Role of RNPP and Agr quorum sensing in the regulation of stationary phase phenotypes in *Clostridium acetobutylicum*

Zak Bean, B. SC

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

September 2016

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.

Zak Bean

September 2016

Abstract

The Gram-positive, endospore forming, strictly anaerobic bacterium *Clostridium aceto-butylicum* is well known for its ability to ferment sugars to organic acids and solvents, most notably butanol. In batch culture, following an initial phase of acid production, the organism undergoes a metabolic and physiological shift to solvent production, granulose accumulation and initiation of endospore formation, usually associated with entry into stationary phase. The regulatory control of this shift and the associated phenotypes is poorly understood and has previously been linked to the master regulator Spo0A, but also the RNPP-like quorum sensing regulators Ca_c0957 and Ca_c0958 and Agr quorum sensing. The main objective of this study was to provide a better understanding of these systems and their interactions with respect to shift-associated phenotypes.

The role of Ca_c0957 and Ca_c0958 as master regulators of stationary phase phenotypes were confirmed through in-frame deletion and chromosomal complementation and both were shown to be essential for solvent formation and sporulation. While previously assumed to act primarily through control of Spo0A phosphorylation, this study demonstrated that some Spo0A activity remained in the absence of Ca_c0957 and served to independently control granulose accumulation. Inactivation of Ca_c0082, a LanC-like peptide cyclase controlled by Ca_c0957, produced a granulose-negative phenotype with reduced sporulation but capable of producing the Agr signalling peptide, thus resembling a Spo0A mutant.

C-terminally tagged Ca_c0957 and Ca_c0958 proteins were constructed and confirmed to be active, allowing for their future structural characterisation, even though initial purification attempts remained unsuccessful.

Agr signalling, controlled by Ca_c0957/Ca_c0958, was shown to be more promiscuous than previously thought, with several synthetic peptides displaying biological activity, including one representing the putative *Clostridium roseum* AgrD signal. Agr signalling influenced sporulation primarily through granulose accumulation, as granulose-deficient mutants were shown to produce fewer spores but generate solvents earlier than the wild type, especially under glucose limitation.

Acknowledgements

This work was made possible by the funding of the University of Nottingham's BBSRC Doctoral Training Partnership (DTP) programme.

I would primarily like to thank my supervisor Klaus Winzer for his useful advice and ideas across many discussions throughout the past 4 years. I would also like to extend that sentiment to the past and present members of the CRG/SBRC who have provided insight, advice or assistance throughout this work (with particular thanks extending to Wouter Kuit as I know he would be the only person likely to read these acknowledgements and be offended if he didn't get a personal mention). I would also like to thank CHAIN Biotech for providing me with a PIPS placement and an insight into industrial *Clostridium* research. I would also like to thank Oliver Severn and the late Victor Tudorica for who as rotation students performed an initial pilot study to characterise a σ^k::CT*ermB* mutant and Jonathan Humphreys who as a rotation student generated and characterised the granulose phenotypes of an initial set of Ca_c0077::CT*ermB* and Ca_c0082::CT*ermB* mutants.

For support and a great social experience throughout my PhD I would also like to thank: multiple past and present CBS employees/students in particular the KRG lads. In addition I extend thanks to the fellow Beeston Dwellers, the Johnson Arms/Star Inn/Crown/Pottle of Blues pubs and the employees and residents of the Kitty Café (especially Baloo, Heathcliffe and Toffee) for making the past 4 years so enjoyable.

Table of Contents

List of Figures

List of Tables

Abbreviations

µF	Micro farads
µg	Microgram
µl	Microliter
3OH-PAME	3-hydroxy palmitic acid methyl ester
5-FC	5-Fluorocytosine
5-FOA	5-Fluoroorotic acid
ABC transporter	ATP-binding cassette transporter
ABE	Acetone butanol ethanol
ACE	Allelic coupled exchange
ADP	Adenosine diphosphate
Agr	Accessory gene regulator
AHL	Acylated homoserine lactones
AIP	Autoinducing Peptides
AMP	Adenosine monophosphate
ATCC	American type culture collection
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool

Bp	Base pair
Cas	CRISPR associated proteins
CBM-S	Clostridial basal medium – sporulation
cDNA	Complementary DNA
CFU	Colony forming unit
CGM	Clostridial growth medium
ChIP-seq	Chromatin immunoprecipitation sequencing
CoA	Coenzyme A
CRISPR	Clustered regularly interspaced short palindromic repeats
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSF	Diffusible signal factor
EDTA	Ethylenediaminetetraacetic acid
EPB	Electroporation buffer
g	gravity
GC	Gas chromatography
gDNA	Genomic DNA
HHQ	2-heptyl-4-quinolone
HRP	Horseradish peroxidase

HSL	Homoserine lactone
HTH	Helix-turn-helix
IPTG	Isopropyl β-D-1-thiogakactopyranoside
kJ	Kilo Joules
kV	Kilo volts
LB	Lysogeny Broth
mg	Milligram
ml	Millilitre
mm	Millimetre
mM	Millimolar
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
Nanoluc	Nano-luciferase
NCBI	National Centre for Biotechnology information
nm	Nanometres
NprR	Neutral protease regulator
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction

PlcR	Phospholipase C regulator
PQS	2-heptyl-3-hydroxy-4-quinolone
RAM	Retro-transposition activated marker
Rap	Response regulator aspartate phosphatases
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RNP	Ribonucleoprotein
RNPP	Rap, NprR, PlcR, PrgX
RPM	Rotations per minute
rRNA	Ribosomal ribonucleic acid
SASPS	Small acid soluble proteins
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electophoresis
SOE	Splicing by overlap extension
SSC	Saline-sodium citrate
TAE	Tris acetate EDTA
TMB	Tetramethylbenxidine
TPR	Tetratricopeptide repeat
UV	Ultra violet
V	Volt

1. Introduction

1.1. Biofuels

1.1.1. The increasing human demand for energy and the increased demand for fossil fuel alternatives

As the worldwide human population, having reached an estimated 7.35 billion as of 2015, continues to climbs (United Nations 2015), its energy demand will continue to rise alongside it (Ehrlich & Holdren 1971). Though renewable energy sources, such as biofuels, currently stand at an all-time high, meeting 2.8% of the rise in demand, this demand has largely been met by the increased utilisation of coal and natural gas (BP 2016).

The continued reliance on fossil fuel supplies to match the expanding energy demands of humanity remains problematic for a multitude of reasons, the most crucial reason being the limited nature of their supply. Though predictions have shown large variations with respect to when fossil fuel reserves will be expended (Shafiee & Topal 2009a; Hughes & Rudolph 2011), it is accepted that as a finite resource both peak supply and eventually complete exhaustion will be achieved in the not too distant future (example calculations by Shafiee & Topal (2009b) predict this to occur within 35, 107 and 37 years for oil coal and natural gas respectively) .

Meeting the expanding energy demand is also likely to require the use of reserves that are subject to more pollutants or are less easily accessible, such as deep sea oil drilling which can occur huge costs not only from the logistics of its extraction but also as a result of damages and spills that may occur as a result (Krupnick et al. 2011). These resources, while yielding a net energy gain from their extraction, have been shown to be approximately 11% less efficient than the ‘clean’ fossil fuel reserves, which consist mostly of existing oil field reserves (Valero & Valero 2012). While speculation causes potentially large fluctuations in the prices of fossil fuels, the increased dependency on less efficient and higher polluting reserves combined with

the rise in energy demand is predicted to ultimately result in the real price of these resources increasing significantly in the future (Shafiee & Topal 2010).

While the increased difficulty in tapping reserves and likely increases in price may aid in the drive for increased renewable sources, it has been argued that due to their present status as being far more economic compared to other energy sources, it is only climate change concerns that will serve as the primary drive towards the expansion of renewable energy sources (Siirola 2014). Carbon dioxide released through the human burning of fossil fuels has been linked to both an increase in atmospheric CO_2 of 40% and an associated rise in the Earths mean surface temperature of 0.4 °C in the past 150 years (Robock 2015). In order to limit greenhouse gas emissions and a continued rise in global temperatures, based on legislative targets, it has been predicted that globally between 33-80% of predicted total fossil fuel reserves should go unused from 2010 to 2050 (McGlade & Ekins 2015).

Due to the many complications of fossil fuels usage and the current dependency on their use, alternate resources have become a necessity in order to meet future growing energy demands. While dependent on the means by which land usage is affected (Searchinger et al. 2008), biofuels represent one such opportunity for a potentially carbon neutral bioenergy generation process. This is theoretically achieved since feed stocks for such processes are produced by the fixation of atmospheric carbon and can potentially result in a near term net decrease in carbon emissions when substituting for fossil fuels (Schmer et al. 2014; Parliamentary Committee 2008).

1.1.2. First generation biofuels

First generation biofuels were developed as a result of the rise in dependency of fossil fuels, as a means to provide convenient and reliable bioenergy sources in a more economical and ecological manner than fossil fuel utilisation (Naik et al. 2010). The three most common forms

of first generation biofuels are: bioethanol, biodiesel and biogas production. Bioethanol is primarily produced from fermentation of starch and/or sugars, often to be used as a fuel source in cars with appropriate engine modification. Biodiesel is a diesel substitute generated through the transesterification of vegetable fats and can serve as a full substitute to diesel with minor engine modifications. Biogas (primarily consisting of methane) is produced by anaerobic fermentation of manure or other digestible feedstock and again can be used in motor vehicles with minor engine adjustments, though the major application is the generation of electricity (Naik et al. 2010).

First generation biofuels have however become associated with a number of production difficulties that have rendered the outcomes of some processes to be less than desirable. One such issue is the fact that all first generation biofuels use resources that are either food stocks or can potentially be used in the generation of food stocks for human/animal consumption. The use of such resources and the associated land consumption for biofuels plays a key role in the 'food vs fuel' competition, which has been observed to drive up prices within the agricultural industry (Murphy et al. 2011), leading to both increased cost and decreased efficiencies in biofuel and food production.

Outside of the competition with food production, various other negative effects have been linked to some first generation biofuels approaches. A major component of this is the requisition of land that has subsequently been cleared for the purposed of cultivating biofuel feed stocks. This land clearing has been shown to often result in an actual considerable net generation of carbon emissions. In some examples carbon emissions generated during land clearing are as large as 17 to 420 times the size of the reductions made as a result of the use of these compounds as bioenergy sources (Fargione et al. 2008). In addition to this the clearing of areas for biofuel production raises environmental concerns with respect to the loss of

biodiversity and the use of nitrogen fertilisers in the intensive cultivation processes (Mohr & Raman 2013). As a result of the complications of first generation biofuels, a limit has been placed on their implementation and a push towards the development and implementation of better design first generation processes and second generation biofuels has been encouraged (EU 2012).

1.1.3. Second generation biofuels

Second generation biofuels attempt to avoid the 'food vs fuel' issue through the use of non-food plant material, primarily lignocellulosic material, which currently is abundantly and cheaply available from a number of plant sources (Gomez et al. 2008). The interest in this process stems from the fact that approximately 75% of plant biomass is comprised of lignocellulosic material, itself composed of a similar levels of polysaccharides potentially representing a large feed stock for biofuel production (Pauly & Keegstra 2008).

Current development of second generation biofuels has been explored through two main routes the first of which involves the degradation of lignocellulosic material to release sugars for subsequent fermentation. Hydrolysis of lignocellulose first requires treatment in order to separate lignin from the hemicellulose (mainly xylose and arabinose containing polymers) and crystalline cellulose fractions that can then be hydrolysed to release the fermentable sugars (Chaturvedi & Verma 2013). This pre-treatment process further complicates potential further enzymatic or microbial fermentation based steps as the process often releases inhibitors, facilitating the need for a detoxification process prior to hydrolysis (Jönsson et al. 2013). Hydrolysis of the pre-treated lignocellulose can be achieved enzymatically, this often being achieved through the use of bacteria which form a highly efficient complex of cellulose enzymes, mounted to a molecular scaffold (Fontes & Gilbert 2010). At present industrial fermentation utilises a cocktail of fungal cellulose enzymes, which while many potent

examples exist, often are relatively expensive to utilise (Pothiraj et al. 2006). The sugars released from the polysaccharides have been primarily used for the generation of bioethanol in a manner similar to first generation production, though these processes are often hindered by the relatively slow rate of cellulose hydrolysis and subsequent fermentation, alongside the aforementioned high prices associated with providing cellulose enzyme cocktails (Jeon et al. 2009; Zertuche & Zall 1982; Pothiraj et al. 2006).

The second approach to this generation of biofuels is the gasification of lignocellulosic material into syngas, a mixture of carbon dioxide, carbon monoxide and hydrogen gas (Becker 1988). Syngas itself can used as a fuel and burned or, through the use of metal catalysts (Chuang 2012) or gas fixing acetogenic bacteria (Ragsdale & Pierce 2008), can be converted into ethanol or other potential biofuels or valuable chemicals. Both processes have their respective limitations, with the use of metal catalysts requiring large energy demands (Chuang 2012) and being potentially susceptible to impurities that often arise in the formation of syngas (Zeisler et al. 2010). Microbial syngas fermentation not requiring such energy intensive conditions and being less prone to impurity susceptibility is itself limited by its relatively slower speed of conversion and its development has largely been restrained as a result of the limited understanding of the organisms' energy metabolism and lack of sophisticated genetic tools for acetogens (Bertsch & Müller 2015).

Though second generation biofuels avoid the 'food vs fuel' conflict and associated land requisition issues created by first generation biofuels due to the existent abundance of lignocellulosic material available (Mabee et al. 2011; Saini et al. 2014; Tye et al. 2016), the processes involving hydrolysis and gasification still largely remain in development. While examples of commercial processes for the generation of biofuels has been achieved through

second generation techniques, at present the majority of biofuel production remains first generation with increased stringency with respect to their environmental and economic impact.

1.1.4. Butanol as a biofuel

Bioethanol is the primary biofuel generated for usage as a transportation fuel, however arguments can be made as to why butanol may serve to be superior especially when considering the fermentation processes both use interchangeable feedstocks and adaptable fermentation processes (Pfromm et al. 2010).

Butanol's primary advantage as a biofuel is the relatively larger energy content by volume compared to ethanol. Butanol has a better heating value of 33.4 kJ g^{-1} as opposed to that of ethanol which is only 27.0 kJ g^{-1} (Weast 1977), placing butanol more in line and therefore substitutable for the 43.4 kJ g^{-1} found in standard transport fuel (Pfromm et al. 2010).

While being more desirable due to its higher energy content biobutanol is also advantageous with respect to its practical application in transport fuel mixtures. In contrast to ethanol which requires significant engine modification in order to be used in high concentration fuel blends, butanol is able to function in existing motors in any concentration up to and including using butanol alone (Szulczyk 2010). Further facilitating its advantageous use as a transportation fuel over bioethanol, biobutanol blends do not undergo separation in the presence of water and due to its much lower miscibility further removes complications in the formation, distribution and use of fuel blends as water responsible for corrosion is excluded (Niven 2005).

While butanol confers several advantages over bioethanol utilisation as a biofuel and the process of microbial butanol production having been long established (Jones & Woods 1986), bioethanol still remains the existing choice of transport biofuel. This is primarily due to the fact that while feed stocks are either the same or similar for both process, biobutanol has been

limited in titre to approximately 2% and as a consequence past applications of the process were relatively inefficient when compared to bioethanol fermentation (Pfromm et al. 2010). However with further study and optimisation of microbial butanol production through the increased availability of generic tools it may be possible to sufficiently increase the efficiency of the existing processes so that they can more efficiently compete with bioethanol production.

1.2. Clostridia

1.2.1. Introduction to the genera *Clostridium*

The genus *Clostridium*, pre 16S-RNA description, compromised all Gram-positive, obligate anaerobic, spore forming, non-sulphate reducing bacteria (Lawson et al. 1993). Since the advent of 16S-RNA analysis, the genus *Clostridium* has since been refined and the evolutionary relationship between these organisms confirmed through more accurate means than previous phenotypic characterisation (Ludwig et al. 2009). The increasing prevalence of whole genome sequencing has further aided in the ongoing process of the reclassification of many organisms identified as belonging to the genus *Clostridium*, such as *Clostridium difficile* which has been reclassified and is now proposed as the type strain for the new genus *Peptoclostridium* (Yutin & Galperin 2013). Though the genus has over time been greatly refined, the current genus *Clostridium* still represents a physiologically and ecologically diverse collection of species.

The genus belongs to the phylum Firmicutes, which generally possess a genomic G+C content of under 50% and a Gram-positive muramic acid-containing cell wall (Ludwig et al. 2009). The phylum is divided into three classes *Bacilli, Clostridia* and *Erysipolotrichia,* which across them include thermophilic, halophilic and mesophilic organisms, spore-forming and non-spore forming bacteria and organism both of importance in healthcare and industry. The genus *Clostridium* belong to the class *Clostridia,* order *Clostridiales* and of the family *Clostridiaceae* Due to the ongoing reclassification amongst the genus, *Clostridium* identified by their

phylogenetic relation to the type strain *Clostridium butyricum* are described as *Clostridium sensu stricto* (Huber & Prangishvili 2006) which have since been shown to conservatively share novel protein signatures unique to the cluster (Gupta & Gao 2009).

Within the *Clostridium sensu stricto* cluster are a number of important pathogenic organisms such as *Clostridium botulinum, Clostridium tetani* and *Clostridium perfringens*. *Clostridium botulinum* is of particular interest as while infection with the organism can result in death through flaccid paralysis, the responsible neurotoxin is also widely used in the treatment of a range of overactive muscle and gland secretion disorders (Münchau & Bhatia 2000). With respect to medical application the organism *Clostridium* sporogenes is also of interest as it has been shown that germination of spores within hypoxic tumours is a safe and potentially effective means of cancer treatment (Heap et al. 2014).

Within the *Clostridium sensu stricto* cluster are also organisms of particular industrial interest, such as the solventogenic *Clostridium acetobutylicum* and carbon monoxide fixing organisms such as *Clostridium autoethanogenum*. The use of waste syngas as a stream for the chemical synthesis, primarily 2-3 butanediol and ethanol, is currently being developed in industry using *C. autoethanogenum* (Köpke et al. 2011). Though not in current use, outside of sporadic use in China and a proposed plant in the US (Chiao & Sun 2007; Fletcher 2016) the use of solventogenic *Clostidium* species, including *C. acetobutylicum* amongst others, have been used in the past for the production of both acetone and biobutanol.

1.2.2. Acetone Butanol Ethanol fermentation

1.2.2.1. The history of ABE fermentation

Though bacterial butanol fermentation was first described by Pasteur in 1861 and, considerably later, acetone fermentation was discovered by Schardinger (1905), the process of ABE (acetone, butanol, ethanol) fermentation was primarily characterised and developed by the

Russian chemist Chaim Weizmann. Chaim Weizmann was responsible for the isolation of a number of butanol producing strains and subsequently patented an industrial fermentation process using the most efficient organism he had isolated, *Clostridium acetobutylicum*, to produce acetone, butanol and ethanol from sugars (Jones & Woods 1986).

Initially the work of Weizmann was centred around the production of butanol for the production of synthetic rubbers, through the conversion of butanol to butadiene (Gabriel 1928), though the focus of his work changed with the outbreak of the first world war. Acetone became the focus of ABE production due to its importance in the manufacturing of the smokeless powder cordite which was required by the British military. Prior to the war acetone was sourced from calcium acetate, partially imported from both Germany and Austria and as a result became limited in supply (Jones & Woods 1986). Weizmann worked with the British government to establish plants for the production of acetone using *C. acetobutylicum* across the British empire/commonwealth and the US until the armistice of 1918 (Jones & Woods 1986).

For his contribution to the war effort Weizmann refused any personal honours or rewards, though as a Zionist he made clear that he desired the creation of a Jewish state in Palestine (Jones & Woods 1986). Though disputed as to whether the creation of Israel in the Balfour declaration was a gift to Chaim Weizmann or an invention of David Lloyd George (Reinharz 1985), upon the founding of Israel Chaim Weizmann became its first president in 1948 (Jones & Woods 1986) cementing the importance of his scientific and political legacy and influence it had played in the first half of the previous century.

Following the first world war the importance of butanol rose once again due to the discovery that butanol was an ideal solvent for use in quick drying lacquers, then in demand by the expanding automobile industry (Jones & Woods 1986). The growth of the ABE fermentation industry was also heavily boosted by the demand for acetone after the outbreak of world war

two. As a result of both these factors by 1945 two thirds of all butanol and one tenth of all acetone in the United states was produced by ABE fermentation (Jones & Woods 1986).

In the 1950's the expanding petrochemical industry became a huge source of competition for ABE fermentation, alongside the escalation in the price of feed stocks due to competition with the cattle industry in the United States in particular (Jones & Woods 1986). As a result ABE fermentation rapidly declined to the point of cessation in the 1960's in the United States and Britain (Jones & Woods 1986). Within the USSR work continued to maintain some interest in ABE fermentation with respect to examining its viability with hydrolysed agricultural waste (Zverlov et al. 2006), while ABE plants in South Africa continued functioning viably until 1983 when a shortage of molasses, due to drought, forced the closure of the last ABE plant (Jones & Woods 1986).

Due to the increased prices and environmental concerns associated with fossil fuels there has been a renewed interest with respect to ABE fermentation. Within China this has led to the reopening of previously closed ABE fermentation plants (Chiao & Sun 2007) and within the USA has led to the first conversion of ethanol fermentation plants to commercially active Biobutanol plants (GreenBiologics 2016). Overall this has caused a large renewed interest in this field of research as the development of new tools and techniques are underway to further increase the efficiency and competitiveness of ABE fermentation (Sauer 2016).

1.2.2.2. Substrates, culture and solvent recovery used in ABE fermentation

Substrates previously used for ABE fermentation have varied from country to country, though the process was initially performed with maize and subsequently horse chestnuts as the first world war created a grain shortage (Jones & Woods 1986). Other substrates that have been used or explored include hydrolysed agricultural waste (Zverlov et al. 2006), molasses (Jones & Woods 1986), Jerusalem artichokes (Sarchami & Rehmann 2015) and date palm (Abd-Alla

& Elsadek El-Enany 2012). The choice of substrate is key in the ABE process as the cost of substrate can account for up to 50% of the running cost of ABE fermentation and thus requires cheap and readily utilisable substrates in order to achieve acceptable profitability (Dürre 2007).

Due to the toxicity of butanol it is common to see lower titres when the organism is grown in batch culture, though often produces the largest concentrations of total solvents (Li et al. 2011). In the laboratory, higher butanol yields and productivity can be obtained through fed-batch culture, in which the substrate is fed as it is utilised, and continuous culture, where upon reaching stationary phase the culture is continuously fed substrate and diluted (Li et al. 2011). While allowing for greater butanol yields and productivity fed-batch cultures and continuous cultures are limited by the time necessary for the organism to switch from acid to solvent production (Li et al. 2011) and may be more susceptible to degeneration in which solvent production and sporulation are lost over extended culturing periods (Kashket & Cao 1995).

Another key technical aspect to ABE fermentation is the recovery of the products from the growth medium. While multiple recovery methods have been reviewed such as distillation and reverse osmosis, the favoured approach for recovery are those that can be implemented *in situ* such as gas stripping, pervaporation and absorption (Kujawska et al. 2015). *In situ* techniques are favourable they allow the reduction of the toxic effects of butanol and acetone upon the cells (Kujawska et al. 2015). Absorption relies upon the preferential attachment of particles from a liquid/gas to a solid surface, which has been shown through one method to extract up to 80% of butanol and increase yields 2-3 fold (Wiehn et al. 2014); while another absorption method has shown significantly increased yields through co-absorption of acetone (Liu et al. 2014). Gas stripping is the separation method by which volatile ABE products are condensed and collected following the sparging of the media. This process can often be partially limited due to the large level of foam formation that can occur (Ezeji et al. 2005), but has also shown

to be largely effective in increasing sugar utilisation and solvent production in the ABE fermentation process (Ezeji et al. 2004). Another *in situ* technique that has been explored is the use of pervaporation, which involves the creation of a driving force by using a vacuum across a membrane selective for butanol recovery. While selective this process can be limited by fouling of the membrane from components of the growth media reducing the effectiveness of this method over time (Fadeev et al. 2000). Though all these in situ processes are still being developed, gas stripping is at presented the only process which has been successfully integrated into large scale ABE fermentation for in situ recovery to increase observed solvent yields in a way deemed to be economically efficient (Kujawska et al. 2015).

1.2.2.3. Strains used in ABE fermentation

As previously discussed *C. acetobutylicum* was the first species used in an industrial ABE fermentation process. However historically multiple different organisms have been used in biobutanol production (Jones & Woods 1986). A confusion arising from the original nomenclature used to describe solventogenic *Clostridia* species, was that the classification of organisms was based primarily on the ratio of solvents they produced (in spite of often being given a different name upon discovery). Under the previous nomenclature most organisms were described as *C. acetobutylicum* when they performed the classical ABE fermentation, as *C. beijerinckii* if they produced isopropanol in the place of acetone and as *C. aurantibutyricum* if they generated both acetone and isopropanol (Jones & Woods 1986). With the advent of 16S rRNA analysis the broad classification of *Clostridia* species has been greatly changed as strains described as one particular species were often shown to be far more closely related to a different type strain. This has led to the reclassification of different organisms so that they now group entirely within their own individual species (Keis et al. 1995; Keis et al. 2001).

Despite the large strain variety originally employed within industry, the primary focus with respect to research has concerned the metabolic engineering of *C. acetobutylicum* (in particular the ATCC 824 type strain) and *C. beijerinckii* (Bahl & Lutke-Eversloh 2011). However, more recent research has branched out into the exploration of other potential industrially relevant strains including those previously used in industrial contexts (ERA IB 2016). One such strain of interest is *C. pasteurianum* for which a number of genetic tools have recently been implemented (Pyne et al. 2013; Pyne et al. 2015; Pyne et al. 2016). This renewed interest is due to the organism's ability to generate butanol from a feedstock of glycerol (Biebl 2001), is a prevalent waste product of biodiesel production (Leoneti et al. 2012).

Though characterisation and optimisation of *C. acetobutylicum* and *C. beijerinckii* type strains is likely to remain prevalent in this area of research, future attempts to increase effectiveness of industrial ABE fermentation will likely use a wider variety of strains based upon: the desired ratio of end products, utilisation of specific substrates as the process demands, the necessary level of provision of extra nutritional supplementation requires and the achievement of economically viable yields, titres and productivities (Jones & Woods 1986). In the past this process was primarily achieved through the isolation and characterisation of new cultures, however the expanding range of genetic tool (discussed in 1.4) will allow for greater manipulation of existing strains in the future development of ABE fermentation processes.

1.2.2.4. Metabolism of ABE fermentation

ABE fermentation is considered a bi-phasic process, in which the organism initially utilise sugars in the generation of ATP through the production of acetic and butyric acid (acidogenesis). Following a shift in metabolism, associated with the entry into stationary phase, solventogenic organisms subsequently begin the reassimilation of acids and production of

acetone, butanol and ethanol (solventogenesis) (Jones & Woods 1986; Bahl & Lutke-Eversloh 2011).

Following the passage of hexose sugars through the Embden-Meyerhof-Parnas pathway and the pentose phosphate pathway, respectively, leading to ATP generation through substrate level phosphorylation, the generated pyruvate is able to enter the ABE pathways. Pyruvate is converted to acetyl-CoA through the activity of pyruvate ferredoxin oxidoreductase (Pfor), a key enzyme in clostridial metabolism, for use in ABE metabolism, thereby releasing CO_2 and generating reduced ferredoxin (Jones & Woods 1986). Pyruvate may instead be directly converted to lactic acid by lactate dehydrogenase (Ldh), though large fluxes to lactate are only observed under certain pH controlled or iron limitation culture conditions (Dabrock et al. 1992; Bahl et al. 1986).

The core of the ABE pathway shared by both acidogenesis and solventogenesis begins with the conversion of two acetyl-CoA to one acetoacetyl-CoA by thiolase (Thl), followed by a reduction to 3-hydroxybutryryl-CoA and a further reduction to butyryl-CoA by the enzymes 3-hydroxybutyryl-CoA dehydrogenase (Hbd), crotonase (Crt) and butyryl-CoA dehydrogenase (Bcd) respectively (Figure 1.1). The activity of Bcd is very exergonic and solventogenic bacteria make use of this through transferring electrons from NADH to ferredoxin (Bahl & Lutke-Eversloh 2011). However unlike some *Clostridium* species such as *C. kluyveri* which displays an Rnf complex conserving the energy in the form of a proton gradient (Herrmann et al. 2008), *C. acetobutylicum* uses the hydrogenase HydA for releasing the electrons through the formation of hydrogen gas, to regenerate ferredoxin and renew cofactors (Bahl & Lutke-Eversloh 2011). HydA gene appears to be of key importance in the regeneration of ferredoxin in the cell as Cooksley et al. (2012) were unable to create inactivation mutants when observing the effect of inactivation's across the ABE pathway.

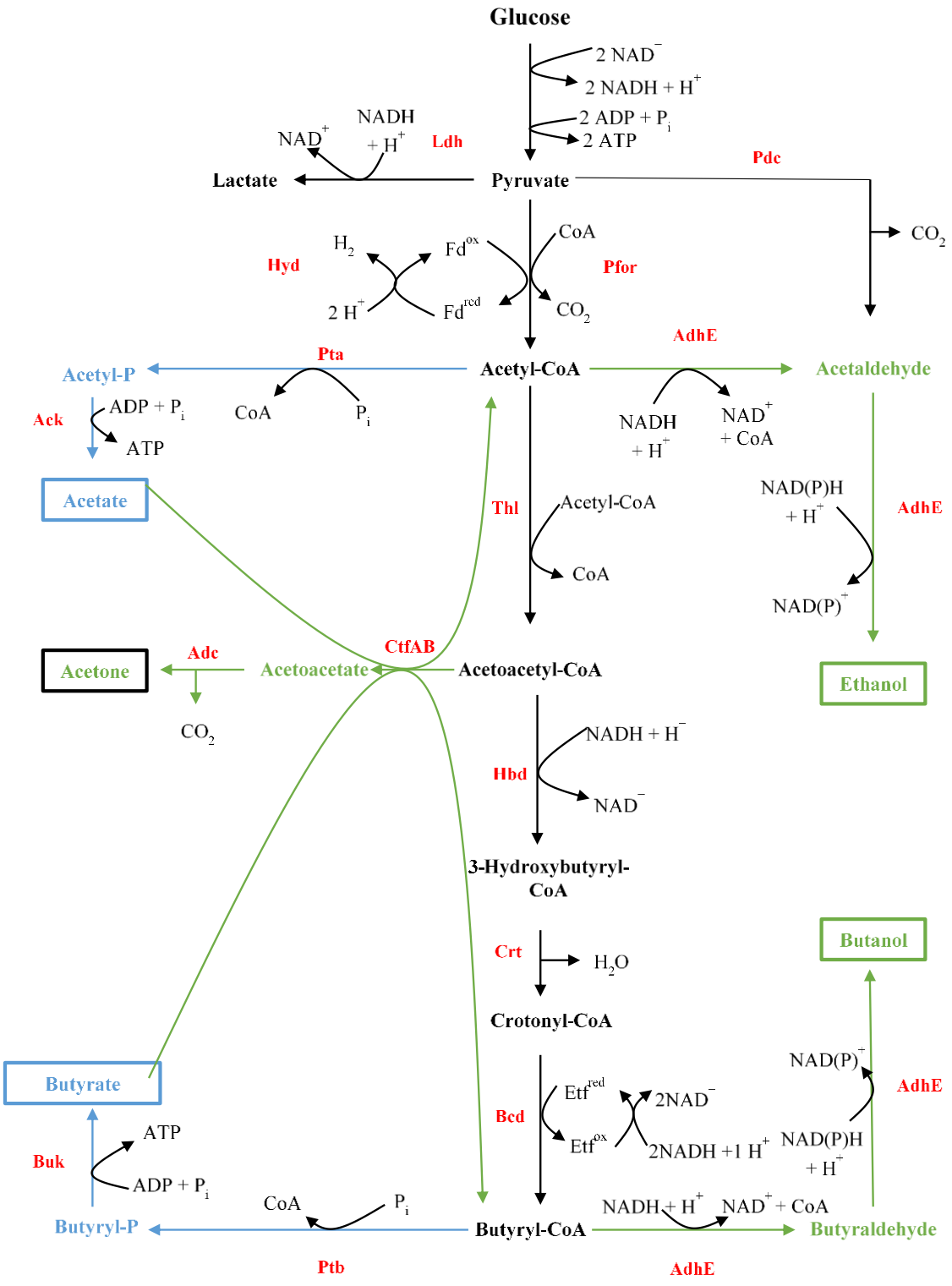


Figure 1.1 ABE fermentation in *C. acetobutylicum*

ABE metabolism of *C. acetobutylicum* adapted from Bahl & Lutke-Eversloh (2011). The red letters show the enzymes involved in the fermentative pathways: Ldh, lactate dehydrogenase; Pdc, pyruvate decarboxylase (gene annotated but pathway unconfirmed); Hyd, hydrogenase; Pfor, pyruvate:ferredoxin oxidoreductase; Fd, ferredoxin; Pta, phosphotransacetylase; Ack, acetate kinase; AdhE, aldehyde/alcohol dehyrogenase; CtfAB, acetoacetyl-CoA:acyl-CoA transferase; Adc, acetoacetate decarboxylase; Thl, thiolase; Hbd, 3-hydroxybutyryl-

CoA dehydrogenase; Crt, crotonase, Bcd, butyryl-CoA dehydrogenase; Etf, electron transfer flavoprotein; Ptb, phosphotransbutyrylase, Buk, butyrate kinase; ox, oxidized; red, reduced. Acidogenic pathways, intermediates and products are coloured blue. Solventogenic pathways, intermediates and products are coloured green.

During acidogenesis, acetyl-CoA and butyryl-CoA are phosphorylated by phosphotransacetylase (Pta) and phosphotransbutyrylase (Ptb) respectively before being converted to acetate and butyrate, by acetate kinase (Ack) and butyrate kinase (Buk) which couple these reactions to the generation of an additional ATP (Bahl & Lutke-Eversloh 2011). The conversion of acetyl-CoA to acetate while generating twice the ATP of butyrate production (as two acetyl-CoA's are used to generate butyrate) allows for a decreased level of NAD cofactors to be regenerated. Regardless formation of acetate and butyrate serves to decrease the pH of the growth medium and if allowed to continue unregulated can be responsible for the collapse of the proton gradient across the cells membrane, due to the diffusion of undissociated acids back into the cell, thus resulting in a loss of ATP formation and a collapse of cellular metabolism (Jones & Woods 1986).

In order to prevent the culture 'acid crashing' (Maddox et al. 2000) the process of solventogenesis is initiated, in which acids are partially reassimilated and subsequently converted to acetone, ethanol and butanol (Figure 1.1). A small amount of ethanol is produced during acidogenic phase though its accumulation is not significant enough to prevent the need for the activation of reassimilation and the formation of acetone and butanol, likely just serving a purpose of maintaining redox balances (Jones & Woods 1986). This shift to solvent production is commonly associated with the organism entering stationary phase and the initiation of a number of exclusive stationary phase phenotypes such as accumulation of the energy storage compound granulose and the initiation of endospore formation (Dürre & Hollergschwandner 2004).

During solventogenesis butyrate and acetate are reassimilated by CftA/CtfB in a reaction coupled to the formation of acetoacetate from acetoacetyl-CoA (Bahl & Lutke-Eversloh 2011), although an alternative, as yet unidentified route may also exist (Millat et al. 2014). Acetoacetate formed in the coupled reaction is subsequently decarboxylated, resulting in the formation of acetone (Figure 1.1). In organisms previously characterised as *C. beijerinckii* or *C. aurantibutyricum* (see 1.2.2.3) an additional NAD(P)H dependent alcohol dehydrogenase (Adh) is present which, while showing some overlapping activity with the existing Adh's, is primarily responsible for the conversion of acetone into isopropanol (Ismaiel et al. 1993).

During solvent formation the reassimilated acids, now present in the form of acetyl-CoA and butyryl-CoA are reduced to aldehydes and subsequently ethanol and butanol by Adh enzymes (Bahl & Lutke-Eversloh 2011). In *C. acetobutylicum* two *adhE* genes are present, *adhE1* and *adhE2* in addition to two specific butanol dehydrogenases *bhdA* and *bdhB* (Walter et al. 1992). AdhE2 is more commonly associated with alcohologenic fermentation (Fontaine et al. 2002) - in which acetone accumulation production is not observed (Girbal et al. 1995), while AdhE1 appears to be associated with standard solventogenic fermentations producing a 2:3:1 ABE product profile (Green & Bennett 1996).

Solvent accumulation eventually becomes toxic to the organisms through increased membrane fluidity causing the inhibition of: internal pH regulation, ATP generation and glucose uptake (Bowles & Ellefson 1985). Though solvents are toxic too, the reassimilation of acids prevents the cells from being killed as a result of an acid crash and allows the metabolism of the organism to continue for considerably longer before lethal levels of a product accumulate in the growth media. In an environmental context the continued metabolism allows the cells time to form endospores in response to the growing toxicity of their environment and ensure the organisms continued survival.

1.3. Physiology of stationary phase *Clostridium acetobutylicum*

Outside of the aforementioned reassimilation of acids and initiation of solventogenesis, other changes in physiology can be observed within stationary phase solventogenic *Clostridium* spp. These are the accumulation of the energy storage polymer granulose, the formation of an extracellular capsule and the initiation of endospore formation (Jones & Woods 1986).

Like *Bacillus* the master regulator of sporulation is Spo0A, but in solventogenic *Clostridium spp.,* it also serves as a master regulator of solventogenesis and has been implied to regulate granulose accumulation indirectly through its activating effect upon σ factors which themselves serve as transcriptional regulators (Ravagnani et al. 2000; Tracy et al. 2011). In *Bacillus spp.* activation of Spo0A is achieved through a two component phosphorelay, in which mechanisms that signal conditions desirable or unfavourable for sporulation act through response regulators to activate/inactivate the phosphorelay system (Burbulys et al. 1991). This systems has not been observed as conserved within any *Clostridium* spp. (Dürre & Hollergschwandner 2004). Instead the regulation of Spo0A/Spo0A-P has been shown to be directly regulated by orphan kinases activity (Underwood et al. 2009; Steiner et al. 2011) though the mechanisms by which these orphan kinases are themselves regulated and activated in order to signal the activation of sporulation remains largely unexplored.

1.3.1. Granulose accumulation

Though not present in all solvent-producing *Clostridium* species the formation of the energy storage compound granulose is seen as highly conserved throughout the butanol producing subset of *Clostridium sensu stricto* strains (Reysenbach et al. 1986). Though the formation of granulose itself is highly conserved some differences in structure can be noted as in *C. acetobutylicum* and *C. pasteurianum* it has been described as being an α-glucose polymer containing only 1-4 glycosidic bonds, while in *C. butyricum* it was shown to possess a

branched structure which incorporated 1-6 glycosidic bonds (Reysenbach et al. 1986; Robson et al. 1974; Hobson & P. Nasr 1951).

Following the initiation of granulose accumulation often coinciding with the initiation of solvent production, peak accumulation can be observed between 45-48 h of growth, at which point cells have been shown to contain a maximum deposition of up to 250 to 300 µg of glucose per 10^8 cells which can account for 40% to 55% of the cells dry weight (Reysenbach et al. 1986; Steiner et al. 2012). The activity of granulose synthase has been shown to be inhibited by ADP and AMP, likely a means to prevent the formation of granulose when cells are more metabolically starved (Robson et al. 1974; Reysenbach et al. 1986).

Following the initiation of endospore formation granulose is broken down and consumed by the organism (Reysenbach et al. 1986). Supporting the theory that granulose serves as a conserved energy source for sporulation is the observation that while solvent production remains largely unaffected, isolated granulose negative mutants have shown an associated reduction in sporulation (Reysenbach et al. 1986). However, due to the not fully characterised mechanism of granulose regulation it remains unclear whether this is a causative or correlative effect (Reysenbach et al. 1986; Steiner et al. 2012).

1.3.2. Endospore formation

Bacterial endospores serve as a survival mechanism for certain species when under unfavourable environmental conditions, as while in the form of a spore little to no metabolic activity occurs and the spore is highly resistant to radiation, mechanical disruption, the effect of toxic chemicals and increased temperatures (Driks et al. 2000). Following the return to a favourable environment a spore is able to germinate into a functional bacterial cell and begins growing once again. While the activation mechanism of sporulation regulator Spo0A shows a significant difference to that present in *Bacillus* spp. the general process of endospore

formation remains largely conserved in *Clostridia* with respect to morphology, biochemical changes and the regulation by σ factors, though with some notable distinctions.

1.3.2.1. Morphological and Biochemical changes during endospore formation

In *Bacillus* and *Clostridium* the first visible event of endospore formation is the segregation of chromosomes by an asymmetric septum, with the larger compartment described as the mother cell and the smaller as the forespore (Piggot & Hilbert 2004; Dürre & Hollergschwandner 2004). The cytoplasmic membrane of the mother cell engulfs that of the forespore in order to create separate environments for the independent regulation of σ factors in the sporulation criss-cross cascade (Piggot & Losick 2002).

In the forespore the segregated chromosome becomes tightly packed in an α-helix due to the binding of small acid soluble proteins (SASP's) (Mohr et al. 1991); while between the two membranes surrounding the forespore two peptidoglycan layers form the cortex (which may contain associated proteins) and the germ cell wall (Driks et al. 2000). Outside of the forespore membrane a protein based coat is synthesised by the mother cell (Driks et al. 2000). Within the forespore a large level of diplicolinic acid production is responsible for the chelation of the spores cations and the generation of 3-phospho-glycerate to provide an energy source post germination alongside the degradation of SASP's (Driks et al. 2000).

Morphologically the formation of *Clostridium* spores shows an additional phase before the formation of the asymmetric septum, in which the cell take on a cigar like swollen shape and accumulate granulose (Woods et al. 1987). It can also be commonly observed that *Clostridium* spores carry appendages that arise from the protein coat of the spore (Samsonoff et al. 1971).

1.3.2.2. Regulation through the σ factor cascade

The regulation of the sporulation cascade has been described as crisscross (Losick et al. 1992), as it is relies on the independent expression of sigma factors and proteins responsible for their expression/activation, which are differentially expressed/activated in the forespore and the mother cell (Figure 1.2).

In *B. subtilis* σ^H is constitutively expressed at low levels in a vegetative cell, though its repression is decreased by the activity of Spo0A-P. Both σ^H and Spo0A-P are required for the expression of anti-anti-sigma factor SpoIIAA, anti-sigma factor SpoIIAB and σ^F. SpoIIAB is responsible for binding σ^F and keeping it inactive and can bind ATP and ADP; with the former phosphorylating SpoIIAA and the later allowing the formation of an ADP-SpoIIAB-SpoIIAA complex (Piggot, P & Losick 2002). The formation of this ADP-SpoIIAB-SpoIIAA complex allows σ^F to become active in the forespore which leads to the transcription of the signalling protein SpoIIR which is secreted into the space between the membranes which separate the fore spore from the mother cell and activates the SpoIIGA protease in the latter. SpoIIGA is responsible for the cleaving of pro-σ^E into the mature σ^E (Piggot & Hilbert 2004). σ^E induces the expression of *spoIIIA* and through an interaction with the products of the *spoIIJ* operon allows the activation of σ^G within the forespore (expressed by the active $\sigma^{F)}$. σ^G is responsible for the maturation of σ^K within the mother cell which is expressed as a result of σ^E activity (Piggot & Losick 2002). Pro- σ^K is processed to σ^K by SpoIVFB upon its release from the SpoIVFA-BofA-SpoIVFB complex due to σ^G activity (Piggot & Hilbert 2004).

This crisscross cascade serves to differentially express genes within the forespore and the mother cell to enact the morphological and biochemical changes previously described. However outside of this, σ factors are also suggested to play roles in the regulation of other late

phase phenotypes of *Clostridium* including: granulose accumulation, solvent production and the expression of virulence factors (Tracy et al. 2011; Saujet et al. 2011).

Outside of its established role as a late phase sigma factor it has recently been suggested that σ^K may also contribute to the regulation of sporulation at a much earlier stage. Kirk et al. (2012) observed that in *C. botulinum* σ^K inactivation leads to a decrease in the transcription of Spo0A and other early σ factors, along with a decrease in the effectiveness of stress response mechanisms (Dahlsten et al. 2013). This suggestion has more recently been assessed in *C. acetobutylicum* by Al-Hinai et al. (2014), who observed the aforementioned decrease in expression of Spo0A and alongside a loss in solvent production and granulose accumulation. While the main components of the sporulation cascade are highly conserved between *Bacillus* and *Clostridium* spp., the implied dual role of σ^K alongside the activation of the cascade independent of a phosphorelay implies *Clostridium* endospore formation may be more divergent from the model *B. subtilis* than was original thought.

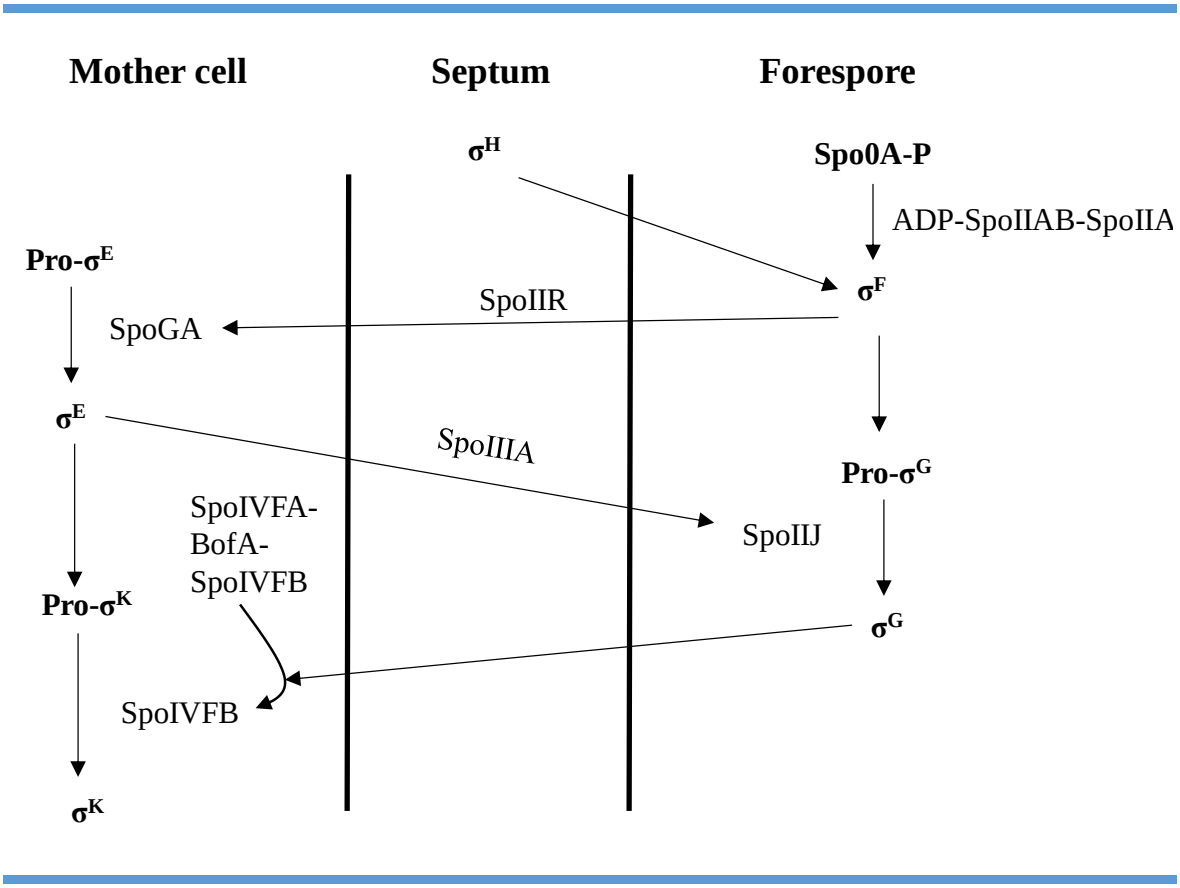


Figure 1.2 Crisscross regulation through sigma factor expression and maturation in *B. subtilis and C. acetobutylicum* based upon Dürre & Hollergschwandner (2004) and Piggot & Hilbert (2004)

Schematic representation of criss-cross regulation, showing the differential expression and activation of sigma factors in the forespore and mother cell.

1.4. Expansion of genetic tools for engineering *Clostridia*

For the majority of the history of ABE fermentation little work was performed with respect to the genetic modification of solventogenic *Clostridia* in order to improve the efficiency of the process through strain engineering, with instead more effort being placed into the isolation of individual strains of phenotypic interest (Jones & Woods 1986). Today however, strain improvement through genetic manipulation is a key concept of much research including: increasing solvent tolerances, increasing cellulosic material utilisation, increasing butanol

selectivity and oxygen tolerance (Papoutsakis 2008), all of which may serve to incrementally increase the viability of the ABE fermentation process.

Though the means to efficiently transform the type strains for *C. acetobutylicum* has long been established (Mermelstein & Papoutsakis 1993), the main hindrance to solventogenic *Clostridia* research was initially gene inactivation which was originally performed by the relatively slow to implement integration of plasmids via single crossover, resulting in mutants that often lacked stability (Green et al. 1996; Harris et al. 2002). It has only been more recently that genetic tools allowing for the faster generation of mutants (both through targeted inactivation/deletion and transposon mutagenesis) and increased chromosomal modification (allowing the uptake of larger volumes of DNA that do not require pressure to maintain plasmid stability), have been successfully established.

1.4.1. Intron Inactivation Tools

A significant development in the development of fast and reliable genetic tools for use in *Clostridia* was based upon the use of group II intron systems (also known as the Targetron). This intron insertion system allows for targeted integration and subsequent disruption of genes within the chromosome to create inactivation mutants (Zhong 2003). Mobile group II introns function through the use of a process termed retrohoming, which involves the use of a catalytically active intron RNA (described as a ribozyme) and an intron-encoded protein which has a reverse transcriptase function (Lambowitz & Zimmerly 2011). The ribozyme is responsible for its own splicing and facilitates the creation of an intron lariat RNA (Zimmerly et al. 1995), which together with the intron-encoded protein form a ribonucleoprotein complex that it able to target a specific DNA site based upon its sequence (Zhong 2003).

The ribonucleoprotein (RNP) complex upon recognising a number of fixed nucleotides in the DNA duplex causes local unwinding of the DNA and allows the RNA to subsequently pair to

the adjacent DNA sequence (Singh & Lambowitz 2001). This interaction involves three short sequences elements in the DNA and RNA interacting to facilitate the insertion of the RNA sequence directly into one strand of the DNA by reverse splicing. While this occurs the intron- optimised through the addition of a retro-transposition-activated selectable marker (RAM), in which an antibiotic resistance encoded proteins is responsible for cleaving the other strand to act as a primer for reverse transcription of the integrated RNA (Zhong 2003). The cDNA created by this process is integrated into the host genome through the use of native DNA recombination and repair enzymes (Cousineau et al. 1998), resulting in the disruption of the original DNA target sequence. Through the appropriate design of the RNA component of the RNP complex it is possible to design group II introns that are either highly target specific (Karberg et al. 2001) or allow for more random gene inactivation (Zhong 2003).

Selection of inactivation mutants has also been refined through the addition of a marker containing a group I intron is present within the group II intron (Zhong 2003). The resistance marker is contained on the intron expressing plasmid on the complementary strand to that of the ribozyme and is disrupted by the presence of a type I intron in the same orientation as the ribozyme sequence (Zhong 2003). Transcription of the ribozyme results in the splicing out of the type I intron and subsequently a complete resistance marker template for cDNA generated by reverse transcription and a selection pressure that is specific for successful intron insertion (Zhong 2003). It is also possible to design the initial group II intron to contain FRT repeat sequences either side of the resistance marker, which upon confirming insertion into the genome can allow the excision of the resistance marker through the use of a flipase protein allowing for the homologous sequences to undergo recombination and remove the resistance marker (Zhong 2003; Heap et al. 2010).

The development of a *Clostridium* equivalent of the Targetron system, the ClosTron (Heap et al. 2007), has been refined over time to include a Flp/FRT excisable *erm*-RAM and has a free

to use intron design tool available for the construction of ClosTron constructs (Heap et al. 2010). The ClosTron system has been exemplified across a range of *Clostridium* species including *C. acetobutylicum*, *C. botulinum*, *C. sporogenes* and *C. beijerinckii* (Heap et al. 2007; Heap et al. 2010; Cooksley et al. 2010).

