenicol. In contrast, dividing cells that carry the *catP* gene on an autonomous plasmid do not always pass it on to their progeny due to the defective replicon of the plasmid employed. Single cross-over integrants do not require exogenous uracil for growth and are sensitive to 5-Fluoroorotic acid (5-FOA) due to the presence of a functional *pyrE* gene. Following the double crossover event, the wildtype *pyrE* allele is replaced with the mutant allele carried by the ACE plasmid. As these cells now become resistant to 5-FOA, they can be selected by plating on media containing this uracil analogue (Heap *et al.*, 2007). However, because they are now uracil auxotrophs, exogenous uracil must also be included in the media.

Similar to the use of *pyrE* gene, the use of *upp* gene of *C. acetobutylicum* (which encodes uracil phosphoribosyl-transferase catalyzing the conversion of uracil into UMP) as a counter selection marker has also been described (Foulquier *et al.*, 2019; Croux *et al.*, 2016).

1.5.5 CRISPR/Cas System

The CRISPR (Clustered Regularly Inter-Spaced Palindromic Repeats) cas system is widely distributed in most archaea and several bacteria and acts as an adaptive immunity against invading extrachromosomal elements such as bacteriophages and plasmids (Makarova *et al*., 2011). The CRISPR locus comprises of distinct arrays of repetitive DNA sequences of up to 50 nucleotides separated by non-repetitive sequences known as spacers (Bolotin *et al*., 2005). CRISPR-associated proteins (Cas) are strictly associated with CRISPR loci and possess DNA-modifying function. The spacers represent regions of DNA from past invaders consistent with a 'cellular memory' of former infections (Bolotin *et al*., 2005; Bhaya *et al*., 2011). The mRNA thus transcribed from these spacers is able to guide the Cas protein to recognise the re-entrance of invasive DNA and break it down, thereby conferring a form of adaptive immunity to the cell.

The type II CRISPR/Cas system from *Streptococcus pyogenes* has been widely used and requires the Cas9 protein (cas9) for function. This protein is able to create a double stranded break (DSB) in DNA at a specific site as directed by the gRNA via base pairing. Endonuclease activity only takes place when target sequence is adjacent to the Protospacer Adjacent Motif (PAM) which is 5'-NGG-3' (Mojica *et al*., 2009). In eukaryotes, DNA repair can occur via Non-Homologous End Joining (NHEJ) which is prone to errors or via the Homology Directed Repair (HDR) which is more error proof (Hsu *et al*., 2014). Prokaryotes are, however, less able to cope with DSBs (Selle and Barrangou, 2015) and gene repair following Cas cleavage occurs only via HDR.

For genome editing in prokaryotes, *S. pyogenes* Cas9 acts as a powerful counter-selection measure against homologous recombination, enabling efficient gene editing (Selle and Barrangou, 2015).

The CRISPR vector essentially consists of the Cas9 endonuclease, single guide RNA (sgRNA), and a homology cassette (figure 1.10A). The Cas9 is able to create a double strand break (DSB) at a specific target site as directed by sgRNA. This DSB is lethal (figure 1.10B), but in the event of a homologous recombination event resulting in removal of the target site, there will be no DSB resulting in cell survival in the presence of the plasmid (figure 1.10C). By selecting for the presence of the vector through antibiotic resistance, it can therefore be assumed that any antibiotic resistant colonies present on selection plates are mutants, assuming Cas9 is targeting and cutting efficiently.

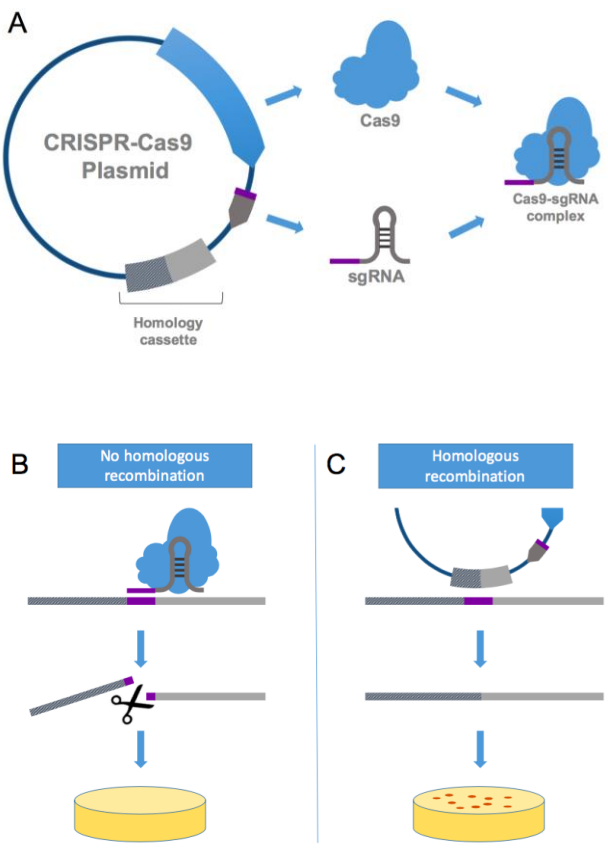


Figure 1. 10: The CRISPR/Cas gene editing.

The CRISPR vector majorly consists of the Cas9 endonuclease, single guide RNA (sgRNA), and a homology cassette (figure 1A). A DSb can be created by cas9 guided by sgRNA. This DSB is lethal (figure 1B), but cell survival occurs if the homologous recombination takes place first (figure 1C). It can therefore be assumed that any antibiotic resistant colonies present on antibiotic selection plates are mutants, assuming Cas9 is targeting and cutting efficiently.

The utility of CRISPR/Cas9 has been demonstrated in *C. difficile*, *C. sporogenes*, *C. botulinum* (Cañadas *et al.*, 2019), *C. acetobutylicum* (Li *et al.*, 2016; Wasels *et al.*, 2017), *C. saccharoperbutylacetonicum* N1-4(HMT) (Atmadjaja *et al.*, 2019), *C. beijerinckii* NCIMB 8052 (Wang *et al.*, 2015; Li *et al.*, 2016), *C. pasteurianum* (Cañadas *et al.*, 2019; Pyne *et al.*, 2016).

1.6 Metabolic Engineering of *C. acetobutylicum*

The first sequenced solventogenic *Clostridium* was *C. acetobutylicum* ATCC 824 and it represents the model organism for clostridial metabolism (Nolling *et al*., 2001). Mermelstein and Papoutsakis (1993) developed a process for DNA methylation to successfully deliver recombinant plasmids into *C. acetobutylicum* while protecting it from degradation by the *Cac824I* restriction endonuclease present in *C. acetobutylicum* (Mermelstein and Papoutsakis, 1993). This development has made the metabolic engineering of *C. acetobutylicum* feasible.

Over the years there have been several attempts to metabolically engineer *C. acetobutylicum* for improved solvent production. The first reported metabolic engineering of Clostridium was in 1993. There, the acetone production genes, (*adc, ctfA, and ctfB*) were overexpressed in *C. acetobutylicum* ATCC 824 using plasmid pFNK6 and a significant change observed in the metabolism compared to the parent strain. 13 g/l butanol was produced, there was increased final solvent concentration, reduced acid concentration and increased acetone/butanol as well as ethanol/butanol ratios (Mermelstein *et al.*, 1993).

Aside from the plasmid-based overexpression of genes, the first reported genomic engineering in clostridia was in *C. beijerinkii* where there was a targeted disruption of the *spo0A* gene using Campbell-like integration (Wilkinson and Young, 1994). In 1996, the ability to selectively inactivate genes in *C. acetobutylicum* using non-replicative plasmids was described; here, the *buk* gene was disrupted with consequent reduction in butyrate production (Green *et al*., 1996).

In subsequent studies, the inactivation of butyrate kinase, *buk*, gene resulted in increased butanol production compared to wild type and yielded 16.7 g/L

butanol, 4.4 g/L acetone, and 2.6 g/L of ethanol although an additional overexpression of the *adhE1* gene did not lead to improvement in butanol production (Harris *et al*., 2000).

The application of antisense RNA for gene down regulation has been explored for metabolic engineering of *C. acetobutylicum* (Desai and Papoutsakis, 1999).

The next major advancement in metabolic engineering was the development of mobile group II intron in *Clostridium*, the ClosTron system (Heap *et al*., 2007, Heap *et al*., 2010). This was a more efficient tool for mutant creation and was utilized by several researchers (Jang *et al*., 2014; Wang *et al*., 2013; Cooksley *et al*., 2012; Mohr *et al*., 2013; Shao *et al*., 2007). This method could be used to create multiple gene insertions. Using the ClosTron mutagenesis, Jang *et al* reported the creation of the BEKW(PptbAAD) strain which was deficient in the *pta, buk and ctfB genes* and with overexpression of *adhE1* gave up to 18 g/l butanol under batch fermentation conditions (Jang *et al*., 2012b).

Similarly, this method was applied in *C. beijerinckii* NCIMB 8052 in the inactivation of the *buk* gene with the production of 12.7 g/l butanol (Wang *et al*., 2013).

Other metabolic engineering of genes apart from the fermentative pathway have also been carried out like the overexpression of the stress gene, *groESL*, in *C. acetobutylicum* for enhanced butanol tolerance up to 17.1 g/l butanol was produced from glucose in a batch fermentation (Tomas *et al*., 2004). Nair *et al*. reported the knock out of the *solR* gene considered to repress the *sol* operon. The mutant so created produced higher butanol titers than wild type. With the overexpression of the *adhE1* in the *solR*- mutant, 17.8 g/l butanol was produced which was higher than that obtained with just the *solR*-mutant (14.6 g/l) (Nair *et*

al., 1999). The ClosTron disruption of the histidine kinase, C_AC3319 (reported as one of the three orphan histidine kinases related to the activation of *spo0A*) was reported to improve butanol titer up to 18.2 g/l (Xu *et al*., 2015).

The development of forward genetics such as transposons to generate libraries of random mutations aimed at understanding specific roles of genes in clostridia metabolism and identification of essential/conditional has been reported (Zhang *et al*., 2016; Zhang *et al*., 2015; Cartman and Minton, 2010).

Table 1.2 gives a summary of some of the metabolic engineering carried out with *Clostridium* towards improved solvent production using sugars. Unless otherwise stated, the engineered organism is *C. acetobutylicum.*

Table 1. 2 Metabolic Engineering of *Clostridium* for Solvent Production

Approach/purpose	**Gene knockout/ overexpression**	**Reference**
Increase butanol yield/titer		
Block or reduce butyrate formation	*Δbuk*	(Harris *et al.*, 2000, Jang *et al.*, 2012a)
	Δptb	(Cooksley *et al*., 2012)
Block or reduce acetate formation	*Δpta*	(Jang *et al*., 2012a)
	Δack	(Cooksley *et al.*; 2012, Kuit *et al*., 2012) (Du *et al.*, 2015; Yu *et al*., 2011)- *C. tyrobutyricum*, a non-solventogenic acetogen
Block or reduce acetone production	*Δadc*	(Jiang *et al*., 2009; Cooksley *et al*., 2012)
	ΔctfAB	(Sillers *et al*., 2009; Cooksley *et al*., 2012)
Increase the re-assimilation of acetate and butyrate	*ctfAB*	(Lu *et al*., 2017) – *C. beijerinkii*
Increase butanol formation	*ΔsolR*	(Harris *et al*., 2001)
	adhE2	(Harris *et al*., 2001; Sillers *et al*., 2009) (Yu *et al*., 2011) - *C. tyrobutyricum*

		(Lu *et al*., 2017) - *C. beijerinkii*
	Thl	(Sillers *et al*., 2009; Mann and Lütke-Eversloh, 2013)
Increase C2 to C4 pathway	*ΔldhA, ΔctfAB, Δptb, Δbuk, thl, hbd*	(Nguyen *et al*., 2018)
Increase intracellular ATP and NADH	*pfkA, pykA*	(Ventura *et al*., 2013)
Enhance strain robustness		
Increase butanol tolerance	*Δhk*	(Xu *et al*., 2015)
	groESL, grpE, htpG	(Tomas *et al*., 2003) (Mann *et al*., 2012)
	Spo0A	(Alsaker *et al*., 2004)
Reduce sporulation	*ΔspoIIE*	(Scotcher and Bennett, 2005)
	ΔCAC1766, ΔCAP0167	(Jones *et al*., 2008)
Improve substrate utilization		
Enhance xylose utilization	*ΔccpA*	(Ren *et al*., 2010)
	xylT, xylA, xylB	(Yu *et al*., 2015) - *C. tyrobutyricum*
Enhance sucrose utilization	*scrB, scrA, scrK*	(Zhang *et al*., 2017) - *C. tyrobutyricum*
Production of other solvents		
Isopropanol-butanol-ethanol	*Sadh, ctfAB, adc, hydG*	(Dai *et al*., 2012; Collas *et al*., 2012; Lee *et al*., 2012; Dusseaux *et al*., 2013; Jang *et al*., 2013; Wang *et al*., 2018)

2,3-butanediol	*Acr*	(Siemerink *et al*., 2011)
1,3-Propanediol	*dhaB1/dhaB2, dhaT, gdh, dhak*	(Gonzalez-Pajuelo *et al*., 2005) (Wischral *et al*., 2016) - *C. beijerinkii*

Table 1.2 was modified from (Cheng *et al*., 2019) and the enzymes/ proteins encoded by the indicated genes are as follows - *buk*, butyrate kinase; *ptb*: phosphotransacetylase; *pta*, phosphotransacetylase; *ack* acetate kinase, *adc*, acetoacetate decarboxylase; *ctfAB*, CoA transferase; *solR*, a repressor of *sol* locus genes; *adhE2*, aldehyde/alcohol dehydrogenase; *thl*, thiolase; *ΔldhA,* lactate dehydrogenase*; hbd*, hydroxybutyryl CoA dehydrogrnase; *pfkA*, 6-phosphofructokinase; *pykA*, pyruvate kinase*; hk*, histidine kinase; *groESL*, heat shock protein, known as hsp10/60; *grpE*, heat shock protein; *htpG*, heat shock protein; *Spo0A*, transcription factor for sporulation; *spoIIE*, stage II sporulation protein E; CAC1766, sporulation-related sigma factor; CAP0167, sporulation-related sigma factors; *ccpA*, catabolite control protein A; *xylA*, xylose isomerase; *xylB*, xylulokinase; *xylT*, xylose proton-symporter; *acr*, acetoin reductase; *scrA*, sucrose-specific PTS; *scrB*, sucrose-6-phosphate hydrolase or sucrose; *scrK*, fructokinase; *Sadh*, secondary-alcohol dehydrogenase; *hydG*, putative electron transfer protein; *acr*, acetoin reductase; *dhaB*, glycerol dehydratase; *dhaT*, 1,3-PD dehydrogenase; *gdh*, glycerol dehydrogenase; *dhak*, *dhaKLM* encoding dihydroxyacetone kinase.

Apart from *Clostridium,* some other organisms have been engineered for butanol production. *E. coli* which is more easily engineered has been metabolically engineered (Atsumi, 2007; Shen and Liao, 2008; Nielsen *et al*., 2009) to produce 0.37 g/l, 1.25 g/l and 0.58 g/l butanol respectively. Also, *Saccharomyces cerevisiae* which is more tolerant to high concentrations of alcohol has been engineered to produce butanol (Steen *et al*., 2008); harbouring isozymes in the butanol biosynthetic pathway *thl, hbd, crt, bcd, adhE2,* it produced up to 2.5 mg/l butanol.

1.6.1 Product Extension in *C. acetobutylicum*: Attempts to Reduce Acetone in ABE Production

The simultaneous production of acetone in the ABE fermentation is quite undesirable as it is a non-fuel and its co-production increases the cost of downstream fuel purification (Lee *et al*., 2012). In order to reduce the acetone production, metabolic engineering to eliminate/downregulate acetone production has been undertaken though this did not result in expected increased butanol titers.

Tummala *et al* (2003b) designed an antisense RNA to downregulate acetoacete decarboxylase, *adc,* expression and it was observed that though downregulation of *adc* was effective, it did not translate to reduction in acetone production. Next, the downregulation of acetoacetyl-CoA:acetate/butyrate:CoA transferase (*ctfA/B*) was targeted, but though the production of acetone was substantially reduced, butanol titers were significantly less compared with the control (Tummala *et al.*, 2003b). In a further work by the same researchers, the alcohol-

aldehyde dehydrogenase gene (*aad*) was overexpressed alongside the *ctfB*-asRNA and this resulted in 2 to 8-fold increase in butanol production compared with the strain containing just *ctfB*-asRNA. However, the acetone production was 4 to 6-fold higher than in strain bearing just *ctfB*-asRNA. (Tummala *et al*., 2003a). The researchers suggested that the downregulation of *ctfB* may have resulted in the degradation of the entire tricistronic *aad-ctfA-ctfB* transcript hence with the overexpression of *aad*, more butanol was produced.

Although this work improved the butanol-acetone ratio, the overall butanol production was reduced. Taking this further, Sillers *et al* (2009) expressed the *aad* gene under the transcriptional control of the *ptb* promoter while the *ctfB*-asRNA downregulated acetone production. This resulted in early alcohol (not particularly butanol) production of 30 g/l as well as an increased alcohol to acetone ratio (Sillers *et al.*, 2009) It seemed the limited availability of butyryl-CoA (due to depleted butyrate observed) was a limiting factor to butanol production hence the thiolase (*thl*) gene was overexpressed to achieve a higher butanol acetone ratio. However, against expectation, product formation was not altered implying the need for more complex metabolic engineering.

1.6.2 Isopropanol – A Viable Fuel Additive

Isopropanol is the simplest secondary alcohol with varied industrial applications including as a cleaning agent. It serves as a precursor for propylene which is used in the synthesis of products such as plastic (Jang et.al., 2013, Collas *et. al*, 2012, Rassadin *et.al*., 2006). In addition, isopropanol is used as an additive in the preparation of high-octane commercial unleaded gasolines. The cost of isopropanol is 2.5 times less than that of additives based on aromatic amines,

hence it is preferable to add isopropanol to commercial gasoline (Rassadin *et al*., 2006).

Compared to butanol in terms of inhibition to bacterial cell, isopropanol is less inhibitory as it is considerably more polar and less chaotropic, hence less inhibitory (Cray *et al.,* 2015). Isopropanol can be naturally produced by some clostridia, such as *C. beijerinckii* and *C. aurantibutyricum* in the Isopropanol-Butanol-ethanol (IBE) fermentation (George *et al*., 1983).

The size of global isopropanol market size was US$ 3.8 Billion in 2018, growing at a CAGR of 5.7% during 2011-2018. The market is estimated to reach US$ 4.8 Billion by 2023, exhibiting a CAGR of nearly 4% during 2019-2024, with a rise in demand for biofuels accelerating growth (IMARC Report, 2019)

1.6.3 IBE Fermentation

Instead of the ABE, some clostridia such as *C. beijerinckii* and *C. aurantibutyricum* carry out the Isopropanol-Butanol-Ethanol (IBE) fermentation (George *et al*., 1983). Acetone is converted in a step to isopropanol by the secondary dehydrogenase, *Sadh* and unlike acetone, isopropanol is an additive for high octane commercial gasoline (Peralta-Yahya and Keasling, 2010). Apart from the natural producers, the more useful fuel mixture, IBE, can be produced by the metabolic engineering of *C. acetobutylicum* to harbour the *Sadh* gene from *C beijerinki* (Wang *et al*., 2018; Bankar *et al*., 2015; Dusséaux *et al*., 2013; Jang *et al*., 2013; Dai *et al*., 2012; Collas *et al*., 2012; Lee *et al*., 2012)

1.7 Background to Project

C. acetobutylicum is an attractive industrial chassis for the production of chemicals and fuels. It already produces the solvents acetone, butanol and ethanol (ABE) but could, through the addition of the necessary pathways, produce further useful products. Moreover, through the manipulation of the existing pathways, it could be possible to improve the yields of the ABE production. The use of ACE technology, in combination with refinements to the ClosTron mutagenesis system will not only provide useful data on the mechanism of solvent production that will inform future metabolic engineering, but will facilitate extending the product streams of *C. acetobutylicum*.

1.8 Aim of Project

This project has two main aims. On the one hand, it is proposed to use a conditional RAM-less ClosTron approach to better understand the regulation of the ABE pathway in terms of solvent production by attempting to knock out genes that were previously impossible to knock out. A second main aim is to put in place a system for introducing a suite of new ACE vectors that can be used to introduce pivotal metabolic enzymes at discrete loci around the chromosome that may be used to implement new metabolic pathways (for example, that for isopropanol production) or to improve existing pathways (for example, butanol production). These alternate loci may also be used to introduce the inducible orthogonal expression system which would control the expression of additional gene sets introduced at the other loci.

Chapter 3 aimed to utilize a conditional RAM-less ClosTron system to knock out genes in *C. acetobutylicum* that had been previously impossible to knock-out using the standard ClosTron mutagenesis, and possibly isolate compensatory mutations that have allowed the mutant viability.

Chapter 4 was directed to create a triple auxotrophic mutant in *C. acetobutylicum* which represent alternative loci to *pyrE* that genes of interest could be readily integrated.

Chapter 5 aimed at testing the utility of the triple auxotrophic mutant created in Chapter 4 by integrating heterologous genes at the *C. acetobutylicum* chromosomal loci to produce isopropanol, utilizing the *tcdR*-P_{tcdB} inducible system.

Chapter 6 is the conclusion and further work.

CHAPTER 2

Materials and Methods

CHAPTER TWO

MATERIALS AND METHODS

2.1 Microbiology

2.1.1 Bacterial Strains

All bacteria strains used in this study are given in table 2.1.

2.1.2 Growth Media, Buffers and Supplements

2.1.2.1 Luria Bertani (LB) Medium

This was prepared according to the method described by (Sambrook *et al.*, 1989) and contained per litre 10 g tryptone, 5 g yeast extract (Difco Laboratories) and 5 g sodium chloride. The pH was adjusted to 7.5 before autoclaving. Agar (10 g/l) was added to make LB agar and medium was supplemented with antibiotics where appropriate.

E. coli TOP 10 or its derivative was grown on LB medium aerobically at 37 ℃ or 30 ℃. Broth cultures were agitated at 200 rpm (Sambrook *et al.,* 1989) during incubation.

2.1.2 .2 Clostridial Growth Medium (CGM)

CGM agar contained, per liter, 1 g yeast extract, 0.5 g KH_2PO_4, 1 g K_2HPO_4, 0.1 g $MgSO_4 \cdot 7H_2O$, 10 mg $MnSO_4 \cdot H_2O$, 15 mg $FeSO_4 \cdot 7H_2O$, 10 mg calcium chloride, 2 mg cobalt chloride, 2 g $(NH_4)_2SO_4$, 2 mg $ZnSO_4$, 2 g tryptone, 50 g glucose and 10 g agar. Medium was adjusted to pH 7.0 before autoclaving. This media is as described by (Hartmanis and Gatenbeck, 1984).

Table 2. 1: List of Bacteria Strains used in this Study

Bacteria Strain	Description	Source
***Escherichia coli* Top10**	Used for cloning and plasmid storage	Invitrogen, Ltd
***Escherichia coli* pAN2**	*E. coli* Top 10 strain carrying pAN2 plasmid for plasmid methylation	Heap *et al*.,2007
***C. acetobutylicum* ATCC 824**	Study Strain	Rostock, Germany
***C. acetobutylicum* adhE2696s::rCT**	*C. acetobutylicum adhE2* RAM-less ClosTron mutant	This Study
***C. acetobutylicum* hbd69s::rCT**	*C. acetobutylicum hbd* RAM-less ClosTron mutant	This Study
***C. acetobutylicum* thl406s::rCT.**	*C. acetobutylicum thl* RAM-less ClosTron mutant	This Study
Pfdx-catP@pyrE	*C. acetobutylicum* with P_{fdx}-catP at *pyrE* locus, WT *argH* and *pheA* loci	This Study
ΔpyrE Pfdx-catP@argH	*C. acetobutylicum* with P_{fdx}-catP at *argH* locus, WT *pyrE* and *pheA* loci	This Study
ΔpyrE Pfdx-catP@pheA	*C. acetobutylicum* with P_{fdx}-catP at *pheA* locus, WT *pyrE* and *pheA* loci	This Study
C. acetobutylicum* Δ*pyrE	*C. acetobutylicum pyrE* mutant created by ACE	CRG,CC

***C. acetobutylicum* ΔpyrEΔargH**	*C. acetobutylicum pyrE and argH double* mutant created by ACE	This Study
***C. acetobutylicum* ΔpyrEΔargHΔpheA**	*C. acetobutylicum pyrE, argH* and *pheA* triple mutant created by ACE	This Study
824BO1	*C. acetobutylicum* with P_{tcdB} -*ctfA/B-adc at pyrE* locus, P_{tcdB} -*SadH* at *pheA* locus, WT *argH*, BglR-P_{bglA}-TcdR on autonomous plasmid, pSOL1 negative	This Study
824BO2	*C. acetobutylicum* with P_{tcdB} -*ctfA/B-adc at pyrE* locus, WT *argH* and *pheA* loci, BglR-P_{bglA}-TcdR on autonomous plasmid, pSOL1 negative	This Study
824BO3	*C. acetobutylicum* with BgaR-P_{bgal}-TcdR at *pyrE* locus, P_{tcdB} –*SadH* at *pheA* locus, WT *argH*, P_{tcdB} -*ctfA/B-adc* on autonomous plasmid, pSOL1 positive	This Study
824BO4	*C. acetobutylicum* with BgaR-P_{bgal}-TcdR at *pyrE* locus, P_{tcdB} –*SadH* at *pheA* locus, WT *argH*, pSOL1 positive	This Study

2.1.2.3 Supplemented Clostridium Basal Medium (CBMS)

CBMS contained, per liter of deionised water 50 g glucose, 200 mg $MgSO_4.7H_2O$, 7.58 mg $MnSO_4.H2O$, 0.5 g KH_2PO_4, 0.5 g K_2HPO_4, 10 mg $FeSO_4.7H_2O$, 5 g $CaCO_3$, 1 mg p-aminobenzoic acid, 2 µg/ml biotin, 1 mg thiamine HCl, 4 g enzymatic casein hydrolysate.

10 g agar was added to prepare CBM agar, and $CaCO_3$ was omitted. This method is as described by (O'Brien and Morris, 1971).

2.1.2.4 CBMS Starch Medium

The CBMS medium as just described was used but 0.5 % glucose was used in addition to 2 % starch. Medium was poured in petri dishes while still hot to avoid solidification. Dried plates were left at 4 ℃ for a week so they turn milky and could be easily used to identify colonies which do not produce amylase. This medium was used to test for the presence or absence of pSOL1 plasmid.

2.1.2.5 Reinforced Clostridium Medium (RCM)

RCM was prepared by adding 21 g RCM agar ready mix (CM0149) to 400 ml of distilled water and contained per liter 13g yeast extract, 10g peptone, 5g glucose, 1 g soluble starch, 5 g NaCl, 3 g Sodium acetate, 0.5 g cysteine hydrochloride and 15 g agar. pH was adjusted to 6.8 before autoclaving.

2.1.2.6 P2 Minimal Medium

P2 medium was prepared by initially preparing four solutions; Solution 1 contained 20 g glucose in 890 ml of water; Solution 2 contained: 0.5 g K_2HPO_4, 0.5 g KH_2PO_4, 2.2 g ammonium acetate in 100 ml water autoclave solution;

Solution 3 contained 2 g $MgSO_4 \cdot 7H_2O$, 0.1 g $MnSO_4 \cdot H_2O$, 0.1 g NaCl, 0.1 g $FeSO_4 \cdot 7H_2O$ in 100 ml water and solution 4 contained 0.1 g p-aminobenzoic acid, 0.1 g thiamine, 0.01 g biotin in 100 ml water. The following made up one litre of P2 medium: 890 ml Solution 1, 100 ml Solution 2, 10 ml Solution 3 and 1 ml Solution 4. This method is as described by (Baer *et al.*, 1987)

2.1.2.7 2x Yeast Tryptone Glucose (2YTG) Broth

2YTG contained per liter 10 g yeast extract, 16 g tryptone (Difco Laboratories) 5g NaCl and 20 g glucose. The pH was adjusted to 5.2 before autoclaving. This method is as described by (Mermelstein *et al*., 1992).

2.1.2.8 Electroporation buffer (EPB)

Electroporation buffer, per liter, contained 92.4 g sucrose, 0.71 g sodium phosphate dibasic, 0.6 g sodium phosphate monobasic, pH was adjusted to 7.4 and solution filter sterilized.

2.1.2.9 TAE buffer

1×TAE buffer contained 40 mM Tris, 1 mM EDTA and 0.1% (v/v) glacial acetic acid.

2.1.2.10 PBS buffer

Phosphate Buffer Saline (PBS) contained, per liter, 8 g sodium chloride, 0.2 g potassium chloride, 1.44 g sodium phosphate dibasic and 0.24 g potassium dihydrogen phosphate. The pH was adjusted to 7.4 before autoclaving.

C. acetobutylicum was grown on CGM, RCM, CBM, P2 or 2YTG anaerobically at 37 ℃ in Don Whitley cabinet model MG1000 Mark II, under atmospheric condition of N_2:H_2:CO_2 (80:10:10). The media were pre-reduced for at least 4 hours prior to use.

Supplements and antibiotics were added where necessary as shown in table 2.2.

Table 2. 2: List of Supplements used in this Study

Supplement	Concentration	Organism
Chloramphenicol	12.5 µg/ml (broth) 25 µg/ml (agar)	*E. coli*
Thaimphenicol	15 µg/ml	*C. acetobutylicum*
Tetracycline	10 µg/ml	*E. coli* pAN2
Erythromycin	10 µg/ml	*C. acetobutylicum*
5 Flourocytosine	500 µg/ml	*C. acetobutylicum*
5 Floroorotic acid	400 µg/ml	*C. acetobutylicum*
Uracil	20 µg/ml	*C. acetobutylicum*
Arginine	10 µg/ml	*C. acetobutylicum*
Phenylalanine	10 µg/ml	*C. acetobutylicum*
Theophylline	5Mm	*C. acetobutylicum*

2.1.3 Storage Conditions

E. coli strains were stored using the Microbank™ Long Term Bacterial and Fungal Storage System (Pro-lab Diagnostics). *C. acetobutylicum* ATCC824 stock cultures were stored at -80 °C in CGM broth with 10% glycerol.

2.1.4 Quantification of Bacterial Growth

Bacteria growth was quantified by measuring the optical density, OD, of culture using Spectrophotometer BioMate 3 (Thermo Fisher Scientific, Madison, USA). A quartz cuvette was used and wavelength was at 600 nm.

2.1.5 Preparation of Electrocompetent *E. coli*

Electrocompetent *E. coli* cells were prepared in the following way – 200 ml LB broth was inoculated with 2 ml of overnight *E. coli* culture (grown on LB with

appropriate antibiotic selection where necessary). The cells were incubated at 37 ℃ shaking at 200 rpm until an OD_{600} of 0.5 – 0.7 (indicative of exponential growth) was reached. The cells were kept on ice for 15 – 30 minutes for harvesting and centrifuged at 4000 x g for 15 minutes at 4 ℃. The pellet was re suspended in 200 ml of ice-cold sterile dH_2O and centrifuged as before. The pellet was re-suspended in 4 ml of 10 % v/v glycerol and centrifuged as before. Pellets were finally re-suspended in 0.8 ml 10 % glycerol. Electro-competent cells were stored in 40µl aliquots at -80 ℃.

2.1.6 Transformation by Electroporation of *E. coli* Cells

4 µl of plasmid was added to 40 µl of electrocompetent cells and mixture was transferred to a pre-chilled 2 mm electroporation cuvette (BioRad). The cuvette was subjected to a pulse of 2.5 kV, 25 µF, 200 Ω using BioRad Gene Pulser according to manufacturer's instruction. 500 µl fresh LB was added immediately transferred to a 1.5 ml eppendorf tube and incubated at 37 ℃ for 60 minutes shaking at 200 rpm. Serial dilution was carried out and spread on LB agar plates with appropriate antibiotics were necessary.

2.1.7 Transformation of *Clostridium acetobutylicum* ATCC 824

This was done according to the method developed by Dr Ying Zhang. Ten (10) ml 2 x YTG was inoculated with a heavy loop of *C. acetobutylicum* and serially diluted (ten-fold) up to 10^{-3} and incubated overnight in anaerobic cabinet at 37 ℃. The highest dilution showing good growth was used to inoculate a flask of 310 ml 2 x YTG and incubated until an OD_{600} of 0.2 – 0.25 was reached. The culture was divided equally into eight (8) 50 ml falcon tubes and centrifuged at 4000 x g, 4 ℃ for 10 minutes. Cells were placed on ice and returned to cabinet. Pellets were re-suspended in 40 ml ice-cold EPB and centrifuged as before.

Pellets were re-suspended and added together in a tube. Competent cells were ready for immediate transformation. 20 µg of methylated plasmid DNA was transferred into a pre-chilled 4 mm electroporation cuvette and 590 µl competent cells was added. Cuvette was incubated on ice for 2 minutes and electroporated using a BIORad Gene Pulser (2.0 kV, 25 F and α Ω). I ml of fresh 2 x YTG was immediately added to the cuvette and contents transferred to a recovery tube containing 9 ml 2 X YTG. This was incubated for at least 4 hours in anaerobic cabinet at 37 ℃. Cells were finally harvested as before at room temperature and pellets were suspended in 0.5 ml fresh 2 x YTG. Cultures were plated on CGM agar plates with appropriate antibiotic selection. Competent cells not used for immediate transformation was stored at -80 ℃ in 590 µl aliquots with 10 % DMSO added.

2.1.8 Plasmid Methylation using *E coli* Top 10 PAN2

In order to overcome the restriction endonuclease Cac824I which is a major barrier to electro-transformation of *C. acetobutylicum*, plasmids were methylated. Methylation of target vectors was carried out by transforming vectors into *E. coli* Top10 cells containing the pAN2 plasmid (Mermelstein and Papoutsakis, 1993, Heap *et al*., 2007). This plasmid expresses a methylase which methylates DNA at target sites thus preventing the degradation of such methylated DNA upon introduction into *C. acetobutylicum.*

2.2 Molecular biology

2.2.1 Chromosomal and Plasmid DNA Extraction and Purification

Chromosomal and plasmid DNA was extracted from overnight cultures using the GenElute™ Bacterial Genomic and Sigma –Aldrich Plasmid Miniprep kit (alkaline lysis) according to manufacturer's instruction and eluted in dH_2O.

2.2.2 Polymerase Chain Reaction (PCR)

DreamTaq Green PCR Master Mix (Thermo Scientific) and KOD Hot Start Master Mix (Merck Millipore) were the polymerases used for DNA amplification in this study.

PCR reactions were set up and carried out according to manufacturer's instructions, using a thermal cycler block (Applied Biosystems Gene Amp PCR system 2400).

2.2.3 Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to analyse DNA fragments. Gels were prepared by dissolving 1 g of agarose in 100 ml of EDTA TAE buffer (VWR International Ltd) and run on the same buffer. SYBR safe was added to the gel at 10 µg/ml for staining. PCR products, plasmid DNA and restriction fragments were separated by gel electrophoresis at 100 V for 50 – 60 minutes. Loading dye (6 x purple, NEB) was added to samples, when required, before loading in a 1:5 (dye:sample) volume ratio and a 2-Log DNA ladder (NEB) was used. DNA bands were visualised using Gel Doc™ XR+ imaging system (Bio-Rad) and where required, DNA fragments were cut from gels under ultra-violet light using a sterile scalpel.

2.2.4 DNA Purification

DNA Gel Extraction kit used was used Monarch® DNA Gel Extraction kit for DNA purification from agarose gels while DNA was purified from PCR reaction mixtures using the Monarch® PCR and DNA Cleanup kit. These were done according to manufacturer's instruction and DNA eluted in dH_2O.

2.2.5 Quantification of DNA

Quantification of DNA was done using a NanoDrop Lite Spectrophotometer (Thermo Scientific™, UK). DNA concentration was measured at 260 nm against a blank of dH_2O and purity determined using the 260/280 ratio. For Illumina sequencing, DNA was measured using a Qubit 4 Fluorometer (Invitrogen) according to the manufacturer's instructions and DNA was run on agarose gel to detect possible RNA contamination.

2.2.6 Splicing by Overlap Extension (SOE)

This was used to create homology arms for gene deletion. A set of four primers were designed namely; LHA_Fwd, upstream of the LHA, LHA_Fwd, downstream of the RHA and two complementary primers, LHA_Rvs and RHA_Fwd which were 40 bp long. These last two primers had 20 bp comprising a portion of the LHA including the start codon of the target gene and the other 20 bp comprised a portion of the RHA with the stop codon of target gene included. In the first round of PCR the LHA_Fwd & LHA_Rvs and RHA_Fwd & RHA_Rvs were used to amplify the LHA and RHA, respectively. Following, gene fragments of homology arms were purified. In a second PCR, both homology arms were joined (the complementary ends introduced by LHA_Rvs

and RHA_Fwd allows the annealing of both homology arms) by using just the LHA_Fwd and LHA_Fwd primers and this was followed by DNA purification. This method is as previously described (Higuchi *et al*., 1988).

2.2.7 Cloning Techniques

2.2.7.1 Restriction Endonuclease DNA Digestion

All restriction endonucleases and buffers used in this study were supplied by NEB and digestion reactions were set up in accordance to the manufacturer's instructions.

2.2.7.2 DNA Blunting using T4 DNA Polymerase

T4 DNA Polymerase (NEB, UK) was used to produce blunt ended DNA after digest with restriction enzymes according to manufacturer's instructions.

2.2.7.3 De-phosphorylation of Linearised DNA Fragments

The Antarctic Phosphatase (NEB, UK) was used for removal of the 5' phosphorylated end of linearised DNA fragments to prevent vector self-ligation and was inactivated by heating at 80 ℃ for 2 minutes. Reactions were set up according to manufacturer's instruction.

2.2.7.4 Ligation and Membrane Dialysis

Digested, purified DNA fragments were ligated using T4 DNA ligase (NEB) at molar ratios of insert to vector of 3:1 and 5:1 or 7.1 in 20µl reaction volumes according to manufacturer's instruction. Ligations were incubated overnight on melting ice. Preceding transformation into electrocompetent cells, ligation product was subjected to a 30-minutes dialysis through an MF-Millipore™ (membrane pore size of 0.025-µm) (Merck) for 20 minutes over dH_2O.

2.2.7.5 Sanger DNA Sequencing

Sanger DNA sequencing was carried out using Source BioScience (Nottingham, UK) Services for plasmid DNA and purified PCR products.

2.2.7.6 Illumina Genome Sequencing

Genome sequencing was carried out by by MicrobesNG (http://www.microbesng.uk), using Illumina-based, Standard Whole-Genome Service.

2.2.8 Plasmid Construction

All plasmids used in this study were constructed based on the modular pMTL80000 series and are listed in table 2.3. The list of oligonucleotides used are given in table 2.4.

2.2.8.1 ClosTron Plasmids

Genes of interest were disrupted by insertion using the ClosTron system as previously described (Heap *et al*., 2007, Heap *et al*., 2010). Retargeted introns were designed using the ClosTon intron design tool available at www.clostron.com and introns were selected based on high score. The incorporation of the retargeted intron sequences into pMTL007C-E2, the standard ClosTron vector, and the plasmid construction was carried out by DNA 2.0 Inc, Menlo Park CA, USA

Table 2. 3: List of Plasmids used in this Study

Plasmid	Relevant Features	Source
ClosTron Plasmids		
pMTL007C-E2	ClosTron vector backbone with pCB102/ColE1 replicons, Cm^R/Tm^R	CRG, CC
pMTL007C-E2::hbd69s	Standard *hbd* ClosTron KO vector	This Study
pMTL007CC-E2	RAM-less ClosTron vector backbone with *codA-catP* fragment for counter-selection with 5-FC, pCB102/ColE1 replicons	CRG, CC
pMTL007CC-E2::adhE2696s	RAM-less ClosTron KO plasmid for *adhE2* gene	CRG, CC
pMTL007CC-E2::hydA936s	RAM-less ClosTron KO plasmid for *hydA* gene	This Study
pMTL007CC-E2::hbd69s	RAM-less ClosTron KO plasmid for *hbd* gene	This Study
pMTL007CC-E2::crt154s	RAM-less ClosTron KO plasmid for *crt* (154) gene	This Study

pMTL007CC-E2::crt307s	RAM-less ClosTron KO plasmid for *crt* (307) gene	This Study
pMTL007CC-E2::thl406s	RAM-less ClosTron KO plasmid for *thl* gene	Ying Zhang
Allelic Exchange/ACE Plasmids		
pMTL8-147-argH	Suicide KO plasmid for *argH* gene with *codA* in backbone; ColE1 replicon; Cm^R/Tm^R	CRG, CC
pMTL8-147-pheA	Suicide KO plasmid for *pheA* gene with *codA* in backbone; ColE1 replicon; Cm^R/Tm^R	CRG, CC
pMTLME6C	*pyrE* complementation plasmid and backbone for *argH* and *pheA* complementation/integration plasmids pIM13/ColE1 replicon; Cm^R/Tm^R	CRG, CC
pMTL_comp_argH.	*argH* complementation plasmid	This Study
pMTL_comp_pheA.	*pheA* complementation plasmid	CRG, CC
pMTL-HZ13-HZ-tcdR	ACE integration vector directing lactose inducible *tcdR* at *pyrE* locus	This Study/ Hengzheng Wang

pMTL-HZ14-PtcdB-ctfAB-adc	ACE integration vector directing acetone operon at *argH* locus	This Study
pMTL-HZ15-PtcdB-SadH	ACE integration vector directing *SadH* gene at *pheA* locus	This Study/ Hengzheng Wang
pMTL-JH36	Plasmid to KO pSOL1 plasmid; Cm^R/Tm^R	CRG, CC
pMTL-SC1-add9-codA-ermB	Used as a source of *codA-erm*	CRG, CC
pMTL-HZ14-PtcdB-ctfAB-adc-codA	pMTL-HZ14-PtcdB-ctfAB-adc with *codA* added	This Study
pMTL-HZ14-PtcdB-ctfAB-adc-suicide	pMTL-HZ14-PtcdB-ctfAB-adc with Gram positive replicon, pIM13, removed	This Study
pMTL HZ14-HZ-tcdR	ACE integration vector directing lactose inducible *tcdR* at *argH*	This Study
pMTL-HZ13-PtcdB-ctfA/B-adc	ACE integration vector directing acetone operon at *pyrE* locus	This Study
pMTL82251-HZ_tcdR	Expression plasmid for the Lactose inducible *tcdR* system; pBRP1/ColE1 replicon; $ermB^R$	CRG, CC

pMTL-HZ13-fdx-catP	ACE integration vector directing fdx-catP at *pyrE* locus	This Study
pMTL-HZ14-fdx-catP	ACE integration vector directing fdx-catP at *argH* locus	This Study
pMTL-HZ15-fdx-catP	ACE integration vector directing fdx-catP at *pheA* locus	This Study
CRISPR/Cas9 Plasmids		
pMTL8315_RibocasE_J23119_MCS original	CRISPR/Cas9 vector backbone; pCB102/ColE1 replicons; Cm^R/Tm^R	Daphne Groothuis
pMTL8315_ RibocasE_pyrE	CRISPR/cas9 knock-out vector for *pyrE*	This Study
pMTL8315_ RibocasE_hbd	CRISPR/cas9 knock-out vector for *hbd*	This Study

**CRG, CC – Clostridia Research Group, Culture Collection*

Table 2. 4: List of Oligonucleotides used in this Study

Primer	Sequence (5′→ 3′)	Template/Purpose
Chapter 3		
adhE2/F	TGTTAAATATCAAACTAATAATAAATTTTATAAAGGAGTG	*adhE2* ClosTron insertion screening
adhE2/R	GAATGTAAAATAGTCTTTGCTTCATTATATTAGC	*adhE2* ClosTron insertion screening
***hbd*/F**	CAGCATTTAGCAGGTATGCAAG	*hbd* ClosTron insertion screening
***hbd*/R**	CCATCTAAAGCATTGACCACAG	*hbd* ClosTron insertion screening
***hydA*/F**	GGCTCCATCAGTAAGAACTGC	*hydA*, ClosTron insertion screening
***hydA*/R**	CCATAGCAGGATCAACTTCACC	*hydA*, ClosTron insertion screening
***thl*/F**	ACCCCGTATCAAAATTTAGGAGG	*thl*, ClosTron insertion screening
***thl*/R**	AAATAACTCTGTAGAACTAATTTATAATTCTACAGAG	*thl*, ClosTron insertion screening
crt/F	AGCCGAGATTAGTACGGTAATG	*crt*, ClosTron insertion screening
crt/R	ATGGGTCAGTATGGTATGATGG	*crt*, ClosTron insertion screening
bcd/F	GATTAGTTGGAATGGGCATGG	*bcd*, ClosTron insertion screening
bcd/R	AGTTCCAGATACAGCGGAAG	*bcd*, ClosTron insertion screening

bcd3F	GTTTGCGGTACTACAGGAG	*bcd*, ClosTron insertion screening
bcd3R2	GCAAGACCTTGGAATTTGTCAAG	*bcd*, ClosTron insertion screening
SCT	GGGTGTAGTAGCCTGTGAAAT	Screen for retargeted site in ClosTron plasmid
EBS Universal	CCATAGCAGGATCAACTTCACC	ClosTron insertion screening (intron-exon junction)
Chapter 4 and 5		
JH14/F	TAGCACAATTGTATTTGGACTTCTTTAAATAAAAACATGG	Screening for deletion/ repair/integration at *pyrE* locus
JH14/R	TTGATGATGTTTGTCTTGATGACTCAACATGC	Screening for deletion/ repair/integration at *pyrE* locus
ACE/F	CAATGGCAAAGTAATCATAGATGATAAAATGAAAGATAG	Sanger sequencing of *pyrE* locus
ACE/3R	CCATTCGCCATTCAGGCTGC	Sanger sequencing of *pyrE* locus
***argH*-ch-F**	TGATGTTGAGGTAATAGCTGTATG	Screening for deletion/ repair/integration at *argH* locus
***argH*-ch-R**	GTGTATTCTCTATATTATCTCTTGC	Screening for deletion/ repair/integration at *argH* locus
ACEHZ14/F	AATGAATTTAAAAATTTCTGTAATCTATTTGATGAAGATGTTTATG	Sanger sequencing of *argH* locus
pheA_ch_F	CATGCAGATATATACATTCCTCTG	Screening for deletion/repair/integration at *pheA* locus

pheA_ch_R	CTCATTGACTTATAATCAAAACGA	Screening for deletion/ repair/integration at *pheA* locus
ACEHZ15/F	GACAATAGATTAAATGAAGCTGTAGAGGCG	Sanger sequencing of *pheA* locus
pyrE-HA/F	TATCATGACGTTGGTGGAGAT	Screening CRISPR KO of *pyrE* gene
pyrE-HA/R	TGTTTATACAGTTACTATCATGCCTTG	Screening CRISPR KO of *pyrE* gene
hbd/HA/F	ACACATAAAAGTACCTTGAATAATATATATC	Screening CRISPR KO of *hbd* gene
hbd/HA/R	GATTTTTCAATGAATAAAGCAGATTACAAGG	Screening CRISPR KO of *hbd* gene
pyrE_gRNA1	GGCACTAGTGGGGATTTCATAACTAAAAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT	sgRNA for *pyrE* locus
hbd_gRNA1	GGCCACTAGTAAGCTTGTAGAGGTAATAAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT	sgRNA for *hbd* locus
sgRNA_AatII_R	GGCCCCGGGACGTCATAAAAATAAGAAGCCTGCAAATGCAGGCTTCTTATTTTTATAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTG	Reverse primer for sgRNA
pyrE_LHA_Fwd	GGCGACGTCGAATAATAAATTGGTTAGATCATTAAAGT	*pyrE* HA cloning
pyrE	GCACATTTATTTTACTATTTTTCCATTTTTAGCCCTCCAT	*pyrE* HA cloning

pyrE_RHA_Fwd	ATGGAGGGCTAAAAATGGAAAAATAGTAAAATAAATGTGC	*pyrE* HA cloning
pyrE_RHA_Rvs	GGCGGCGCGCCGCAAAAATTAAATTTGCAGATCTTG	*pyrE* HA cloning
hbd _LHA_Fwd	GGCGACGTCTAGATTTCTTATCAAGTCAAAAACCTC	*hbd* HA cloning
Hbd_LHA_Rvs	TAGGGAGGTCTGTTTAATGAAAAAATAAGTTTACAAGAATC	*hbd* HA cloning
Hbd_RHA_Fwd	GATTCTTGTAAACTTATTTTTTCATTAAACAGACCTCCCTA	*hbd* HA cloning
Hbd_RHA_Rvs	GGCGGCGCGCCGAATTAACAGCTGTTTTACTTGGA	*hbd* HA cloning
M13F	TGTAAAACGACGGCCAGT	Screening/sequencing for modular plasmids
M13R	CAGGAAACAGCTATGACC	Screening/sequencing for modular plasmids
CatP_F1	GATAAATAGTTAACTTCAGGTTTGTC	Screening/sequencing for modular plasmids
CatP_R1	CAAGTTTATCGCTCTAATGAAC	Screening/sequencing for modular plasmids
pCB102-R1	CTGTTATGCCTTT TGACTATC	CRISPR screening/sequencing

2.2.8.2 Conditional RAM-less ClosTron Plasmids

The RAM-less ClosTron plamids were made by cloning in the insert (from standard ClosTron plasmid) using *Hin*d*III* and *BSr*G*I* into the pMTL007CC-E2 plasmid backbone and then deleting the RAM element (1167 bp) by a using the *Mlu*I restriction enzyme.

The plasmid maps for standard ClosTron, pMTL007C-E2 and the conditional RAM-less ClosTron, pMTL007-E2-RAM^{-}, are given in figure 2.1.

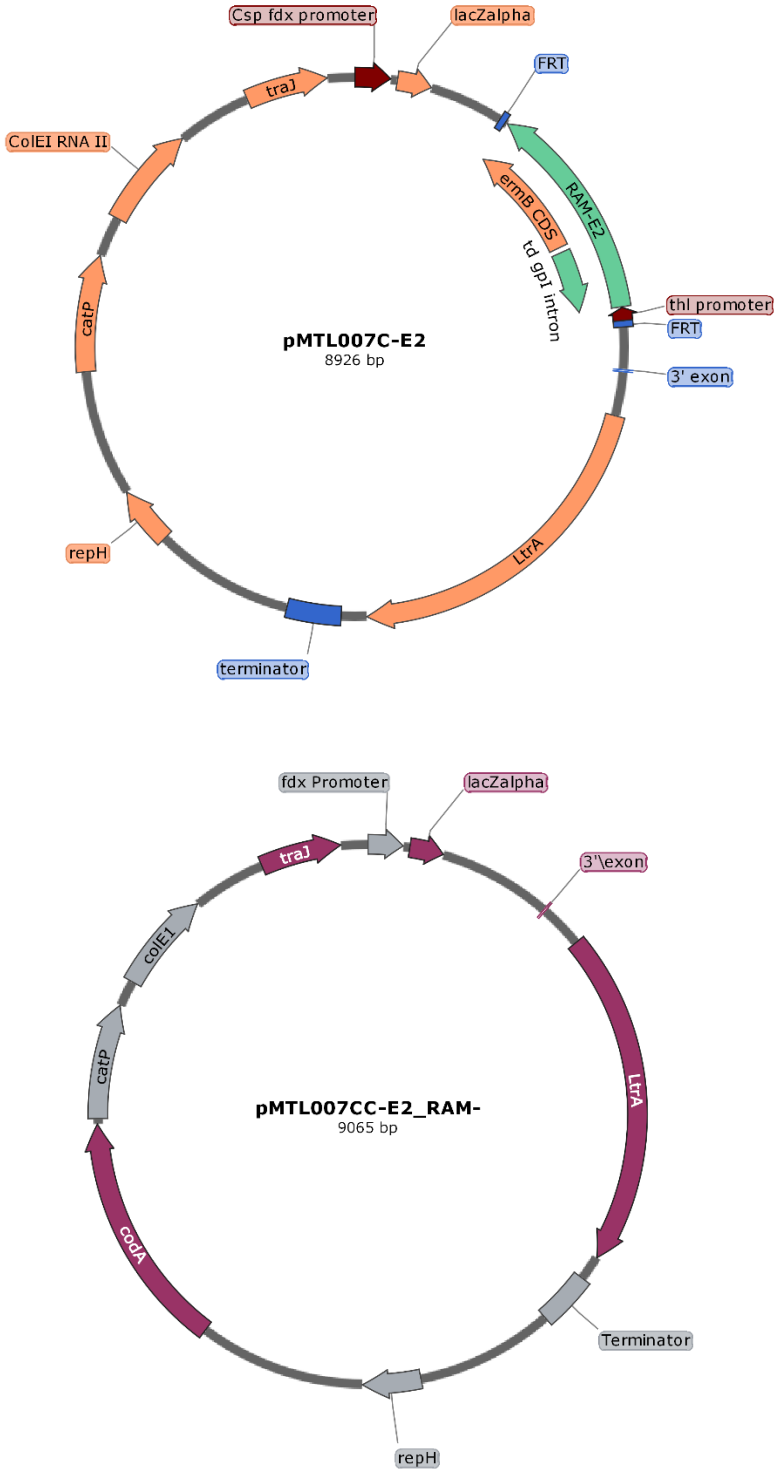


Figure 2. 1: The ClosTron plasmid backbone.
Standard ClosTron plasmid backbone, pMTL007C-E2 and conditional RAM-less ClosTron plasmid, pMTL007-E2-RAM^{-}. Plasmids consist of *C. sporogenes fdx* promoter followed by *lacZ* region where retargeted intron can be inserted, Retrotransposition Activated Marker, RAM-E2, *ltrA* gene, a terminator, Gram-positive replicon, *repH*, antibiotic marker, *catP*, Gram-negative replicon, ColE1. In the case of pMTL007-E2-RA M^{-}, the RAM-E2 is removed and it contains the *codA* gene.

2.2.8.3 Suicide Knock-Out Plasmids

The KO plasmids for *argH* and *pheA* genes were pMTL8-147-argH and pMTL8-147-pheA respectively. These plasmids are suicide, bearing no Gram-positive replicon.

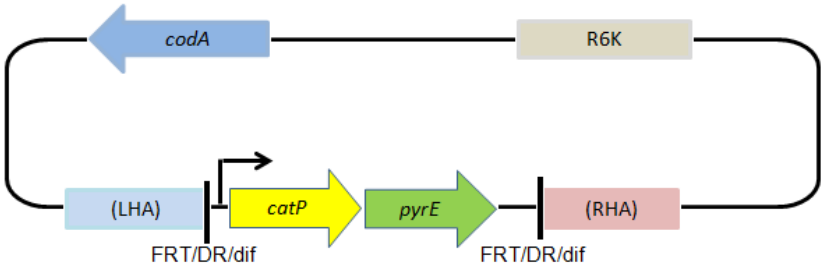


Figure 2. 2: Representative of the suicide KO vectors, pMTL8-147-argH and pMTL8-147-pheA.

The *codA* gene and a functional *pyrE* gene served as counter selection markers. The left and right homology arms were designed to direct the deletion of 406 and 301 bp from the 3'end of *argH* and *pheA* genes respectively. The Gram-negative replicon was R6K and *catP*, the antibiotic marker. These plasmids were obtained from the Clostridia Research Group Culture Collection (CRG, CC).

2.2.8.4 Allele Coupled Exchange (ACE) Repair Plasmids

Thee repair plasmids (pMTL_comp_argH and pMTL_comp_pheA) for *argH* and *pheA* genes were made based on the pMTLME6C *pyrE* ACE complementation plasmid (Ehsaan *et al*., 2016) with the homology arms replaced to direct recombination at the *argH* and *pheA* sites.. The LHA and RHA were amplified from the *C. acetobutylicum* genome. The LHA represented the previously deleted 3' regions and flanked by restriction sites for *Sbf*I and *Not*I

while the RHA was 1200 bp downstream of the *argH/pheA* gene and was bound by *Nhe*1 and *Asc*1 restriction recognition sites.

2.2.8.5 Allele Coupled Exchange Knock-In Plasmids

The ACE expression vectors pMTL-HZ13-HZ-tcdR, pMTL-HZ14-PtcdB-ctfAB-adc and pMTL-HZ15-PtcdB-SadH were similar to the repair vectors pMTLME6C, pMTL_comp_argH and pMTL_comp_pheA respectively but had the gene/operon to be expressed at these sites cloned in between the homology arms in each case. For the pMTL-HZ13-HZ-tcdR vector, the

621 bp fragment encoding the *tcdR* gene was cut from pMTLME6C-tcdR (CRG, CC) and pasted into pMTL-ME6C::HZ(Erm) (CRG, CC) via the *Nde*I and *Nco*I restriction sites.

To clone pMTL-HZ14-PtcdB-ctfAB-adc (CRG, CC), the sequential cloning of the *C. difficile* P_{tcdB} promoter, *ctfA/B* and *adc* genes into pMTL_comp_argH was carried out. *ctfA/B* and *adc* were amplified from the pSOL1 plasmid of *C. acetobutylicum*. Following a terminator, P_{tcdB} was flanked by the *Not*I and *Nde*I restriction sites, *ctfA/B* by *Nde*I and *Spe*I while *adc* and stop codon was flanked by *Spe*I and *Sal*I.

pMTL-HZ15-PtcdB-SadH was made by cloning sequentially the P_{tcdB} promoter and the *SadH* gene from *C. beijerinkii* into pMTL_comp_pheA. P_{tcdB} was flanked by restriction sites *Not*I and *Nde*I and *Nde*I and *Xho*I flanked *SadH* gene.

2.2.8.6 CRISPR/Cas9 Gene Deletion Plasmids

Gene deletion by CRISPR/Cas9 was carried out by using the vector backbone, pMTL8315_RibocasE_J23119_MCS (CRG, CC). This vector consists of the

Cas9 gene from *Streptococcus pyogenes* under the transcriptional control of the promoter of the *C. sporogenes* ferredoxin gene, fused to a theophylline-responsive riboswitch (P_{fdx}-rbE). It also contains a synthetic promoter, J23119 which drives the sgRNA expression. Using the Benchling platform, sgRNA was selected based on high on- and off- target scores as calculated (Doench *et al*., 2016, Doench *et al*., 2014, Hsu *et al*., 2014). sgRNAs were generated as primer dimers using a sgRNA specific forward primer, and a reverse primer, sgRNA_AatII_R. The generated sgRNA was flanked by the *Spe*I and AatII restriction sites. Homology arms were designed to flank the region to be deleted, amplified form the chromosome and joined together by SOE PCR. The homology arms were flanked by *Aat*II and *Asc*I restriction sites. The sgRNA was digested with *Spe*I and *Aat*II, homology arms with *Aat*II and *Asc*I and ligated together with the *Spe*I and *Asc*I digested vector backbone, pMTL8315_RibocasE_J23119_MCS, in one step.

2.2.9 Targeted Genomic Mutagenesis of *C. acetobutylicum*

2.2.9.1 ClosTron Mutagenesis

Methylated plasmids were electroporated into *C. acetobutylicum*, carried out as described in section 2.1.6. Transformed cells were plated on CGM plates supplemented with erythromycin, incubated for 24 – 48 hours and integrants were selected based on the acquisition of antibiotic resistance. PCR screening of putative integrants was carried out using two sets of primers; gene-specific flanking primers as well as EBS and gene-specific flanking (forward or reverse depending on the orientation of the gene). The later pair amplified the intron–exon junctions. Plasmid loss was screened based on antibiotic sensitivity. DNA

fragments were subjected to Sanger sequencing to confirm the nucleotide sequence was as expected.

For the RAM-less ClosTron, transformed cells were plated on CGM plate supplemented with thiamphenicol (Tm) for 24 – 48 hours. 96 colonies were used to inoculate a 96-well microtitre plate containing 2 xYTG media. Cells were pooled and PCR screened using two sets of primers as previously mentioned. Cells with the expected DNA fragment were grown in 2xYTG overnight, diluted up to 10^{-3} and plated on CBM+FC plates and plasmid loss selected for based on thiamphenicol sensitivity.

2.2.9.2 Suicide Knock-Out of Genes

Upon transformation of *C. acetobutylicum* with methylated plasmids, single cross integrants could be selected for as faster growing antibiotic resistant colonies on CGM+Tm plates. Colonies were restreaked on same plates and double cross was screened for by growing colonies overnight in 2xYTG diluting up to 10^{-3} and plating on CBM 5FC plates. Plasmid loss was selected for based on thiamphenicol sensitivity and PCR screening was done using gene flanking primers. *catP-pyrE* flip was selected for by plating on CGM 5-FOA plates.

2.2.9.3 ACE Gene Repair and Knock-In

Methylated ACE plasmids were transformed into auxotrophic *C. acetobutylicum* mutants, plated on CGM+erm plates appropriately supplemented based on the specific auxotrophic mutant (uracil, arginine or phenyalanine) and incubated for 24 – 48 hours. To screen for gene repair, colonies were streaked on unsupplemented plates and colonies that had become prototrophic were PCR

screened using appropriate gene flanking primers. Plasmid loss was selected for based on erythromycin sensitivity.

2.2.9.4 CRISPR/Cas9 Gene Deletion

CRISPR/Cas9 plasmids were methylated and electroporated into competent cells of *C. acetobutylicum* and plated onto CGM+Tm (non-induced) and CGM+Tm+Theophiline (5 mM) plates. After 48 hours, thiamphenicol resistant transformants from the non-induced were re-streaked onto induced plates and incubated for 24 – 48 hours. Colonies that grew on induced plates were PCR screened using appropriate primers flanking the homology arms.

2.3 CAT Enzyme Assay

A full loop of the strain was inoculated in 10 ml 2xYTG broth, dilutions made to 10^{-5} and incubated overnight. 100 ml 2xYTG was inoculated with overnight pre-culture to give a final OD_{600} of 0.05. Samples were collected at 4 and 8 hours, centrifuged at 13,000 x g for 10 minutes at 4 ℃ and pellets were stored at -80 ℃. Pellets were re-suspended in 1 ml PBS supplemented with Roche complete protease inhibitor cocktail tablets, transferred on ice and lysed using a bioruptor. Lysates were centrifuged at 13,000 x g for 30 minutes at 4 ℃ and supernatant retained for analysis

A quartz cuvette was prepared containing 540 µl of 100 mM Tris buffer (pH 7.8), 200 µl of 2.5 mM DTNB solution (5,5-Dithio-bis(2-Nitrobenzoic Acid) in100 mM Tris buffer, (pH 7.8), 200µl of 5.0 mM freshly prepared acetyl coenzyme A solution in deionized water and 10 µl of cell lysate. The cuvette

was pre-warmed to 25 ℃ and reaction initiated by adding 10 µl of 0.3 % chloramphenicol solution. Increase in absorbance at 412 nm was recorded. Units/ml in lysate was calculated:

$$\text{Units/ml enzyme} = \frac{(\Delta A_{412nm}/\text{min Test} - \Delta A_{412nm}/\text{min Blank})\ (3)\ (\text{dilution factor})}{(0.0136)\ (0.05)}$$

(3) – Total volume (ml) of assay
(0.0136) – Micromolar extinction coefficient for TNB
(0.05) – Volume (ml) of enzyme used

2.4 Fermentation and Solvent Analysis

Fermentation was set up in flasks using 60 ml CBMS appropriately supplemented. Medium was inoculated using overnight cultures to achieve an initial OD of 0.05. Samples for solvent analysis were collected at intervals, pH and OD were also measured. For GC and HPLC sample collection, 1 ml of sample was centrifuged at 13,000 x g for 5 minutes at 4 ℃, supernatants stored in screw top tubes at – 80 ℃ until analysis performed.

2.4.1 Gas Chromatography, GC

Cell free supernatant was thawed on ice and 500 µl of sample transferred to an Eppendorf tube. 5 µl of H_2SO_4 was added and mixed shortly after which 500 µl of propyl propionate (100 ml containing 549 µl of internal standard, valerate) and vortexed. Solution was centrifuged and 300 µl from top (organic) layer transferred into a GC vial containing an insert and screw cap**.** Butanol, ethanol, acetone, acetate and butyrate were analysed using Thermo Focus GC equipped with a TR-FFAP (30 m) column with internal diameter of 0.25 mm. and a Flame Ionization Detector (FID). The carrier gas was hydrogen and flow rate was

0.8 ml/minute. For peak resolution, the starting temperature was set at 40 ℃ for 2 minutes followed by a ramp of 80 ℃/ minute until 150 ℃ and finally a 45 °C/minute ramp until 210 ℃ which was held for 1 minute. The injector temperature was 240 ℃ while the detector was at 270 ℃. Flame maintenance was by nitrogen at 30 ml/minute, compressed air at 350 ml/minute and hydrogen at 35 ml/minute.

2.4.2 High Performance Liquid Chromatography, HPLC

Cell free supernatant was thawed on ice and mixed with equal (300 µl) volume of internal standard (80 mM valeric acid [Sigma-Aldrich, UK] in 0.005 M H_2SO_4) and vortexed. Solutions were filtered through 0.2 µm syringe filters directly into HPLC vials containing 300 µl inserts with split caps. Glucose and lactose were analysed using a Dionex UltiMate 3000 HPLC system (Thermo Fisher Scientific, UK) supplied with a Bio-Rad Aminex HPX-87H (BioRad, UK) column, a refractive index (RI) and diode array detector (DAD) at UV 210 nm at an isocratic flow rate of 0.5 ml/minute of 0.005 M H_2SO_4 as mobile phase and a column temperature of 35 ℃ for 55 minutes. Range of standard concentrations were 0 – 300 mM. The Chromeleon 7.2 Chromatography Data System (Thermo Fisher Scientific, UK) was used to for signal analysis.

2.5 Bioinformatics Tools

2.5.1 Plasmid Map Construction and Visualization

Plasmid map construction and visualization was performed using SnapGene® software (from GSL Biotech; available at snapgene.com). The software was also used for sequence data handling.

2.5.2 Primer Design

Manual primers design was carried out using the SnapGene software and their characteristics verified upon ordering from the Sigma-Aldrich online platform.

2.5.3 Analysis of Illumina Sequencing Data

Whole genome sequencing from Illumina was analysed using the CLC Genomics Workbench version 10. The Basic Variant Detection tool was used for identification of single nucleotide variants, insertions and deletion.

2.5.4 Sequence Database Search

Alignment of gene and protein sequences together with database searches were carried out using BLASTn and BLASTp of the National Centre for Biotechnology Information (NCBI).

2.5.5 Data analysis and Visualization

Microsoft Excel 2016 (Microsoft, USA) and GraphPad Prism version 8 for Windows (GraphPad Software, USA) was used for data analysis, visualization and graphical presentation of data.

CHAPTER 3

Targeted Mutagenesis using Conditional RAM-less ClosTron

CHAPTER 3

Targeted Mutagenesis using a Conditional RAM-less ClosTron

3.1. Introduction

The branched metabolic pathway of ABE fermentation involves complicated cellular processes and many details of the regulatory mechanisms in operation remain poorly understood (Birgen *et al*.; 2019, Liu *et al*., 2016; Jones and Woods, 1986). The generation of mutants for improved understanding and manipulation of solvent pathways is necessary in the development of an industrial solvent producing strain.

3.1.1 Previous Mutational Analysis of the ABE Pathway

Cooksley *et al*. (2012) used ClosTron mutagenesis to knock out *adhE1*, *adhE2* (alcohol/aldehyde dehydrogenase), *bdhA*, *bdhB* (butanol dehydrogenases), *ack* (acetate kinase), *adc* (acetoacetate decarboxylase), *ctfA*, *ctfB* (CoA transferase), CAP0059 (putative alcohol dehydrogenase) and *ptb* (phosphotransbutyrylase) in *Clostridium acetobutylicum* ATCC 824. The genes *hydA* (main hydrogenase) and *thl* (thiolase) could not be knocked out despite repeated attempts and construction of a minimum of two retargeted ClosTron plasmids. This large-scale knock out was carried out to provide information on the fundamental process of solvent production in *C. acetobutylicum* to see the possibility of altering the ABE ratio in favour of butanol as a step towards the generation of an industrial and efficient butanol producing strain.

The isolation of the *ptb* mutant proved difficult; it involved repeated attempts and testing of three different retargeted ClosTron plasmids. The single *ptb* mutant eventually obtained, produced no butyrate but increased amounts of acetic and lactic acid. Only a little butanol was produced (7 mM), no acetone was formed while ethanol was produced in large quantities (113 mM). However, the *ptb* mutant could not be complemented. The sequencing of the *ptb* mutant genome was carried out and it revealed a frameshift mutation in the *thl* gene which explained the phenotypic observations. This is quite interesting as the independent knock-out of the *thl* gene was not possible but became viable in combination with a *ptb* mutant. Was *ptb* knock-out only possible because of a frameshift in *thl*, or *vice versa*?

Inactivation of *ptb* must cause a build-up of butryl-CoA (substrate), and loss of butyrate phosphate (product). One of these metabolites is most likely responsible for the presumed lethality of inactivation of *ptb*. Inactivation of thiolase would block formation of butyryl-CoA.

Subsequent to the study of Cooksley and co-workers, Lehmann *et al*. (2012) reported the generation of a *ptb* mutant of the same strain of *C. acetobutylicum* with a butyrate-negative phenotype. The major fermentation products without pH control were acetate, lactate and butanol while ethanol and butanol were the products when pH was regulated at ≥ 5.0. However, the authors made no attempt to complement nor did they sequence the genome. Accordingly, it cannot be definitely stated that the observed phenotype is due to disruption of *ptb* gene alone, or whether ancillary mutations are also present.

Lehmann and Lutke-Eversloh (2011) had also previously demonstrated knock-out of the *hbd* gene using ClosTron technology. Activities of thiolase and

crotonase were significantly reduced compared to the wild type. The disruption in the butyrate/butanol pathway was reported, converting *C. acetobutylicum* to an ethanol producer. However, as with their *ptb* study, these authors did not complement their *hbd*::int (69) mutant. Accordingly, as they had not established Molecular Koch's postulates, the observed physiological and phenotypic differences observed in their putative mutant cannot be assumed to be solely due to intron-mediated inactivation of *hbd.*

Jang and co-workers have reported the ClosTron-mediated, multiple knock-out of genes involved in solventogenesis in *C. acetobutylicum* (Jang *et al*., 2014) including genes such as *hydA* that could not be inactivated in the study of Cooksley *et al.* (2012). The explanation for this is simple. The authors (Jang *et al*., 2014) used RAM-less ClosTrons and directed the insertion of the intron into the sense strand of the DNA target and used a selective strategy that ensured the intron delivery vector (which carries the *ltrA* gene) remained within the cell. As a consequence, the inserted intron would be spliced from transcribed mRNA by the vector encoded LtrA protein leading to the translation of full-length protein of the 'inactivated gene'. In short, they would not be null mutants but only knock-down mutants (figure 3.1)

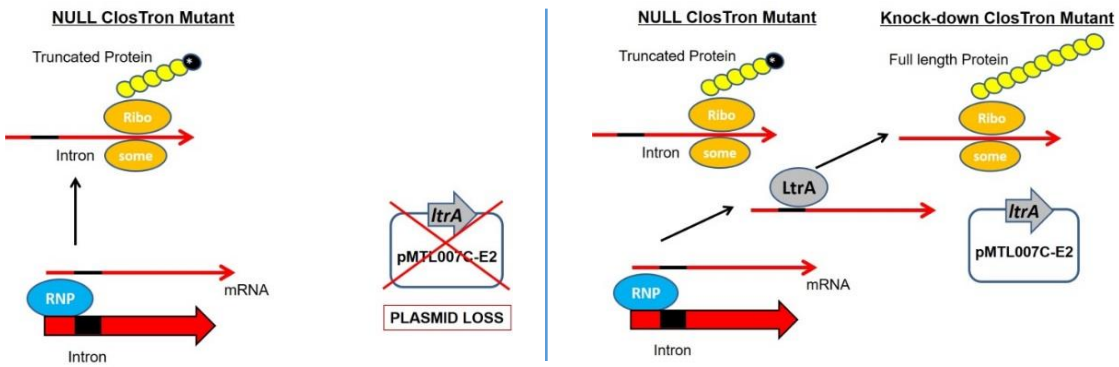


Figure 3. 1: Null and Knock-down RAMless ClosTron Mutants.

The ClosTron is an insertional mutagenesis tool for gene interruption and contains the RAM element (1162 bp). Without the RAM, the intron is a smaller insert and can consequently be easily spliced out of the transcribed mRNA in the presence of the LtrA protein in the plasmid, allowing full length protein translation (knock-down ClosTron mutant). However, with plasmid loss the intron remains inserted, protein is truncated, creating a NULL ClosTron mutant.

The RAMless Clostron has been employed by previous researchers for the disruption of lactate dehydrogenase and malate dehydrogenase genes in *Clostridium cellulolyticum* (Li *et al.*, 2012) as well as phosphotransacetylase and butyrate kinase in *Clostridium beijerinckii* (Wang *et al.*, 2013). In this thesis, the use of the conditional RAMless ClosTron which allows the initial knockdown of a gene prior to the isolation of a null mutant is explored. This system could allow the knock out of an otherwise essential gene.

3.1.2 Improving Understanding of the ABE Pathway Using Conditional RAM-less ClosTrons

Based on the previous data of Cooksley *et al.* (2012), it is clear that certain genes cannot be inactivated, notably *thlA* and *hydA*. Moreover, it is likely that *ptb* cannot be inactivated under normal conditions, and that the intron insertion clone

isolated in this study was due to the presence of the *thl* frameshift. If thiolase cannot be inactivated, then it is difficult to see why any of the genes encoding the enzymes involved in the conversion of acetyl-CoA to butryl-CoA (*thl,* thiolase; *hbd,* 3-hydroxybutyryl CoA dehydrogenase; *crt,* crotonase, and; *bcd,* butyryl-CoA deyhdrogenase) could be knocked-out. Whilst the inactivation of *hbd* was reported (Lehmann and Lutke-Eversloh, 2011), as discussed above, the authors did not complement their mutant or undertake NGS on the strain. It is conceivable that like the *ptb* mutant of Cooksley *et al.* (2012), compensatory mutations(s) had arisen (for example the *thl* frame-shift) that made the mutant viable. Moreover, all attempts at Nottingham to knock-out the *bcd* gene of ATCC 824 with the ClosTron have been unsuccessful (N. Minton, personal communication), while a new allelic exchange method led to the isolation of a cell line which appeared to carry both the wild type and mutant allele (Ryan Hope, personal communication).

Further evidence in support of this hypothesis comes from the fact that the Minton group have recently obtained the *ptb* mutant of Lehmann and Lutke-Eversloh (2011) and sequenced its genome. Of the 7 SNPs and Indels it carries distinct to the parent, ATCC 824 COSMIC sequence, 1 SNP is in the *hbd* gene. Whilst a conservative substitution (Ile237Met) it seems like more than just a coincidence that one of the 4 genes involved in acetyl-CoA to butryl-CoA conversion carries a mutation. One would predict that this variant *hbd* enzyme shows substantive differences in terms of activity compared to the parent strain.

To test this general hypothesis, the CRG have developed a novel conditional derivative of the ClosTron (pMTL007C-A2) which is both RAM-less (lacks the *ermB*::RAM) and carries a copy of the counter selection marker, the *codA* gene

(encodes cytosine deaminase) in the vector backbone. Cells carrying this plasmid will be unable to grow in the presence of 5-FC (5-Flourocytosine) as it is converted to the highly toxic 5-FU (5-Flourouracil) by cytosine deaminase. It follows that only cells that lose the plasmid will survive. To use this plasmid to isolate compensatory mutations, its intron is targeted against the gene of choice (for example *ptb*), and intron insertions isolated through appropriate screening of cells which have received the retargeted, RAM-less ClosTron plasmid. Until this point all cultures are supplemented with thiamphenicol to ensure the ClosTron plasmid, and encoded LtrA, are retained. This ensures splicing of the intron insertion from the target gene mRNA, and prevents the mutation from becoming a null mutation. Once a pure intron knock-down mutant is isolated it may be plated on media containing 5-FC. This selects for cells that have lost the ClosTron plasmid, and *ltrA,* thereby preventing splicing. If the null mutant created is lethal (for example, *ptb*), then the only way cells can survive is if they have acquired a compensatory mutation (for example, *thl*) that overcomes this lethality (see figure 3.1).

3.2 Aim of Chapter

To utilize the conditional RAM-less ClosTron system to knock out genes in *C. acetobutylicum* that had been previously impossible to knock-out using the standard ClosTron mutagenesis, and possibly isolate compensatory mutations that allow the mutant viability. The generation of these mutants and manipulation of solvent pathways is necessary in the development of an industrial solvent producing strain.

3.3 Results and Discussion

As the Knock-out (KO) of *thlA* and *ptb* using the conditional RAM-less ClosTron approach is already being undertaken within the CRG, two other genes that had previously proven impossible to KO using the standard ClosTron were chosen as targets. These were *hydA*, coding for hydrogenase and *hbd*, encoding 3-hydroxybutyryl CoA dehydrogenase as well as the crotonase (*crt*) and butyryl CoA dehydrogenase (*bcd*) genes. In order to ensure that the RAM-less procedure was working, a gene was targeted that previously proved possible to KO using the standard ClosTron (Cooklsley *et al*., 2012), namely the *adhE2* gene encoding alcohol dehydrogenase.

3.3.1 Experimental Control using *adhE2* (alcohol dehydrogenase)

The *adhE2* gene encoding alcohol dehydrogenase had previously been knocked out using ClosTron and was therefore included as a positive control. Following removal of the RAM from previously made *adhE2* retargeted ClosTron plasmid (Cooksley *et al*., 2012), the RAM-less variant, pMTL007CC-E2::adhE2696s, was transformed into *C. acetobutylicum* ATCC 824 and thiamphenicol resistant transformants selected on CGM+Tm plates. Following re-streaking to purity, four colonies were grown up in 2xYTG media and PCR screened for the presence of cells containing the intron insertion using two gene flanking primers (amplifies entire insertion) or intron specific primer (EBS) and a flanking primer (amplifies an intron junction). Plasmid curing was achieved by plating on CBM+FC plates and cells that had lost plasmid selected for by thiamphenicol

sensitivity. This analysis revealed that 3 of the 4 colonies were apparently pure intron mutants (figure 3.2) and confirmed the functionality of the RAM-less ClosTron system.

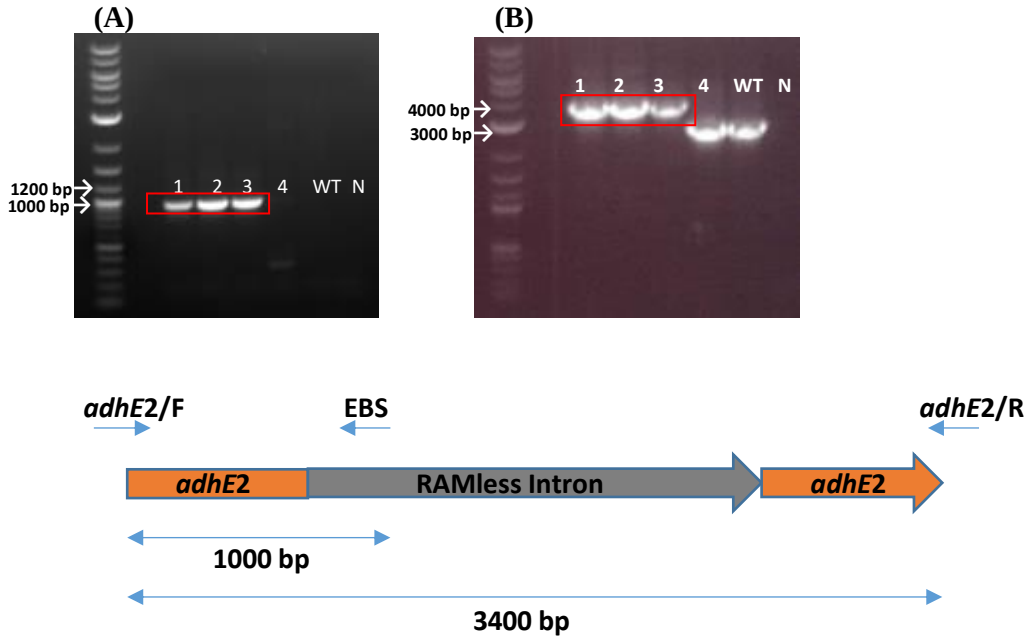


Figure 3. 2: PCR screening of *adhE2* intron insertions

PCR screening of *adhE2* intron insertions using two sets of primers: (A) EBS universal & adhE2/F; Lanes 1- 4 are genomic DNA of potential *adhE2* mutants. 1-3 give DNA fragment size of 1000 bp while wild type, WT, and negative control (just water), N, does not give any DNA fragment (B) flanking primers, adhE2/F & R; lanes 1-3 are pure *adhE2* intron insertions, DNA fragment size of 3400 bp, Wild type, WT, at 2666 bp and just water, (N) gives no fragment.