While the availability of intron inactivation has been of clear significance in allowing the generation of rapid and stable mutants (Kuehne & Minton 2012), the technique still has a number of limitations. The first is that due to the targeting nature of the group II intron it may be difficult or impossible to identify a target sequence to allow for the effective inactivation of smaller target genes (Steiner et al. 2012). Equally due to potential incorrect insertion mutants arising strains require confirmation by Southern blot to ascertain whether insertion at a second site may have occurred (Zhong 2003). Though complementation will confirm whether this is the case, when using a RAM with its own promoter it is also possible to see phenotypic changes as a result of altered expression of flanking genes due to a polar effect (Sigma-Aldrich 2008). Furthermore while an effective inactivation tool, as a potential delivery system for foreign DNA intron based techniques have been shown to be highly limited with respect to the size of the cargo they are effectively able to carry (Klein et al. 2004).

1.4.2. Allelic exchange tools

Due to the aforementioned limitations in intron inactivation based techniques a significant development in *Clostridium* genetic tools has stemmed from the development of allelic exchange vectors systems that allow for integration into and excision of DNA sequences from chromosome. Due to these techniques requiring multiple recombination events these processes are reliant on the use of negative/counter selection markers in *Clostridium* in order to generate effective tools for the selection of double cross-over events.

1.4.2.1. Negative/Counter selection markers

Due to the limitation of available antibiotics employable in organisms such as *C. difficile* (Sebaihia et al. 2006) and the capability to screen for effective double crossover events, recombination tools require effective negative/counter selection markers. These selection markers have only been more recently developed, with only three main systems based upon *pyrE, codA* and *mazF*, having been demonstrated for their viability in both integration and deletion of chromosomal DNA in *Clostridium* spp. (Ng et al. 2013; Ehsaan et al. 2016; Heap et al. 2012; Al-Hinai et al. 2012).

1.4.2.1.1. *pyrE* based chromosomal modification

In bacteria, *pyrE* encodes the native orotate phosphoribosyl transferase, an enzyme which catalyses the fifth reaction in the *de novo* synthesis of pyrimidine nucleotides and is essential for uracil production (Ghim et al. 1994). As a result of its inactivation, a Δ*pyrE* mutant strain is a uracil auxotroph that is no longer subject to 5-fluoroorotic acid toxicity (Widlund & Davis 2005). The use of uracil deficiency and 5-FOA selection has been used to create a Δ*pyrE* strain of *C. acetobutylicum*, that in place of a clean removal of the gene contains instead a truncated gene through the removal of the final 300bp of the sequence (Heap et al. 2012). This serves to provide the phenotype of a *pyrE* mutant required for counter selection but also provides the basis for larger chromosomal integration techniques (Heap et al. 2012).

The *pyrE* mutant was first exemplified for its basis as a chromosomal integration system by Heap et al. (2012), in a process described as allelic coupled exchange (ACE). During this process large volumes of DNA can be stably integrated into a *pyrE* mutant through the use of homologous recombination at the *pyrE* locus (Figure 1.3), which in turn repairs the *pyrE* truncation and restore its functionality (Heap et al. 2012). The success of this technique is accredited to the use of unevenly sized homology arms, with the forward arm only containing

the 300 bp needed to repair the truncated *pyrE* gene, while the reverse arm contains 1200 bp of the sequence downstream of the *pyrE* gene (Heap et al. 2012). Due to the shorter nature of the forward homology arm single recombination event are more likely to occur and be stable when occurring with the longer homology arm. Therefore selection of a restored *pyrE* gene is more likely to represent successful double cross over mutants as opposed to single cross-over events (Heap et al. 2012).

The Δ*pyrE* mutant also serves as the basis for in-frame deletion techniques based around uracil/5-FOA selection to identify single and double cross over mutants (Figure 1.4). In this in-frame deletion process homology arms flanking a target region are present on a plasmid backbone that contains either an unstable or an absent replicon and an a *pyrE* gene from another *Clostridium* on the backbone (Ng et al. 2013; Ehsaan et al. 2016). Selection for uracil autotrophy identifies single cross over mutations through the activity of the plasmid backbone *pyrE* gene (due to dissimilarity from the native *pryE* repair to the truncation cannot occur) allowing for uracil autotrophy (Ng et al. 2013; Ehsaan et al. 2016). Following this a double cross over mutants can be selected through the use of 5-FOA resulting in either the removal of the plasmid in its original form or following the excision of the native gene on to the unstable plasmid an in-frame deletion (Ng et al. 2013; Ehsaan et al. 2016). After screening to confirm successful in-frame deletion mutants, the generated strains are then able to be subject to further deletions using the same technique or can be complemented using ACE as previously described.

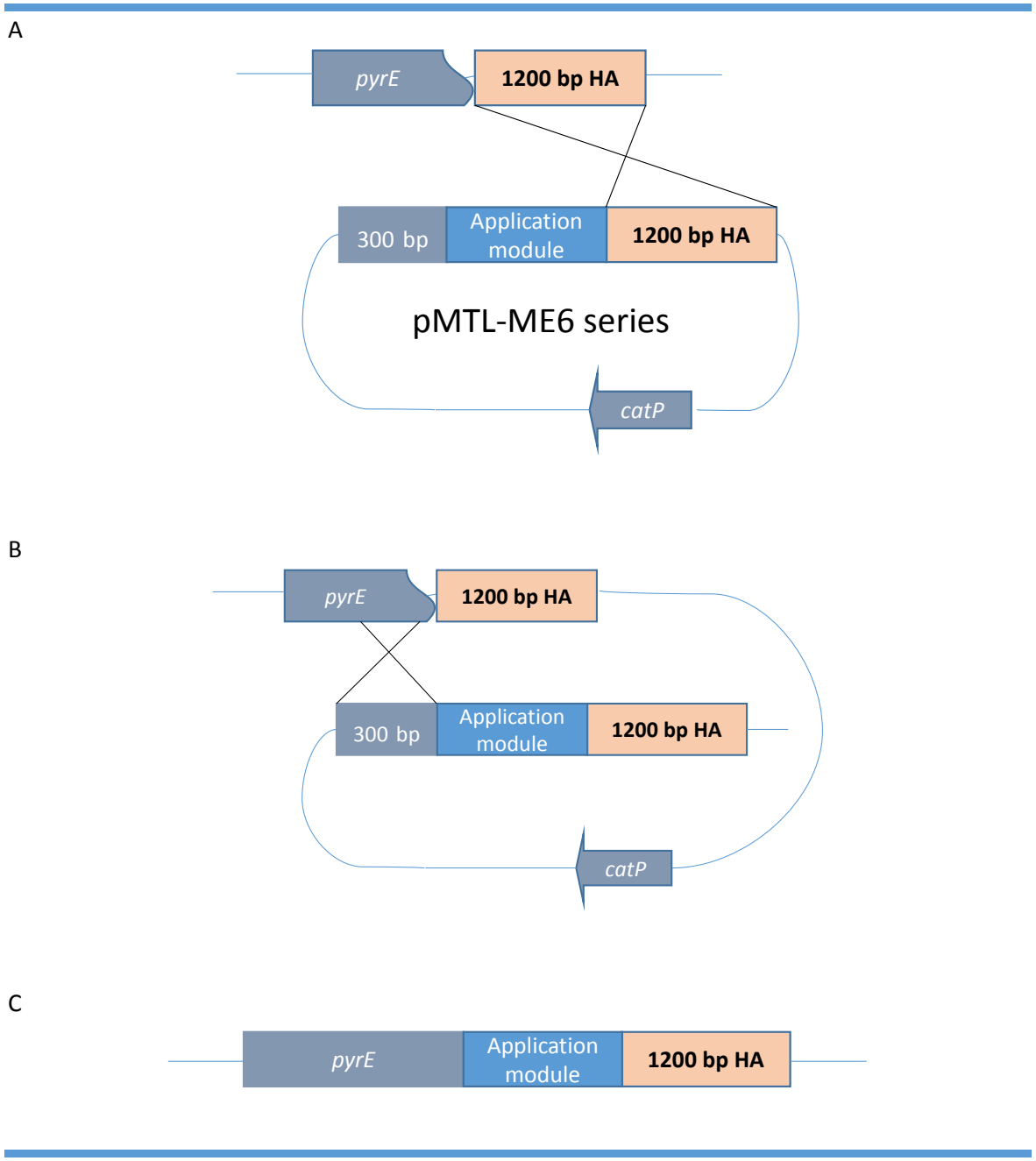


Figure 1.3 Chromosomal insertion and *pyrE* repair using an allelic couple exchange method

A *C. acetobutylicum* mutant containing a *pyrE* truncated by 300 bp is transformed with an allelic coupled exchange vector (pMTL-ME6 series) which contains homology arm for the final 300 bp of the truncated *pyrE* sequence an application module (dependent on the ACE vector used) and a 1200 bp sequence homologous the downstream region of the *pyrE* locus (A). Following selection with thiamphenicol (resistance conferred by *catP* on the ACE vector backbone), single cross-over mutants observed (B) are highly favoured to be as a result of homologous recombination using the longer homology arm. Selection for restored uracil autotrophy is used to select for double cross over mutants, as the *pyrE* locus is repaired by the second recombination event (C). In repairing of the *pyrE* ACE vectors also allow for the stable chromosomal insertion of application modules placed between the 300 bp and 1200 bp homology arms.

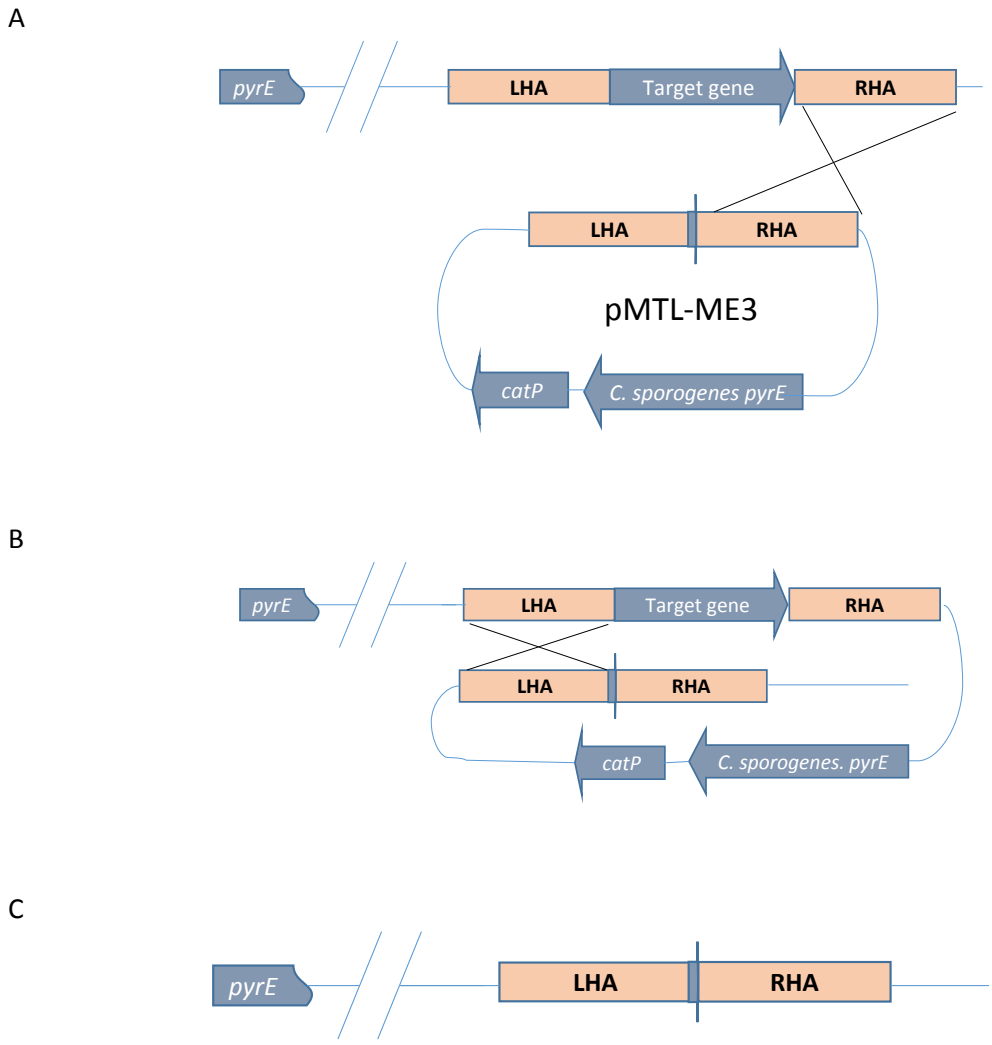


Figure 1.4 The generation of in-frame deletion mutants using *pyrE* selection in *C. acetobutylicum*

(A) An in-frame deletion vector (pMTL-ME3 series) containing 2-3 codons following the start and proceeding the stop codon of the target gene, between a left homology arm (LHA) and a right homology arm (RHA) each of length approximately 750 bp is transformed into a *C. acetobutylicum* Δ*pyrE* mutant. Selection using thiamphenicol resistance and uracil autotrophy (encoded for by a *catP* and a copy of the *C. sporogenes pyrE* gene on the vector backbone) is used to isolate successful single cross-over mutants as in-frame deletion vectors carry an unstable gram positive replicon (B). Following counter selection with 5-FOA a second recombination event using the other homology arm allows for a potential double cross-over event resulting in the substitution of the target gene on the chromosome with the 4-6 codon variant introduced from the in-frame deletion vector (C).

This technique allows for effective chromosomal modification of *Clostridium* with both respects to both deletion and integration of desired sequences out and on to the chromosome. A potential caveat of this negative selection marker is that studies in yeast have suggested 5-FOA to induce chromosomal alterations primarily in the formation of single nucleotide polymorphisms (Wellington et al. 2006). At present this has not been observed in *C. acetobutylicum* Δ*pyrE* mutants. which have shown no evidence of chromosomal alterations compared to their parent strains (Ehsaan et al. 2016).

1.4.2.1.2. *codA* and *mazF* based chromosomal modification

codA is a cytosine deaminase that catalyses the conversion of cytosine to uracil, although the strictness of its substrate specificity is loose enough to facilitate the conversion of 5-fluorocytosine into 5-fluorouracil (O'Donovan & Neuhard 1970). 5-fluorouracil is subsequently highly toxic following its conversion by uracil phosphoribosyltranferase, resulting through a series of steps in the irreversible inhibition of thymidylate synthase and the incorporation of fluorinate nucleotides into DNA and RNA (Longley et al. 2003).

MazF is an mRNA interferase toxin from *E. coli* that in the native organism is neutralised by an anti-toxin MazE (Yamaguchi & Inouye 2009). In the absence of MazE, whose expression is often lost due to its instability during various stresses, MazF activity initiates a programmed cell death pathway (Engelberg-Kulka et al. 2005). The *mazF* system developed by Al-Hinai et al. (2012) relies on placing this toxic gene behind a lactose inducible promoter, reducing MazF toxicity to a minimal level prior to its use in counter selection. Also unlike the previously described *codA* and *pyrE* based deletion techniques this system choses to place an erythromycin resistance cassette flanked by FRT sites between the homology arms to allow for easier selection of favoured double crossover mutants (Al-Hinai et al. 2012).

These techniques have been exemplified as a means to create in-frame deletions in a similar way described using *pyrE* based selection, however due to the lack of a functional *codA/mazF* within the exemplified *Clostridium* the process can be performed using the wild type strains. Selection for the initial single cross over mutant is based upon the speed of growth of antibiotic resistant colonies due to loss of the burden of plasmid replication (Al-Hinai et al. 2012; Cartman et al. 2012; Ehsaan et al. 2016). As with *pyrE*-based modifications, double crossovers are isolated through counter selection, either the provision of 5-FC for *codA* or through induction of *mazF* expression in order to facilitate removal of the plasmid backbone. In the *mazF* system the substitution of the desired gene for an erythromycin resistant cassette allows for the specific selection of double mutants that have undergone the desired double crossover recombination event (Al-Hinai et al. 2012). The *codA* technique, as was described with the *pyrE* based deletions, requires additional PCR screening to distinguish wild type reverted colonies (Cartman et al. 2012; Ehsaan et al. 2016). To facilitate additional chromosomal modification the erythromycin cassette inserted by *mazF* selection requires removal; the resistant marker can be subsequently flipped out by an Flp/FRT system (Al-Hinai et al. 2012).

Unlike the *pyrE* based system both these chromosomal techniques can be performed within the wild type and can be used to successfully generate complementation strains. They have however not been demonstrated for the capacity to introduce large segments of DNA on to the chromosome as has been exemplified by ACE.

1.4.3. CRISPR based gene deletion

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISP-associated (Cas) proteins form the equivalent of an adaptive immune response in bacteria and archaea to foreign nucleic acids, primarily that of bacteriophages (Sorek et al. 2013). Between the short CRISPR repeat sequence of an organism are short DNA sequence that are responsible for

determining CRISPR targeting (Sorek et al. 2013). Upon translation of the CRISPR array, the subsequent CRISPR-RNA is able to direct Cas proteins to target sites resulting in the cleaving of the target sequence (Chylinski et al. 2013). CRISPR-Cas systems show a high diversity in both their evolution and mechanism of action, with systems such as Cas9 only requiring a single effector module for RNA binding and target cleavage, while systems such as Cas3 require the action of multiple subunits (Makarova et al. 2015).

Through CRISPR-Cas systems are largely still being developed for use in all *Clostridium* spp. as some organisms lack the supporting apparatus for certain Cas proteins, the use of CRISPR-Cas has been shown as an effective genome editing tool in both *C. beijerinckii* and *C. pasteurianum* (Pyne et al. 2016; Y. Wang et al. 2015). Through the expression of Cas proteins and targeted CRISPR arrays it has been shown that a CRISPR-Cas system can act as an effective counter selection marker against a sequence targeted for deletion resulting in significantly increased levels of successful double crossover recombination event (Wang et al. 2015; Pyne et al. 2016).

Though not as universally applicable as other genome editing tools due to the varying compatibility of an individual CRISPR-Cas system with specific organisms, CRISPR-Cas systems appear to overcome a large portion of the difficulties of associated with isolating double cross over mutants using established methods and will likely be an important genetic tool in the future manipulation of solventogenic *Clostridium*.

1.4.4. Transposon Mutagenesis

Alongside the described reverse genetic techniques, developments have been made towards the generation of effective forward genetic tools through the means of universal *Clostridium* transposon systems. These systems are based upon a mariner based transposon, Himar1, which

is a mobile genetic element able to facilitate random gene disruption across a genome with no supplementary elements required (Lampe et al. 1996).

The use of a mariner transposon in *C. difficile* was first exemplified by Cartman & Minton (2010), showing the effective random transposon mutagenesis across the organisms genome. This original system, like *pyrE* and *codA* in-frame deletion techniques, was dependent on the loss of a pseudo-suicide plasmid, the replicon instability of which eventually resulted in the loss of free plasmid (Cartman & Minton 2010). Though effective this system was deemed undesirable due to the serial passage needed to facilitate the loss of the transposon expressing plasmid and the subsequent risk of multiple transposon insertions (Zhang et al. 2015). A newer version of this system has been exemplified in both *C. acetobutylicum* and *C. sporogenes* that overcomes this issue through the use of IPTG inducible plasmid loss, the result of increased transcription of the pCB102 replicon causing a near total loss in its stability (Zhang et al. 2015).

Previous attempts at transposon systems in *Clostridia* have often been limited by complications such as organism specific delivery methods (Lanckriet et al. 2009) or integration site bias (Wang et al. 2006). The recent establishment of forward genetic techniques provide a far greater opportunity for the discovery of genes of interest in solventogenic *Clostridia* not previously identified/characterised by reverse genetic approaches.

1.5. Quorum Sensing

1.5.1. Introduction

Quorum sensing is a process by which diffusible cell-to-cell signal molecules are produced by an organism and subsequently accumulate within the environment; upon reaching a threshold level a subsequent response is activated through a change in transcriptional regulation throughout the culture as a result of sufficient signal recognition (Ramsey et al. 2009). Due to the necessity of signal accumulation to a threshold level in order to activate the associated

response, quorum sensing is described as being cell-to-cell signalling density dependent as a certain number of signal producing cells is usually required to reach signal threshold concentrations (Fuqua et al. 1994).

Quorum sensing is distinct from simple detection and response to extracellular metabolites due to its reliance of specific cell-to-cell signal molecules (also termed autoinducers). Autoinducers are distinguished from other signal response mechanism by being: produced during specific phases of growth/environmental conditions/physiological conditions, their extracellular accumulation and recognition by specific receptors, inducing a response upon reaching a critical threshold and said response extending beyond physiological changes required to metabolise or detoxify the signal molecule (Winzer et al. 2002).

It has been proposed that quorum sensing is likely influenced or can be entirely explained by general diffusion sensing to determine the threshold level needed to be reached in a particular culture (Redfield 2002). It is therefore necessary to distinguish quorum sensing from other diffusion sensing phenomena such as compartment sensing, in which signals do not diffuse far from the producing cell, allowing them accumulate and lead to a response (Winzer et al. 2002). Quorum sensing has been suggested to be best considered a form of diffusion sensing though is distinguished by its requirement for high enough cell densities to produce a signal molecule, as a single cell should be unable to achieve a threshold level alone (Williams et al. 2007).

1.5.2. The importance of quorum sensing regulation

Quorum sensing serves as a mechanism by which bacterial populations are able to exert a coordinated response primarily based upon their density, this can serve many different purposes aiding the population's survival. The phenotypes regulated by quorum sensing are often considered to be achievable or beneficial only when performed by larger populations but not individual cells (Darch et al. 2012). In the case of pathogenic organisms this often takes the

form of coordinated regulation of virulence factors and biofilm formation (de Kievit & Iglewski 2000; De Kievit 2009) increasing survival through the generation of more difficult to clear infections (Costerton et al. 1999). In other organisms quorum sensing has been suggested to aid survival through the increased regulation of sporulation and may serve to increase cell competence and survival under stressful conditions, often arising as a result of high cell densities (Schultz et al. 2009; Jabbari et al. 2013).

Due to the role of quorum sensing in aiding bacterial survival by mediating a coordinated response, inhibition of quorum sensing mechanism is a growing area of research with respect to clearing resilient bacterial infections (LoVetri & Madhyastha 2010; Hentzer & Givskov 2003; Otto 2004). The interference/exploitation of quorum sensing can be seen abundantly through nature as similar bacteria have been shown to provide variant signals molecules for cross inhibition of quorum sensing (Otto et al. 2001), or degrade the signals produced by other species (Chen et al. 2013). Even eukaryotes have shown the ability to both create 'mimic' signals to reduce activity and recognise some quorum sensing signals and shift their behaviours in response in order to exploit quorum sensing bacteria (Bauer & Mathesius 2004).

Population-wide activities such as those coordinated by quorum sensing can be subject to social cheating (Diggle et al. 2007). For instance in *Pseudomonas aeruginosa* the emergence of quorum sensing receptor deficient mutants have been shown to arise and benefit from the activity of the bacterial populations (Jiricny et al. 2014). Free of the metabolic burden placed upon them by quorum sensing based expression, these quorum sensing mutants are shown to be able to better compete with the wild type at the cost of attenuated virulence (Wilder et al. 2011;Sandoz et al. 2007). However in *P. aeruginosa* quorum sensing has also been proposed as a means of policing the emergence of social cheaters through the regulation of cyanide

production which usually kill *lasR* deficient mutant strains at a faster rate than quorum sensing cooperating cells (Wang et al. 2015).

In summary quorum sensing regulation has been shown to be a key mechanism by which a range of bacteria control factors to increase survival and competence in response to their increased cell densities. Conversely quorum sensing is of importance due to the means by which is can be inhibited or exploited by other organism or social cheaters within a culture (or may be used to govern social cheaters) and therefore remains a field of interest with respect to organisms showing interesting cell density dependent phenotypes.

1.5.3. Signal molecules used in quorum sensing

Quorum sensing was the term used to originally describe the cell-to-cell communication controlled regulation of gene expression by acyl-homoserine lactone signals (Fuqua et al. 2001), though overtime numerous more cell-to-cell signal molecules have been identified. In Gram-negative bacteria 4-quinolones, fatty acids and fatty acid methyl esters have also been identified as quorum sensing signal molecules, though due to their chemically diverse nature many more may exist (Williams et al. 2007). By contrast gram positive bacteria have been shown to almost exclusively use auto inducing peptides (AIPs), of varying sizes and structure as their quorum sensing signalling molecules (Williams et al. 2007).

1.5.4. Gram negative quorum sensing

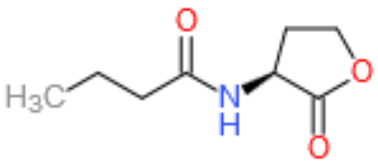
1.5.4.1. Acyl-homoserine lactone quorum sensing

When the N-(3-oxohexanoyl) homoserine lactone was first isolated from *V. fischeri* as an autoinducer of bioluminescence (Eberhard et al. 1981) it was originally believed that acylhomoserine lactone (AHL) mediated signalling was unique to this organism. However following the discovery of 3-oxo-C6-HSL as a quorum sensing signal regulating the

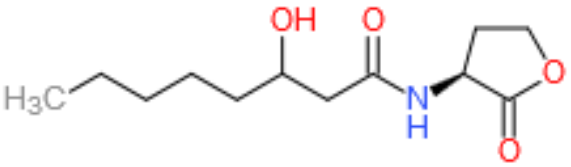
biosynthesis of carbapenem in *Erwinia carotovora* (Bainton et al. 1992), AHL quorum sensing has been observed in a diverse range of Gram-negative proteobacteria of the subdivisions α, β, and γ (Chhabra et al. 2005; Williams et al. 2007).

AHL signalling is characterised by a homoserine lactone ring which is *N*-acylated with a fatty acyl chain of varying length at the α position (Williams et al. 2007) with modifications often performed on the C3 of the chain (see Figure 1.5). This side chain has not been observed to be shorter than a C4 chain likely due to ring instability at lower pH's, while the largest AHLs discovered show up to C18 chains (Cao & Meighens 1989; Wagner-Döbler et al. 2005; Yates et al. 2002). These chains also show a tendency toward the inclusion of an even number of carbons, although both C5 and C7 exceptions have been identified (Lithgow et al. 2000; Horng et al. 2002).

A C4-HSL



B 3-hydroxy-C8-HSL



C 3-oxo-C12-HSL

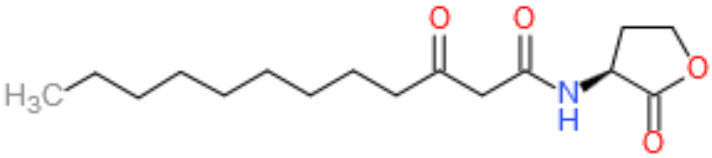


Figure 1. 5 Examples of AHL quorum sensing molecules used in gram negative quorum sensing

AHL's contain a conserved homoserine lactone ring which is N-acetylated with a fatty acid chain at the α position. In the case of the the C4-HSL (A) signalling molecule utilised by organisms such as *Aeromonas hydrophila* this is a 4 carbon length chain with no modifications made to the third carbon of the chain (Kirke et al. 2004). The 3-hyrdoxy-C8-HSL (B) found in organisms such as *Burkholderia pseudomallei* (Lumjiaktase et al. 2017) demonstrates a longer carbon chain and the addition of a hydroxyl group to the third carbon of the chain. 3-oxo-C12-HSL (C) produced by organisms such as *P. aeruginosa* shows a double bond to an oxygen at the third carbon and a fatty acid chain extending as up to 12 carbons long to function as its LasI signalling molecule (Williams et al. 2007).

Synthesis of AHLs is typically through the activity of homologues of the LuxI protein (Williams et al. 2007) though it has been shown that AHL synthesis can occur through the activity of LuxM homologues (Bassler et al. 1993). In addition a third potential independent AHL synthase HdtS belonging a family unrelated to any Lux proteins has been identified (Laue

et al. 2000). AHLs are recognised by LuxR homologues and are subsequently responsible for changes in expression of multiple target genes, often including the corresponding AHL synthase in order to create a positive feedback loop (Williams et al. 2007).

Organisms often carry multiple LuxR/LuxI systems each producing and recognising their own specific tail length AHL signal molecule and controlling their own subset of phenotypes and/or form a hierarchy of regulation with one system in turn activating/inactivating another (Williams et al. 2007).

1.5.4.2. Quinolone quorum sensing

Following the characterisation of the AHL quorum sensing signalling pathways in *P. aeruginosa* a third system of cell-to-cell signal was discovered in the form of 2-heptyl-3-hydroxy-4-quinolone (PQS) and its precursor 2-heptyl-4-quinolone (HHQ), which have been shown to act as autoinducers (Pesci et al. 1999; Diggle et al. 2007). HHQ has also been observed as an autoinducer in *Burkholderia pseudomallei* (Diggle et al. 2006).

In *P. aeruginosa* the synthesis of PQS is the result of condensations of anthranilate and β-keto dodecanoate and requires the products of *pqsA-D* genes, which generate a number of 2-alkyl-4-quinolones including HHQ (Lépine et al. 2004). The final component of the operon *pqsE* has an unknown function but has been suggested to play a role in signal response (Gallagher et al. 2002). The quinolone quorum sensing signalling of *P. aeruginosa* serves to fit into the existing regulatory hierarchy of the AHL quorum sensing systems and has been proposed as the means by which the organisms two AHL systems are linked (McKnight et al. 2000).

1.5.4.3. Fatty acid and fatty acid methyl ester quorum sensing

The final two main quorum sensing systems identified in Gram-negative bacteria are based upon fatty acids and fatty acid methyl ester signalling, i.e. the diffusible signal factor (DSF)

and 3-hydroxy palmitic acid methyl ester (3OH-PAME) respectively (Williams et al. 2007). Both these systems rely on a two component system response regulator, a mechanism also seen conserved in many Gram-positive bacterial quorum sensing systems.

The first DSF molecule to be identified was a likely fatty acid derived signal molecule required for the regulation of pathogenicity of *Xanthomonas campestris* (Barber et al. 1997) though it has since been well studied in *Xanthomonas*, *Burkholderia*, *Xylella* and *Stenotrophomonas* (Cheng et al. 2010; Deng et al. 2009; Ryan et al. 2015; Huedo et al. 2015). The DSF signal is synthesised through the activity of the fatty acyl-CoA ligase RpfB and DSF synthase RpfF, while signal recognition is the result of the hybrid histidine kinase receptor and effector RpfC and its cytoplasmic regulator element target RpfG (Ryan et al. 2015).

3OH-PAME was identified as an autoinducer of virulence in the plant pathogen *Ralstonia solanacearum* (Flavier, Clough, et al. 1997), which like fatty acid signal recognition is dependent on a two component histidine kinase response regulator pair PhcS-PhcR, which also serves to regulate AHL quorum sensing (Flavier, Ganova-Raeva, et al. 1997). This system has not been described outside of *Ralstonia solanacearum* however it has been shown that many organisms such as *P. aeruginosa* and *Stenotrophomonas maltophillia* are capable of degrading 3OH-PAME and thus act as 'quorum quenchers' against *R. solanacearum* signalling (Achari & Ramesh 2015).

1.5.5. Gram positive quorum sensing

Quorum sensing in Gram-positive bacteria is mediated through the production of peptide signal molecules and can be generally classified into two categories. Unlike Gram-negative quorum sensing molecules, auto inducing peptides (AIP's) are not variations upon a single core molecule and are transcribed and translated in an unprocessed/unmodified form directly from their encoding genes (Ng & Bassler 2009) and the two main classifcations are distinguished by

the structure of their signal peptides and whether the system requires multiple proteins to form the response regulator (Figure 1.6).

The better characterised of these two categories involves a two-component response regulator system consisting of: a membrane bound receptor that the AIP interacts and a response regulator which is phosphorylated by the former upon AIP binding, leading to changes in transcriptional regulation (Sturme et al. 2002). The signalling peptides used in two component signalling are post-translationally modified to produce distinct structures often containing distinct peptide rings. The other main class of Gram-positive quorum sensing, the so-called RNPP systems, conversely use linear signalling peptides as AIPs, which are matured by basic peptide cleaving and have not yet been observed to be post-translationally modified. These AIPs are able to enter the cell through generic oligopeptide permeases in order to interact with their associated receptor protein.

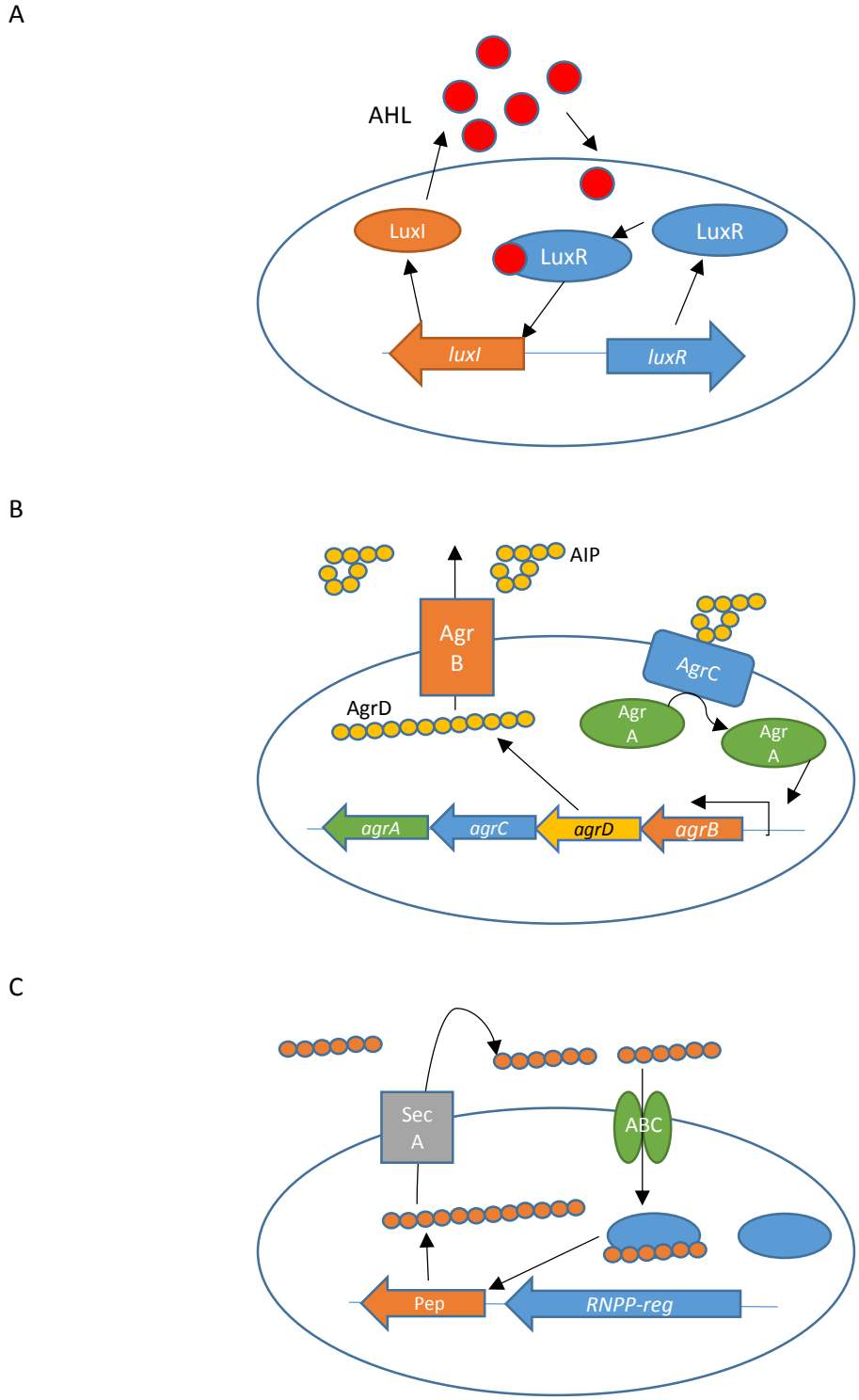


Figure 1.6 Gram positive quorum generation and detection of peptide quorum sensing molecules in contrast to a LuxI/LuxR system found in gram negative bacteria.

Quorum sensing systems rely on the accumulation of an autoinducer both extracellularly and intracellularly using a positive feedback loop to increase expression. In gram positive bacteria this is commonly performed using a

LuxI/LuxR based system (A) in which the LuxI is responsible for generating AHL's which diffuse out of and in to cells and activate the LuxR regulator, which in turn increases *luxI* expression. In gram positive organisms the signal molecule is instead a peptide transcribed from the genome. In two component systems such as the Agr quorum sensing systems (B) a translated peptide forms a circular structure and exported by the activity of AgrB. Upon recognition of this AIP by AgrC, the membrane bound kinase is able to activate AgrA which in turn up regulates the expression of the *agr and* other target operons. In RNPP type systems (C) the peptide signal remains in a linear form and is transported from the cell by SecA dependent transporter systems. The peptide is then further processed by the activity of extracellular proteases before re-entering cells through ABC transporters, after which it can directly interact with its corresponding RNPP regulator, allowing it to directly bind DNA or phosphorylate target genes in order to change gene expression.

The two component response regulator and RNPP systems are involved in a large variety of regulatory processes within gram positive bacteria that contribute towards the fitness of their respective organisms, in particular under more challenging environmental or growth conditions. One such example is the quorum sensing control of natural competence in *Streptococcus pneumoniae* in which a sub fraction of cells within the population undergo lysis in order to facilitate the natural competence based recombination induced by quorum sensing in the remaining population (Steinmoen et al. 2002). The recombination resulting from this process subsequently causes an increased long term fitness of the organisms when grown for extended periods under stress conditions (Engelmoer 2012). Other examples of quorum sensing increased fitness are often associated with the regulation of virulence factors in pathogenic organisms such as *S. aureus, C. botulinum* and *E. faecalis* (Ji et al. 1995; Cooksley et al. 2010; Pinkston et al. 2011). The regulation of virulence factors by quorum sensing may also be used to prioritise selective advantages to virulence as opposed to long term survivability of the organism, as was shown in *B. anthracis* where the presence of a Rap phosphatase system on a virulence plasmid served to down-regulate endospore formation, a process which while increasing survival in challenging environments is postulated as inhibiting the organisms toxin production (Bongiorni et al. 2006).

Alternatively gram positive quorum sensing has also been suggested to play a role with respect to how organisms are able to compete with one another when growing in the same

environments. While this may include interference in the activity of the quorum sensing system of a competing organism (Otto et al. 1999), this also extends to the production of antimicrobials. Outside of bacteriocin based quorum sensing (discussed in 1.5.5.1.1) other products toxic to bacteria, such as cytolysin produced by *E. faecalis* have been shown to be under the control of quorum sensing system (Haas et al. 2002). Interestingly while antimicrobials such as bacteriocins are largely shown to be up regulated by their associated quorum sensing systems (Ramsey et al. 2009) the production of cytolysin is in fact down regulated a two component system suggesting that the organism selectively produces the toxin under a low cell density (potentially due to the influence of a competing organism) and subsequently halt the production when a high cell density population is established (Haas et al. 2002).

The regulation by quorum sensing in gram positive organisms commonly targets genes associated with survival (such as natural competence or endospore formation) and virulence/toxin production. However the role of the quorum sensing systems can vary greatly as examples above have shown some organisms to use quorum sensing system to selectively express or inhibit these systems as is deemed favourable to the respective organism favoured phenotypes or growth conditions.

1.5.5.1. Two component response regulator quorum sensing

Though acting through similar two component response regulator systems, the AIPs used by Gram-positive bacteria shows a significantly larger level of diversity and specificity when compared to the signals employed by Gram-negative bacteria. AIPs have been show to vary in size between 5 to 34 amino acids and often shown unusual structure based upon post translational modification (Williams et al. 2007). Included amongst AIPs that rely on two

component signalling are bacteriocins, isoprenylated tryptophan peptides and lactone or thiolactone ring based AIPs.

1.5.5.1.1. Bacteriocin quorum sensing

Bacteriocins such as nisin produced by *Lactococcus lactis* are a class of peptide antibiotic produced in order to inhibit the growth of competing bacteria, although some of them also play a secondary role as a quorum sensing AIPs (Ramsey et al. 2009). Bacteriocin AIPs are generally exported through ABC-transporters and are matured to an their active signalling structures by cleaving and potentially further post-translational modification (Heng et al. 2007).

Bacteriocins are generally classified by their size and mode of action into one of four groups. Class I are known as the lantibiotics and are post translationally modified, class II are small (≤10kDa) heat stable membrane-active bacteriocins, class III are larger (≥30kDa) and are heat label and class IV are complex bacterioicins which contain additional lipid or carbohydrate moieties (Heng et al. 2007). Of these classes it is the lantibiotics which have provided examples of quorum sensing regulation in addition to their established antimicrobial roles.

In a number of examples, lantibiotic synthesis has been shown to be growth-phase dependent and under the control of a conserved two-component signal transduction system consisting of a membrane bound receptor histidine kinase and a response regulator that becomes phosphorylated on a conserved aspartate residue leading to transcriptional activation (Chatterjee et al. 2005). The genes encoding the two component response regulator are commonly found in clusters alongside those encoding the lantibiotic precursor, the enzymes responsible for its modification, transport and processing of the lantibiotic, in addition to other conserved genes of unknown function (Chatterjee et al. 2005). The two component regulation system has generally been observed to form as a positive feedback loop to increase the

expression of the respective lantibiotic (Chatterjee et al. 2005). However this is not universally conserved across lantibiotic synthesis as some organism lack the two component regulatory system in their lantibiotic gene clusters, while others have shown repression of lantibiotic production in response to the activation of their two component system (Chatterjee et al. 2005).

1.5.5.1.2. Isoprenylated tryptophan peptide quorum sensing

Bacillus subtilis uses a two component system ComP and ComA, in which ComP is the membrane bound histidine kinase and ComA a cytoplasmic response regulator, in order for regulate natural competence (Dubnau et al. 1994). The ComX-derived AIP responsible to activation of this two component systems has been shown to be highly polymorphic between *B. subtilis* strains, with many clear examples of cross inhibition and activation profiles seen between different signals and corresponding response regulators (Ansaldi et al. 2002).

These AIP structures are general 7-10 amino acids in size and appear to vary with respect to post translational modifications of their backbone structures (Ansaldi et al. 2002). Conserved across the ComX AIP structures is a conserved tryptophan residue that shows the addition of a geranyl group resulting in the formation of a ring structure shown to be essential for signal recognition (Okada et al. 2005)

1.5.5.1.3. Lactone and thiolactone-containing AIP quorum sensing

Lactone and thiolactone containing AIP quorum sensing systems are the most extensively studied across Gram-positive bacteria. These includes the accessory gene regulator (*agr*) systems of *Staphylococcus aureus* and its homologues, such as the *Enterococcus faecalis fsr* system, which are present throughout the firmicutes (Qin et al. 2001; Wuster & Babu 2008). Though the respective AIPs observed across the system are highly variable with respect to peptide sequence and signal size, *agr* homologues are characterised by the inclusion of lactone

or thiolactone bond (dependent on the inclusion of either conserved cysteine or serine residues) resulting in the formation of a ring structure (Nakayama et al. 2001; McDowell et al. 2001).

The archetype organism for Agr quorum sensing is *Staphylococcus aureus*, in which the system was initially described for the regulation of the organisms accessory and virulence genes, with most homologous systems in other organisms named after it (Yarwood & Schlievert 2003; Wuster & Babu 2008). In *S. aureus* the *agr* cluster is characterised as containing two operons, the *agrBDCA* and RNA III transcripts, each driven by two divergent promoters P2 and P3 (Koenig et al. 2004). While this is the archetype, outside of *S. aureus* differences in the cluster have been observed to include the separation of the *agrBD* into its own operon and the exclusion of the RNA III transcript (Steiner et al. 2012; Martin et al. 2013; Cook & Federle 2014).

Of the genes in the cluster the *agrD* provides the unmodified peptide sequence which is both processed and exported by the AgrB membrane bound enzyme to produce an AIP with the associated lactone/thiolactone ring, the tail sequence may be subject to further cleaving by extracellular proteases to form the mature AIP (Ji et al. 1995). AgrC serves as the sensory kinase which responds the AgrD derived signal and in turn phosphorylates AgrA (Wuster & Babu 2008). AgrA is responsible for DNA binding resulting in the up regulation of the P2 and P3 transcripts (Koenig et al. 2004). RNA III is a regulatory RNA responsible for regulating the expression of a large number of genes within *S. aureus* through interaction with their mRNA's (Boisset et al. 2007; Chevalier et al. 2010) although a subset of genes has also been shown to be regulated directly by phosphorylated AgrA (Queck et al. 2008). In Agr systems outside of *S. aureus* where no RNA III is present the alteration of gene expression is likely under the direct regulation of AgrA.

1.5.5.1.3.1. Agr quorum sensing in *Clostridia*

While evidence exists for the wide distribution of Agr-like quorum sensing systems in *Clostridium* spp. (Wuster & Babu 2008), examination of these systems has mostly been focused upon their role in pathogenic strains *C. difficile, C. botulinum* and *C. perfringens* (Ohtani et al. 2009; Martin et al. 2013b; Cooksley et al. 2010). Outside of pathogenic strains, Agr quorum sensing has only been identified and characterised in *C. acetobutylicum* and has yet to be investigated in any other solventogenic strain (Steiner et al. 2012).

In previous studies multiple virulence factors including toxin expression in *C. difficile* (Martin et al. 2013; Darkoh et al. 2015), neurotoxin production in *C. botulinum* (Cooksley et al. 2010) and theta/alpha and kappa –toxins within *C. perfringens* (Ohtani et al. 2009), have all been shown to be regulated by their respective Agr systems. Outside of the expression of virulence factors, most Agr systems studied within *Clostridium* spp. have been shown to regulate the formation of endospores including within *C. perfringens, C. botulinum* (though more significantly through a different AgrBD signalling system than that responsible for neurotoxin regulation) and *C. acetobutylicum* (J. Li et al. 2011; Cooksley et al. 2010; Steiner et al. 2012). While sporulation and virulence factors (particular toxin production) has been shown to be largely conserved with respect to regulation by Agr systems, outside of these phenotypes both *C. difficile* and *C. acetobutylicum* show regulation of flagella assembly and granulose accumulation, respectively to also be under the control of each of their Agr systems (Martin et al. 2013; Steiner et al. 2012).

The diversity within proposed AIP structures also varies greatly between the studied organisms as *C. acetobutylicum* was described as showing maximum activity with a six membered ring while *C. perfringens* has shown maximum activity with a 5 membered ring (Steiner et al. 2012; Ma et al. 2015). Furthermore the recent study by Ma et al. (2015) showed that the *C.*

perfringens cyclic AIP could be substituted by a 5-amino acid linear version and still show biological activity (despite the absence of a thiolactone ring structure). To emphasise the level of variation observed in the Agr quorum sensing of *Clostridium,* it should be noted that in terms of its amino acid sequence the *C. acetobutylicum agrBD* locus clusters together with homologues found in *Listeriaceae* as opposed to those present in other *Clostridium sensu stricto*, potentially suggesting horizontal gene transfer between the different genera (Wuster & Babu 2008) as opposed to shared evolution with other *Clostridium* spp.

Overall Agr quorum sensing has yet to be well characterised outside a small number of *Clostridium* species of interest. While characterisation has identified key phenotypes controlled by these Agr systems, this has yet to be explored to the level of the more archetypal Agr quorum sensing organisms. With the suggestion that some late phase processes such as solventogenesis may be linked to a quorum sensing like regulation in some *Clostridium* strains of interest (Kosaka et al. 2007), Agr quorum sensing may be found to play a significant role upon further exploration of regulation and physiology of these solventogenic strains.

1.5.5.2. RNPP quorum sensing

The RNPP group, named for the Rap histidine kinases found in *Bacillus*, the NprR regulator of neutral proteases in *B. subtilis*, the sex pheromone receptor PrgX of *E. faecalis* and the PlcR transcription regulator of the *Bacillus cereus* group (Rocha-Estrada et al. 2010), are a family of structurally similar quorum sensing regulators whose associated AIPs are short linear peptides. Regulator and peptide encoding genes are often co-located with the peptide genes residing downstream of their associated RNPP response regulator gene. The encoded precursor peptides are processed by generic peptide cleaving enzymes either within the cell or the external environment and are able to enter cell through generic peptide oligopermease structures (Rocha-Estrada et al. 2010).

Unlike two component systems an RNPP regulator is directly responsible for both binding the AIP peptide and acting as the response regulator (Rocha-Estrada et al. 2010). Binding occurs through the C-terminal tetratricopeptide repeat domain (TPR) and allows activation of either a N-terminal phosphatase/kinase domain or causes a conformational shift to allow an N-terminal helix-turn-helix domain to bind to directly DNA sequences (Zouhir et al. 2013a; Rocha-Estrada et al. 2010). The activation of the phosphatase/kinase activity of Rap proteins has been shown to directly feed into the phosphorelay of *Bacillus* spp. allowing for the regulation of sporulation, while the DNA binding regulators have been shown to associate with a number of phenotypes across the studied organisms including virulence, sporulation and conjugative plasmid transfer (Rocha-Estrada et al. 2010; Kozlowicz et al. 2004; Cook & Federle 2014).

Based upon the sequence and structure of Rap proteins, putative RNPP-like quorum sensing systems have been identified within multiple *Bacillus* spp. but have also been suggested to be widely distributed throughout *Clostridium*, though showing the helix-turn-helix N-terminal motif in place of the phosphatase/kinase domain (Rocha-Estrada et al. 2010).

1.5.5.2.1. RNPP quorum sensing in *Clostridia*

At the time of writing only two studies of RNPP-like systems have been performed within the family *Clostridiaceae* and with only one of these within a *Clostridium sensu stricto*. In *C. difficile* two putative RNPP regulators CD2121 and CD3668 were identified by Edwards et al. (2016), with the latter of the two genes being shown to be a positive regulator of sporulation and a negative regulator of toxin production and motility through the regulation of σ^D expression.

With respect to solventogenic *Clostridium* spp. 10 putative RNPP-like receptor gene were identified within *C. acetobutylicum* (Kotte 2013), 8 of which were co-located with a probable signal peptide, while the remaining two were transcribed in the same operon. Of these 10

systems the Ca_c0324 was shown to be a true quorum sensing system that was responsible for repression of solvent production unless in the presence of the AIP derived from the co-located peptide encoding gene (Kotte 2013). Inactivation of the Ca_c0957 and Ca_c0958 RNPP-like regulator genes revealed them to be responsible for a sporulation, Agr quorum sensing, and solvent production as all these phenotypes were lost in the mutant strains (Kotte 2013). These phenotypic observations resembled those observed upon the inactivation of the master regulator of sporulation and solvent production *spo0A* (Ravagnani et al. 2000; Latonia M. Harris et al. 2002). Due to the importance of RNPP regulators in the regulation of sporulation of *Bacillus* (Pottathil & Lazazzera 2003; Perchat et al. 2011) and the identification of RNPP like regulators within *C. acetobutylicum* that play a key role in stationary phase regulation (Kotte 2013), RNPP quorum sensing represent a key area of interest in understanding and manipulating the physiology of solventogenic *Clostridium* for future strain development.

1.6. **Aims and Objectives**

The fundamental objective of this work was to build upon the initial previous knowledge of quorum sensing systems in *C. acetobutylicum* in order to gain a better understanding of: the mechanisms through which the systems function, interactions that may occur between the different regulatory systems and understand what limitation/potential they may provide for future strain exploitation.

C. acetobutylicum and other solventogenic *Clostridium* spp. have previously had importance in industrial biobutanol production and have recently seen a resurgence in interest with the increased availability of genetic tools. However largely unexplored are factors that control the master regulation of the switch from acid to solvent production and the resultant limitation on the time taken to enter solventogenesis in industrial cultures (particularly in large fed-batch or continuous cultures).

Recent evidence has shown that the Agr system plays a role in both sporulation and granulose accumulation of *C. acetobutylicum* though it has not been explored in any greater detail as to the mechanisms through which this is done. Furthermore previous work has been limited as it has been performed in more energy rich conditions in which the accumulation of granulose may be less relevant with regards to the associated phenotypes when excess glucose serves to counterbalance the effect of changes in granulose accumulation.

RNPP systems have also been identified as potential master regulators of sporulation, solvent production and granulose accumulation through the generation of insertional inactivation mutants. While this work has shown strong evidence of Ca_c0957 and Ca_c0958 as master regulators of these processes and suggested an associated with Spo0A regulation this work requires confirming through the generation of clean in-frame deletion mutants, which can also serve as the basis for future work characterising these regulators. Due to the lack of clear signal peptide these systems should be considered putative quorum sensing systems and should be investigated for the presence of a binding peptide (which can subsequently be identified on the genomes). Finally while Ca_c0957/Ca_c0958 are prime candidates for being master regulators of stationary phase phenotypes, the extent to which their effects are mediated by controlling Spo0A phosphorylation, for instance through histidine kinases or other early phase regulators such as the proposed σ^k, and the genes which they direct control remain in need of clarification, as is highlighted in Figure 1.7.