3.3.2 *hydA* (Hydrogenase) Conditional RAM-less knock-out

Following electroporation of *C. acetobutylicum* ATCC 824 with a *hydA*-retargeted conditional, RAM-less ClosTron plasmid, pMTL007CC-E2::hydA936s, transformed cells were plated onto CGM agar supplemented with thiamphenicol and 96 colonies used to inoculate a 96-well microtitre plate containing 2xYTG media. Following growth of the inoculated wells, the 96

isolates were pooled into 12 columns (1-12) and 8 rows (A-H), making a total of 20 separate pools (figure 3.3A) which were PCR screened using gene flanking primers (forward and reverse, *hyd*A/F & *hyd*A/R) as well as EBS universal and forward flanking primer.

The former pair of primers should PCR amplify a 1250 bp DNA product if the intron has inserted, or a 536 bp fragment if no intron is present. The latter pair of primers will give a 471 bp fragment if the intron has inserted, and no PCR product in its absence (figure 3.3B).

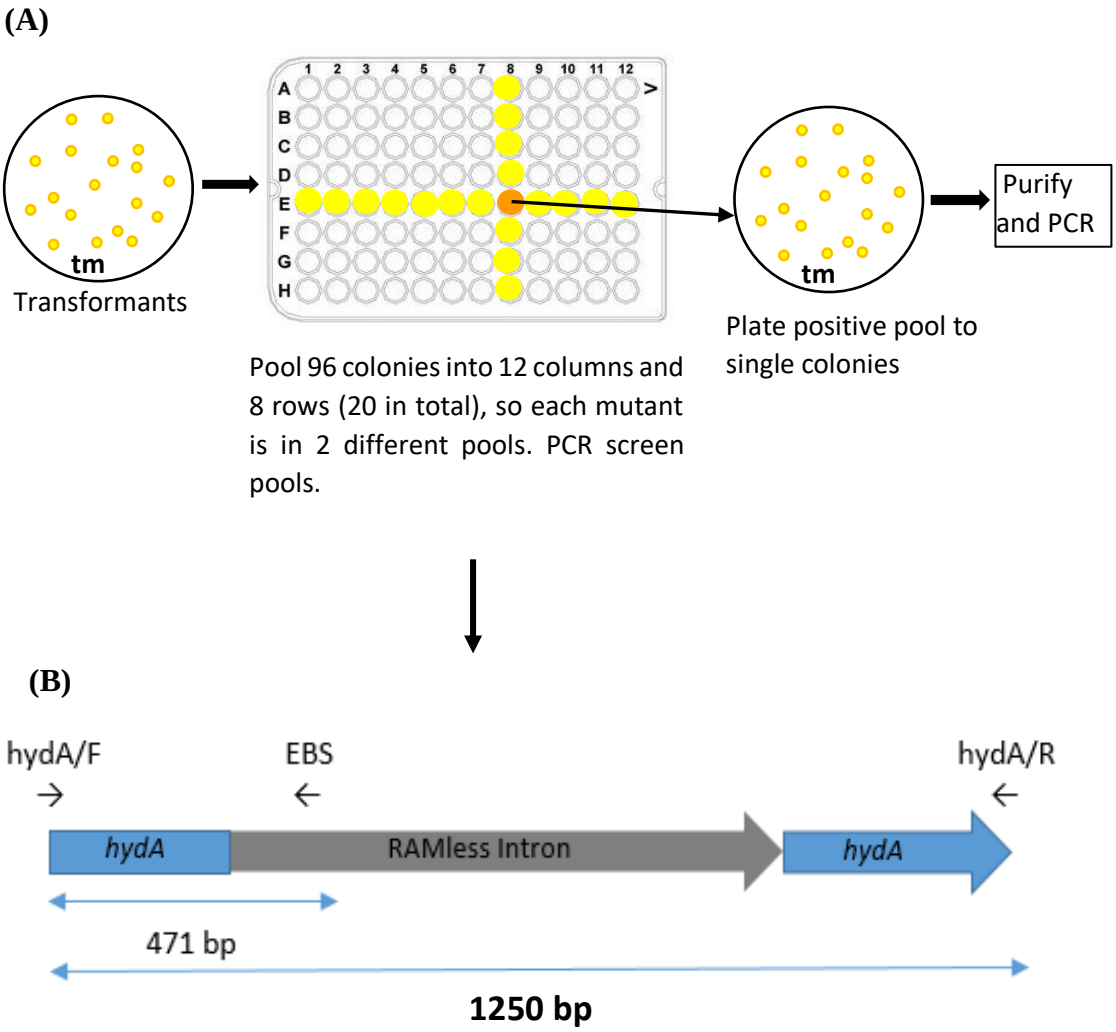


Figure 3. 3: Schematic of isolate pooling (A) and PCR screening (B) of RAM-less ClosTron insertion in *hydA* gene.

(A) Schematic of pooled isolates for PCR screening for *hydA* RAMless intron insertion. 96 colonies were inoculated into a 96-well plate and pooled into 12 columns and 8 rows, making a total of 20 and each isolate was present in 2 different pools. Following the identification of a positive pool, the pool is streaked out on a fresh plate to obtain single colonies which can then be purified and PCR screened to obtained pure mutants.

(B) Using gene flanking primers *hyd*A/F & *hyd*A/R, a DNA fragment of 1250 bp is expected while using the *hyd*A/F and EBS, a product of 471 bp is expected. The expected size of WT with no intron insertion is 1380 bp

These analyses showed that every pool of cells screened contained a population of cells that had an intron insertion. Thus, in just about every case, the EBS

universal primer and *hyd*A/F generated a 471bp fragment (figure 3.4 Ai, Bi, Ci). However, all pools also generated a wildtype sized DNA fragment consistent with the presence of cells that did not contain an intron insertion. The pools were, therefore, composed of mixtures of mutant and wildtype cells. What was unexpected, however, was the absence of a 1250 bp fragment in those PCR reactions that used the flanking primers *hyd*A/F and *hyd*A/R. This should have been obtained if the pooled cells are a mixed population. The explanation is not clear. Either the mutants are present at such a low level that they cannot be detected by PCR or the fragment being generated is too large to be effectively amplified. The latter seems unlikely given the EBS and *hyd*A/F primers generated a PCR product.

Given that the majority of the pools contain intron insertions, an attempt was made to isolate a pure clone. Cells from the Row 5 pool (figure 3.4 Ai, Bi, Ci) were plated out to single colonies on CGM agar supplemented with thiamphenicol and cells from 20 colonies subjected to PCR screening using the same two PCR primer pairs as before. It was assumed that if a cell gave an intron-specific DNA fragment with the EBS and *hyd*A/F primer pair, and no wildtype sized DNA fragment with the *hyd*A/F and *hyd*A/R primer pair, then it must be a pure intron insertion mutant. However, in every case the screened cells all appeared to be composed of a mixture of mutant and wildtype (figure 3.4 Aii). Two further attempts were made to isolate a pure intron insertion mutant, through the serial dilution of a randomly selected mixed population to single colonies and the subsequent screening of a random 20 colonies (figure 3.4 Bii and Cii). In no instance was evidence obtained for the presence of a pure intron insertion clone. While re-streaking failed to isolate pure *hydA* insertion mutants, it would have been useful to screen the isolates prior to pooling. At this

stage it would appear that the proportion of cells that carry an intron insertion in *hydA* is too low to be isolated.

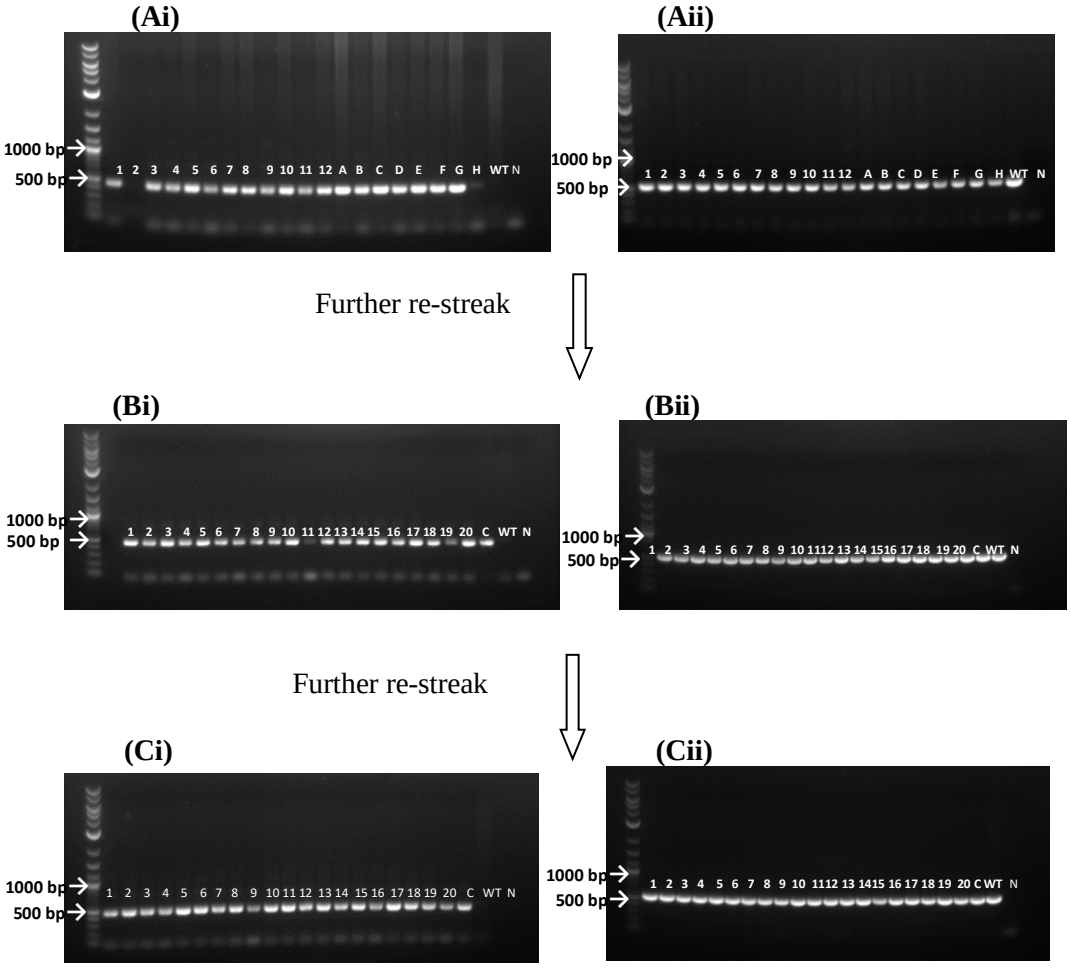


Figure 3. 3: PCR screening for *hydA* intron insertion

PCR screening for *hydA* RAMless intron insertion. The 20 lanes are genomic DNA of 20 separate pools made; consisting of 12 columns (1-12) and 8 rows (A-H). Screening was done using two sets of primers: (1) EBS universal & hydA/F (Ai, Bi, Ci); DNA fragment is seen at 472 bp in all 20 lanes indicating intron insertion except in wild type, WT. (2) Using flanking primers, hydA/F& hydA/R, (Aii, Bii, Cii) there is no evidence of insertion as expected DNA fragment should be at 1250 bp. Expected WT DNA fragment size is 536 bp. Figures B and C represent further passages of colonies from A aimed at isolating a pure clonal population.

3.3.3 *hbd* (3-hydroxybutyryl CoA dehydrogenase) Conditional RAM-less ClosTron Knock-out

The RAM-less ClosTron plasmid retargetted for *hbd*, pMTL007CC-E2::hbd69s, was electroporated into *C. acetobutylicum* ATCC 824 using the transformation procedure previouly described. In other work involving isolation of the *ptb* RAM-less mutant, intron insertions had been isolated with relative ease (Ying Zhang, personal communication). Therefore, rather then begin by screening pools of clones, initial work focused on screening single clones. Accordingly, twelve transformants from CGM+Thiamphenicol plates were picked and PCR screened using two pairs of primers; hbd/F & EBS and hbd/F & hbd/R. All colonies showed evidence of the presence of a population of cells carrying intron insertions as well as wildtype cells using both primers (figure 3.5 A & B). Colonies 1,8 and 11 were selected, plated out to single colonies and PCR screened in order to isolate a pure *hbd* clonal population (figure 3.5 C).

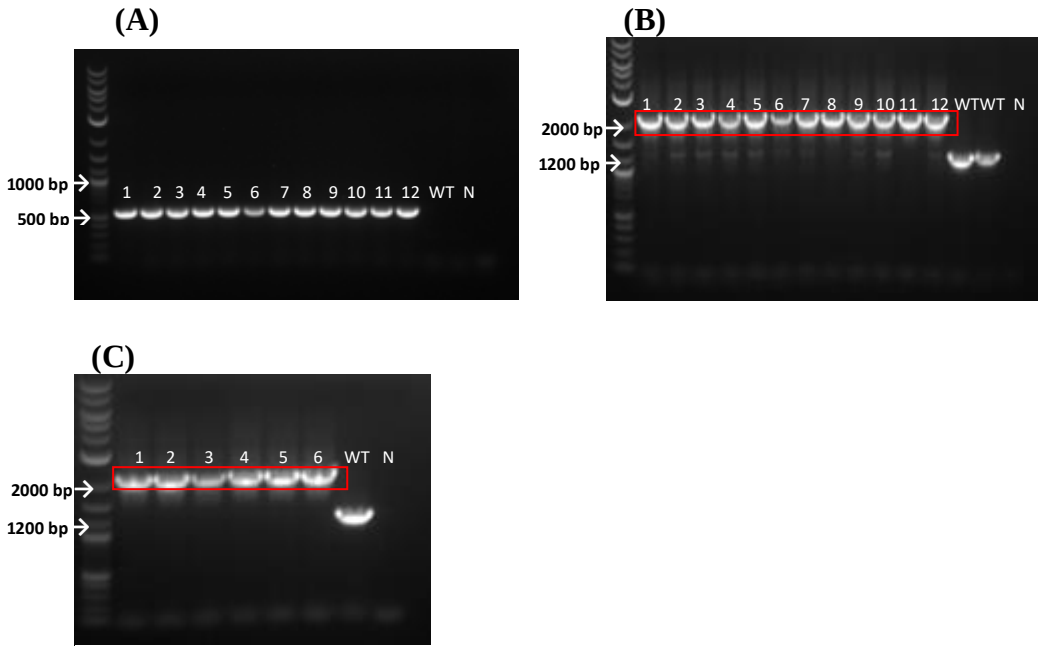


Figure 3. 4: (A) PCR screening of RAM-less ClosTron knock-out of *hbd* gene

PCR screening of RAM-less ClosTron knock-out of *hbd* gene using two sets of primers. Lanes represent genomic DNA of potential mutants. (A) EBS & hbd/F; lanes 1-12 give DNA fragment size of 548bp while wild type, WT, and negative control, N, do not give any DNA fragment. (B) PCR screening using *hbd* gene flanking primers, hbd/F and hbd/R; Lanes 1-12 give DNA fragment size of 2007 bp, Wild type, WT, at 1285bp and negative control, (N) gives no band. (C) Further purification of colonies from (B) and PCR screening with flanking primers give pure *hbd* insertion as seen in lanes 1-6 which give DNA fragment size of 2007 bp. WT at 1285 bp and no band for negative control, N.

Colonies 1, 3 and 5 were further plated out to get single colonies. Plasmid (containing *codA*) loss was selected by plating on CBM+FC plates using CBM only as control. About 1-2 log decrease in cell numbers was observed when adding FC and patch plating on CGM±Tm was undertaken. The *hbd* mutant which had become sensitive to thiamphenicol was selected, designated hbd69s::rCT and subjected to batch fermentation. Solvent analysis and Illumina sequencing data are discussed in section **3.3.12.**

3.3.4 *crt* (crotonase) Conditional RAM-less Knock-out

Retargeted RAM-less plasmid for *crt*, pMTL007CC-E2::crt154s, was cloned and two independent transformations carried out. After electroporation, 96 transformants were picked into 96-well plate containing 2 x YTG+Tm. PCR screening was done using two sets of primers: EBS universal and forward flanking primer (*crt*/F) and flanking primers, crt/F & *crt*/R. Using EBS and *crt*/F, all 20 lanes showed correct intron insertion of DNA fragment size 560 bp and no band for WT as expected. However, using gene flanking primers, *crt*/F and *crt*/R, no intron insertion was observed at 2200 kb and all lanes appear as wild type. This is shown in figure 3.6.

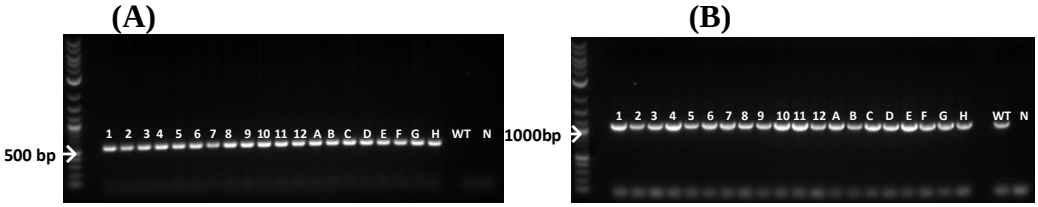


Figure 3. 5: (A) PCR screening of RAM-less ClosTron knock out of *crt* gene

PCR screening of RAM-less ClosTron knock out of *crt* gene. Lanes 1-12 and A-H are genomic DNA of 96 isolates pooled into columns and rows (20 pools in total). Two sets of primers were used: (A) EBS & crt/F; lanes 1-H give DNA fragment size of 560 bp while wild type, WT, and negative control, N, do not give any DNA fragment. (B) Using two gene flanking primers; *crt*/F and *crt*/R; no expected intron insertion at 2200 kb was observed (1-H), rather all screened colonies appear as WT. Water was used as negative control, N. DNA fragments from A were Sanger sequenced and confirmed to have the expected nucleotide sequence.

Selected lanes were further passaged in order to isolate pure mutants, 20 colonies picked and PCR screened using flanking primers. Similar results as previously

obtained for *hydA* gene was observed with no evidence of insertion using flanking primers.

3.3.5 *crt* 307s Conditional RAM-less Knock-out

Aside the crt 154s retargeted RAM-less plasmid, another target site, crt 307s, was chosen to determine if this will make the knock-out possible. The retargeted *crt* plasmid pMTL007CC-E2::crt307s was cloned, methylated and transformed into *C. acetobutylicum*. Using the 96-well plate, transformants were selected and PCR screened for insertions using both using flanking primers *crt*/F and *crt*/R and EBS universal & *crt*/F. However, isolation of a pure intron mutant was still not possible (Data not shown).

3.3.6 *bcd* (butyryl CoA dehydrogenase) Conditional RAM-less Knock-out

Retargeted RAM-less KO plasmid for *bcd*, pMTL007CC-E2::hbd496s was cloned and appropriately methylated. Two independent transformations into *C. acetobutylicum* were carried out. After electroporation, 96 transformants were picked into 96-well plate containing 2YTG + Tm. PCR screening was done using two sets of primers: EBS universal and forward flanking primer (*bcd*/F) and flanking primers, bcd/F & *bcd*/R.

Using EBS universal and forward flanking primer (*bcd*/F), all 20 lanes showed correct intron insertion at 1063 bp with no insertion for wild type. However, using two gene flanking primers; *bcd*/F and *bcd*//R, no intron insertion was observed as all DNA fragments appear just as wildtype at 1552 bp. Two lanes were selected, further passaged, 96 colonies picked and PCR screened but still no evidence using flanking primers. This is shown in figure 3.7.

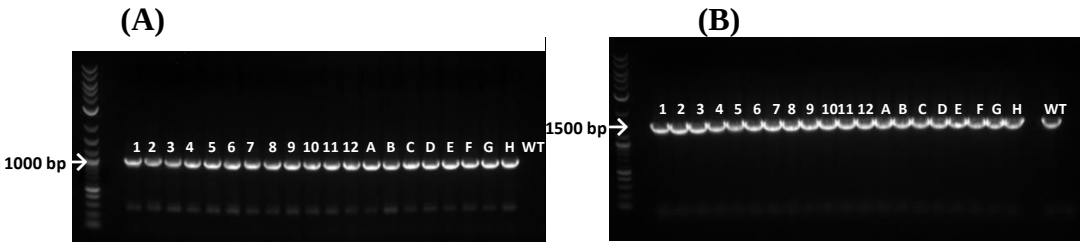


Figure 3. 6: PCR screening of RAM-less ClosTron knock out of *bcd* gene

PCR screening of RAM-less ClosTron knock out of *bcd* gene. Lanes 1-12 and A-H are genomic DNA of 96 isolates pooled into columns and rows (20 pools in total). Two sets of primers were used: (A) EBS & bcd/F; lanes 1-H give DNA fragment size of 1063 bp. while wild type, WT does not give any DNA fragment. (B) Using two gene flanking primers; *bcd*/F and *bcd*/R, all screened colonies do not show evidence of insertion but appear at 1552 bp just as WT. Water was used as negative control, N.

Two lanes were further passaged and PCR screened but results came out similar to previous screening with no evidence of insertion using gene flanking primers. To confirm that the EBS and bcd/F primers had correctly amplified the intron insertion, six (6) of the DNA fragments were cut, recovered and sent out for Sanger sequencing using the same primers. Nucleotide sequence was as expected confirming correct intron insertion.

3.3.7 Re-screening of *bcd* intron insertion using different primers

A second set of flanking primers, bcd3F and bcd3R2 were designed to give smaller bands compared with bcd/F and bcd/R. it was thought that perhaps if the mutant bands to be amplified were just about 1,300 bp rather than 2,300, these rare mutant occurrences may be easier to isolate. In figure 3.8, using the flanking primers, there is still no evidence of insertion.

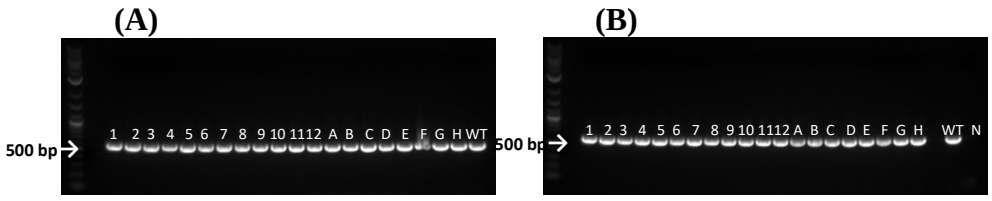


Figure 3. 7: *bcd* gene intron insertion screening

PCR screening for *bcd* gene intron insertion using a new set of flanking primers, bcd3F and bcd3R2 Lanes 1-12 and A-H are genomic DNA of 96 isolates pooled into columns and rows (20 pools in total). All bands after transformation (A) and further passaging (B) appear as WT at 580 bp. With the right intron insertion, DNA fragment of 1300 bp was expected.

3.3.8 Further Evidence of the Functionality of the Conditional RAM-less ClosTron System

So far in this chapter the conditional RAM-less ClosTron KO of *hydA, hbd, crt* and *bcd* genes have been attempted with only the *hbd* KO being successful. As a control that the methodologies carried out were being performed correctly, the knock-out of the *thl* gene that had previously been impossible to knock-out (Cooksley *et al.*, 2012) but has recently been achieved in the SBRC (Ying Zhang) using the RAM-less ClosTron, was carried out.

3.3.9 Control thiolase (*thl*) RAM-less KO

Methylated *thl* RAM-less plasmid was transformed into *C. acetobutylicum*. 96 transformants were picked and screened using EBS and thl/F and flanking primers, thl/F and thl/R. There is evidence of insertion using both sets of primers (figure 3.9 A & B). Faint mutant bands larger than wild type are observed using flanking primers (figure 3.12 B) and these were further passaged to isolate pure mutants (figure 3.9 C). RAM-less ClosTron *thl* mutants were designated *C. acetobutylicum* thl406s::rCT

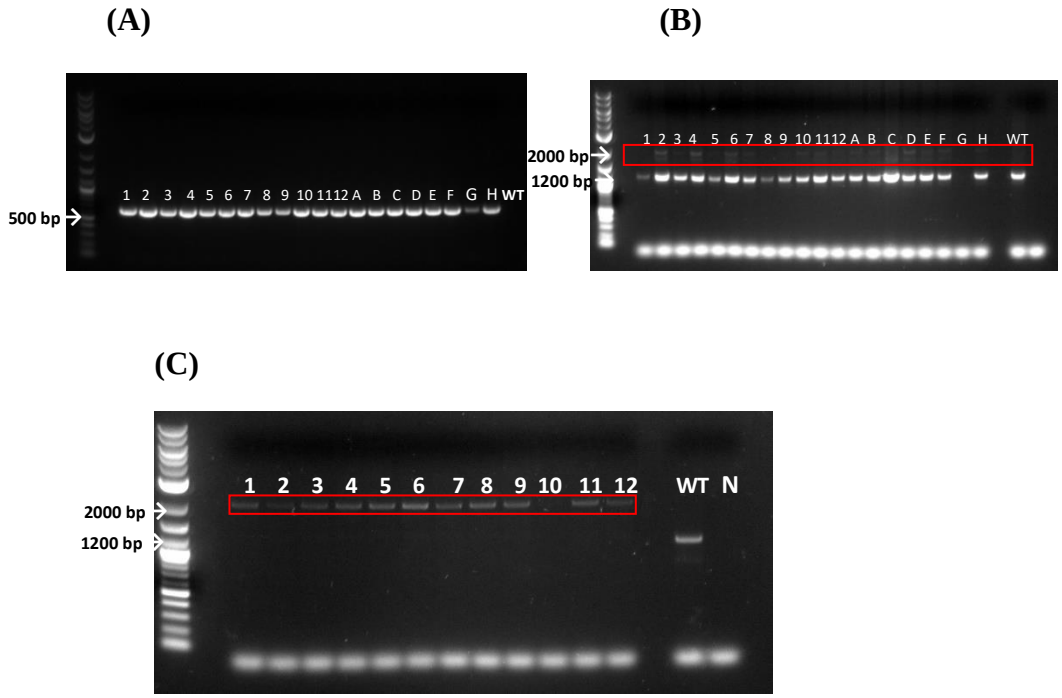


Figure 3. 8: (A) PCR screening of RAM-less ClosTron knock out of *thl* gene

PCR screening of RAM-less ClosTron knock out of *thl* gene. Lanes are genomic DNA of potential mutants. In (A) and (B), Lanes 1-12 and A-H are genomic DNA of 96 isolates pooled into columns and rows (20 pools in total).Using two sets of primers: (A) EBS & thl/F, lanes 1-H give DNA fragment size of 500 bp while wild type, WT does not give any DNA fragment. (B) Using two gene flanking primers; thl/F and thl/R, faint mutant DNA fragments in addition to the wild type is observed. (C) Pure *thl* mutants obtained by passaging the mixed population (4D and 2F). DNA fragments are of size 1987 bp while WT is 1267 bp. Purified DNA fragments were Sanger sequenced and nucleotide sequence were as expected. Water was used as negative control, N.

In addition to the *thl* RAM-less ClosTron KO, it was also important to show again that the *hbd* gene which has now been knocked out by the RAM-less ClosTron could not be knocked out by other means hence the KO of *hbd* was attempted using the standard ClosTron as well as CRISPR/Cas9 systems.

3.3.10 Standard ClosTron *hbd* Knock-out

Two independent transformations into *C. acetobutylicum* using pMTL007C-E2::hbd69s was carried out and twenty (20) Em^R transformants PCR screened using flanking primers hbdF/R but there was no evidence of intron insertion observed (figure 3.10).

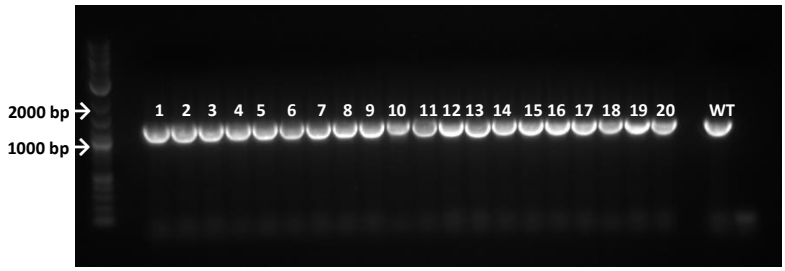


Figure 3. 9: *hbd* standard ClosTron screening

Screening of standard ClosTron insertion in the *hbd* gene. Genomic DNA of 20 potential candidates were screened using gene flanking primers hbdF/R. There is no evidence of insertion as all colonies screened have DNA fragment size similar to wild type (1285 bp). With insertion, expected DNA fragment size should be 3169 bp.

3.3.11 Using the CRISPR/Cas9 System

As a CRISPR/Cas9 knock out system had recently been developed within the SBRC (Cañadas *et al.*, 2019), an attempt to knock out *hbd* using CRISPR was undertaken. But first, it was important to demonstrate that the developed CRISPR/Cas9 system was functional in *C. acetobutylicum*. As a target, therefore, the *pyrE* gene was selected as it has been previously shown to be easily knocked out using Allele Coupled Exchange, ACE.

3.3.11.1 CRISPR KO of Control Gene, *pyrE*

To test that CRISPR/Cas9 works in *C acetobutylicum*, the *pyrE* gene was targeted. pMTL8315_RibocasE_pyrE plasmid (figure 3.11) was generated by

cloning in sgRNA and homology arms (between sites *Spe*I and *Asc*I) targeted at *pyrE* gene in plasmid, pMTL8315_RibocasE_J23119_MCS originally created by Ines Canadas. pMTL8315_RibocasE_pyrE was methylated and electroporated into competent cells of *C. acetobutylicum*. Transformants were plated onto CGM Tm (non- induced) and CGM Tm Theophiline (Thp) induced plates (5mM). After 48 hours, there were 6 colonies on non-induced plates and 1 colony on induced plate. Two types of colonies were observed on the plates. On the non-induced plates colonies were flat and seemed not to be growing so well (indicating colonies had now become auxotrophs) while on the induced plates, the colonies were raised and larger.

Colonies from the non-induced were re-streaked onto induced plates and PCR screened using primers flanking the homology arms, pyrE-HA/F&R. Of the 14 candidates screened, 2 had clean *pyrE* deletions with DNA fragment size at 1533 bp indicating successful deletion of the 675bp *pyrE* gene (figure 3.12). These two candidates were Sanger sequencing and nucleotide sequence shown to be exactly as expected.

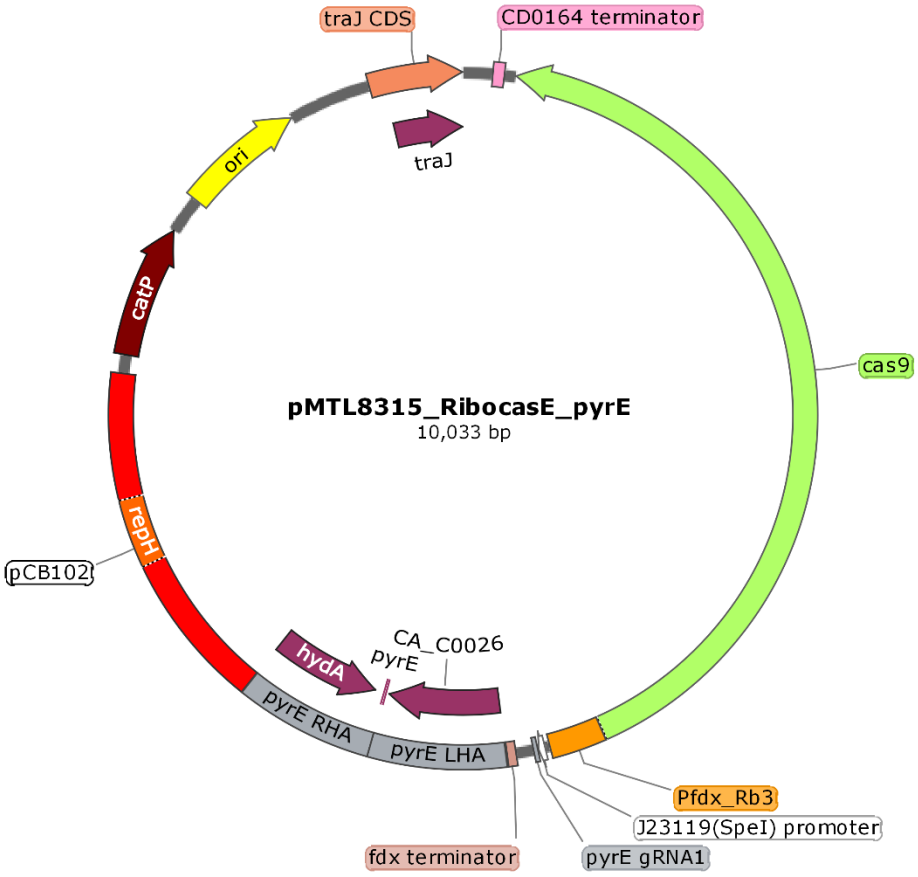


Figure 3. 10: pMTL8315_ RibocasE_pyrE; CRISPR/cas9 knock-out vector for *pyrE* gene

Comprises the Cas9 nuclease downstream of P_{fdx} fused to a riboswitch, J23119 promoter which drives the sgRNA expression; *pyrE* RHA and *pyrE* LHA are regions of homology flanking the *pyrE* gene to be deleted; pCB102 is Gram-positive replicon; *catP* is chloramphenicol antibiotic marker; ColE1 is Gram-negative replicon and *traJ* is transfer gene.

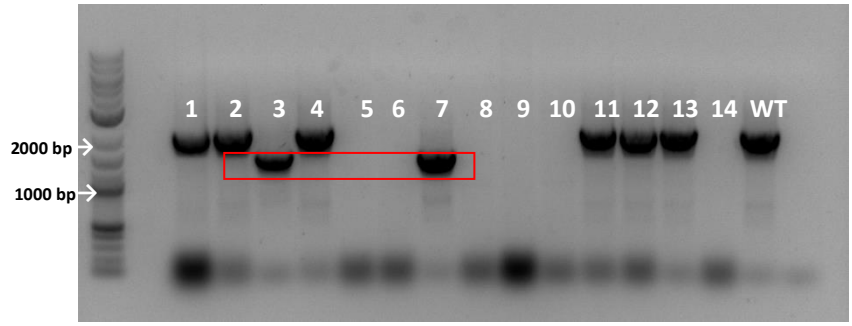


Figure 3. 11: CRISPR/Cas9 *pyrE* deletion screening

Candidates for CRISPR/Cas9 *pyrE* deletion were PCR screened using primers pyrE-HA/F&R, flanking both HAs. Lanes are genomic DNA of potential mutants. Lanes 3 and 7 give smaller DNA fragment size at 1533 bp indicating deletion 675bp *pyrE* gene while other candidates are of same DNA fragment size as WT (2196 bp). 3 and 7 were subjected to Sanger sequencing and showed nucleotide sequence exactly as expected.

Having established that CRISPR KO could be used to effectively knock out a gene in *C. acetobutylicum*, the knock-out of *hbd* gene was attempted.

3.3.11.2 *hbd* CRISPR KO

hbd CRISPR vector, pMTL8315_ RibocasE_hbd was created by replacing the sgRNA and homology arms (between sites *SpeI* and *AscI*) targeted at *hbd* in pMTL8315_pyrE. This was transformed into *C. acetobutylicum* and plated onto CGM Tm (non-induced) and CGM Tm Thp (5mM induced). About 1-2 colonies grew per CGM Tm plate and were re-streaked on CGM Tm Thp plates. Colonies were PCR screened using primers flanking the homology arms, hbd/HA/F and R. Figure 3.13 shows that all screened candidates had no evidence of deletion as they all have similar DNA fragment size as the wild type parent strain.

While the knock out of the *pyrE* gene by the CRISPR approach was successful at first attempt, the *hbd* gene could not be inactivated using the CRISPR system neither could it be knocked out using the standard ClosTron methodology, except by the conditional RAMless ClosTron. This again goes to show that the

hbd gene like the other genes in the C4 central metabolic pathway in *C. acetobutylicum* (*thl*, *crt* and *bcd*) may be impossible to knock-out except some compensatory mutation have arisen which allowed the viability of the mutant.

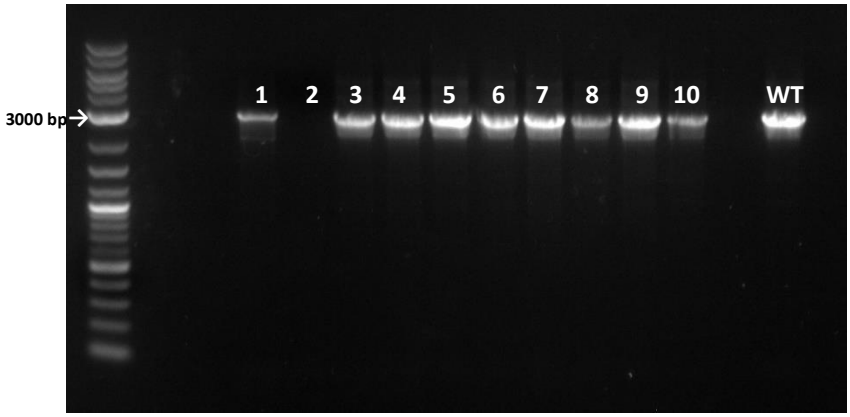


Figure 3. 12: CRISPR/Cas9 *hbd* deletion screening

PCR screening of genomic DNA of 10 candidates for CRISPR/Cas9 deletion in *hbd* gene using primers hbd/HA/F and R. All screened candidates show no evidence of deletion and have similar DNA fragment size as the wild type parent strain, 3072 bp. Expected DNA size with *hbd* gene deletion should be 2235 bp.

3.3.12 Characterisation of *hbd* KO Mutant, hbd69s::rCT

Batch fermentation in 60 ml CBMS was undertaken with the *hbd* knock-out mutant obtained by conditional RAM-less ClosTron, hbd69s::rCT, and samples taken at intervals for measurement of OD, pH and solvent profile analysis. HPLC and GC were used to analyse solvents. Results obtained are given in figure 3.14.

It was observed that hbd69s::rCT initially grew slower than the wild type *C. acetobutylicum* but went on to higher OD values. There was no butyrate produced by hbd69s::rCT compared with the wild type which produced up to 28 mM butyrate at the peak of acidogenesis. More acetate (6. 25 mM) was produced

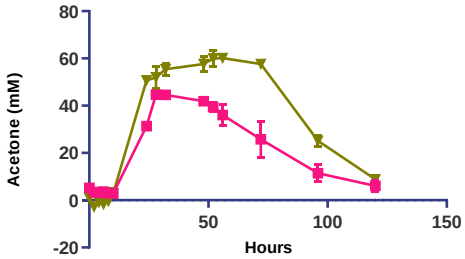
in this mutant likely to make up for the lack of ATP generation through butyrate production.

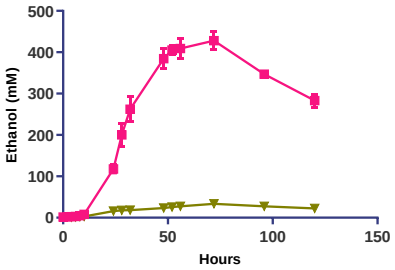
For solvent production, worthy of note is the greatly increased ethanol titre in hbd69s::rCT; up to 427 mM was produced compared with wild type which made 33 mM ethanol. As the C4 metabolic pathway had been inactivated with the knock-out of *hbd* gene, the flux went into the production of more ethanol. Previous researchers have reported that *C. acetobutylicum* could be switched into an ethanol producer by the disruption of the butyrate pathway through the knock down of the *hbd* gene (Lehmann and Lutke-Eversloh, 2011). Ethanol was also reported to be the major fermentation product (113 mM) in the *C. acetobutylicum ptb* ClosTron mutant which carried a frameshift mutation in *thl* gene (Cooksley *et al*., 2012).

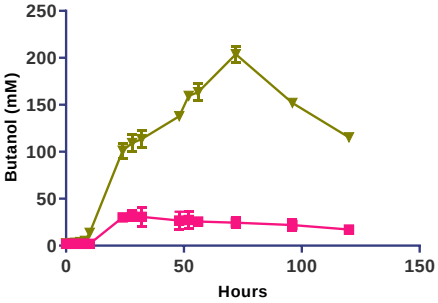
For butanol production, in comparison with wild type which produced 203 mM butanol, hbd69s::rCT produced 32 mM. This was quite surprising as hbd69s::rCT was not expected to be able to produce butanol considering that the *hbd* gene had been inactivated. A possible explanation for this could be that in *C. acetobutylicum*, there are two chromosomal genes coding for 3-hydroxybutyryl-CoA dehydrogenase, responsible for the conversion of acetoacetyl-CoA to 3- hudroxybutyryl-CoA: 3-hydroxybutyryl-CoA dehydrogenase, *hbd,* (CAC2708) and 3-hydroxyacyl-CoA dehydrogenase, *mmgB*, (CAC2009) (Nolling *et. al*., 2001). Grimmler and co-workers noted that while *hbd* was significantly induced during acidogenesis and repressed with a drop in pH, *mmgB* was induced during the acidogenesis-solventogenesis switch and remained expressed all through solventogenesis (Grimmler *et al* 2011). This possibly explains the absence of butyrate (produced during acidogenesis) in the

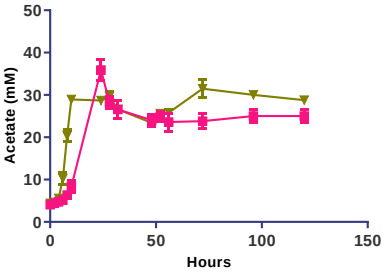
hbd mutant while some butanol (produced during solventogenesis) was detected likely due to the activity of *mmgB.*

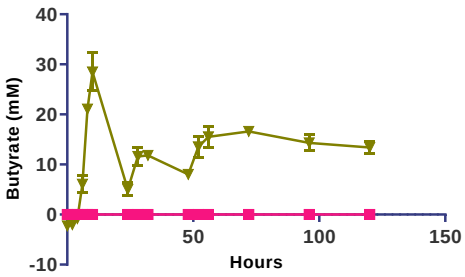
Acetone production was lower in the mutant. Here, the production of acetone may be less desirable as there is a need to produce more reduced products. This could also be related to a possible decrease in the activity of thiolase, the first enzyme in the *C. acetobutylicum* C4 metabolic pathway which branches off to the production of acetoacetate and decarboxylation to acetone. A decrease in the thiolase activity and general reduction of flux towards the C4 pathway has been reported in the *hbd* knock down created by other researchers. (Lehmann and Lutke-Eversloh, 2011)

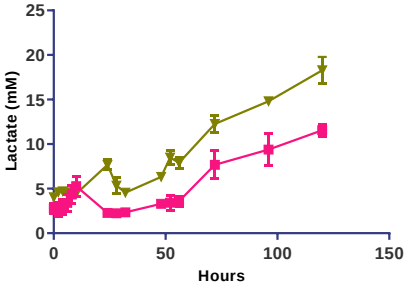

Acetone (mM)
80
60
40
20
0
-20
50
100
150
Hours

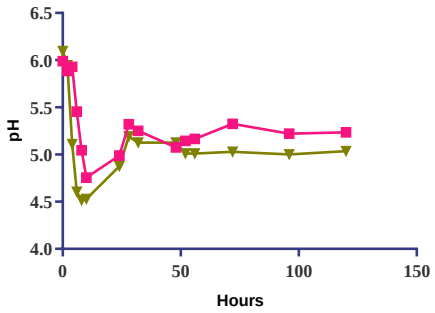

Ethanol (mM)
500
400
300
200
100
0
0
50
100
150
Hours

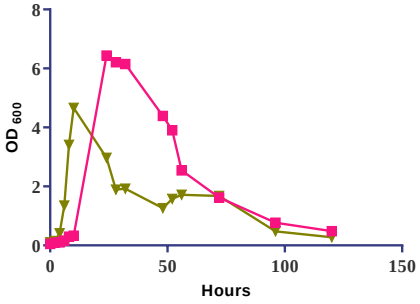

Butanol (mM)
250
200
150
100
50
0
0
50
100
150
Hours

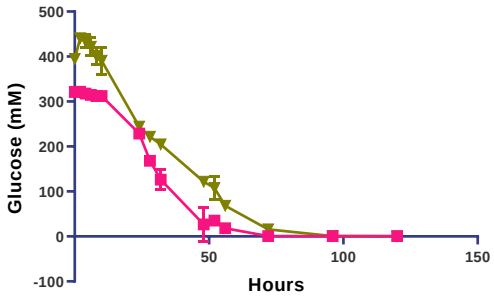

Acetate (mM)
50
40
30
20
10
0
0
50
100
150
Hours


Butyrate (mM)
40
30
20
10
0
-10
50
100
150
Hours


Lactate (mM)
25
20
15
10
5
0
0
50
100
150
Hours


pH
6.5
6.0
5.5
5.0
4.5
4.0
0
50
100
150
Hours


OD 600
8
6
4
2
0
0
50
100
150
Hours


Glucose (mM)
500
400
300
200
100
0
-100
50
100
150
Hours


hbd69s::CT
Wild Type

Figure 3. 13: Solvent profile of hbd69s::rCT

Solvent profile of batch fermentation of *hbd* mutant, hbd69s::rCT, obtained by RAM-less ClosTron insertion. Samples were taken at varying time points during fermentation to monitor growth, pH and OD. Solvents were analysed using HPLC and GC. Values are mean of duplicate values**.**

Having obtained pure *hbd* mutants, the next step was to sequence the genomes for identification of any compensatory mutations that may have arisen. Unfortunately, sequencing results revealed a big deletion downstream of the *abrB* gene (regulators of stationary/sporulation gene expression) and this was also observed in the parent wild type *C. acetobutylicum* used. This deletion was also PCR confirmed. It is not known what effect this deletion might have on the mutants, consequently, an additional four independent transformations were carried out with a new wild type strain. Pure *hbd* knock-out mutants were isolated and seven mutants from the four independent transformations subject to paired-end Illumina sequencing. The sequences obtained were mapped to the annotated genome reference sequence of ATCC 824 and the results presented in table 3.1.

Table 3. 1: Genomic Sequencing of *hbd* Knock-out Mutants

1

Ref Positn	Type	Ref	Allele	Overlapping annotations	Coding region change	Amino acid change	Gene Function
363789	Insertion	-	G				Part of a deletion downstream of AbrB
363790	SNV	T	A				Part of a deletion downstream of AbrB
363794	Deletion	A	-				Part of a deletion downstream of AbrB
506312	SNV	C	A	CDS: CA_C0437, Gene: CA_C0437	NP_347077.1:c.1855C>A	NP_347077.1:p.Leu619Ile	Histidine kinase
1217953	SNV	A	G	CDS: CA_C1065, Gene: CA_C1065	NP_347700.1:c.9A>G		Uncharacterised protein
1971128	SNV	T	A				Downstream phosphocarrier protein
2802560	SNV	G	T	CDS: CA_C2681, Gene: CA_C2681	NP_349288.1:c.166G>T	NP_349288.1:p.Glu56*	Hypothtical protein
3778758	SNV	C	T	CDS: CA_C3579, Gene: CA_C3579	NP_350162.1:c.272G>A	NP_350162.1:p.Arg91Lys	MarR transcriptional regulator

2

Ref Positn	Type	Ref	Allele	Overlapping annotations	Coding region change	Amino acid change	Gene Function
90941	SNV	C	A	CDS: CA_C0082, Gene: CA_C0082	NP_346727.1:c.1175C>A	NP_346727.1:p.Ala392Asp	Predicted membrane downstream agrA
363789	Insertion	-	G				Part of a deletion downstream of AbrB
363790	SNV	T	A				Part of a deletion downstream of AbrB
363794	Deletion	A	-				Part of a deletion downstream of AbrB
3778319	SNV	G	A	CDS: fabH, Gene: fabH	NP_350161.1:c.242C>T	NP_350161.1:p.Thr81Ile	3-oxoacyl-[acyl-carrier-protein] synthase 3

3

Ref Positn	Type	Ref	Allele	Overlapping annotations	Coding region change	Amino acid change	Gene Function
1209596	SNV	G	A	CDS: CA_C1053, Gene: CA_C1053	NP_347689.1:c.1401G>A		LPS glycosyltransferase
2013807	SNV	G	T				Upstream hypothetical protein
2802565	Insertion	-	GTTATATAGCAACA	CDS: CA_C2681, Gene: CA_C2681	NP_349288.1:c.171_184dup	NP_349288.1:p.Asn62fs	Hypothetical protein

4

Ref Positn	Type	Ref	Allele	Overlapping annotations	Coding region change	Amino acid change	Gene Function
88358	SNV	C	T	CDS: agrC, Gene: agrC	NP_346725.1:c.742C>T	NP_346725.1:p.His248Tyr	Histidine kinase
348714	SNV	T	G				Downstream hypothetical protein
348937	SNV	C	A				Downstream hypothetical protein
363789	Insertion	-	G				Part of a deletion downstream of AbrB
363790	SNV	T	A				Part of a deletion downstream of AbrB
363794	Deletion	A	-				Part of a deletion downstream of AbrB
2802622	Deletion	A	-	CDS: CA_C2681, Gene: CA_C2681	NP_349288.1:c.230delA	NP_349288.1:p.Asn77fs	Hypothetical protein
3007065	SNV	C	T	CDS: CA_C2873, Gene: CA_C2873	NP_349476.1:c.52G>A	NP_349476.1:p.Gly18Arg	Acetyl CoA acetyltransferase (thiolase)
3736454	Deletion	A	-				Hypothetical protein

5

Ref Positn	Type	Ref	Allele	Overlapping annotations	Coding region change	Amino acid change	Gene Function
88358	SNV	C	T	CDS: agrC, Gene: agrC	NP_346725.1:c.742C>T	NP_346725.1:p.His248Tyr	Histidine kinase
348714	SNV	T	G				Downstream hypothetical protein
348937	SNV	C	A				Downstream hypothetical protein
363789	Insertion	-	G				Part of a deletion downstream of AbrB

363790	SNV	T	A				Part of a deletion downstream of AbrB
363794	Deletion	A	-				Part of a deletion downstream of AbrB
3006990	SNV	T	A	CDS: CA_C2873, Gene: CA_C2873	NP_349476.1:c.127A>T	NP_349476.1:p.Ile43Leu	Acetyl CoA acetyltransferase (thiolase)
3775907	SNV	C	T	CDS: fabD, Gene: fabD	NP_350158.1:c.400G>A	NP_350158.1:p.Ala134Thr	Malonyl CoA-acyl carrier protein transacylase

6

Ref Positn	Type	Ref	Allele	Overlapping annotations	Coding region change	Amino acid change	Gene Function
3006696	SNV	C	A	CDS: CA_C2873, Gene: CA_C2873	NP_349476.1:c.421G>T	NP_349476.1:p.Asp141Tyr	Acetyl CoA acetyltransferase (thiolase)

7

Ref Positn	Type	Ref	Allele	Overlapping annotations	Coding region change	Amino acid change	Gene Function
2802463	Deletion	A	-	CDS: CA_C2681, Gene: CA_C2681	NP_349288.1:c.72delA	NP_349288.1:p.Lys24fs	Hypothetical protein
3007050	SNV	C	A	CDS: CA_C2873, Gene: CA_C2873	NP_349476.1:c.67G>T	NP_349476.1:p.Asp23Tyr	Acetyl CoA acetyltransferase (thiolase)

Key

Yellow	Common to 4 mutants from 4 independent transformations
Blue	Common to 4 mutants from 2 independent transformations
Green	Common to 4 mutants form 2 indpendent transformations
White	Unique to individual transformation

There were seven mutants from four independent transformations. Mutants 1 and 2 were from transformation one, 3 from transformation two, 4 and 5 from transformation three while 6 and 7 were from the fourth transformation. Further details on the genes with mutations are given below.

CA_C0437 – This is one of the five orphan histidine kinase in *C. acetobutylicum* (Nolling *et al.*, 2001) which control the phosphorylation of *Spo0A* transcription factor which regulates the endospore formation (Durre, 2011). It has been suggested that CAC_0437 could be a modulator which prevents sporulation as a *C. acetobutylicum* CAC_0437 mutant was observed to be a hyper-sporulator (Steiner *et al.*, 2011) while sporulation reduced to 0.02 % of wild type level when multiple copies of CAC_0437 were introduced into wild type.

CA_C1065 - Uncharacterized protein, related stage III sporulation protein AH of *Bacillus* sp

CAC_2681 - In multiple *Clostridium* species including *C. pasteurianum, C. felsinium C. amyloliticum C. akagii,* this protein is a DUF4325 STAS-like domain-containing protein which is functionally uncharacterized (Marchler-Bauer *et al.,* 2017). In *E coli*, the STAS domain of the YchM protein (of the SLC26, SulP family), has been shown to interact with acyl carrier protein (ACP), which carries acyl intermediates during acylation processes such as fatty acid biosynthesis, FAB (Babu *et al.,* 2010).

CAC_3579 - Is a transcriptional regulator, MarR/EmrR family possessing DNA-binding transcription factor activity

CA_C0082 - Predicted membrane protein downstream of *agrA* gene

***fabH* gene** - The *fabH* gene also known as 3-oxoacyl-[acyl-carrier-protein] synthase 3 catalyzes the condensation reaction of malonylacyl carrier protein and an acyl-CoA substrate. This is the first condensation reaction in the fatty acid biosynthesis and determines the synthesis of straight or branched-chain fatty acids based on substrate specificity. It belongs to the thiolase-like superfamily (Khandekar *et al*., 2001; Tomas *et al*., 2004).

***fabD* –** This is also known as Malonyl CoA-acyl carrier protein transacylase catalyses the conversion of malonyl-CoA to malonyl-ACP which is the second step in fatty acid biosynthesis after acetyl-CoA is converted to malonyl-CoA by acetyl-CoA carboxylase (Zhang *et al*., 2012).

agrC - Is a histidine kinase-like ATPase and constitutes a part of the agr system originally discovered in *Staphylococcus aureus* and has been described to regulate granulose formation and sporulation in *C. acetobutylicum* (Steiner *et al*., 2012). The *agr* locus is made up of *agrB*, *agrD*, *agrC* and *agrA* genes and constitutes the cell to cell signalling system while the RNAIII is the actual effector for the regulation of target gene (Novick and Geisinger, 2008). The *agrD* contains the Autoinducing Peptide (AIP) while *agrB* is a membrane associated protein. Following extracellular AIP accumulation, *agrC*, which is the histidine kinase, phosphorylates the response regulator, *agrA* which goes on to activate the *agrBDCA* operon and induce expression of RNAIII. In addition, the agr-system also controls other target genes independently of RNAIII (Queck *et al*., 2008).

Acetyl-CoA acetyltransferase, *thlA* (EC:2.3.1.9) - Catalyzes the condensation of acetoacetyl coenzyme A from two acetyl coenzyme A molecules. As a thiolase, it can also catalyse the reverse reaction by splitting acetoacetyl coenzyme A into two acetyl CoA molecules. It is a central metabolite and involved in mevalonate pathway, fatty acid pathway and biosynthesis of secondary metabolites (NCBI, 2018, Mann and Lütke-Eversloh, 2013).

abrB - A common deletion at reference positions 363789, 363790 and 360794 is observed in four mutants obtained from two independent transformations. This deletion is downstream of the *abrB* gene (regulators of stationary/sporulation gene expression) as well as the CA_C0309 gene (cell wall-associated hydrolase) and appears to be a terminator. The absence of a terminator at this position can result in interference between the genes as they are being transcribed. The expression of *abrB* was shown to be opposite *spo0A* (Tomas *et al*., 2004) and *spo0A* has been reported to negatively regulate *abrB* in *B. subtilis* (Fawcett *et al.,* 2000)

Ribonuclease J - CAC_1683 - An RNase that has 5'exonuclease and possibly endonuclease activity. It is involved in the maturation/stability of rRNA or in some organism's mRNA maturation or decay. It has a well conserved central region related to the beta-lactamase family at the amino terminal (Madhugiri and Evguenieva-Hackenberg, 2009)

A compensatory mutation occurs when the loss of fitness caused by one mutation is remedied by a second mutation at another site in the genome. The effect of the first mutation is usually dependent on the presence or absence of these

compensatory mutations, a phenomenon, known as epistasis (Rojas Echenique *et al.*, 2019). For instance, mutations that endow resistance abilities to an antimicrobial agent are often deleterious to the organism in the drug-free environment, however with the acquisition of a compensatory mutation survival rates are enhanced (while retaining the initial mutation) rather than a back mutation with loss of the resistance-conferring mutation (Maisnier-Patin and Andersson, 2004).

To be able to conclude that a mutation is compensatory for the knock-out of *hbd*, we expect it to be common to all the *hbd* mutants. Seven *hbd* KO mutants from four independent transformations were subject to paired-end Illumina sequencing and the sequences obtained mapped to the annotated genome reference sequence of ATCC 824.

In all seven mutants sequenced, a mutation in genes either directly involved in or related to the fatty acid biosynthesis pathway was observed; figure 3.15 shows the fatty acid operon in *C. acetobutylicum*. These genes are, *fabD, fabH, marR* (CAC_3579), acetyl CoA acetyltransferase (thiolase) and CAC_2681 (may interact with acyl carrier protein in FAB). In addition to these, mutations in or around genes related to endospore formation (CAC_1065, CAC_0437, deletion downstream *abrB* gene) and the *agr* quorum sensing system (*agrC* and CAC_0082) were observed. The second frequently occurring mutation following the FAB pathway genes, was the deletion downstream the *abrB* gene which occur in 4 mutants from two independent transformations. Mutations in CAC_1053 (LPS Glycosyltransferase) and CAC_1683 (ribonuclease J) were observed in one mutant each.

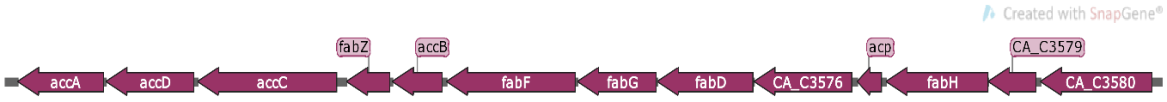


Figure 3. 14: Fatty acid operon in *C. acetobutylicum*.

Mutations were observed in four of the genes - *fabD, fabH,* CAC_3579 *(marR),* and CAC_2681 (may interact with acyl carrier protein, ACP) in the sequenced *hbd* mutant.

Cell membrane fluidity and loss of membrane potential is an established result of exposure to solvents such as ethanol and butanol (Huffer *et al*., 2011, Dyrda *et al*., 2019) and the alteration in cell membrane fatty acid composition is a mechanism cells employ to tolerate high solvents levels (Zhao *et al*., 2003).

A reduced expression of the fatty acid biosynthesis operon genes in a solvent-tolerant strain of *C. acetobutylicum* 824 (pGROE1) and the wild type when exposed to a 0.75 %, (vol/vol) butanol challenge was previously reported (Tomas *et al.*, 2004). The mutations in the FAB genes observed in all mutants may not be unrelated to increased ethanol production in the *hbd* mutant (figure 3.14) compared to the parent wild type strain.

A SNP in *thl* gene was observed in hbd69s::rCT and the metabolite analysis are consistent with a reduction in thiolase activity. Thiolase precedes Hbd in the metabolic pathway, therefore a disruption of the *hbd* gene may have resulted in reduced thiolase activity based on the general principle of substrate:product ratio control of metabolic flux. With a reduced thiolase activity, there is a disruption of the C4 pathway and either acetate or ethanol can be produced. A reduction in acetone is also expected. In this case, ethanol becomes the preferred product as there is the need to produce more reduced products for regeneration of $NAD(P)^+$.

This explains why the *hbd* mutant, hbd69s::rCT produced very high amounts of ethanol up to 427 mM. In a hbd knock down mutant previously described, biochemical analyses exhibited reduced thiolase activity in cell extracts compared to wild type *C. acetobutylicum*. A reduction in crotonase (which catalyses the conversion of crotonyl-CoA to butyryl-CoA) which follows Hbd in the metabolic pathway was also observed (Lehmann and Lutke-Eversloh, 2011).

It is quite interesting though to observe a mutation in the *thl* gene occurring with the knock-out of *hbd* gene. Previously it has been shown that the *thl* gene cannot be knocked out (Cooksley *et al.*, 2012) except together with the knock-out of *ptb*, hence the question was raised; was *ptb* knock-out only possible because of a frameshift in *thl*, or *vice versa*? Consequently, we sought to understand if there are compensatory mutations that would allow the viability of *C. acetobutylicum* bearing otherwise detrimental mutations. While the knock-out of *thl* gene has now been achieved using this conditional RAM-less ClosTron system (Ying Zhang, Personal Communication) and compensatory mutations identified, we sought to do same with the *hbd* gene.

To be able to confirm that these observed mutation(s) is/are responsible for the viability of the *hbd* KO mutant, it is important to replicate these mutations (individually or together) on a *C. acetobutylicum* wild type background and attempt to knock out *hbd* using the standard ClosTron system, which before now was shown to be impossible. If this becomes achievable, it will confirm the hypothesis that knock-out of the *hbd* gene is only feasible because of these compensatory mutations that have arisen, thereby making the mutant viable. Though this was the part of the plan at the beginning of this thesis, due to time

constraints, these mutations could not be replicated on the wild type background to confirm that these acquired mutations are responsible for the viability of the *hbd* muatant which could not be knocked out earlier using standard ClosTrom mutagenesis or the CRISPR/Cas9 system.

In addition, due to time restrictions, the solvent profile of these seven mutants remain to be analysed and matched with the sequencing results. For example, it will be expected that that mutant no 1 which had mutation in CAC_0437 (histidine kinase in endospore formation control) would produce more spores than wild type as a *C. acetobutylicum* CAC_0437 mutant has been previously reported to be a hypersporulator (Steiner *et al.*, 2011)

3.4. Conclusion

The creation of mutants for improved understanding and manipulation of solvent pathways is important in the development of an industrial solvent producing strain. In this chapter, we sought to utilize the conditional RAM-less ClosTron system to knock out genes in *C. acetobutylicum* that were previously impossible to knock-out and where possible isolate compensatory mutations that have made the mutant to be non-lethal. The *hydA, hbd, crt* and *bcd* genes were targeted using appropriate RAM-less ClosTron plasmids. For the *hydA, crt* and *bcd* genes, it was not possible to isolate a knock-out mutant. Although evidence of intron insertion was observed when candidates were PCR screened using EBS and a flanking primer, in each case, no mutant DNA fragment was observed

when PCR screening was done with gene flanking primers perhaps because the mutants are present at such a low level that they cannot be detected by PCR.

The *hbd* gene which could not be knocked out using the standard ClosTron or CRISPR/Cas9 system was knocked out using the conditional RAM-less ClosTron and pure insertional mutants isolated, designated as hbd69s::rCT. Solvent profile analysis showed that hbd69s::rCT produced more ethanol (427 mM) compared with wild type (33mM) and there was no butyrate produced by this mutant. Less acetone, butanol and more acetate were produced by hbd69s::rCT. When seven *hbd* mutants from four independent transformations were subjected to Illumina sequencing, results showed a mutation in genes involved in or related to the fatty acid biosynthesis pathway in all seven. Mutations in genes related to endospore formation, *agr* quorum sensing system, LPS Glycosyltransferase and ribonuclease J were also observed. However, to be able to conclude that these mutations have resulted in the viability of the *hbd* knock-out mutant, they have to be replicated on a wild type background and thereafter, the knock-out of *hbd* using the standard ClosTron system should be possible. However, due to time constraints, these mutations could not be replicated on the wild type background as planned. Also, as a further work, the solvent profile of these seven mutants need to be analysed and matched with the sequencing results for better understanding.

CHAPTER 4

Creation of Alternative Integration Sites to *pyrE*

CHAPTER 4

Creation of Alternative Integration Sites to *pyrE*

4.1 Introduction

The development of genetic tools for targeted mutagenesis of *Clostridium acetobutylicum* is essential to metabolic engineering strategies that seek to direct metabolism in favour of desired products. Allele Coupled Exchange, ACE, developed here at the CRG (Heap *et al*., 2012a; Ng *et al*., 2013; Ehsaan *et al*., 2016; Minton *et al*., 2016), readily allows the creation of mutants and their rapid complementation without the need for antibiotic markers and avoids the phenotypic effects associated with high copy plasmids. It has been demonstrated to effectively generate mutants in many clostridial species, including *C. acetobutylicum, C. difficile, C. botulinum, C. beijerinckii, C. perfringes, C. autoethanogenum C. pasteurianum, C. sporogenes* and even *Gerbil's thermoglucosidasius* (Minton *et al*., 2016)

The ACE system utilizes the *pyrE* gene (which codes for orotate phosphoribosyl transferase required for the *de novo* synthesis of pyrimidine) for the rapid deletion and insertion of desired DNA.

4.1.1 The *pyrE* Strain

Compared to gene expression via autonomous plasmids, the integration of operons into the genome results in a more stable strain suitable for industrial applications (Friehs, 2004). Though plasmid expression has the advantage of

high copy number, it is segregationally unstable and mostly has to be maintained using antibiotics which are expensive and cause environmental concerns (Hagg *et al*., 2004). The *pyrE* mutant strain created using ACE has two main functions:

i. It allows knock-out (KO) of other genes in the chromosome using standard allelic exchange procedures in which the KO plasmid carries a heterologous, functional copy of *pyrE* that is used as a negative selection marker (Ng *et al.*, 2012; Ehsaan *et al.*, 2016).

ii. Concomitant with restoration of the *pyrE* allele to wildtype, cargo DNA can be introduced into the genome by inserting it between the Long Homology Arm (LHA) and Short Homology Arm (SHA)

4.1.2 Use of *pyrE* Strains for Gene Knock-out

An example of an unpublished KO vector that uses a functional *pyrE* gene as a counter selection marker is shown in Figure 4.1. The vector also makes use of a second counter selection marker, *codA*. This encodes cytosine deaminase which converts the innocuous 5-Fluorocytosine (5FC) into the toxic 5-Fluorouracil (5FU).

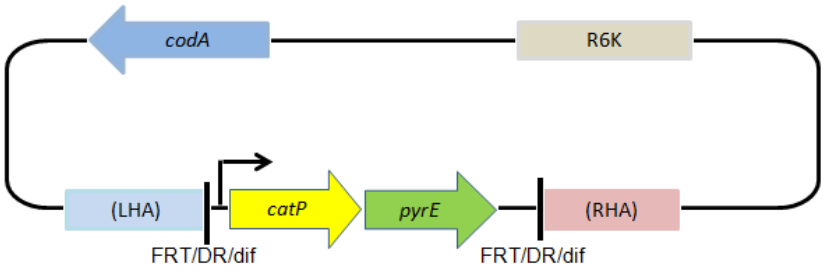


Figure 4. 1: Suicide KO plasmid

Suicide KO plasmid comprised *cod*A negative selection (5- flourocytosine, 5FC), R6K, Gram-ve replicon only replicates in *E. coli* strains producing π protein, LHA (Left

Homology Arm), *catP* (thiamphenicol resistance gene from *C. perfringens*), *pyrE* (prototrophic selection, uracil +ve), RHA (Right Homology Arm)

Following the integration of the suicide plasmid into the host chromosome through recombination between one of the two regions of homology, single crossover integrants can be selected as faster growing colonies on solid media containing thiamphenicol. These are then plated onto solid media containing 5FC, which is toxic to cells that carry the plasmid, and its *codA* gene, integrated into the genome. This allows the selection of colonies in which the plasmid has excised and lost the plasmid (and *codA*). Dependant on which homology arm mediates the excision event, progeny can be either the desired KO mutant, or the original wildtype cell. These can be distinguished on the basis of their sensitivity to thiamphencol. Mutant cells are resistant, as they carry the *catP*::*pyrE* cassette. Wildtype cells are sensitive. Isolated, 5FC resistant, thiamphenicol resistant cells are then plated on media containing 5-FOA. The integrants are sensitive to this analogue because of the presence of the functional *pyrE* gene. Cells in which this *catP*::*pyrE* cassette flips out, due to recombination between the direct repeats that flank the cassette (FRT/dif/DR) become resistant to 5-FOA, as now the only *pyrE* allele in the cell is the original mutant one. Figure 4.2 below is a schematic of the knock out using *pyrE* gene/suicide plasmid.

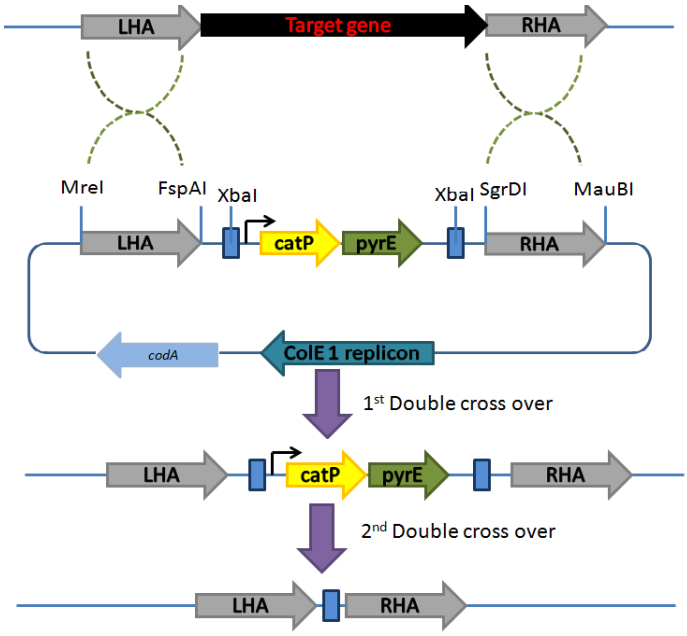


Figure 4. 2: In-frame deletion of target gene using *pyrE* gene/suicide plasmid. Following the first recombination event, entire plasmid is integrated in the chromosome and the target gene is replaced by *catP-pyrE* after the second recombination occurs (which is the 1st double cross over). The 2nd double cross over, giving rise to *catP-pyrE* flip, can be screened for by plating of FOA plates which serve as a counter selection marker for *pyrE* gene.

4.1.3 Use of *pyrE* Strains for Knock-in

Once generated, an ACE correction vector may be used to restore the *pyrE* allele to wild type, in which the missing 3'-end of the *pyrE* gene is combined with the LHA. Here, following the isolation of a single crossover integrant, the required double crossover strain is selected by plating on media lacking uracil as the strain generated becomes prototrophic for uracil. The *pyrE* corrected cells also become FOA sensitive (^{S}). The phenotype of the gene knocked-out may then be determined in a cell that is otherwise wildtype. If the *pyrE* gene was not corrected, the effects could be partially due to the *pyrE* mutation.

The ACE system can also be used to complement the knocked-out gene. In this case, the ACE complementation vector has restriction sites positioned between

the LHA (3'-end of the *pyrE* gene) and the RHA (the 1200bp region from downstream of the 3'-end of *pyrE*). This allows a functional copy of the gene that has been knocked-out to be cloned between the LHA and RHA, either with its own promoter or with no promoter. In the latter case the gene is expressed from the promoter responsible for *pyrE* expression. The cloned gene becomes inserted into the gene following ACE at the same time as the *pyrE* gene is restored to wild type. The gene can also be expressed from the strong P_{fdx} promoter (derived from the ferredoxin gene of *C. sporogenes* using an appropriate ACE vector that contain this promoter upstream of the restriction sites into which the gene is inserted. Figure 4.3 shows ACE correction, complementation and overexpression vectors

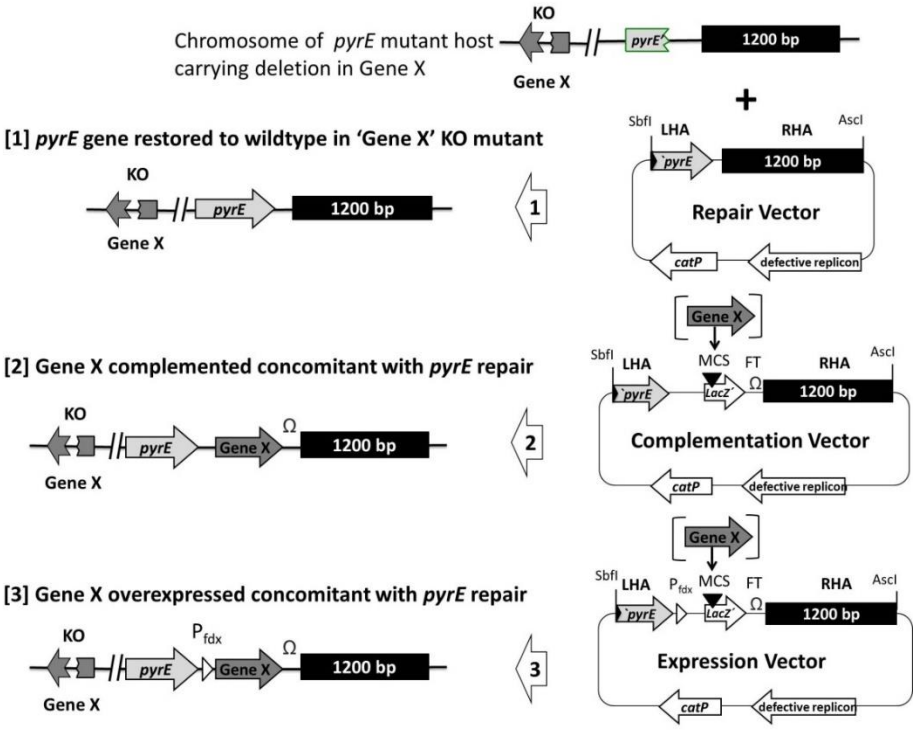


Figure 4. 3: ACE Correction, Complementation and Expression Vectors.

Figure taken from Minton *et al.*, 2016.

4.1.4 Alternative Loci to *pyrE*

One feature that would facilitate the engineering of the *C. acetobutylicum* chassis would be the provision of alternative loci to *pyrE* at which heterologous or homologous genes could be rapidly inserted. Thus, if the desired pathway required the addition of more than one operon or gene set, then one could be inserted at *pyrE* and the other could be inserted elsewhere at an equivalent locus through the construction of a series of ACE vectors.

Ideally, like *pyrE*, such loci should: (i) represent the last gene of an operon, thereby preventing polar effects on downstream genes following insertion of cargo DNA; (ii) should result in an easily screenable phenotype (ie., uracil auxotrophy), the correction of which is selectable (ie., uracil prototrophy), and; (iii) if possible should also confer on the host a selectable phenotype when inactivated (FOA^R). Two promising candidates identified by the CRG are *pheA* (CAC0217, prephenate dehydratase) and *argH* (CAC0974, argininosuccinate lyase). The former is predicted to be moncistronic and its inactivation should lead to phenylalanine auxotrophy and may become resistant to the phenylalanine analogue, β-2-thienylalanine (Nelms *et al*., 1992). Inactivation of the latter should lead to arginine auxotrophy. In-frame deletion mutants of both could be made using our heterologous *pyrE*-based counter selection, allelic exchange system. Mutants will be made in the *pyrE* host, and the *pyrE* allele corrected to prototrophy using an ACE Correction vector. Following isolation of each mutant, ACE Correction vectors specific to each allele was made, and shown to be capable of correcting each specific mutant back to the wildtype phenotype (eg., arginine prototrophy in the case of *argH*) using the ACE method.

4.2 Aim of Chapter

To create a triple auxotrophic mutant in *C. acetobutylicum* which represent alternative loci to *pyrE* that genes of interest could be readily integrated.

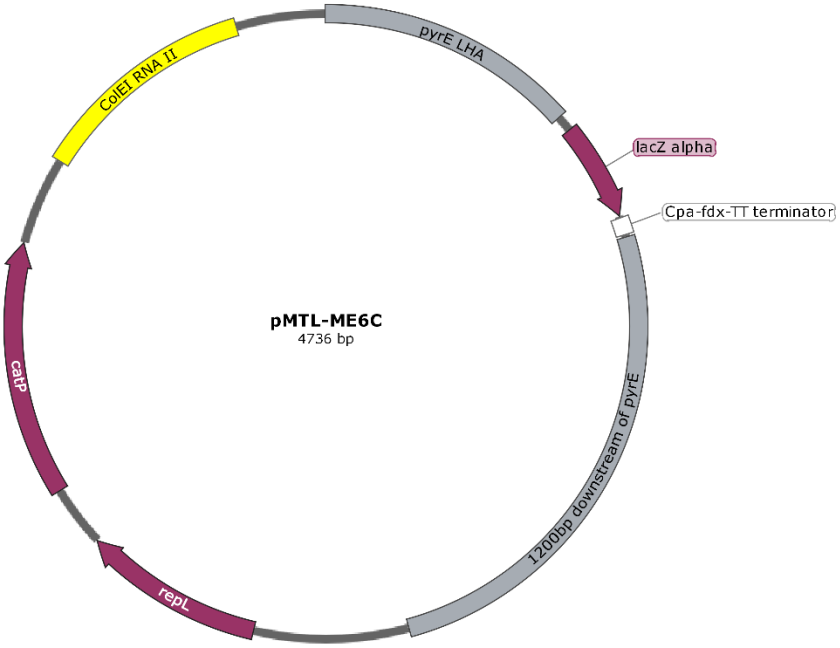
4.3 Results and Discussion

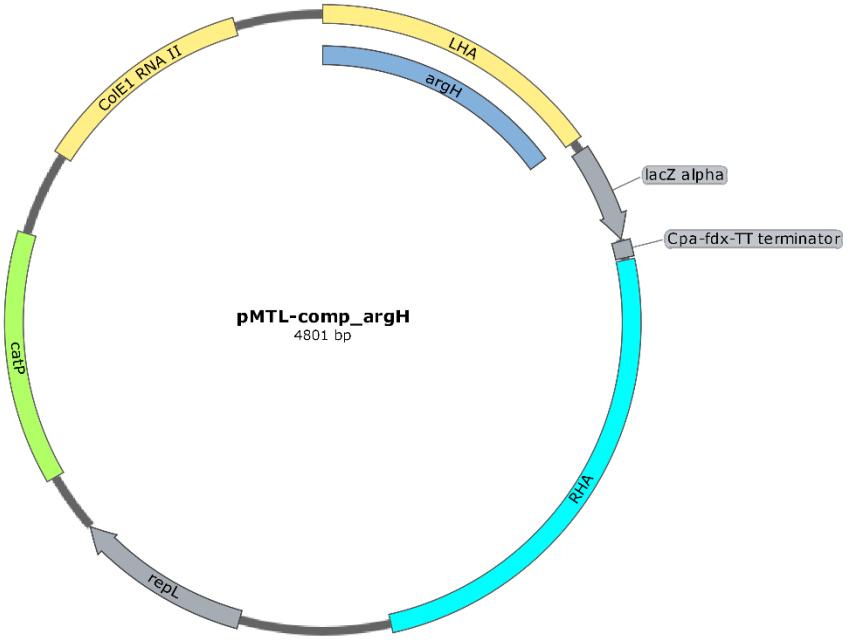
4.3.1 Construction of an ACE integration system based on *argH*

Previous work by Hengzheng Wang and Jon Baker had resulted in the creation of a specific in-frame deletion mutant of the *argH* locus and *pheA* locus in the *pyrE* mutant background of *C. acetobutylicum.* The nature of the deletion was such that they resembled the *pyrE* mutant in that 406 and 301 bp were removed from the 3'-end of the *argH* and *pheA* genes, respectively. It follows that it should be possible to restore both the *argH* and *pheA* locus to wildtype using an ACE correction vector built to the same principles as the *pyrE* ACE correction vector but in which the *pyrE* specific RHA and LHA are replaced with the equivalent regions of the *argH* or *pheA* locus.