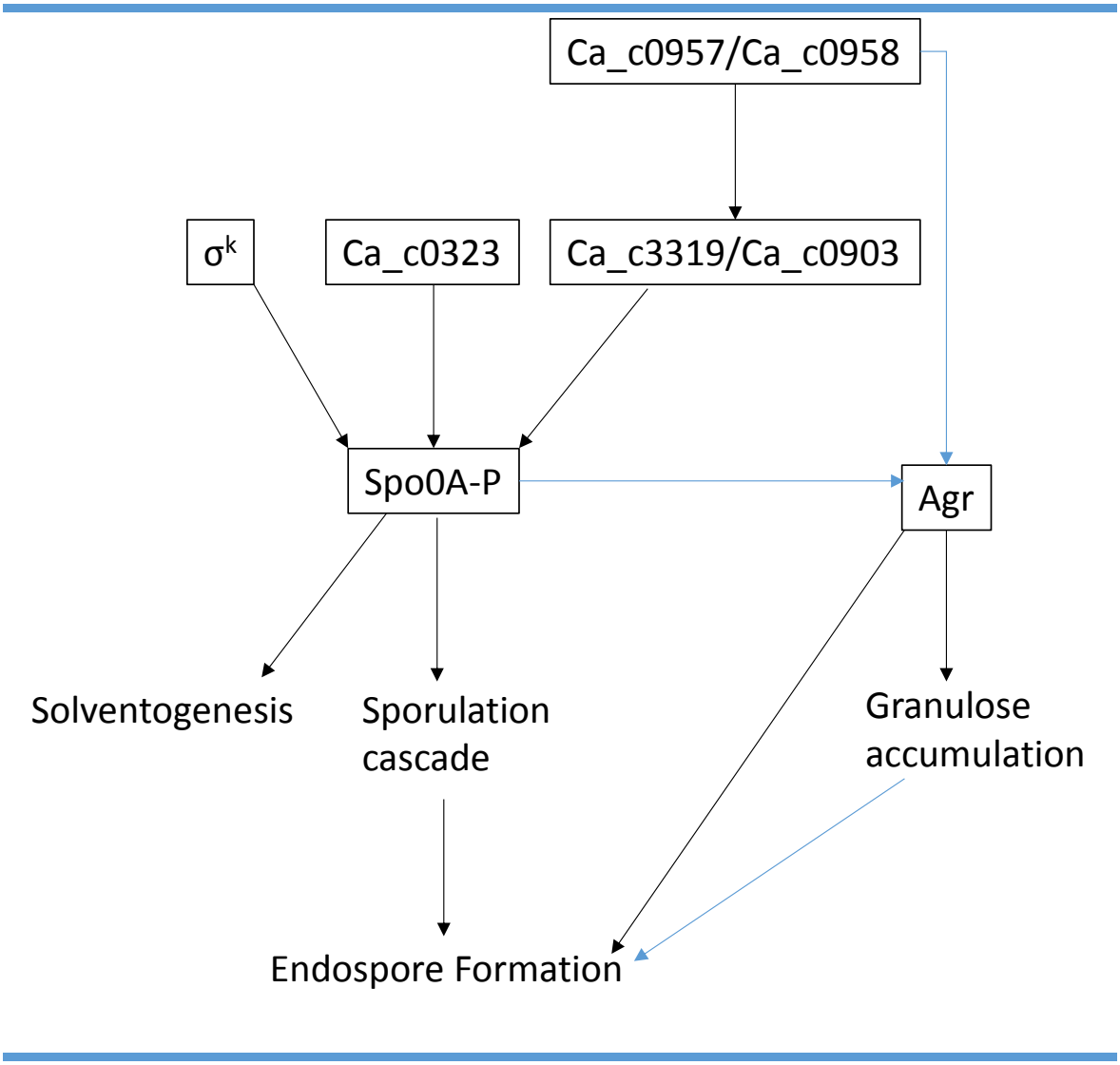


Figure 1. 7 The proposed regulator pathway of stationary phase phenotypes in *C. acetobutlicum* prior to this study

Spo0A is known to be the master regulator of both the sporulation cascade resulting in endospore formation and solventogenesis (Ravagnani et al. 2000) and is activated once phosphorylated. Phosphorylation of Spo0A is performed by orphan kinases Ca_c0323, Ca_c3319 and Ca_c0903 (Steiner et al. 2011), in addition the expression of Spo0A is suggested to be regulated by σ^k (Al-Hinai et al. 2014). The RNPP-like regulators Ca_c0957/Ca_c0958 have been shown to control the expression of Ca_c3319 and Ca_c0903 and show a phenotype matching a Spo0A mutant (Kotte 2013). In addition to this Ca_c0957/Ca_c0958 show the ability to regulate *C. acetobutylicum's* Agr system (Kotte 2013), itself known to regulate granulose accumulation and influence endospore formation (Steiner et al. 2012). It is unknown if Ca_c0957/Ca_c0958 directly regulate *agrBD* expression as suggested by Kotte (2013) or whether this is a downstream effect of the loss of Spo0A activity which directly controls the Agr system. It is also unknown to what extent the influence upon endospore formation by the Agr system is the result of its regulation of granulose or whether it acts through a more direct regulatory mechanism.

Thus the overarching aim to further explore and characterise Agr and RNPP systems in *C. acetobutylicum* for their role in stationary phase phenotypes breaks down to the following specific objectives:

(i) Clarifying the role of the RNPP systems and creation of in-frame deletion strains to further investigate the structural/expression patterns of the regulators.

(ii) Investigating the regulatory hierarchy of the different quorum sensing systems, Spo0A and other potential regulators of stationary phase processes.

(iii) Exploring the mechanisms by which Agr quorum sensing is effectively controlled/inhibited by the use of synthetic peptides and the relative importance of its phenotypes under conditions more varied glucose rich/poor culture conditions.

2. Materials and Methods

2.1. Materials

Sigma-Aldrich was used for the supply of chemicals and biochemical used in this study unless otherwise specified.

2.2. List of bacterial strains, plasmids, oligonucleotides and peptides

The bacterial strains of wild type and genetically modified *E. coli* and *C. acetobutylicum* are listed in Table 2.1, unmodified and engineered plasmids are listed in Table 2.2 oligonucleotides in Table 2.3 and peptides in Table 2.4.

Table 2.1 Bacterial strains used in this study

Strain	Relevant properties	Source
E. coli TOP10	*F- mrcA* Δ*(mrr-hsd*RMS0*mcr*BC) Φ80*lacz*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen
E. coli TOP10 pAN2	*E. coli* TOP10 with the pAN2 methylation plasmid	(Heap et al. 2007)
E. coli BL 21 Star (DE3)	F-*ompT hsdSB(rB-mB-) gal dcm rne 131* (DE3) - Used for expression of T7 promoter based expression systems of pET vectors	Themo Scientific
E. coli BL 21 Star (DE3) pET21d	*E. coli* BL 21 Star (DE3) with a control pET21d expression plasmid	This Study
E. coli BL 21 Star (DE3) pET21d Ca_c0957 His	*E. coli* BL 21 Star (DE3) used for the overexpression of histidine tagged Ca_c0957	This Study
E. coli BL 21 Star (DE3) pET21d Ca_c0958 His	*E. coli* BL 21 Star (DE3) used for the overexpression of histidine tagged Ca_c0958	This Study

E. coli BL 21 Star (DE3) pET21d Ca_c0958 Strep	*E. coli* BL 21 Star (DE3) used for the overexpression of streptavidin tagged Ca_c0957	This Study
E. coli BL 21 Star (DE3) pET21d Ca_c0957 Strep	*E. coli* BL 21 Star (DE3) used for the overexpression of streptavidin tagged Ca_c0960	This Study
C. acetobutylicum ATCC 824 SysMO	*C. acetobutylicum* wild type derived from a spore stock generated by the Bahl laboratory at the University of Rostock; principle strain used for systems biology work by the trans-European COSMIC I and II consortium	Bahl laboratory University Rostock
*C. acetobutylicum agrB::*CT*ermB*	*C. acetobutylicum* ATCC 824 *agrB* ClosTron mutant	(Steiner et al. 2012)
*C. acetobutylicum agrC::*CT*ermB*	*C. acetobutylicum* ATCC 824 *agrC* ClosTron mutant	(Steiner et al. 2012)
C. acetobutylicum Δ*glgA* pMTL-ME6	*C. acetobutylicum* granulose synthase in-frame deletion mutant with *pyrE* locus repaired using pMTL-ME6	(Ehsaan et al. 2016)
C. acetobutylicum Δ*spo0A* pMTL-ME6	*C. acetobutylicum spo0A* in-frame deletion mutant with pyrE locus repaired using pMTL-ME6	(Ehsaan et al. 2016)
C. acetobutylicum Δ*spo0AΔpyrE*	*C. acetobutylicum spo0A* in-frame deletion mutant with a truncated *pyrE*	(Ehsaan et al. 2016)
C. acetobutylicum Δ*pyrE*	*C. acetobutylicum* mutant with a truncated *pyrE*	(Ehsaan et al. 2016)
C. acetobutylicum Ca_c0077*::*CT*ermB*	*C. acetobutylicum* Ca_c0077 ClosTron mutant	This Study

C. acetobutylicum Ca_c0082::CT*ermB*	*C. acetobutylicum C*a_c0082 ClosTron mutant	This Study
C. acetobutylicum Ca_c0082::CT*ermB* pMTL85141 Ca_c0082	*C. acetobutylicum* Ca_c0082 ClosTron mutant with plasmid-based complementation	This Study
C. acetobutylicum Ca_c0082::CT*ermB* pMTL85141 *agrCA*	*C. acetobutylicum* Ca_c0082 ClosTron mutant with plasmid-based expression of *agrCA*	This Study
C. acetobutylicum Δ*pyrE*ΔCa_c0957	*C. acetobutylicum* Ca_c0957 in-frame deletion mutant with a truncated *pyrE*	This Study
C. acetobutylicum ΔCa_c0957	*C. acetobutylicum* Ca_c0957 in-frame deletion mutant with the *pyrE* locus repaired using pMTL-ME6 integration plasmid	This Study
C. acetobutylicum ΔCa_c0957 pMTL-ME6C Ca_c0957	*C. acetobutylicum* Ca_c0957 in-frame deletion mutant with Ca_c0957 complemented at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study
C. acetobutylicum ΔCa_c0957 pMTL-ME6X Ca_c0957	*C. acetobutylicum* Ca_c0957 in-frame deletion mutant with Ca_c0957 overexpressed at the *pyrE* locus using pMTL-ME6X integration plasmid	This Study
C. acetobutylicum ΔCa_c0957 pMTL-ME6C Ca_c0957 His	*C. acetobutylicum* Ca_c0957 in-frame deletion mutant with Ca_c0957 complemented with a histidine tagged fusion at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study
C. acetobutylicum ΔCa_c0957 pMTL-ME6X Ca_c0957 His	*C. acetobutylicum* Ca_c0957 in-frame deletion mutant with Ca_c0957 overexpressed with a histidine tagged fusion at the *pyrE* locus using pMTL-ME6X integration plasmid	This Study

C. acetobutylicum ΔCa_c0957 pMTL-ME6C Ca_c0957 Strep	*C. acetobutylicum* Ca_c0957 in-frame deletion mutant with Ca_c0957 complemented with a streptavidin tagged fusion at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study
C. acetobutylicum ΔCa_c0957 pMTL-ME6X Ca_c0957 Strep	*C. acetobutylicum* Ca_c0957 in-frame deletion mutant with Ca_c0957 overexpressed with a streptavidin tagged fusion at the *pyrE* locus using pMTL-ME6X integration plasmid	This Study
C. acetobutylicum Δ*pyrE*ΔCa_c0957 Δ*spo0A*	*C. acetobutylicum* Δ*pyrE*ΔCa_c0957 mutant with a *spo0A* in-frame deletion	This Study
C. acetobutylicum ΔCa_c0957Δ*spo0A*	*C. acetobutylicum* Δ*pyrE*ΔCa_c0957 mutant with a *spo0A* in-frame deletion with the *pyrE* locus repaired using pMTL-ME6 integration plasmid	This Study
C. acetobutylicum ΔCa_c0957*Δspo0A* pMTL-ME6C spo0A	*C. acetobutylicum* Δ*pyrE*ΔCa_c0957 mutant with a *spo0A* in-frame deletion with *spo0A* complemented at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study
C. acetobutylicum Δ*spo0A*Δ*pyrE* ΔCa_c0957	*C. acetobutylicum* Δ*spo0A*Δ*pyrE with a* Ca_c0957 in-frame deletion	This Study
C. acetobutylicum Δ*spo0A*ΔCa_c0957	*C. acetobutylicum* Δ*spo0AΔpyrE* with a Ca_c0957 in-frame deletion with *pyrE* locus repaired using pMTL-ME6 integration plasmid	This Study
C. acetobutylicum Δ*spo0A*ΔCa_c0957 pMTL-ME6C Ca_c0957	*C. acetobutylicum* Δ*spo0AΔpyrE* with a Ca_c0957 in-frame deletion with Ca_c0957 complemented at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study

C. acetobutylicum Δ*pyrE*ΔCa_c0958	*C. acetobutylicum* Ca_c0958 in-frame deletion mutant with a truncated *pyrE*	This Study
C. acetobutylicum ΔCa_c0958	*C. acetobutylicum* Ca_c0958 in-frame deletion mutant with the *pyrE* locus repaired using pMTL-ME6 integration plasmid	This Study
C. acetobutylicum ΔCa_c0958 pMTL-ME6C Ca_c0957	*C. acetobutylicum* Ca_c0958 in-frame deletion mutant with Ca_c0958 complemented at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study
C. acetobutylicum ΔCa_c0958 pMTL-ME6X Ca_c0957	*C. acetobutylicum* Ca_c0958 in-frame deletion mutant with Ca_c0958 overexpressed at the *pyrE* locus using pMTL-ME6X integration plasmid	This Study
C. acetobutylicum ΔCa_c0958 pMTL-ME6C Ca_c0958 His	*C. acetobutylicum* Ca_c0958 in-frame deletion mutant with Ca_c0958 complemented with a histidine tagged fusion at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study
C. acetobutylicum ΔCa_c0958 pMTL-ME6X Ca_c0958 His	*C. acetobutylicum* Ca_c0958 in-frame deletion mutant with Ca_c0958 overexpressed with a histidine tagged fusion at the *pyrE* locus using pMTL-ME6X integration plasmid	This Study
C. acetobutylicum ΔCa_c0958 pMTL-ME6C Ca_c0958 Strep	*C. acetobutylicum* Ca_c0958 in-frame deletion mutant with Ca_c0958 complemented with a streptavidin tagged fusion at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study
C. acetobutylicum ΔCa_c0958 pMTL-ME6X Ca_c0958 Strep	*C. acetobutylicum* Ca_c0958 in-frame deletion mutant with Ca_c0958 overexpressed with a streptavidin tagged fusion at the *pyrE* locus using pMTL-ME6X integration plasmid	This Study

C. acetobutylicum Δ*pyrE*Δσ^K	*C. acetobutylicum* σ^K in-frame deletion mutant with a truncated *pyrE*	This Study
C. acetobutylicum Δσ^K	*C. acetobutylicum* σ^K in-frame deletion mutant with the *pyrE* locus repaired using pMTL-ME6 integration plasmid	This Study
C. acetobutylicum Δσ^K pMTL-ME6C σ^K	*C. acetobutylicum* σ^K in-frame deletion mutant with σ^K complemented at the *pyrE* locus using pMTL-ME6C integration plasmid	This Study
C. acetobutylicum pMTL85143	*C. acetobutylicum* with an empty overexpression plasmid using the P-*fdx*	This Study
C. acetobutylicum pMTL85143 *nluc* 1	*C. acetobutylicum* containing a plasmid-based overexpression of a codon optimised Nanoluc (version 1)	This Study
C. acetobutylicum pMTL85143 *nluc* 2	*C. acetobutylicum* containing a plasmid-based overexpression of a codon optimised Nanoluc (version 2)	This Study
C. acetobutylicum pMTL85143 *mCherry* 1	*C. acetobutylicum* containing a plasmid-based overexpression of a codon optimised mCherry (version 1)	This Study
C. acetobutylicum pMTL85143 *mCherry* 2	*C. acetobutylicum* containing a plasmid-based overexpression of a codon optimised mCherry (version 2)	This Study

Table 2.2 Plasmids used in this study

Plasmid	Relevant properties	Source
pAN2	Plasmid containing the φ3TI methyltransferase	(Heap et al. 2007)

pET-21d	Protein expression vector containing an IPTG inducible T7 promoter system and including an optional C-terminal Histidine Tag sequence	Novagen
pET21d Ca_c0957 His	Protein expression vector containing an IPTG inducible T7 promoter system, uses the native Histidine tag sequence to create a Ca_c0957 histidine tag fusion	This Study
pET21d Ca_c0958 His	Protein expression vector containing an IPTG inducible T7 promoter, and uses the native Histidine tag sequence to create a Ca_c0958 histidine tag fusion	This Study
pET21d Ca_c0958 Strep	Protein expression vector containing an IPTG inducible T7 promoter system to express a Ca_c0957 streptavidin tag fusion	This Study
pET21d Ca_c0957 Strep	Protein expression vector containing an IPTG inducible T7 promoter system to express a Ca_c0958 streptavidin tag fusion	This Study
pMTL007C-E2::Ca_c0077 716-717s	ClosTron plasmid retargeted for interruption of the sense strand of Ca_c0077	This Study
pMTL007C-E2::Ca_c0077 548-549a	ClosTron plasmid retargeted for interruption of the anti-sense strand of Ca_c0077	This Study
pMTL007C-E2::Ca_c0082 1119-1120s	ClosTron plasmid retargeted for interruption of the sense strand of Ca_c0082	This Study
pMTL007C-E2::Ca_c0082 538-539a	ClosTron plasmid retargeted for interruption of the anti-sense strand of Ca_c0082	This Study

pMTL85141	Clostridium modular plasmid containing a pIM13 Gram-positive replicon, a *catP* selection marker, a *ColE1* gram negative replicon and a *lacZ* multiple cloning site	(Heap et al. 2009)
pMTL85141 Ca_c0082	pMTL85141 with Ca_c0082 and suspected promoter region cloned into the multiple cloning site at NotI and NheI sites	This Study
pMTL85141 *agrCA*	pMTL85141 with *agrCA* and its promoter region cloned into the multiple cloning site	(Steiner et al. 2012)
pMTL85143	Clostridium modular plasmid containing a pIM13 Gram-positive replicon, a *catP* selection marker, a *ColE1* gram negative replicon and a *lacZ* multiple cloning site containing a strong *fdx* promoter	(Heap et al. 2009)
pMTL85143 *nluc* 1	pMTL-85143 with Biomatik codon optimised Nanoluc synthetic gene (version 1) placed behind the *fdx* promoter by EcoRI and NheI cloning	This Study
pMTL85143 *nluc* 2	pMTL-85143 with manually codon optimised Nanoluc synthetic gene (version 2) placed behind the *fdx* promoter by EcoRI and NheI cloning	This Study
pMTL85143 *mCherry* 1	pMTL-85143 with Ransom et al (2015) codon optimised mCherry synthetic gene placed behind the *fdx* promoter by EcoRI and NheI cloning	This Study

pMTL85143 *mCherry* 2	pMTL-85143 with Biomatik codon optimised mCherry synthetic gene placed behind the *fdx* promoter by EcoRI and NheI cloning	This Study
pMTL-ME3	*C. acetobutylicum* in-frame deletion vector with a *catP* marker, pIM13 gram positive replicon, ColE1 gram negative replicon and a *C. sporogenes pyrE* counter selection marker	(Ehsaan et al. 2016)
pMTL-ME3 Δ*spo0A*	pMTL-ME3 with a *spo0A* in-frame deletion cassette	(Ehsaan et al. 2016)
pMTL-ME3 ΔCa_c0957	pMTL-ME3 with a Ca_c0957 in-frame deletion cassette cloned using NotI and BamHI restriction/	This Study
pMTL-ME3 ΔCa_c0958	pMTL-ME3 with a Ca_c0958 in-frame deletion cassette cloned using NotI and BamHI restriction	This Study
pMTL-ME3 $\Delta\sigma^K$	pMTL-ME3 with a *sigK* in-frame deletion cassette cloned using NotI and NheI restriction	This Study
pMTL-ME3 ΔCa_c0323	pMTL-ME3 with a *Ca_c0323* in-frame deletion cassette cloned using NotI and NheI restriction	This Study
pMTL-ME6	*C. acetobutylicum* ACE correction vector containing a *catP* marker, pIM13 gram positive replicon, ColE1 gram negative replicon and homology arms for the repair of the *pyrE* locus	(Ehsaan et al. 2016)

pMTL-ME6C	pMTL-ME6 with the addition of a *lacZ* multiple cloning site between the *pyrE* repairing homology arms to allow use as a complementation vector	(Ehsaan et al. 2016)
pMTL-ME6C Ca_c0957	pMTL-ME6C with Ca_c0957 and an upstream region cloned into the MCS by NotI and NheI restriction	This Study
pMTL-ME6C Ca_c0958	pMTL-ME6C with Ca_c0958 and an upstream region cloned into the MCS by NotI and NheI restriction	This Study
pMTL-ME6C Ca_c0957 His	pMTL-ME6C with Ca_c0957 histidine tagged fusion and an upstream region cloned into the MCS by NotI and NheI restriction	This Study
pMTL-ME6C Ca_c0958 His	pMTL-ME6C with Ca_c0958 histidine tagged fusion and an upstream region cloned into the MCS by NotI and NheI restriction	This Study
pMTL-ME6C Ca_c0957 Strep	pMTL-ME6C with Ca_c0957 streptavidin tagged fusion and an upstream region cloned into the MCS by NotI and NheI restriction	This Study
pMTL-ME6C Ca_c0958 Strep	pMTL-ME6C with Ca_c0958 streptavidin tagged fusion and an upstream region cloned into the MCS by NotI and NheI restriction	This Study
pMTL-ME6C σ^K	pMTL-ME6C with σ^K and an upstream region cloned into the MCS by NotI and NheI restriction	This Study
pMTL-ME6C *spo0A*	pMTL-ME6C *spo0A* complementation vector	(Ehsaan et al. 2016)

pMTL-ME6X	pMTl-ME6 with the addition of an *fdx* promoter *lacZ* multiple cloning site between the *pyrE* repairing homology arms to allow use as an overexpression vector	(Ehsaan et al. 2016)
pMTL-ME6X Ca_c0957	pMTL-ME6C with Ca_c0957 cloned into the MCS by NotI and NheI restriction behind the *fdx* promoter	This Study
pMTL-ME6X Ca_c0958	pMTL-ME6C with Ca_c0958 cloned into the MCS by NotI and NheI restriction behind the *fdx* promoter	This Study
pMTL-ME6X Ca_c0957 His	pMTL-ME6C with Ca_c0957 histidine tagged fusion cloned into the MCS by NotI and NheI restriction behind the *fdx* promoter	This Study
pMTL-ME6X Ca_c0958 His	pMTL-ME6C with Ca_c0958 histidine tagged fusion cloned into the MCS by NotI and NheI restriction behind the *fdx* promoter	This Study
pMTL-ME6X Ca_c0957 Strep	pMTL-ME6C with Ca_c0957 streptavidin tagged fusion cloned into the MCS by NotI and NheI restriction behind the *fdx* promoter	This Study
pMTL-ME6X Ca_c0958 Strep	pMTL-ME6C with Ca_c0958 streptavidin tagged fusion cloned into the MCS by NotI and NheI restriction behind the *fdx* promoter	This Study

Table 2.3 Oligonucleotides used in this study

Oligo nucleotide	Relevant use	Sequence
Generic primers		
M13-R	Modular plasmid forward sequencing primer	AGGAAACAGCTATGAC CAT
T7-F	pET21d forward sequencing primer	TAATACGACTCACTATA GGG
EBS2	ClosTron Probe Generation	GCAAGTTTCTAATTTCG GTTCTCATCCGATAGAG GAAAGTGTCT
Intron SalI-R1	ClosTron Probe Generation	TGTGACTGGTTTGCACC ACCCTCTTCG
Mutant confirmation and sequencing primers		
Cac0026-sF2	Confirmation of *pyrE* correction/integration at the *pyrE* locus	TAGCACAATTGTATTTG GACTTCTTTAAATAAAA ACATGG
Cac-*hydA*-sR2	Confirmation of *pyrE* correction/integration at the *pyrE* locus	TTGATGATGTTTGTCTT GATGACTCAACATGC
Cac-*spo0A*-sF2	Confirmation of Spo0A in-frame deletion	CTCACCCTTTCTTTCCA TCAC
Cac-*spo0A*-sR2	Confirmation of Spo0A in-frame deletion	CCTAAAAGTATGATGAT TAAAGTTACAGATGC
Ca_c0957 -sF1	Confirmation of Ca_c0957 in-frame deletion	GTAGAATGATAAGGGA TACTGCAG

Ca_c0957 -sR1	Confirmation of Ca_c0957 in-frame deletion	CTCAATCTTTTCCTTATAGTAATCAAATTG
Ca_c0958 -sF1	Confirmation of Ca_c0958 in-frame deletion	GAATAAGAACGAGTTAGGGGTAG
Ca_c0958 -sR1	Confirmation of Ca_c0958 in-frame deletion	GTTTTTTAACCCCTCTGGGAC
Cac-σ^K-sF1	Confirmation of σ^K in-frame deletion	CTTTGATTTAGAGTCG
Cac-σ^K -sR1	Confirmation of σ^K in-frame deletion	GGTTTTATCCGCACCAATAGT
Ca_c0323 -sF1	Confirmation of Ca_c0363 in-frame deletion	AAGATTGGTGCGCTTGTTGT
Ca_c0323 -sR1	Confirmation of Ca_c0363 in-frame deletion	AAAACATGCCTGACGATTAT
Ca_c0077 -sF1	Confirmation of Ca_c0077 ClosTron insertion	AACGCTTGAACTTTC
Ca_c0077 -sR1	Confirmation of Ca_c0077 ClosTron insertion	GGCATTGGACTTGAACTG
Ca_c0082 -sF1	Confirmation of Ca_c0082 ClosTron insertion	AAGTGGTAATAGTAGACG
Ca_c0082 -sR1	Confirmation of Ca_c0082 ClosTron insertion	CTTCTATCCCATCTAGCA

Generation of in-frame deletion cassette primes		
Ca_c0957 LHA -F	Generation of LHA for Ca_c0957 in-frame deletion vector. Used in SOE PCR to create in-frame deletion cassette	TATACGCGGCCGCTTTTTAATTGTACAG
Ca_c0957 LHA -R	Generation of LHA for Ca_c0957 in-frame deletion vector	TATTTCCATACGTATCTTCCTTTA
Ca_c0957 RHA -F	Generation of RHA for Ca_c0957 in-frame deletion vector	GTAAAGGAAGATACGTATGGAAATAGATTTGTAATCGGATAGGTGTG
Ca_c0957 RHA -R	Generation of RHA for Ca_c0957 in-frame deletion vector. Used in SOE PCR to create in-frame deletion cassette	AGAAGGATCCTTACACTATATATTTTTGC
Ca_c0958 LHA -F	Generation of LHA for Ca_c0958 in-frame deletion vector. Used in SOE PCR to create in-frame deletion cassette	ATTTAGCGGCCGCTAATTCATCTTGTAGTTAT
Ca_c0958 LHA -R	Generation of LHA for Ca_c0958 in-frame deletion vector	TATTTCCATCCTTCTCACACCTATCCGA
Ca_c0958 RHA -F	Generation of RHA for Ca_c0958 in-frame deletion vector	ATGGAAATAAAATGTAAATTTTATTGCTCATGTAA
Ca_c0958 RHA -R	Generation of RHA for Ca_c0958 in-frame deletion vector. Used in SOE PCR to create in-frame deletion cassette	ACAATTCTCTCTGCGGATCCCTCTACTAT

sigK LHA -F	Generation of LHA for σ^K gene in-frame deletion vector. Used in SOE PCR to create in-frame deletion cassette	GTTTCTTCGAATTCGCGGCCGCACTAGAGAGTCGGAGGGAAGACTTGA
sigK LHA -R	Generation of LHA for σ^K gene in-frame deletion vector	ACAATACATTACAAACACAGGTAATCCTCC
sigK RHA -F	Generation of RHA for σ^K gene in-frame deletion vector	GATTACCTGTGTTTGTAATGTATTGTTAAAATATAAAAACATGCTTTGTAAC
sigK RHA -R	Generation of RHA for σ^K gene in-frame deletion vector. Used in SOE PCR to create in-frame deletion cassette	GTTTCTTCCTGCAGCGGCCGCGCTAGCTTTCCGATTTCCCTTTGATTT
Ca_c0323 LHA -F	Generation of LHA for Ca_c0363 in-frame deletion vector. Used in SOE PCR to create in-frame deletion cassette	GTTTCTTCGAATTCGCGGCCGCACTAGAGTGCGCTTGTTGTCTGTGAG
Ca_c0323 LHA -R	Generation of LHA for Ca_c0363 in-frame deletion vector	TTACTCCAAATAATCCATTTTAATCGGTATATATTAGTAATTAA
Ca_c0323 RHA -F	Generation of RHA for Ca_c0363in-frame deletion vector	GAGAGGGATTAAAATGGATTATTTGGAGTAA
Ca_c0323 RHA -R	Generation of RHA for Ca_c0363 in-frame deletion vector. Used in SOE PCR to create in-frame deletion cassette	GTTTCTTCCTGCAGCGGCCGCGCTAGCGGATTGATGAAGCTGCTGAAA

Complementation and overexpression construct primers		
Ca_c0957-p-F	Cloning of Promoter region up stream of Ca_c0957	GTTTCTTCGAATTCGCGGCCGCTGATAAGGGATACTGCAGAAGA
Ca_c0957-p-R	Cloning of Promoter region up stream of Ca_c0957	ACGTATCTTCCTTTACTCGCTC
Ca_c0957 -F	Cloning of Ca_c0957	GTTTCTTCGAATTCGCGGCCGCACTAGATAAAGGAAGATACGTATGGAAATA
Ca_c0957 R	Cloning of Ca_c0957	GTTTCTTCCTGCAGCGGCCGCGCTAGCCACACCTATCCGATTACAAATC
Ca_c0957 His -R	Cloning of Ca_c0957, introducing a C-terminal histidine tag introducing an XhoI site in place of the original stop codon	GGGCCCGCTAGCTTAGTGGTGGTGGTGGTGGTGCTCGAGATCCTTCAATACACCTATCTTTTTG
Ca_c0957 Strep -R	Cloning of Ca_c0957, introducing a C-terminal streptavidin tag introducing an XhoI site in place of the original stop codon	GTTTCTTCCTGCAGCGGCCGCTTATTTTTCAAATTGAGGATGACTCCAGAGCTCATCCTTCAATACACCTATCTTTTTG
Ca_c0958-P-SOE	Cloning of Ca_c0958 with an overlap for SOE PCR to the promoter region of Ca_c0957	GAGCGAGTAAAGGAAGATACGTTCGGATAGGTGTGAGAAGGATGGAAATA

Ca_c0958-F	Cloning of Ca_c0958	GTTTCTTCGAATTCGCGGCCGCACTAGATCGGATAGGTGTGAGAAGGATGGAAATA
Ca_c0958-R	Cloning of Ca_c0958	GTTTCTTCCTGCAGCGGCCGCGCTAGCTTACATGAGCAATAAAATTTACATT
Ca_c0958 His -R	Cloning of Ca_c0958, introducing a C-terminal histidine tag introducing an XhoI site in place of the original stop codon	GGGCCCGCTAGCTTAGTGGTGGTGGTGGTGGTGCTCGAGTTTATTGCTCATGTTAATTGCTAAATTAAAATATC
Ca_c0958 Strep -R	Cloning of Ca_c0958, introducing a C-terminal streptavidin tag introducing an XhoI site in place of the original stop codon	GTTTCTTCCTGCAGCGGCCGCTTATTTTTCAAATTGAGGATGACTCCAGAGCTCTTTATTGCTCATGTTAATTGCTAAATTAAAATATC
Cac-*sigK*-p-F	Cloning of σ^K gene with its promoter region	GTTTCTTCGAATTCGCGGCCGCACTAGAGAGTCGGAGGGAAGACTTGA
Cac-*sigK*-R	Cloning of σ^K gene	GTTTCTTCCTGCAGCGGCCGCGCTAGCCCGGCTGAACATTAACAGAAATATG

Ca_c0082 -F	Cloning of Ca_c0082 with its promoter region	GTTTCTTCGAATTCGCGGCCGCACTAGACATTCTGTAAATAGGCAAATAGAA
Ca_c0082 -R	Cloning of Ca_c0082 with its promoter region	GTTTCTTCCTGCAGCGGCCGCGCTAGCTTATTCCAACCATAAAATTTTCGGAAC

Table 2.4 Synthetic peptides used in this study

Peptide	Relevant properties	Sequence
R6T0-Ca	*C. acetobutylicum* AgrD derived signal peptide described by Steiner et al. (2012)	CVLVTL*
Amid-R6T0-Ca	*C. acetobutylicum* AgrD derived signal peptide described by Steiner et al. (2012) having undergone a proposed molecular rearrangement to form an amide bond	CVLVTL**
Ac-R6T0-Ca	*C. acetobutylicum* AgrD derived signal peptide described by Steiner et al. (2012) with the addition of an acetyl group to prevent a molecular rearrangement	CVLVTL*
Ac-R6T0-Cr	A signal molecule derived from the putative AgrD operon of *C. roseum*, acetylated as the Ac-R6T0-Ca	CVLTTL*
Ac-R6T1-Cr	A signal molecule derived from the putative AgrD operon of *C. roseum*, with a 1-amino acid length tail	ICVLTTL*

Ac-R6T4-Cr	A signal molecule derived from the putative AgrD operon of *C. roseum*, with a 1-amino acid length tail	ATDICVLTTL*
R6T0-Ca-L	Linearised R6T0-Ca peptide	CVLVTL
R6T1-Ca-L	Linearised R6T0-Ca peptide with a 1-amino acid tail	KCVLVTL
R6T4-Ca-L	Linearised R6T0-Ca peptide with a 4-amino acid tail	ANGKCVLVTL

***note this is a cyclic structure formed by a thiolactone bond between the SH group of the cysteine and the C-terminal carboxy group**
****note this is a cyclic structure formed by an amid bond between the N-terminal and C-terminal carboxy groups**

2.3. Microbiological Methods

2.3.1. Media

Escherichia coli was grown in LB (lysogeny broth = 10 g/l tryptone, 5 g/l yeast extract, 5 g/l sodium chloride, pH=7.2), with the addition of 15 g/l agar for solid agar medium. The medium was supplemented with 25 µg/ml chloramphenicol, 10 µg/ml tetracycline, 500 µg/ml erythromycin and 50 µg/ml ampicillin when appropriate for antibiotic selection.

C. acetobutylicum was grown in CBM-S broth for phenotypic analysis and on CBM-S agar for the selection of single crossover mutants and granulose staining (O'Brien & Morris 1971). CBM-S broth contained:– 10 ml of 20 g/l $MgSO_4$ x 7 H_2O, 4g casamino acids, 379 µl of 20 g/l $MnSO_4$ x H_2O, 500 µl of 20 g/l $FeSO_4$ x $7H_2O$, 1ml of 1 g/l p-amino-benzoic acid , 1ml of 1 g/l thiamine-HCl and 20 µl of 0.1 g/l biotin . The resulting solution was made up to a volume of 750ml using reverse osmosis water and following autoclaving the following was added: 20ml of 5 g/l $CaCO_3$ (autoclaved), 10 ml of sterile filtered 50 g/l K_2HPO_4 anhydrous and 10 ml 50 g/l KH_2PO_4 anhydrous, a variable amount of a 500 g/l glucose to create the desired glucose

concentration and distilled water used to create a final volume of 1 litre. For solid agar $CaCO_3$ was replaced by distilled water and prior to autoclaving 1.5% (w/v) technical agar 1 was added.

For transformation and genome preparation *C. acetobutylicum* was grown in liquid 2xYTG (16 g/l tryptone, 10 g/l yeast extract, 5 g/l sodium chloride, pH=5.2) made up to 950 ml before autoclaving, after which 50 ml of a 500 g/l glucose solution was added. Enumeration plating for heat resistant cfu's and non-uracil dependent selection was performed with CGM agar (50 g glucose, 2 g tryptone, 2g $(NH_4)_2SO_4$, 1 g yeast extract, 1 g K_2HPO_4, 0.5 g KH_2PO_4, 0.1 g $MgSO_4$ x 7 H_2O, 750 µl 20 g/l $FeSO_4$ x 7 H_2O, 500 µl $CaCl_2$ 20 g/l, 500 µl 20 g/l $MgSO_4$ x H_2O, 100 µl 20 g/l $CoCl_2$, 100 µl 20g/l $ZnSO_4$, 10.2 g agar and reverse osmosis water to make a total of a litre)

Before the addition of post autoclave components media were stored at room temperature for no more than 1 month. Agar plates were reduced for a minimum of 4 hours and liquid cultures for a minimum of 16 hours inside the anaerobic work station (MG1000 anaerobic work station, Don Whitley Scientific). Media for *C. acetobutylicum* were supplemented with the following as appropriate no earlier than 12 hours before commencing an experiment: thiamphenicol 15 µg/ml, erythromycin 50 µg/ml, uracil 20 µg/ml or 2 µg/ml in the presence of 5-FOA and 5-FOA 200 µg/m

2.3.2. Growth

E. coli was grown in LB medium at 37 °C with liquid cultures having shaking agitation of 200 rpm. *C. acetobutylicum* was grown at 37 °C in an anaerobic workstation (Don Whitely) in an atmosphere of 80% nitrogen, 10% hydrogen and 10% carbon dioxide on CBM-S, 2xYTG and CGM media.

E. coli from bead stocks was used to directly inoculate LB broth overnights for pre-cultures or plasmid preparations. *C. acetobutylicum* was first grown on plates either by streaking from a glycerol stock or from a volume of 50-100 µl of a spore stock that was heat treated at 80 °C. Inoculation of *C. acetobutylicum* liquid overnight cultures involved resuspending several colonies scraped from an agar plate in 10 ml of an appropriate liquid medium. When a growing culture was required this neat overnight culture would be serially diluted 1:10 in fresh medium to create a dilution series down to 10^{-3}.

To inoculate large cultures overnight cultures showing an OD_{600} of between 0.4 and 0.8 were selected to use cells still undergoing exponential growth to inoculate larger volumes to the required target OD_{600}.

2.3.3. Storage

E. coli strains and any *C. acetobutylicum* strain incapable for forming spores were stored at -80 °C as using beads or glycerol stocks respectively. *E. coli* stocks were formed by selecting single colonies from a plate using an inoculation loop and resuspending it with in a glycerol bead stock vial, after which the vial was inverted approximately 10 times to allow cell adhesion to the beads and the supernatant removed prior to storage. *C. acetobutylicum* was stored as a glycerol stock at -80 °C through the addition of 200 µl 50% (v/v) glycerol to 1 ml of an exponential phase culture before storing.

All *C. acetobutylicum* strains capable of doing so were stored a spore stocks to eliminate the risk of reviving degenerated strains. Spore stocks were produced by inoculating 10 ml of CMB-S broth 6% (w/v) glucose with the desired strain and allowing the culture to grow for 5 days. After 5 days the suspension was poured off to remove any remaining $CaCO_3$ and then centrifuged at 10,000g to create a pellet. The pellet was subsequently resuspended in phosphate

buffered saline (PBS = 0.001 M phosphate buffer, 0.0027 M KCl, 0.137 M NaCl, pH 7.4) with 10% (v/v) glycerol and stored at - 20°C.

2.4. Molecular biological methods

2.4.1. Isolation of DNA

2.4.1.1. Isolation of plasmid DNA from *E. coli*

Plasmid DNA was isolated from a 5 ml overnight culture of *E. coli* using the Sigma-Aldrich GenElute Plasmid Miniprep kit, with the only alteration being the reduction of the elution volume to 50 µl. Resultant plasmid preparation were stored at -20 °C.

2.4.1.2. Isolation of genomic DNA from *C. acetobutylicum* by column purification

gDNA was isolated from 5 ml of *C. acetobutylicum* 2xYTG overnight culture using the Sigm-Aldrich GenElute Bacterial Genomic DNA kit, with the alteration being the pre-treatment of the cells with 100 µg/ml of lysozyme and the reduction of the final elution volume to 50 µl. Resultant gDNA preparation were stored at -20 °C.

2.4.1.3. Isolation of genomic DNA from *C. acetobutylicum* by phenol chloroform extraction

In order to increase the gDNA yields for effective southern blotting, phenol-chloroform extraction was used to overcome the binding capacity limitation of column preparation when isolating *C. acetobutylicum* gDNA. For this process 10 ml of an overnight culture in exponential phase growth was centrifuged at 10,000g at room temperature for 5 minutes and the supernatant decanted. Following this the pellet was resuspended in 180 µl of a lysis buffer (PBS with 10 mg/ml lysozyme) and incubated at 37 °C for 30 minutes. Following this 20 µl of

a sigma RNAase solution was added and the culture incubated at 37 °C for an additional 10 minutes. The sample was then treated with 25 µl of proteinase K solution (Qiagen) and 110 µl 10% (w/v) SDS solution and incubated at 65 °C for 30 minutes.

DNA samples were supplemented with an equal volume of phenol-chloroform-isoamyl alcohol pH 8 and inverted approximately 10 times before being transferred to a phase lock tube. Following 3 minutes of centrifugation at 16,000g at room temperature, the aqueous top layer was removed to a new phase lock tube at which point the supplementation and centrifugation was repeated. Following yet another aqueous phase removal, supplementation with phenol-chloroform-isoamyl alcohol pH 8 and centrifugation the final aqueous phase was transferred to a 1.5 ml reaction tube and mixed with 40 µl of 3 M sodium acetate and 800 µl of ice cold 100% ethanol and the DNA precipitated at -80 °C for 30 minutes. The precipitated DNA was pelleted by centrifugation at 16,000g at room temperature and washed with 1 ml of 70% ethanol. The sample was again centrifuged at the supernatant removed by pipetting, with the remainder of the supernatant left to air dry for at least 30 minutes. After this the DNA pellet was gently resuspended in 50 µl of distilled water and stored at -20 °C.

2.4.2. Methods of DNA Transfer

2.4.2.1. Preparation and transformation of electrocompetent *E. coli*

E. coli strains for the generation of electrocompetent cells were grown overnight in 10ml LB liquid medium, which was used to inoculate a 200 ml of LB in a 500 ml conical flask. The culture was subsequently grown at 37 °C shaking at 200 rotations per minute (RPM). The OD_{600} of the initial culture and subsequent hourly samples were measured until the cells achieved an OD_{600} of 0.6-0.8.

Cells were harvested and spun down at 4,000 g at 4 °C for 10 minutes and the supernatant removed. Cell pellets were suspended in 80 ml of ice cold distilled water and centrifuged again. The resulting cell pellet was resuspended one further time in 20 ml of ice cold distilled water and centrifuged again before the final pellet was resuspended in ice old 1 ml 10% (v/v) glycerol (dissolved in distilled water). The electrocompetent cell suspension was stored in 50 µl aliquots at -80 °C until they were needed.

When needed electrocompetent cell aliquots were thawed on ice and mixed with 100-500 ng of a plasmid preparation. The aliquots were subsequently transferred to chilled 2 mm electroporation cuvettes and electroporation was carried out at 200 Ω, 2.5 kV and 25 µF using a Bio-Rad Micro Pulse. Immediately following electroporation, 1 ml of LB broth was added to the aliquot and was gently mixed by pipetting. The recovering cells were transferred to 1.5 ml reaction tubes and incubated for 1 hour at 37 °C shaking at 200 RPM, before being plated on to LB agar containing the appropriate antibiotic. After overnight incubation colonies were selected to confirm successful plasmid uptake.

2.4.2.2. Preparation and transformation of *C. acetobutylicum* competent cells

Following an overnight 1:10 dilution series (down to 10^{-3}) in 2xYTG, an overnight culture showing an OD_{600} of between 0.4-0.8 was used to inoculate 75ml of 2xYTG. Upon this culture reaching an OD_{600} of 0.6-0.8 the cells were harvested and centrifuged at 4,000 g at °C for 10 minutes before returning to the anaerobic work station. The supernatant was then removed and the cells resuspended in 10 ml of a sterile, ice cold, reduced electroporation buffer (EPB), consisting of 19 parts 284 mM sucrose solution and 1 part 100 mM sodium phosphate buffer (made up from 9.5 mM monobasic sodium phosphate and 40.5 mM dibasic sodium phosphate).

Resuspended cells were again centrifuged at 4,000g at °C for 10 minutes before returning to the anaerobic workstation and being resuspended in a final volume of 2 ml of EPB.

Aliquots of 540 µl were either mixed with 60 µl of DMSO and stored at -80 °C until they were needed, or aliquots of 570 µl were mixed with 10-30 µl (containing 5-10 µg) of prepared plasmid the methylation strain pAN2 (Heap et al. 2007) for direct transformation. Aliquots were transferred to 4 mm electroporation cuvettes and pulsed at 1 Ω, 2. kV and 25 µF, after which the aliquot was inoculated into 10 ml of 2xYTG media and left for 3-4 hours for recovery. After recovery the cells were centrifuged at 4000g at room temperature for 10 minutes before being transferred back to the anaerobic work station. After decanting the supernatant, cell pellets were resuspended in 1 ml of 2xYTG and subsequently plated on CGM agar containing the appropriate supplements.

2.4.3. DNA manipulation and analysis techniques

2.4.3.1. Quantification and purity determination of nucleic acids

Nucleic acids preparation were measured using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific). Samples were quantified at 260 nm and extinction values were quantified using the known value of 50 ng/µl of double stranded DNA having an extinction coefficient of 1. DNA purity was tested through measuring the ratios of 260 nm/230nm and 260 nm/280 nm. A ratio of below 1.8 implied contaminations with protein/phenol and carbohydrates for the respective ratio of the measurements.

2.4.3.2. Polymerase chain reaction for the amplification of DNA

PCR's for the initial screening of double cross over events or ClosTron insertions was carried out in 20 µl reaction volumes using: 100-200 ng of gDNA, 10 µl of DreamTaq master mix (Thermo Scientific), 1 µl 0.1 mM forward and reverse primer and up to 10 µl distilled water.

The program structure used was 95 °C for 2 minutes followed by 30 cycles of 95 °C for 30 seconds, annealing at 5 °C below the melting temperature or at 50 °C for 30 seconds and elongation at 72 °C for 1 minute per 1 kb (dropping this temperature to 68°C if extension lasted longer than 3 minutes). DNA for sequencing and cloning application was amplified using proof-reading Phusion polymerase (New England Biolabs) using 10 µl Phusion 2x master mix, 100-200 ng of template DNA, µl 0.1 mM forward and reverse primer and up to 10 µl distilled water. The following program structure was used 95 °C for 2 minutes followed by 30 cycles of 95 °C for 30 seconds, annealing at 5 °C below the melting temperature or at 50 °C for 10 seconds and elongation at 70 °C for 30 seconds per 1 kb. All PCR reactions were performed in a T3000 Thermocycler (Biometra) or a Sensoquest labcyler (Geneflow).

Splicing by overhang extension (SOE) PCR was performed in 20 µl volumes using the following: 10 µl Phusion master mix, 10-20 ng of each fragment of template DNA, 0.5 µl of 0.1 mM forward and reverse primer and up to 10 µl distilled water. The following program structure was used: 95 °C for 2 minutes followed by 30 cycles of 95 °C for 30 seconds, annealing at 50 °C for 10 seconds and elongation at 70 °C for 1.5 minutes.

2.4.3.3. Restriction Digests of DNA

DNA was specifically cleaved using Fastdigest restriction enzymes (Thermo Scientific) according to the instruction of the manufacturer. Reactions were generally performed in 20 µl volumes using up to 1 µg of DNA, 10 µl of 2x reaction buffer (green), 1 µl of restriction enzyme and distilled water up to 10 µl. Reactions were incubated for 15-30 minutes at 37 °C.

2.4.3.4. Agarose Gel Electrophoresis

DNA fragments were separated by size through the use of gel electrophoresis at 90-110 V in 1x TAE running buffer (40 mM Tris, 1 mM EDTA and 0.1% (v/v) glacial acetic acid) for 50-

70 minutes. Electrophoresis was carried out using a Compact gel tank (Biometra) and a PowerPac 300 (Bio-Rad). Gels were made of 1x TAE buffer containing 1% (w/v) agarose and 0.01% (v/v) SYBR SAFE DNA gel stain. DNA was visualised with the use of a blue light transilluminator and band size was determine through the use of a Generuler 1kb ladder (Thermo Scientific).

2.4.3.5. DNA purification

DNA fragments were purified and/or concentrated from PCR reaction with the use of a Zymo DNA Clean & Concentrator kit (Zymo Research) and was performed as per the manufacturer's instructions. DNA was purified from agarose gel fragments through the use of Zymoclean Gel DNA Recovery kit (Zymo Research) as per the manufacturer's instructions.

2.4.3.6. DNA ligation

DNA fragments were ligated using 1 µl T4 ligase (Promega), 2 µl of 10x ligation buffer, a general ratio of 3:1 of insert to vector with 100-200 ng of vector and distilled water for a final volume of 20 µl. Ligations were incubated at room temperature for 2 hours or at 4 °C for 14-16 hours.

2.4.3.7. Dialysis of DNA

Excess salts were removed from ligation reactions prior to electroporation through the use of a 0.0025 µm nitrocellulose membrane (Millipore). The sample was dialysed against 15-20 ml of distilled water for 15 minutes inside a Petri dish.

2.4.3.8. DNA sequencing

Sanger sequencing to confirm DNA samples was performed by Source BioScience Gene Service (Nottingham). Samples were prepared as per the services instructions and collected for delivery in 2 ml Cryotubes.

2.4.3.9. Methylation of Plasmid DNA

Plasmid DNA for transformation of *C. acetobutylicum* was methylated in vivo using *E. coli* TOP10 containing the pAN2 methylation plasmid (Heap et al. 2007). Plasmids were transformed into the methylation strain, *E. coli* TOP10 pAN2, as previously described. Methylated plasmid was harvested through the scraping of a confluent agar plate following overnight incubation and the resuspension of the cells in 400 µl of resuspension buffer (Sigma). Plasmid preparation was performed using a Sigma Genelute miniprep kit, however due to high cell densities volumes of resuspension, lysis and neutralisation buffer were doubled. Following binding to the column plasmid preparation was carried out as per the manufacturer's instructions.

2.4.3.10. Southern blotting and probe hybridisation

2-5 µg of *C. acetobutylicum* gDNA isolated by phenol chloroform extraction as previously described was digested overnight using HindIII restriction enzyme at 37 °C. A wild type negative control and a ClosTron plasmid positive control were digested in the same manner. Following restriction the samples were run on a 1% (w/v) agarose gel as previous described before the DNA was transferred onto a positively charged Hybond nylon membrane (Roche) by alkaline transfer using 0.4 M sodium hydroxide. For this process approximately 20-30 mm of a container was filled with 0.4 M sodium hydroxide and an inverted electrophoresis tray placed inside the container. Covering the tray was a wick already soaked with 0.4 M sodium

hydroxide made of chromatography paper and contacting the sodium hydroxide pool on both sides of the container. Any air bubbles were removed from the wick by rolling with a 10 ml glass pipette before the agarose gel was place on the wick. The nylon membrane, cut down to match the size of the gel and separately pre-soaked in 0.4 M sodium hydroxide was placed over the gel and covered with a pre-wetting blotting pad (Sigma). Dry blotting pads were layered above this and a weight place on top of the assembly. The apparatus was left for 3 hours in order to allow DNA transfer to occur.

Following DNA transfer the membrane was washed twice in 2 x saline-sodium citrate solution (SSC, 3 M sodium chloride 300 mM tri-sodium-citrate) inside a sealed plastic container for 2 minutes (per side of the membrane) and placed under a UV transilluminator for two minutes on either side to UV-fix the DNA to the membrane. The membrane was place in a hybridisation tube and was incubated for an hour with 10 ml DIG-Easy Hyb (Roche) at 42 °C for pre-hybridisation treatment. The pre-hybridisation solution was replaced with a second DIG-Easy Hyb solution containing approximately 0.1 µg/ml of a ClosTron specific DIG-labelled probe. This probe was generated by PCR using the EBS2 and Intron SalI-R1 primers from a ClosTron plasmid template and was labelled by denaturing the PCR product at 100 °C and incubating it overnight following the addition of 4 µl of DIG-High Prime. This reaction was stopped by heat inactivation at 65 °C for 10 minutes. Probes were stored at -20 °C and denatured at 100 °C for 5 minutes prior to immediate use.

Hybridisation at 42 °C occurred overnight, after which the membrane was washed twice with 100 ml of 2x washing solution (2x SSC and 0.1% (w/v) SDS) for 5 minutes at room temperature and then twice with 0.5x washing solution (0.5x SSC and 0.1% (w/v) SDS) for 15 minutes at 68 °C. A 5x maleic acid buffer (0.5 M maleic acid and 0. 75 M sodium chloride, pH 7.5) was used to make 50ml 1x maleic acid buffer with 0.3% (v/v) Tween 20 as an equilibration buffer,

used to equilibrate the membrane at room temperature for 1 minute. Membrane were then washed in 30 ml of blocking solution consisting of 1x maleic acid buffer and 1x blocking buffer (Roche) for 30 minutes at room temperature. The membrane was exposed to anti-DIG antibodies buffer (Roche) at a 1:10000 dilution in blocking buffer for 30 minutes at room temperature then subsequently washed twice in equilibration buffer and incubated with detection buffer (100 mM Tris-HCl and 100 mM sodium chloride, pH 9.5) for 2 minutes The membrane was removed from the hybridisation tube and placed on an acetate sheet (Q-connect) with the DNA side facing upwards. 1.5 ml CSDP-ready to use (Roche) was directly pipetted onto the membrane and an acetate sheet was placed above the membrane. Air bubble were removed gently with the aid of a paper towel before the acetate sheets were sealed with tape and the membrane was incubated at 37 °C for 10 minutes. Following this samples were ready for exposure to photographic film.

2.4.3.11. Photographic film development

Working in a dark room (using red light) chemiluminescence film was placed on top of the membrane inside the acetate sheet envelope. After re-sealing, the film was exposed for a variable time dependent on how intense previous exposures had been shown. After exposure the film was developed using developing solution (KODAK) for 30 seconds before washing in tap water for 30 seconds and fixing solution (KODAK) for 1 minute. After this the film was washed a final time in tap water and left to air dry.

2.4.4. Expression and analysis of proteins

2.4.4.1. Expression of tagged proteins from *E.coli*

BL21 Star (DE3) *E.coli* containing pET-21d expression vectors were grown over night in 5 ml of LB with ampicillin, to be used to inoculate 100 ml of LB to an OD_{600} of 0.05-0.1. OD_{600} was

measured and upon reaching 0.4-0.5 expression of the tagged protein was induced through the addition of 500 µl of a 100 mM IPTG stock solution. OD_{600} was continued to be monitored until the cells were harvested at approximately OD_{600} 1.0. Cell pellets were harvested following centrifugation at 10,000 g and then stored frozen at -20 °C.

2.4.4.2. Expression of tagged and reporter proteins from *C. acetobutylicum*

C. acetobutylicum reporter strains with genes encoding tagged proteins under the control of the *fdx* promoter were grown overnight in CBM-S 6% glucose and when at exponential phase used to inoculate a 200 ml CBM-S 6% glucose main culture. OD_{600} values were measured hourly and cells harvested between values of 0.8-1.0 by centrifugation at 10,000 g. Tagged protein cell pellets were stored frozen at -20 °C.

2.4.4.3. Lysis of cell pellets for protein extraction

Following freeze thaw lysis of cell pellets, *E. coli* cell pellets were resuspended in 500 µl of BugBuster Protein Extraction Reagent (Novagen) and incubated at 37 °C for 30 minutes. Cell debris were separated by centrifugation at 10,000g and supernatants removed for subsequent protein work. *C. acetobutylicum* cell pellets following freeze thawing were subject to resuspension in 500 µl of 10 mg/ml lysozyme solution for 1 hour at 37 °C, following centrifugation at 10,000 g to remove debris and the removal of supernatant for subsequent work.

2.4.4.4. Column purification of His- and Strep-tagged fusion proteins

Following cell lysis His-tagged fusion constructs were pulled down by nickel affinity column purification using a HisPur Ni-NTA spin purification kit (Thermo) and following the protocol specified within. Strep-tagged fusion purification was performed using a Strep-tag Starter 3C kit (IBA-lifesciences) and following the protocol specified within.

2.4.4.5. SDS-PAGE

10 µl of the prepared protein samples were mixed with 5 µl of 4x NuPAGE LDS Sample Buffer (Thermo Scientific), 2 µl of 10x NuPAGE Reducing Agent (Thermo scientific) and 3 µl of deionised water to create a final sample size of 20 µl in a 1.5 ml reaction tube. These samples were heated to 100 °C for 5 minutes to ensure protein denaturing before loading onto a NuPAGE 4-12% Bis-Tris mini gel (Thermo Scientific).

50 ml 20x NuPAGE MES Running Buffer (Thermo Scientific) was added to 950 ml of deionised water to create a 1x Running Buffer. Prior to loading Bis-Tris mini gels were washed in 1x Running Buffer and loaded into an XCell SureLock Mini-Cell gel running tank (Thermo Scientific) as per the manufacturer's instructions. The upper chamber of the tank was filled with approximately 200 ml of 1x Running Buffer while the lower chamber was filled with approximately 600 ml of 1x Running Buffer. 10 µl of samples were loaded onto the gels with blank wells being filled with a blank samples prepared as previously described but with the substitution of the protein sample with deionised water. One lane was loaded with a Colour Prestained Protein Standard, Broad Range (New England Biosciences) in order to allow size determination. Gels were run at 150 V for 1-1.5 hours dependent on clear separation of the protein ladder bands. Mini gels were removed from their plastic casing by separation of the plastic at the corners and subsequent transfer of the gels using forceps.

2.4.4.6. Coomassie Blue Staining

Following SDS-PAGE mini gels were transferred into a square petri-dish and incubate at room temperature in 30-50 ml of Coomassie Blue Staining solution (0.25% (w/w) coomassie blue R-250, 50% (v/v) deionised water, 40% (v/v) methanol and 10% (v/v) acetic acid) for 1 hour. Following this the gel was washed with and subsequently left to incubate at room temperature

overnight in a de=staining solution (50% (v/v) deionised water, 40% (v/v) methanol and 10% (v/v) acetic acid) or until bands were visible in the absence of background coomassie staining.

2.4.4.7. Western Blotting

Following SDS-PAGE the mini-gel was placed upon pre-soaked anode stack and nitrocellulose membrane of a Trans-Blot Turbo Transfer System Pack (BIO-RAD) using forceps and air bubbles were removed by the use of a roller. Following this the cathode stack was placed above the gel and air bubbles again removed using a roller. The cassette was transferred to a Trans-Blot Turbo Start System (BIO-RAD) and run as per the manufacturers mixed molecular weight protocol (5-150 kD, 7 minutes, 1.3-2.5A constant dependent on the use of 1-2 gels and a maximum of 25 V.

After transfer the cassette was dissembled, the nitrocellulose membrane transferred into a square petri dish and incubated at room temperature for 1 hour in a blocking solution (PBS with 5% (w/v) milk powder (Marvel)). Following a rapid wash with PBS the membranes were left overnight at 4 °C to incubate in blocking solution containing a 1:10000 dilution of the appropriate antibody-HRP conjugate (either anti-histidine or anti-streptavidin). Following a rapid wash in PBS + 0.05% tween 20, the membrane was inoculated with approximately 5ml of TMB solution (Thermo Scientific) and membranes were observed for the formation of blue pigments.

2.5. Use of synthetic oligonucleotide and peptide constructs

2.5.1. Codon optimisation of synthetic Nluc and mCherry gene constructs

For the Nanoluc synthetic genes codon optimisation of the first construct was performed by the provider, Biomatik, for use in *C. acetobuytlicum* while the second construct was determined using simply the codons abundantly used in *C. acetobutylicum* highly expressed genes (Musto

et al. 2003). mCherry constructs were a copy of the codon optimised *C. difficile* construct previously tested (Ransom et al. 2015) and a *C. acetobutylicum* codon optimised structure devised by Biomatik. All codon optimised constructs were purchased and synthesised from Biomatik.

2.5.2. Design and synthesis of re-targeted pMTL007C-E2 plasmids

Design of re-targeted pMTL007C-E2 constructs was performed using an intron targeting tool described by Heap et al. (2010). Target sites were chosen to include both: the highest scores for either a sense or antisense insertion site (whichever of the two was greatest) and direct disruption of a conserved domain. Re-targeted pMTL007C-E2 constructs were purchased and synthesised by DNA 2.0.

2.5.3. Preparation of synthetic peptides

Synthetic peptides were manufactured by Peptide Synthetics (Cambridge) with a minimum purity of ≥95%. The peptides were only resuspended prior to their first use to maintain stability and were stored at -20 °C. Peptides were dissolved in DMSO to a stock concentration of 20 mM and were subsequently again stored at -20 °C. Due to potential instability peptide stocks older than 6 month were subsequently reordered before any additional experiments were performed.

2.6. *C. acetobutylicum* genetic manipulation

2.6.1. Generation, isolation and confirmation of *C. acetobutylicum* ClosTron mutants

C. acetobutylicum electrocompetent cells were transformed with 5-10 µg of methylated pMTL007C-E2 and allowed to recover as previously described. Transformants were selectively plated on CGM agar containing thiamphenicol to screen for successful uptake of

the pMTL007C-E2 plasmid. Following overnight growth single colonies were re-streaked onto CGM agar containing erythromycin and again allowed overnight growth. These colonies were subsequently cultured overnight in 2xYTG liquid medium (containing erythromycin) for the preparation of gDNA using column preparation.

Mutant gDNA was screened by PCR using primers flanking the anticipated target site and a primer across the intron junction and the results were analysed by gel electrophoresis in order to confirm insertion into the correct target site. Following this single insertion was confirmed through the use of Southern hybridisation of a phenol-chloroform purified gDNA sample using an *ermB* based probe (Heap et al. 2007).

2.6.2. Generation, isolation and confirmation of *C. acetobutylicum* in-frame deletion mutants and chromosomal insertion mutants

C. acetobutylicum Δ*pyrE* was transformed with 5-10 µg of the appropriate pMTL-ME3 in-frame deletion vector and left to recover as previously described (2.4.2.2). Following this transformants were plated onto CGM agar containing thiamphenicol and allowed overnight growth. Larger colonies, more likely to contain successful single crossover events were subsequently re-streaked onto CBM agar and allowed overnight growth. Suspected single cross over mutants were re-streaked on to CGM agar containing 5-FOA and uracil and left to grow overnight. Following this suspected double crossover mutants were selected for growth in 2xYTG (supplemented with uracil) for gDNA preparation. The colonies were also re-streaked onto CGM agar containing thiamphenicol to confirm the loss of the pMTL-ME3 plasmid

Following gDNA extraction through column preparation, the desired target region was screened by PCR and gel electrophoresis using flanking primers and internal primers for the deleted gene region. Subsequent mutants showing successful deletion and no inclusion of the internal region were confirmed through Sanger sequencing.

2.7. Phenotypic analysis of *C. acetobutylicum*

2.7.1. Preparation of *C. acetobutylicum* cultures for phenotypic assays

C. acetobutylicum colonies were picked and resuspended in 10 ml CBM-S broth 6% glucose before subsequently being diluted to provide overnight pre-cultures (as previously described). Pre cultures were measured to determine which had OD_{600} values of between 0.4-1.0 to be used for the inoculation of main cultures to a starting OD_{600} of between 0.01 and 0.02 for growth curve analysis and 0.05 and 0.1 for other purposes using the same medium. Culture volumes were either 30 ml CBM-S broth 6% (w/v) glucose within a 50 ml reaction flask or 75 ml CBM-S broth 6% glucose (in a 200 ml conical flask). At the point of inoculation samples were taken for initial solvent and heat resistant cfu/ml values. All strains were grown in technical and biological (mutants generated from independent transformations) triplicates.

2.7.1.1. Preparation of *C. acetobutylicum* cultures for assays under altered media compositions

Due to large number of replicates and/or the limited volume of supplement available, both the synthetic peptide and glucose concentration assays were performed in smaller volume cultures for their respective phenotypic characterisation to be able to maintain biological and technical triplicates.

For the assays regarding of the effect of synthetic peptides 10 ml CBM-S broth 6% (w/v) glucose serial dilutions over nights were prepared down to 10^{-3} for overnight growth. Selecting an overnight culture showing an OD_{600} of 0.4-0.8 a 250 ml culture of CBM-S 6% (w/v) glucose broth was inoculated to an OD_{600} between 0.05 and 0.1. Upon reaching OD_{600} 0.6 the cultures were distributed into 10 ml or 500 µl volumes (in 15 ml and 2 ml reaction tubes respectively) and inoculated with a volume of 20 mM peptide DMSO solution to create the desired

concentration and a volume of DMSO to ensure equal final concentrations in each culture. Samples were taken at 0 h, 48 h and 120 h time points for heat resistant cfu/ml determination. All assays were performed alongside a DMSO control.

Assays designed to determine the effect of glucose on mutant phenotypes were set up similarly, however CBM-S for the overnight pre-cultures and subsequent 250 ml experimental culture contained only 0.5% (w/v) glucose. Upon the 250 ml culture reaching an OD_{600} of 0.6 the culture was again divided however into volumes of 7.5 ml each (in 15 ml reaction tubes). These cultures were subsequently made up to a total volume of 10 ml through the addition of a volume of 500 g/l glucose solution to create the final desired glucose concentration and an additional amount distilled water to achieve the desired volume. These cultures were also sampled at 0 h, 48 h and 120 h time points.

2.7.2. OD measurements

Cell growth was monitored through measurement of the OD_{600} of a culture at 600 nm, using a 1 cm gap cuvette and a BioMate 3 spectrophotometer using a sterile liquid medium sample as a blank. Samples with an OD_{600} in excess of 0.8 were diluted 1:10 with the respective sterile liquid medium. Before samples were taken, cultures were mixed through vortexing and pipetting for 3 seconds. Cultures containing $CaCO_3$ were left for approximately 1 minute to allow the buffering agent to visibly settle out before taking OD_{600} samples.

2.7.3. Microscopy

To visualise cells and phase bright endospores, *C. acetobutylicum* was grown in liquid culture and a 10 µl sample places on a microscope carrier covered with a light cover slip. The cells were analysed using an inverted light microscope (Nikon Eclipse TE200) at 40 fold magnification.

2.7.4. Granulose staining and cross-streaking of *C. acetobutylicum*

Cultures for granulose staining were serially diluted in liquid medium for overnight growth and cultures showing an OD of between 0.4-1.0 were selected for use in cross streaking. Strains were vortexed and using a 10 µl sterile inoculation loop a streak was made back a forth to well distribute the droplet of culture on a CBM agar plate containing 5-6% (w/v) glucose, this process was then repeated 3-4 times to ensure high cell numbers in the streak. Cross-streaking to observe interactions were done so that streaks of desired strains ended approximately 5 mm from one another, allowing the effective diffusion of signalling peptide to occur. In the provision of synthetic signalling peptides antibiotic disks was placed within approximately 5 mm of the streak and inoculated with 5 µl of up to 10 mM of the peptide dissolved in DMSO. Plates were left for 36-48 hours for incubation before staining, with a wild type control initially being screened to confirm observable levels of granulose accumulation.

Cultures which successfully dug into the agar were flooded with Gram's iodine solution, drained and allowed time for excess unbound iodine solution to evaporate in a vented hood. Cultures which failed to dig into the media and could be displaced by plate flooding, were instead stained using iodine vapour. Iodine crystals were placed on an inverted petri dish lid upon a heating block set to 37 °C, with the base over the lid to hold the subsequent iodine vapour. Plates were exposed to the vapour until the streaks had turned visibly black or yellow at which point they were left open to allow the sublimation of excess iodine from the plate before photographing. Pictures were taken using a Nikon COOLPIX S3700 digital camera.

2.7.5. Gas chromatography

Acetate, butyrate, acetone, butanol and ethanol from *C. acetobutylicum* cultures were quantified by the use of gas chromatography. Samples collected from *C. acetobutylicum* cultures previously described were centrifuged for 5 minutes at 4 °C at 16,000g to obtain a cell-

free supernatant before being stored at -20 °C for no longer than one month. Samples were defrosted on ice and 500 µl extracted for use in the extraction process. 500 µl samples were acidified with 5 µl of 10M H_2SO_4 before the addition of 500 µl of propyl propionate containing 50 mM of valeric acid as an internal standard. The mixture was vortexed for 15 seconds before being centrifuged for 5 minutes at 4 °C at 16,000g. Of the organic (upper) phase 300 µl was transferred into a 2 ml sample vial containing a 300 µl glass insert and the vials were closed with the use of screw-top lips. Samples were analysed chromatographically with the use of a Thermo Focus gas chromatograph equipped with a 30 meter TR-FFAP column (0.25 mm internal diameter) and a flame ionisation detector. The carrier gas supplied at 0.8 ml/min was hydrogen and the flame was maintained by compressed air (350 ml/min), hydrogen (35 ml/min) and nitrogen (30 ml/min). The injector temperature was 240 °C and the detector temperature was 270 °C. The column profile used in peak resolution was 40 °C for an initial 2 minutes, an 80 °C/min ramp to 150 °C and a 45 °C/min ramp to 210 °C at which point the temperature remained for 1 minute. Sample injection volume was 1 µl which was performed by an auto sampler.

For each individual GC run two independent sets of standards were analysed in order to provide adequate calibration. Standards were prepared from 1 M stock standards containing acetone, butanol, butyrate, ethanol and acetate, which were diluted using ELGA water to provide the following standard concentrations 250mM, 100mM, 50mM, 25mM, 15mM, 10mM, 5mM, 2.5mM and ELGA water/CBM-S blanks. Extraction and processing of these standards was performed as described for supernatant samples.

2.7.6. Determination of heat resistant cfu/ml

Quantification of sporulation was measured by enumerating heat-resistant colony-forming units. Upon reaching a desired time point cultures of *C. acetobutylicum* were vortexed and a

100 µl sample transferred into a 1.5 ml reaction tube or one well of a 96 – well PCR plate. Sample were serially 1:10 diluted down to 10^{-4} using PBS through transferring 10 µl to an adjacent tube/well following vigorous mixing through pipetting. Samples were then placed in a heating block at 80 °C for 10 minutes before being transferred back to an anaerobic work station.

CBM plates were divided into four quarters and 10 µl spots of each dilution of a sample was placed into the quarters (with the 10^{-4} dilutions of each sample placed on to a new plate together). Spots were left to dry before inverting the plates and allowing 24 hours before heat resistant CFU's were enumerated.

2.7.7. Activity assay for nanoluc reporters

E. coli and *C. acetobutylicum* cells were grown overnight on LB/CGM agar or in LB or CBM-S broth 6% (w/v) glucose to an OD_{600} of 1.0. The cells were harvested and where appropriate exposed to atmospheric oxygen conditions. Colonies on plates were inoculated with a few drops of a 20 µM furimazine (Promega) solution and placed within a dark room to visualise fluorescence. Colonies from liquid culture were resuspended in PBS containing 50 µM of furimazine and emissions of light at 460 nm measured using a lumino-meter.

2.7.8. Activity assay for mCherry constructs

E. coli and *C. acetobutylicum* cells were grown overnight on LB/CGM agar or in LB or CBM-S broth 6% (w/v) glucose to an OD_{600} of 1.0. Cells on plate were observed for red pigmentation while cells in liquid culture were resuspended in PBS and measured for emission of light at 610 nm when excited at a wavelength of 554 nm using a microplate fluorometer (BMG CLARIOstar). *C. acetobutylicum* cells were measured for activity post oxygen exposure and

additional cell pellets were observed after overnight exposure to oxygen when stored at room temperature and 4 °C.

2.8. In silico software and tools

Nucleotide sequences and identification of homologues was performed using the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) databases and the Basic Local Alignment Search Tool (BLAST). The default nBLAST algorithm i.e. 10 expected chance matches in a random model and a 28 base pair exact match as required minimum seed to initiate an alignment, was used in order to identify homologous genes. Identification of functional domains in the design of ClosTron vectors were identified with the NCBI Pfam online tool using the UniProt reference proteome.

All raw numerical data was stored, calibrated and analysed in Microsoft Excel 2013 (Microsoft). Data was presented graphically and statistical significance levels determined using GraphPad Prism 6 or 7 (GraphPad Software). Significance levels were determined through T-test analysis to identify significant differences between data sets.

3. Characterising the role of RNPP-like proteins Ca_c0957 and Ca_c0958 as stationary phase regulators

3.1. Introduction

RNPP-type quorum sensing systems, widely distributed across Gram-positive bacteria consist of an intracellular RNPP family regulator and a corresponding linear peptide. RNPP regulator are characterised by their tetratricopeptide repeat domains (TPR) which directly interacts with the linear peptide signal following its uptake via ABC transporters (Gallego del Sol & Marina 2013). This interaction results in a change in protein conformation often resulting in multimerisation (Declerck et al. 2007). This, dependent on the N-terminal domain present, can affect gene regulation in several ways, either by allowing DNA binding via a helix-turn-helix motif, by dissociating from a previously bound response regulator thereby unblocking its DNA binding site or by inhibiting phosphatase activity intrinsic to the RNPP protein which would otherwise dephosphorylate a target response regulator (Rocha-Estrada et al. 2010).

In the genera Bacillus the Rap-phosphatases are of key interest for the understanding of stationary phase phenotypes of their respective organisms, as they have been shown to be directly responsible for feeding into the phosphorelay responsible for the activation the master regulator of sporulation Spo0A (Pottathil & Lazazzera 2003). While no evidence for a complete phosphorelay has been observed in any Clostridium (Dürre & Hollergschwandner 2004), the control of stationary phase phenotypes in *C. acetobutylicum* through quorum sensing regulation would be of little surprise as the activation of granulose accumulation, solventogenesis and sporulation phenotypes all occur upon the organism reaching the high cell densities of the stationary phase (Jones & Woods 1986) and dilution of cultures can extend the time required to observe a switch from acidogenic to solventogenic metabolism (S. Y. Li et al. 2011).

Within the *C. acetobutylicum* genome genes for a total of 10 potential RNPP-like regulators were previously identified (Kotte 2013), 8 of which are located alongside a short peptide

sequence suggested to be the precursor to the peptide signal molecule their respective receptors are responsible for binding. The remaining two systems, Ca_c00957 and Ca_c00958 have been suggested by RNA-seq data to be co-transcribed in the same operon at a relatively consistent level throughout the examined growth phases of the organism (Venkataramanan et al. 2013).

In a previous study the 10 quorum sensing systems were inactivated, through ClosTron mutagenesis to assess their effect upon the late phase phenotypes of *C. acetobutylicum* (Kotte 2013). While some effects upon solventogenesis and sporulation could be observed in mutants of the Ca_c0324 RNPP-like regulator, the most significant changes could be observed from the Ca_c0957 and Ca_c0958 inactivation mutants, which displayed a complete loss of granulose accumulation and sporulation phenotypes alongside a dramatic reduction in solvent formation (Kotte 2013). The phenotypes match those described for a *spo0A* mutant (Harris et al. 2002) and Kotte (2013) postulated that Ca_c0957/Ca_c0958 may act as master regulators of stationary phase processes through their regulation of orphan kinases responsible for regulating Spo0A activity. This hypothesis was based upon a preliminary RNA-seq study for the Ca_c0957::CT*ermB* mutant showing the observed down-regulation of two genes encoding Spo0A phosphorylating histidine kinases. However some evidence for Spo0A-independent regulation was also obtained: the loss of granulose could be compensated for by the provision of an AgrD peptide (the signalling molecule employed by the Agr quorum sensing system, see 1.5.5.1.3), suggesting that peptide production is positively controlled by these proteins, and was thus absent in the mutants (in contrast to Spo0A mutants which continue to produce peptide, see chapter 4). This hypothesis was further supported by the massively reduced expression of the *agrBD* operon in a Ca_c0957::CT*ermB* mutant (Kotte 2013).

The characterisation of Kotte (2013) was performed using an intron based technique to insertionally inactivate the RNPP-like genes. However following this study the genetic tools

available for use in *C. acetobutylicum* were further developed to allow for more precise genome editing through in-frame deletion and chromosomal complementation (Ehsaan et al. 2016). This is desirable in the case of Ca_c0957 and Ca_c0958 where polar effects of intron insertion may impact on the expression of the downstream *clpB*, encoding a chaperone with ATPase activity known and in wide variety of studied organisms plays a role in the stress conditions (Allan et al. 1998; Grandvalet et al. 1999; Pan et al. 2012; Shih & Lai 2007).

The lack of a downstream associated peptide and the suggestion that expression of Ca_c0957/Ca_c0958 remains consistent across the growth of the organism, concluded from supplementary data of Venkataramanan et al. (2013), highlights the need to further characterise the means by which these regulators undergo the necessary conformational shift and become active in order to induce the stationary phase phenotypes linked to them. Resolution of the protein structures and/or isolation of bound peptide ligands would serve to confirm the proposed regulator to be a true quorum sensing system and allow for the greater manipulation of the system through the use of system induction using synthetic peptides. Equally the identification of potential binding sites, through techniques such as ChIP-seq would allow the identification of directly controlled genes, i.e. core regulon, and thus the means by which Ca_c0957 and Ca_c0958 cause the large transcriptional alterations observed in inactivation mutants (Kotte 2013).

3.1.1. Aims of this study

The aims of this study were to expand on the previous work carried out by Kotte (2013) to confirm the importance of Ca_c0957 and Ca_c0958 as key regulators of stationary phase phenotypes in *C. acetobutylicum*, using more recently available genetic tool, and develop strains and tools to help further characterise these RNPP-like regulators. Therefore the study had the specific aims:

- To construct clean in-frame deletions for Ca_c0957 and Ca_c0958 and to analyse the phenotypic consequence and corroborate, or disprove previous findings

  With an expanded range of more precise genetic tools can the previous observations with respect to granulose accumulation, solvent production and sporulation, which were based on insertional mutations, be confirmed?

- To construct tagged Ca_c0957 and Ca_c0958 proteins and introduce the encoding genes into *C. acetobuitylicum* for future analysis of protein functionality.

  Does the addition of a peptide tag disrupt or interfere with the activity of the RNPP regulator? Which tag is most effective for Ca_c0957 and Ca_c0958 pull downs for use in future analysis by X-ray crystallography, ChIP-seq or to directly identify a bound ligand?

- To test the suitability of Nanoluc luciferase and mCherry as translational reporters in *C. acetobutylicum*.

There is only a limited range of reporter genes available for solventogenic *C. acetobutylicum* species. Found to be the most reliable in our group is a system based on the chloramphenicol acetyl transferase gene *catP* (Bullifent et al. 1995). While exemplified this system is inconvenient compared to luminescence and fluorescence based systems due to requiring the

collection of cell lysate. Is it possible to use a whole cell reporter system to overcome these limitations to allow easier characterisation of gene expression in *C. acetobutylicum* strains, in particular those regulated by Ca_c0957 and Ca_c0958?

3.2. Results

3.2.1. The role of Ca_c0957 and Ca_c0958 in the regulation of sporulation

To confirm the importance of Ca_c0957 and Ca_c0958 in the regulation of sporulation, in-frame deletion mutants of Ca_c0957 and Ca_c0958 were generated in *C. acetobutylicum* strain ATCC 824 as described in 2.6.2 and demonstrated in Figure 3.1. Subsequently, these mutants were further modified for the purpose of complementation and overexpression, using both native and tagged Ca_c0957 and Ca_c0958 gene variants. Over expression strains (denoted by an O in figures 3.2-3.4) was achieved using the *C. sporogenes fdx* promoter due to its integration into the modular system from which these constructs were derived (Heap et al. 2009) and its previous characterisation as a strong constitutive promoter (Pennington 2006). Importantly, these modifications were based on chromosomally inserted genes via the use of pMTL-ME6C/X constructs (see Table 2.2) negating the need for empty vector controls. In-frame deletion mutants were grown in CBM-S containing 6% (w/v) glucose and sporulation assessed (by measuring heat resistant cfu/ml) at 48 h and 120 h time points alongside the complementation, overexpression and histidine and streptavidin C-terminally tagged variants.

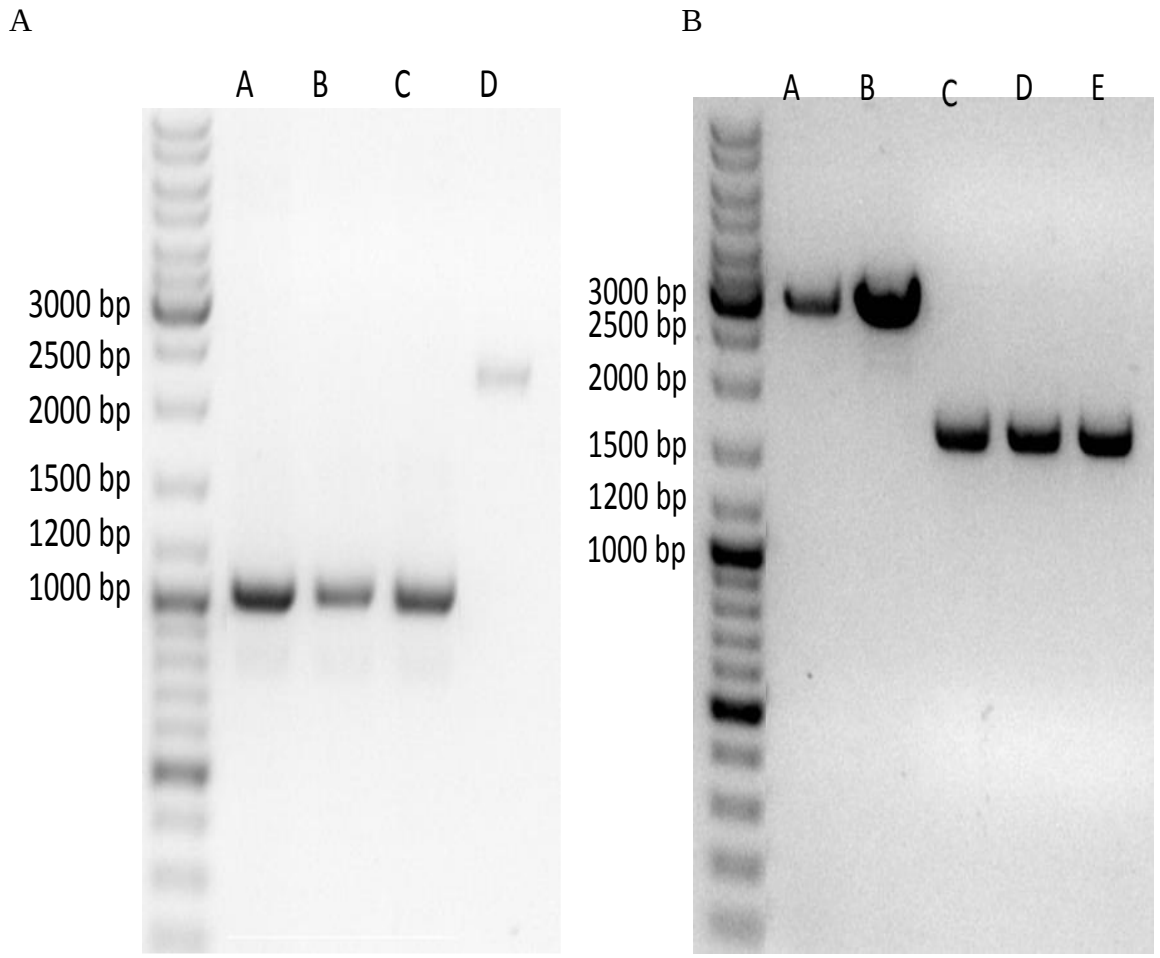
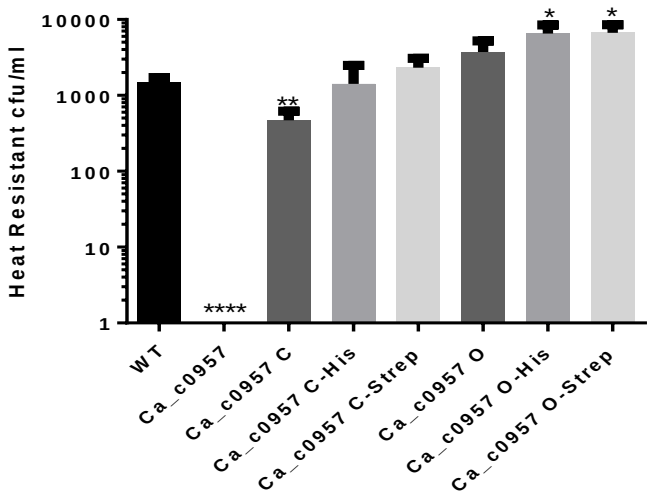


Figure 3. 1 PCR confirmation of in-frame deletion in ΔCa_c0957 and ΔCa_c0958 mutants

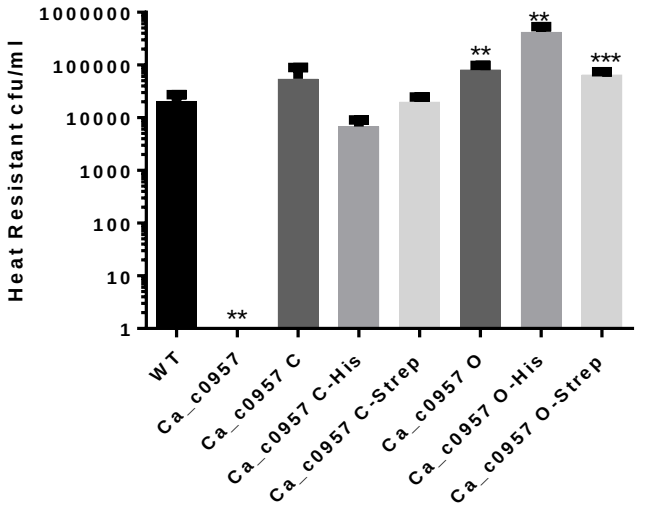
PCR amplification from genomic DNA followed by gel electrophoresis was used to detect in-frame deletion mutants comparing band size to a Gene Ruler mix DNA ladder. (A) Amplification of wild type control gDNA (lane D) using primers Ca_c0957-sF1 and Ca_c0957-sR1 provide a band corresponding to the expected 2360 bp while in-frame deletion mutants, lanes A-C, show amplification of bands of approximately 1073 bp corresponding to the in-frame deletion of Ca_c0957. (B) Amplification of wild type control gDNA (lane A) using primers Ca_c0958-sF1 and Ca_c0958-sR1 shows a band appearing to match the expected band of 2988 bp, lane B shows a selected mutant that failed to undergo a successful double cross over event, while lanes C-E show a band of approximately 1740 bp corresponding to the in-frame deletion of Ca_c0958. All results were confirmed by Sanger sequencing

As can be observed from Figure 3.2, the ΔCa_c0957 and ΔCa_c0958 mutants showed a complete loss of detectable sporulation, with no heat resistant cfu's observed from mutants grown in liquid culture. This was further supported from observations under light microscopy (data not shown) which showed no free spores or phase bright forespores present in the cultures after 120 h. Attempts to recover spores from re-suspending colonies grown on both CGM and CBM-S containing 6% (w/v) glucose agar plates in PBS 10% glycerol also failed to produce any detectable heat resistant cfu's (data not shown).

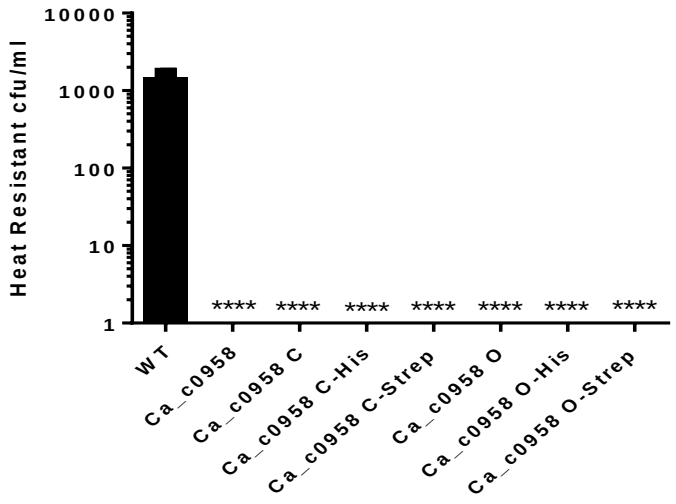
A


Heat Resistant cfu/ml
10000
1000
100
10
1
****
**
*
*
WT
Ca_c0957
Ca_c0957 C
Ca_c0957 C-His
Ca_c0957 C-Strep
Ca_c0957 O
Ca_c0957 O-His
Ca_c0957 O-Strep

B


Heat Resistant cfu/ml
1000000
100000
10000
1000
100
10
1
**
**
**
***
WT
Ca_c0957
Ca_c0957 C
Ca_c0957 C-His
Ca_c0957 C-Strep
Ca_c0957 O
Ca_c0957 O-His
Ca_c0957 O-Strep

C


Heat Resistant cfu/ml
10000
1000
100
10
1
****
****
****
****
****
****
****
WT
Ca_c0958
Ca_c0958 C
Ca_c0958 C-His
Ca_c0958 C-Strep
Ca_c0958 O
Ca_c0958 O-His
Ca_c0958 O-Strep

D

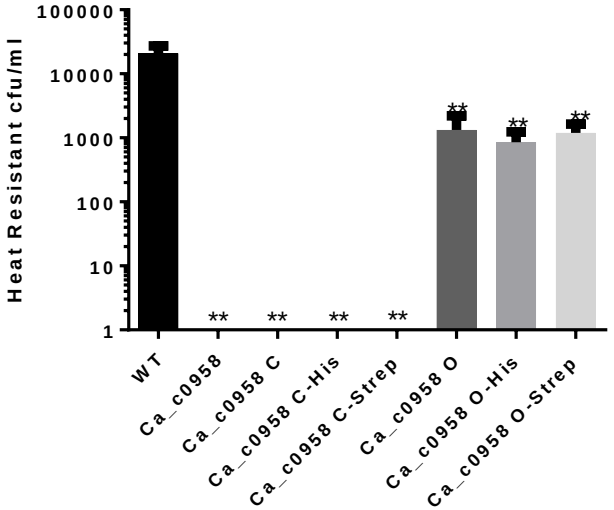

Heat Resistant cfu/ml
100000
10000
1000
100
10
1
**
**
**
**
**
**
**
WT
Ca_c0958
Ca_c0958 C
Ca_c0958 C-His
Ca_c0958 C-Strep
Ca_c0958 O
Ca_c0958 O-His
Ca_c0958 O-Strep

Figure 3.2 Sporulation of ΔCa_c0957 and ΔCa_c0958 mutants and corresponding complementation and overexpression strains

In-frame mutants ΔCa_c0957 (A, B) and ΔCa_c0958 (C, D) were complemented with chromosomally inserted, native Ca_c0957 and Ca_c0958 genes, respectively, as well as their corresponding C-terminally His- and Strep-tagged derivatives. The native and a strong *fdx* promoter were used to achieve expression and overexpression of the genes, respectively. Sporulation levels were assessed in CBM-S, 6% (w/v) glucose at 48 h (A, C) and 120 h (B, D). Expression in ΔCa_c0957 using the native promoter and native, His-, and Step-tagged Cac_0957: Cac_0957 C, Cac_0957 C-His, Cac_0957 C-Strep; overexpression using the fdx promoter and native, His-, and Step-tagged Cac_0957: Cac_0957 O, Cac_0957 O-His, Cac_0957 O-Strep. Expression in ΔCa_c0958 using the native promoter and native, His-, and Step-tagged Cac_0958: Cac_0958 C, Cac_0958 C-His, Cac_0958 C-Strep; overexpression using the fdx promoter and native, His-, and Step-tagged Cac_0958: Cac_0958 O, Cac_0958 O-His, Cac_0958 O-Strep. Results were analysed by T-test analysis against a wild type control, results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

Complementation of Ca_c0957 through integration of the regulator behind an upstream region predicted to host a promoter, at the *pyrE* locus of a ΔCa_c0957 mutant was performed as this complementation strategy has been previously exemplified by Ehsaan et al. (2016). The complementation mutant showed sporulation restored to levels no longer significantly different of that of the wild type control at 120 h Figure 3.2B). Though complementation restored end point sporulation levels, it can be noted that the variation of sporulation levels observed for the complemented strains was generally larger than that of the wild type; furthermore the complemented (untagged) strain at the 48 h time point appeared to show significantly less spore formation Figure 3.2A). Over-expression of Ca_c0957 using the *fdx* promoter appeared to result in sporulation levels that were shown to be significantly higher than that observed in the wild type. This phenomenon that has also been observed when the known stationary phase master regulator Spo0A is relatively overexpressed (Harris et al. 2002) in agreement with the idea that the regulators act through the control Spo0A activity.

ΔCa_c0958 was in parallel engineered to have the Ca_c0958 under the Ca_C0957 promoter shown to be effective in restoring wild type phenotype in the ΔCa_c0957 strain, as published RNA-seq data (Venkataramanan et al. 2013) suggests the genes to be expressed as a single operon and the short region between the genes appeared to host no independent promoter

elements. However, as can be observed in Figure 3.2C and Figure 3.2D, the expression of Ca_c0958 using this promoter region was shown to be insufficient to allow the restoration of the sporulation phenotype in the ΔCa_c0958 mutant background. Sporulation was detectable when the assumed native promoter region was instead substituted for the strong *fdx* promoter, though the number of heat resistant cfu's detected was still significantly lower than that of the wild type, with no spores observed until after 48 h of growth suggesting a delay in endospore formation. These results may suggest that compared to the Ca_c0957 expression some increased level of promoter activity is required for the expression of Ca_c0958 in order to achieve biological activity.

As can be seen from Figure 3.2 the sporulation profiles created by a clean in-frame deletion, support previous observations for ClosTron mutants. ΔCa_c0957 and ΔCa_c0958 show a complete loss in detectable spores, both by light microscopy and through detection of heat resistant cells. No significant variations could be observed between the growth of the wild type *C. acetobutylicum* and the mutant strains (Appendix Figure 1) suggesting the loss of heat viable cells was solely due to a change in the regulation of a stationary phase processes. Complementation experiments provided the evidence to support this by showing the restoration of sporulation in Ca_c0957 mutants. However this sporulation study has observed the necessity of increased expression of Ca_c0958 beyond the level provided by the assumed shared promoter.

Alongside the above complementation and overexpression experiments ΔCa_c0957 and ΔCa_c0958 stains complemented with variant genes encoding C-terminally histidine or streptavidin tagged proteins, controlled by either the native or the *fdx* promoter were analysed for functionality. Figure 3.2A and B confirmed that the tagged Ca_c0957 variants were active as they restored sporulation using the same effective upstream promoter to no longer show a

significant difference from the wild type at both the 48 h and 120 h time points. Overexpression using the stronger *fdx* promoter also resulted in significant increases in sporulation compared to what was observed in the wild type control as had been seen for the native gene's expression, showing up to a 10-fold increase in heat resistant cfu's/ml compared to the former.

As can be observed in Figure 3.2C and D the histidine and streptavidin tagged variants of the Ca_c0958 showed no biological activity when expressed from the native promoter, as was observed for the native gene. Overexpression using the *fdx* promoter was sucessful in restoring sporulation however as was observed with the native gene, the sporulation level were signficantly lower showing an approximate 10-fold decrease in spore levels at 120 h and no evidence of spore formation after 48 h.

3.2.2. The role of Ca_c0957 and Ca_c0958 in the regulation of solvent production

To support the previous claims that Ca_c0957 and Ca_c0958 act as global regulators of stationary phase phenotypes, in a manner similar to the master regulator Spo0A, the solvent profiles of the ΔCa_c0957 and ΔCa_c0958 alongside complemented and overexpression strains were examined after 120 h of growth in CBM-S 6% (w/v) glucose. Strains were primarily assessed for their changes in acetone and butanol concentrations as these had been highlighted as being most heavily affected in previous inactivation mutants (Kotte 2013).

Figure 3.3 confirms the importance of both the RNPP-like regulators importance in the regulation of solventogenesis, as the deletion of these systems resulted in the complete loss of detectable acetone production and a drastic reduction in butanol production. The loss of both these solvent is comparable to what has been observed in Spo0A inactivation mutants (Harris et al. 2002).

Ca_c0957 biological activity was successfully able to restore solvents to levels that were no longer significantly distinguishable from the wild type using the previously described predicted upstream promoter region (Figure 3.3A and B). Unlike sporulation however, overexpression of Ca_c0957 in the ΔCa_c0957 background did not result in an increase in the solvent phenotype in contrast to what had been observed for sporulation. With respect to acetone, overexpression of Ca_c0957 resulted in no significant changes in final acetone concentrations compared to either the wild type or complemented strain. However, butanol concentrations observed when Ca_c0957 was overexpressed were shown to be significantly lower than the wild type showing a fall from 47 mM to 22 mM. Interestingly, this effect resembles that observed for overexpression of *spo0A* (Harris et al. 2002), which also resulted in decreased butanol but similar levels of acetone production.

Similar to the observations for sporulation, the predicted upstream promoter region of Ca_c0957 showed no clear restoration of solvent production when used to attempt to complement the ΔCa_c0958 mutant. Overexpression of Ca_c0958 with the *fdx* promoter provided a slightly differing effects with respect to restoring acetone and butanol concentrations in the ΔCa_c0958 mutant. While the acetone concentration was increased through the use of overexpression of Ca_c0958 it remained significantly lower than levels the levels observed for the wild type. Butanol concentrations after 120 h of growth showed a greater level of standard deviation and were not shown to be significantly different from that observed for the wild type.

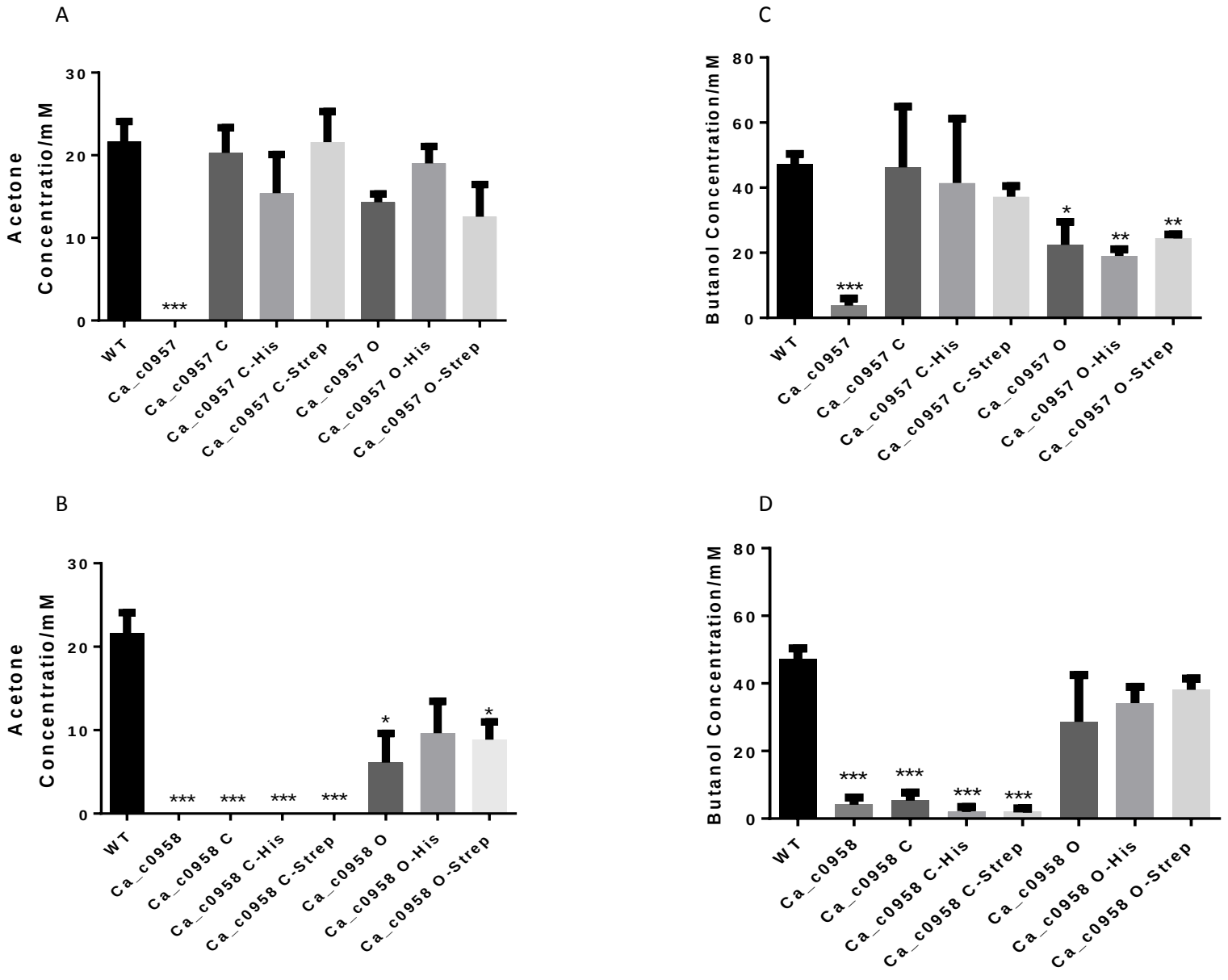

A
Acetone Concentratio/mM
30
20
10
0
***
WT
Ca_c0957
Ca_c0957 C
Ca_c0957 C-His
Ca_c0957 C-Strep
Ca_c0957 O
Ca_c0957 O-His
Ca_c0957 O-Strep
C
Butanol Concentration/mM
80
60
40
20
0
***
*
**
**
WT
Ca_c0957
Ca_c0957 C
Ca_c0957 C-His
Ca_c0957 C-Strep
Ca_c0957 O
Ca_c0957 O-His
Ca_c0957 O-Strep
B
Acetone Concentratio/mM
30
20
10
0
***
***
***
***
*
*
WT
Ca_c0958
Ca_c0958 C
Ca_c0958 C-His
Ca_c0958 C-Strep
Ca_c0958 O
Ca_c0958 O-His
Ca_c0958 O-Strep
D
Butanol Concentration/mM
80
60
40
20
0
***
***
***
***
WT
Ca_c0958
Ca_c0958 C
Ca_c0958 C-His
Ca_c0958 C-Strep
Ca_c0958 O
Ca_c0958 O-His
Ca_c0958 O-Strep

Figure 3.3 Solvent production of ΔCa_c0957 and ΔCa_c0958 mutants and corresponding complementation and overexpression strains

In-frame mutants ΔCa_c0957 (A, B) and ΔCa_c0958 (C, D) were complemented with chromosomally inserted, native Ca_c0957 and Ca_c0958 genes, respectively, as well as their corresponding C-terminally His- and Strep-tagged derivatives. The native and a strong *fdx* promoter were used to achieve expression and overexpression of the genes, respectively. Acetone and butanol concentrations were measured following growth in CBM-S, 6% (w/v) glucose at 48 h (A, C) and 120 h (B, D). Expression in ΔCa_c0957 using the native promoter and native, His-, and Step-tagged Cac_0957: Cac_0957 C, Cac_0957 C-His, Cac_0957 C-Strep; overexpression using the fdx promoter and native, His-, and Step-tagged Cac_0957: Cac_0957 O, Cac_0957 O-His, Cac_0957 O-Strep. Expression in ΔCa_c0958 using the native promoter and native, His-, and Step-tagged Cac_0958: Cac_0958 C, Cac_0958 C-His, Cac_0958 C-Strep; overexpression using the fdx promoter and native, His-, and Step-tagged Cac_0958: Cac_0958 O, Cac_0958 O-His, Cac_0958 O-Strep. Results were analysed by T-test analysis against a wild type control, results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

Tagged Ca_c0957 and Ca_c0958 variants continued to show biological activity at levels similar to the native genes with respect to solvent phenotypes though showed a lower standard deviation in butanol concentrations than the native genes when both were overexpressed (Figure 3.3B and D), overall supporting the hypothesis that the C-terminal tag does not interfere significantly the regulators' with biological activity.

3.2.3. The role of Ca_c0957 and Ca_c0958 in the regulation of granulose accumulation and Agr signalling

To confirm the granulose and Agr phenotypes of ΔCa_c0957 and ΔCa_c0958 mutants, i.e. the postulated absence of granulose accumulation due to a lack of AgrD-derived signalling peptide, the strains were grown on CBM-S agar containing 6% (w/v) glucose for 48 h and granulose within visualised through the use of iodine staining. To observer Agr-signalling dependent interactions, strains were cross-streaked against an *agrB*::CT*ermB* mutant to detect the presence of the diffusible AgrD signal peptide for the strains examined.