Accordingly, the RHA and the LHA of the *pyrE* ACE complementation vector pMTL-ME6C were replaced with *argH* specific DNA. These were derived by the PCR amplification of the appropriate region from the *C. acetobutylicum* chromosome using PCR primers that incorporated restriction enzyme recognition sites that allowed their cloning into pMTL-ME6C in place of the *pyrE* specific regions. In the case of the LHA, the 300 bp region incorporated the 3'-end of the *argH* gene and was bounded by restriction sites for *Sbf*I and *Not*I. The RHA represented a 1200 bp region immediately downstream of the

argH gene and was flanked by *Nhe*1 and *Asc*1 restriction recognition sites. These two regions were sequentially cloned into pMTL-ME6C, replacing the *pyrE* specific DNA between *Sbf*I and *Not*I (LHA) and between *Nhe*1 and *Asc*1 (RHA). The final *argH* and *pheA* complementation vector was designated, pMTL_comp_argH and pMTL_comp_pheA respectively (figure 4.4).





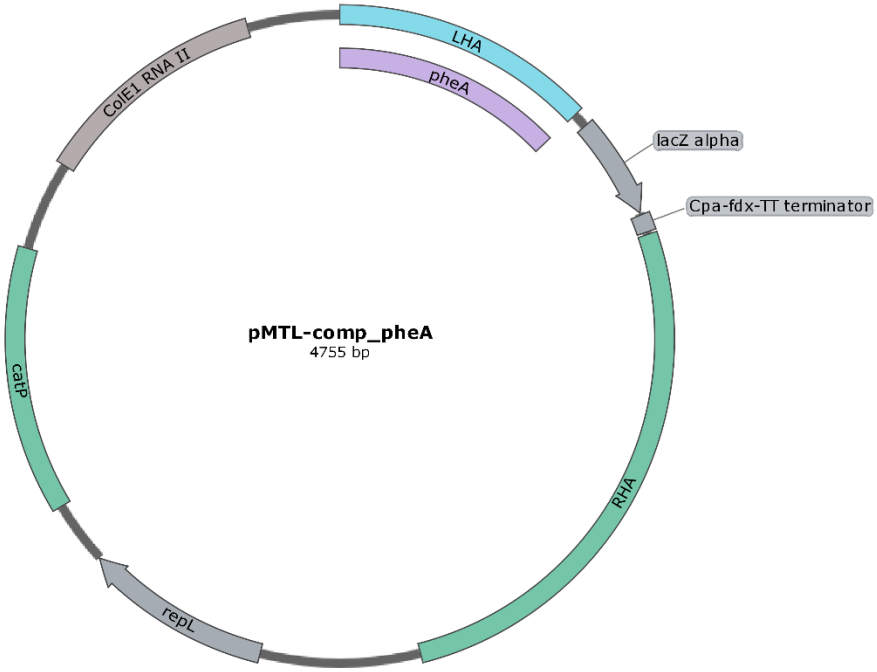


Figure 4. 4: ACE complementation vectors

For the three loci, *pyrE* (pMTLME6C*)*, *argH* (pMTL_comp_argH) and *pheA* (pMTL_comp_pheA). RHA is the long homology arm downstream of the deleted gene and initiates first recombination process; LHA is the short homology arm, carries the deleted 3' segment of the gene to be repaired and mediates second crossover event which repairs the deletion; ColE1 is Gram-negative replicon; *catP* is chloramphenicol antibiotic marker; *repL* - pIM13 Gram-positive defective replicon; *lacZ* alpha encodes a multicloning site where cargo genes of interest could be cloned in to make expression vectors.

Having obtained pMTL_comp_argH, it was transformed into the *C. acetobutylicum* Δ*pyrE*Δ*argH* double mutant strain initially selecting for thiamphenicol resistant colonies, and then streaking them onto minimal media supplemented with uracil but lacking arginine. Three independent putative repaired strains were selected (Figure 4.5) and PCR undertaken using appropriate flanking primers to confirm that the *argH* gene had been restored to

wildtype. A PCR product of the predicted size was obtained (Figure 4.6) and shown by subsequent Sanger sequencing to be exactly as expected.

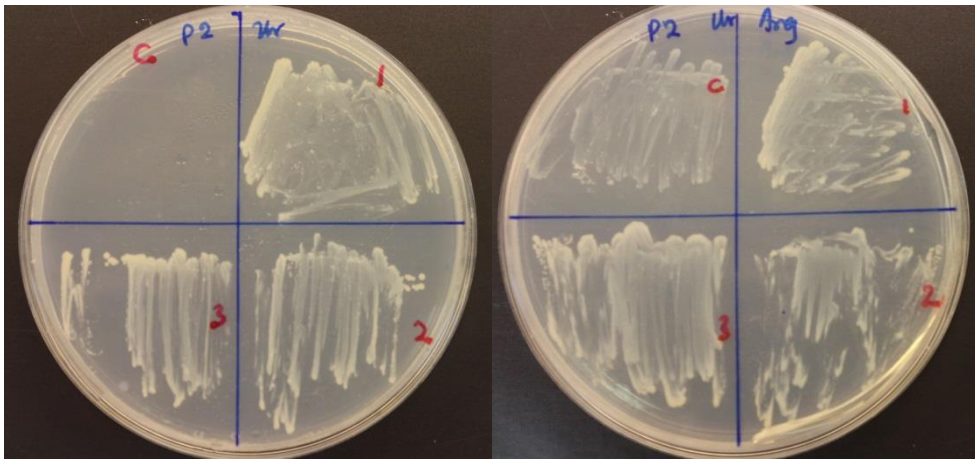


Figure 4. 5: Repair of *argH* gene

argH repair in *C. acetobutylicum* Δ*pyrE*Δ*argH* using pMTL_comp_argH. There is growth of *argH* repaired mutant (1, 2 and 3) on both arginine supplemented and unsupplemented P2 minimal plates while the parent strain, *C. acetobutylicum* Δ*pyrE*Δ*argH* (control, C) grows only on supplemented inimal plates.

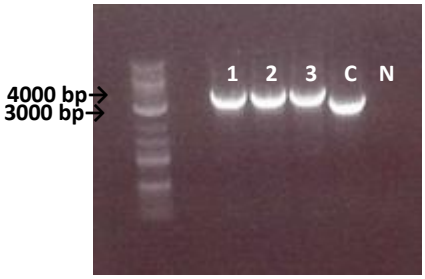


Figure 4. 6: PCR screening of *argH* repair

Gel picture of PCR screening of *argH* repair using flanking primers argh_ch_F&R. Lanes are genomic DNA of potential complemented mutants. Lanes 1, 2, 3 are complemented *argH* mutant with expected DNA fragment size of 3547bp; *C. acetobutylicum* Δ*pyrE*Δ*argH* double mutant control (C) with fragment size of 3186bp and no band for negative control (N). DNA Ladder is 2-log.

4.3.2 Characterisation of the genome sequence of the *C. acetobutylicum* *ΔpyrEΔargH* and *ΔpyrEΔpheA* double mutants

The eventual goal of this work is to also have a third locus for introducing genes based on the *pheA* gene (phenylalanine auxotrophy). However, before moving to the construction of a triple mutant strain (Δ*pyrE*, Δ*argH*, Δ*pheA*) it was important to show that both double mutants did not carry any significant SNPs (Single Nucleotide Polymorphisms) or InDels. Therefore, chromosomal DNA from both double mutants was prepared and subject to paired-end Illumina sequencing. The sequences obtained were then mapped to the annotated genome reference sequence of ATCC 824. Based on results shown in Table 4.1 below, strain argH 10_2 was selected to go forward with to create the triple mutant strain.

Table 4. 1: Genomic Sequencing Results of Double mutants, Δ*pyrE*Δ*argH* and Δ*pyrE*Δ*pheA*

CACM_1	**pheA 2_1**						
Mapping	**Reference Position**	**Type**	**Reference**	**Allele**	**Overlapping annotations**	**Coding region change**	**Amino acid change**
	38051	Insertion	-	C		**the end of pyrE, not a real SNV**	
	162936	SNV	A	G			
	345917	SNV	A	G			
	345993	SNV	C	T	rRNA: rRNA, Gene: CA_Cr027		
	2294876	SNV	A	G	CDS: CA_C2201, Gene: CA_C2201	NP_348818.1:c.34T>C	
CACM_2	pheA 21_1				**no coverage of pSOL**		
Mapping	**Reference Position**	**Type**	**Reference**	**Allele**	**Overlapping annotations**	**Coding region change**	**Amino acid change**
	38051	Insertion	-	C		**the end of pyrE, not a real SNV**	
	162936	SNV	A	G			
	345917	SNV	A	G			
	345993	SNV	C	T	rRNA: rRNA, Gene: CA_Cr027		
CACM_3	**pheA 15_1**				**no coverage of pSOL**		
Mapping	**Reference Position**	**Type**	**Reference**	**Allele**	**Overlapping annotations**	**Coding region change**	**Amino acid change**
	38051	Insertion	-	C		**the end of pyrE, not a real SNV**	
	162936	SNV	A	G			
	345917	SNV	A	G			
	345993	SNV	C	T	rRNA: rRNA, Gene: CA_Cr027		
	348937	SNV	C	A			
	459698	SNV	C	A	CDS: kdgK, Gene: kdgK	NP_347035.1:c.280C>A	NP_347035.1:p.Gln94Lys

CACM_4	**argH 10_2**						
Mapping	**Reference Position**	**Type**	**Reference**	**Allele**	**Overlapping annotations**	**Coding region change**	**Amino acid change**
	38051	Insertion	-	C		**the end of pyrE, not a real SNV**	
	162936	SNV	A	G	inter rRNA regions		
	173271	SNV	T	C	inter rRNA regions		
	345917	SNV	A	G	inter rRNA regions		
	345993	SNV	C	T	rRNA: rRNA, Gene: CA_Cr027		

CACM_5	**argH 30_1**				**no coverage of pSOL**		
Mapping	**Reference Position**	**Type**	**Reference**	**Allele**	**Overlapping annotations**	**Coding region change**	**Amino acid change**
	38051	Insertion	-	C		**the end of pyrE, not a real SNV**	
	162936	SNV	A	G			
	345917	SNV	A	G			
	345993	SNV	C	T	rRNA: rRNA, Gene: CA_Cr027		
	2824877	SNV	A	G	between *guaB* and CA_C2702		

CACM_6	**argH 30_7**						
Mapping	**Reference Position**	**Type**	**Reference**	**Allele**	**Overlapping annotations**	**Coding region change**	**Amino acid change**
	38051	Insertion	-	C		**the end of pyrE, not a real SNV**	
	162936	SNV	A	G			
	345917	SNV	A	G			
	345993	SNV	C	T	rRNA: rRNA, Gene: CA_Cr027		
	2539244	SNV	G	A	CDS: CA_C2422, Gene: CA_C2422	NP_349037.1:c.500C>T	NP_349037.1:p.Thr167Ile

4.3.3 Creation of Triple Mutant

From the results of genomic sequencing of second double mutants, argH_10-2 (Δ*pyrE*Δ*argH)* was selected to make the triple mutant. This had no significant SNPs and pSOL1 plasmid was still intact.

The suicide KO plasmid to knock out the *pheA* gene, pMTL-8-147-pheA was designed just as shown in figure 4.1 but the homology arms flanked the 301 bp region to be deleted in the *pheA* gene. The plasmid was methylated and transformed into Δ*pyrE*Δ*argH* (argH_10-2). Single cross integrants were plated on CBM- FC ur+arg+phe+Th plates (to screen for excised plasmid and *codA* loss) and patch plated on P2 ur+arg+Th +/- phe. Double cross was obtained and confirmed using primers, pheA_ch_F and pheA_ch_R. This is shown in figure 4.7.

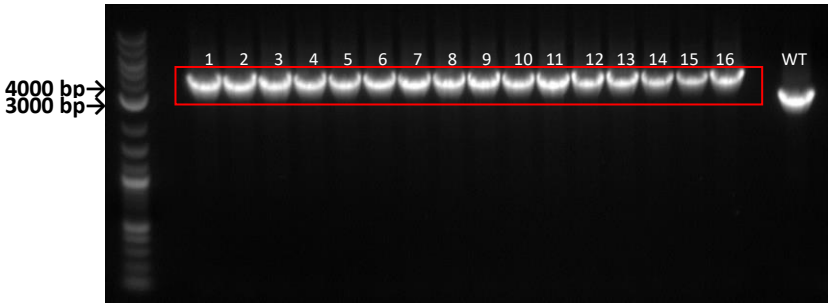


Figure 4. 7: PCR screen of *pheA* KO

PCR screen of *pheA* knock out using gene flanking primers pheA_ch_F and pheA_ch_R. Lanes are genomic DNA of potential *pheA* KO candidates. Double cross should give a DNA fragment size of 4229 bp (1-16), while Wild type should be 3090 bp.

To screen for *catP-pyrE* flip, colonies were plated on 5-FOA Ur+arg+phe plates. Colonies were PCR screened with same flanking primers and results shown in figure 4.8.

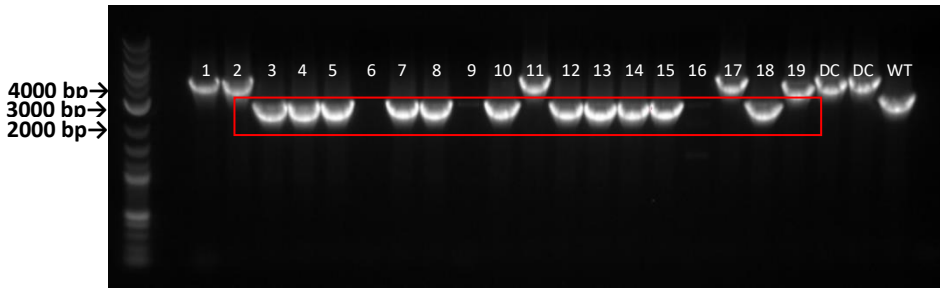


Figure 4. 8: PCR screen for *catP-pyrE* flip in *pheA* KO mutants

PCR screen for *catP-pyrE* flip in *pheA* KO using flanking primers pheA_ch_F and pheA_ch_R. Genpmic DNA of 19 colonies were screened and 11 (highlighted in red) show expected DNA fragment size at 2817 bp. Double cross (DC) gives fragment size of 4229 bp and WT, 3090 bp.

Putative *catP-pyrE* clean colonies were streaked onto minimal medium, P2 ur+arg ± phe. The created triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*, can grow only on minimal plates supplemented with phenylalanine while the Control, (Δ*pyrE*Δ*argH*, arg 10-1) can grow both on minimal plates with or without phenylalanine. This is shown in figure 4.9.

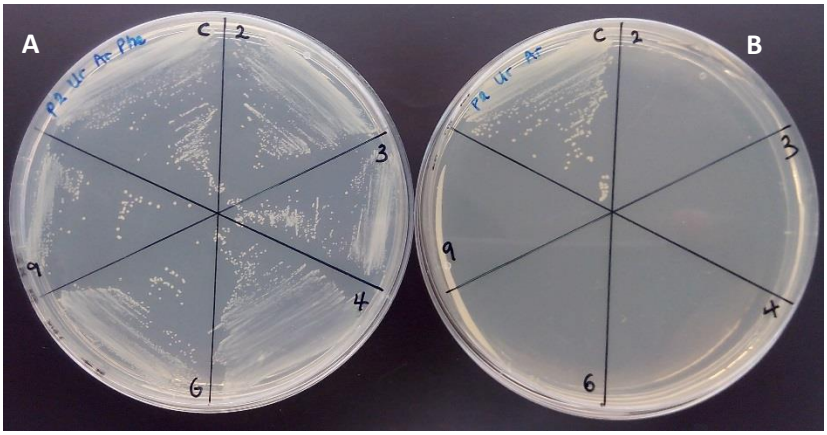


Figure 4. 9: Triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*

Triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*, colonies 2, 3, 4, 6, 9, can grow only on minimal plates supplemented with phenylalanine (plate A) while Control (double mutant *C. acetobutylicum* Δ*pyrE*Δ*argH)* can grow both on minimal plates with/without phenylalanine (plates A & B)

4.3.4 *pheA* Repair

To show that the *pheA* gene in triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*, could be repaired, the cells were transformed with the *pheA* cargo complementation vector, pMTL_comp_pheA. Integrants were PCR screened using primers, pheA_ch_F and pheA_ch_R. These are shown in the figures 4.10 & 4.11.

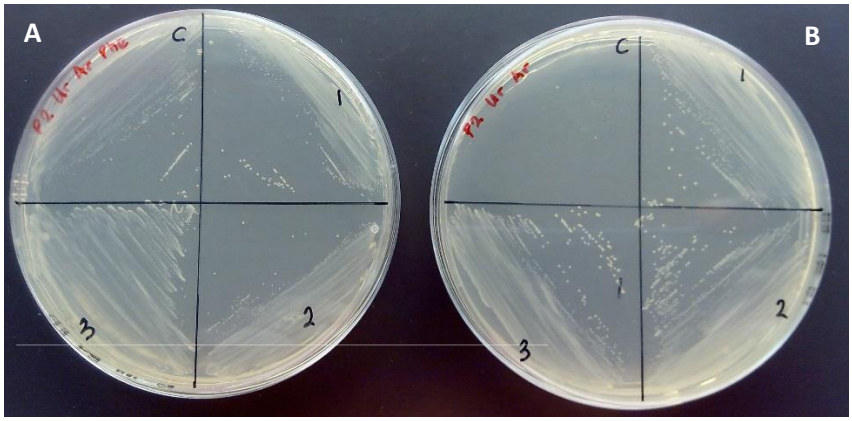


Figure 4. 10: *PheA* gene repair in triple mutant

PheA gene repaired in triple mutant, colonies 1, 2, 3 can grow on minimal plates with/without phenylalanine supplementation (plates A & B) while triple mutant *C. acetobutylicum ΔpyrEΔargHΔpheA* (C) can only grow on minimal plates supplemented with phenylalanine (plate A).

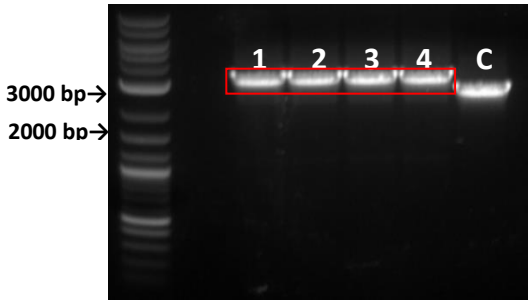


Figure 4. 11: PCR screening of repaired *pheA* gene in *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA.*

PCR screening of *pheA* gene repair in triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA* using flanking primers pheA_ch_F and pheA_ch_R. Lanes are genomic DNA of potential complemented mutants. Lanes 1, 2, 3 and 4 are complemented *pheA* mutants with expected DNA fragment size of 3090 bp while triple mutant with unrepaired strain (C) shows gives a size of 2817 bp.

4.3.5 Confirmation of Presence of pSOL1 plasmid in Δ*pyrE*Δ*argH*Δ*pheA* mutants

Just to be sure, the presence of pSOL1 plasmid in triple mutant was confirmed (figure 4.12) by amplifying the *adhE2* gene using primers adhe2F/R and using wildtype *C. acetobutylicum* as control. *adhE2* is one of the genes borne on the pSOL1 plasmid.

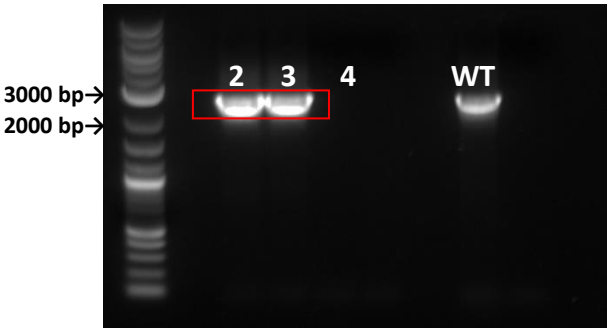


Figure 4. 12: Screening for presence of pSOL1

Screening of genomic DNA of triple mutant, Δ*pyrE*Δ*argH*Δ*pheA*, for the presence of pSOL1 plasmid using flanking primers for the *adhE2* gene. DNA fragment size for *adhE2* gene (2666 bp) is seen for 2 and 3 indicating presence of pSOL1 plasmid just as in wild type used as control.

4.3.6 Genomic Sequencing Results of Triple mutant *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*

As previously performed, chromosomal DNA from *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA* was prepared and subjected to paired-end Illumina sequencing. The sequences obtained were then mapped to the annotated genome reference sequence of ATCC 824 using CLC Genomics Workbench to check for significant SNPs or InDels. The results are presented in table 4.2. In addition to the SNPs in the parent double mutant, argH 10_2 (highlighted in yellow), a SNP in CDS: CA_C2843 (aminopeptidase) was observed, this was an amino acid

change from Asp to Tyr. The aminopeptidase catalyzes the cleavage of amino acids from the amino (N) - terminus of polypeptides, they are widely distributed in the animal and plant kingdoms and exist in subcellular organelles, cell membrane and in cytoplasm. (Gonzales and Robert-Baudouy; 1996, Taylor, 1993). This mutation was however not in the protein active site.

Table 4. 2: Genomic Sequencing Results of *C. acetobutylicum* triple mutants, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*

P2

Position	Type	Ref	Allele	Freq	Overlapping annotations	Coding region change	Amino acid change
38051	Insertion	-	C	100	None	None	None
162936	SNV	A	G	48.18	None	None	None
173271	SNV	T	C	100	None	None	None
345917	SNV	A	G	100	None	None	None
345993	SNV	C	T	100	Gene: CA_Cr027, rRNA: rRNA	None	None
2975353	SNV	G	T	100	CDS: CA_C2843, Gene: CA_C2843	NP_349447.1:c.1096G>T	NP_349447.1:p.Asp366Tyr

P3

Position	Type	Ref	Allele	Freq	Overlapping annotations	Coding region change	Amino acid change
38051	Insertion	-	C	100	None	None	None
162936	SNV	A	G	54.67	None	None	None
173271	SNV	T	C	100	None	None	None
345993	SNV	C	T	100	Gene: CA_Cr027, rRNA: rRNA	None	None
2975353	SNV	G	T	97.5	CDS: CA_C2843, Gene: CA_C2843	NP_349447.1:c.1096G>T	NP_349447.1:p.Asp366Tyr

4.3.7 Locus Specific Effects on Gene Expression

At this stage it is not clear what specific effects the loci will have on the expression of genes inserted at the three envisaged sites (*pyrE, argH* or *pheA*). To analyse this, an expression cassette was made, comprising a transcriptional terminator, followed by the promoter (P_{fdx}) of the *C. sporogenes* ferredoxin gene and finally a promoter-less copy of the *C. perfringens catP* gene. This cassette was cloned into the three different ACE complementation vectors, pMTL-ME6C, pMTL_comp_argH and pMTL_comp_pheA between the homology arms, using *Not*I and *Pvu*I to give pMTL-HZ13-fdx-catP, pMTL-HZ14-fdx-catP and pMTL-HZ15-fdx-catP respectively. The resultant vectors were transformed into the mutants with the concomitant repair of the genes. The *pyrE* mutation was also repaired in the double mutants.

The data obtained (figure 4.13) suggests that there is a difference in the level of *catP* expression obtained dependent on the site of insertion in the chromosome ($p < 0.05$ at 4 hours and $p < 0.0001$ at 8 hours). The highest level of expression was obtained at the *pyrE* locus, then at the *pheA* and finally from the *argH*. The simplest explanation for this observation relates to the position of these loci relative to the chromosomal origin of replication. Those nearest to the origin will have more copies than those that are more distal due to their closer proximity to the multiple replication forks that will be present in an actively growing cell (Couturier and Rocha, 2006). Similar effect of gene dosage has been reported in *E. coli* (Sousa *et al*., 1997) and *Bacillus subtilis* (Sauer *et al*., 2016). Consistent with the relative expression levels observed, the position of the three genes relative to the origin are 37,376 (*pyrE*), 241,803 (*pheA*) and 1,121,793 (*argH*)**.**

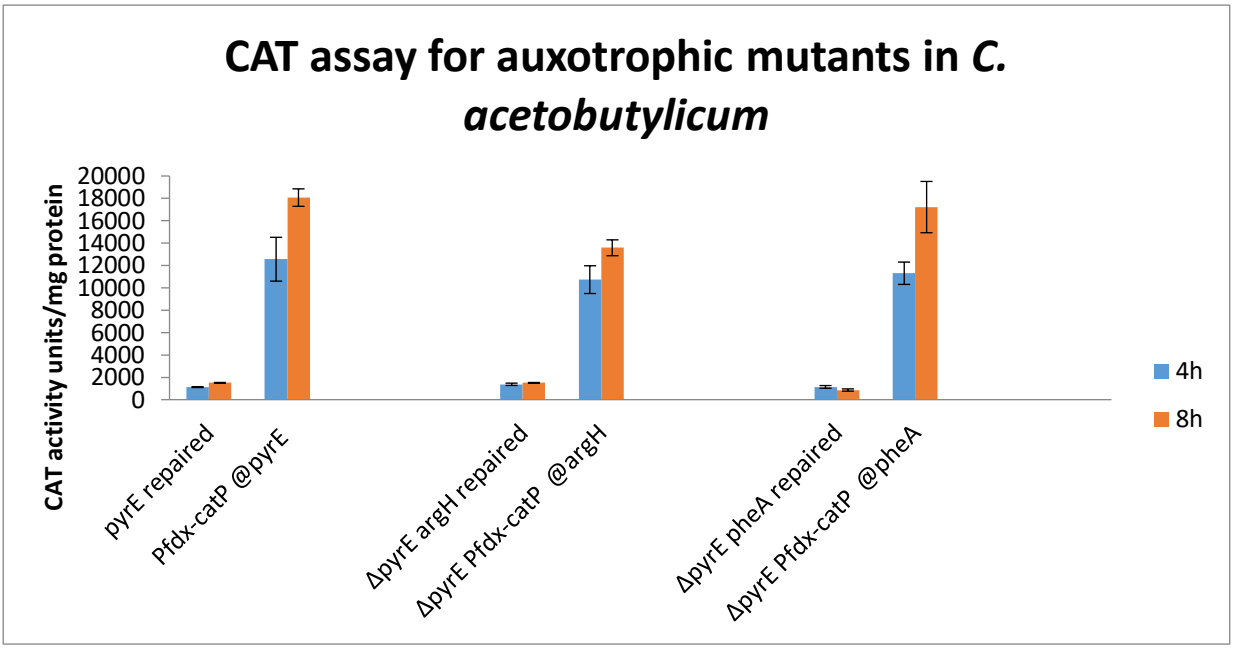


Figure 4. 13: Locus specific (*pyrE, argH and pheA)* effects on CAT expression in *C. acetobutylicum*.

Mutants were grown in CBMS, triplicate samples taken at 4 and 8 hrs and CAT assay carried out. Highest level of expression is observed at the *pyrE* locus, next at *pheA* and then *argH* and this could be based on the position of these loci relative to the chromosomal origin of replication.

4.4 Conclusion

The successful creation of a triple auxotrophic mutant in *C. acetobutylicum* was achieved. This was carried out by the sequential deletion of 406 and 301 bp from the 3' end of the *argH* and *pheA* genes in a *pyrE* background. This was achieved by using a suicide vector which utilized a heterologous *pyrE* gene as a counter selection marker. These deleted genes could be repaired by designing plasmids made similar to the *pyrE* repair plasmid pMTLME6C, where the homology arms had been replaced to direct integration at the *argH* and *pheA* loci.

Mutants were subjected to paired-end Illumina sequencing and the final triple mutant had one mutation distinct from the parent strain. This was an amino acid change from Asp to Tyr in the aminopeptidase gene (CA_C2843). It is not yet clear if this mutation will have any effect on the integration and expression of genes at these loci but this mutation was not in the protein active site. In addition, *C. acetobutylicum* carries other similar genes (Aminopeptidase, CA_C2195 and predicted aminopeptodase, CA_C3508) that could possibly complement for this mutation in CA_C2843. Locus specific effect of gene expression showed that the highest level of *catP* expression was at the *pyrE* locus followed by *pheA* and then *argH*. This could be related to the position of these loci relative to the chromosomal origin of replication with the nearest having most expression. With the availability of alternative loci to *pyrE* where genes of interest could be readily inserted in the chromosome, product extension in *C. acetobutylicum* was investigated by integrating heterologous genes at these loci.

CHAPTER 5

Towards Product Extension in *Clostridium acetobutylicum*

CHAPTER 5

Towards Product Extension in *Clostridium acetobutylicum*

5.1 Introduction

Over and above the use of the *pyrE* locus for complementation studies, is its major use as a locus for inserting native or heterologous DNA encoding beneficial functions. These can include those genes encoding the enzymes involved in the production of a novel or existing chemical (Heap *et al*., 2012); hydrolases for degrading complex carbohydrates (Kovacs *et al*., 2013); therapeutic genes in cancer delivery vehicles (Heap *et al*., 2014); an *ermB* gene to improve the reproducibility of the virulence of the NAPI/B1/027 epidemic strain R20291 in the hamster model of infection (Kelly *et al*., 2016), and; an antibiotic resistance gene to or a sigma factor to allow deployment of a *mariner* transposon (Zhang *et al*., 2015).

Figure 5.1 shows the use of the *pheA* locus for integration of cargo gene into chromosome.

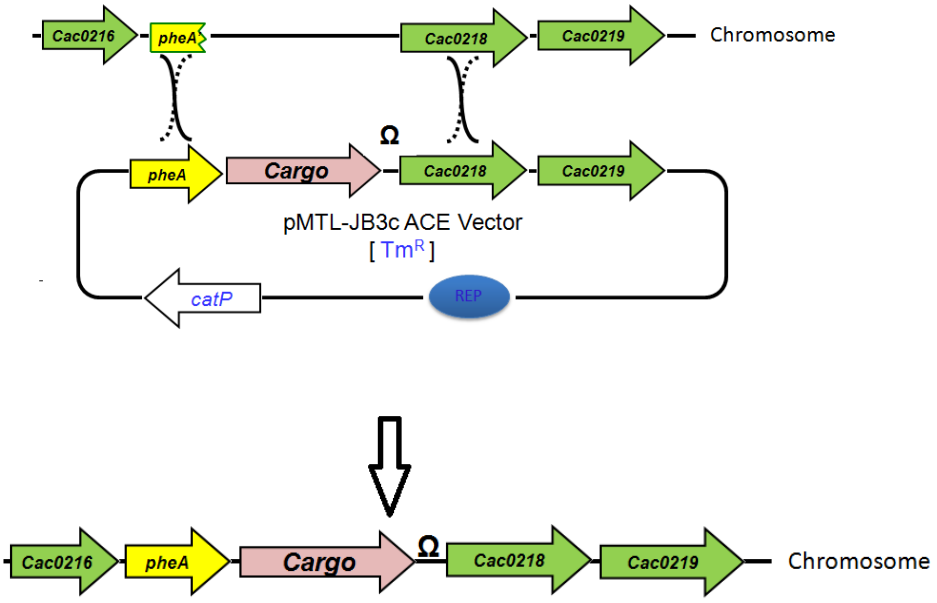


Figure 5. 1: The use of *pheA* locus for integration of cargo genes via ACE.

First integration occurs through the Long Homology Arms (LHA) while the second crossover event occurs via the Short Homology Arms (SHA) with the concomitant repair of the *pheA* gene and integration of gene into chromosome.

5.1.1 Acetone – An Undesired By-product in ABE Fermentation

In the ABE fermentation, it is desirable to eliminate or reduce the production of acetone as it is corrosive and a non-fuel by-product (Wang *et al*., 2018; Lee *et al*., 2012). In addition, its simultaneous production with butanol reduces the butanol ratio and increases the cost of downstream purification making the process economically non-viable (Nicolaou *et al*., 2010).

Attempts to reduce acetone production have been reported by several researchers over the years and this has mostly been associated with reduced butanol titers and are discussed below. Two enzymes are involved in the conversion of acetoacetyl-CoA to acetone: acetoacetyl-CoA:acetate/butyrate CoA transferase (*ctfA/B*) and acetoacete decarboxylase (*adc*) (Jones and Woods, 1986).

Janati-Idrissi reported the generation of *C. acetobutylicum* 2-BrBu1 from chemical mutation, which was deficient in acetone production but also showed a reduced butanol ratio (Janati-Idrissi *et al*., 1987). Tummala *et al* designed an antisense RNA to downregulate *adc* expression and it was observed that though downregulation of *adc* was effective, it did not translate to reduction in acetone production indicating that *adc* was not the limiting step in acetone production. Next, the downregulation of *ctfA/B* was targeted, but though the production of acetone was substantially reduced, butanol titers were significantly less compared with control (Tummala *et al*., 2003b). In a further work by the same researchers, the alcohol-aldehyde dehydrogenase gene (*aad*) was overexpressed alongside the ctfB-asRNA and this resulted in 2 to 8-fold increase in butanol production compared with the strain containing just ctfB-asRNA. In addition however, acetone production was 4 to 6-fold higher than in strain bearing just ctfB-asRNA. (Tummala *et al*., 2003a). The researchers suggested that the downregulation of *ctfB* resulted in the degradation of the entire *aad-ctfA-ctfB* transcript hence with the overexpression of *aad*, more butanol was produced.

Although this work improved the butanol-acetone ratio, the overall butanol production was reduced. Taking this further, Sillers *et al* (2009) expressed the *aad* gene using the *ptb* promoter while the ctfB-asRNA downregulated acetone production. This resulted in early alcohol production of 30 g/l as well as an increased alcohol to acetone ratio (Sillers *et al*., 2009). It seemed the limited availability of butyryl-CoA (due to depleted butyrate) was a limiting factor to butanol production, so, the thiolase (*thl*) gene was overexpressed to achieve a higher butanol acetone ratio. However, against expectation, product formation was not altered implying the need for more complex metabolic engineering. In

another work, the strain, MGC Δ*cac1502*Δ*upp*ΔctfA/B which lacked the type II restriction endonuclease encoded by *CA_C1502, the upp* gene (CA_C2879) and the *ctfA/B* genes was reported to produce no acetone, increased acids (especially acetate) but butanol titers were also reduced compared with wild type (Croux *et al*., 2016).

Jiang and co-workers reported that the disruption of the *adc* gene using targetron technology resulted in an increase in butanol ratio form 70 – 80% and a reduction of acetone to approximately 0.2 g/l (Jiang *et al*., 2009).

5.1.2 An Alternative Route – Converting Acetone to Isopropanol

An alternative route to reduce/eliminate acetone production in ABE fermentation is the reduction of acetone to isopropanol in just one step by the secondary dehydrogenase, *Sadh* which could be obtained from the natural isopropanol producer, *C. beijerinckii* NRRL B593 (Lee *et al*., 2012; Chen and Hiu, 1986). This represents a more efficient use of the acetone pathway. Unlike acetone, isopropanol is a simple secondary alcohol which has been used as an additive for high octane gasoline preparation (Peralta-Yahya and Keasling, 2010). Its other uses include as a cleaning agent, in paint thinning, plasticizer in plastic industry (Jang *et al*., 2013). Converting acetone to isopropanol creates the so called IBE mixture as opposed to ABE. The IBE mixture is more attractive for fuel applications as it can be used directly as a fuel since isopropanol is a fuel compound (Collas *et al*., 2012; Li *et al*., 2019). The use of IBE mixture as a fuel additive in the production of gasoline or diesel oil has already been reported (Peralta-Yahya and Keasling, 2010; Li *et al*., 2019)

5.1.3 IBE from Natural Producers

Some clostridia naturally produce IBE such as *C. beijerinckii* and *C. aurantibutyricum* (George *et al.*, 1983). Until 2018 when (Zhang *et al.*, 2018) reported the isolation of a *C. beijerinckii* strain BGS1 producing 10.21 g/l butanol and 3.41 g/l isopropanol, the IBE production has been low (5.87 g/l) and productivity slow (90.12 g/l/h) (Survase *et al.*, 2011). Even then, due to the availability of more developed genetic tools for metabolic engineering, higher IBE values have been achieved using *C. acetobutylicum*.

5.1.4 IBE from *C. acetobutylicum*

A summary of the reports of IBE production from *C. acetobutylicum* from 2011 till date (Wang *et al.*, 2018; Bankar *et al.*, 2015; Dusséaux *et al.*, 2013; Jang *et al.*, 2013; Dai *et al.*, 2012; Collas *et al.*, 2012; Lee *et al.*, 2012) is given in table 5.3 (later on in his chapter). Reported strategies of IBE production from *C. acetobutylicum* include the metabolic engineering of *C. acetobutylicum* to harbour plasmids bearing the *C. beijerinckii SadH* gene, synthetic acetone genes under regulation of various promoters, regulation of intracellular NAD(P)H levels and the use of hyper-producing or *buk- C. acetobutylicum* strains. Fermentations have been performed in flasks, bioreactors and there is a report of the use of a 200-l pilot scale fermentation (Jang *et al.*, 2013). So far, the highest IBE reported is 27.9 g/l by Jang *et al.* (2013). This particular study used a hyper-producing *C. acetobutylicum buk*$^{-}$ strain harbouring P_{thl}-*Sadh-hydG* on plasmid, BKM19 (pIPA100) in a controlled batch fermentation with 100g/l glucose.

Table 5.3 gives a summary of strategies for strain engineering/process optimization carried out in *C. acetobutylicum* and the IBE titres recorded.

5.1.5 Orthogonal Inducible System

The capacity to independently regulate multiple genes in parallel is desirable in synthetic biology. Orthogonal systems are uncoupled form evolutionary constraints and independent of cellular regulation making use of heterologous elements like bacteriophage-encoded T7 RNA polymerase for transcription (An and Chin, 2009; An and Chin, 2011). The key advantage of using such polymerase is the high specificity to target promoters. As a replacement for the T7 RNAP which could be associated with toxicity (Temme *et al*., 2012), here in the Clostridium Research Group (CRG) an orthogonal gene expression system based on the *C. difficile* sigma factor TcdR has been constructed. The TcdR belongs to the σ^{70} family of sigma factors (a novel sub category, group 5) which only recognises promoters with specific conserved non-canonical -35 (or -10) recognition sequences (Dupuy *et al*., 2006). The functionally interchangeable homologues of TcdR are conserved in other toxigenic clostridia such as *C. botulinum* (BotR), *C. tetani* (TetR) and *C. perfringens* (UviA) (Dupuy *et al.*, 2006). Activation of target promoters (P_{tcdA} and P_{tcdB}) is only possible in the presence of TcdR. Figure 5.2 gives a schematic of the *tcdR* loci in *C. difficile.*

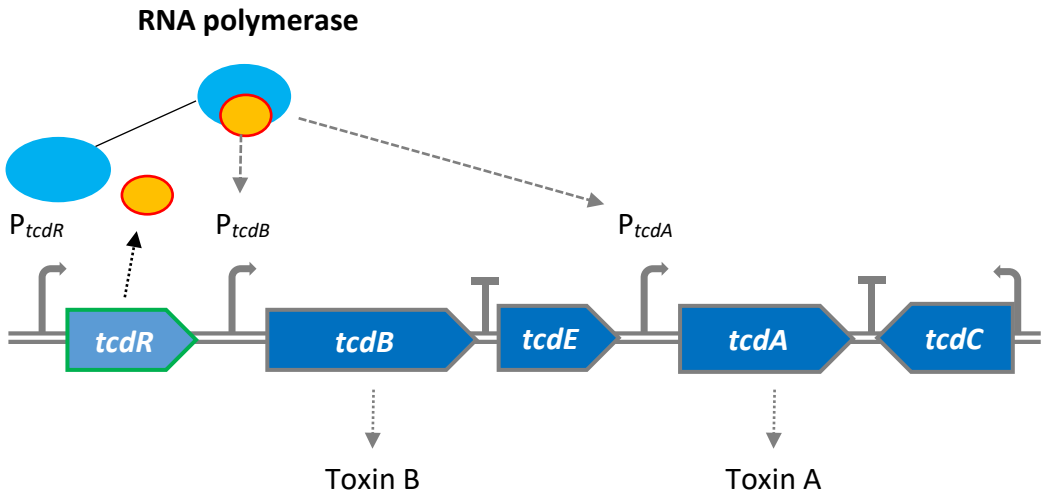


Figure 5. 2: The *tcdR* locus in *C. difficile*.