Through the use of iodine staining it was observed that ΔCa_c0957 and ΔCa_c0958 mutants showed an inability to accumulate the energy storage compound granulose (Figure 3.4). As expected previously, complementation and overexpression mutants derived from the

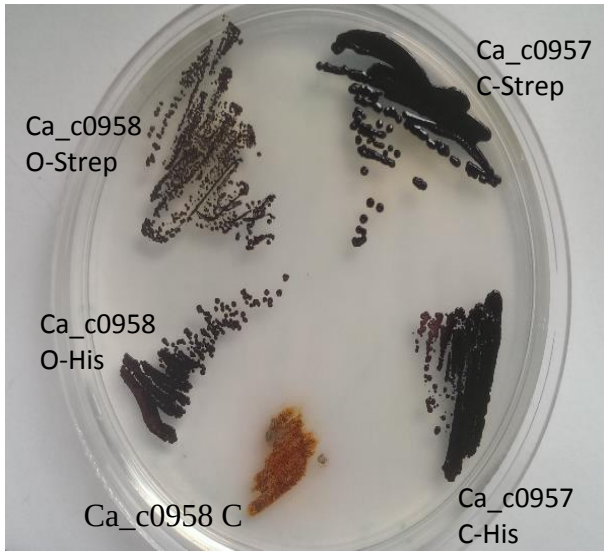
ΔCa_c0957 strain were subsequently shown to restore the lost phenotype and accumulated granulose (Figure 3.4B). However similar to sporulation and solvent formation, ΔCa_c0958 showed no granulose accumulation when the complementing gene was integrated at the *pyrE* locus and expressed under the control of the Ca_c0957 promoter region. Granulose accumulation was however restored when expression of Ca_c0958 was placed under the control of the *fdx* promoter. Due to the nature of the granulose staining, the relative levels of granulose accumulation could not be accurately quantified to determine if complementation or overexpression mutants synthesised levels of granulose significantly different relative to those observed in the wild type. Figure 3.3 also shows that this loss in granulose accumulation was directly linked to the known quorum sensing regulation previously described in *C. acetobutylicum*, the Agr system (Steiner et al. 2012).

By streaking ΔCa_c0957 and ΔCa_c0958 in the presence of a wild type strain, granulose accumulation could be seen as observed in the region of the streak with the closest proximity to the wild type *C. acetobutylicum* (Figure 3.4A). This showed that the wild type is able to produce the necessary signal molecule that the ΔCa_c0957 and ΔCa_c0958 mutants are deficient in and was responsible for the loss of granulose accumulation. The ΔCa_c0957 and ΔCa_c0958 were shown to still maintain expression of the signal response regulator required to induce granulose accumulation, which is a phenotype matched by what was observed in an *agrB*::CT*ermB* mutant (Steiner et al. 2012).

A


WT
Ca_c0957
WT
Ca_c0958

B


Ca_c0957
C-Strep
Ca_c0958
O-Strep
Ca_c0958
O-His
Ca_c0958 C
Ca_c0957
C-His

C


agrB::CTermB
Ca_c0957

D


agrB::CTermB
ΔCa_c0957 C

Figure 3.4 Granulose and Agr signalling in ΔCa_c0957 and ΔCa_c0958 mutants and corresponding complementation and overexpression strains

(A) ΔCa_c0957 and ΔCa_c0958 mutants show a granulose deficient phenotype complemented by Agr signalling from a wild type. (B) Granulose accumulation in a ΔCa_C0958 overexpressing His- and Strep- tagged ΔCa_c0958, ΔCa_c0957 expressing His- and Strep-tagged Ca_c0957 from the native promoter. ΔCa_c0958 expressing Ca_c0958 from the native promoter does not show granulose accumulation. (C) ΔCa_c0957 shows deficiency of Agr quorum sensing signalling as it fails to induce granulose accumulation when cross-streaked against an *agrB*::CT*ermB* mutant. (D) Restoration of Agr signalling in a ΔCa_c0957 mutant expressing Ca_c0957 from the native promoter produces the AgrD derived peptide signal and induces granulose accumulation in an *agrB*::CT*ermB* mutant. Strains were grown in liquid CBM-S containing 6% (w/v) glucose to mid exponential phase before being streaked on to CBM-S agar containing 6% (w/v) glucose and incubation for 48 h. Granulose was visualised by iodine staining. Expression in ΔCa_c0957 using the native promoter and native, His-, and Step-tagged Cac_0957: Cac_0957 C, Cac_0957 C-His, Cac_0957 C-Strep; overexpression using the fdx promoter and native, His-, and Step-tagged Cac_0957: Cac_0957 O, Cac_0957 O-His, Cac_0957 O-Strep. Expression in ΔCa_c0958 using the native promoter and native, His-, and Step-tagged Cac_0958: Cac_0958 C, Cac_0958 C-His, Cac_0958 C-Strep; overexpression using the fdx promoter and native, His-, and Step-tagged Cac_0958: Cac_0958 O, Cac_0958 O-His, Cac_0958 O-Strep

The deficient signalling was further confirmed as that of the Agr quorum sensing system through the inability of the in-frame deletion mutants to induce granulose accumulation in an *agrB*::CT*ermB* mutant (Figure 3.4C). The complemented strains, which accumulated granulose, also induced granulose accumulation within the *agrB*::CT*ermB* mutant, indicating a restoration of AgrD signalling peptide production.

The other significant phenotype described in *agr* mutants by Steiner et al. (2012), is a 10-20 fold decrease in sporulation levels when grown in liquid culture and the almost complete elimination of detectable spores when grown on agar. Growth of ΔCa_c0957 and ΔCa_c0958 in the CBM-S broth 6% (w/v) glucose with the addition of 20 µM of a synthetic AgrD derived peptide Ac-R6T0-Ca (described in 5.2.1) was attempted to observe whether sporulation could be partially restored, however no restoration of sporulation could be observed (data not included).

As can be seen from Figure 3.4B the use of tagged Ca_c0957 was shown to provide biological activity at a sufficient level to restore granulose accumulation when under the control of its native promoter or the *fdx* promoter. The Ca_c0958 tagged variants were again only shown as

restoring the stationary phase phenotype under control of the *fdx* promoter as had been observed for sporulation and solvent production. Though not quantifiable the results agree with the previous conclusion that the tagged showed activity similar to the native genes.

3.2.4. The Use of His-tagged Ca_c0957 and Ca_c0958 proteins

Having confirmed their biological activity, the initial aim was to purify the histidine-tagged Ca_c0957 and Ca_c0958 variants after heterologous production in *E. coli* for confirmation of hetero-dimer formation and use in crystallography. Purification from *C. acetobutylicum* in order to, resolve the structure of the ligand binding domain (in the event the correct folding was dependent upon the lower internal pH of the *C. acetobutylicum* cytoplasm), attempt to co-purify any ligand peptides bound to the TPR domain and identify the formation of Ca_c0957-Ca_c0958 multimers in the native background. A histidine tag was initially chosen at the recommendation of an X-ray crystallography specialist due to their familiarity with the system and ease of forming fusion constructs in *E. coli* pET expression vectors.

The purification of histidine-tagged fusions from *E. coli* was achieved through growth in LB using 0.5 mM IPTG induced expression from a pET-21d vector containing a Ca_c0957 or Ca_c0958 histidine-tagged fusion. Following the concentration of histidine-tagged proteins by affinity tag purification the protein samples were subject to SDS-PAGE and visualised by Coomassie blue staining with an NEB colour plus (10-230 kDa) ladder used to determine protein sizes (with the predicted sizes for Ca_c0957-His and Ca_c0958-His being 51.29 kDa and 51.22 kDa respectively . As can be seen from Appendix Figure 2, this process allowed for the effective isolation of both tagged Ca_c0957 and Ca_c0958 from *E. coli*.

To purify tagged proteins from *C. acetobutylicum* overexpression strains (ΔCa_c0957 ME6X Ca_c0957 and ΔCa_c0958 ME6X Ca_c0958), alongside a wild type control were grown in 100 ml CBM-S 6% glucose until they reached a target OD_{600} of 0.8-1.0. These cultures were subject

to the same column purification and staining process previously described. No clear band could be observed for either protein though this might be due to limited column binding due to the presence of a large background band observed in all *C. acetobutylicum* samples between the 150 kDa and 230 kDa markers on the protein ladder (Appendix Figure 2).

Variation of time points with respect to harvesting the cells was ineffective in reducing the prevalence of the background bands, which remained highly expressed and limited the purification of the His-tagged Ca_c0957 and Ca_c0958 fusion proteins. It was therefore decided to switch to a new protein tag in *C. acetobutylicum* and for consistency in any future structural analyses. It was also decided to switch the tag in the *E. coli* expression system for consistency between the two systems.

3.2.5. The use of Streptavidin-tagged Ca_c0957 and Ca_c0958 proteins

A streptavidin based tag and purification system was chosen as a replacement, as this had been effectively demonstrated in *C. acetobutylicum* in previous work by Girbal et al. (2005). As with the histidine tag fusion, the streptavidin tagged fusions were introduced to the ΔCa_c0957 and ΔCa_c0958 mutants to assess their functionality as described in 3.2.1 to 3.2.3. To maintain consistency with *E. coli* work streptavidin tagged variants were also cloned into pET-21d for expression from *E. coli*. Initial attempts to purify Strep-tagged Ca_c0957 and Ca_c0958 proteins from *E. coli* was unsuccessful (data not shown) and due to time constraints was not repeated.

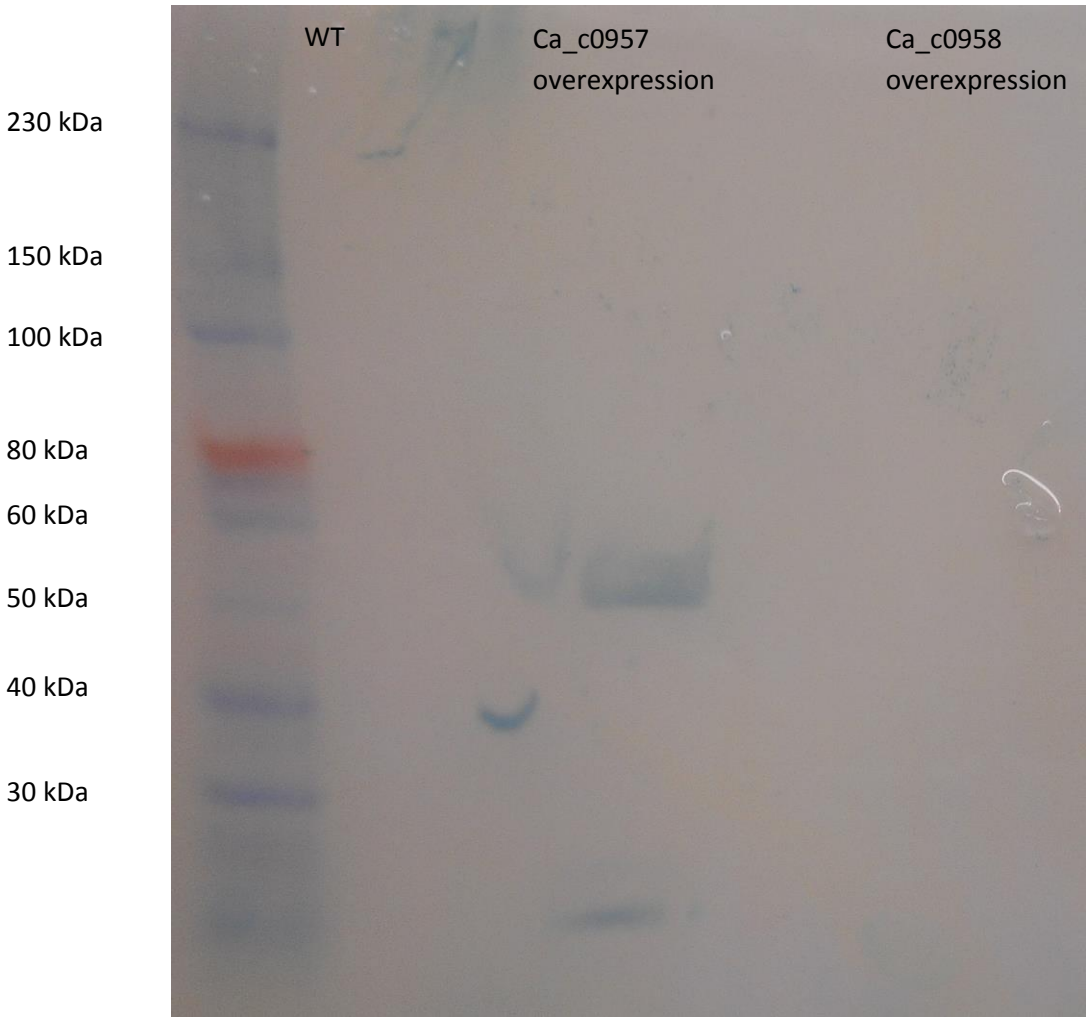


Figure 3.5 Detection of Streptavidin-tagged proteins in *C. acetobutylicum* crude lysates through Western blotting

Lysates of *C. acetobutylicum* ATCC 824 wild type (lane 2), a Ca_c0957 overexpression strain (ΔCa_c0957 ME6X Ca_c0957) (lane 3) and Ca_c0958 overexpression strain (ΔCa_c0958 ME6X Ca_c0958) (lane 4) were subject to SDS-PAGE and western blotting using an Anti-Strep-HRP conjugate antibody. Lane 1 NEB colour Plus (10-230 kDa) protein ladder. The expected molecular weights were 51.51 kDa and 51.44 kDa for Ca_c0957 and Ca_c0958 respectively

Initial attempts to purify the streptavidin tagged Ca_c0957/Ca_c0958 proteins from *C. acetobutylicum*, using the overexpression constructs (grown as described previously in 3.2.4) failed to show detectable yields when subject to SDS-PAGE and Coomassie blue staining. Due to time constraints subsequent attempts to repeat strep-tag purification to increase yields were limited, however western blot analysis as shown by Figure 3.5 did show the detectable presence of a 51.51 kDa band corresponding to the streptavidin-tagged Ca_c0957 protein from crude lysate of the overexpression strain (though this was not observed for Ca_c0958 which should

display a band of corresponding to 51.44 kDa). Figure 3.5 shows that the Strep-tag sequence is likely a more viable route for purification from *C. acetobutylicum* as no additional sequences were recognised from the Ca_c0957-Strep over expressing strain outside of the expected protein band (with no other bands detected for the other lanes, with the exception of some occurrence of leaking from the Ca_c0957-Strep lysate well). The use of Streptavidin tagging has been established within *Clostridium* and from the Western blot results the system appears a viable route for the future isolation of fusion tagged proteins for subsequent structural analysis.

3.2.6. Creation of reporter constructs

Available RNA-seq data (Venkataramanan et al. 2013) analysis suggested a constitutive expression of the Ca_c0957 and Ca_c0958 operon throughout the growth of *C. acetobutylicum*, even though as a proposed quorum sensing system it would likely have a positive feedback loop (Haas et al. 2000). With no known signal peptide it would be of interest to investigate whether any up regulation of the Ca_c0957 and Ca_c0958 receptors directly occurs throughout the growth of *C. acetobutylicum* using reporter constructs. These reporter constructs would also be of use should future ChIP-seq allow the identification of Ca_c0957 and Ca_c0958 binding sites and thus target genes to confirm the changes in gene expression directly caused by these regulators.

While a CatP based reporter system exists for use in *Clostridium* (Bullifent et al. 1995), it has primarily been used in past studies to limited observation of activity to exponential phase growth, during which stationary phase regulators Ca_c0957 and Ca_c0958 are likely to have a lower activity. Due to difficulty experienced in obtaining *C. acetobutylicum* cell lysates from late stationary phase cultures it was decided to investigate the use of techniques that did not require cell lysis and would be more suited to rapid sample analysis (as they would not require

complex sample processing) . It was therefore decided to investigate the use of potential who cell reporters, which at time of writing have not yet been established within solventogenic *Clostridium*.

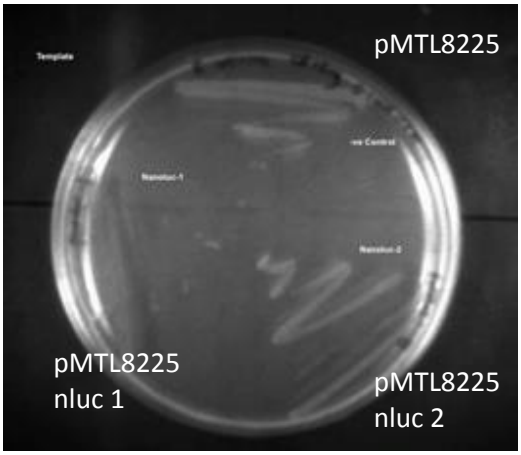
3.2.6.1. Testing nanoluc and mCherry based whole cell reporter

Due to the limitations of the CatP reporter and the general scarcity of published reporter systems available in *C. acetobutylicum* it was decided to test both a nano-luciferase (nanoluc) and mCherry derived systems. Although reliant on oxygen and therefore unsuitable for real-time monitoring, these systems were considered promising candidates for the creation of reporters that would not require less extensive sample preparation.

The nanoluc system was investigated as potential reporter due to its relatively small size (compared to other luciferase enzymes), high physical stability and no evidence of requiring post-translational modification or formation of di-sulphide bonds (Hall et al. 2012) suggesting it may be more likely to correctly fold and show activity in *C. acetobutylicum* compared to other luciferase systems. Two synthetic codon optimised nanoluc reporters, the first optimised by Biomatik and the second by simple selection of most frequently used codons in *C. acetobutylicum* (Musto et al. 2003), were obtained from the commercial provider, were placed under the control of the *fdx* promoter and introduced into *C. acetobutylicum* on a modular plasmid pMTL85143 (Heap et al. 2009). The functionality of the construct was first assessed in *E. coli* TOP10 grown in LB medium, this work was carried out as part of an associated Msc project by Kameswari Tallapragada. Expression of both nanoluc constructs was assessed on plate (Figure 3.6) and in liquid (not shown) and shown to be active. Following this the activity of the nanoluc constructs in *C. acetobutylicum* was assayed both by overnight growth on both CBM-S broth and plate for 16 hours and grown in liquid culture to an OD_{600} of 1.0 before being incubated aerobically with the substrate furimazine. Light emissions were assayed at 460 nm

for liquid cultures or were observed directly in a dark room for plate based assays. Despite both constructs showing functionality in *E. coli* both tested methods and nanoluc constructs were not shown to be functional in *C. acetobutylicum* (data not shown).

A



B

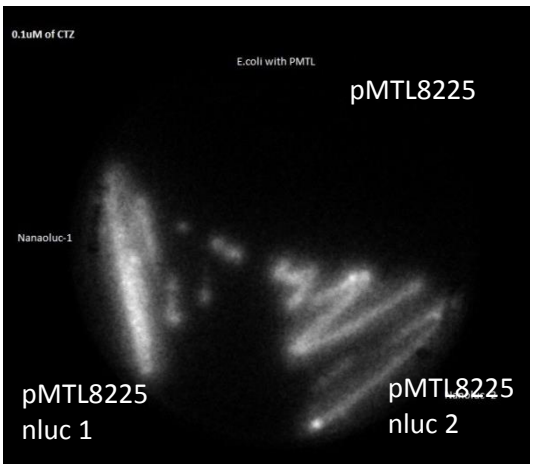


Figure 3.6 Activity of nanoluc reporters expressed in *E. coli* TOP10.

Figure provided by Kameswari Tallapragda

E. coli TOP10 colonies containing the overexpression vector pMTL82252-nluc1, pMTL82252-nluc2 and plasmid control pMTL82252 were grown on LB agar plates and tested for luminescence following the addition of the furimazine substrate. A faint glow from the colonies was visible using the naked eye when colonies on an LB agar plate (A) were placed in a dark room at higher substrate concentrations (100 μM). (B) The light output of individual colonies visualised at a substrate concentration of 0.1 μM using a light camera box (HAMAMATSU light box).

The mCherry reporters which were selected based upon a system recently demonstrated within *C. difficile* (Ransom et al. 2015) and an additional variant codon optimised by the supplier Biomatik were also cloned behind an *fdx* promoter and introduced into *C. acetobutylicum* using a modular plasmid pMTL85143. Though not screened *E. coli* TOP10 carrying the pMTL85143-mCherry1 and pMTL85143-mCherry2 were observed as displaying a red colouration when grown on LB agar or broth (data not shown). As with the nanoluc reporter the cultures were grown on both plate for 16 h and in liquid culture to an OD_{600} of 1.0 before being exposed to oxygen. Plates were left in excess of 24 hours and colonies observed for any potential change in colour. While cells in liquid culture were centrifuged at 10,000g to form a pellet and resuspended in volumes of PBS to create a dilution gradient. These cell suspensions were excited at a wavelength of 554 nm and below and potential emissions were detected at a wavelength of 610 nm and above every 15 minutes for up to 8 h.

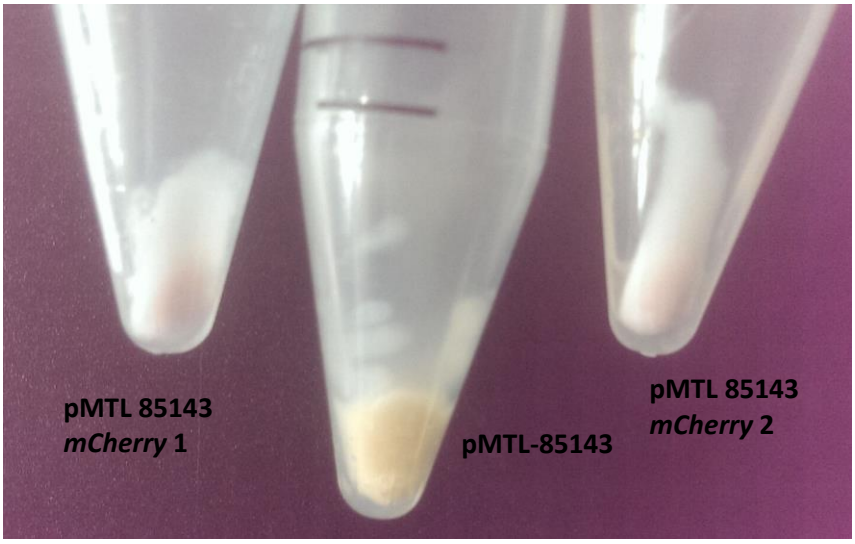


Figure 3.7 mCherry activity in *C. acetobutylicum*

mCherry activity could be observed in the indicated *C. acetobutylicum* strains expressing codon-optimised *mcherry* constructs 1 and 2 (left and right tube respectively) expressed from a modular plasmid pMT85143 using the *fdx* promoter, while a control strain containing an empty pMTL85143 plasmid showed no change in the colour of the pellet.. Strains were grown overnight in CBM-S 6% glucose before being centifuged at 10000g and left overnight to become aerobic to allow for mCherry activation.

The mCherry constructs were visibility active and expressed in *E. coli*, and post oxygen exposure eventually showed activity in *C. acetobutylicum* when grown in liquid culture though for unknown reasons this was not observed on plate. However the use of this reporter was still limited, as unlike Ransom et al. (2015) which observed peak mCherry activity in *C. difficile* following only 4 h of exposure to an aerobic environment, *C. acetobutylicum* displayed no detectable activity within 8 h post oxygen exposure (data not shown) despite undergoing the same treatment process. Evidence of mCherry activity was only initially observed in culture cell pellets (and later cell suspensions) that were left exposed to oxygen overnight for 16-18 h, at which point a red/pink decolouration could be observed as shown in Figure 3.7. Attempts to keep cultures at 4 °C in order to slow processes in the oxygen exposed dying cell leading to protein degradation also delayed the period required for mCherry to mature beyond 16 h. Due to the ill-defined and unrepresentative state of the cells after 16 h, when they showed peak mCherry activity, it was decided to disregard the use of this protein as a potential reporter system for this study.

3.3. Discussion

Ca_c0957 and Ca_c0958 have been shown to be critically important regulators with respect to granulose accumulation (through their regulation of Agr signalling), solventogenesis and sporulation. Deregulation of these phenotypes is clearly observed in the absence of either Ca_c0957 or Ca_c0958 and is similar to the changes observed for the loss of the function of Spo0A with respect to sporulation and solventogenesis. The comparative phenotypes of ΔCa_c0957 and Δ*spo0A* mutants with respect to the regulation of granulose accumulation is further expanded upon in chapter 4. This work has cemented the assertions made by Kotte (2013) through the use of in-frame deletion to show that the previously observed phenotypes were not the result of polar activity.

Previous RNA-seq data generated for Ca_c0957::CT*ermB* mutants (Kotte 2013) supports the hypothesis that Ca_c0957/Ca_c0958 and Spo0A activity are interlinked, as it showed a down regulation in the orphan kinases Ca_c3319 and Ca_c0903. which have previously been characterised as a directly phosphorylating Spo0A (Steiner et al. 2011). Inactivation of both kinases has been shown to eliminate sporulation (Steiner et al. 2011), a phenotype matching that observed in ΔCa_c0957 and ΔCa_c0958 mutants and providing an explanation for a mechanism by which Spo0A activity is decreased and justify the loss of sporulation and solventogenesis phenotypes in these mutants. Though a likely mechanism for the down regulation of stationary phase phenotypes it is not clear whether the down-regulation of Spo0A phosphorylation is absolute within the ΔCa_c0957 and ΔCa_c0958 mutants, or whether some level activity remains. The potential activity of Spo0A in a ΔCa_c0957 background is further explored in chapter 4.

The effects associated with Ca_c0957 overexpression support the hypothesis that Ca_c0957/Ca_c0958 exert their effect through Spo0A as similarities to a Spo0A overexpression strain can be observed (Harris et al. 2002). Harris et al. (2002) describe a Spo0A overexpressing strain to show an increase in sporulation at the cost of solvent production, when compared to a wild type, hypothesising that overexpression leads to accelerated spore formation at the expense of overall solvent production. Ca_c0957 over-expression shows significantly higher sporulation levels than the wild type after 120 h, while showing significant reductions in butanol concentrations, similar to that described in a Spo0A overexpression mutant with acetone remaining at similar levels to the wild type. The similar changes in sporulation and solvent levels observed in Ca_c0957 and Spo0A deletion and overexpression mutants strongly supports the overall hypothesis that Ca_c0957 primarily regulates stationary phenotypes through histidine kinase expression and the subsequent level of Spo0A activity within the cell.

Evidence from the RNA-seq data of Kotte (2013) alongside the in-frame deletion mutant phenotypic observations both confirm that the loss of granulose accumulation, in ΔCa_c0957 and ΔCa_c0958 mutants, occurs through the severe down regulation of the *agrBD* operon resulting in the loss of AgrD derive peptide signalling. What remains unclear however, is whether this regulatory mechanism is a direct result of the loss of activity of Ca_c0957/Ca_c0958 or a subsequent downstream effect upon a reduction in Spo0A phosphorylation. Spo0A is considered the master regulator of sporulation, which granulose accumulation is closely associated with, and Spo0A inactivation has been observed to lead to a loss of granulose accumulation in *C. beijerinckii* (Dürre & Hollergschwandner 2004; Harris et al. 2002). It may therefore be possible that the observed repression of Agr signalling is simply a consequence of the downregulation of Spo0A activity, as opposed to a more direct form of regulation by Ca_c0957 and Ca_c0958. Identifications of regulatory mechanisms, outside of the known orphan kinases Ca_c0903 and Ca_c3319, are also of interest investigation in order to fully understand the regulatory hierarchy that governs stationary phase phenotypes with many potential genes of interest, described as showing large changes in expression profiles by Kotte (2013), that have yet to be investigated. The regulatory interactions, hierarchy and other potential regulatory targets of interest to these systems has been further explored in chapter 4.

Previous (insertional) Ca_c0958::CT*ermB* mutants were complemented through the expression of the gene behind the strong *fdx* promoter on a multi-copy plasmid (Kotte 2013), which showed phenotypic restoration. However this would have been predicted to result in large overexpression of the gene compared to the promoter region upstream of Ca_c0957. The use of in-frame deletion and chromosomal complementation in this study however revealed that the use of this promoter region, while effective for the complementation of ΔCa_c0957 mutants is ineffective in restoring Ca_c0958 expression to biologically active levels. However, the use

of a single copy of the strong *fdx* promoter at the *pyrE* locus did result in the restoration of granulose accumulation and also partial restoration of sporulation and solvent phenotypes.

The ΔCa_c0958 mutant only showing phenotypic restoration under over-expression conditions would suggest some form of regulation of the system in addition to its known transcription in an operon alongside Ca_c0957. The demonstration that over-expression using a known strong *fdx* promoter only served to partially restore the lost phenotypes, while Kotte (2013) showed restoration when using multiple copies of this expression system, may imply that the expression of Ca_c0958's independent regulation plays a significant role in the regulatory mechanism of the Ca_c0957 and Ca_c0958 system. In quorum sensing systems high levels of expression of system components are controlled by a positive feedback loop that induces continual high expression of constituent parts of the system (Fuqua et al. 2001). While no elements within the Ca_c0957 gene have been able to be identified as distinct regulatory elements, from what is seen from the inability to fully complement using the upstream region and the inferred high expression required in order to restore the lost phenotypes, may be explained by the presence of a positive feedback loop responsible for upregulating Ca_c0958 activity. However as protein production also relies on both efficient transcription and translation, it is also possible the latter was affected in the generated constructs as the gene was taken out of its native sequence context.

Through further analysis of the section of the operon between the initial promoter region and the Ca_c0958 gene, through the use of sequence extension and subsequent analysis of phenotype restoration, it may be possible in the future to identify the mechanism(s) by which Ca_c0958 is presumably upregulated in order to achieve biologically functional levels of expression within *C. acetobutylicum*.

In order to further characterise the mechanism by which Ca_c0957 and Ca_c0958 act as master regulators it was decided to create tagged fusions variants of the proteins. With respect to the choice of terminus to tag, the N-terminus was rejected due to its close proximity to the helix-turn-helix domains of each regulator and therefore the increased probability of the tag disrupting the regulators DNA binding functions. The TPR domain is a greater distance from the C-terminus and the latter was therefore chosen as the site for the proteins tags. The tagged proteins were also checked for biological activity in ΔCa_c0957 and ΔCa_c0958 mutants, due to a hypothesis previously raised, namely that the downstream peptides of RNPP regulators may have evolved from peptide structures at the C-terminus of their respective regulators (Declerck et al. 2007). The lack of a downstream peptide and this proposed hypothesis raised the possibility that these RNPP-like regulators may still interact with their C-terminal amino acids via folding into their TPR domains as opposed to using a signal peptide. While without subsequent analysis it cannot be shown that this interaction does not occur within the Ca_c0957 and Ca_c0958 regulators, the tagged protein variants showed biological activities at levels matching those of the native structure. This suggests that either the ligand peptide is not derived from the C-terminus of Ca_c0957/Ca_c0958 or is sufficiently upstream of the amino acid tags used in this study to allow any processing and/or folding required to bind to the TPR domain to occur unhindered by the addition of the tagging sequence.

Though functionally active, the subsequent analysis planned for use with the histidine tagged variant constructs was deemed unfavourable for purifying larger amounts of the Ca_c0957 and Ca_c0958 proteins due to the occurrence of contaminating proteins when purification was attempted from the native *C. acetobutylicum* however, this was not observed in *E. coli*. In spite of attempts to modify the procedure through variation of harvesting points and culture conditions this was not able to be successfully overcome. The decision was therefore taken to switch to a streptavidin based tagged fusion, due to its previous successful use in *Clostridium*

(Girbal et al. 2005). Western blot analysis has indicated that no Strep-tag like sequence present in the wild type that would likely contaminate column purification samples and successfully detected the Strep-tagged Ca_c0957 from its overexpression strain (though failed to do so for the Strep-tagged Ca_c0958). The streptavidin tagged expression strains will be the basis of the future work described below in further characterising the structure, peptide ligands and DNA binding sites of Ca_c0957 and Ca_c0958.

Due to the proximity of the Ca_0957 and Ca_c0958 genes and the matching phenotypes observed upon their deletion it is probable that these genes act upon the same regulatory pathways likely due to the formation of a heterodimeric structure with one another. Resolving dimerisation would be possible through the use of native PAGE for analysis of pulled down structures, as well as identification of proteins contained within these structures by mass spectrometry (Brymora et al. 2001). Antibodies against the tags would also potentially allow immunoprecipitation of Ca_c0957/Ca_C0958 proteins required for ChIP-seq analyses to enable identification of the binding sites of the helix-turn-helix motifs identification of the binding sites of the helix-turn-helix motifs (Furey 2012), from which sites of potential interest, informed by previous RNA-seq data, can be analysed (Kotte 2013). The identification of binding site would be of use in: characterising genes directly regulated by Ca_c0957, i.e. the core regulon and those altered by Spo0A activity. The identification of relevant promoter regions might provide the means by which regulation of stationary phase phenotypes can be manipulated (i.e. induced at earlier time points or lower cell densities) for better strain generation.

Previously published RNA-seq data suggests the Ca_c0957-Ca_c0958 operon remains at a near consistent level of expression throughout the growth of *C. acetobutylicum* (Venkataramanan et al. 2013). This suggests that the activation of these regulators, leading to the up regulation of

granulose accumulation, solventogenesis and sporulation must occur due to a change in the activity of the regulators, likely due to the a conformational shift as a result of a ligand binding in the TPR domain. With no known ligand identified, tagged proteins would be potentially able to be purified and their structure resolved by X-ray crystallography in order to suggest a potential ligand that would fit into the binding pocket. Conversely it may be able to directly identify a signalling peptide bound to a TPR domain through the pull down of active Ca_c0957 and Ca_c0958 from *C. acetobutylicum*. Identification of a binding peptide would not only help to better characterise these systems but also allow for the more effective control over the initiation of stationary phase phenotypes from *C. acetobutylicum*. A precedence exists in the form of the AI-2 quorum sensing receptor LuxP, which after purification and crystallisation revealed the structure of its AI-2 ligand which had remained in complex with the protein during the purification process (Chen et al. 2002). Furthermore purified functional Ca_c0957, Ca_c0958 and potential heteromeric complexes may also be used to test whether they bind any of the eight known signalling peptides produced by other, complete RNPP quorum sensing systems present in *C. acetobutylicum*. Such studies could be undertaken using synthetic peptide and purified protein and methods such as Isothermal Titration Calorimetry (ITC) or Surface Plasmon Resonance (Zouhir et al. 2013b; Patching 2014).

As mentioned above currently available RNA-seq data showing little change in the expression level of Ca_c0957 and Ca_c0958 (Venkataramanan et al. 2013) though this is partially due to the limited number of time points available and the respective culture conditions tested. Future analysis of Ca_c0957/Ca_c0958 gene expression or other targets of interest in *C. acetobutylicum* through a reporter (which are currently limited in availability for *C. acetobutylicum*) would allow more convenient analysis of expression levels across a larger range of time points and conditions.

Analysis of transcription by proxy of a reporter using the CatP system is well documented in *Clostridium* (Takamizawa et al. 2004), however previous studies have primarily been centred on the identification of strong promoters active as early as the exponential phase of the organism. This system does however require extensive sample treatment to extract cell lysate and is therefore less convenient for use in extensive reporter assay experiments.

At present no whole cell luminescent based reporter system has been effectively established for use in *C. acetobutylicum*, as systems often fail to function under the anaerobic, highly reduced or lower pH environment created by the organism or are undermined by the native emission spectrum of *C. acetobutylicum* which accumulates riboflavin (Katagiri et al. 1958). Although reporter proteins for in vivo fluorescence without oxygen have been described almost a decade ago (Drepper et al. 2007), and can be purchased for use in *Clostridium* spp (https://www.jenabioscience.com/images/495a55f185/FP-21060-21066.pdf) these and related proteins were previously tested in our laboratory (Minton and Winzer, personal communication) and found not to display fluorescence above the background level. It was decided to test the Nanoluc system from NEB due to its small size and lack of di-sulphide bonds suggesting it may be less prone to mis-folding and an mCherry system that had been optimised and shown to be effective in *C. difficile* (Ransom et al. 2015).

The Nanoluc system showed no activity when expressed within *C. acetobutylicum* despite showing functionality in *E.coli* and was subsequently discarded. The reasons for the lack of functionality were not established, but may be linked to insufficient expression due to inappropriate codon usage or the mis-folding in the (acidic) cytoplasm of *C. acetobutylicum* cell. The mCherry system was shown to be able to correctly fold in *C. acetobutylicum* cells following sufficient oxygen exposure, however unlike the previous studies performed in *C. difficile* optimal fluorescence was achieved far in excess of the 4 hours post sampling

previously described by Ransom et al. (2015) Due to the impracticality of the time required for oxygen exposure at temperatures above 4 °C this system was not further investigated for use in this study.

In conclusion, the role of Ca_c0957 and Ca_c0958 as master regulators of granulose accumulation, solvent production and sporulation are now well established, providing a basis to explore regulatory hierarchy in chapter 4. Observation of deletion mutants and overexpression of Ca_c0957 suggests these genes primarily operate through the regulation of Spo0A activity due to the respective similarities between Spo0A inactivation and overexpression mutants with respect to sporulation and solvent phenotypes. Through the use of clean in-frame deletion mutants and chromosomal complementation it has also established that Ca_c0958 requires expression in excess of what was achieved when coupling the gene to its operon's native promoter, a limitation that had previously gone unobserved due to use of plasmid-based complementation using the strong *fdx* promoter. C-terminally tagged fusion proteins have been shown to be functionally active and can serve as the basis for future structural function analysis of Ca_c0957/Ca_c0958 alongside identification of potential ligand peptides and DNA binding sites for future characterisation and strain manipulation. Attempts to establish whole cell reporter systems to use as a proxy for gene expression remain limited in *C. acetobutylicum*, with the currently favoured reporter system, CatP, requires extensive sample preparation in order to extract cell lysates and the inconvenience associated with this. The fluorescent Nanoluc and mCherry systems explored in this study were found to be ineffective or limited for the purpose of an effective whole cell activity reporter system. Further exploration of reporters in the future may be of use for measuring expression of genes, such as Ca_c0957 and Ca_c0958, as the lack of such a system still limits work in this field.

4. Understanding the role of RNPP-type regulators Ca_c0957 and Ca_c0958 in the hierarchy of stationary phase regulation

4.1. Introduction

As demonstrated in chapter 3, the RNPP regulators Ca_c0957 and Ca_c0958 are of clear importance in the regulation of granulose accumulation, solventogenesis and sporulation within *C. acetobutylicum*. Previous studies have shown, through RNA-seq, analysis that a Ca_c0957::CT*ermB* mutant showed considerable changes in its transcriptional profile compared to the wild type in agreement with the loss of stationary phase phenotypes observed (Kotte 2013). However the precise role(s) that Ca_c0957 and Ca_c0958 play in the regulation of these phenotypes remains uncertain, although some evidence suggest they may be involved in the activation of Spo0A (Kotte 2013).

Spo0A in *C. acetobutylicum* has long been known to act as a master regulator of the characteristic late phase phenotypes of solventogenesis and sporulation (Ravagnani et al. 2000), with *spo0A* mutants showing phenotypes (Harris et al. 2002), matching those described for ΔCa_c0957 and ΔCa_c0958 mutants suggesting a direct regulatory link between the three proteins. Spo0A activation has been shown in *C. acetobutylicum* to be directly the result of the activity of orphan histidine kinases (Steiner et al. 2011), two of which were observed as being down regulated in a Ca_c0957::CT*ermB* mutant (Kotte 2013). Ca_c3319 and Ca_c0903 when inactivated have been shown to cause a reduction in sporulation, that becomes an absolute loss in sporulation when the genes are simultaneously inactivated (Steiner et al. 2011). Though these histidine kinase mutants were not assessed for their granulose and solventogenic phenotypes, the hypothesis suggested by Kotte (2013) is that these phenotypes are likely also lost or at least drastically reduced when Ca_c3319 and Ca_c0903 are down regulated in a Ca_c0957::CT*ermB* mutant.

The Ca_c0957::CT*ermB* mutant also showed an approximately 900 fold down regulation in the expression of *agrBD* operon (Kotte 2013), which accounts for the loss in granulose

accumulation and the ability to respond to the provision of an AgrD derived signal as described in chapter 3. Inactivation of Spo0A in solventogenic *C. beijerinckii* and σ^G in *C. acetobutylicum* show granulose negative phenotypes (Ravagnani et al. 2000; Tracy et al. 2011), suggesting that its regulation is part of or linked to the classical endospore regulatory cascade. What is not clear however, is whether the regulation of the Agr system by Ca_c0957 is the result of its dependence on Spo0A activation and subsequent regulatory cascade or whether *agrBD* expression is independently regulated by Ca_c0957 and Agr signalling would remain active in a Δ*spo0A* mutant.

Furthermore the attempt to place Ca_c0957 and Ca_c0958 as a master regulator system acting above Spo0A is further complicated by the recent suggestion that σ^k, classically believed to be a late phase regulator of sporulation, acts as a regulator of granulose accumulation, solventogenesis and sporulation through altering the expression of *spo0A* (Al-Hinai et al. 2014). The creation of $\Delta\sigma^k$ mutants in that study resulted in a phenotype matching characteristics observed in ΔCa_c0957, ΔCa_c0958 and Δ*spo0A* mutants, although they were not rescued with respect to Agr signalling.

Interestingly it can also be observed from RNA-seq data obtained from Ca_c0957::CT*ermB* mutants that a significant down regulation occurred in the genes either side of the *agrBD* and *agrCA* loci (Kotte 2013). The Ca_c0077 and Ca_c0082 genes when analysed for conserved domain through BLAST are suggested to be a CAAX-like membrane endopeptidase and a LanC like peptide cyclase respectively. It was decided to investigate whether these genes also played a role in Agr signalling as their co-localisation and implied regulation by Ca_c0957 may indicate an association with peptide signalling due to their potential roles as peptide cleavage/modification enzymes.

4.1.1. Aims of this study

The aims of this study were to investigate the questions described above in order to expand our knowledge of the regulatory hierarchy of late phase regulators and the specific genes and phenotypes under their respective controls. Specifically, this study aimed to answer the following:

- Is the lossof granulose accumulation in Ca_c0957 mutants the result of decreased Spo0A activity?

  Is the observed decrease in *agrBD* expression seen in Ca_c0957 mutants also observed in Δ*spo0A* or is this directly regulated by Ca_c0957?

- If *agrBD* expression is controlled by Ca_c0957 but not Spo0A does the latter independently control granulose accumulation via a different mechanism?

  If Spo0A mutants are known to be granulose negative but remain capable of Agr signalling does granulose accumulation require both regulatory pathways in order to be enabled?

- Does σ^k play a role in early phase regulation of stationary phase phenotypes?

  If σ^k is confirmed as early phase regulator of sporulation, solventogenesis and granulose accumulation is there any indication as to its potential interaction with Ca_c0957 regulation and Agr signalling?

- Do the Ca_c0077 and Ca_c0082 genes either side of the *agr* locus play a role in cell to cell signalling?

Does the down regulation of these genes in a Ca_c0957::CT*ermB* mutant, also deficient in Agr signalling also imply the potential involvement of these genes in Agr signalling/regulation?

4.2. Results

4.2.1. Regulation of granulose accumulation by Spo0A

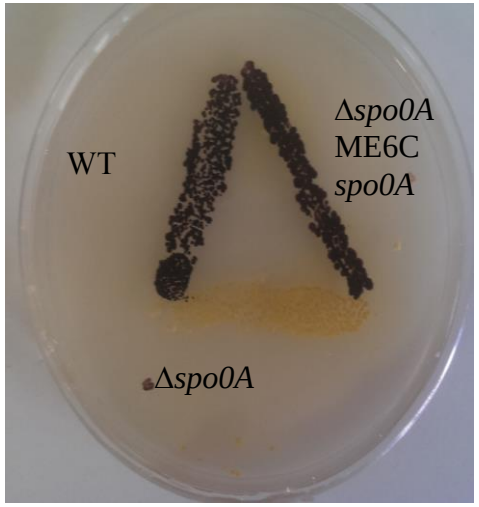
The phenotypic characteristics of ΔCa_c0957 and ΔCa_c0958 mutants, alongside the previously gathered RNA-seq data had suggested that RNPP regulators govern the late phase phenotype primarily through the phosphorylation of Spo0A (Kotte 2013). To test the necessity of Spo0A for the induction of granulose accumulation a provided in-frame deletion Δ*spo0A* mutant and the complemented Δ*spo0A* ME6C *spo0A* stain (Ehsaan et al. 2016) were grown on CBM-S 6% (w/v) glucose and stained using iodine after 48 h of growth. Furthermore to test whether a loss of granulose accumulation was a consequence of a loss in *agrBD* expression as observed in ΔCa_c0957 and ΔCa_c0958 mutants, the Δ*spo0A* mutant was streaked against an *agrB*::CT*ermB* mutant.

As can be seen from Figure 4.1A and B, Δ*spo0A* mutant showed no granulose accumulation but was still able to induce a response in an *agrB*::CT*ermB* mutant, showing the mutant was capable of producing the Agr signal. This discrepancy between the *Δspo0A* and the ΔCa_c0957/ΔCa_c0958 deletion phenotypes suggested that Spo0A may play an essential but independent role in the formation of granulose in contrast to the suggested control by Ca_c0957 through the Agr system. For this purpose *spo0A* was excised by in frame deletion from ΔCa_c0957 mutants to confirm the independent nature of Spo0A mediated control through observing changes in granulose accumulation when cross-streaked against appropriate mutants.

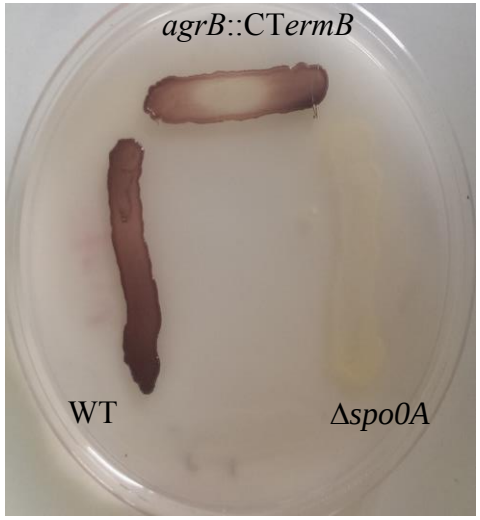
Figure 4.1C showed that the new double knockout strain (generated by removing *spo0A* from the ΔCa_c0857 mutant) resulted in: a granulose negative phenotype, continued repression of

agrBD expression and loss of the ability to respond to an Agr signal and subsequently initiate granulose production. Like the Δ*spo0A* mutant, ΔCa_c0957Δ*spo0A* also lacked the ability to form phase bright endospores or heat resistant cfu's (data not shown). As already suggested for the Δ*spo0A* mutant, Spo0A was confirmed as essential for the initiation of granulose accumulation, as restoration of *spo0A* by allelic coupled exchange restoring Spo0A at the *pyrE* locus of the ΔCa_c0957Δ*spo0A* double mutant allowed the observation of granulose accumulation in the presence of the wild type generated Agr signal (Figure 4.1)

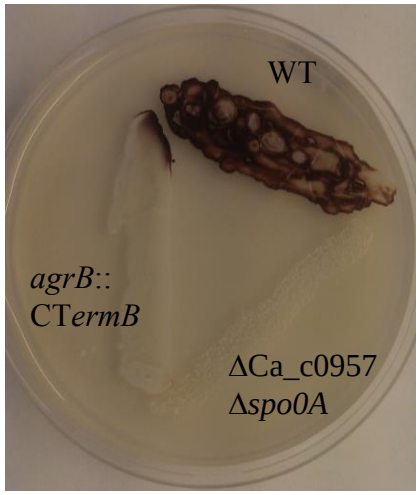
A


WT
Δ*spo0A*
ME6C
spo0A
Δ*spo0A*

B


agrB::CT*ermB*
WT
Δ*spo0A*

C


WT
agrB::
CT*ermB*
ΔCa_c0957
Δ*spo0A*

D

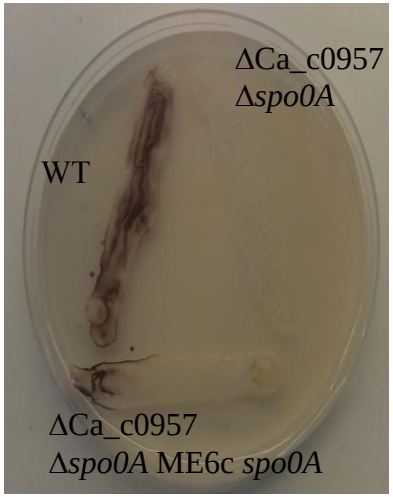

ΔCa_c0957
Δ*spo0A*
WT
ΔCa_c0957
Δ*spo0A* ME6c *spo0A*

Figure 4.1 Spo0A control of granulose production independent of Agr signal generation

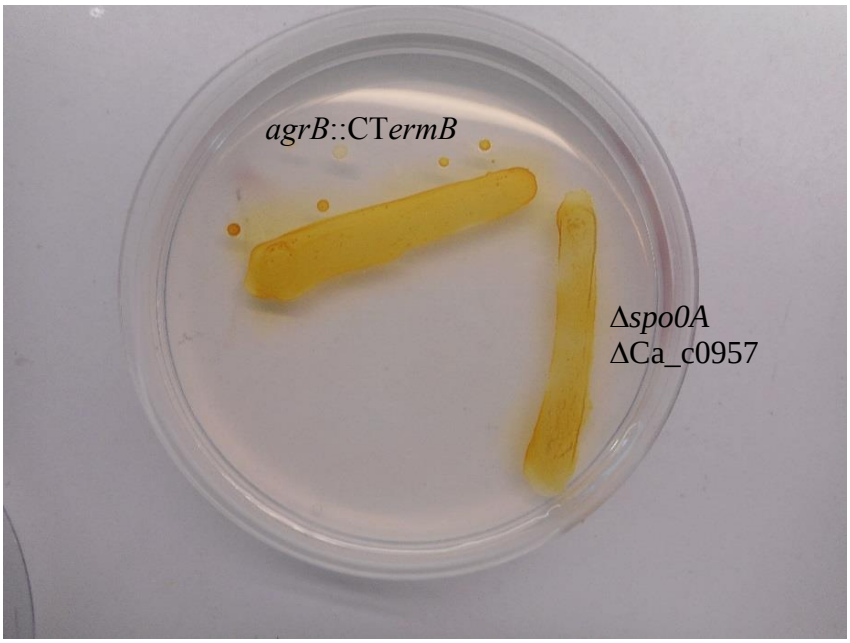
(A) Δ*spo0A* mutants are deficient in granulose production (B) Δ*spo0A* mutants produce the Agr signal, (C) the deletion of *spo0A* in a ΔCa_c0957 background results in the loss of response to Agr signalling and (D) restoration of Agr signalling in a ΔCa_c0957Δ*spo0A* double mutant following its complementation with *spo0A.* Strains were grown in liquid CBM-S containing 6% (w/v) glucose to mid exponential phase growth before being streaked on to CBM-S agar with containing 6% (w/v) glucose and incubated for a further 48 h. Granulose was visualised by iodine staining.

4.2.2. Ca_c0957 regulation of granulose accumulation independent of Spo0A activity

While Spo0A was shown to be essential for granulose accumulation independent of Agr signalling, the latter is known to play a crucial role (Steiner et al. 2012) and in chapter 3 has been proposed as the primary reason of the loss of granulose accumulation in ΔCa_c0957 and ΔCa_c0958 mutants. To confirm the hypothesis that Ca_c0957 regulation of the Agr serves as an independent, but still critical, secondary regulatory pathway in the activation of granulose accumulation, the ability of a Δ*spo0A* strains to induce granulose accumulation in *agrB*::CT*ermB* mutants was assessed following the excision and repair of Ca_c0957 by in-frame deletion and allelic coupled exchange.

An independently generated Δ*spo0A*ΔCa_c0957 double mutant showed a phenotype matching the ΔCa_c0957Δ*spo0A* double mutant described previously, confirming the observed phenotypes of both accumulating no granulose and failing to show evidence of Agr signalling. The shared phenotypes from both background strains of the double mutants confirms the change in phenotypes observed was not the consequence of random mutagenesis or degeneration of cultures due to the repeated streaking involved in the generation of multiple phases of double cross-over mutants. Upon complementation of Ca_c0957at the *pyrE* locus, the restoration of the Agr signalling could clearly be observed through the induction of granulose accumulation when cross streaked with an *agrB*::CT*ermB* mutant.

A



B

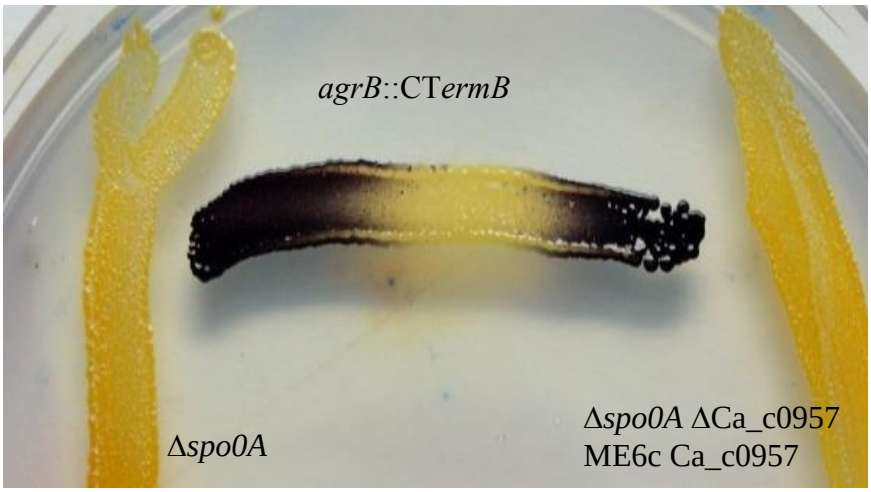


Figure 4.2 Ca_c0957 regulation of granulose accumulation indirectly through Agr signalling

(A) Deletion of Ca_c0957 resulting in the loss of Agr signal production in a Δ*spo0A* mutant and (B) the restoration of signalling in a Δ*spo0A* ΔCa_c0957 upon complementation with Ca_c0957. Granulose accumulation in the signalling peptide deficient *agrB*::CT*ermB* mutant served as a phenotypic marker for successful signalling Strains were grown in liquid CBM-S containing 6% (w/v) glucose to mid exponential phase growth before being streaked on to CBM-S agar with containing 6% (w/v) glucose and incubated for a further 48 h. Granulose was visualised by iodine staining.

4.2.3. The role of σ^k as a regulator of stationary phase phenotypes

4.2.3.1. Growth of a $\Delta\sigma^k$ mutant

$\Delta\sigma^k$ mutants were generated by in-frame deletion (see Figure 4.3) to test whether σ^k is involved in the regulation of stationary phenotypes as previously claimed (Al-Hinai et al. 2014) and whether this regulation is linked in some manner to that exerted by Ca_c0957/Ca_c0958. To test the claim that $\Delta\sigma^k$ mutants were unable to regulate acid production to the extent that strains studied by Al-Hinai et al. (2014) acid crashed within 24 h when grown in batch culture $\Delta\sigma^k$ mutants were grown in CBM-S broth containing 6% (w/v) glucose and the OD_{600} used to monitor growth and indicate the occurrence of cell death.

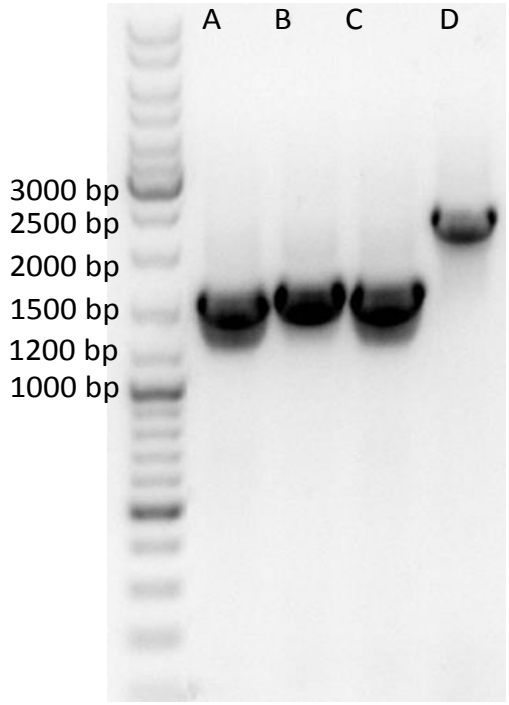


Figure 4. 3 PCR confirmation of in-frame deletion of $\Delta\sigma^k$ mutants

PCR amplification from genomic DNA followed by gel electrophoresis was used to detect in-frame deletion mutants, with band sizes estimated using a Gene Ruler mix DNA ladder. Amplification of wild type control (lane D) using primers Cac-σ^K-sF1and Cac-σ^K-sR1provided a band corresponding to the expected band of 2217 bp while in-frame deletion mutants, lanes A-C, show amplification of bands of 1073 bp corresponding to the in-frame deletion of σ^K.

The $\Delta\sigma^k$ mutants' growth profile showed no significant difference with respect to the rate of exponential growth or change in optical density following entry into stationary phase compared to what was observed in the wild type. This was in clear contrast to what had previously been described by Al-Hinai et al. (2014), which claimed that batch culture was not viable for fermentations over 24 h and due to the accumulation of toxic levels of acids, pH controlled batch fermentations were the only viable means of growing the mutant strain. Direct comparison of $\Delta\sigma^k$ mutants growth profiles was not possible beyond this as outside of the acid crash description no growth profiles were provided by Al-Hinai et al. (2014)

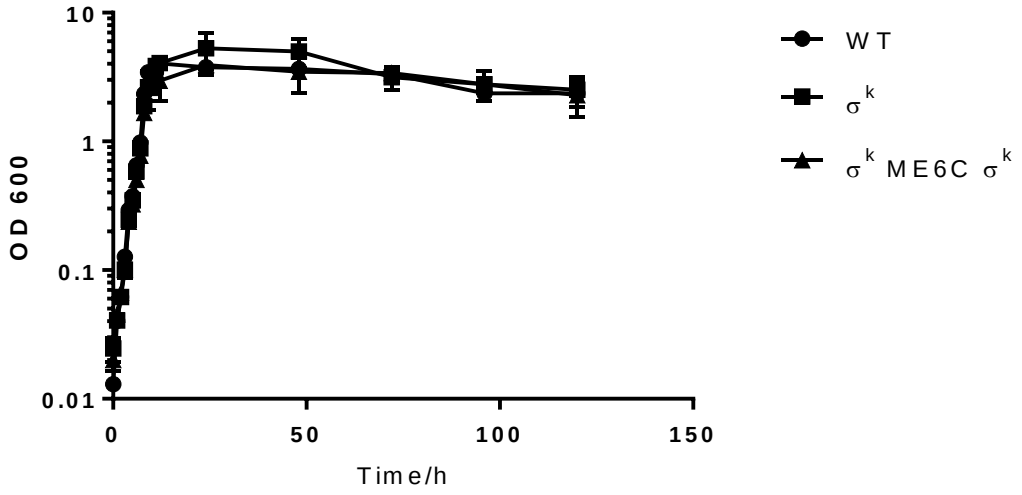


Figure 4.4 Growth profile of $\Delta\sigma^k$ mutants

Strains grown in CBM-S broth 6% (w/v) glucose following inoculation from overnight cultures to an approximate starting OD_{600} of 0.01 and OD_{600} measurements were taken every hour until the cultures appeared to enter stationary phase, after which measurements were made every 24 h. Included were the ATCC 824 parent strain (WT), the $\Delta\sigma^k$ (σ^k) mutant and the complemented mutant (σ^k ME6C σ^k). All strains were grown in biological and technical triplicates

Throughout the growth experiments no evidence of acid crashing was observed within 120 h, as shown in Figure 4.4 as no sudden drop in OD_{600} could be observed after entry into stationary phase. Should the strains have undergone an acid crash as describe by Al-Hinai et al. (2014), the cells would have been expected to outperformed the buffering capability of the calcium carbonate and as a result following an initial period of exponential growth undergone cell lysis causing a drop in the cultures $OD_{600.}$ While the pH was not measured in these cultures, the initial 0.5% calcium carbonate supplied within the medium was never visually observed as having been exhausted supporting the optical density observations. The discrepancy between the growth/acid crashing profiles suggests that the mutants generated by Al-Hinai et al. (2014) showed an inability to down regulate acid production and/or initiate solvent production and acid reassimilation, which is unlikely to have occurred within the $\Delta\sigma^k$ mutants generated in study.

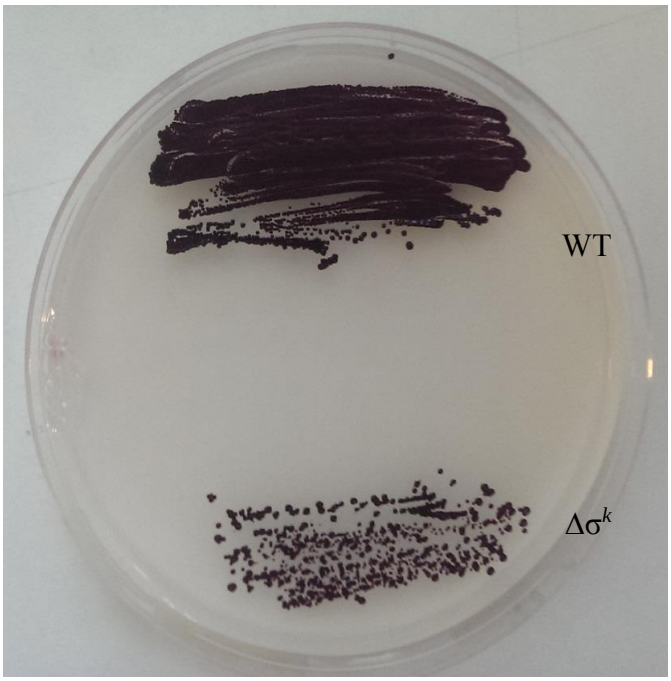
4.2.3.2. Granulose accumulation and Agr signalling in a $\Delta\sigma^k$ mutant

Due to the identification of the two independent regulatory mechanisms for granulose accumulation, as previously described in 4.2.1 and 4.2.2 , $\Delta\sigma^k$ mutants which had been suggested by Al-Hinai et al. (2014) to be granulose negative, were streaked in the presence of both an *agrB*::CT*ermB* mutants and wild type *C. acetobutylicum* and grown for 48 h on CBM-S agar containing 6% (w/v) glucose. This would have indicated whether the proposed loss of granulose accumulation occurred and was a result of Agr signalling deficiency.

However as can be seen from Figure 4.5, the $\Delta\sigma^k$ mutants after 48 h growth on CBM-S agar, displayed no loss of granulose accumulation. Furthermore the $\Delta\sigma^k$ mutants were also shown to have a functional Agr signalling system as they showed the capability to induce granulose accumulation in an *agrB*::CT*ermB* mutant. The occurrence of both these phenotypes within the $\Delta\sigma^k$ mutants would suggest that no significant disruption of either Ca_c0957 or Spo0A

granulose regulation pathways, both of which have been shown previously to be essential for granulose formation, occurs within the mutants.

A



B

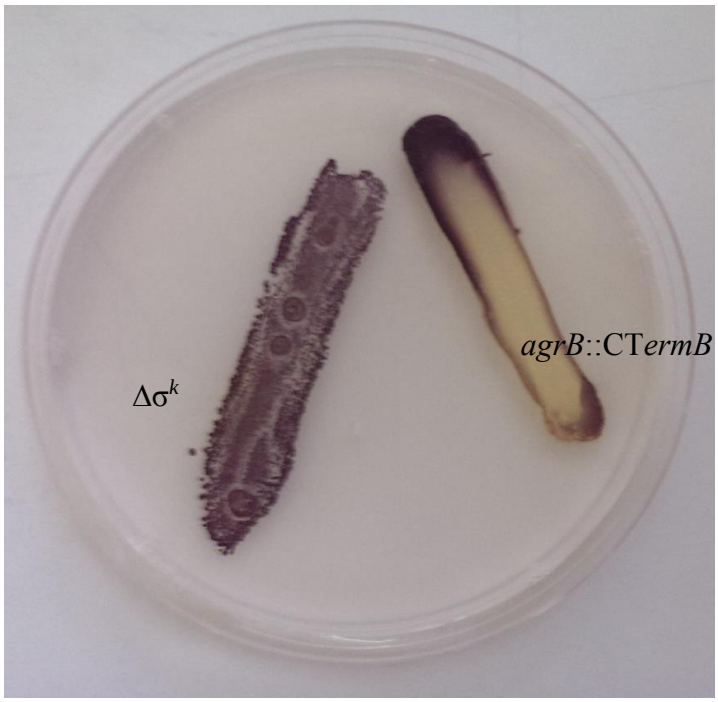


Figure 4.5 Granulose accumulation and Agr signalling in σ^k mutants

(A) Granulose accumulation and (B) Agr signal production by a *C. acetobutylicum* σ^k mutant. Strains were grown in liquid CBM-S containing 6% (w/v) glucose to mid exponential phase growth before being streaked on to CBM-S agar containing 6% (w/v) glucose and incubated for a further 48 h. Granulose was visualised by iodine staining.

4.2.3.3. Solvent profile of a Δσ^{k} mutant

As mentioned above, Δσ^{k} mutants generated by Al-Hinai et al. (2014) were unable to grow in non-pH controlled batch culture, due to an observed acid crash within 24 h of growth, though growth within a pH controlled culture was achieved and the Δσ^{k} mutants were observed as producing acetone and butanol levels comparable to a Δ*spo0A* mutant. The lack of reassimilation and solvent production was also reported to lead to the production of excess butyrate and acetate. The decrease in the solvent profile and lack of any evidence of acid reassimilation was claimed to be the likely result of *spo0A* down regulation.

Δσ^{k} mutants generated in this study were able to be grown in traditional batch culture and were subsequently cultured in CBM-S 6% (w/v) glucose for 120 h after which their solvent levels were assessed. Unlike observations of Al-Hinai et al. (2014), these mutants showed no significant change in acetone or butanol levels following the completion of these fermentations (Figure 4.6). In support of this and the growth observation in 4.2.3.1 the accumulation of acetate and butyrate was never shown as significantly exceeding what was observed in the wild type (Appendix Figure 3).

Thus, the Δσ^{k} mutants generated in this study seemed fully capable of switching to solvent production and reassimilating butyrate and acetate, in contrast to what was observed by Al-Hinai et al. (2014). The results of the solvent profile analysis of the Δσ^{k} mutants supports the observations of the granulose accumulation assays, in that they suggested that Spo0A and Ca_c0957/Ca_c0958 appear to remain biologically active based upon their associated phenotypes following the deletion of σ^{k}.

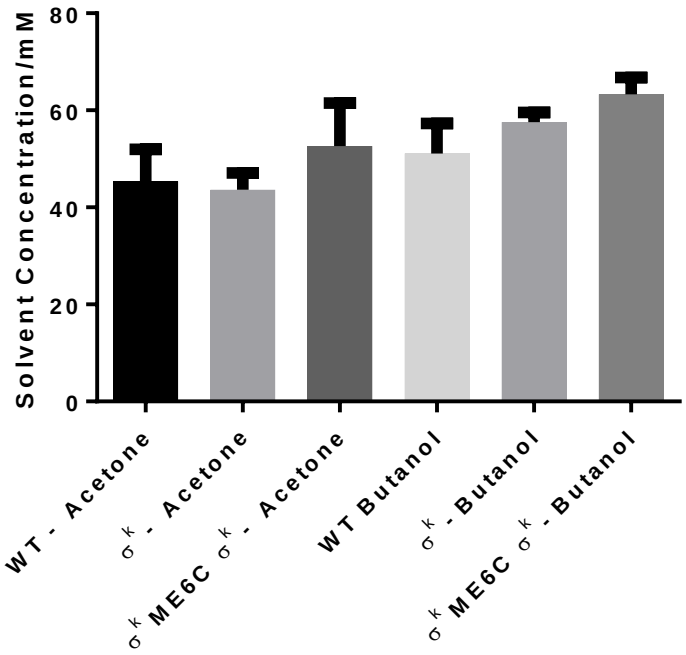


Figure 4.6 Acetone and Butanol production by σ^k mutants

Acetone and Butanol levels of strains inoculated from overnight cultures of CBM-S 6% glucose and grown for 120 h in CMB-S broth 6% (w/v) Included were the ATCC 824 parent strain (WT), the $\Delta\sigma^k$ (σ^k) mutant and the complemented mutant (σ^k ME6C σ^k). No significant differences were found from the wild type control in either of the strains. No significant difference where observed. n=9

4.2.3.4. Sporulation in a $\Delta\sigma^k$ mutant

Outside of the early regulatory role described by Al-Hinai et al. (2014), σ^k mutants were expected to show no capacity to form complete endospores due to their role in the criss-cross regulation cascade of endospore formation (Dürre & Hollergschwandner 2004). To confirm that this remained the case the $\Delta\sigma^k$ mutant strains were grown in CBM-S 6% glucose and sporulation assessed by heat resistant cfu/ml at 48 h and 120 h.

$\Delta\sigma^k$ mutants showed a complete loss of detectable heat resistant cfu's and no phase bright spores were seen using light microscopy (Figure 4.7 and 4.8). These result confirmed the

expected role of the sigma factor and also that the deletion of σ^k was successful as it is a known conserved component of the sporulation cascade of *Clostridium* and *Bacillus* spp. (Dürre & Hollergschwandner 2004). Complementation of the strain through allelic coupled exchanged restored sporulation to levels statically indistinguishable from the wild type. These results are in agreement to what was observed by Al-Hinai et al. (2014) who showed that restoration of sporulation was only achieved through the reintroduction of σ^k, independent of the role they claimed it to play in *spo0A* regulation, supporting that it still acted as a late phase regulator.

A

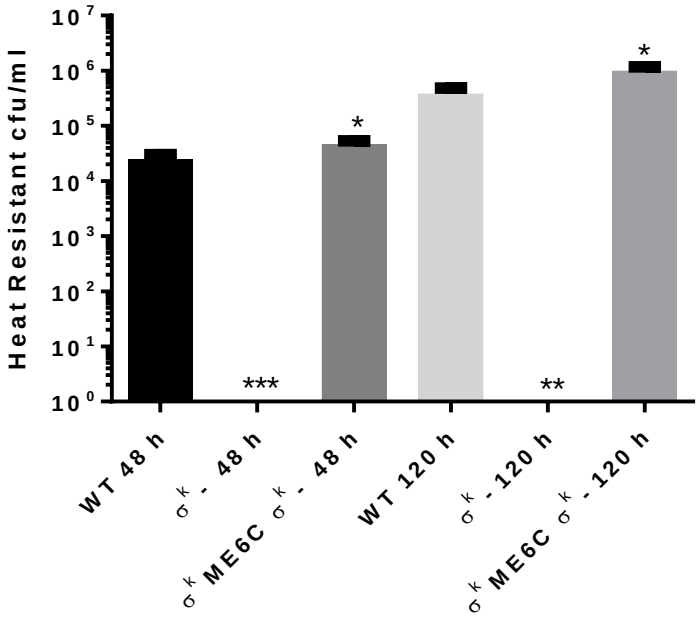
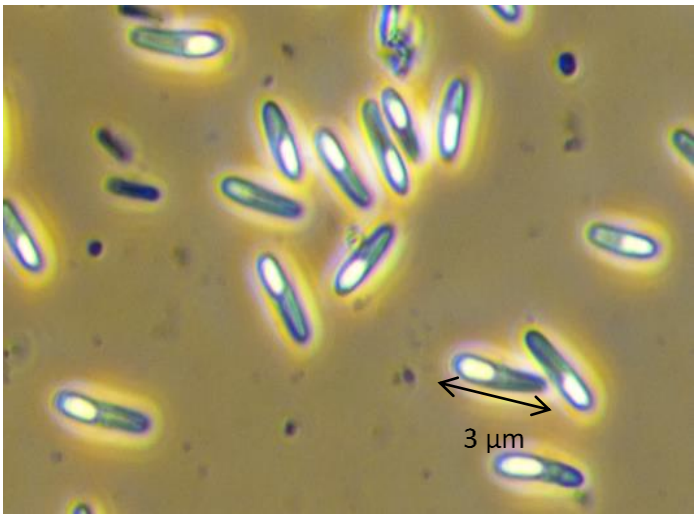


Figure 4.7 Sporulation of Δσ^k mutants

The ATCC 824 parent strain (WT), the Δσ^k (σ^k) mutant and the complemented mutant (σ^k ME6C σ^k) were grown in CBM-S 6% glucose with samples taken to measure sporulation (through heat resistant cfu/ml) at 48 h and 120 h. Stains were analysed for statistical significance from the through T-test analysis, results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

A.



B

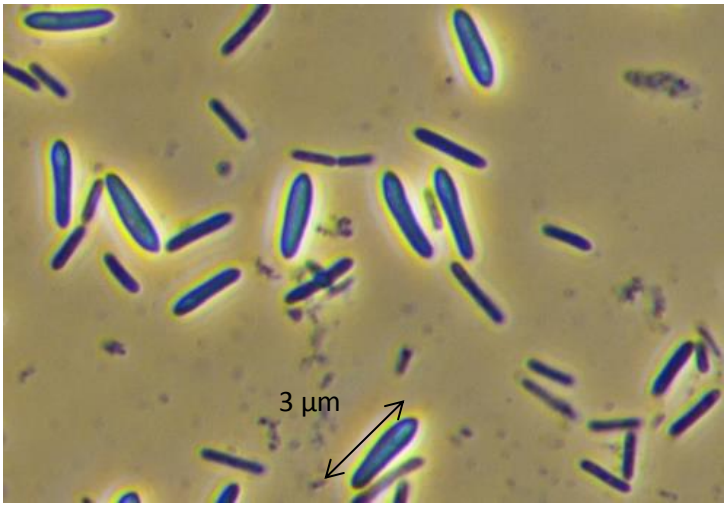


Figure 4. 8 Sporulation negative Δσ^{k} mutants visualised by light microscopy

(A) Formation of phase bright forespores in *C. acetobutylicum* ATCC 824 wild type. (B) Δσ^{k} mutants show no formation of phase bright forespores. Cells were grown in CBM-S broth 6% glucose (w/v) for 48 h, after which 10 µl was visualised by light microscopy at an 800 fold magnification.

From the mutants generated in this study, σ^{k} has only been successfully confirmed for its initial assumed role as a late phase regulator in the sporulation cascade, with only sporulation eliminated in $\Delta\sigma^{k}$ mutants and restored upon complementation and providing no evidence to

suggest any significant interactions with other early phase regulators such as Spo0A or Ca_c0957.

4.2.4. Generation of a Ca_c0323 orphan kinase mutant

Due to the essentiality of Spo0A in granulose accumulation and the implied level of activity it must still have within a ΔCa_c0957 mutant, it was decided to investigate the impact of Spo0A on granulose formation and other relevant phenotypes dependent on its level of phosphorylation. For ease of study it was decided to observe the previously identified orphan histidine kinase Ca_c0323 (Steiner et al. 2011), which had previously displayed complete loss in sporulation when inactivated in a similar manner to the co-inactivation of Ca_c3319 and Ca_c0903. Generation of a strain severely affected in Spo0A phosphorylation and the observation of its granulose and solvent phenotypes would likely give an indication as to how much Spo0A phosphorylation affected these phenotype compared to a ΔCa_c0957 mutant. For this purpose it was decided to attempt in frame deletion of Ca_c0323.

While single cross-over mutants were observed, no double cross over mutants could be successfully generated by in-frame deletion. Personal communication with the lead author of the original publication (E. Steiner) revealed that the ClosTron inactivated mutants as difficult to obtain and were instable (i.e. were not able to be successfully revived from glycerol stocks) necessitating the creation of fresh mutants for each experiment. It was therefore decided to forego characterising the *agrBD*, granulose accumulation and solvent profile of this potential mutant.

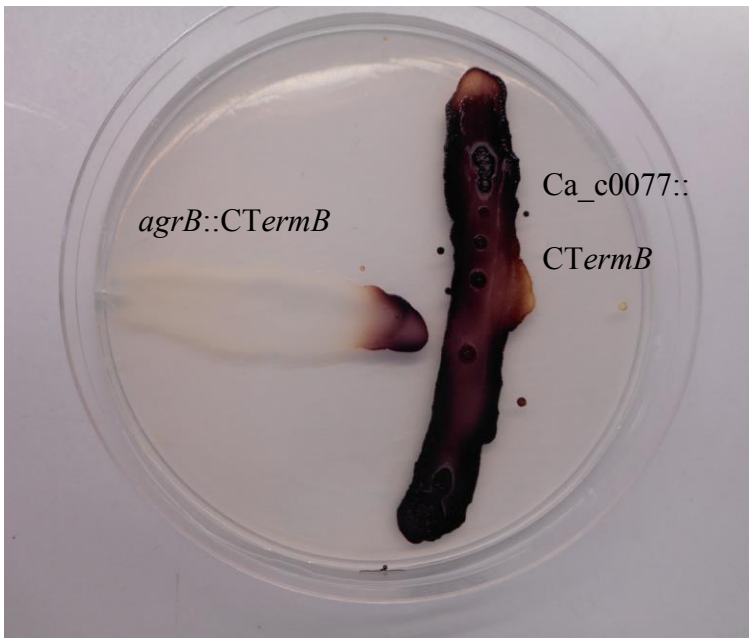
4.2.5. CAAX-like membrane protease Ca_c0077

To help further characterise the role of Ca_c0957/Ca_c0958 in the regulation of granulose accumulation the CAAX-like membrane protease upstream of the *agrBD* operon, shown to be down regulated in a Ca_c0957::CT*ermB* mutant, was inactivated through the use of ClosTron mutagenesis. Due to the suggested peptidase activity of this enzyme and co-localisation with the *agrBD* operon it was hypothesised to play a role in AgrD peptide processing.

The Ca_c0077::CT*ermB* mutants were screened for granulose accumulation and implied changes in Agr signalling through growth on CBM-S 6% (w/v) glucose agar followed by iodine staining to visualise granulose accumulation and sporulation levels assessed through growth in CBM-S 6% (w/v) glucose broth and detection of heat resistant cfu/ml at 48 h and 120 h time points, with both of which phenotypes known to be influenced by Agr quorum sensing (Steiner et al. 2012).

As can be observed from Figure 4.9 the Ca_c0077::CT*ermB* mutants were shown to produce the AgrD derived signal, accumulate granulose and showed a sporulation profile comparable to the wild type control. Inactivation of the CAAX-like membrane showed no significant change in phenotype with respect to Agr functionality and suggested that either the enzyme plays no role in the processing of the AgrD derived peptide or that the loss of activity is compensated through the activity of other protease enzymes. For this reason despite its down regulation in a Ca_c0957::CT*ermB* mutant and its proximity upstream of the *agrBD* operon; it was decided to not perform any further investigations with respect to the role of this protease.

A



B

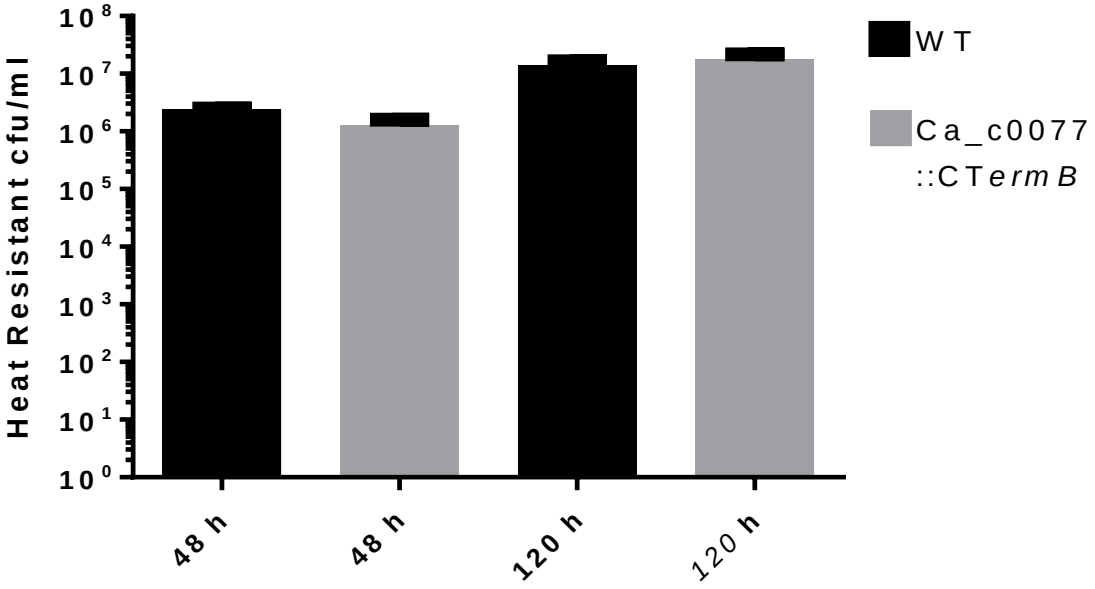


Figure 4.9 Characterisation of Ca_c0077::CTermB mutants

(A) Granulose accumulation of a Ca_c0077::CT*ermB* mutant grown alongside an *agrB*::CT*ermB* strain. Strains were grown in liquid CBM-S containing 6% (w/v) glucose to mid exponential phase growth before being streaked on to CBM-S agar with containing 6% (w/v) glucose and incubated for a further 48 h. Granulose was visualised by iodine staining. (B) Sporulation was assessed by heat resistant cfu after 48 h and 120 h growth in CBM-S broth 6% (w/v) glucose. No significant differences were observed from the wild type level of sporulation. n=9

4.2.6. LanC like peptide cyclase Ca_c0082

4.2.6.1. Granulose accumulation and Agr quorum sensing in a Ca_c0082::CT*ermB* mutant

To further explore the potential regulation of the Agr system by Ca_c0957, inactivation of the LanC-like peptide cyclase Ca_c0082, shown to be down regulated approximately 100 fold upon Ca_c0957 inactivation (Kotte 2013), was performed. Due to this gene being located downstream of the *agrCA* operon and the role of LanC-like enzymes in peptide modification in lantibiotic production and quorum sensing it was hypothesised that the enzyme may play a role in Agr peptide modification.

Inactivation of the Ca_c0082 was performed by ClosTron mutagenesis (see 2.6.1) using pMTL007C-E2::Ca_c0082 1119-1120s and pMTL007C-E2::Ca_c0082 538-539a to create 2 sense and one anti-sense independent mutants strains. This inactivation was observed as producing a significant effect with respect to granulose accumulation, failing to show any detectable granulose accumulation, as can be seen in Figure 4.10. The cause of this however was shown not to be due to the loss of production of the AgrD derived signalling peptide. Cross streaking the strain against a ΔCa_c0957 or *agrB*::CT*ermB* mutant showed the restoration of granulose accumulation in these two strains. Furthermore provision of an AgrD derived peptide signal, supplied from cross-streaking with a wild type *C. acetobutylicum* or direct addition of peptides (see chapter 5) failed to produce a granulose accumulation response or restore sporulation in a Ca_c0082::CT*ermB* mutant.

Importantly when a plasmid-based copy of Ca_c0082 was introduced to the inactivation strain the granulose phenotype was restored confirming that the phenotype was not the consequence of polar effects upon the upstream two component system responsible for peptide recognition *agrCA*. Ca_c0082 clearly plays a key role in granulose accumulation, however the means by

which this is achieved is independent of *agrBD* expression in peptide signalling, though no further evidence ascertaining to its activity were obtained in this study.

A

B

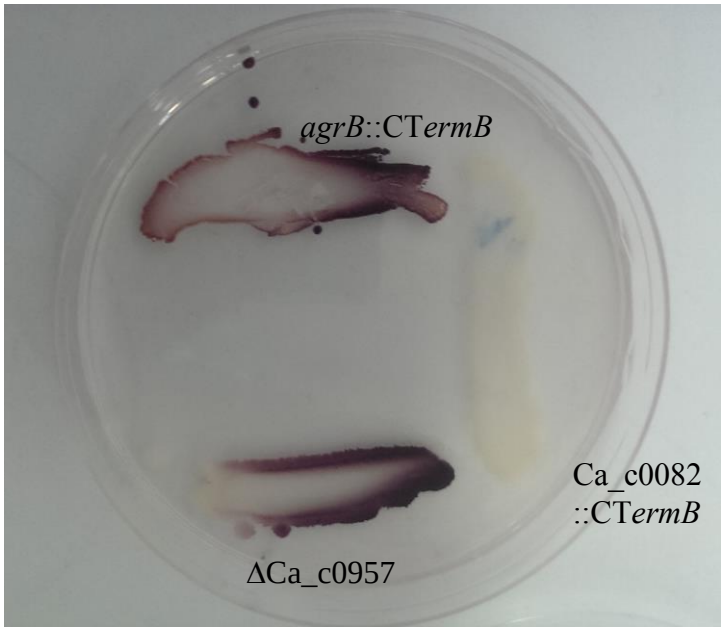


Figure 4.10 Granulose accumulation and Agr signalling of a Ca_c0082::CT*ermB* mutant

(A) Granulose accumulation is lost when Ca_c0082 is inactivated and restored upon complementation and (B) an active Agr peptide is produced by a Ca_c0082 mutant restoring granulose formation in an *agrB*::CT*ermB* mutant. Strains were grown in liquid CBM-S containing 6% (w/v) glucose to mid exponential phase growth before being streaked on to CBM-S agar with containing 6% (w/v) glucose and incubated for a further 48 h. Granulose was visualised by iodine staining.

4.2.6.2. Sporulation of a Ca_c0082::CT*ermB* mutant

Due to the loss of granulose also being observed in Agr mutants (Steiner et al. 2012) and the proposed role of granulose as energy and carbon storage for endospore formation (Reysenbach et al. 1986; Jones & Woods 1986), the level of endospore formation was investigated in Ca_c0082::CT*ermB* mutants (through assessing heat resistant cfu/ml) following 120 h of growth in CBM-S 6% (w/v) glucose.

The inactivation of Ca_c0082 resulted in an approximate 10 fold decrease in sporulation when compared to the wild type as can be seen in Figure 4.11. It can also be observed that the Ca_c0082 mutants' sporulation profile shows no significant difference to that of an *agrC*::CT*ermB*. To negate the possibility of a polar effect on the expression of *agrCA* due to intron insertion the two component response regulator was expressed from a plasmid in the Ca_c0082::CT*ermB* mutant. No change in the sporulation levels were observed as a result of *agrCA* expression and only complementation with the Ca_c0082 gene on a plasmid resulted in a restoration of sporulation to wild type levels.

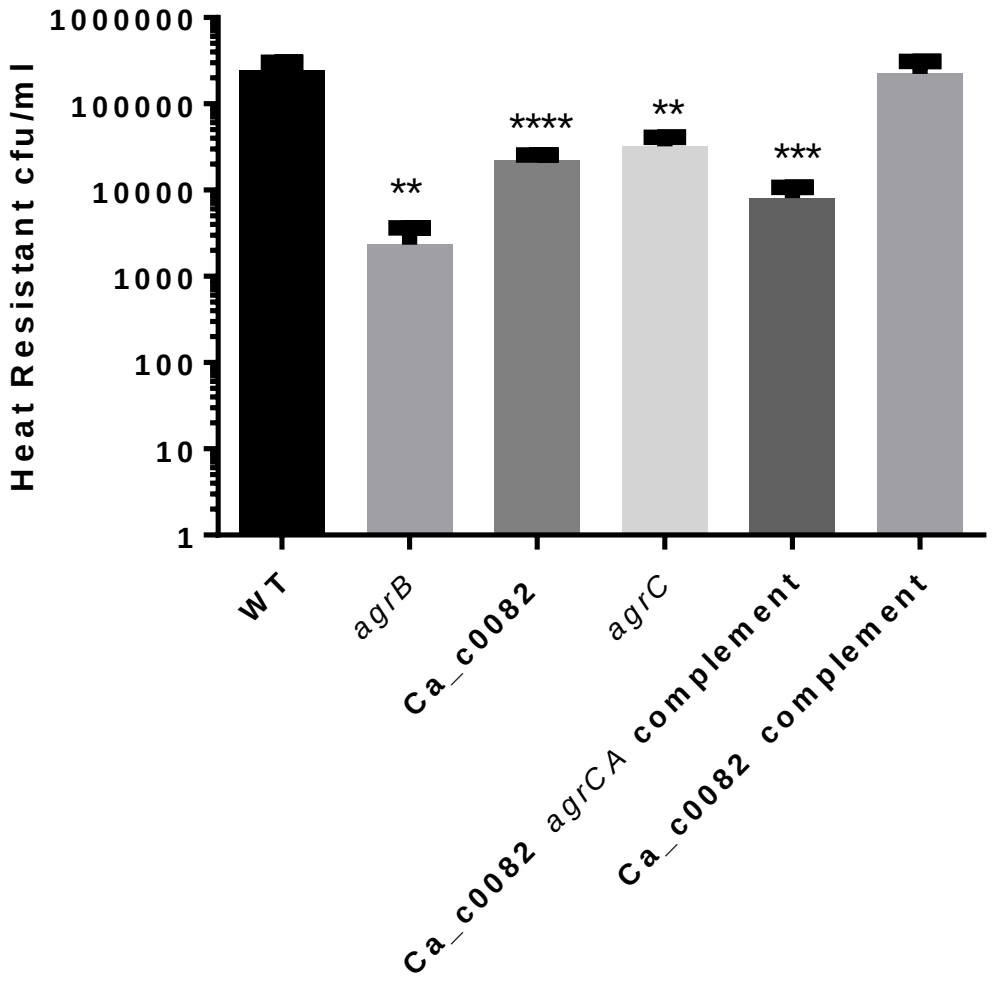


Figure 4.11 Sporulation of Ca_c0082 mutant and *agr* mutants and complemented strains

Sporulation levels of the generated Ca_c0082::CTermB mutants alongside *agr* mutants measured by heat resistant cfu/ml following 120 h growth in CBM-S 6% (w/v) glucose. Ca_c0082 was complemented directly and for potential polar effects by expressing Ca_c0082 and AgrCA respectively on modular pMTL85141 plasmids. Stains were analysed for statistical significance from the wild type through T-test analysis, results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9..

.

4.3. Discussion

Whilst the role of Spo0A as the master regulator of sporulation and solventogenesis has been established in *C. acetobutylicum* (Harris et al. 2002) and inactivation of *spo0A* in *C. beijerinckii* had resulted in a loss in granulose accumulation (Ravagnani et al. 2000), this study has served to confirm the critical role of Spo0A as a master regulator of granulose accumulation in *C. acetobutylicum*. Deletion of *spo0A* has been shown to completely remove *C. acetobutylicum*'s capability to initiate granulose accumulation, to an extent observed in an *agr* or *glgA* (granulose synthase) mutants (Steiner et al. 2012; Ehsaan et al. 2016) yet remained capable of producing the AgrD derived signalling peptide. Though not studied directly in the past, the occurrence of this phenotype is not unexpected as previous studies have identified mutants of σ^G, part of the sporulation regulatory cascade, as being unable to synthesise granulose (Tracy et al. 2011).

Previously generated RNA-seq data from Ca_c0957::CT*ermB* mutants had suggested that the change in stationary phase phenotypes observed in the strains was primarily the consequence of a down regulation of histidine kinases Ca_c3319 and Ca_c0903 (Kotte 2013), which according to Steiner et al. (2011) results in a near total reduction in Spo0A phosphorylation. However this study revealed that the granulose phenotype observed in ΔCa_c0957 and Δ*spo0A* mutants are the result of alterations to two independent but equally essential pathways responsible for granulose synthesis regulation.

Through the generation of double mutants it has been possible to observe that granulose accumulation requires the presence of Spo0A regulation in order to be successfully induced within *C. acetobutylicum*. Conversely while Ca_c0957/Ca_c0958 regulation remains functional, the expression of *agrBD* for the production of the Agr signal peptide is maintained and is equally necessary in order to induce granulose accumulation through quorum sensing. Elimination of either regulatory pathway results in the loss of granulose accumulation within a

strain, but Spo0A regulation has been shown to clearly exert a more absolute regulatory control when compared to the Ca_c0957-Agr based regulation in terms of hierarchical importance due to the ability of ΔCa_c0957 to respond to an Agr signal only when Spo0A functionality is still possible within the cell.

The fact that *spo0A* is still required to facilitate granulose production in the presence of the Agr signal in a ΔCa_c0957 mutant suggests that the initial hypothesis by Kotte (2013) that Spo0A is completely inactive within this mutant as unlikely to be correct. While both the Ca_c3319 and Ca_c0903 histidine kinases are down regulated and both sporulation and solventogenesis phenotypes match that of a Δ*spo0A* mutant, the data presented here suggests that some level of Spo0A activity likely still occurred within the ΔCa_c0957 mutant.

An initial approach made was to examine the role of Spo0A in a potential phosphorylation negative background, through the removal of Ca_c0323 a histidine kinase previously shown to result in the total loss of sporulation when inactivated (Steiner et al. 2011). Observing granulose accumulation, solvent production and *agrBD* expression in this background and both within a wild type strain and a ΔCa_c0957 mutant may have provided evidence for the presence of Spo0A-P in a ΔCa_c0957 mutant or a potential role of unphosphorylated Spo0A. However this work was discontinued as in-frame deletion mutants could not be successfully generated. ClosTron insertion mutants were considered though personal communication with the lead author of the previous study who reported that they were instable, could not be stored at -80 °C and were required to be freshly generated for each experiment. While this may have been possible to repeat, an intron inactivation mutant was deemed further undesirable due to any potential polar effects on the adjacently located RNPP-like regulator Ca_c0324, which has been shown to act as a repressor of solventogenesis and sporulation (Kotte 2013). Due to the difficulties in removing the activity of the histidine kinase entirely, an alternate approach in the

future may be to introduce the Ca_c0323 gene under the control of an inducible promoter either through allelic exchange or prior to attempts to inactivate the gene as a means of studying a low Spo0A phosphorylation background, while potentially avoiding the instability resulting from the genes removal. It was however decided to instead further examine other genes of potential interest with respect to either acting as early phase regulators of stationary phase phenotypes shown to be differentially expressed in Ca_c0957::CT*ermB* mutants

A recently suggested early regulator of clostridial sporulation is σ^k, which had previously been believed to act as a late phase regulator of sporulation according to its role in *Bacillus* spp. (Dürre & Hollergschwandner 2004). Transcriptional analysis of σ^k inactivation mutants in *C. botulinum* has revealed down regulation of both Spo0A and σ^f and has been shown to play a role in stress response (Kirk et al. 2012; Dahlsten et al. 2013), suggesting it acts as an early phase regulator of the sporulation cascade. Following the observations made in *C. botulinum*, Al-Hinai et al. (2014) claimed a similar effect in the down regulation of *spo0A* transcripts within *C. acetobutylicum*. Furthermore the loss in *spo0A* expression was linked in that study to a subsequent loss of granulose accumulation, solvent production and sporulation comparable to what has been observed in both Δ*spo0A* and ΔCa_c0957/ΔCa_c0958 mutants. For this reason it was decided to investigate the role of σ^k further to determine if the change in phenotypes observed by Al-Hinai et al. (2014) could be linked to Ca_c0957 activity and/or played a direct role in regulation of Spo0A activity. As expected of a loss in a sigma factor conserved in the endospore regulatory cascade, sporulation was abolished in $\Delta\sigma^k$ mutants. However granulose accumulation and solventogenesis losses described by Al-Hinai et al. (2014) were not observed in $\Delta\sigma^k$ mutants generated in this study.

The work by Al-Hinai et al. (2014) described the $\Delta\sigma^k$ strain as showing a near complete loss in butanol and acetone production and subsequently leading to an acid crash of normal batch

cultures, while reintegration of σ^k at its native locus and overexpression of Spo0A in the mutant strain from a plasmid resulted in solvent production and subsequently viable growth in batch culture (though no direct comparative cultures of the $\Delta\sigma^k$ mutant and complimented strains had solvent profiles characterised for growth under the same conditions). Despite this the mutants generated in this study showed no significant changes from the wild type or complemented strains with respect to their ability to reassimilated acids and produce solvents.

In this study it was unnecessary to pH regulate $\Delta\sigma^k$ mutants when grown in batch culture, though Al-Hinai et al. (2014) described pH control as a necessity in order to culture beyond 24 h. Though the down regulation of solvent production can lead to the greater accumulation of butyrate and acetate within the media, it was also be observed for Ca_c0957/Ca_c0958 mutants and Δ*spo0A* mutants that no additional buffering was required, outside of the initial calcium carbonate provided, in order to successfully batch culture these strains (Ehsaan et al. 2016). While *spo0A* expression may be down regulated when σ^k is deleted, a clean deletion of *spo0A*, showing no acid crashing phenotype suggests that the role of σ^k as an early phase regulator would be unlikely to cause the occurrence of an acid crash simply due to the down regulation of *spo0A*.

Ca_c0957 and Spo0A work described here revealed the necessity for both systems in order to initiate granulose accumulation. $\Delta\sigma^k$ mutants generated in this study showed a granulose positive phenotype and the ability to produce the Agr signalling peptide, suggesting that both regulatory systems were active. Supporting this idea that both Ca_c0957 and Spo0A global regulation remained active was the observation that solvent production appeared unaffected in the $\Delta\sigma^k$ mutants. While Spo0A expression was claimed to be reduce in σ^k mutants to a large extent by Al-Hinai et al. (2014) the proposed hypothesis that ΔCa_c0957 mutants still maintains some level of Spo0A activity would support the formation of granulose in the $\Delta\sigma^k$

mutants as both would share a low activity of Spo0A and therefore could respond to Agr signalling to produce granulose but unlike the ΔCa_c0957 the $\Delta\sigma^k$ would still maintain *agrBD* expression.

Furthermore the observation by Al-Hinai et al. (2014) that granulose accumulation was absent may be questioned as the TEM images which formed the basis of the claim were taken after 5 days of growth and are limited with respect to the number of images taken of the wild type and $\Delta\sigma^k$ strains. Previous studies have shown that granulose production is usually observed as beginning at late exponential phase of growth and peak between 24 h and 48 h (Reysenbach et al. 1986; Steiner et al. 2012), and while evidence of the wild type showing granulose accumulation was provided it was also noted that free spores were also plentiful in the culture (as observed by microscopy and through sporulation assays) implying that many wild type cells had already consumed their accumulated granulose. It may therefore be the case that a large proportion of $\Delta\sigma^k$ mutants observed by TEM had consumed their accumulated granulose by the 120 h time point and therefore granulose accumulation was not detected. Alternately granulose accumulation may have been reduced in the $\Delta\sigma^k$ mutants, instead of entirely eliminated which may further increase the difficulty of detecting its presence by the 120 h time point as opposed to the earlier time points presented in this study.

While the previous work of Al-Hinai et al. (2014) showed a down regulation of *spo0A* that was hypothesised to cause a loss in the described stationary phase phenotypes it should be noted that the studies transcriptional analysis of *spo0A*, limited to the first 26 hours, does not indicate a total loss of *spo0A* expression. Detection of Spo0A by Western blot analysis in the same study extended as far as 48 hour though is not directly quantifiable. Though a reduction in active Spo0A will likely lead to a loss of activity and may impact granulose accumulation and solventogenesis as was proposed for the Ca_c0957 and Ca_c0958 mutants, past studies of

histidine kinases have shown that partial reductions in active Spo0A may only result in a partial reduction in sporulation (Steiner et al. 2011). While this study cannot disprove the reported loss of *spo0A* expression and the result matches that seen by Kirk et al. (2012) there is no necessary justification that the change in active Spo0A levels would directly result in the inhibition of stationary phase processes outside of sporulation (for which σ^k is part of the regulatory cascade). Furthermore it can be noted that in *B. subtilis* the activation of genes by Spo0A vary with respect as to how high Spo0A activity needs to be in order to induce their expression, with some genes active at lower levels of Spo0A activity (Fujita et al. 2005). With no available evidence as to the level of Spo0A activity required to enable the activation of all the described stationary phase phenotypes and the lack of observational change of any stationary phase phenotype outside of the sporulation cascade in *C. botulinum* (Kirk et al. 2012), this study would suggest that the observed transcriptional changes described by Al-Hinai et al. (2014) may not be sufficiently severe to justify the loss of solventogenesis and granulose accumulation.

Further RNA-seq data obtained for the Ca_c0957::CT*ermB* mutants (Kotte 2013) revealed that the genes flanking the *agrBD* and *agrCA,* loci were severely down regulated. These genes were therefore investigated for what role they may play with respect to how Ca_c0957 regulates granulose accumulation and sporulation through Agr signalling. Aside from their co-localisation with the Agr loci the genes were identified as being of interest due to the conserved domains of these genes identifying them as a CAAX –like membrane protease and a LanC like peptide cyclase, both potentially involved in processing of peptide structures of AIPs.

Inactivation of Ca_c0077 showed no significant change in the Agr linked phenotypes of granulose accumulation or sporulation and were shown to effectively express functional AgrD derived peptide. Though it is co-localised near the *agrBD* operon, its function in peptide

maturation of the signal peptide appeared to be unimportant. However, a reason for this may be the existence of homologues of Ca_c0077 within *C. acetobutylicum*, as an initial BLAST analysis has identified Ca_c0420, Ca_c0482, Ca_c2478 and Ca_c3028 as other potential CAAX-like membrane proteases. It therefore may be the case that the inactivation of one system is simply compensated for by one or more redundant systems. While it's down regulation in ΔCa_c0957 (and presumably ΔCa_c0958) mutants may still indicate a role Agr signalling, the unequivocal proof would have required the inactivation of one or more of the other systems and it was therefore decided not to explore the activity of Ca_c0077 any further.

Lantibiotics are a group of bacteriocins that contain unusual amino acids such as lathionine which are generated by post-translational modification. These often polycyclic peptides have been well studied for their use as antimicrobials (Chatterjee et al. 2005) but also as a cell-to-cell signalling mechanism (Ramsey et al. 2009). *C. acetobutylicum* appears to lack most of the genes required to form lantibiotic-like compounds but does contain a *lanC* peptide cyclase homologue Ca_c0082 (or potentially based upon structural similarity a *lanM* homologue, which is also has peptide cyclase like activity). As the organism appears to lack the ability to synthesise lantibiotics, the localisation of Ca_c0082 and the *agrCA* operon and its down regulation of the gene observed in a Ca_c0957::CT*ermB* mutant, Ca_c0082 was investigated for what role it may play in Agr signalling under the hypothesis that it may be involved in peptide modification.

Inactivation of the gene and subsequent complementation analysis showed a distinct granulose negative phenotype alongside a decrease in the organisms sporulation profile similar to that observed in an *agr* mutant (Steiner et al. 2012). Similarity of the gene to cyclases would at first suggest that it plays a role in AgrD signal generation, however whilst granulose formation was abolished and sporulation reduced the mutant was still capable of producing the AgrD signal

as shown by cross-streaking with an *agrB*::CT*ermB* mutant. Furthermore a wild type generated Agr signal were unable to complement in cross-streaks suggesting no negative effect upon Agr signal production following Ca_c0082 inactivation.

An alternate hypothesis would be that Ca_c0082 is required for signal recognition in conjunction with the AgrCA two component system, which would justify why the Ca_c0082::CT*ermB* does not respond to the provision of an Agr signal. However the fact that the AgrD derived signal is produced shows that the expression of *agrBD* still occurs within the Ca_c0082::CT*ermB* mutant, in contrast to *agrA*::CT*ermB* and *agrC*::CT*ermB* mutants (Steiner et al. 2012), suggesting that the positive feedback loop (leading to increased *agrBD* expression and thus sufficient peptide production) is still occurring. Instead the similar sporulation phenotypes may be the result of a loss of sporulation as a consequence of the absence of granulose accumulation (regulated by mechanisms independent to Agr signalling). The extent to which the loss of granulose is responsible for the sporulation phenotypes in granulose deficient mutants was investigated further in chapter 5.

Ca_c0082 inactivation appears instead to show more similarity to the regulation of granulose accumulation by Spo0A activity and the sporulation cascade, with both Ca_c0082 and *spo0A* mutants having shown no ability to form granulose while still maintaining production of the AgrD derived peptide, which may suggest that both Spo0A and Ca_c0082 operate as part of the same regulatory pathway. If this was the case the down regulation of Ca_c0082, observed when Ca_c0957 was inactivated, may actually be a consequence of the loss of Spo0A activity.

An alternate hypothesis based on its genetic localisation may be that Ca_c0082 is regulated by Agr quorum sensing as the means by which it regulates granulose accumulation, justifying why Agr quorum sensing appears functional in a Ca_c0082::CT*ermB* mutant but maintains a granulose negative phenotype. It is also possible that Ca_c0082 still plays a role with respect

to peptide modification and cell-to-cell signalling, although based on the results obtained in this study this peptide is unlikely to be derived from the Agr system.