TcdR allows specific recognition of *tcdA* and *tcdB* toxin genes promoters by RNA polymerase, as well as recognition of its own promoter. A holin-like protein is encoded by *tcdE* and *tcdC* is a negative regulator of toxin expression

Till now, in the CRG, the inducible *tcdR*-P_{tcdB} orthogonal system has given the highest levels of expression of heterologous genes even compared with one of our strongest constitutive promoters, P_{fdx}. This was demonstrated in the expression of *catP* in *C. acetobutylicum* (Wang Hengzheng, 2014), nitrofurazone reductase *in C. sporogenes* (Zygoroupoulou Maria, 2019) and *C. autoethanogenum* (Christopher Humphreys, Personal Communication).

5.2 Aim of Chapter

With the creation of a triple auxotrophic mutant in *C. acetobutylicum* representing three different loci at which genes could be inserted into the chromosome (chapter 4), and the availability of the *tcdR* system, the aim of this chapter was to test the utility of the inducible *tcdR*-P_{tcdB} system in the *C.*

acetobutylicum triple mutant, *C acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*, for the production of isopropanol.

5.3 Resullts

5.3.1 Creation of ACE plasmids

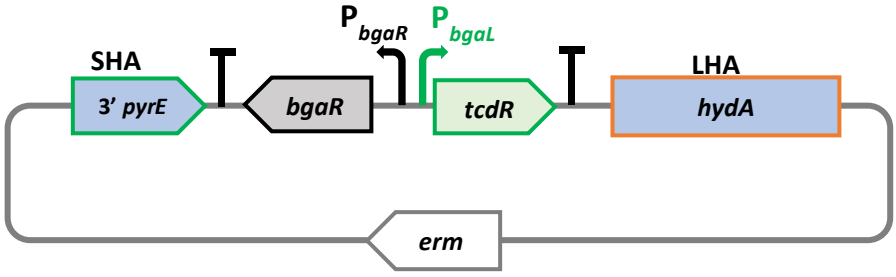
The following ACE plasmids were created to be integrated at the *pyrE, argH and pheA* loci. As described in chapter 4.4.1, expression vectors were designed to direct the integration of the following genes at the three loci created:

- ***pyrE* locus** - Lactose inducible *tcdR* system
- ***argH* locus** - Acetone operon (P_{tcdB}-*ctfA/B-adc*)
- ***pheA* locus** - Secondary dehydrogenase (P_{tcdB} –*SadH*)

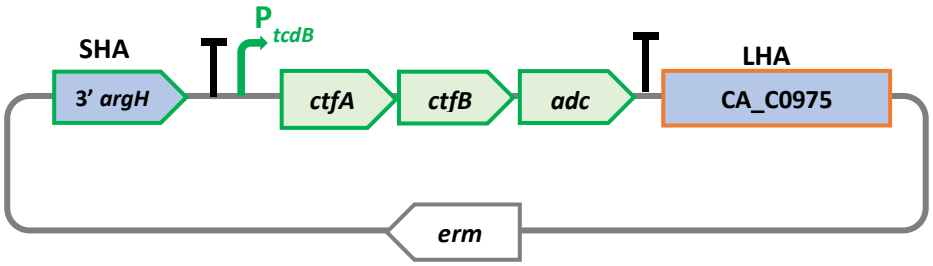
The ACE expression vectors incorporated a Left Homology Arm (LHA) which comprised the 3'-end of the gene being targeted (*pyrE, argH* or *pheA*) and a Right Homology Arm (RHA) which encompassed the 1200bp region from downstream of the 3'-end of the gene target (*pyrE, argH* or *pheA*). On the basis of size, the LHA was the Short Homology Arm (SHA) while the RHA was the Long Homology Arm (LHA). The genes to be expressed were cloned between the homology arms to give plasmids pMTL-HZ13-HZ-tcdR, pMTL-HZ14-PtcdB-ctfAB-adc and pMTL-HZ15-PtcdB-SadH. For plasmid pMTL-HZ13-HZ-tcdR, the lactose inducible *tcdR* system, *bgaR-tcdR* was cut from template, pMTL-ME6C::HZ-tcdR. For plasmid pMTL-HZ14-PtcdB-ctfAB-adc, the *ctfAB* and *adc* were amplified from the pSOL1 plasmid in *C. acetobutylicum* 824 and placed under the transcriptional control of the *tcdB* promoter, P_{tcdB}. For plasmid pMTL-HZ15-PtcdB-SadH, the *SadH* gene was amplified from *C. beijerinckii*

and also positioned under the transcriptional control of the P_{tcdB} promoter. Homology arms were flanked by restriction sites *Sbf*1 and *Asc*1. Figure 5.3 (A), (B) and (C) below show the cloned plasmids.

(A) pMTL-HZ13-HZ-tcdR to enable insertion of the lactose inducible *tcdR* gene at the *pyrE* locus



(B) pMTL-HZ14-PtcdB-ctfAB-adc to mediate insertion of the acetone operon at the *argH* locus



(C) pMTL-HZ15-PtcdB-SadH to direct integration of the *SadH* gene at the *pheA* locus

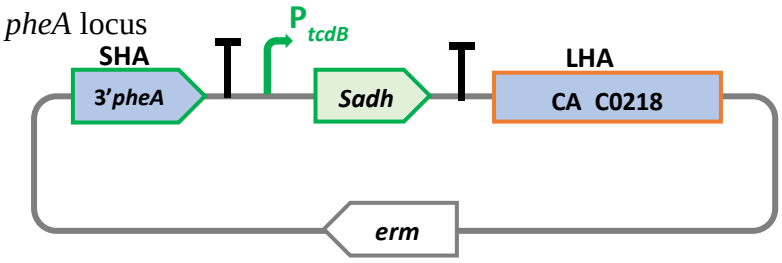


Figure 5. 3: ACE integration plasmids

A, B and C are plasmids pMTL-HZ13-HZ-tcdR, pMTL-HZ14-PtcdB-ctfAB-adc and pMTL-HZ15-PtcdB-SadH to direct integrations at the *pyrE*, *argH* and *pheA* loci respectively. The P_{tcdB} was under the control of the lactose inducible *tcdR* system.

5.3.2 Introduction of Lactose inducible *tcdR* system, HZ-tcdR at the *pyrE* Locus

Competent cells of triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*, were transformed with methylated pMTL-HZ13-HZ-tcdR plasmid and plated on CGM+Ur+arg+phe+erm (erythromycin) plates. After 2 – 3 days, erm^R colonies were restreaked onto CBM+arg+phe plates and incubated for 24 – 48 hrs. Patch plate on CGM+arg+phe ± erm was carried out and a total of 12 colonies (6 each from two independent transformations,) that had become sensitive to erm were selected and PCR screened using gene flanking primers JH14/F and JH14/R. Results are shown in figure 5.4.

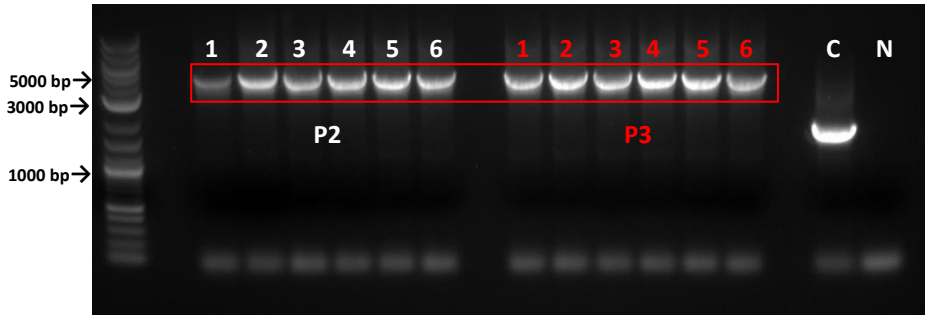


Figure 5. 4: PCR screening of HZ13-HZ-tcdR integration at *pyrE* locus.

PCR screening of HZ13-HZ-tcdR integration at *pyrE* locus using JH14/F and JH14/R. Lanes are genomic DNA of potential candidates. All 12 colonies gave an amplified DNA fragment of 4268 bp, which is consistent with insertion of the lactose inducible *tcdR* gene. In contrast, the control (C), comprising the parental, triple mutant (Δ*pyrE*Δ*argH*Δ*pheA*), gave the expected smaller (1989 bp) amplified DNA fragment. The negative control (N) used water as the PCR template. P2 and P3 represent two independent transformations. Gel extracted bands were subjected to Sanger sequencing using primers ACE/F and ACE/3R and shown to be composed of the expected nucleotide sequence

5.3.3 Introduction of Acetone Operon, P_{tcdB}-ctfA/B-adc at the *argH* Locus

The plasmid pMTL-HZ14-ptcdB-ctfA/B-adc was transformed into the strain obtained above, (which already had *tcdR* integrated at the *pyrE* locus) and plated on CGM+arg+phe+erm. However, colonies did not grow when streaked onto minimal plates lacking arginine supplementation. Two further attempts were made, during which a few colonies appeared on unsupplemented plates after seven days of incubation. However, the size of the DNA fragment PCR amplified from their isolated DNA using *argH* gene flanking primers arg_ch_F/R was the same size as the WT control (figure 5.5).

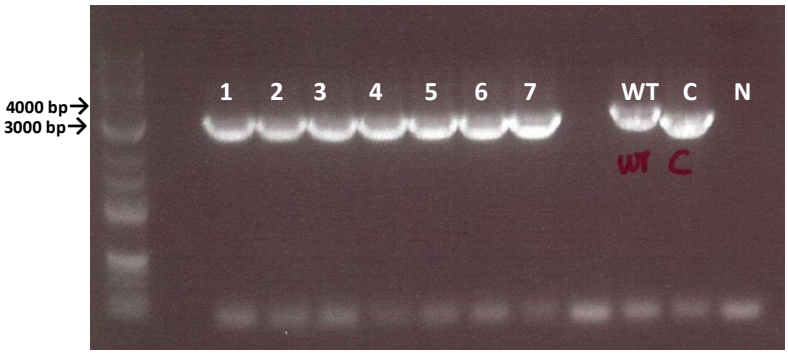


Figure 5. 5: PCR screening of ptcdB-ctfA/B-adc integration at *argH* locus.

PCR screening of ptcdB-ctfA/B-adc integration at *argH* locus. Lanes are genomic DNA of potential candidates. All 7 colonies did not give DNA fragments of the expected size, 4268 bp, but instead gave a DNA fragment with a size equivalent to the triple mutant, Δ*pyrE*Δ*argH*Δ*pheA*, used as a control (C), at 3186bp. WT was at 3547bp and water used as the negative control (N).

5.3.3.1 Investigating the Difficulties Associated with Integration at *argH* Locus

5.3.3.1.1 Control Repair of *argH* gene

As the integration of the acetone operon at the *argH* had proven difficult, the ability to simply repair the *argH* gene in the mutant cells being used was tested using the empty *argH* complementation vector, pMTL_comp_argH. Using this plasmid, prototrophic colonies (able to grow on minimal media lacking exogenous arginine) were readily obtained. Template DNA made from these colonies were found to generate a DNA fragment, following PCR using flanking primers, that was of a size entirely consistent with repair of the *argH* locus (figure 5.6).

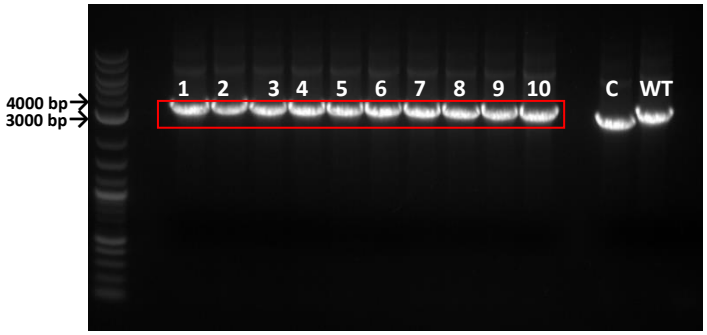


Figure 5. 6: Control repair of *argH* gene.

PCR screening of *argH* gene repair using gene flanking primers. Lanes are genomic DNA of potential candidates. The ten repaired mutants (lanes 1-10) generate a DNA fragment of equivalent size (3547 bp) to the WT using the flanking primers arg_ch_F/R. The control (C) used was the parental strain carrying the *argH* deletion, and gave a PCR fragment of some 3186 bp.

As it has been shown that the repair of the *argH* locus was possible, there is the possibility that perhaps the difficulty in integrating the acetone operon at *argH*

locus was due to the presence of the pSOL1 plasmid (which carries the same acetone genes as in the ACE plasmid) with which the ACE plasmid may preferentially recombine. To test this hypothesis, it was decided to eliminate the pSOL1 plasmid from the cell.

5.3.3.1.2 Deletion of pSOL1

The ACE plasmid pMTL-JH36 carries a promoter-less copy of an *erm* gene that is preceded by an allelic exchange cassette that shares homology to the pSOL1 megaplasmid. Specifically, the DNA comprises a 1200 bp LHA from upstream of the promoter region of the pSOL1 replication gene *repA*, and a RHA (300 bp) comprising a proximal region of the *repA* gene itself, but lacking the start of the gene. In between the LHA and RHA on pMTL-JH36 is positioned a transcriptional terminator. As the downstream *ermB* gene is promoter-less, the presence of this terminator means that the *ermB* gene is not expressed. Accordingly, plasmid pMTL-JH36 does not confer resistance to erm on the cell. When pMTL-JH36 is transformed into the clostridial cell carrying pSOL1, selection is initially on thiamphenicol. The plasmid integrates into pSOL1 via the larger LHA. These cells are then plated on erm plates. The only way the cells can become resistant to erm is if the integrated plasmid excises and the plasmid formed carries the pSOL1 *repA* gene promoter upstream of *ermB*. In parallel, the *repA* promoter region is exchanged for the terminator formerly present on pMTL-JH36. As a consequence, *repA* is no longer expressed, and pSOL1 is lost from the cell (figure 5.7). Simplistically, cells in which pSOL1 has been lost are selected on the basis of acquisition of resistance to erm. The

plasmid, pMTL-JH36, was obtained from the Clostridia Research Group Culture Collection and was made by John Heap.

The plasmid pMTL-JH36 was transformed into the *C. acetobutylicum* triple mutant (Δ*pyrE*Δ*argH*Δ*pheA*) which had HZ-tcdR integrated at the *pyrE* locus. To check for the presence of pSOL1, a simple starch plate test was performed by dispensing 20µl of an overnight culture onto agar plates containing starch. For comparative purposes, the WT cells were also included as a positive control. The pSOL1 bears the *amyP* gene (CA_P0168) which produces extracellular α-amylase when grown on medium containing starch as the sole carbon source (Ehsaan *et al.*, 2016). Cells that have lost the pSOL1 plasmid should not form a halo on starch plate (figure 5.8).

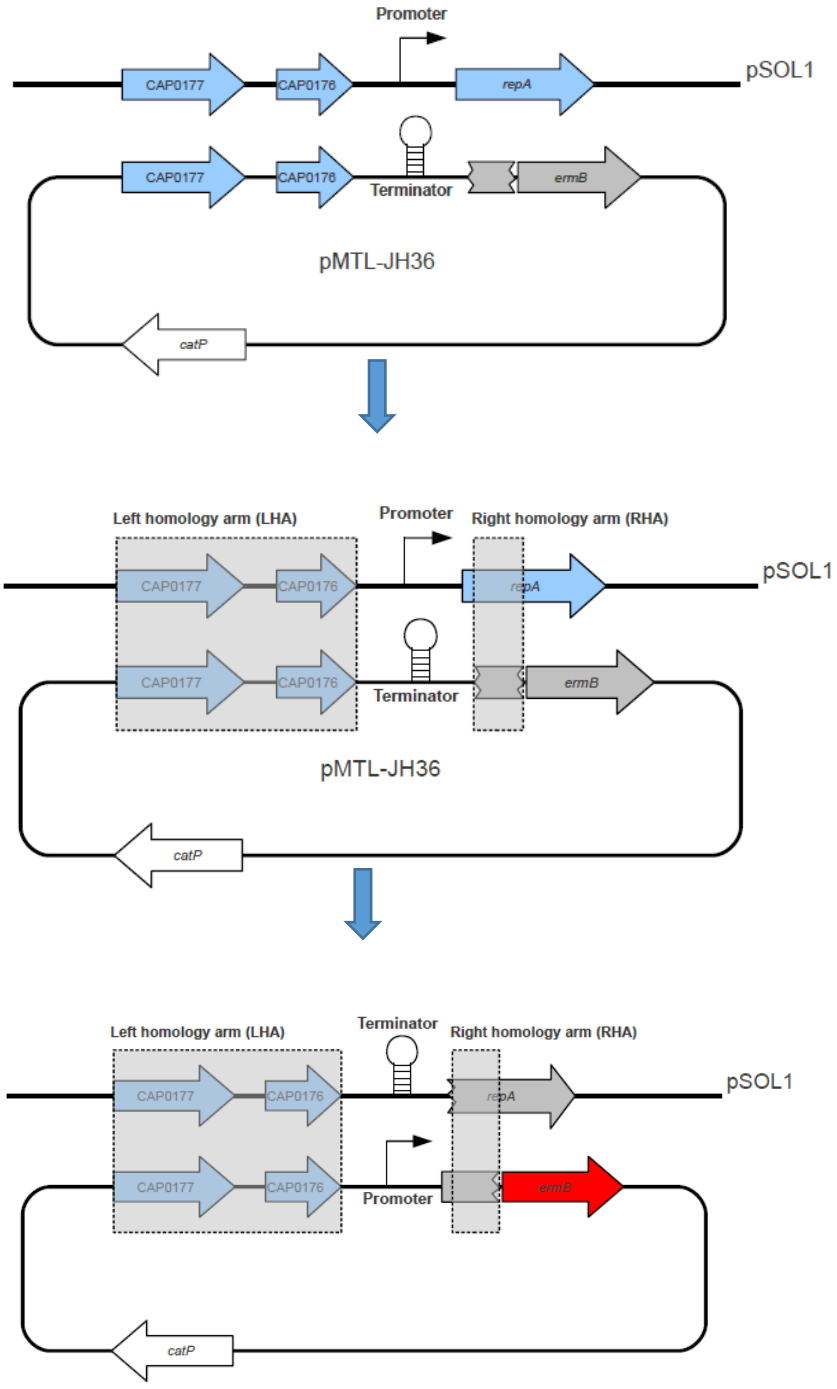


Figure 5. 7: pSOL1 removal using ACE.

First integration occurs via LHA and second recombination is through the RHA which directs plasmid excision and activates the *erm* gene which is now downstream of the *repA* promoter. The *repA* gene of pSOL1 becomes inactivated and pSOL1 is lost as it cannot replicate**.**

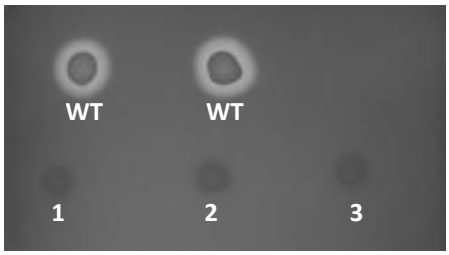


Figure 5. 8: A simple starch plate test for loss of pSOL1 plasmid.

Starch test for presence of pSOL1 plasmid in mutants; 20µl of an overnight culture was dropped onto agar plates containing starch and mutants that have lost pSOL1 plasmid (1, 2, 3) do not form a halo, whereas the wildtype (WT) does.

Having eliminated pSOL1, the mutant strain was again transformed with plasmid pMTL-HZ14-ptcdB-ctfA/B-adc.

Once again, it proved extremely difficult to obtain a strain in which the acetone operon had apparently integrated at the *argH* locus. Eventually, a number of candidate clones were obtained which had become prototrophic for arginine and which gave DNA fragments of the correct size in a PCR using flanking primers (see figure 5.9). However, inexplicably, it subsequently proved impossible to lose the ACE plasmid. Thus, when integrants were restreaked onto CBM+Ph plates and then patch plated onto the same media with and without erm, no clone had become sensitive to this antibiotic, indicative of plasmid loss. Equivalent results were obtained when cells were cultured in minimal P2 liquid medium, before being diluted and patch plated onto CGM+Ph ± erm. No plasmid loss was observed.

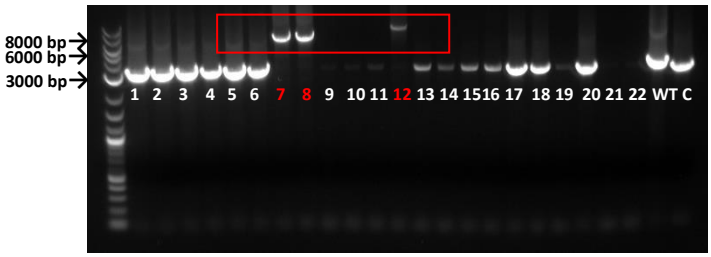


Figure 5. 9: Screening for integration of acetone operon at the *argH* locus

Screening of more candidates (*C. acetobutylicum* + tcdR + ctfA/B-adc) for integration of the acetone operon at the *argH* locus. Lanes are genomic DNA of potential candidates. Lanes 7, 8 &12 show evidence of a larger DNA fragment of 6353 bp in size, indicative of insertion of the acetone operon. The other clones produce a similar smaller sized (3186 bp) DNA fragment consistent with the *argH* deletion of the parental control (C). The WT sample produces the expected slightly larger 3547 bp amplified DNA fragment**.**

5.3.3.1.3 Construction and Evaluation of Plasmid Variants

In order to overcome the challenge of non-plasmid loss, two variants of the plasmid, pMTL-HZ14-ptcdB-ctfA/B-adc were constructed, namely:

(1) pMTL-HZ14-PtcdB-ctfAB-adc-suicide (a suicide plasmid variant) – by cutting out the Gram-positive replicon between *Asc*1 and *Fse*1 (figure 5.10) so plasmids will not be able to replicate autonomously in cell and will be lost. This was achieved by digesting with *Asc*1 and *Fse*1 restriction enzymes followed by blunt end ligation.

(2) pMTL-HZ14-PtcdB-ctfAB-adc-codA - to include *codA* as a negative selection marker to allow the direct selection of plasmid loss. This was achieved by replacing the erm marker in pMTL-HZ14-ptcdB-ctfA/B-adc with erm-*codA* from pMTL-SC1-add9-codA-ermB using restriction enzymes *Pme*1 and *Fse*1.

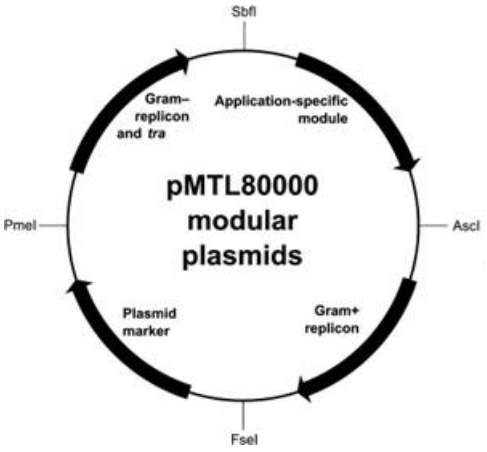


Figure 5. 10: pMTL80000 modular plasmids

The general structure of the pMTL80000 modular plasmids on which all the plasmids in this study are based. (www.clostron.com)

Both plasmid variants were methylated and transformed into triple mutant, *C acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*. Using the suicide plasmid variant, no transformants were observed on CGM+Ur+arg+phe+erm plate. Using the *codA* plasmid variant, 5 transformants were observed per plate (CGM+Ur+arg+phe+erm) which could grow on plates lacking arginine indicating possible integration at the *argH* locus. These candidates were PCR confirmed using flanking primers argH_ch_F and argH_ch_R, subjected to Sanger sequencing using primers ACEHZ14/F and ACE/3R and shown to be composed of the expected nucleotide sequence. Colonies confirmed to have acetone operon integrated at the *argH* locus were plated on 5-Flouro cytosine plates (CBM+Ur+Phe+FC) as the *codA* gene encoding cytosine deaminase will convert 5 FC to the highly toxic 5-FU (5-Flourouracil). This means that only cells that have lost the plasmid should be able to survive in the presence of 5-FC. Just a few colonies (compared to previous experience using 5- FC) grew on

FC plate and were shown to still be resistant to erm when patch plate was done on plates with and without erm indicating continued presence of plasmid.

5.3.3.2 Swapping Integration Loci

To test if the difficulty encountered in obtaining integration at this locus was in any way related to the presence of the HZ-*tcdR* already integrated at the *pyrE* locus, an attempt was made to integrate the acetone operon into the *argH* locus in the *C. acetobutylicum* triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA* lacking this insertion. No integrants were obtained.

To further investigate if this problem was peculiar to the integration of acetone operon at the *argH* locus, a new plasmid, pMTL HZ14-HZ-tcdR was constructed to direct the integration of the lactose inducible *tcdR* system at the *argH* locus. This plasmid was transformed into *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA* but isolation of the desired integrants proved not to be possible (figure 5.11), and all of the tested clones contained no insert of the correct size.

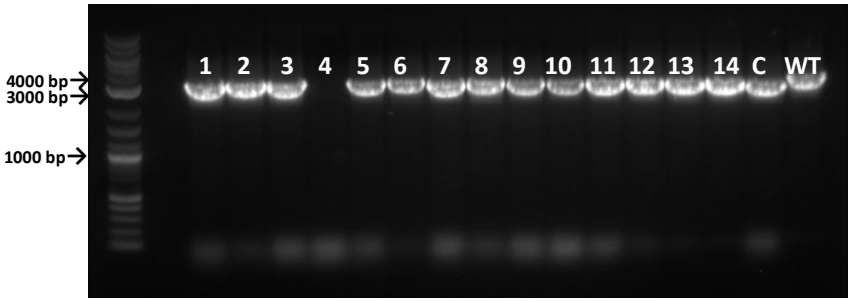


Figure 5. 11: Integration of HZ-tcdR at the *argH* locus
PCR screening of HZ-tcdR integration at the *argH* locus using flanking primers, arg_ch_F/R. Lanes are genomic DNA of potential candidates. All colonies screened generated a DNA fragment 3186 bp in size same as the triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA* used as Control (C) compared to the 3547 bp fragment in the WT. The predicted size of the DNA fragment amplified if integration of the acetone operon

had occurred was 5826 bp. DNA fragments were Sanger sequenced and nucleotide sequence were as the parent strain, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*

5.3.4 Change of Integration Strategy

At this point, it is not very clear why the integration of acetone operon at the *argH* locus was met with much difficulty however we speculate that the insert size could be a factor as we were previously able to integrate *catP* gene (about 650 bp) at this locus but the larger genes/operons (HZ-tcdR and ptcdB-ctfA/B-adc about 3000 bp) proved difficult even after several attempts. When very few integrants were isolated, they were unstable, non-reproducible and plasmid loss impossible.

In the light of the difficulties encountered in the integration at the *argH* locus, which was confirmed to also be the case in *Acetobacterium woodii* (personal communication, Jon Baker), the original plan was adjusted as follows:

1. Acetone operon (P_{tcdB} -ctfA/B-adc) to now go at the *pyrE* locus
2. Secondary dehydrogenase (P_{tcdB} –adh) at *pheA* locus
3. Repair *argH* locus
4. Introduce Lactose inducible *tcdR* system (HZ- tcdR) on plasmid basis.

Consequently, new plasmid, HZ13-PtcdB-ctfA/B-adc, was created to direct integration of acetone operon at the *pyrE* locus (figure 5.12).

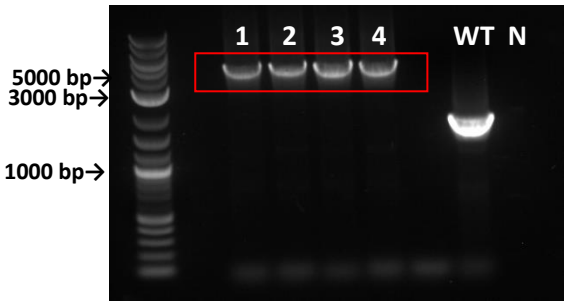


Figure 5. 12: Integration of acetone operon at the *pyrE* locus of the triple mutant.

Integration of acetone operon at the *pyrE* locus of the triple mutant using plasmid HZ13-tcdR-ctfA/B-adc. Lanes are genomic DNA of potential candidates. Transformants that could grow on non-supplemented minimal plate, P2+Ar+Ph were patch plated on P2+Ar+Ph ± Erm. Colonies that had lost plasmid and sensitive to erm were PCR confirmed using primers JH14/F and JH14/R. The four clones screened (lanes 1-4) all generated a DNA fragment of the expected size (4703 bp) in PCR, compared to the WT fragment of 1989 bp.

Next, the *SadH* gene was integrated at the *pheA* locus using pMTL-HZ15-ptcdB-SadH as shown in figure 5.13.

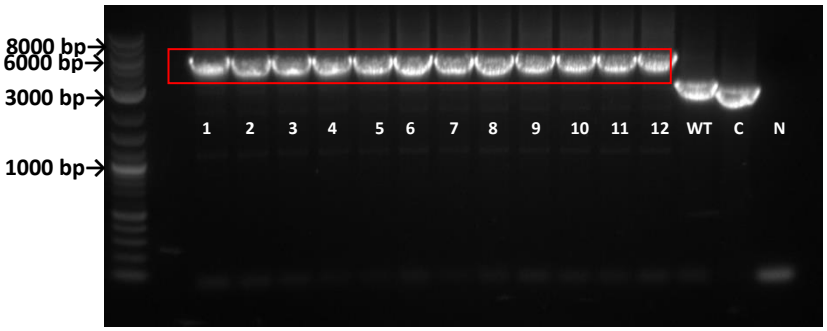


Figure 5. 13: *SadH* integration at the *pheA* locus of triple mutant

SadH integration at the *pheA* locus of *C. acetobutylicum* Δ*pyrE*Δ*argHpheA* with acetone operon integrated at the *pyrE* locus. Lanes are genomic DNA of potential candidates. PCR screening was undertaken with the flanking primers pheA_ch_F/R. Expected band sizes are 6390 bp with integration in lanes 1 - 12); 2817 bp for *pheA* deletion (as in the control, C) and 3090 bp for Wild Type (WT)

Following, the isolation of a clone containing the acetone operon at the *pyrE* locus and *SadH* gene at *pheA* locus, the *argH* mutant allele was restored to wildtype using the pMTL_comp_argH plasmid and selection for prototrophy. To complete the construction of the strain, plasmid pMTL82251-HZ_tcdR bearing the lactose inducible *tcdR* gene was introduced into the strain as an autonomous plasmid.

The final strain was designated **824BO1** and characterised by:

- a genome integrated acetone operon (P_{tcdB} -ctfA/B-adc) at the *pyrE* locus
- a genome integrated secondary alcohol dehydrogenase gene (P_{tcdB} –adh) at the *pheA* locus;
- a repaired *argH* locus (corresponding essentially to WT);
- carriage of an autonomous plasmid carrying a lactose inducible *tcdR* system (HZ-*tcdR*), and;
- the absence of the pSOL1 megaplasmid.

For comparative purposes, an equivalent strain to **824BO1** was made which did not carry the *SadH* gene at the *pheA* locus, rather the *pheA* locus was simply restored to WT. This strain was designated **824BO2.**

Finally, a third strain was made, **824BO3**, which retained pSOL1 and in which the lactose inducible *tcdR* system was integrated at the *pyrE* locus and the acetone operon was localised to the autonomously replicating plasmid.

The essential features of the strains created are summarised in Table 5.1

Table 5. 1 Table 5.1: Summary of Engineered Strains

Strain	*pyrE*	*pheA*	*argH*	Autonomous Plasmid	pSOL1	Products
824BO1	P_{tcdB} -*ctfA/B-adc*	P_{tcdB} -*SadH*	WT	BgaR-P_{bgal}-TcdR	NO	Acetone & IPA
824BO2	P_{tcdB} -*ctfA/B-adc*	WT	WT	BgaR-P_{bgal}-TcdR	NO	Acetone
824BO3	BglR-P_{bglA}-TcdR	P_{tcdB} -*SadH*	WT	P_{tcdB} -*ctfA/B-adc*	YES	Acetone & IPA
824BO4	BglR-P_{bglA}-TcdR	P_{tcdB} -*SadH*	WT	None	YES	Acetone & IPA

5.3.5 Preliminary Fermentation Tests

Preliminary fermentation tests were undertaken with a strain in which the acetone operon had not been added, only the *SadH* gene needed for isopropanol production and in which this gene was under the transcriptional control of the inducible *tcdR* gene. This equated to the *C acetobutylicum* triple mutant (*C. acetobutylicum* Δ*pyrE*Δ*argHpheA*) in which: (i) the inducible *tcdR* gene had been integrated at the *pyrE* locus; (ii) the *SadH* gene under the control of the P_{tcdB} promoter was integrated at the *pheA* locus, and; (iii) the *argH* locus had been restored to prototrophy. This strain was designated 824BO4. The solvent profile of two replicates of this 824BO4 strain, 2 and 6, is shown in figure 5.14.

Batch fermentation in 60 ml CBMS was undertaken with the 824BO4 strains and samples taken at intervals for measurement of OD, pH and solvent profile analysis. Induction with 50 mM lactose was performed at 8 hrs. This concentration of lactose was initially used based on previous data (Wang, Hengzheng, 2014) which demonstrated that optimum induction was observed at this level of lactose induction. The 8 hrs time point was chosen for this first test because that is typically the period of exponential growth and onset of

acidogenesis and it was thought to express the *SadH* gene early so that it is ready to convert acetone to isopropanol upon production. The early production of isopropanol and minimized accumulation of acetate has been reported when either the constitutive P_{thl} or the P_{ptb} (which is active during acidogenesis) promoters were used to express the acetone genes, *ctfA/B-adc* and *SadH* in *C. acetobutylicum* (Dusséaux *et al*., 2013). Figure 5.14 shows the results obtained.

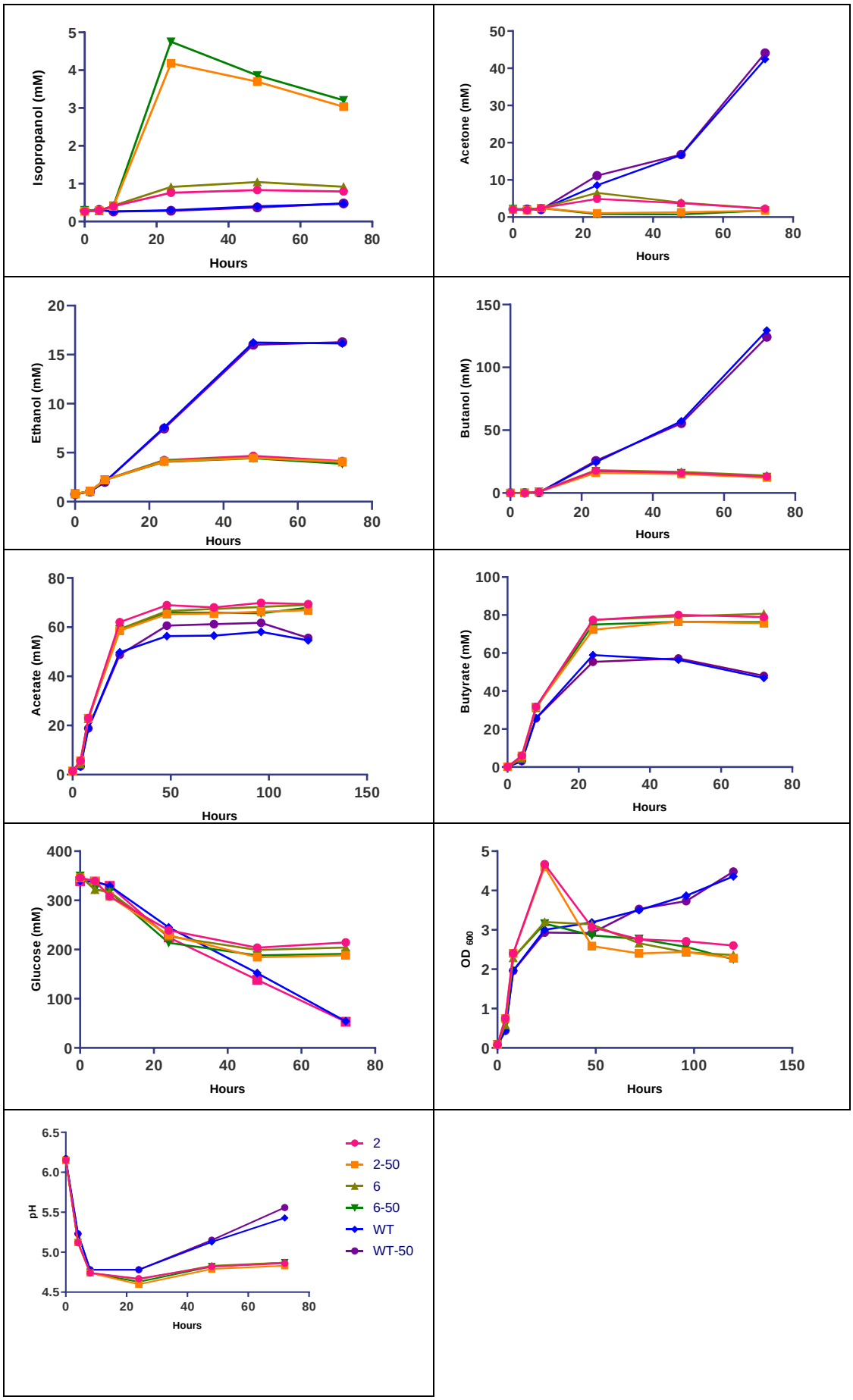

Isopropanol (mM)
Hours
Acetone (mM)
Hours
Ethanol (mM)
Hours
Butanol (mM)
Hours
Acetate (mM)
Hours
Butyrate (mM)
Hours
Glucose (mM)
Hours
OD 600
Hours
pH
Hours
2
2-50
6
6-50
WT
WT-50

Figure 5. 14: Solvent profile analysis of 824BO4-2 and 824BO4-6.

2 and 6 are biological replicates and 2-50, 6-50 indicate 50 mM lactose induction at 8 hours. WT is *C. acetobutylicum* ATCC 824. Pink circle, non-induced 824BO4-2; orange square, induced 824BO4-2; olive triangle, non-induced 824BO4-6; green triangle, induced 824BO4-6; blue diamond, non-induced WT; purple circle, induced WT.

From figure 5.2, with the introduction of the *SadH* gene, the successful conversion of acetone to isopropanol at up to 4.8 mM was achieved with strain 6 at 24 hours. Solvent concentrations were lower in 824BO4-2 & 6 than in the wild type. The *SadH* gene is NADPH dependent just as the butanol dehydrogenase is in the production of butanol (Gheshlaghi *et al*., 2009). The introduction of *SadH* may have resulted in decreased availability of NADPH due to increased demand (Wang *et al*., 2018) which may have resulted in the reduced butanol titres observed. As expected, the acetone titer was lower in 824BO4-2 & 6 because of its conversion to isopropanol by *SadH*.

More acids were produced in strains 2 and 6 and the typical acid re-assimilation pattern was not observed compared with wild type. Similar results have been reported where there was a decrease in glucose consumption, total solvents produced (Collas *et al*., 2012; Lee *et al*., 2012) as well as an unexpected increase in residual acids (Lee *et al*., 2012) with the expression of the *SadH* gene in *C. acetobutylicum.* These researchers further noted that with the introduction of the synthetic acetone pathway genes, there was increased solvent production as well as glucose utilization.

Having shown that the inducible *tcdR* system can be used to produce isopropanol from the metabolically engineered *C. acetobutylicum* strain, the following fermentation tests were carried out.

5.3.6 Fermentation Tests using 824BO1 and 824BO2

Batch fermentation in flasks was carried out similar to the preliminary tests except that CGM was used now and lactose concentrations used were 0, 10 and 100 mM.