In conclusion the regulation of stationary phase phenotypes by Ca_c0957 and Spo0A has been further characterised based on phenotypic differences observed with respect to the means by which granulose was shown to be regulated. Ca_c0957 was shown to regulate granulose accumulation, primarily through controlling the expression of the *agrBD* operon and the subsequent production of signal peptide and its recognition by AgrCA. Ca_c0077 up regulation by Ca_c0957 may play a role in this signal production however without further investigation of other potentially redundant CAAX-like peptidases its contribution to AgrD processing remains unclear. Ca_c0957 is also to influence the activation of the granulose accumulation via Spo0A, as evidenced by its low sporulation and solvent profiles. However, data obtained in this study suggest that Spo0A remains partially active in Ca_c0957 mutants due to granulose restoration being possible within the strain when in the presence of an AgrD signalling peptide. Control of granulose accumulation regulated by Spo0A may also involve the *lanC* like peptide cyclase Ca_c0082, as its mutation creates a phenotype matching a *spo0A* mutant with respect to granulose activity and Agr signalling. Alternatively it can be hypothesised that the Ca_c0082 may instead be regulated by the Agr system. Despite its down regulation in Ca_c0957::CT*ermB* mutant Ca_c0082 has been shown to be a necessity in granulose accumulation and if under the control of the Agr system provision of the AgrD signalling peptide would allow for the increased expression of Ca_c0082. In the future it may be of interest to express/inactivate Ca_c0082 in Δ*spo0A* and ΔCa_c0957 backgrounds to observe what effect this would have upon granulose accumulation and potentially identify through which regulatory mechanism Ca_c0082 acts.

Attempts to reconcile the shared sporulation and solvent phenotypes of Δ*spo0A* and ΔCa_c0957/ΔCa_c0958 mutants have yet to incorporate any other potential early phase regulators and signals. Examining a Spo0A phosphorylation negative Ca_c0323 mutant was attempted but not achieved, presumably due to mutant instability, though could be examined further in the future with the use of expression under a controllable regulatory element. Examining the proposed early phase regulator of *spo0A* expression, σ^k, to determine any potential interaction with the Ca_c0957/Ca_c0958 system revealed that mutants did not displaying the previously published phenotypes. Mutants generated in this study only showed a loss of their sporulation phenotype and appeared unaffected with respect to solvent production, *agrBD* expression and granulose accumulation. It was therefore concluded that even if σ^k regulates *spo0A* at a transcriptional, as previously postulated, level, its contribution would be negligible with respect to any phenotype outside of sporulation under the employed conditions.

5. Structure-function analysis of Agr signalling and characterising the significance of granulose accumulation in sporulation and solvent production

5.1. Introduction

Agr quorum sensing was first described and extensively characterised in *Staphylococcus aureus* (Ji et al. 1995; Kong et al. 2006), and has been studied in a number organisms across the genus Clostridium, including *C. botulinum, C. difficile, C. perfringens* and *C. acetobutylicum* (Cooksley et al. 2010; Martin et al. 2013; J. Li et al. 2011; Steiner et al. 2012) The system is commonly associated with controlling sporulation within these organisms alongside other stationary phase processes; in *C. acetobutylicum* this takes the form of granulose accumulation. These previous studies have subsequently raised multiple questions with respect to Agr quorum sensing and the structure of the associated peptide molecule. These include: defining the precise structure of *C. acetobutylicum* signal peptides, potential cross communication between different *Clostridium* species and the mechanism by which the Agr system regulates granulose accumulation and its influence upon the process of sporulation.

Previous work attempted to identify the most active, and therefore most probable, structure the *C. acetobutylicum* AgrD derived peptide would take following processing by AgrB and any other potential peptidase activity (Steiner et al. 2012). The relative recognition of these peptides by the AgrC, facilitating AgrA phosphorylation, was measured by their capacity to restore granulose accumulation and sporulation. Of the known AgrD derived peptides, the majority form five membered rings with a short N-terminal tail due to the formation of a thiolactone or lactone bond to the C-terminal carboxyl-group of the precursor peptide. Through the use of synthetic peptides Steiner et al. (2012) proposed that the most effective structure of a *C. acetobutylicum* Agr signal peptide would be a six membered ring lacking an N-terminal peptide tail. However, while the proposed peptide was shown to be active, due to the lack of an N-terminal tail it was likely that the ring structure was able to undergo a rearrangement reaction

under physiological conditions (as demonstrated in Figure 5.1). With no tail present it is possible for the thiolactone bond to be replaced with an amide bond between the N-terminal cysteine and the C-terminal carboxyl group, creating a stable and irreversible new ring structure. It was therefore plausible that the mature *C. acetobutylicum* AgrD peptide may have actually originated from this proposed derivative, instead of the originally proposed six membered thiolactone ring structure.

Another new and interesting suggestion with respect to clostridial Agr quorum sensing is that the AgrD derived peptides need not necessarily be circular in order to show activity. For *C. perfringens* it was demonstrated that several linear versions of the AgrD peptide have also possess biological activity, albeit slightly weaker than the circular structure, with respect to restoring virulence (Ma et al. 2015). This suggests the need for further exploration of the differences between clostridial and staphylococcal AgrD signalling and processing.

In *S. aureus* the phenomenon of cross inhibition or activation of quorum sensing systems can occur between strains/species that use similar AgrD derived peptides (Otto et al. 2001). With the recently increased availability in completed clostridial genomes, identification of similar Agr systems within organisms closely related to *C. acetobutylicum* such as *Clostridium roseum* has been made possible. Thus, it has become possible to study genetically related and physiologically similar species that occupy the same natural habitats with respect to quorum sensing based interactions

As shown by Steiner et al. (2012) the Agr signalling of *C. acetobutylicum* contributes to the regulation of sporulation and granulose accumulation. A possibility is that endospore formation is indirectly regulated by the Agr system as a result of the regulation of granulose. The accumulation of this starch-like storage compound has been a known characteristic of *C. acetobutylicum* and other species of the genus for a long time (Robson et al. 1974; Reysenbach

et al. 1986; Jones & Woods 1986); however the role of granulose accumulation has served mostly as a point of speculation. Proposed as means of energy storage, in preparation for the formation of spores (Reysenbach et al. 1986), the most prominent evidence is the decrease in sporulation observed when the Agr system is inactivated (Steiner et al. 2012; Jabbari et al. 2013) or when granulose is no longer observed in Ca_c0082 mutants (See Chapter 4.2.6). It can also be noted that in *C. acetobutylicum* mutants the decrease in sporulation was observed as being far more drastic in plate-grown colonies as opposed to the moderate change observed in liquid culture (Steiner et al. 2012). However, in strains such as *C. botulinum* or *C. perfringens* inactivation of the Agr system shows a change in sporulation profile (Cooksley et al. 2010; J. Li et al. 2011), even if the strains do not accumulate granulose, suggesting a potential direct regulation outside of granulose accumulation. Furthermore the occurrence of granulose accumulation in some solventogenic clostridia has only been described as occurring during conditions of limited carbon and energy (Reysenbach et al. 1986). While supporting its role as an energy storage compound, the work of Steiner et al. (2012) by using an excess of 5% (w/v) glucose for the previous characterisation of the *C. acetobutylicum* Agr system may not have allowed for an expansive enough assessment of the importance of granulose accumulation in endospore formation.

This raises the possibility that the effect of the Agr system, on sporulation and potentially solvent formation, through the loss of granulose accumulation may have in the past been understated, as a result of energy and carbon rich growth conditions. If this were to be the case this may suggest that ΔCa_c0957 and ΔCa_c0958 regulation in *C. acetobutylicum* occurs as a result of compounding the phenotypic effects of decreased Spo0A activation and repressed *agrBD* expression.

5.1.1. Aims of this study

The aims of this study were to investigate the questions described above, in order to expand the knowledge of the role and function of Agr quorum sensing in *C. acetobutylicum*, therefore this study aimed to answer the following:

- What is the biologically active structure of the processed AgrD peptide?

  Is the proposed thiolactone ring the most biologically active form of the peptide or is functionality linked to a ring structure formed by an amide bond or a linear peptide? This was investigated through the use of synthetic peptides

- Can evidence be found for cross-species communication?

  This was addressed through the use of *C. roseum* a close relative of *C. acetobutylicum* and synthetic peptides derived from its predicted AgrD sequence.

- Does Agr regulation of granulose influence the formation of endospores and potentially solventogenesis?

  This was investigated using a number of Agr and granulose deficient mutants subject to a range of glucose (w/v) concentrations ranging from limitation to excess.

5.2. Results

5.2.1. The effect of AgrD peptide isomerisation on its role in biological activity

To further examine the functionality of the Agr system and for subsequent experimental design involving the use of artificial peptides, the active structure of the AgrD derived peptide was re-examined. The peptides that were synthesised (as shown in Figure 5.1) represented the initial predicted arrangement of the six member ring (R6T0-Ca) and the hypothesised rearrangement (Amid-R6T0-Ca) which would be structurally stable. In addition an acetylated variant of the original ring structure was synthesised, which would be unable to undergo the molecular rearrangement (Ac-R6T0-Ca). To examine the effectiveness of each peptide *agrB*::CT*ermB* mutants were grown in the presence of each variant peptide in order to test for the restoration of granulose accumulation and sporulation. An *agrB*::CT*ermB* mutant was chosen as it had been previously shown to be incapable of producing the Agr signal peptide and as a consequence showed no granulose accumulation and an approximate 10-fold decrease in sporulation.

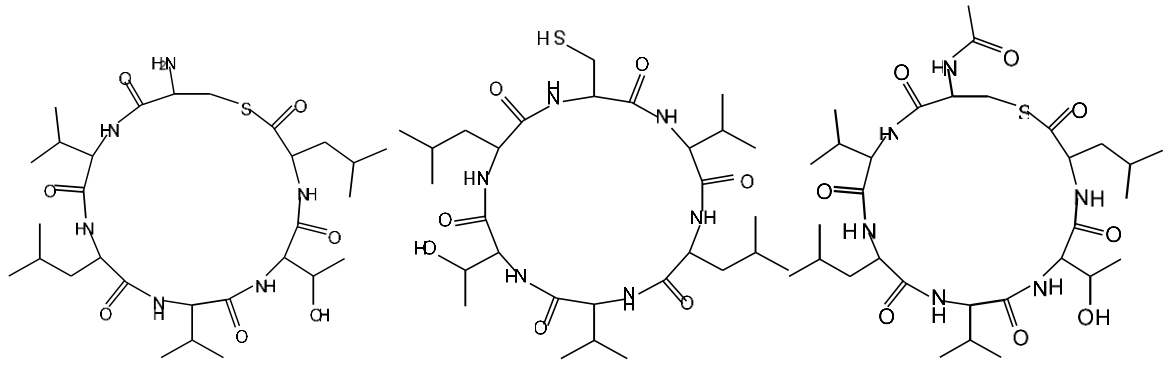


Figure 5.1 Structure of the AgrD derived signalling peptide and its use in this study

The original *C. acetobutylicum* six membered ring structure propose by Steiner et al. (2012), R6T06-Ca may undergo a rearrangement reaction, as represented by the new structure Amid-R6T0-Ca, but will maintain its original structure through acetylation as seen in Ac-R6T0-Ca.

As can be seen from Figure 5.2, all three peptides were capable of restoring granulose accumulation in an *agrB*::CT*ermB* mutant confirming the biological activity of both the original proposed six membered thiolactone ring and the amide bond rearrangement. The granulose assay lacked the sensitivity to demonstrate any difference in activity between the three peptide structures; however the fact granulose accumulation was restored, within an *agrB*::CT*ermB* strain, showed the peptides to have significantly more biological activity than ring structures with a peptide tail tested by Steiner et al. (2012) which have failed to elicit a granulose response.

Due to the nature of peptide binding specificity in Agr quorum sensing, which often is sensitive to the degree that a single amino acid substitution in the ring can result in loss of activity by competitive inhibition (Otto et al. 1999), it was predicted that one of the two isomeric ring structures would likely show a greater activity than the other. To examine this further a sporulation assay was performed, which was considered more sensitive than the granulose stain, as in the previous work this had successfully shown the relative activity of six membered peptide rings containing a tail, all of which showed no activity with respect to granulose accumulation (Steiner et al. 2012). To assess this the *argB*::CT*ermB* mutant was grown in CBM-S with 6% (w/v) glucose in the presence of the three peptides (Figure 5.1) and their sporulation profiles analysed after 120h of growth.

As can be seen from Figure 5.3, as was seen with granulose accumulation all three peptides variants were active and provided some level of restoration of the sporulation phenotype across the range of concentrations tested. However, with the increased sensitivity of the sporulation assay it became evident that the acetylated structure was the most effective in restoring sporulation across the entire range of concentrations tested, showing a 58 fold increase in sporulation at 20 µM concentrations.

Figure 5.2 Ability of AgrD-derived synthetic peptides to induce granulose accumulation in a *C. acetobutylicum agrB*:::CT*ermB* mutant

(A) The original R6T0-Ca synthetic peptide shown in Fig 5.1A, (B) the hypothesised new structure Amid-R6T0-Ca shown in Fig 5.1B and the acetylated variant Ac-R6T0-Ca shown in Fig 5.1C all show the capacity to restore granulose accumulation in an *agrB*::CT*ermB* mutant. The *agrB*::CT*ermB* mutant was grown on CMB-S medium with 6% glucose for 48 h alongside filter disks inoculated with 5 µl of a DMSO control (bottom) and 20 µM of synthetic peptide (top). Granulose in the cells was visualised by iodine staining

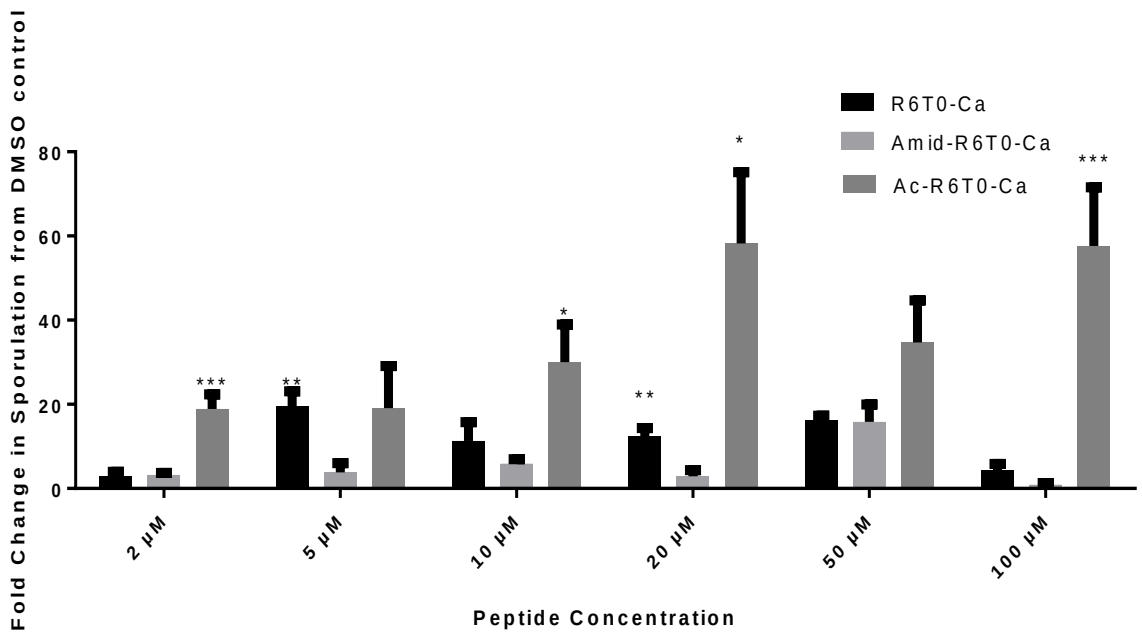


Figure 5.3 the ability of AgrD derived peptides to induce sporulation in *agrB*::CT*ermB* mutants

The fold change in sporulation (observed as heat resistant cfu/ml) for a *C. acetobutylicum agrB*::CT*ermB* mutant after 120 h of growth in CBM-S 6% (w/v) glucose in the presence of the synthetic peptides R6T0-Ca, Amid-R6T0-Ca and Ac-R6T0-Ca (see Figure 5.1 for structures). The *agrB*::CT*ermB* mutant was inoculated with peptide upon reaching an OD_{600} of 0.6 and all strains were compared to a culture inoculated with a DMSO control. The differences between R6T0-Ca and Ac-R6T0-Ca data compare to Amid-R6T0-Ca were analysed for statistical significance through T-test analysis, results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

The hypothesis of instability in the tail-less six membered ring leading to a structural rearrangement is highly supported by the increases in sporulation observed with the use of the R6T0-Ca and Ac-R6T0-Ca peptides compared to that of the Amid-R6T0-Ca peptide. With the exception of the of 5 µM and 50 µM concentrations respectively, the Amid-R6T0-Ca peptide is shown as causing a significantly lower level of heat resistant cfu's/ml and fold change in

sporulation from a DMSO control compared to the Ac-R6T0-Ca peptide. The levels induced by the original R6T0-Ca structure were either similar (2 μM, 10 μM, 50 μM, 100 μM) or higher (5 μM and 20 μM) than those of the Amid-R6T0-Ca but never exceeded those of the Ac-R6T0-Ca suggesting the following general order in terms of biological activity: Ac-R6T0-Ca>R6T0-Ca>Amid-R6T0-Ca

Though both forms of the peptide were shown to be active, the clearly superior sporulation response observed from of the Ac-R6T0-Ca lead to the decision that all future peptide work using similar six membered ring constructs would be based on acetylated variants in order to achieve the maximum response through the maintenance of the most biologically active isomer.

5.2.1.1. Confirming the loss of Agr response in a Ca_c0082 mutant through the addition of synthetic peptide

With the establishment of the most effective peptide structure, the observations regarding the Ca_c0082::CT*ermB* mutants were confirmed through the use of synthetic peptides. As established in chapter 4 the Ca_c0082::CT*ermB* mutant showed an inability to accumulate granulose and a lower sporulation profile than the wild type but nevertheless produced a signal to which an *agrB*::CT*ermB* mutant was able to respond. In order to test the mutant was not producing an incomplete form of the AgrD derived peptide, which receiving strains could modify through functional Ca_c0082, the synthetic acetylated peptide previously described was tested to observe whether provision of a complete peptide ring could induce granulose accumulation or sporulation.

As can be seen from Figure 5.4 the provision of the synthetic peptide ring was unable to induce a fold change in sporulation and was unable to induce granulose accumulation (data not included), as can be observed in the *agrB*::CT*ermB* mutant when supplemented with the same peptide (Figure 5.2). The lack of response to the Ac-R6T0-Ca peptide supports the hypothesis

suggested in chapter 4, that the expression of *agrBD* and the subsequent production of signal peptide is unaffected by the inactivation of Ca_c0082. Instead the findings lends support to the hypothesis that Ca_c0082 may instead regulated granulose accumulation through the same or similar pathway as Spo0A or be regulated by the *agr* system. Though the decrease in sporulation was shown to be similar in scale to an *agrC* or *agrA* mutant, the question was raised as to whether this loss of sporulation was directly governed by the Agr system or whether it was a side effect of the loss of the energy storage compound.

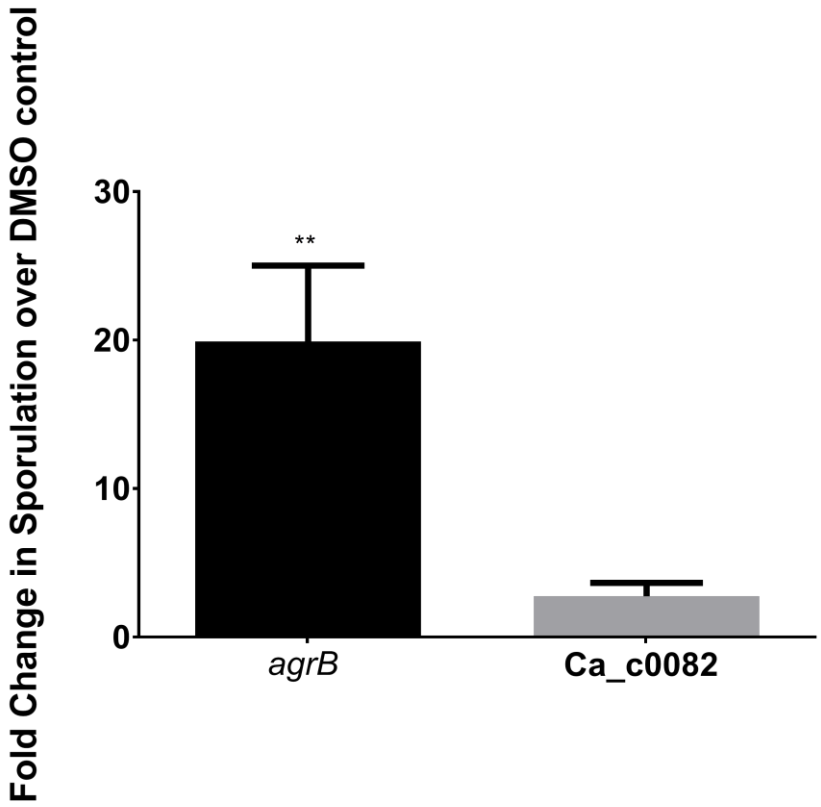


Figure 5.4 Induction of sporulation by peptide addition to Ca_c0082 mutants

Fold change in heat resistant cfu/ml observed for *agrB*::CT*ermB* and Ca_c0082::CT*ermB* mutants grown in 10 ml of CBM-S 6% glucose for 120 h with the addition of 20 µM of Ac-R6T0-Ca peptide compared to a DMSO control of each strain. Results were analysed for statistical through T-test analysis, results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

5.2.2. The use of linear peptides to stimulate an Agr response in a *agrB*::CT*ermB* mutant

The observation of linear peptide activity in *C. perfringens* with respect to restoring Agr function (Ma et al. 2015) was also cross examined for *C. acetobutylicum* strains. Due to the limited activity of tailed peptides structures with respect to granulose production (Steiner et al. 2012) and the presence of alternate ring structures, the *agrB*::CT*ermB* mutants were examined at the more sensitive sporulation level to determine if linear peptide structures, representing the peptide sequence of the six membered ring and potential tail amino acids, created an AgrCA mediated response. These included synthetic linear peptides R6T0-Ca-L, R6T1-Ca-L and R6T4-Ca-L containing a linearised six membered ring and 0, 1 and 4 amino acids of the proposed tail sequence, respectively. These peptides were added to exponential phase *C. acetobutylicum agrB*::CT*ermB* mutants growing in CBM-S 6% (w/v) glucose and heat resistant cfu/ml were determined after 120 h.

As can be seen from Figure 5.5 the only linear peptide to show any statically significant change in sporulation from the *agrB*::CT*ermB* control was the R6T4-Ca-L peptide when present at a concentration of 2 µM, showing a less than 2 fold increase in sporulation. However, when increased to 20 µM no significant change in the sporulation profile was seen for any of the linear peptide at 120 h end point assays.

As was shown in Figure 5.3 and Figure 5.7B the addition of circular peptide Ac-R6T0-Ca had increased sporulation by 15 and 58 fold in an *agrB*::CT*ermB* mutant at 2 µM and 20 µM concentrations respectively. For *C. perfringens* Ma et al. (2015) reported linear peptide activity as restoring Agr functionality to near levels obtained from the addition of the cyclic form; however, in *C. acetobutylicum* where 20 µM of a circular peptide had been shown to be an effective complementation concentration (Figure 5.3) no significant increases or decreases

were observed from the DMSO control. Though *C. acetobutylicum* quorum sensing has been shown to accept rings of slightly altered structure and showed minor activity when interacting with a ring with an amino acid substitution (see 5.2.1 and 5.2.3), this study has shown that its AgrCA two component system is unable to recognise linear peptides in contrast to what has been observed in *C. perfrigens*.

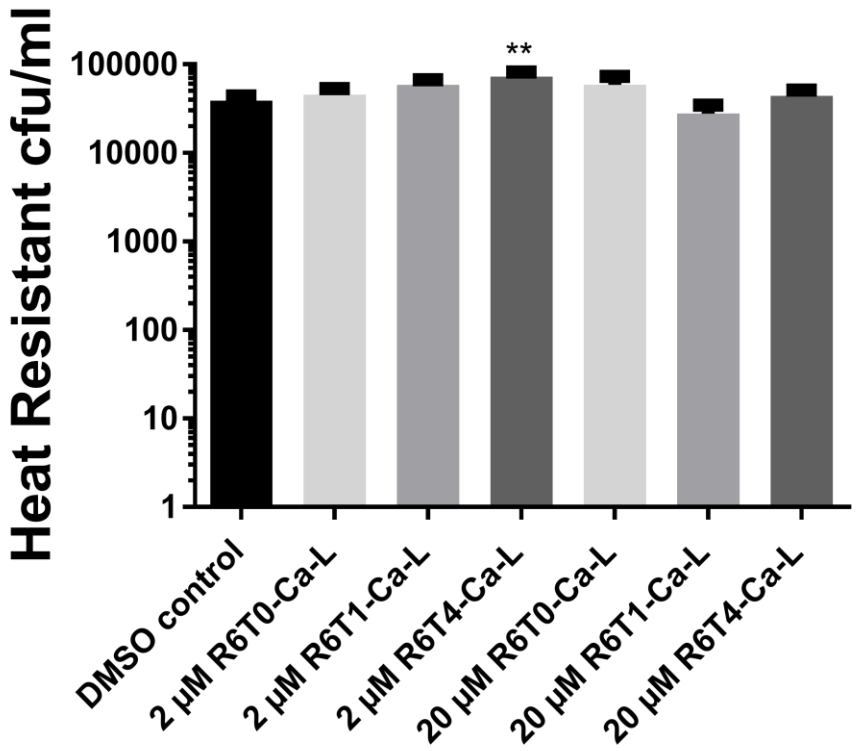


Figure 5.5 Effect of linear AgrD peptides on sporulation in *C. acetobutylicum agrB*::CT*ermB*

Sporulation (heat resistant cfu/ml) levels were assessed after 5 days of growth in CBM-S 6% glucose in the presence of linear peptides R6T0-Ca-L, R6T1-Ca-L and R6T4-Ca-L at a concentration of either 2 µM or 20 µM. Results were analysed for statistical significance from the DMSO control through T-test analysis, results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001, where n=9.

5.2.3. Agr cross signalling between *C. acetobutylicum* and *C. roseum*

Following the establishment of effective ring structures the potential for cross inhibition/activation of the *C. acetobutylicum* Agr system with respect to the closely related *C. roseum* was examined. The recently sequenced genome of *C. roseum* (Poehlein, University of Gottingen, unpublished) revealed an Agr system predicted to produce a peptide with a six membered ring almost identical to that of *C. acetobutylicum* but with a single amino acid substitution. As has been seen for *S. aureus,* single amino acid substitution can lead to an inhibition of an Agr response and a means by which related species often compete (Otto et al. 2001).

C. acetobutylicum AgrD
MNLKEQLNKVNDKFIKGLGKASMKIGEQANGKCVLVTLYEPKMPEELLKENIDK
MNLKKQVNKISEKLIEGIGNVSIKVGEQATDICVLTTLYEPKMPEELLKGNIDK
****:*:**:.:*:*:*:*:.*::*:****..***.*************.****
C. roseum putative AgrD

Figure 5.6 AgrD peptide sequence for *C. acetobutylicum* and *C. roseum*

Alignment of *C. acetobutylicum agrD* and *C. roseum* putative *agrD* peptide sequences. Highlighted amino acids show the conserved thiolactone ring structure. Conserved amino acids in the *C. roseum* AgrD are depicted with a (*), semi conserved substitutions with a (:) and substitutions with a (.).

5.2.3.1. Growth of *C. roseum*

Initial attempts to cross streak *C. roseum* were complicated by its aggregating slimy phenotype which did not facilitate consistent streaks. This aggregating phenotype also created inconsistencies in initial attempts to assay the organisms' sporulation profile generating variable cfu/ml values between technical replicates. Furthermore attempts to analyse a granulose phenotype through iodine staining were hampered by the organism producing a red/pink pigment (for which it is named), that under oxygen exposure converted to a dark brown/black colour similar to that observed in granulose staining. Due to the difficulty in culturing the organism it was decided to not analyse any cross inhibition of *C. acetobutylicum* Agr signalling upon *C. roseum*, instead it was decided to study any potential effect of *C. roseum* signalling upon that of *C. acetobutylicum* through the use of synthetic peptides.

5.2.3.2. The effect of a synthetic *C. roseum* signal peptides on granulose accumulation and sporulation of *C. acetobutylicum*

The *C. roseum* peptides synthesised from the predicted AgrD sequence (Figure 5.6) were designed as a six membered ring with no tail and acetylated (Ac-R6T0-Cr) to avoid intermolecular rearrangement, as this had been found to be the most effective synthetic peptide structure in *C. acetobutylicum* (see 5.2.1), and as ring structures with 1 and 4-amino acid tails (R6T1-Cr and R6T4-Cr respectively).

Addition of the *C. roseum* derived peptides appeared to have no effect upon either the accumulation of granulose of a *C. acetobutylicum* wild type or *agrB*::CT*ermB* strain (data not shown). It was therefore decided to test any potential effect upon the strains with respect to sporulation.

The *C. roseum* peptides were added to *C. acetobutylicum* at various concentrations alongside Ac-R6T0-Ca as a control. Interestingly the tailless Ac-R6T0-Cr and the Ac-R6T0-Ca peptides both appeared to inhibit sporulation of the wild type but not of the *agrB*::CT*ermB* mutant. As can be seen from Figure 5.7A spore levels of the wild type were generally reduced, although only statistically significantly for two of the conditions (2 μM and 20 μM for the Ac-R6T0-Ca and Ac-R6T0-Cr respectively), suggesting that signal peptides in excess of that produced by the wild type had a detrimental effect upon sporulation. By contrast when added to the *agrB*::CT*ermB* mutant, the tail-less *C. roseum* derived peptide was able to cross activate *C. acetobutylicum* AgrCA albeit only at higher concentrations (20 μM). Addition of R6T1-Cr and R6T5-Cr peptides resulted in no significant changes in sporulation of the wild type (Appendix Figure 4) and *agr* mutants (data not shown) at any of the concentrations tested and were therefore not studied beyond this point.

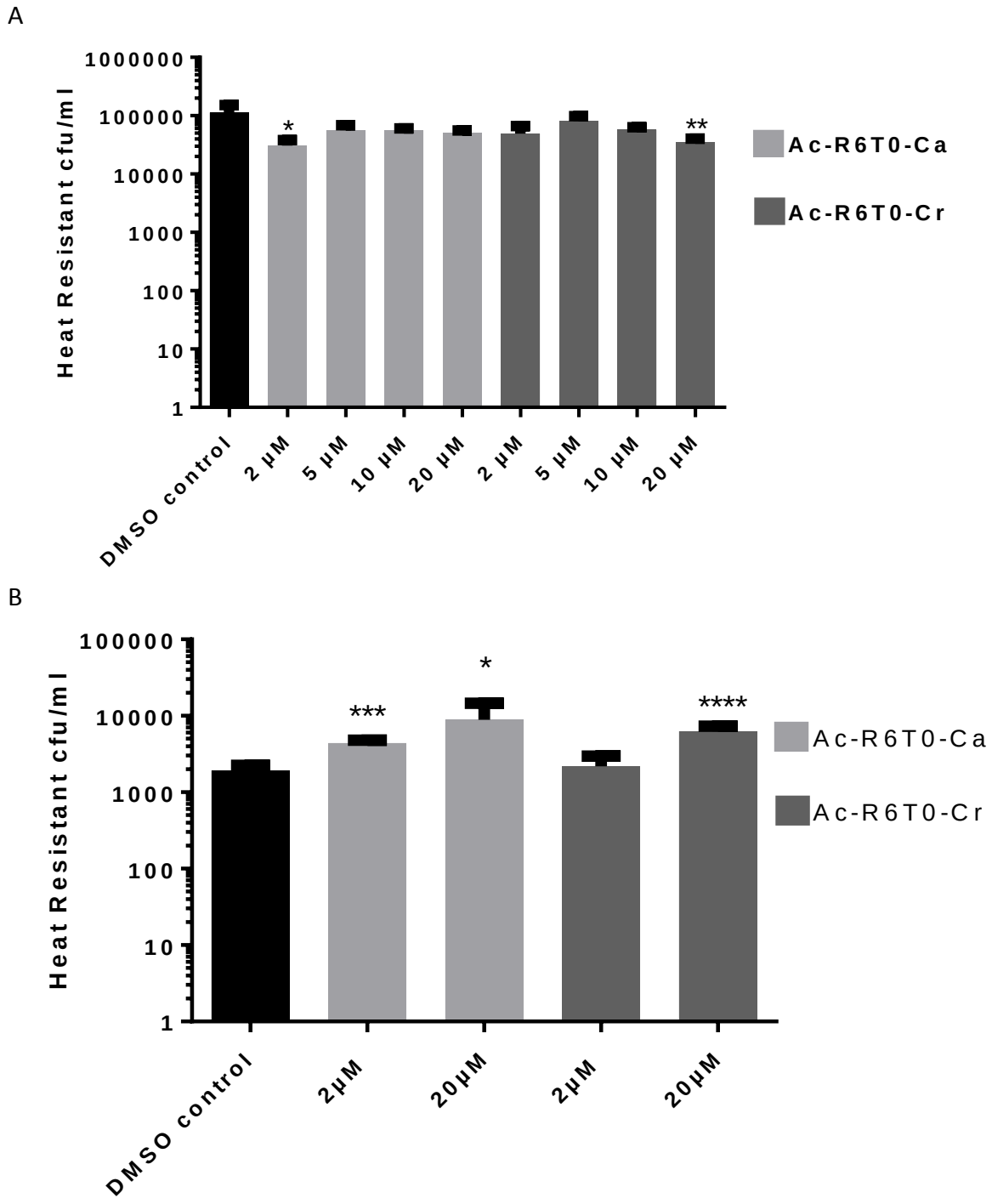


Figure 5.7 Effect of *C. roseum* derived AgrD peptide on sporulation of *C. acetobutylicum*

ATCC 824 wild type (A) and *agrB*::CT*ermB* mutant (B) changes in sporulation as a result of inoculation with synthetic *C. acetobutylicum* (Ac-R6T0-Ca) and *C. roseum* (Ac-R6T0-Cr) AgrD derived synthetic peptides. Sporulation was measured as heat resistant cfu/ml after growth in 10 ml CBM-S 6% (w/v) glucose for 120h. Ac-R6T0-ca was used as a positive control and DMSO as a negative control. Results were compared to the DMSO

control through T-test analysis- results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

5.2.3.3. Sporulation of a *C. acetobutylicum agrB*::CT*ermB* mutant in the presence of *C. acetobutylicum* and *C. roseum* derived synthetic peptides

Given that the Ac-R6T0-Ca and Ac-R6T0-Cr displayed similar activating or inhibiting properties dependent on whether the native Agr system was functional, it was decided to observe the sporulation of an *agrB*::CT*ermB* mutant in the presence of defined ratios of these peptides. An initial hypothesis was that while still showing biological activity with respect to spore formation, the weaker inducing effect observed in an *agrB*::CT*ermB* mutant by the Ac-R6T0-Cr peptide may imply that *C. roseum* signalling still serves to act as a minor competitive inhibitor to *C. acetobutylicum* Agr signalling. However due to previous observations that peptide concentrations of Ac-R6T0-Ca in excess of 20 µM appeared to show a detrimental effect with respect to end point spore formation (Figure 5.3), it was decided to limit total peptide concentrations to 20 µM to attempt to avoid the effect caused by peptide saturation resulting in the appearance of competitive inhibition.

Figure 5.8 shows that when added to a total concentration of 20 µM mixtures of Ac-R6T0-Ca and Ac-R6T0-Cr peptides appeared to have no effect on the overall sporulation level of the *agrB*::CT*ermB* mutant when compared to single additions. Of the concentrations tested only the sole provision of the Ac-R6T0-Cr peptide was shown to have an effect that was significantly less than that observed for the Ac-R6T0-Ca peptide control.

Of note is that all provided peptide mixtures increased sporulation to approximately the same degree as 20 µM of pure Ac-R6T0-Ca peptide even when they contained as little as 2 µM of this peptide. Thus paradoxically, the presence of the Ac-R6T0-Cr peptide increased sporulation of the *agrB*::CT*ermB* mutant to the same degree as Ac-R6T0-Ca so long as the latter was present, which contradicts findings for the wild type (Figure 5.7A).

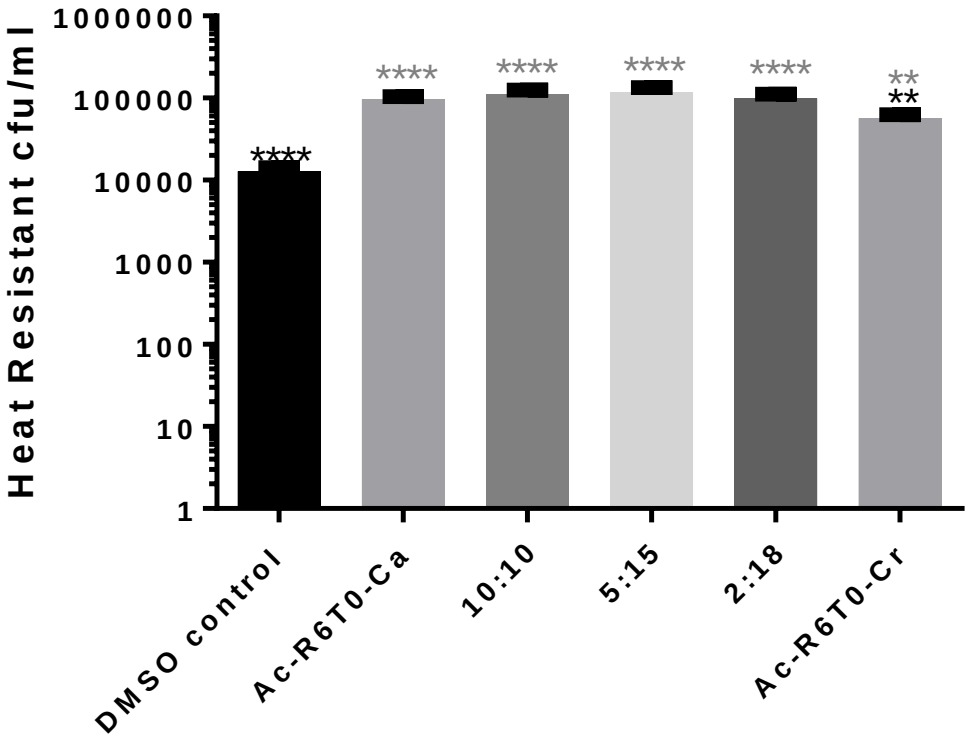


Figure 5.8 Sporulation of *Clostridium agrB*::CT*ermB* mutant in the presence of *C. acetobutylicum* Ac-R6T0-Ca and *C. roseum* Ac-R6T0-Cr peptide mixtures

Heat resistant cfu/ml values were determined for *agrB*::CT*ermB* mutants grown in 10 ml CBM-S 6% (w/v) glucose for 120 in the presence of 20 µM of the indicated peptide mixtures (neat or ratios of Ac-R6T0-Ca:Ac-R6T0-Cr). Results were analysed by T-test comparing DMSO control (grey) and the addition of 20 µM of Ac-R6T0-Ca (black) through T-test analysis, results above show P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

5.2.4. The role of granulose accumulation in relation to glucose availability

The annotated genome of *C. acetobutylicum* contains a gene cluster responsible for the synthesis of granulose from glucose. Of this cluster the *glgA* gene (Ca_c2239) is predicted to encode the synthase responsible for forming the α 1-4 glucan chains from ADP-glucose. With the availability of a Δ*glgA* mutant (Ehsaan et al. 2016) and the existing hypothesis that granulose accumulation and Agr regulations are a mechanism of energy conservation (Jabbari et al. 2013), it was decided to characterise granulose deficient mutants in the presence of varying glucose concentrations. This could be used to determine to what extent granulose accumulation is important in a carbon and energy limited environment and whether reduced sporulation in *agr* and Ca_c0082 mutants was due at least in part to a loss in granulose accumulation.

To examine the role of granulose accumulation with respect to solvent and sporulation phenotypes and in the context of the Agr system, *C. acetobutylicum* wild type was grown for 120 h on glucose ranging from 1-6% (w/v) alongside granulose deficient mutants: *agrC*::CT*ermB*, Ca_c0082::CT*ermB* and Δ*glgA*. By assessing the solvent and sporulation profiles of the strains at 48 h and 120 h time points the aim was to determine what role the Agr system and granulose accumulation of *C. acetobutylicum* might play with respect to solvent production and sporulation outside of the context of a carbon rich environment that they have formerly been examined within.

5.2.4.1. Sporulation profiles of granulose deficient mutants in relation to glucose availability

As can be seen from Figure 5.9 at both the 48 h and 120 h time points, across all tested glucose ranges, the Δ*glgA* mutant and all other granulose deficient mutants were shown to have significantly lower sporulation levels (heat resistant cfu/ml) compared to the wild type. It can

also be observed that across the range of glucose concentrations the granulose deficient mutants showed a greater variation in the level of sporulation. This supports the hypothesis that granulose accumulation serves a role in energy storage for the consistent formation of spores and is likely a key component of the decreased sporulation observed within *agr* and Ca_c0082 mutants.

It can also be noted that compared to the wild type all the granulose deficient strains required sufficiently higher levels of glucose in order to reach the observed peaks in spore levels. The wild type required as low as 2% glucose in order to see near maximum sporulation at both the 48 h and the 120 h time points while the granulose deficient mutants reached their significantly lower peaks at 3-4% glucose concentrations.

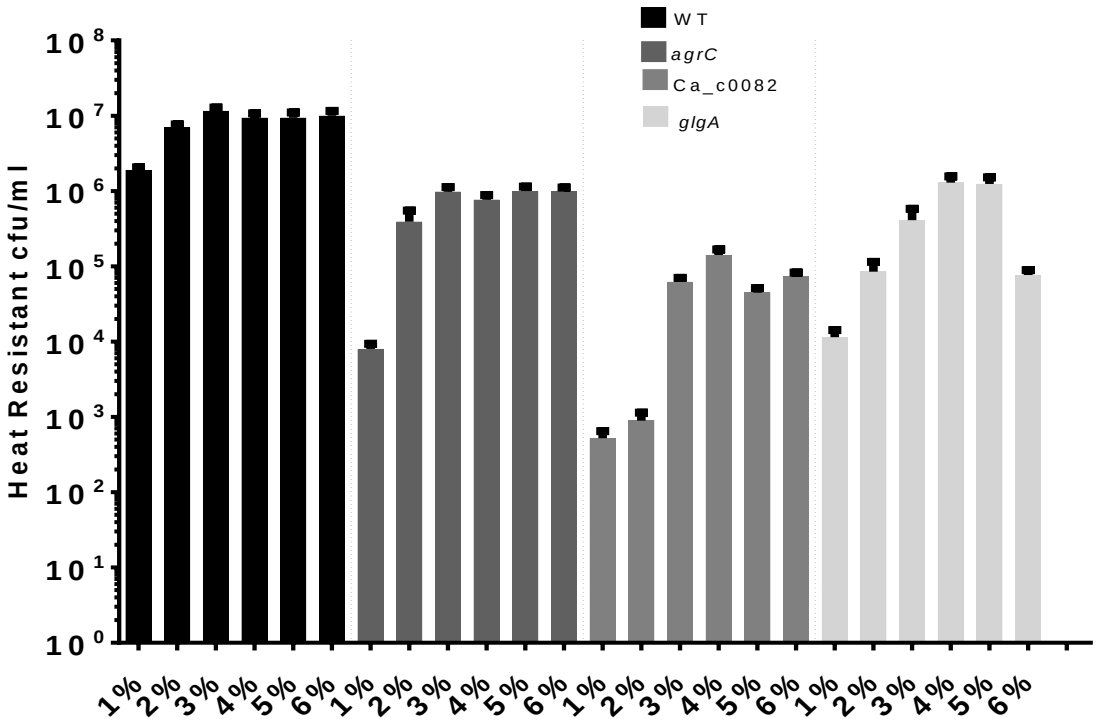
The granulose deficient mutants are further distinguished from the wild type through the increases in sporulation levels observed as glucose levels are raised and the resulting fold difference in the sporulation levels between these strains and the wild type. At 1% glucose while the wild type failed to achieve maximum levels of sporulation it produced over a 100 fold increase in heat resistant cfu/ml than the granulose deficient strains grown at the same concentration. As glucose concentration was increased only a 17 fold increase was observed for the wild type before sporulation levels reached near optimum levels, with a large proportion of this increase occurring when the concentration was raised to 2% glucose. The granulose deficient mutants however showed increases in the range of 100 to 1000 fold in sporulation levels at the 120 h time point (dependent on the mutant strain) as glucose concentration was increased up to 3-4%, showing staggered increases compared to the large single increase observed in the wild type.

As can be seen from Figure 5.9, the strains deficient in granulose accumulation showed markedly lower spore levels compared to the wild type after 48 h. Interestingly these levels

generally did not change much for the 120 h time point suggesting that the process of sporulation had largely been completed after 48 h for the granulose negative strains. While under carbon limitation at 1%, the wild type also showed no further increases in spore formation after 48 h.

Though the granulose deficient strains behaved similarly when compared to the wild type some clear differences could be observed when comparing behaviour of the individual strains. Most noticeable is that under glucose limitation of 1% both the *agrC*::CT*ermB* and the Ca_c0082::CT*ermB* mutants showed significantly lower final sporulation levels than the Δ*glgA* mutant. Though generally few significant differences between the strains were observed at the 120 h point, at 48 h the *agrC*::CT*ermB* mutant was shown as consistently producing a larger number of spores than the Ca_c0082::CT*ermB* across all tested glucose concentrations. Similarly, the Δ*glgA* mutant at 48 h showed significantly higher spore levels at 4% and 5% glucose compared to the Ca_c0082::CT*ermB* mutant though again by the 120 h time point both mutants generally showed few differences in sporulation levels.

A 48 h


WT
agrC
Ca_c0082
glgA
Heat Resistant cfu/ml
10 8
10 7
10 6
10 5
10 4
10 3
10 2
10 1
10 0
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

B 120 h

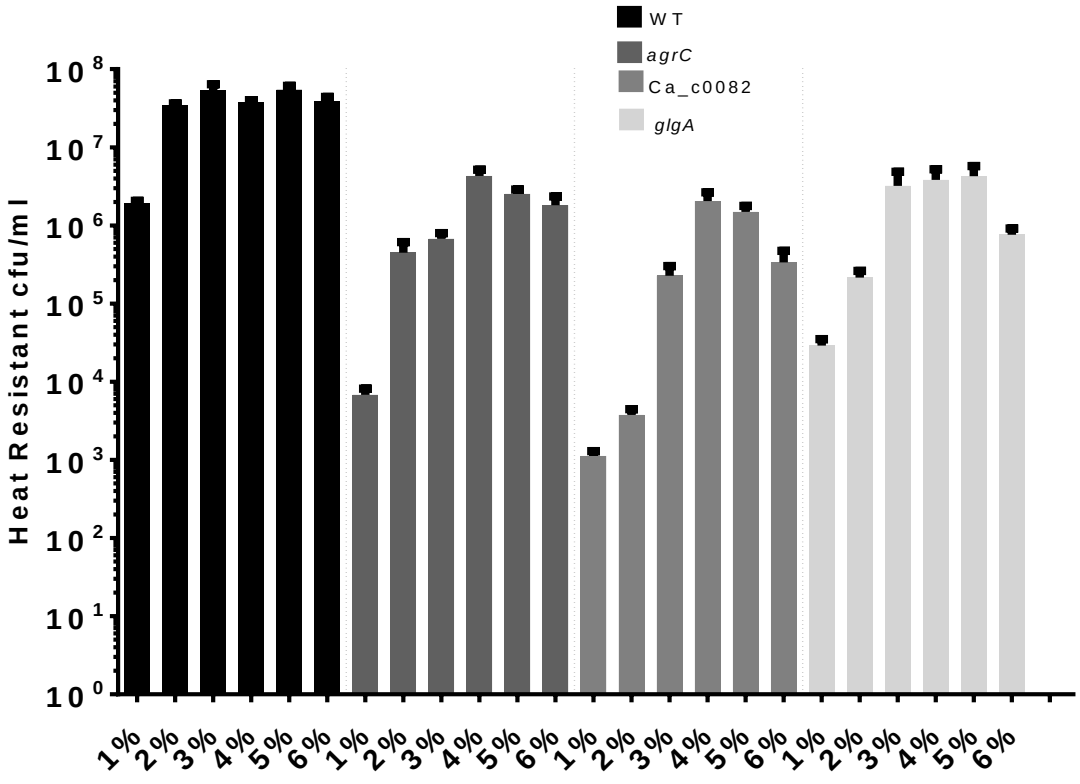

WT
agrC
Ca_c0082
glgA
Heat Resistant cfu/ml
10 8
10 7
10 6
10 5
10 4
10 3
10 2
10 1
10 0
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

Figure 5.9 Sporulation (heat resistant cfu/ml) of *C. acetobutylicum* wild type, *agrC*::CT*ermB*, Ca_c0082::CT*ermB and* Δ*glgA* at different glucose concentrations

Following growth in 10 ml of CBM-S at the indicated glucose (w/v) concentrations samples were taken at (A) 48 h and (B) 120 h time points and analysed for sporulation levels by assessing heat resistant cfu/ml. Statistical analysis through T-tests are summarised in Table 5.1 where n=9.

Table 5.1 Statistical significance of sporulation levels of *C. acetobutylicum* wild type, *agrC*::CT*ermB*, Ca_c0082::CT*ermB* and Δ*glgA* at different glucose concentrations

	48h						120h					
	1%	2%	3%	4%	5%	6%	1%	2%	3%	4%	5%	6%
Δ*glgA* vs:												
WT	****	****	****	****	****	****	****	****	****	****	****	****
***agrC* vs:**												
WT	****	****	****	****	****	****	****	****	****	****	****	****
Δ*glgA*	**	*	**	*	ns	ns	**	ns	ns	ns	ns	*
Ca_c0082 vs:												
WT	****	****	****	****	****	****	****	****	****	****	****	****
Δ*glgA*	**	ns	ns	****	**	ns	**	ns	ns	*	ns	ns
agrC	***	**	***	*	***	****	***	*	ns	ns	ns	**

Statistical analysis through T-tests are summarised above with P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

5.2.4.2. Acetone production in granulose deficient mutants in relation to glucose availability

Alongside sporulation levels, the solvent profiles for the different strains were analysed for all glucose concentrations. As can be seen from Figure 5.10 the final acetone concentrations showed no consistent significant differences from what was observed in the wild type compared to the granulose deficient mutants across all of the glucose concentrations tested. At 120 h, all strains showed increasing levels of acetone concentration when grown on medium containing 1-5% glucose alongside an increase in the deviation of the final concentration measured. At 6% glucose both the *agrC*::CT*ermB* and the Δ*glgA* mutant showed a small in

acetone concentration compared to growth using 5% glucose. However while the wild type and Ca_c0082::CT*ermB* mutant showed a decreases in 120 h acetone concentration upon exceeding 5% glucose the concentrations observed were not shown to be statically significantly different to the concentrations observed in *agrC*::CT*ermB* and the Δ*glgA* when provided 6% glucose.

What can be observed however, in Appendix Figure 6, is that production of acetone in the granulose deficient strains and the wild type varied with respect to when it occurred. By comparing the 48 h and 120 h data, it is evident that a certain degree of evaporative loss of acetone had occurred over the course of the fermentation, especially for the mutant cultures which had completed production earlier (see below) this is clear from the observed reduction in acetone concentration in some samples at the later time point due to the use on an "open system" which permitted gas exchange.

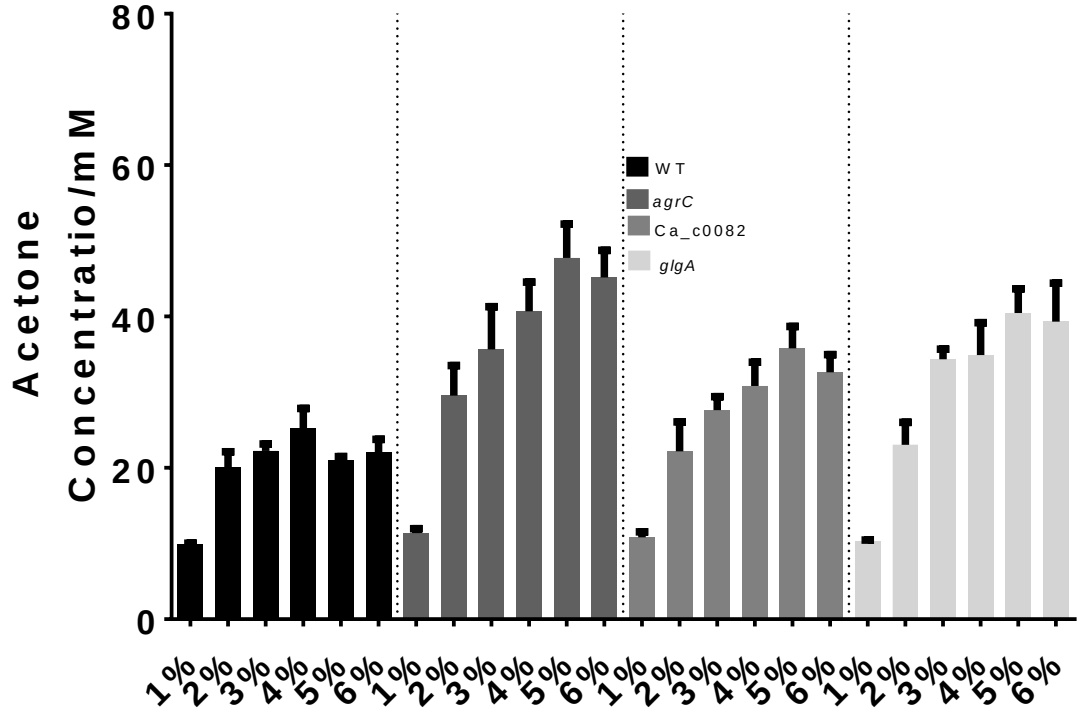
Nevertheless, a number of general conclusions can be drawn. After 48 h while there were no significant differences at low glucose concentrations (1-2%) all granulose negative mutants had produced more acetone than the wild type at higher concentration ranges (this being statically significant at 5-6%). Furthermore similar to what had been observed for sporulation, acetone concentrations did not change dramatically after 48 h, apart from the aforementioned evaporative loss, suggesting that acetone production had been completed or almost completed at this time point.