Isopropanol was produced in a dose-dependent manner when *tcdR* was induced with lactose. Up to 18.65 mM isopropanol was produced at 24 hours (824BO1 -100 in figure 5.15) with residual acetone of 2.73 mM. In the control strain, 824BO2 which harboured just the *ctfA/B-adc* genes but not the *SadH* gene, up to 20.05 mM acetone was produced (824BO2 -100 in figure 5.15) and no isopropanol. There was little or no butanol observed in both 824BO1 and 824BO2 and the acids produced did not show the typical re-assimilation pattern during solventogenesis owing to the absence of the pSOL1 plasmid. Initially, the pSOL1 was deleted from the strains because of the difficulty encountered with integration at the *argH* locus as earlier discussed in this chapter. It was then thought that in the absence of the pSOL1 plasmid, the acids produced would be re-assimilated by the integrated *ctfA/B* to form acetone and subsequently isopropanol. It has been shown that the uptake of acids is coupled with the production of acetone, (Hartmanis *et al.*, 1984).

However, the total solvent production from 824BO1 and 824BO2 was poor. Solvent production in *C. acetobutylicum* serves primarily for acid detoxification. In strains 824BO1 and 824BO2 which lacked the pSOL1 plasmid, there was acid production until 24 hours and no further increase after that, the same was observed for the solvents which actually decreased after 24 hours and glucose consumption which ceased at about 24 hours. A possible reason for this could

be that the cells had become metabolically inactive if not largely dead after this time point due to high concentrations of acids.

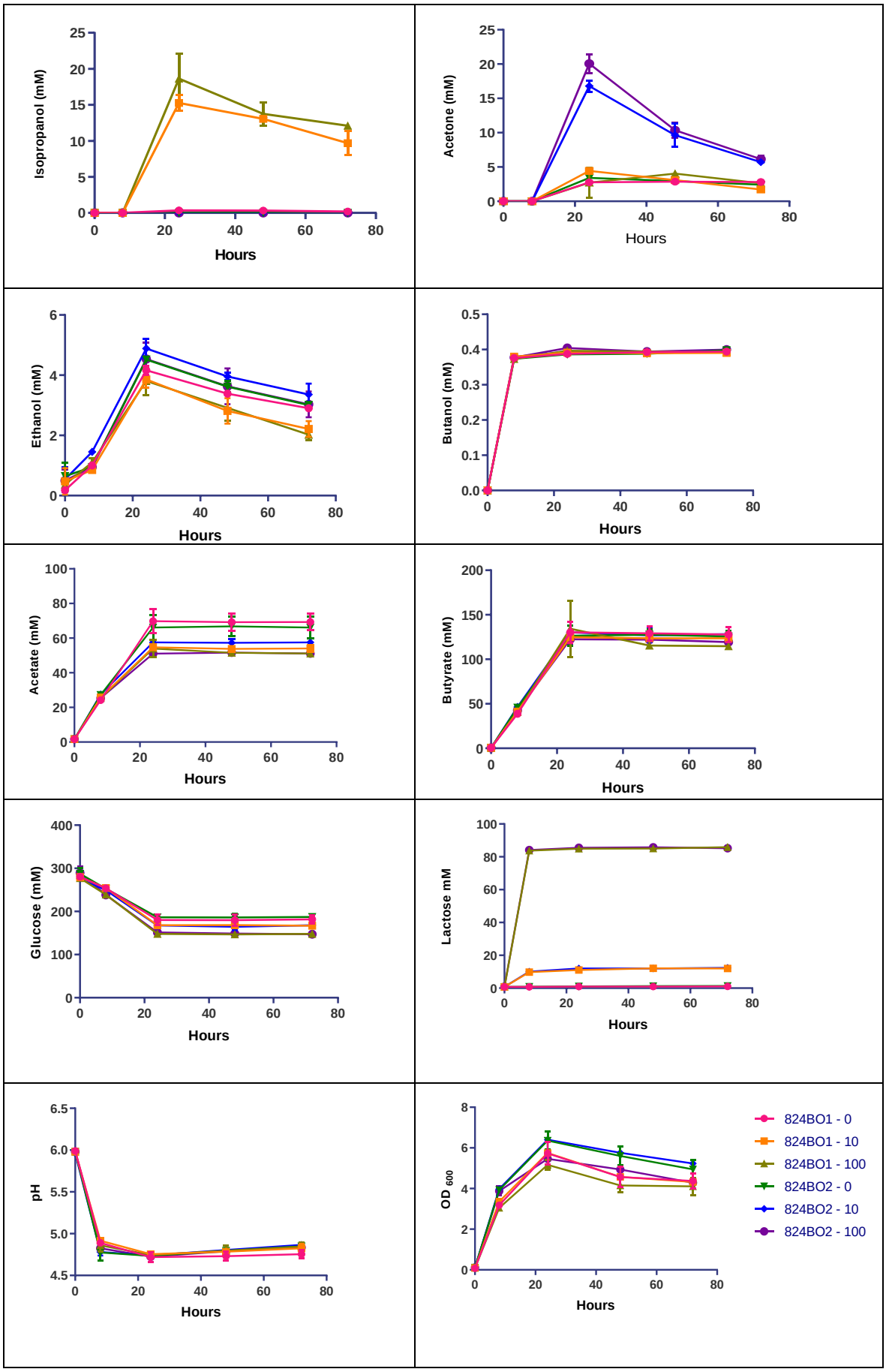

Isopropanol (mM)
Hours
Acetone (mM)
Hours
Ethanol (mM)
Hours
Butanol (mM)
Hours
Acetate (mM)
Hours
Butyrate (mM)
Hours
Glucose (mM)
Hours
Lactose mM
Hours
pH
Hours
OD 600
Hours
824BO1 - 0
824BO1 - 10
824BO1 - 100
824BO2 - 0
824BO2 - 10
824BO2 - 100

Figure 5. 15: Solvent profile analysis for strains 824BO1 and 824BO2

Lactose induction was implemented at 8 hours using 0, 10 and 100 mM concentrations and production of solvent is observed in a dose-dependent manner. Pink circle, non-induced 824BO1; orange square, 10 mM induced 824BO1; olive triangle, 100 mM induced 824BO1; green triangle, non-induced 824BO2; blue diamond, 10 mM induced 824BO2; purple circle, 100 mM induced 824BO2.

5.3.7 Fermentation Tests using 824BO3 and 824BO4

Batch fermentation was done using 824BO3 (which had *tcdR* and *SadH* integrated in chromosome, *argH* repaired, harboured *ctfA/B* on plasmid and had pSOL1 plasmid intact). As a control, 824BO4-6 (which had *tcdR* and *SadH* integrated in chromosome, *argH* repaired and had pSOL1 plasmid intact) was used. For comparison, solvents and acids profiles in g/l are shown in table 5.2.

Early lactose induction at 4 and 8 hours was considered as it has been reported that early availability of the *SadH* gene resulted in early production of isopropanol and less residual acids. The product profile at these times are given in table 5.2. At 8 hours, higher total solvents was achieved than at 4 hours (18.47 g/l and 13.35 g/l respectively). Also, from the two previous fermentation tests, CGM gave better solvent production than CBM (minimal medium) and CGM is the medium that most previous researchers have used (Lee *et al*., 2012, Dai *et al*., 2012, Jang *et al*., 2013, Dusséaux *et al*., 2013). Consequently, fermentations were carried out in CGM medium with 50 mM lactose induction.

With 824BO3, there was an increase in butanol titer and the total solvents produced (IBE - 18.47 g/l) compared with wild type (ABE - 11.28 g/l). This increase in total solvent could be due to the overexpression of the acetone genes (Mermelstein *et al*., 1993) and the complete utilization of glucose was also

observed. On the other hand, a decrease in total solvents was observed with 824BO4-6 (IBE - 2.25 g/l) which had only the *SadH* gene integrated and just about half of the glucose was used up. In a similar experiment by Lee *et al.,* 2012, total solvent concentration was slightly lower in *C. acetobutylicum,* 824 (pIPA1) harbouring *SadH* on plasmid while higher concentration was observed with 824 (pIPA3) bearing *ctfA/B-adc* in addition to *SadH,* when compared with wild type. Previous work by (Mermelstein *et al*., 1993) noted that overexpressing the *ctfA/B* and *adc* genes led to increased production of total solvents and decrease in residual carboxylic acids concentration.

Table 5. 2: Solvent and Acid profile of strains 824BO3 and 824BO4-6 in g/l

Strain	Lactose Induction (mM)	Glucose Utilized g/l	Acetone g/l	Isopropanol g/l	Butanol g/l	Ethanol g/l	Acetate g/l	Butyrate g/l	IBE/ABE g/l
824BO3	0	25.77 ± 0.08	0.9 ± 0.03	0.1 ± 0.0	2.68 ± 0.31	0.71 ± 0.03	3.72 ± 0.07	5.74 ± 0.42	3.49
	50 (4 hrs)	52.24 ± 0.39	0.93 ± 0.32	2.98 ± 0.23	9.43 ± 1.03	0.94 ± 0.28	3.66 ± 0.30	1.87 ± 0.19	13.35
	50 (8 hrs)	**54.32 ± 0.38**	**0.27 ± 0.09**	**8.47 ± 0.29**	**9.78 ± 0.10**	**0.22 ± 0.07**	**1.75 ± 0.13**	**0.96 ± 0.23**	**18.47**
824BO4-6	0	27.50 ± 1.32	0.75 ± 0.11	0.12 ± 0.0	3.87 ± 0.05	0.60 ± 0.09	3.84 ± 0.4	4.41 ± 0.58	4.59
	50	**24.48 ± 1.36**	**0.2 ± 0.07**	**0.56 ± 0.03**	**1.53 ± 0.15**	**0.16 ± 0.06**	**3.85 ± 0.02**	**7.30 ± 0.24**	**2. 25**
Wild Type		**41.16 ± 2.70**	**2.02 ± 0.13**	**N/A**	**7.65 ± 0.43**	**1.61 ± 0.11**	**4.06 ± 0.17**	**2.55 ± 0.03**	**11.28**

*__N/A-__ Not applicable

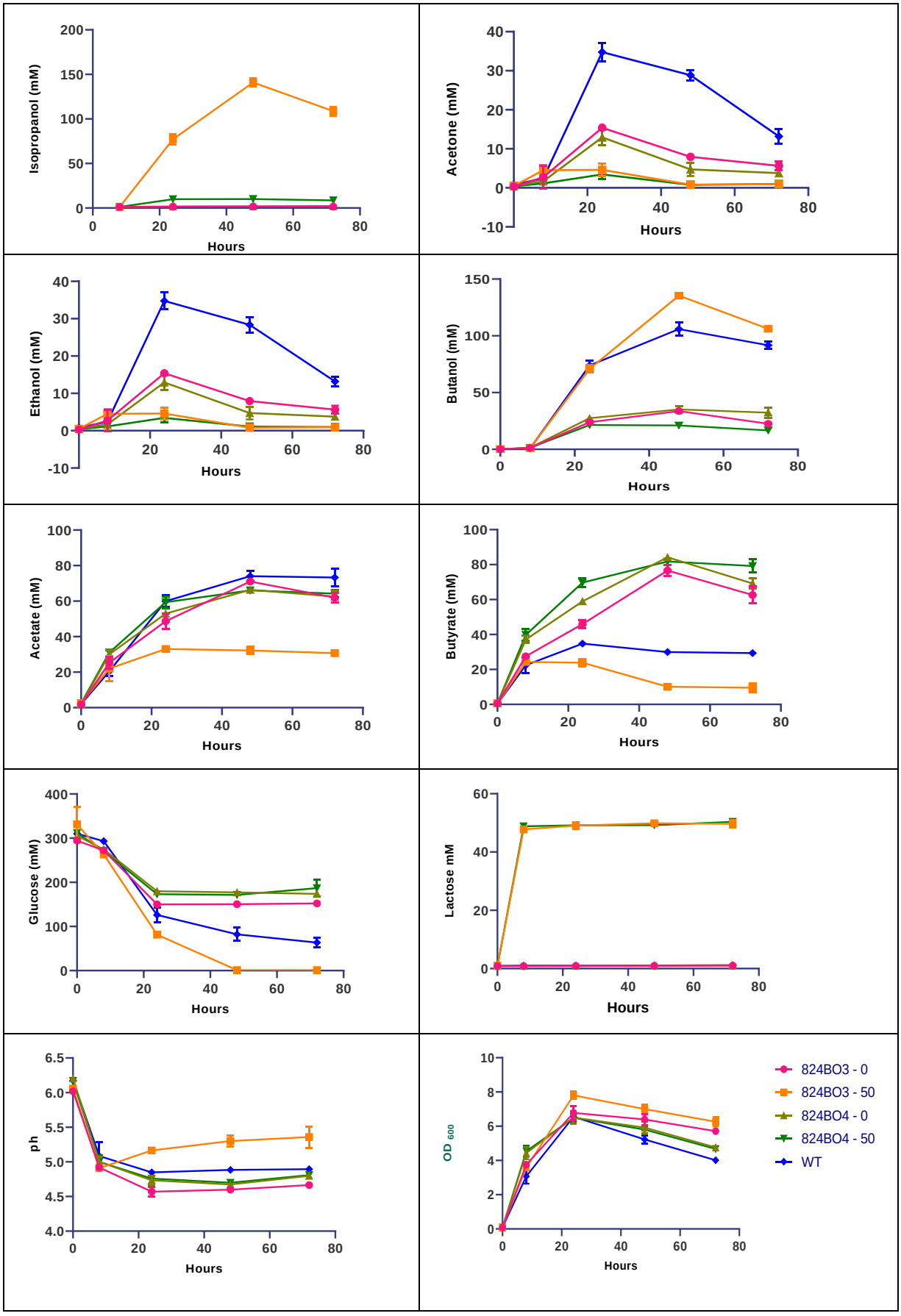

Isopropanol (mM)
Hours
Acetone (mM)
Hours
Ethanol (mM)
Hours
Butanol (mM)
Hours
Acetate (mM)
Hours
Butyrate (mM)
Hours
Glucose (mM)
Hours
Lactose mM
Hours
ph
Hours
OD 600
Hours
824BO3 - 0
824BO3 - 50
824BO4 - 0
824BO4 - 50
WT

Figure 5. 16: Solvent profile analysis for strains 824BO3 and 824BO4-6.

These values are expressed in g/l in table 5.2. Lactose induction was done at 50 mM. Pink circle, non-induced 824BO3; orange square, induced 824BO3; olive triangle, non-induced 824BO4-6; green triangle, induced 824BO4-6; blue diamond, WT.

Less butanol was produced with 824BO4-6 perhaps due to the increase in demand for NADPH as both butanol dehydrogenase responsible for reduction of butyraldehyde to butanol and the secondary dehydrogenase introduced to convert acetone to isopropanol are NADPH dependent enzymes (Gheshlaghi *et al*., 2009). *C. acetobutylicum* is unlike many bacteria that are able to utilize the oxidative pentose phosphate pathway (PP) and the main route for NADPH generation is through reduced ferredoxin produced by the phosphoroclastic conversion of pyruvate to acetyl-CoA via pyruvate ferredoxin oxidoreductase (PFOR). Depending on the cellular state, reduced ferredoxin can produce H_2, NADH or NADPH (Petitdemange *et al.*, 1977). The decrease (0.17 to 0.07 µmol/g DCW) in NADPH with recombinant *C. acetobutylicum* strain harbouring the *SadH* gene compared to wild type which was maintained NADPH at a stable level, has also been reported (Wang *et al*., 2018).

A re-assimilation of acids as solvent production increased was observed for 824BO3 and there was less residual acids compared with wild type. This was expected with the overexpression of the *ctfA/B* genes which are responsible for the reutilization of acids. An increase in acetone and subsequent isopropanol production was not unexpected as the 824BO3 strain bore additional acetone formation genes. The re-assimilation of acids is also coupled with acetone production (Jones and Woods, 1986). The isopropanol/butanol ratio was 0.86 for 824BO3 while acetone/butanol ratio was 0.27 for wild type.

For strain 824BO4-6, in addition to the low solvent titers, the typical acid re-assimilation pattern was not observed, this could be linked to the glucose utilization which stopped at about 24 hours. However, with the introduction of the acetone genes, there was increased acid assimilation as well as solvent production. This pattern has also been reported by other researchers in *C. acetobutylicum* engineered to harbour the *SadH* gene (Wang *et al.*, 2018, Lee *et al*., 2012)

Isopropanol titre and yield from 824BO3 were 8.47 ± 0.29 g/l and 0.15g/g at 48 hrs with residual acetone of 0.27 ± 0.09 g/l. As expected, less acetone was produced on induction due to conversion to isopropanol as the *SadH* gene was turned on and little or no isopropanol production was observed without induction.

From this study, the total IBE produced by 824BO3 after 48 hrs from 55 g/l glucose in flask fermentation was 18. 47 g/l. In comparison with *Clostridium beijerinckii*, a natural producer of IBE, the results from this study is higher. In 2000, Shaheen *et al*., in a comparative fermentation study of industrial Clostridia strains showed that *C. beijerinckii* NRRL B592 produced 16.2 g/l total solvents from maize mash (80 g/l). Survase *et al.,* 2011 reported the batch production of 5.9 g/l IBE (3.7 g/l butanol and 2.2 g/l isopropanol) while continuous fermentation yielded 7.51 g/l. In a more recent study, *C. beijerinckii* BGS1 was reported to produce 13.6 g/l IBE (10.2 g/l butanol and 3.4 g/l isopropanol) with negligible ethanol from 60 g/l glucose (Zhang *et al.*, 2018).

Table 5.3 gives a summary of IBE production from *C. acetobutylicum* as published so far in literature.

5. 3: Summary of IBE production from *C. acetobutylicum* as in published Literature till date

Authors/Year	Isopropanol (g/l)	Butanol (g/l)	Ethanol (g/l)	IBE (g/l)	Strain Engineering/Process optimization
This Work	8.47	10.02	0.21	18.70	*C. acetobutylicum* with Lac Inducible *TcdR* and P_{tcdB}-SadH integrated via Alelle Coupled Exchange (ACE) at *pyrE* and *pheA* loci respectively and harbouring P_{tcdB} -ctfA/B-adc on plasmid Uncontrolled flask fermentation with 55g/l glucose
Wang *et al.*, 2018	6.62	10.51	1.24	17.77	*C. acetobutylicum* harbouring plasmid the *SadH* gene with *ptb* promoter. Regulation of intracellular NAD(P)H levels via NADK activated by $CaCO_3$. Controlled batch fermentation with 60g/l glucose
Banker *et al.*, 2015	2.51	10.78	NS	18.00	Alelle-Coupled Exchange used to integrate *SadH* gene with *adc* promoter. Controlled batch fermentation.
Dusseaux *et al.*, 2013	2.00 (P_{Ptb}) 2.05 (P_{thl})	NS	NS	8.95(P_{Ptb}) 7.96(P_{thl})	*C. acetobutylicum buk*- strain harbouring P_{Ptb}/P_{thl} – ctfA/B-SadH-adc on plasmid Uncontrolled batch fermentation with 50g/l glucose
Dusseaux *et al.*, 2013	4.75	14.63	1.01	20.40	*C. acetobutylicum buk*$^-$ strain harbouring P_{Ptb} – ctfA/B-SadH-adc on plasmid pH (5) controlled batch fermentation with 80g/l glucose

Jang *et al.*, 2013	3.6	14.8	9.5	27.9	Hyper-producing *C. acetobutylicum buk*- strain harbouring P_{Thl} -S*adh-hydG* on plasmid - BKM19(pIPA100) Controlled batch fermentation with 100g/l glucose
Dai *et al.*, 2012	7.6	15	1.28	23.88	Butanol tolerant *C. acetobutylicum* habouring P_{thl}-*SadH* on plasmid Rh8(*Sadh*) pH (≥5) controlled batch fermentation.
Collas *et al.*, 2012	8.0	NS	NS	24.4	*C acetobutylicum* 824 harbouring P_{thl} -SadH-ctfA/B-adc on plasmid Non-pH controlled 60g/l in batch fermentation bioreactors
Lee *et al.*, 2011	5.1	8.0	0.8	13.9	824(pIPA3): *C. acetobutylicum* harbouring P_{adc} –ctfA/B-adc-SadH Uncontrolled flask fermentation
Lee *et al.*, 2011	6.1	10.2	0.8	17.1	824(pIPA3): *C. acetobutylicum* harbouring P_{adc} –ctfA/B-adc-SadH Controlled (pH 5) batch fermentation in bioreactors with 80g/l glucose
Lee *et al.*, 2011	4.4	14.1	1.9	20.4	824(pIPA3-Cm2): *C. acetobutylicum buk*– strain harbouring P_{adc} –ctfA/B-adc-SadH Controlled (pH 5) batch fermentation in bioreactors with 80g/l glucose

*__NS__- Not stated in paper

From the above table 5.3 the highest IBE so far reported is 27.9 g/l by Jang *et al.* 2013. There, a hyper-producing *C. acetobutylicum buk*- strain harbouring P_{thl}-*Sadh-hydG* on plasmid - BKM19 (pIPA100) was used in a controlled batch fermentation with 100g/l glucose. On a g/g basis, their lab scale fermentation had a 0.27 g/g yield while for this study, it is 0.34 g/g. It is therefore promising that higher titers could be attained when fermentation is undertaken using our strain in pH-controlled batch fermentations in bioreactors. In addition, the little residual acetone means this strain has potentials for large scale IBE production.

5.4 Conclusion

The introduction of genes in the previously created *C. acetobutylicum* triple mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA*, was carried out to extend product streams in *C. acetobutylicum* from ABE to IBE, driven by the lactose inducible *tcdR* system. Integration at the *pyrE* and *pheA* loci was relatively easily accomplished. In contrast, integration at the *argH* locus proved difficult which may have been due to the size of insert or the effect of integration on the gene downstream of *argH.* This therefore required a change of plan where just the first two loci were used, the *argH* locus repaired and the gene that should have gone in at that *argH* was expressed on plasmid basis.

Strains 824BO4-2 & 6 were the initial constructs and contained the *tcdR* and *SadH* gene integrated and did not harbour any autonomous plasmids. Preliminary test was done with this to show that the system works and isopropanol was being produced.

Strains 824BO1 and 824BO2 were made next and had the *ctfA/B* and *SadH* integrated with the *tcdR* on plasmid and had lost the pSOL1 plasmid. More isopropanol was produced compared to preliminary tests but there was poor glucose utilization, solvent production and acid re-assimilation.

Strain 824BO3 had the *tcdR* and *SadH* integrated and carried the acetone genes, *ctfA/B-adc* on an autonomous plasmid. The total solvents (IBE) produced by 824BO3 was 18. 47 g/l from 55 g/l glucose (8.47 g/l isopropanol, 10.02 g/l butanol, 0.21 g/l ethanol) and this seems like the highest recorded isopropanol titer from *C. acetobutylicum* under the conditions applied.

CHAPTER 6

Conclusion and Further Work

CHAPTER 6

CONCLUSION AND FURTHER WORK

Clostridium acetobutylicum represents the model clostridia for solvent production and in the so-called ABE fermentation produces acetone, butanol and ethanol. Over the years a range of genetic tools have been developed that have enabled metabolic engineering of this organism as a means of improving solvent production. Strategies explored include: to enhance the efficiency of the C2 to C4 pathway (Nguyen *et al*., 2018); block or reduce acetate formation (Cooksley *et al*., 2012; Kuit *et al*., 2012); block or reduce acetone production (Jiang *et al*., 2009; Cooksley *et al*., 2012); increase butanol tolerance (Tomas *et al*., 2003); (Mann *et al*., 2012), and; improve substrate utilization (Ren *et al*., 2010).

The aim of this project was first to use a conditional RAM-less ClosTron system to knock out genes that were previously impossible to knock out and possibly identify the compensatory mutations that have allowed its viability. A second aim was to create a system that allows the introduction of genes/operons of interest at discrete loci around the *C. acetobutylicum* chromosome using ACE vectors, which together were placed under the control of a common orthogonal expression system.

6.1 Targeted Mutagenesis using RAM-less ClosTron System

The identification of those factors that influence solvent production through mutation of the encoding genes is key to the rational metabolic engineering and

development of an industrial solvent producing *C. acetobutylicum* strain. Using the ClosTron system, Cooksley *et al*. (2012) reported the large-scale targeted mutagenesis of the genome, generating in the process a set of ten mutants. In a number of instances, it proved impossible to knock-out certain genes. In particular, the *thl* gene could not be inactivated, while only a rare single *ptb* mutant was obtained despite repeated attempts. Moreover, the mutant obtained could not be complemented. NGS analysis of the *ptb* mutant revealed a frameshift mutation in *thl* gene. This led to the hypothesis that perhaps there are compensatory mutations that occur to allow the viability of an otherwise lethal mutation.

Consequently, attempts were made to knock out genes that had been previously impossible to knock-out, and possibly identify any compensatory mutations that may have arisen, using a the RAM-less ClosTron system which allows an initial knock-down of a gene, allowing its isolation and thereafter the selection of a null mutant with the loss of plasmid bearing the LtrA protein.

The knock-out of *hydA*, *hbd, crt* and *bcd* genes were attempted but it only proved possible to obtain a null mutant in the case of *hbd.* Although evidence of intron insertion was observed when putative mutants of *hydA, crt* and *bcd* were PCR screened using EBS and a flanking primer, in each case, no mutant DNA fragment was observed when PCR screening was done with gene flanking primers perhaps because the mutants are present at such a low level that they cannot be detected by PCR.

The *hydA* gene is the major hydrogenase in *C. acetobutylicum* and is responsible for the regeneration of NAD (in addition to the NADH consuming pathway

leading to butyryl production) via hydrogen production during acidogenesis and under the present condition a *hydA* mutant was non-viable. The Soucaille research group had reported the inactivation of the *hydA* and *thlA* genes independently using the intron-based insertion system however these mutants were not viable as it was impossible to lose the plasmid in each case, confirming gene essentiality. The presence of plasmid means that the intron insertion in the mRNA can be spliced out allowing gene expression and cell survival. However, the simultaneous inactivation of *hydA* and *thlA*, creating a double mutant was viable producing mainly ethanol, lactate with little amount of glycerol (Nguyen, 2016). Similarly, Jang *et al*., 2014 could not achieve a null *hydA* mutant in the *C. acetobutylicum* quintuple mutant (*pta, buk, ctfB, adhE1, hydA)* created as there was the continued presence of *ltrA* in the cell. Though the inactivation of the *hydA* gene in *C. pasteurianum* has been reported, this could be due to the presence of a number of [FeFe]-hydrogenases and [NiFe]-hydrogenases present in this organism though not much is known about their specific roles in hydrogen production (Schwarz *et al*., 2017). Though two [FeFe]-hydrogenases, *hydA* and *hydB* are present in *C. acetobutylicum*, and a [NiFe]-hydrogenase which resides in the pSOL1 plasmid (Nolling *et al*., 2001), the *hydA* is the main hydrogenase responsible for hydrogen formation and its function key in the maintenance of redox balance for cell metabolism (King *et al*., 2006; Demuez *et al.*, 2007)

The *crt* and *bcd* genes are involved in the central metabolic pathway which converts acetyl-CoA to butyryl-CoA for butyrate and butanol formation. It has not been possible to knock-out these genes previously with the use of ClosTron system and when the conditional RAM-less ClosTron was applied, the isolation of a pure knock-out mutant could not be achieved indicating that in common

with *hydA*, these genes are essential and cannot be inactivated under these conditions.

The *hbd* gene which could not be knocked out using the standard ClosTron or CRISPR/cas9 system could be knocked out using the conditional RAM-less ClosTron and was characterised. Worthy of note was the production of up to 427 mM ethanol compared with wild type which produced 33 mM. Subsequent NGS analysis revealed a mutation in genes involved in or related to the fatty acid biosynthesis pathway in all seven mutants from four independent transformations. Mutations in genes related to endospore formation, *agr* quorum sensing system, LPS Glycosyltransferase and ribonuclease J were also observed.

At this point, it cannot be concluded that these observed mutations have been compensatory for the *hbd* knock-out because ideally, they have to be replicated on a wildtype background and the knock-out of *hbd* using the standard ClosTron system should then be possible. Though this was included in the original thesis plan, it could not be achieved in this project due to time constraints and is a further work to be done. The identified genes could be deleted individually or in combination using the ACE technology or the CRISPR/Cas9 system and attempt to knock out the *hbd* gene, using the standard ClosTron or CRISPR/Cas9 which was previously not possible, carried out. The ability to identify these compensatory mutations that allow the viability of otherwise lethal mutations will improve understanding of the complex regulation of metabolic processes in *C. acetobutylicum* and inform more rational future genetic engineering towards improved solvent formation.

6.2 Towards Product extension in *Clostridium acetobutylicum* Using a Triple Auxotrophic Mutant and an Orthogonal Expression System

This other aspect of this thesis was aimed at creating a system for introducing pivotal metabolic enzymes at discrete loci around the chromosome, similar to *pyrE*, that may be used to implement new metabolic pathways and to exemplify its utility by extending product formation in *C. acetobutylicum*, to include isopropanol using the inducible *tcdR* system. The idea was to use *tcdR* which would simultaneously control different gene sets, and because they have disparate locations, recombination between the identical promoters is not an issue (as when they are tandemly arranged). If it was not desirable for the different gene sets to be expressed at the same level, there is the option of using the *tcdA* promoter or modifying P_{tcdB} to make it slightly weaker.

Using a suicide KO system, a triple auxotrophic mutant, *C. acetobutylicum* Δ*pyrE*Δ*argH*Δ*pheA* was created and it represented three different loci where gene/operons of interest could be inserted in the chromosome. Based on relative distance of locus from chromosomal origin of replication, gene expression was greatest at the *pyrE* locus followed by the *pheA* locus and then the *argH* locus indicating that genes closest to the origin of replication will have more copies than those that are more distal in an actively growing cell. Accordingly, a lactose inducible *tcdR* system, acetone formation genes and the secondary dehydrogenase from *C. beijerinkii* were integrated at the various loci.

Integration at the *pyrE* and *pheA* loci were relatively easily accomplished compared to the *argH* locus which proved difficult despite several attempts and it was speculated that this could be related to the size of the insert being

introduced or possibly the effect of the integration on downstream genes. Consequently, only the *pyrE* and *pheA* loci were employed for integrations.

The final strain 824BO3 which had *tcdR* and *SadH* integrated and carried the acetone genes on an autonomous plasmid produced 18.47 g/l IBE (8.47 g/l isopropanol, 10.02 g/l butanol, 0.21 g/l ethanol). This combination of the lactose inducible *tcdR* system, chromosomal-integrated and plasmid-based expression seems to have produced the highest recorded production of isopropanol from *C. acetobutylicum* till date. The development of an IBE producing *C. acetobutylicum* strain is particularly valuable as the IBE unlike the ABE mixture can be used directly as fuel doing away with the extra cost required for downstream purification of ABE as the acetone simultaneously produced is a non-fuel component. Furthermore, this strain produced minimal residual acetone (0.27 g/l) making it an ideal potential candidate for large scale IBE production.

In addition, a system has now been put in place for introducing genes at discrete loci around the *C. acetobutylicum* chromosome that may be used to implement new metabolic pathways as desired and which could be further expanded for use in other Clostridia. While the incorporation of operons into autonomous plasmids has the advantage of higher copy number, guaranteeing higher yields of products, there is the problem of segregational instability where not all daughter cells get a copy of the plasmid during cell division. This results in plasmid loss and reduced productivity which constitutes a major industrial problem (Friehs, 2004, Hagg *et al.*, 2004). These plasmids can be stabilised by including a selectable antibiotic resistance gene in their backbone, but the addition of such genes could lead to the spread of antibiotic genes between

bacterial populations. This gives rise to antibiotic resistant bacteria creating environmental and health problems (Hagg *et al*., 2004). Besides, the addition of antibiotics to cultures is expensive and also not desirable. Integration of gene sets into the chromosome overcomes all of these drawbacks, producing strains that are segregationally stable and more suitable for industrial applications.

The original plan was to have at least three difference chromosomal loci for gene integration, however the *argH* locus was shown to be quite problematic. Another possible locus has been identified which could serve as an alternative to *argH,* the *purD* (phosphoribosylamine–glycine ligase). This gene is involved in the *de novo* synthesis of purine nucleotide (Truong *et al*., 2015) and is at the end of the *pur* operon. The *purD* mutant has also been shown to require additional added purine to grow in *Brucella abortus,* (Truong *et al*., 2015) and this screenable phenotype is desirable for this purpose.

Table 6.1: Brief Summary of Major Outputs, Limitations and Future Work of Thesis

Chapter	Outputs	Limitations	Future Work
3	➢ Exemplified the utility of the conditional RAM-less ClosTron in the knock out of the *adhE2* gene in *C. acetobutylicum.* ➢ The *hbd* gene which could not previously be knocked-out using the standard ClosTron or CRISPR/Cas9 system was knocked out using the conditional RAM-less ClosTron, creating hbd69s::rCT. ➢ hbd69s::rCT produced up to 427 mM ethanol compared with wild type (33mM) ➢ Demonstrated the knock out of the *thl* gene using this RAMless ClosTron system. ➢ Described the knock-out of *pyrE* gene in *C. acetobutylicum* using the CRISPR RiboCas system.	➢ It was not possible to isolate knock-out mutants for *hydA*, *crt* and *bcd* genes using the conditional RAMless ClosTron, though evidence of intron insertion was demonstrated in all cases.	➢ Replicate the mutations observed in the hbd69s::CT on a *C. acetobutylicum* wild type background and attempt to knock out *hbd* using the standard ClosTron system ➢ Carrry out solvent profile analysis of the seven mutants obtained and match them with sequencing results.
4	➢ Creation of alternative integration sites to *pyrE* by the construction of an ACE integration system based on *argH and pheA* and the demonstration of gene knock-out and complementation for both genes. ➢ Creation of double (*ΔpyrEΔargH* and *ΔpyrEΔpheA*) and triple auxotrophic (*ΔpyrEΔargHΔpheA)* mutants	Nil	The creation of these alternative loci to *pyrE* was done as an enabling system for integration of genes/operons of interest in the chromosome towards product extension in *C. acetobutylicum*. This was done in Chapter 5.

	in *C. acetobutylicum* and characterisation of genome sequence. ➢ Description of locus specific effects on gene expression using *catP* assay.		
5	➢ Creation of ACE plasmids for the introduction of Lactose inducible *tcdR*, acetone operon and *SadH at the pyrE, argH* and *pheA* loci. ➢ Successful integration of *tcdR* and *SadH* at *pyrE* and *pheA* loci respectively was done. ➢ Creation of four strains - 824BO1, 824BO2, 824BO3 and 824BO4 bearing either *tcdR, SadH* and/or acetone operon integrated in chromosome or on an autonomous plasmid. ➢ Isopropanol could be produced by *C. acetobutylicum* with introduction of *SadH* ➢ Up to 18.47 g/l IBE was produced by 824BO3 with a yield of 0.34 g/g ➢ The isopropanol titer (8.47 g/l) and yield (0.15g/g) obtained from this study looks to be the highest reported to date with residual acetone of 0.27 g/l.	➢ Introduction of acetone operon at *argH* locus was not successful ➢ The elimination of pSOL1 or construction of two plasmid variants (suicide and to include *codA* in backbone) did not result in successful integration of acetone operon at *argH* locus ➢ Swapping of integration loci did not result in successful integration at the *argH* locus	➢ Use the *purD* locus in place of the *argH* locus to achieve a strain with three different chromosomal loci for gene integration to achieve a final strain bearing no plasmid. This strain will represent an industrially viable platform for enhanced solvent production/ product extension in *C. acetobutylicum.*

REFERENCES

REFERENCES

ALBERT COTTON, G. W., CARLOS A. MURILLO, MANFRED BOCHMANN 1999. *Advanced Inorganic Chemistry*, 6th Edition 1376 pp

ALSAKER, K. V. & PAPOUTSAKIS, E. T. 2005. Transcriptional program of early sporulation and stationary-phase events in *Clostridium acetobutylicum*. *J Bacteriol,* 187**,** 7103-7118.

ALSAKER, K. V., PAREDES, C. & PAPOUTSAKIS, E. T. 2010. Metabolite stress and tolerance in the production of biofuels and chemicals: gene-expression-based systems analysis of butanol, butyrate, and acetate stresses in the anaerobe *Clostridium acetobutylicum. Biotechnol Bioeng,* 105**,** 1131-47.

ALSAKER, K. V., SPITZER, T. R. & PAPOUTSAKIS, E. T. 2004. Transcriptional analysis of spo0A overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. *J Bacteriol,* 186**,** 1959-1971.

AN, W. & CHIN, J. W. 2009. Synthesis of orthogonal transcription-translation networks. *Proc Nat Acad Sc,* 106**,** 8477-8482

AN, W. & CHIN, J. W. 2011. Chapter five - Orthogonal Gene Expression in *Escherichia coli*. *Methods in Enzymology.* Academic Press. 497,115-134

ATMADJAJA, A. N., HOLBY, V., HARDING, A. J., KRABBEN, P., SMITH, H. K. & JENKINSON, E. R. 2019. CRISPR-Cas, a highly effective tool for genome editing in *Clostridium saccharoperbutylacetonicum* N1-4(HMT). *FEMS Microbiol Lett,* 366.

ATSUMI, S. 2007. Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab Eng.*

BABU, M., GREENBLATT, J. F., EMILI, A., STRYNADKA, N. C., REITHMEIER, R. A. & MORAES, T. F. 2010. Structure of a SLC26 anion transporter STAS domain in complex with acyl carrier protein: implications for E. coli YchM in fatty acid metabolism. *Structure,* 18**,** 1450-1462.

BAER, S. H., BLASCHEK, H. P. & SMITH, T. L. 1987. Effect of Butanol Challenge and Temperature on Lipid Composition and Membrane Fluidity of Butanol-Tolerant *Clostridium acetobutylicum*. *Appl Environ Microbiol,* 53**,** 2854-61.

BAHL, H., GOTTWALD, M., KUHN, A., RALE, V., ANDERSCH, W. & GOTTSCHALK, G. 1986. Nutritional Factors Affecting the Ratio of Solvents Produced by *Clostridium acetobutylicum*. *Appl Environ Microbiol,* 52**,** 169-172.

BANKAR, S. B., JURGENS, G., SURVASE, S. A., OJAMO, H. & GRANSTRÖM, T. 2015. Genetic engineering of *Clostridium acetobutylicum* to enhance isopropanol–butanol–ethanol production with an integrated DNA-technology approach. *Renew Energy,* 83. 1076–1083

BANKAR, S. B., SURVASE, S. A., OJAMO, H. & GRANSTRÖM, T. 2013. Biobutanol: the outlook of an academic and industrialist. *RSC Adv,* 3. 24734-24757

BHAYA, D., DAVISON, M. & BARRANGOU, R. 2011. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet,* 45**,** 273-97.

BIRGEN, C., DÜRRE, P., PREISIG, H. A. & WENTZEL, A. 2019. Butanol production from lignocellulosic biomass: revisiting fermentation performance indicators with exploratory data analysis. *Biotechnol Biofuels,* 12**,** 167.

BOLOTIN, A., QUINQUIS, B., SOROKIN, A. & EHRLICH, S. D. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology,* 151**,** 2551-61.

BP 2016. BP Statistical Review of World Energy, 2016.

BUCKEL, W. & THAUER, R. K. 2013. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation. *Biochim Biophys Acta,* 1827**,** 94-113.

BUCKEL, W. & THAUER, R. K. 2018. Flavin-Based Electron Bifurcation, Ferredoxin, Flavodoxin, and Anaerobic Respiration With Protons (Ech) or NAD(+) (Rnf) as Electron Acceptors: A Historical Review. *Front Microbiol,* 9**,** 401-401.

CAÑADAS, I. C., GROOTHUIS, D., ZYGOUROPOULOU, M., RODRIGUES, R. & MINTON, N. P. 2019. RiboCas: A universal CRISPR-based editing tool for Clostridium. *ACS Synthetic Biology,* 8**,** 1379-1390.

CARTMAN, S. T. & MINTON, N. P. 2010. A mariner-based transposon system for in vivo random mutagenesis of Clostridium difficile. *Appl Environ Microbiol,* 76**,** 1103-1109.

CHEN, J.-S. & HIU, S. F. 1986. Acetone–butanol–isopropanol production by *Clostridium beijerinckii* (synonym, *Clostridium butylicum*). *Biotech Lett,* 8.

CHENG, C., BAO, T. & YANG, S.-T. 2019. Engineering Clostridium for improved solvent production: recent progress and perspective. *Appl Microbiol and Biotechnol,* 103**,** 5549-5566.

COLLAS, F., KUIT, W., CLEMENT, B., MARCHAL, R., LOPEZ-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**,** 45.

COOKSLEY, C. M., ZHANG, Y., WANG, H., REDL, S., WINZER, K. & MINTON, N. P. 2012. Targeted mutagenesis of the *Clostridium acetobutylicum* acetone-butanol-ethanol fermentation pathway. *Metab Eng,* 14**,** 630-41.

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.

COUSINEAU, B., SMITH, D., LAWRENCE-CAVANAGH, S., MUELLER, J. E., YANG, J., MILLS, D., MANIAS, D., DUNNY, G., LAMBOWITZ, A. M. & BELFORT, M. 1998. Retrohoming of a bacterial group II intron: mobility via complete reverse splicing, independent of homologous DNA recombination. *Cell,* 94**,** 451-62.

COUTURIER, E. & ROCHA, E. P. 2006. Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. *Mol Microbiol,* 59**,** 1506-18.

CRAY, J.A., STEVENSON A., BALL P., BANKAR S.B., ELEUTHERIO

E.C.A., EZEJI T.C., SINGHAL R.S., THEVELEIN J.M., TIMSON D.J. & HALLSWORTH J.E. 2015. Chaotropicity: a key factor in product tolerance of biofuel-producing microorganisms. *Curr Opin Biotechnol*, 33, 228-259.

CROUX, C., NGUYEN, N.-P.-T., LEE, J., RAYNAUD, C., SAINT-PRIX, F., GONZALEZ-PAJUELO, M., MEYNIAL-SALLES, I. & SOUCAILLE, P. 2016. Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. *Biotechnol Biofuels,* 9**,** 23.

DAI, Z., DONG, H., ZHU, Y., ZHANG, Y., LI, Y. & MA, Y. 2012. Introducing a single secondary alcohol dehydrogenase into butanol-tolerant *Clostridium acetobutylicum* Rh8 switches ABE fermentation to high level IBE fermentation. *Biotechnol Biofuels,* 5**,** 44.

DEMUEZ, M., COURNAC, L., GUERRINI, O., SOUCAILLE, P. & GIRBAL, L. 2007. Complete activity profile of *Clostridium acetobutylicum* [FeFe]-hydrogenase and kinetic parameters for endogenous redox partners. *FEMS Microbiol Lett,* 275**,** 113-21.

DESAI, R. P. & PAPOUTSAKIS, E. T. 1999. Antisense RNA strategies for metabolic engineering of *Clostridium acetobutylicum*. *Appl Environ Microbiol,* 65**,** 936-45.

DOENCH, J. G., FUSI, N., SULLENDER, M., HEGDE, M., VAIMBERG, E. W., DONOVAN, K. F., SMITH, I., TOTHOVA, Z., WILEN, C., ORCHARD, R., VIRGIN, H. W., LISTGARTEN, J. & ROOT, D. E. 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol,* 34**,** 184-191.

DOENCH, J. G., HARTENIAN, E., GRAHAM, D. B., TOTHOVA, Z., HEGDE, M., SMITH, I., SULLENDER, M., EBERT, B. L., XAVIER, R. J. & ROOT, D. E. 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol,* 32**,** 1262-7.

DU, Y., JIANG, W., YU, M., TANG, I. C. & YANG, S. T. 2015. Metabolic process engineering of *Clostridium tyrobutyricum* Deltaack-adhE2 for enhanced n-butanol production from glucose: effects of methyl viologen on NADH availability, flux distribution, and fermentation kinetics. *Biotechnol Bioeng,* 112**,** 705-15.

DUNLOP, M. J. 2011. Engineering microbes for tolerance to next-generation biofuels. *Biotechnol Biofuels,* 4**,** 32.

DUPUY, B., RAFFESTIN, S., MATAMOUROS, S., MANI, N., POPOFF, M. R. & SONENSHEIN, A. L. 2006. Regulation of toxin and bacteriocin gene expression in Clostridium by interchangeable RNA polymerase sigma factors. *Mol Microbiol,* 60**,** 1044-57.

DURRE, P. 2007. Biobutanol: an attractive biofuel. *Biotechnol J,* 2**,** 1525-34.

DURRE, P. 2011. Ancestral sporulation initiation. *Mol Microbiol,* 80**,** 584-7.

DUSSEAUX, S., CROUX, C., SOUCAILLE, P. & MEYNIAL-SALLES, I. 2013. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for the high-yield production of a biofuel composed of an isopropanol/butanol/ethanol mixture. *Metab Eng,* 18**,** 1-8.

DUSSÉAUX, S., CROUX, C., SOUCAILLE, P. & MEYNIAL-SALLES, I. 2013. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for the high-yield production of a biofuel composed of an isopropanol/butanol/ethanol mixture. *Metab Eng,* 18.

DYRDA, G., BONIEWSKA-BERNACKA, E., MAN, D., BARCHIEWICZ, K. & SŁOTA, R. 2019. The effect of organic solvents on selected microorganisms and model liposome membrane. *Mol Biol Rep,* 46**,** 3225-3232.

EHSAAN, M., KUIT, W., ZHANG, Y., CARTMAN, S. T., HEAP, J. T., WINZER, K. & MINTON, N. P. 2016. Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. *Biotechnol Biofuels,* 9**,** 4.

ELLISON, M. J., KELLEHER, R. J., 3RD & RICH, A. 1985. Thermal regulation of beta-galactosidase synthesis using anti-sense RNA directed against the coding portion of the mRNA. *J Biol Chem,* 260**,** 9085-7.

EUROSTAT 2019. Eurostat Statistics Explained: Renewable Energy Statistics. *Online Publication, Data extracted January 2019.*

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**,** 1697-712.

EZEJI, T., QURESHI, N. & BLASCHEK, H. P. 2007. Production of acetone–butanol–ethanol (ABE) in a continuous flow bioreactor using degermed corn and *Clostridium beijerinckii*. *Process Biochemistry,* 42**,** 34-39.

FAWCETT, P., EICHENBERGER, P., LOSICK, R. & YOUNGMAN, P. 2000. The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A,* 97**,** 8063-8.

FONTAINE, L., MEYNIAL-SALLES, I., GIRBAL, L., YANG, X., CROUX, C. & SOUCAILLE, P. 2002. Molecular characterization and transcriptional analysis of adhE2, the gene encoding the NADH-dependent aldehyde/alcohol dehydrogenase responsible for butanol production in alcohologenic cultures of *Clostridium acetobutylicum* ATCC 824. *J bacteriol,* 184**,** 821-830.

FOULQUIER, C., HUANG, C.-N., NGUYEN, N.-P.-T., THIEL, A., WILDING-STEEL, T., SOULA, J., YOO, M., EHRENREICH, A., MEYNIAL-SALLES, I., LIEBL, W. & SOUCAILLE, P. 2019. An efficient method for markerless mutant generation by allelic exchange in *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* using suicide vectors. *Biotechnol biofuels,* 12**,** 31-31.

FRIEHS, K. 2004. Plasmid copy number and plasmid stability. *Adv Biochem Eng Biotechnol,* 86**,** 47-82.

GEORGE, H. A., JOHNSON, J. L., MOORE, W. E., HOLDEMAN, L. V. & CHEN, J. S. 1983. Acetone, isopropanol, and butanol production by *Clostridium beijerinckii (syn. Clostridium butylicum*) and *Clostridium aurantibutyricum*. *Appl Environ Microbiol,* 45.

GHESHLAGHI, R., SCHARER, J. M., MOO-YOUNG, M. & CHOU, C. P. 2009. Metabolic pathways of clostridia for producing butanol. *Biotechnol Adv,* 27**,** 764-781.

GIRBAL, L., CROUX, C., VASCONCELOS, I. & SOUCAILLE, P. 1995. Regulation of metabolic shifts in *Clostridium acetobutylicum* ATCC 824. *FEMS Microbiology Reviews,* 17**,** 287-297.

GONZALES, T. & ROBERT-BAUDOUY, J. 1996. Bacterial aminopeptidases: Properties and functions. *FEMS Microbiology Reviews,* 18**,** 319-344.

GONZALEZ-PAJUELO, M., MEYNIAL-SALLES, I., MENDES, F., ANDRADE, J. C., VASCONCELOS, I. & SOUCAILLE, P. 2005.

Metabolic engineering of *Clostridium acetobutylicum* for the industrial production of 1,3-propanediol from glycerol. *Metab Eng,* 7**,** 329-36.

GREEN, E. M., BOYNTON, Z. L., HARRIS, L. M., RUDOLPH, F. B., PAPOUTSAKIS, E. T. & BENNETT, G. N. 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. *Microbiol,* 142.

GRIMMLER, C., JANSSEN, H., KRAUSSE, D., FISCHER, R. J., BAHL, H., DURRE, 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.

HAGG, P., DE POHL, J. W., ABDULKARIM, F. & ISAKSSON, L. A. 2004. A host/plasmid system that is not dependent on antibiotics and antibiotic resistance genes for stable plasmid maintenance in *Escherichia coli. J Biotechnol,* 111**,** 17-30.

HARRIS, L. M., BLANK, L., DESAI, R. P., WELKER, N. E. & PAPOUTSAKIS, E. T. 2001. Fermentation characterization and flux analysis of recombinant strains of *Clostridium acetobutylicum* with an inactivated solR gene. *J Ind Microbiol Biotechnol,* 27**,** 322-8.

HARRIS, L. M., DESAI, R. P., WELKER, N. E. & PAPOUTSAKIS, E. T. 2000. Characterization of recombinant strains of the *Clostridium acetobutylicum* butyrate kinase inactivation mutant: need for new phenomenological models for solventogenesis and butanol inhibition? *Biotechnol Bioeng,* 67.

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. *J Bacteriol,* 184**,** 3586-97.

HARTMANIS, M. G. & GATENBECK, S. 1984. Intermediary Metabolism in *Clostridium acetobutylicum:* Levels of Enzymes Involved in the Formation of Acetate and Butyrate. *App Environ Microbiol,* 47**,** 1277-1283.

HARTMANIS, M. G. N., KLASON, T. & GATENBECK, S. 1984. Uptake and activation of acetate and butyrate in *Clostridium acetobutylicum*. *App Microbiol Biotechnol,* 20**,** 66-71.

HEAP, J. T., EHSAAN, M., COOKSLEY, C. M., NG, Y.-K., CARTMAN, S. T., WINZER, K. & MINTON, N. P. 2012a. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucl Acids Res,* 40**,** e59-e59.

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. *J Microbiol 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.

HEAP, J. T., PENNINGTON, O. J., CARTMAN, S. T. & MINTON, N. P. 2009. A modular system for Clostridium shuttle plasmids. *J Microbiol Methods,* 8.

HIGUCHI, R., KRUMMEL, B. & SAIKI, R. K. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res,* 16**,** 7351-67.

HSU, C. S. R., PAUL R. 2017. Handbook of Petroleum Technology (2nd, illustrated ed.). Springer. p. 3-6.

HSU, P. D., LANDER, E. S. & ZHANG, F. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell,* 157**,** 1262-78.

HUFFER, S., CLARK, M. E., NING, J. C., BLANCH, H. W. & CLARK, D. S. 2011. Role of Alcohols in Growth, Lipid Composition, and Membrane Fluidity of Yeasts, Bacteria, and Archaea. *App Environl Microbiol,* 77**,** 6400.

IMARC Market Research Report, (2019) **"**Isopropyl Alcohol Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2019-2024**"** Published February 11, 2019

JANATI-IDRISSI, R., JUNELLES, A. M., EL KANOUNI, A., PETITDEMANGE, H. & GAY, R. 1987. Sélection de mutants de

Clostridium acetobutylicum défectifs dans la production d'acétone. *Annales de l'Institut Pasteur / Microbiologie,* 138**,** 313-323.

JANG, Y.-S., IM, J. A., CHOI, S. Y., LEE, J. I. & LEE, S. Y. 2014a. Metabolic engineering of *Clostridium acetobutylicum* for butyric acid production with high butyric acid selectivity. *Metabol Eng,* 23**,** 165-174.

JANG, Y.-S., LEE, J. Y., LEE, J., PARK, J. H., IM, J. A., EOM, M.-H., LEE, J., LEE, S.-H., SONG, H., CHO, J.-H., SEUNG, D. Y. & LEE, S. Y. 2012a. Enhanced Butanol Production Obtained by Reinforcing the Direct Butanol-Forming Route in *Clostridium acetobutylicum mBio,* 3**,** e00314-12.

JANG, Y. S., LEE, J. Y., LEE, J., PARK, J. H., IM, J. A., EOM, M. H., LEE, S. H., SONG, H., CHO, J. H., SEUNG DO, Y. & LEE, S. Y. 2012b. Enhanced butanol production obtained by reinforcing the direct butanol-forming route in *Clostridium acetobutylicum*. *MBio,* 3.

JANG, Y. S., MALAVIYA, A., LEE, J., IM, J. A., LEE, S. Y., LEE, J., EOM, M. H., CHO, J. H. & SEUNG DO, Y. 2013. Metabolic engineering of Clostridium acetobutylicum for the enhanced production of isopropanol-butanol-ethanol fuel mixture. *Biotechnol Prog,* 29**,** 1083-8.

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**,** 284-91.

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.

KAZEMI SHARIAT PANAHI, H., DEHHAGHI, M., KINDER, J. E. & EZEJI, T. C. 2019. A review on green liquid fuels for the transportation sector: a prospect of microbial solutions to climate change. *Biofuel Res J,* 6**,** 995-1024.

KHANDEKAR, S. S., GENTRY, D. R., VAN ALLER, G. S., WARREN, P., XIANG, H., SILVERMAN, C., DOYLE, M. L., CHAMBERS, P. A.,

KONSTANTINIDIS, A. K., BRANDT, M., DAINES, R. A. & LONSDALE, J. T. 2001. Identification, substrate specificity, and inhibition of the *Streptococcus pneumoniae* beta-ketoacyl-acyl carrier protein synthase III (FabH). *J Biol Chem,* 276**,** 30024-30.

KIM, B. H., BELLOWS, P., DATTA, R. & ZEIKUS, J. G. 1984. Control of Carbon and Electron Flow in Clostridium acetobutylicum Fermentations: Utilization of Carbon Monoxide to Inhibit Hydrogen Production and to Enhance Butanol Yields. *Appl and Environ Microbiol,* 48**,** 764-770.

KING, P. W., POSEWITZ, M. C., GHIRARDI, M. L. & SEIBERT, M. 2006. Functional studies of [FeFe] hydrogenase maturation in an *Escherichia coli* biosynthetic system. *J Bacteriol,* 188**,** 2163-72.

KLAUS WINZER, K. L. A. P. D. 1997. Acetate kinase from *CIostridium acetobutylicum* : a highly specific enzyme that is actively transcribed during acidogenesis and solventogenesis. *Microbiology,* 143**,** 3279-3286

KOLESINSKA, B., FRACZYK, J., BINCZARSKI, M., MODELSKA, M., BERLOWSKA, J., DZIUGAN, P., ANTOLAK, H., KAMINSKI, Z. J., WITONSKA, I. A. & KREGIEL, D. 2019. Butanol Synthesis Routes for Biofuel Production: Trends and Perspectives. *Materials (Basel),* 12.

KUEHNE, S. A. & MINTON, N. P. 2012. ClosTron-mediated engineering of Clostridium. *Bioengineered,* 3(4):247-54

KUIT, W., MINTON, N. P., LOPEZ-CONTRERAS, A. M. & EGGINK, G. 2012. Disruption of the acetate kinase (ack) gene of *Clostridium acetobutylicum* results in delayed acetate production. *Appl Microbiol Biotechnol,* 94**,** 729-41.

LEANG, C., UEKI, T., NEVIN, K. P. & LOVLEY, D. R. 2013. A genetic system for Clostridium ljungdahlii: a chassis for autotrophic production of biocommodities and a model homoacetogen. *Appl Environ Microbiol,* 79**,** 1102-9.

LEE, J., JANG, Y. S., CHOI, S. J., IM, J. A., SONG, H., CHO, J. H., SEUNG DO, Y., PAPOUTSAKIS, E. T., BENNETT, G. N. & LEE, S. Y. 2012. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for isopropanol-butanol-ethanol fermentation. *Appl Environ Microbiol,* 78**,** 1416-23.

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.

LEHMANN, D. & LUTKE-EVERSLOH, T. 2011. Switching *Clostridium acetobutylicum* to an ethanol producer by disruption of the butyrate/butanol fermentative pathway. *Metab Eng,* 13**,** 464-73.

LI, F., HINDERBERGER, J., SEEDORF, H., ZHANG, J., BUCKEL, W. & THAUER, R. K. 2008. Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri*. *J Bacteriol,* 190**,** 843-850.

LI, G., LEE, T. H., LIU, Z., LEE, C. F. & ZHANG, C. 2019. Effects of injection strategies on combustion and emission characteristics of a common-rail diesel engine fueled with isopropanol-butanol-ethanol and diesel blends. *Ren Ener,* 130**,** 677-686.

LI, Q., CHEN, J., MINTON, N. P., ZHANG, Y., WEN, Z., LIU, J., YANG, H., ZENG, Z., REN, X., YANG, J., GU, Y., JIANG, W., JIANG, Y. & YANG, S. 2016. CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. *Biotechnol J,* 11**,** 961-72.

LI Y., TSCHAPLINSKI T.J., ENGLE N.L., HAMILTON C.Y., RODRIGUEZ M.J.R., LIAO J.C., SCHADT C.W., GUSS A.M., YANG Y. & GRAHAM D. E. 2012. Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenases genes substantially increases ethanol yield from cellulose and switchgrass fermentations. *Biotechnol Biofuels,* 5(1), 2

LIAO, C., SEO, S. O., CELIK, V., LIU, H., KONG, W., WANG, Y., BLASCHEK, H., JIN, Y. S. & LU, T. 2015. Integrated, systems metabolic picture of acetone-butanol-ethanol fermentation by *Clostridium acetobutylicum*. *Proc Natl Acad Sci U S A,* 112**,** 8505-10.

LIU, H., HUANG, D. & WEN, J. 2016. Integrated intracellular metabolic profiling and pathway analysis approaches reveal complex metabolic regulation by *Clostridium acetobutylicum*. *Microbial Cell Fact,* 15**,** 36.

LU, C., YU, L., VARGHESE, S., YU, M. & YANG, S. T. 2017. Enhanced robustness in acetone-butanol-ethanol fermentation with engineered *Clostridium beijerinckii* overexpressing adhE2 and ctfAB. *Bioresour Technol,* 243**,** 1000-1008.

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. *Bioresour Technol,* 104**,** 380-7.

LÜTKE-EVERSLOH, T. & BAHL, H. 2011. Metabolic engineering of *Clostridium acetobutylicum*: recent advances to improve butanol production. *Curr Opin in Biotechnol,* 22**,** 634-647.

MADHUGIRI, R. & EVGUENIEVA-HACKENBERG, E. 2009. RNase J is involved in the 5'-end maturation of 16S rRNA and 23S rRNA in Sinorhizobium meliloti. *FEBS Lett,* 583**,** 2339-42.

MAISNIER-PATIN, S. & ANDERSSON, D. I. 2004. Adaptation to the deleterious effects of antimicrobial drug resistance mutations by compensatory evolution. *Res Microbiol,* 155**,** 360-9.

MAKAROVA, K. S., HAFT, D. H., BARRANGOU, R., BROUNS, S. J., CHARPENTIER, E., HORVATH, P., MOINEAU, S., MOJICA, F. J., WOLF, Y. I., YAKUNIN, A. F., VAN DER OOST, J. & KOONIN, E. V. 2011. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol,* 9**,** 467-77.

MANN, M. S., DRAGOVIC, Z., SCHIRRMACHER, G. & LUTKE-EVERSLOH, T. 2012. Over-expression of stress protein-encoding genes helps *Clostridium acetobutylicum* to rapidly adapt to butanol stress. *Biotechnol Lett,* 34**,** 1643-9.

MANN, M. S. & LÜTKE-EVERSLOH, T. 2013. Thiolase engineering for enhanced butanol production in *Clostridium acetobutylicum. Biotechnol Bioeng,* 110**,** 887-897.

MARCHLER-BAUER, A., BO, Y., HAN, L., HE, J., LANCZYCKI, C. J., LU, S., CHITSAZ, F., DERBYSHIRE, M. K., GEER, R. C., GONZALES, N. R., GWADZ, M., HURWITZ, D. I., LU, F., MARCHLER, G. H., SONG, J. S., THANKI, N., WANG, Z., YAMASHITA, R. A., ZHANG, D., ZHENG, C., GEER, L. Y. & BRYANT, S. H. 2017.

CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res,* 45**,** D200-d203.

MERMELSTEIN, L. D. & PAPOUTSAKIS, E. T. 1993. In vivo methylation in Escherichia coli by the Bacillus subtilis phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol,* 59**,** 1077-81.

MERMELSTEIN, L. D., PAPOUTSAKIS, E. T., PETERSEN, D. J. & BENNETT, G. N. 1993. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for increased solvent production by enhancement of acetone formation enzyme activities using a synthetic acetone operon. *Biotechnol Bioeng,* 42**,** 1053-1060.

MERMELSTEIN, L. D., WELKER, N. E., BENNETT, G. N. & PAPOUTSAKIS, E. T. 1992. Expression of Cloned Homologous Fermentative Genes in *Clostridium Acetobutylicum* ATCC 824. *Bio/Technology,* 10**,** 190-195.

MINTON, N. P., BREHM, J. K., SWINFIELD, T.J., WHELAN, S. M., MAUCHLINE, M.L., BODS-WORTH, N. & OULTRAM, J. D. 1993. Clostridial cloning vectors. In: Woods DR, editor. *The clostridia and Biotechnology*. Stoneham: Butterworth-Heinemann. . p 119-150.

MINTON, N. P., EHSAAN, M., HUMPHREYS, C. M., LITTLE, G. T., BAKER, J., HENSTRA, A. M., LIEW, F., KELLY, M. L., SHENG, L., SCHWARZ, K. & ZHANG, Y. 2016. A roadmap for gene system development in Clostridium. *Anaerobe,* 41**,** 104-112.

MOHR, G., HONG, W., ZHANG, J., CUI, G. Z., YANG, Y., CUI, Q., LIU, Y. J. & LAMBOWITZ, A. M. 2013. A targetron system for gene targeting in thermophiles and its application in *Clostridium thermocellum. PLoS One,* 8**,** e69032.

MOJICA, F. J., DIEZ-VILLASENOR, C., GARCIA-MARTINEZ, J. & ALMENDROS, C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology,* 155**,** 733-40.

MOON, B. K. 2016. United Nations, The Paris Agreement, accessed April 25, 2016.

MOON, H. G., JANG, Y. S., CHO, C., LEE, J., BINKLEY, R. & LEE, S. Y. 2016. One hundred years of clostridial butanol fermentation. *FEMS Microbiol Lett,* 363.

NAIR, R. V., GREEN, E. M., WATSON, D. E., BENNETT, G. N. & PAPOUTSAKIS, E. T. 1999. Regulation of the *sol* locus genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 by a putative transcriptional repressor. *J bacteriol,* 181**,** 319-330.

NCBI 2018. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res,* 46**,** D8-d13.

NELMS, J., EDWARDS, R. M., WARWICK, J. & FOTHERINGHAM, I. 1992. Novel mutations in the pheA gene of *Escherichia coli* K-12 which result in highly feedback inhibition-resistant variants of chorismate mutase/prephenate dehydratase. *Appl Environ Microbiol,* 58**,** 2592-8.

NG, Y. K., EHSAAN, M., PHILIP, S., COLLERY, M. M., JANOIR, C., COLLIGNON, A., CARTMAN, S. T. & MINTON, N. P. 2013. Expanding the repertoire of gene tools for precise manipulation of the Clostridium difficile genome: allelic exchange using pyrE alleles. *PLoS One,* 8(2):e56051.

NGUYEN, N.-P.-T., RAYNAUD, C., MEYNIAL-SALLES, I. & SOUCAILLE, P. 2018. Reviving the Weizmann process for commercial n-butanol production. *Nat Comm,* 9**,** 3682.

NGUYEN, N. P. T. 2016. Metabolic engineering of clostridium acetobutylicum for the production of fuels and chemicals. *Microbiology and Parasitology, INSA de Toulouse,* English. ffNNT : 2016ISAT0051.

NICOLAOU, S. A., GAIDA, S. M. & PAPOUTSAKIS, E. T. 2010. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation. *Metab Eng,* 12**,** 307-331.

NIELSEN, D. R., LEONARD, E., YOON, S.-H., TSENG, H.-C., YUAN, C. & PRATHER, K. L. J. 2009. Engineering alternative butanol production platforms in heterologous bacteria. *Met Eng,* 11**,** 262-273.

NOLLING, J., BRETON, G., OMELCHENKO, M. V., MAKAROVA, K. S., ZENG, Q. D., GIBSON, R., LEE, H. M., DUBOIS, J., QIU, D. Y., HITTI, J., WOLF, Y. I., TATUSOV, R. L., SABATHE, F.,

DOUCETTE-STAMM, L., SOUCAILLE, P., DALY, M. J., BENNETT, G. N., KOONIN, E. V., SMITH, D. R. & FINISHING, G. S. C. P. 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol,* 183.

NOVICK, R. P. & GEISINGER, E. 2008. Quorum sensing in staphylococci. *Annu Rev Genet,* 42**,** 541-64.

O'BRIEN, R. W. & MORRIS, J. G. 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *J Gen Microbiol,* 68**,** 307-18.

PAPOUTSAKIS, E. T. 2015. Reassessing the progress in the production of advanced biofuels in the current competitive environment and beyond: what are the successes and where progress eludes us and why. *Industr Eng Chem Res,* 54**,** 10170-10182.

PERALTA-YAHYA, P. P. & KEASLING, J. D. 2010. Advanced biofuel production in microbes. *Biotechnol J,* 5**,** 147-62.

PETITDEMANGE, H., CHERRIER, C., BENGONE, J. M. & GAY, R. 1977. Study of the NADH and NADPH-ferredoxin oxidoreductase activities in *Clostridium acetobutylicum*. *Can J Microbiol,* 23**,** 152-60.

PURDY, D., O'KEEFEE,T.A.T.,ELMORE,M., HERBERT, M., MCLEOD, A., BOKORI-BROWN, M.,OSTROWSKI, A. & MINTON, N.P. 2002. Conjugative transfer of Clostridia shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier *Mol Microbiol. ,* 46, 439-452.

PYNE, M. E., BRUDER, M. R., MOO-YOUNG, M., CHUNG, D. A. & CHOU, C. P. 2016. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in Clostridium. *Sci Rep,* 6**,** 25666.

QUECK, S. Y., JAMESON-LEE, M., VILLARUZ, A. E., BACH, T. H., KHAN, B. A., STURDEVANT, D. E., RICKLEFS, S. M., LI, M. & OTTO, M. 2008. RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus. Mol Cell,* 32**,** 150-8.

QURESHI, N., SAHA, B. C., DIEN, B., HECTOR, R. E. & COTTA, M. A. 2010. Production of butanol (a biofuel) from agricultural residues: Part I – Use of barley straw hydrolysate. *Biom Bioener,* 34**,** 559-565.

RASSADIN V., SHLYGIN O., LIKHTEROVA N., SLAVIN V. & ZHAROV A. 2006 Problems in production of high-octane, unleaded automotive gasolines. *Chem Tech Fuels Oil* 42, 235-242.

REN 2019. REN 21 Renewables 2019 Global Status Report.

REN, C., GU, Y., HU, S., WU, Y., WANG, P., YANG, Y., YANG, C., YANG, S. & JIANG, W. 2010. Identification and inactivation of pleiotropic regulator CcpA to eliminate glucose repression of xylose utilization in *Clostridium acetobutylicum*. *Met Eng,* 12**,** 446-454.

RODRÍGUEZ-FERNÁNDEZ, J., HERNÁNDEZ, J. J., CALLE-ASENSIO, A., RAMOS, Á. & BARBA, J. 2019. Selection of blends of diesel fuel and advanced biofuels based on their physical and thermochemical properties. *Energies,* 12, 2034

ROJAS ECHENIQUE, J. I., KRYAZHIMSKIY, S., NGUYEN BA, A. N. & DESAI, M. M. 2019. Modular epistasis and the compensatory evolution of gene deletion mutants. *PLoS Genet,* 15**,** e1007958.

SABERI, F., KAMALI, M., NAJAFI, A., YAZDANPARAST, A. & MOGHADDAM, M. M. 2016. Natural antisense RNAs as mRNA regulatory elements in bacteria: a review on function and applications. *Cell Mol Biol Lett,* 21**,** 6.

SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. 1989. *Molecular Cloning: a Laboratory Manual,* Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

SAUER, C., SYVERTSSON, S., BOHORQUEZ, L. C., CRUZ, R., HARWOOD, C. R., VAN RIJ, T. & HAMOEN, L. W. 2016. Effect of genome position on heterologous gene expression in *Bacillus subtilis:* an unbiased analysis. *ACS Synth Biol,* 5**,** 942-7.

SAUER, M. 2016. Industrial production of acetone and butanol by fermentation-100 years later. *FEMS microbiology letters,* 363**,** fnw134.

SCHWARZ, K. M., GROSSE-HONEBRINK, A., DERECKA, K., ROTTA, C., ZHANG, Y. & MINTON, N. P. 2017. Towards improved butanol

production through targeted genetic modification of Clostridium pasteurianum. *Met Eng,* 40**,** 124-137.

SCOTCHER, M. C. & BENNETT, G. N. 2005. SpoIIE regulates sporulation but does not directly affect solventogenesis in *Clostridium acetobutylicum* ATCC 824. *J Bacteriol,* 187**,** 1930-6.

SEEDORF, H., FRICKE, W. F., VEITH, B., BRUGGEMANN, H., LIESEGANG, H., STRITTMATTER, A., MIETHKE, M., BUCKEL, W., HINDERBERGER, J., LI, F., HAGEMEIER, C., THAUER, R. K. & GOTTSCHALK, G. 2008. The genome of *Clostridium kluyveri,* a strict anaerobe with unique metabolic features. *Proc Natl Acad Sci U S A,* 105**,** 2128-33.

SELLE, K. & BARRANGOU, R. 2015. Harnessing CRISPR-Cas systems for bacterial genome editing. *Trends Microbiol,* 23**,** 225-32.

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**,** 115-24.

SHAO, L., HU, S., YANG, Y., GU, Y., CHEN, J., YANG, Y., JIANG, W. & YANG, S. 2007. Targeted gene disruption by use of a group II intron (targetron) vector in. *Cell Research,* 17**,** 963.

SHEN, C. R. & LIAO, J. C. 2008. Metabolic engineering of Escherichia coli for 1-butanol and 1-propanol production via the keto-acid pathways. *Met Eng,* 10**,** 312-320.

SIEMERINK, M. A., KUIT, W., LOPEZ CONTRERAS, A. M., EGGINK, G., VAN DER OOST, J. & KENGEN, S. W. 2011. D-2,3-butanediol production due to heterologous expression of an acetoin reductase in *Clostridium acetobutylicum*. *Appl Environ Microbiol,* 77**,** 2582-8.

SILLERS, R., AL-HINAI, M. A. & PAPOUTSAKIS, E. T. 2009. Aldehyde-alcohol dehydrogenase and/or thiolase overexpression coupled with CoA transferase downregulation lead to higher alcohol titers and selectivity in *Clostridium acetobutylicum* fermentations. *Biotechnol Bioeng,* 102**,** 38-49.

SOUSA, C., DE LORENZO, V. & CEBOLLA, A. 1997. Modulation of gene expression through chromosomal positioning in *Escherichia coli*. *Microbiology,* 143(6)**,** 2071-8.

STEEN, E. J., CHAN, R., PRASAD, N., MYERS, S., PETZOLD, C. J., REDDING, A., OUELLET, M. & KEASLING, J. D. 2008. Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microb Cell Fact,* 7**,** 36.

STEINER, E., DAGO, A. E., YOUNG, D. I., HEAP, J. T., MINTON, N. P., HOCH, J. A. & YOUNG, M. 2011. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. *Mol Microbiol,* 80**,** 641-654.

STEINER, E. J., SCOTT, J. C., MINTON, N. P. & WINZER, K. 2012. An agr quorum sensing system that regulates granulose formation and sporulation in *Clostridium acetobutylicum*. *Appl Environ Microbiol,* 78.

SURVASE, S. A., JURGENS, G., HEININGEN, A. & GRANSTRÖM, T. 2011. Continuous production of isopropanol and butanol using *Clostridium beijerinckii* DSM 6423. *Appl Microbiol Biotechnol,* 91.

TAYLOR, A. 1993. Aminopeptidases: structure and function. *Faseb j,* 7**,** 290-8.

TEMME, K., HILL, R., SEGALL-SHAPIRO, T. H., MOSER, F. & VOIGT, C. A. 2012. Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nucleic Acids Res,* 40**,** 8773-81.

TOMAS, C. A., BEAMISH, J. & PAPOUTSAKIS, E. T. 2004. Transcriptional analysis of butanol stress and tolerance in Clostridium acetobutylicum. *J Bacteriol,* 186**,** 2006-2018.

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. *App Env Microbiology,* 69**,** 4951-4965.

TRUONG, Q. L., CHO, Y., BARATE, A. K., KIM, S., WATARAI, M. & HAHN, T. W. 2015. Mutation of purD and purF genes further attenuates *Brucella abortus* strain RB51. *Microb Pathog,* 79**,** 1-7.

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. *J Bacteriol,* 185**,** 3644-53.

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*. *J Bacteriol,* 185**,** 1923-34.

VANDECASTEELE, R. M. B. P. 1985. Industrial optimization of acetone-butanol fermentation: a study of the utilization of Jerusalem artichokes. *App Microbiol Biotechnol,* 23**,** 92–98.

VENTURA, J. R., HU, H. & JAHNG, D. 2013. Enhanced butanol production in Clostridium acetobutylicum ATCC 824 by double overexpression of 6-phosphofructokinase and pyruvate kinase genes. *Appl Microbiol Biotechnol,* 97**,** 7505-16.

WANG, C., XIN, F., KONG, X., ZHAO, J., DONG, W., ZHANG, W., MA, J., WU, H. & JIANG, M. 2018. Enhanced isopropanol-butanol-ethanol mixture production through manipulation of intracellular NAD(P)H level in the recombinant *Clostridium acetobutylicum* XY16. *Biotechnol Biofuels,* 11**,** 12.

WANG, H., 2014. The Acetone-Butanol-Ethanol Pathway of *Clostridium Acetobutylicum*, PhD Thesis. University of Nottingham, Nottingham, UK

WANG, Y., GUO, W., CHENG, C. L., HO, S. H., CHANG, J. S. & REN, N. 2016. Enhancing bio-butanol production from biomass of *Chlorella vulgaris* JSC-6 with sequential alkali pretreatment and acid hydrolysis. *Bioresour Technol,* 200**,** 557-64.

WANG, Y. LI, X., MILNE, C.B., JANSSEN, H., LIN, W., PHAN, G., HU, G., JIN, Y., PRICE, N.D. & BLASCHEK, H.P. 2013 Using mobile group II introns (targetron) and genetic disruption of acid production pathways in *Clostridium beijerinkii. Appl Environ Microbiol,* 79, 5853–5863

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**,** 5853-63.

WANG, Y., ZHANG, Z. T., SEO, S. O., CHOI, K., LU, T., JIN, Y. S. & BLASCHEK, H. P. 2015. Markerless chromosomal gene deletion in

Clostridium beijerinckii using CRISPR/Cas9 system. *J Biotechnol,* 200**,** 1-5.

WASELS, F., JEAN-MARIE, J., COLLAS, F., LÓPEZ-CONTRERAS, A. M. & LOPES FERREIRA, N. 2017. A two-plasmid inducible CRISPR/Cas9 genome editing tool for *Clostridium acetobutylicum*. *J Microbioll Methods,* 140**,** 5-11.

WIETZKE, M. & BAHL, H. 2012. The redox-sensing protein Rex, a transcriptional regulator of solventogenesis in *Clostridium acetobutylicum*. *Appl Microbiol Biotechnol,* 96**,** 749-61.

WILKINSON, S. R. & YOUNG, M. 1994. Targeted gene integration of genes into the *Clostridium acetobutylicum* chromosome. *Microbiol,* 140. 40(1), 89-95

WISCHRAL, D., ZHANG, J., CHENG, C., LIN, M., DE SOUZA, L. M. G., PESSOA, F. L. P., PEREIRA, N., JR. & YANG, S. T. 2016. Production of 1,3-propanediol by *Clostridium beijerinckii* DSM 791 from crude glycerol and corn steep liquor: Process optimization and metabolic engineering. *Bioresour Technol,* 212**,** 100-110.

XU, M., ZHAO, J., YU, L., TANG, I. C., XUE, C. & YANG, S. T. 2015. Engineering *Clostridium acetobutylicum* with a histidine kinase knockout for enhanced n-butanol tolerance and production. *Appl Microbiol Biotechnol,* 99**,** 1011-22.

XUE, C., ZHAO, J., CHEN, L., YANG, S. T. & BAI, F. 2017. Recent advances and state-of-the-art strategies in strain and process engineering for biobutanol production by *Clostridium acetobutylicum*. *Biotechnol Adv,* 35**,** 310-322.

YAMAGISHI, A., TANIMOTO, T., SUZUKI, T. & OSHIMA, T. 1996. Pyrimidine biosynthesis genes (pyrE and pyrF) of an extreme thermophile, Thermus thermophilus. *App Environ Microbiol,* 62**,** 2191.

YOO, M., BESTEL-CORRE, G., CROUX, C., RIVIERE, A., MEYNIAL-SALLES, I. & SOUCAILLE, P. 2015. A Quantitative System-Scale Characterization of the Metabolism of *Clostridium acetobutylicum mBio,* 6**,** e01808-15.

YU, L., XU, M., TANG, I. C. & YANG, S. T. 2015. Metabolic engineering of *Clostridium tyrobutyricum* for n-butanol production through co-utilization of glucose and xylose. *Biotechnol Bioeng,* 112**,** 2134-41.

YU, M., ZHANG, Y., TANG, I. C. & YANG, S. T. 2011. Metabolic engineering of *Clostridium tyrobutyricum* for n-butanol production. *Metab Eng,* 13**,** 373-82.

ZHANG, C., LI, T. & HE, J. 2018. Characterization and genome analysis of a butanol–isopropanol-producing *Clostridium beijerinckii* strain BGS1. *Biotechnol Biofuels,* 11**,** 280.

ZHANG, J., YU, L., LIN, M., YAN, Q. & YANG, S. T. 2017. n-Butanol production from sucrose and sugarcane juice by engineered *Clostridium tyrobutyricum* overexpressing sucrose catabolism genes and adhE2. *Bioresour Technol,* 233**,** 51-57.

ZHANG, X., AGRAWAL, A. & SAN, K. Y. 2012. Improving fatty acid production in *Escherichia coli* through the overexpression of malonyl coA-acyl carrier protein transacylase. *Biotechnol Prog,* 28**,** 60-65.

ZHANG, Y., GROSSE-HONEBRINK, A. & MINTON, N. P. 2015. A universal mariner transposon system for forward genetic studies in the genus Clostridium. *PLoS One,* 10**,** e0122411.

ZHANG L., NIE X., RAVCHEEV D.A., RODIONOV, D.A., SHENG, J., GU, Y., YANG, S., JIANG, W. & YANG, C. 2014. Redox-responsive repressor Rex modulates alcohol production and oxidative stress tolerance in *Clostridium acetobutylicum. J Bacteriol,* 196(22), 3949-3963.

ZHANG, Y., XU, S., CHAI, C., YANG, S., JIANG, W., MINTON, N. P. & GU, Y. 2016. Development of an inducible transposon system for efficient random mutagenesis in *Clostridium acetobutylicum. FEMS Microbiol Lett,* 363

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. *Appl Environ Microbiol,* 69**,** 2831-41.

ZYGOUROPOULOU, M., 2019. Advancing the Clinical Implementation of Clostridial Cancer Therapies, PhD Thesis. University of Nottingham, Nottingham, UK