A significant difference can be observed with respect to acetone profiles at 48 hours, as the wild type appears to be limited to the production of acetone up to a value of approximately 20 mM, achieved already after the addition of only 2% glucose. The granulose deficient mutants however showed consistent increases in the level of acetone accumulating within the growth medium as the concentration of glucose in the media was increased; they also showed more variability in the level of acetone produced, reflected by larger standard deviations, compared

to the more consistent levels observed, reflected by the larger standard deviations, between replicates of the wild type.

At 48 h, the *agrC*::CT*ermB* mutant showed the largest mean acetone values of all strains with statistically significant differences to the wild type at glucose concentrations as low as 3%. The *agrC*::CT*ermB* mutant was also shown to produce significantly more acetone than the Ca_c0082 mutant at 5-6% glucose at this time point.

A 48 h


Acetone Concentratio/mM
80
60
40
20
0
WT
agrC
Ca_c0082
glgA
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

B 120 h

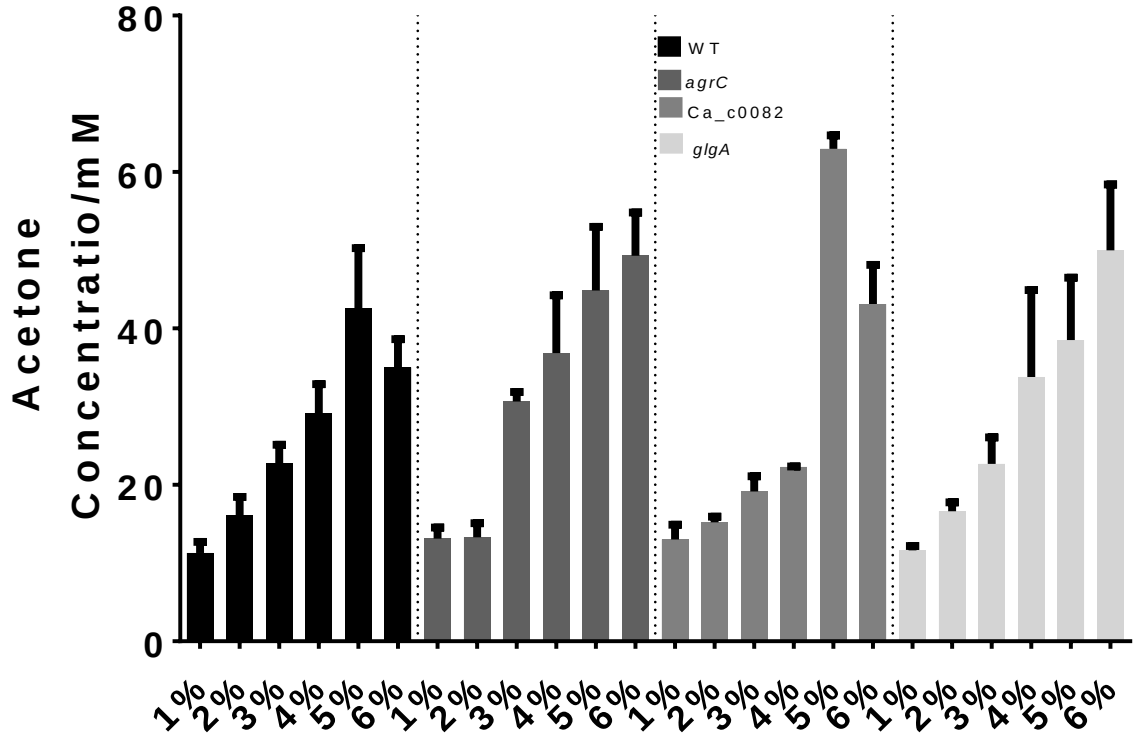

Acetone Concentratio/mM
80
60
40
20
0
WT
agrC
Ca_c0082
glgA
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

Figure 5.10 Acetone concentrations of *C. acetobutylicum* wild type, *agrC*::CT*ermB*, Ca_c0082::CT*ermB* and ΔglgA mutants at different glucose concentrations

Following growth in 10 ml of CBM-S at the indicated glucose concentrations (1%-6%) samples were taken at (A) 48 h and (B) 120 h time points and acetone concentrations determined. Statistical analysis through T-tests are summarised in table 5.2 where n=9.

Table 5.2 Statistical significance of acetone concentrations of *C. acetobutylicum* wild type, *agrC*::CT*ermB*, Ca_c0082::CT*ermB* and Δ*glgA* at different glucose concentrations

	48h						120h					
	1%	2%	3%	4%	5%	6%	1%	2%	3%	4%	5%	6%
Δ*glgA* vs												
WT	ns	ns	*	ns	**	**	ns	ns	ns	ns	ns	ns
***agrC* vs**:												
WT	ns	ns	*	*	***	**	ns	ns	ns	ns	ns	ns
Δ*glgA*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Ca_c0082 vs:												
WT	ns	ns	ns	ns	**	*	ns	ns	ns	ns	ns	ns
Δ*glgA*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	*	ns
agrC	ns	ns	ns	ns	*	*	ns	ns	*	ns	ns	ns

Statistical analysis through T-tests are summarised above with P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

5.2.4.3. Butanol production in granulose deficient mutants in relation to glucose availability

As with acetone concentration 120 h butanol concentrations were observed in all samples increasing as the provided glucose concentration increased from 1% to 5% in all strains alongside an increase in the standard deviation of the final butanol concentrations observed. The Δ*glgA* and *agrC*::CT*ermB* mutants again showed a continued increase in butanol production as glucose was increased to 6%, with the Δ*glgA* mutant showing a final butanol concentration that was significantly greater than that of the wild type (though not significantly different to any other granulose deficient strain). The Ca_c0082::CT*ermB* and wild type strains,

as was observed previously for 120 h acetone concentrations, have a lower mean butanol yield when provided with 6% glucose when compared with that observed at the provision of 5%.

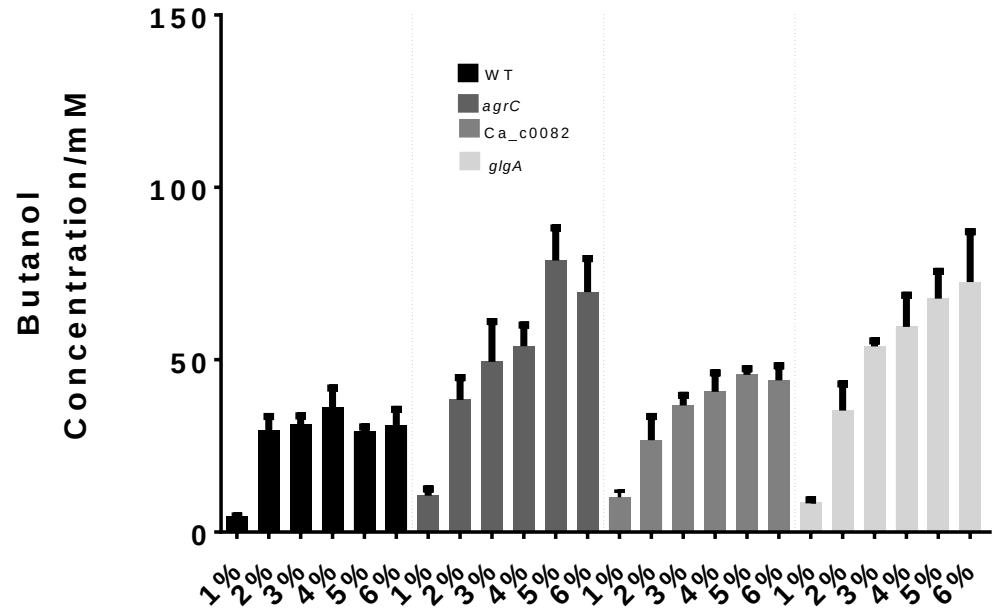
Though more variable, the 48 h wild type butanol concentrations showed a similar pattern to that of acetone with a marked increase observed between 1% and 2% glucose and then a fairly consistent level maintained (at approximately 30-35 mM) even as the available glucose was increased. A difference to note is that while both acetone and butanol are low at 1% in the wild type, butanol was particularly low undergoing over a 4 fold increase when glucose was raised to 2% (as opposed to the approximate 2 fold increase observed with acetone). The repression of butanol production in the wild type was so pronounced at 1% that it showed a statically significant difference to the values obtained for *agrC*::CT*ermB* and Ca_c0082::CT*ermB* which like the Δ*glgA* strain, also showed very low butanol concentrations under these conditions.

At the 48 h time point the granulose deficient strains showed consistent increases in mean butanol yield up to 5%, with Δ*glgA* showing this increase extending up to 6%. At 5% and above all the deficient strains showed significantly higher butanol production than the wild type, as was observed with acetone, though with respect to butanol the Δ*glgA* mutant showed significantly higher butanol concentration at glucose concentrations as low as 3% glucose. As was observed with acetone the Ca_c0082 mutant showed lower mean concentrations of butanol at this time point than the other granulose negative strains, though with butanol these difference became statistically significant at 4%. At 5% the difference is most pronounced with *agrC*::CT*ermB* showing an average of 78mM of butanol while the Ca_c0082::CT*ermB* showed only 45mM of butanol.

In summary, the butanol profile presented the same general patterns as seen for the acetone, showing higher concentrations at the 48 h time point at higher glucose concentrations for granulose deficient mutants. The wild type again reaches a maximum product concentration at

approximately 2% glucose at 48 h while the other mutants continue to show incrementally increased yields. As described for acetone a certain evaporative loss of butanol was observed for the mutant cultures for lower glucose concentrations (1-2%) although this was less pronounced.

A 48 h


Butanol Concentration/mM
150
100
50
0
WT
agrC
Ca_c0082
glgA
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

B 120 h

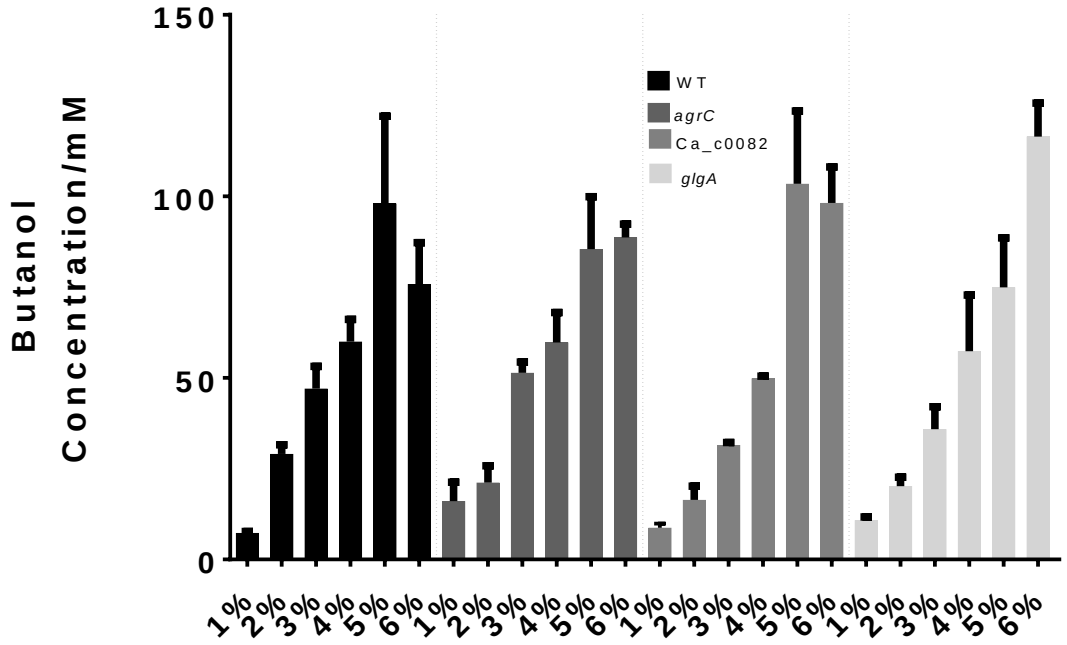

Butanol Concentration/mM
150
100
50
0
WT
agrC
Ca_c0082
glgA
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

Figure 5.11 Butanol concentrations of *C. acetobutylicum* wild type, *agrC::*CTermB, *Ca_c0082*::CTermB and Δ*glgA* mutant at different glucose concentrations

Following growth in 10 ml of CBM-S at the indicated glucose concentrations (1%-6%) samples were taken at (A) 48 h and (B) 120 h time points and butanol concentration determined. Statistical analysis through T-tests are summarised in Table 5.3 where n=9.

Table 5.3 Statistical significance of butanol concentrations of *C. acetobutylicum* wild type, *agrC*::CT*ermB*, Ca_c0082::CT*ermB* and Δ*glgA* at different glucose concentrations

	48h						120h					
	1%	2%	3%	4%	5%	6%	1%	2%	3%	4%	5%	6%
Δ*glgA* vs												
WT	ns	ns	*	*	**	*	ns	ns	ns	ns	ns	*
***agrC* vs**:												
WT	*	ns	ns	ns	***	*	ns	ns	ns	ns	ns	ns
Δ*glgA*	ns	ns	ns	ns	ns	ns	ns	ns	*	ns	ns	ns
Ca_c0082 vs:												
WT	*	ns	ns	ns	**	*	ns	*	*	ns	ns	ns
Δ*glgA*	ns	ns	ns	*	ns	ns	ns	ns	ns	ns	ns	ns
agrC	ns	ns	ns	**	*	ns	ns	ns	*	ns	ns	ns

Statistical analysis through T-tests are summarised above with P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001 where n=9.

5.3. Discussion

This study extended the work of Steiner et al. (2012) with respect to the active form of the AgrD derived peptide ring used a the signal molecule in *C. acetobutylicum* quorum sensing. Though organisms such as *L. plantarum* also utilise a tail-less signal peptide (Sturme et al. 2007), the lack of an N-terminal tail may in higher pH environments create instability in thiolactone structures in favour of the formation of an amide bond (Sewald & Jakubke 2009). Using an acetylated variant to the original six membered ring, Ac-R6T0, as well as a synthetic peptide in the amide rearranged form, it was possible to assess the relative biological activity of both structures and determine the likely form responsible for AgrCA activation. Restoration

of granulose accumulation and the sporulation was shown to be present for all forms of the peptide, however highest biological activity was observed for acetylated form of the original ring structure, Ac-R6T0-Ca. The unmodified original structure R6T0-Ca showed activation of sporulation between that seen for the Ac-R6T0-Ca and the rearranged structure, suggesting that a certain proportion of it had undergone rearrangement to the less active Amid-R6T0-Ca form. In the context of the wild type, the AgrD derived signal peptide is likely to exist in both the original and amid conformation, the extent likely varying as the pH of the culture medium both falls and then subsequently rises (Jones & Woods 1986).

The ability of the system to recognise both forms of the peptide is likely an evolutionary advantage gained by the organism to allow Agr quorum sensing to remain effective across a larger range of pH values in its natural environment. The system remaining functional when the peptide is more likely to take the amide form, allows Agr quorum sensing to continue in the absence of the pH drop observed when the organisms is grown in monoculture. This would still allow granulose accumulation and sporulation to occur when the organism reaches high cell densities, despite any potential off-setting effect of other organisms in the environment which may serve to limit the extent to which *C. acetobutylicum* may lower the environmental pH. Conversely when the environmental pH does begin to show a significant fall and Agr quorum sensing will likely primarily use the thiolactone structure, the organism will have increased capacity for the synthesis of granulose and the formation of spores in response to the growing toxicity of the acidic environment.

The clarification of the optimal signal peptide structure was of use in itself, as it provided an informed basis for the structural design of a *C. roseum* derived AgrD synthetic peptide. The acetylated peptide was also used to further confirm the role of Ca_c0082's in granulose accumulation was not through altering Agr signalling through peptide modification.

The recent observation that a linear variant of the usually circular AgrD signalling peptide could successfully be used for the complementation of *C. perfringens agrB* deficient strains (Ma et al. 2015), was not confirmed for *C. acetobutylicum* Agr quorum sensing. Using a linear equivalent to the opened ring structure as described by Ma et al. (2015) and variants including the predicted tail region, to account for the size difference of rings and potential peptide cleavage/processing sites, an *agrB*::CT*ermB* mutant failed to show any restoration of sporulation. While other variants of the linear peptide could be explored with respect to tail length, the tested peptides failed to achieve the activity observed by tailed ring structures (Steiner et al. 2012) suggesting further exploration would be unlikely to yield any significant biological activity.

Interestingly the closely related solventogenic organism *C. roseum* is predicted to produce an AgrD derived signal similar to that of *C. acetobutylicum* varying by a single amino acid in the ring structure, suggesting the possibility of interspecies communication. Unfortunately *C. roseum* could not be grown in mixed culture with *C. acetobutylicum* due to unfavourable growth characteristics including extensive aggregation and slime production. However the use of a synthetic peptide predicted from its AgrD sequence was of use in acquiring some evidence to this potential cross-species interaction. While little conclusive effect could be observed with the *C. acetobutylicum* wild type, for unknown reasons but likely due to the existing high background of native peptide, it was possible to show the stimulation of an Agr response in the *agrB*::CT*ermB* mutant.

The response for this *C. roseum* derived peptide was weaker than for the *C. acetobutylicum* native peptide, as stimulation of sporulation was not detectable at 2μM, shown to be effective for the latter and it was unable to stimulate granulose accumulation to a level detectable through iodine staining. Provision of 20 μM of the *C. roseum* peptide however showed significant

increases in the sporulation profile of an *agrB*::CT*ermB* mutant, suggesting a degree of interaction between the two organisms Agr systems.

The occurrence of cross inhibition between organisms due to similar peptide structures has been shown as a means of repressing virulence between *Staphylococcus* species (Otto et al. 2001). This phenomenon however does not appear to occur, with respect *C. roseum* and the effect of its peptide upon *C. acetobutylicum* Agr functionality. This cross-interaction became obvious when *C. roseum* and *C. acetobutylicum* peptides were added in combination. Interestingly the *C. roseum* peptide which had shown a much lower activity when provided in isolation it clearly showed increased sporulation to similar level to that observed for 20 µM of the native peptide when supplied in tandem with the latter further supporting the idea that the *C. roseum* AgrD derived peptide did not effectively competitively inhibit *C. acetobutylicum* Agr signalling.

In an environmental context the ability to effectively respond to the *C. roseum* peptide is likely an evolutionary advantage to *C. acetobutylicum* as its Agr system still serves to functionally respond to high cell densities and subsequently increases the rate of sporulation even when growing in the presence of a mixed population with *C. roseum*. This would serve to increase the likelihood of its survival as nutrient sources are exhausted and/or toxic products are formed by the ABE fermentation of each organism. However, it may also be the case that *C. roseum* effectively stimulates sporulation in *C. acetobutylicum* in order to divert its use of nutrient sources away from the production of toxic products of ABE fermentation and towards granulose accumulation and spore formation (as has been suggested to occur in Figure 5.9-Figure 5.11).

Due to the difficulties culturing *C. roseum* it was not possible to establish whether *C. acetobutylicum* is able to cross interact with the Agr system of *C. roseum*. In the future it would

be worth investigating whether in a mixed culture context *C. acetobutylicum* can effectively socially cheat (West et al. 2006; Diggle, Griffin, et al. 2007), using the *C. roseum* AgrD derived peptide to stimulate its own sporulation while using the production of its signal peptide to potentially inhibit these processes in *C. roseum*. Should a strain of *C. roseum* that is not so prone to aggregation and slime production be available (either by mutagenesis or by use of a different, more amenable, strain with the confirmed AgrD sequence), it would be of interest to co-culture the organism and observe the relative ratio of sporulation (by heat resistant cfu/ml) of each organism (distinguishing *C. roseum* colonies by their distinct pigment). Equally a pigment-less strain of *C. roseum* may be of use in cross streak analysis as granulose staining would no longer be obscured and potential cross inhibition or activation more easily detectable.

In previous work, Steiner et al. (2012) showed that the Agr quorum sensing system is a major regulator of granulose production in *C. acetobutylicum*. However the role of the system in sporulation remains less clear, with the inactivation of the system providing only a moderate 10-20 fold decrease in spore levels. As granulose has been postulated as a storage compound used to provide energy and carbon for the complex process of endospore formation (Reysenbach et al. 1986; Jones & Woods 1986), it was hypothesised that the reduction in spore formation in *agr* mutants was not the result of direct regulation but instead indirectly caused by the loss of granulose formation. The evidence provided here generally supports this view, however it also identifies key differences between the different mutants strains tested.

Studying the Δ*glgA* mutant has shown that in the absence of granulose accumulation alone, *C. acetobutylicum* is greatly hindered with respect to spore formation across all the growth conditions tested, although the degree to which this was observed has been shown to be dependent upon the availability of glucose within the culture. The wild type was able to show consistently high levels of sporulation from the provision of as little as 2% glucose, while the

provision of extra glucose to granulose deficient strains was shown to lead to direct increases in sporulation up to concentrations of 5-6% glucose. It is evident that the wild types ability to store excess glucose as granulose allows cells to better conserve and manage the energy and carbon needed for spore formation. It therefore greatly outperformed the deficient strains even when grown in 1% glucose, showing a near 100 fold increase in spores over the Δ*glgA* mutant under these conditions. Unable to effectively store carbon and energy in the form of granulose, the deficient mutants require supplementing with excess glucose (4-5%) in order to reach their peaks in terms of endospore. While excess glucose served to reduce to fold difference in sporulation from the wild type, the sporulation profiles of the characterised granulose deficient mutants still served to match what had previously been observed under the relevant conditions (Steiner et al. 2012).

At 120 h solvent profiles for granulose deficient mutants showed no consistent significant differences from the wild type, with the exception of butanol production at 1% glucose, likely the result of granulose deficient mutants having completed their fermentations considerably earlier than the wild type. By examining the 48 h solvent profiles it is however evident that glucose is being differentially utilised by the granulose deficient mutants. For the wild type at the 48 h time point the provision of glucose above 2% does not serve to increase the level of either acetone or butanol, while granulose deficient mutants show increasing levels of solvent production as the availability of glucose increases.

It is therefore reasonable to assume that in the wild type excess glucose is being effectively stored as granulose within cells, at most concentrations used in this study, resulting in a limitation of solvent production to the levels observed at 48 h due to competition between the different biological processes for this carbon source, a finding that challenges the classical view of granulose accumulating cells representing the most prolific solvent producing stage (Jones

et al. 1982). As the majority of sporulation is seen occurring between 48 h and 120 h the wild type subsequently utilises the granulose when it is required in order to provide the effective energy necessary for sporulation, likely resulting in the more the consistent sporulation profiles observed.

Being unable to effectively control this timed utilisation of glucose, the granulose deficient mutants might metabolise glucose at a faster rate, as indicated by the higher yields of acetone and butanol observed at 48 h. Upon reaching the phase of sporulation at which the wild type would likely utilise the granulose to provide carbon and energy for sporulation, the granulose deficient mutants are limited to the remaining glucose available in the media. In all cases this was shown to limit the amount of heat resistant spores that could be successfully formed, but is highlighted in the lower glucose conditions where sporulation is significantly reduced compared to the wild type. Growth under richer glucose concentrations allows for more glucose to remain available as cells require energy for sporulation and as a result the formation of spores in granulose deficient strains is far less reduced, compared to the differences seen between the mutants and the wild type at lower glucose concentrations. Supporting this are the observed differences in evaporative losses of acetone at low glucose concentrations between 48 h and 120 h in the granulose deficient strains and the wild type. The loss of acetone is far more pronounced in the deficient strains suggesting an early conclusion of acetone production in these strains, from which point on they would have been subject to evaporation in contrast to the wild type that continues the formation of acetone beyond 48 h.

While all granulose deficient mutants showed inferior sporulation and increased early solvent production compared to the wild type, both the *agrC*::Ct*ermB* and the Ca_c0082::CT*ermB* displayed profiles to differentiate themselves from a Δ*glgA* mutant, suggesting that their changed phenotypes were not simply due to a loss in granulose accumulation alone.

The *agrC*:::CT*ermB* mutant while behaving similar to the Δ*glgA* mutant with respect to sporulation at high glucose concentrations, produced significantly less spores when only 1% glucose was available. The greater loss in sporulation under 1% carbon limitation accompanied by no significant change in solvent production suggests that the Agr system is also likely responsible for regulating sporulation through a mechanism outside of granulose accumulation. These proposed other mechanisms clearly are secondary in importance compared to the regulation of granulose considering that *agrC*::CT*ermB* mutant and Δ*glgA* mutants showed no significant differences in sporulation profile when excess glucose is provided in order overcome the limitations on sporulation resulting from a loss of granulose accumulation.

Ca_c0082::CT*ermB* showed an even greater reduction of sporulation at low glucose concentrations than the *agrC*::CT*ermB* strain, and also showed the lowest levels of acetone, butanol and sporulation at 48 h of all the granulose deficient strain. The reduced sporulation and solvent profile at this earlier time point suggest the Ca_c0082::CT*ermB* mutant has a more delayed stationary phase metabolism and physiology than the *agrC*::CTermB mutant, likely due to independent regulation by the two systems.

Taken together these findings suggest that AgrCA and Ca_c0082 have regulatory roles affecting metabolism beyond what results from the loss of granulose accumulation. To follow up on these observations it would be interesting to further investigate independently regulated targets of AgrCA and Ca_c0082 to further determine their role in granulose accumulation, sporulation and solvent production. It would also be of interest to expand the time points at which granulose deficient mutants are examined in order to determine if the loss of granulose accumulation results in a faster switch to solvent production, substrate utilisation and the formation of the final product profile. It may also be of interest to expand the extent of glucose limitation upon the cultures to observe the relative effects upon the granulose deficient mutants,

though an initial study for the required level glucose to observe solvent and spore formation in the wild type would first be required.

In summary, this study has expanded the existing knowledge of the Agr system in *C. acetobutylicum* through confirming the adaptability of the AgrCA receptor system, showing that both a peptide having undergone an intramolecular rearrangement due to potential molecular instability and peptide derived from a related *C. roseum* are able to induce levels of activity. This serves to display adaptability the of this particular Agr system, as peptide signalling systems are often considered to be very strict with respect to specific peptide binding. These adaptations would likely allow the organism to better respond to high cell densities by increased sporulation under a greater range of environmental conditions (such as non-acidic or competitive environments). The system still however showed preferential responses to the 'optimal' peptide ring structure, producing reduced responses with the *C. roseum* and isomeric rearrangement variants and showing no response to synthesised linear peptides variants.

The study has also demonstrated the role of granulose accumulation with respect to energy and carbon storage, and how its absence results in decreased spore formation likely the result of the earlier utilisation of glucose in ABE fermentation. Previous results (see 4.2.6.2) showing similar losses of sporulation between *agr* and Ca_c0082 mutant have been distinguished from each other and from a Δ*glgA* mutant through examining the solvent and sporulation profiles across a range of glucose concentrations. This has provided evidence for AgrC and Ca_c0082 as regulating sporulation independently of granulose accumulation, though this appears very much as a secondary effect and may only be relevant under conditions of carbon limitation. The Ca_c0082::CT*ermB* mutant showed a slowed metabolism of acetone butanol and reduced sporulation levels compared to other granulose deficient mutants. While this suggests a slower utilisation of glucose, the hypothesised limitation to spore formation caused by the inability to

store carbon and energy for spore formation remains in effect as the Ca_c0082 still requires excess glucose to show maximum yields of spores and never reaches levels observed in the wild type.

6. General Discussion and Future Research

6.1. Background

The metabolic and physiological changes occurring during the stationary phase of solventogenic *Clostridium* spp. have long been known (Amador-Noguez et al. 2011; Grimmler et al. 2011) and since the expansion of available genetic tools over the last decade, the manipulation of *C. acetobutylicum* has yielded improved strains with respect to product profiles, substrate utilisation and increased tolerances to the toxic products of metabolism (Bahl & Lutke-Eversloh 2011; Cai & Bennett 2011; Willson et al. 2016; Zhu et al. 2011). Conversely strain optimisation through manipulation of regulatory pathways has been largely limited by the narrow understanding of the regulation processes of the stationary phase physiology. The initiation of sporulation and solvent production by phosphorylated Spo0A does not utilise a complex phosphorelay as in *Bacillus* (Eswaramoorthy et al. 2010; Jung et al. 2012), instead the mechanisms of Spo0A activation in *Clostridium* spp. are not as evidently conserved and remain largely uncharacterised (Dürre & Hollergschwandner 2004; Talukdar et al. 2015).

Though considerable effort has been invested into identifying the means by which stationary phase phenotypes in *C. acetobutylicum* are controlled, this has mostly centred around the confirmation of Spo0A as the master regulator of late phase phenotypes (Ravagnani et al. 2000) and its direct activation through the activity of orphan histidine kinases (Steiner et al. 2011) and sigma factors involved in the criss-cross regulatory cascade of endospore formation (Tracy et al. 2011; Al-Hinai et al. 2014). Outside of this only metabolic regulators such as the catabolite control protein CcpA (Ren et al. 2012) and the redox sensing Rex (Wietzke & Bahl 2012) have been investigated for their role in controlling stationary phase phenotypes. Due to the association of high cell densities and the activation of stationary phase metabolic and physiologic changes (Jones & Woods 1986; Tashiro et al. 2005; Choi et al. 2012), past investigations have linked Agr and RNPP-like quorum sensing systems to the regulatory

control of granulose accumulation, sporulation and/or solvent production (Steiner et al. 2012; Kotte 2013).

While these initial studies have identified these systems in their broad regulatory roles, the need to better understand both how these systems function and how they interact with respect to fitting into the global regulatory hierarchy of the metabolic and physiological shifts during stationary phase still requires deeper investigation. This study has therefore attempted to clarify, expand and integrate the working knowledge of these systems in order to strive towards a more complete and therefore manipulable model for *C. acetobutylicum* stationary phase regulation.

6.2. Key findings and future work

6.2.1. Ca_c0957 and Ca_c0958 as regulators of sporulation solventogenesis and granulose accumulation.

Initial investigation by Kotte (2013) identified two RNPP-like regulators, Ca_c0957 and Ca_c0958, as responsible for the regulation of granulose accumulation, sporulation and solventogenesis, and this has been confirmed in this study through the generation of in-frame deletion mutants. While precise deletion of the genes confirmed their proposed role as master regulators the increased accuracy of chromosomal techniques has provided new insight in to the level of expression of Ca_c0958 required for biological activity compared to what was seen by Kotte (2013). Through plasmid-based complementation using the strong *fdx* promoter Kotte (2013) was able to show phenotypic restoration of sporulation and solventogenesis to wild type levels in a Ca_c0958::CT*ermB* mutant, though the organisation of the gene on an operon with Ca_c0957 would suggest that expression was driven by a shared promoter. In this study the predicted promoter region upstream of Ca_c0957 was shown at a single copy to be effective in phenotypic restoration of a ΔCa_c0957 mutant but not a ΔCa_c0958 mutant. Furthermore the

complemented ΔCa_c0958 only showed partial phenotypic restoration when using a single chromosomally integrated Ca_c0958 under the control of the strong *fdx* promoter. These results combined with the observations of Kotte (2013) suggest that the level of expression required for Ca_c0958 activity is significantly greater than that of Ca_c0957 and due to the high level of expression required it may be hypothesised to form part of the positive feedback loop process that is commonly associated with quorum sensing systems (Williams et al. 2007). However, more experimentation is required to establish whether cellular Ca_c0957 and Ca_c0958 expression levels truly differ or whether the observed requirement for a strong promoter in the ΔCa_c0958 complementation experiments was caused by less efficient translation of the transcribed gene when taken out of native context. If a requirement for higher expression is confirmed it would be interesting to investigate whether the Ca_c0957 coding region harbours a second promoter responsible for higher Ca_c0958 expression.

The confirmation of biologically activity of Strep-tagged Ca_c0957 and Ca_c0958 opens up further opportunities to analyse the function and activity of the regulators. With respect to one opportunity is the resolution of the protein structures by X-ray crystallography which has already been performed for the PgrX, PlcR and NprR (Shi et al. 2005; Declerck et al. 2007; Zouhir et al. 2013a) direct DNA binding RNPP regulators. As these previous studies have shown this would be of use in understanding the mode of action of Ca_c0957 and Ca_c0958 with respect to the complexes they would form, as they were able to identify PlcR as forming a dimer, PgrX as forming a tetramer and NprR as switching from the former to the latter upon ligand binding (Shi et al. 2005; Declerck et al. 2007; Zouhir et al. 2013a). In addition to identify potential multimeric structures Ca_c0957 and Ca_c0958 may form, X-ray crystallography would also allow for the identification of heterodimeric complexes that the two proteins are suspected to form due to their similar phenotypes. Any potential structural resolution of Ca_c0957 and Ca_c0958 may also allow for the co-purification of bound proteins/ligands,

which has been previously exemplified with a quorum sensing system (Chen et al. 2002), which could subsequently be analyse by mass-spectrometry(Kool et al. 2011) or X-ray crystallography. Outside of co-purification the structural resolution of the TPR ligand binding pocket and similarities to existing resolved RNPP structures may further aid in the identification of a peptide ligand, through the direct co-purification of a bound signal peptide if isolated from *C. acetobutylicum*. Identification of a signalling peptide, and its provision through overexpression or supplementation using synthetic peptides, may represent a new means by which stationary phase phenotypes, such as solventogenesis, can be induced at earlier time points than is currently achievable.

Antibody targeting of the Strep-tagged Ca_c0957 and Ca_c0958 and subsequent immunoprecipitation would allow for the identification of DNA binding sites of the helix-turn-helix domains of the respective proteins through ChIP-seq analyses (Furey 2012). A similar approach has been used to identify gene directly controlled by LuxR, quorum sensing regulator, in *V. harveyi* using a similar approach based on a FLAG-tagged protein variants (Kessel et al. 2013). In *C. acetobutylicum* identification of the core regulon, in combination with the RNA-seq data previously obtained for a Ca_c0957::CT*ermB* mutant, would provide the potential for the identification of promoters that specify greater activity in stationary phase (and may be desirable for the expression of potentially toxic genes). Alternatively in tandem with the discovery of a signalling peptide these promoters may be of use in the creation of inducible expression systems for use in *C. acetobutylicum*.

The RNA-seq data obtained by Kotte (2013) suggested that the down regulation of histidine kinases Ca_c3319 and Ca_c090, responsible for Spo0A phosphorylation (Steiner et al. 2011), was the primary means by which Ca_c0957 (and likely Ca_c0958) were able to act as master regulators of stationary phase processes and explained their mutants' phenotypes similarity to

that of a *spo0A* mutant. However this study has shown that this mechanism alone is not sufficient to entirely eliminate all levels of functional Spo0A activity. ΔCa_c0957 and ΔCa_c0958 mutants showed a loss in granulose accumulation that was demonstrated as being the result of a loss in Agr signalling, as granulose accumulation could be observed in the presence of an AgrD derived signal peptide. Conversely a Δ*spo0A* mutant while still showing a granulose negative phenotype, showed functional Agr signalling and failed to produce granulose under any tested conditions. This is likely due the suggested regulatory role of σ^E, part of the endospore regulatory cascade, initiated by the activity of phosphorylated Spo0A, as a regulator of granulose formation (Tracy et al. 2011). The necessity of Spo0A demonstrated in this study and the ability to restore granulose accumulation in ΔCa_c0957 and ΔCa_c0958 mutants suggests that some level of Spo0A activity and subsequent activation of the sporulation cascade still occurs within the mutants, albeit not enough to facilitate the complete initiation of solvent production and sporulation. A hypothesis would be that regulation of sporulation and solventogenesis require a higher degree of Spo0A phosphorylation, and thus activation, that of granulose accumulation. The phenomenon of Spo0A gene regulation depending on the level of Spo0A phosphorylation has been previously described in *B. subtilis* (Fujita et al. 2005). An initial attempt to investigate the effect of the phosphorylation state of Spo0A upon the phenotype of *C. acetobutylicum* was hindered as the deletion of the histidine kinase Ca_c0323 was not achieved, due to strain instability. An alternate approach to consider in the future may be to express different kinases from an inducible promoter prior to deleting their native genes to generate a more stable mutants that can be observed for phenotypic characterisation under varying levels of Spo0A activity. A strategy to achieve this has recently been exemplified in *C. difficile* (Walker et al. 2016).

Through BLAST analysis it has been possible to identify that operons similar to the Ca_c0957-Ca_c0958 found in *C. acetobutylicum* are also found across multiple other *Clostridium sensu*

stricto group members. The conservation of the XRE HTH and TPR domains of the Ca_c0957 and the HTH_19 domain of the Ca_c0958 have been shown to be present in adjacent genes in multiple Clostridia species, as demonstrated in Figure 6.1. Figure 6.1 only represents a few of the identified conserved genes as in addition evidence of the conservation of a Ca_0957/Ca_0958 system can be further observed in *C. autoethanogenum* (CLAU_2639 and CLAU_2640), *C. carboxydivorans* (Ccar_22755 and Ccar_22756) and *C. pasteurianum* (CLPA_c14500 and CLPA_c14501) amongst other strains of industrial interest. Through in-frame deletion/ClosTron inactivation of these systems across multiple *Clostridium* spp. it may be possible to identify a conserved mechanism by which stationary phase phenotypes are regulated across a diverse range of organisms previously poorly understood with respect to how these phenotypes are regulated outside of the activation of Spo0A by studied orphan kinases and the influence of metabolic regulators (Talukdar et al. 2015).

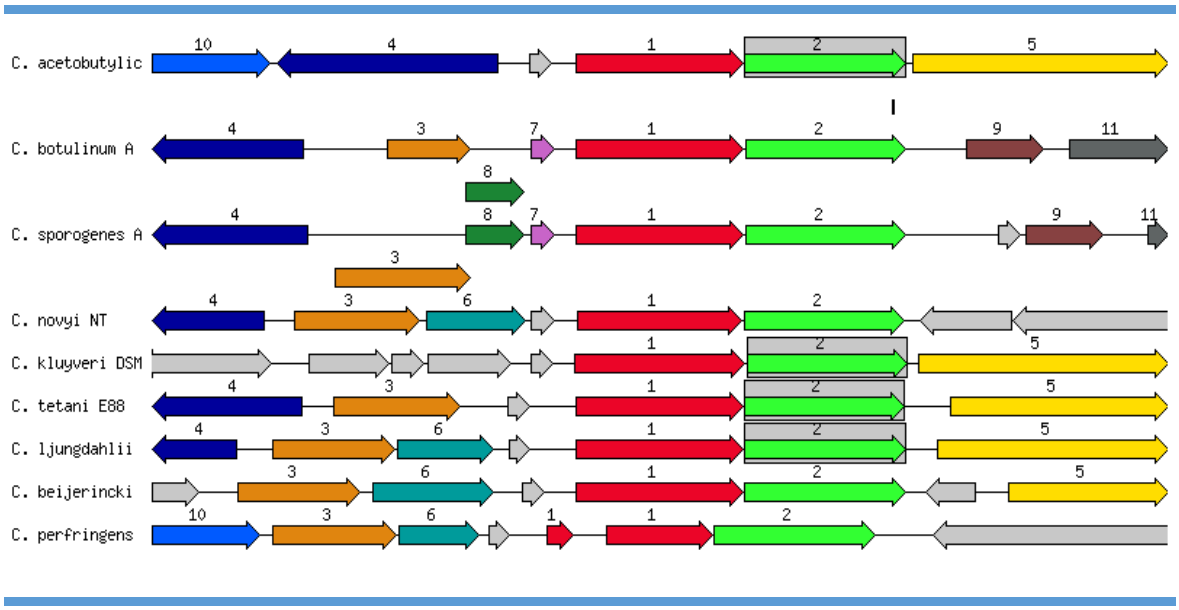


Figure 6. 1 Examples of the conservation of Ca_c0957 and Ca_c0958 in the genus *Clostridium*

Comparison of conserved regions surrounding Ca_c0957 in *C. acetobutylicum* to homologues found in *C. botulinum* ATCC 3502, *C. sporogenes* ATCC 15579, *C. novyi* NT, *C. tetani* E88, *C. ljundahlii* DSM 13528, *C. beijerinckii* 8052 and *C. perfringens* ATCC 13124 using The SEED Viewer. Genes labelled 1and 2 represent homologues to Ca_c0957 and Ca_c0958 respectively. Other genes found conserved within 8000 bp of the Ca_c0957 homologues across these organisms include (3, 7, 11) hypothetical proteins, (4) cell division protein

FtsH, (5) ClpB, (6, 8) ribonuclease HI-related proteins, (9) an intracellular nuclease, (10) and a rhomboid family serine protease.

In addition the Ca_c0957/Ca_c0958 regions conservation across *Clostridium* sensu stricto also includes the downstream *clpB* gene (see Figure 6.1), encoding a chaperone that rescues proteins from an aggregated state (Saibil 2013; Doyle et al. 2013). Interestingly, this gene is predicted to be co-transcribed with Ca_c0957 and Ca_c0957 (Venkataramanan et al. 2013), raising the question as to whether is role may be linked to their activation and proper functioning and/or stresses that arise as the result of solvent formation. In *E. coli clpB* is required for the activation of the RK2 plasmid replication initiation protein TrfA (Konieczny & Liberek 2002). As some other *Clostridium* have been observed to co-locate this *clpB* gene downstream of the paired RNPP-like regulators, it may be of interest to investigate whether *clpB* could play similar activating role with respect to the functionality of the Ca_c0957/Ca_c0958 system. This could be initially investigated through in-frame deletion of the *clpB* gene and observing if any significant change is observed in the stationary phase phenotypes associated with Ca_c0957 and Ca_c0958 occur.

Due to potential influences on stationary phase regulation, and thus the processes controlled by Ca_c0957 and Ca_c0958, a recent publication of Al-Hinai et al. (2014) was investigated, who reported the a role of σ^k in the regulation of *spo0A* expression, solventogenesis and granulose accumulation. It was decided to investigate the potential role σ^k may serve in the hierarchy of the quorum sensing, histidine kinase and Spo0A-based regulatory control of stationary phase phenotypes. After extensive investigation the granulose and sporulation phenotypes observed by Al-Hinai et al. (2014) could not be observed in the $\Delta\sigma^k$ mutants generated in this study. The $\Delta\sigma^k$ mutants also showed functioning Agr quorum sensing and the reported phenomenon of acid crashing was never observed during batch culture growth, instead showing the successful reassimilation of acids and initiation of solvent production. $\Delta\sigma^k$ mutants showed only a

sporulation-negative phenotype, and while the transcriptional expression of *spo0A* was not determined in this study, no evidence could be observed for a loss of the master regulators biological activity, supporting its original proposed role, being primarily late phase regulator of sporulation.

6.2.2. LanC-like peptide cyclase Ca_c0082 and its role in the regulation of granulose accumulation

The Ca_c0082 gene downstream of the *agrCA* operon, was shown in RNA-seq data to be down regulated approximately 100-fold when Ca_c0957 was inactivated (Kotte 2013). Inactivation of Ca_c0082 in this studied revealed that it plays a role in the regulation of granulose accumulation and sporulation of *C. acetobutylicum*. Due to proximity of the gene to the *agr* locus and the similarity the gene shares with peptide cyclases involved in lantibiotic synthesis, the original hypothesis was that this gene played a role in Agr quorum sensing signalling in *C. acetobutylicum* as no other homologues for lantibiotic production and signalling genes were identified. LanC peptide cyclases act alongside LanB dehydratases to mature lanthipeptides, through the formation of thioether bridges created through the dehydration of serine and threonine residues, followed by the steroeselective intramolecular Michael-type addition of Cys thiols to the newly formed dehydroamino acids (Zhang et al. 2012). Evolutionarily LanC peptide cyclases share a common ancestry with LanM proteins which alone are responsible for catalysing the two reaction described previously and maintain the conserved zinc binding motif seen in LanC peptide cyclases, though bare no homology to LanB dehydratases (Zhang et al. 2012). Though baring most similarity to a LanC like peptide cyclase, which would be expected to have little function without the LanB dehydratase, Ca_c0082 may instead function more like a LanM enzyme and solely be able to catalyse this reaction as *C. acetobutylicum* lacks a LanB homologue.

A role in the modification of the AgrD derived peptide in the signalling pathway was originally postulated as AgrD derived peptides are known to to undergo significant processing and maturation by enzymes including the AgrB and SpsB signal peptidase (Thoendel 2012). This process still remains not fully understood and evidence has even recently suggested in *S. aureus* the N-terminal leading sequence undergoes subsequent processing to become a biologically active peptide (Schwartz et al. 2014). Due to the unique nature of *C. acetobutylicum* Agr signalling peptide being identified as a tail-less six membered ring (Steiner et al. 2012) it was hypothesised that the LanC like peptide cyclase activity may facilitate the formation of this uncommon signalling peptide form.

This hypothesis was eliminated however, due to the observation that Ca_c0082*::*CT*ermB* mutants were able to induce granulose accumulation in an *agrB*::CTermB strain, despite themselves being granulose-negative, phenotypically matching what was observed with respect to granulose accumulation, but not solvent formation, in a Δ*spo0A* mutant. The continued expression of *agrBD*, likely due to the positive feedback loop, implies that the AgrCA two component response regulator system also remains active in a Ca_c0082 mutant, as *agrCA* mutants do not produce the AgrD signalling peptide (Steiner et al. 2012). It is currently unknown through which mechanism Ca_c0082 influences granulose accumulation, though at present the likely hypotheses are: that it acts in the same regulatory pathway as Spo0A due to their similar granulose and Agr phenotypes or is regulated by the Agr system and is the effector by which this quorum sensing system regulates granulose accumulation. It would therefore be of interest to delete and/or overexpress Ca_c0082 in the Δ*spo0A* and ΔCa_c0957 and backgrounds in order to determine if the granulose negative phenotypes these mutants display are in part due to their influence upon the expression of Ca_c0082. With respect to investigating whether the AgrCA controls the expression of Ca_c0082 this could be tested through RT-PCR

of the gene in an *agrB*::CT*ermB* mutant, both in the presence and absence of a synthetic AgrD peptide, Ac-R6T0-Ca.

However it is also possible that Ca_c0082 is responsible for modifying another peptide required for cell-to-cell signalling and granulose accumulation. Ca_c0082 mutants were shown to have a decreased endpoint sporulation levels similar to that of other granulose deficient mutants when grown in energy/glucose rich conditions suggesting the majority of its influence on this phenotype is through its influence on granulose accumulation. If this hypothesis is true it would be of interest to examine the secretome of the Ca_c0082:CT*ermB* mutant with that of the wild type and *agrB*::CT*ermB* mutant, in order to identify the presence/absence of potential peptide signalling molecules, through techniques such as multistage mass-spectrometry which have been successfully used to identify and sequence unique cyclic peptide structures (Mohimani et al. 2011) or the occurrence of unusual peptide signalling molecules such as a nine membered lactone ring AgrD based peptide from *S. intermedius* (Kalkum et al. 2003)

6.2.3. Sporulation and solvent production in granulose deficient mutants of *C. acetobutylicum*

Increased examination as to the role of granulose accumulation upon the sporulation and solvent profiles of granulose deficient mutants and the wild type, under varying glucose availability, has revealed that the loss of granulose accumulation is the main contributor to the phenotypes seen in *agr* and Ca_c0082 mutants.

The wild-type was observed at 48 h and 120 h as consistently producing a 10-fold greater level of sporulation than the granulose deficient mutants when under rich glucose conditions. However while the wild type maintained its higher level of sporulation down to 2% glucose and only showing a small decrease at 1% glucose, all granulose deficient strains showed regular decreases in sporulation after dropping below 3-4% glucose. Under glucose concentrations of

1% this resulted in a 100-1000 fold decrease in the granulose deficient mutants compared to the wild type, reinforcing the previously postulated hypothesis that the compound primarily serves as an energy storage mechanism for sporulation (Reysenbach et al. 1986).

The accumulation of granulose also likely coincides with a slowed fermentation process in the wild type as glucose is stored in granulose in place of being metabolised. This was supported by the 48 h acetone and butanol concentrations which upon reaching 2% were seemingly at their peak for that time point in the wild type, as excess glucose failed to increase the concentrations of either solvent. Conversely the granulose deficient mutants do not show this limitation to their production of solvents, instead showing regular increases in both acetone and butanol production, exceeding the levels observed in the wildtype, as the availability of glucose is increased. At the 120 h time point the differences between the acetone and butanol concentrations observed for the granulose deficient mutants and the wild type are insignificant suggesting that the wild type has subsequently metabolised the stored granulose to produce solvents but also the aforementioned increase in sporulation.

The use of energy storage compounds to prolong metabolism and therefore aid survival has also been observed in *S. mutans*, which has been suggested to accumulate its glycogen like intracellular polysaccharide not only to provide energy in the event of starvation but also maintain acid production to maintain a favourable environmental pH to assist the organism's persistence (Busuioc et al. 2009). Therefore the role of granulose plays in potentially limiting solvent production through spore formation would be of interest in further examination, particularly with respect to how larger scale fermentations should be optimised (either through substrate limitation or the use of granulose deficient strains) in order to avoid a reduced productivity due to the organism prioritising carbon and energy for survival as opposed to a maximised product yield.

Between granulose deficient mutants some minor differences were observable with respect to both their solvent and sporulation phenotypes, though these were primarily only observable under specific glucose conditions. Both the *agrC*::CT*ermB* and Ca_c0082::CT*ermB* mutant showed significantly lower granulose accumulation than a Δ*glgA* mutant when grown at a concentration of only 1% glucose. In addition at 48 h it was observed that the *agrC*::CT*ermB* mutant showed significantly higher butanol and acetone concentrations than the Ca_c0082::CT*ermB* mutant when higher glucose concentrations were available in the medium. These factors suggest that outside of granulose accumulation both the Agr system and Ca_c0082 are responsible for influencing different regulator targets, which may be of interest in characterising in the future through transcriptional analysis techniques (such as RT-PCR or RNA-seq) using these mutants when grown in both limiting and excessive glucose levels.

6.2.4. Agr peptide signalling

The AgrD structure-function analyses presented here extends beyond the work of Steiner et al. (2012), who had identified a six membered tail-less ring as the optimal structure for the induction of sporulation and granulose accumulation in an *agrB*::CT*ermB* mutant. Although the proposed ring structure was confirmed as the most biologically active form, the AgrC receptor was shown to possess a degree of tolerance with respect to the signal peptides it was capable of responding to. This may have evolved in response to a molecular rearrangement that the ring structure may undergo as a result of thiolactone ring instability (which it may suffer from in a less acidic environment due to the lack of a peptide tail), resulting in a similar but amide bond based ring structure. Though less effective than a stabilised version of the original six ringed structure (Steiner et al. 2012), the amide bond based ring representing a stable molecular rearranged structure was shown to effectively induce granulose accumulation and increase sporulation in an *agrB*::CT*ermB* mutant. This likely represents an evolutionary

advantage for the organism as it allows Agr signalling to continue when grown in an environment where the activity of other organisms may not facilitate a drop in pH needed for thiolactone ring stability. A subsequent fall in the pH of the environment, would then cause an increase in AgrCA activity allowing the cells to effectively limit further metabolism through diversion of carbon towards granulose accumulation and increased survival through increased spore formation.

The system also showed flexibility with respect to its response to a peptide signal derived from the predicted *C. roseum* AgrD, varying by only an amino acid to that of the *C. acetobutylicum* peptide, was shown to be less effective but still resulted in an increase in sporulation of an *agrB::*CT*ermB* mutant when provided in a sufficient concentration or when supplemented alongside the native peptide. While it was not possible to assess the effects of *C. acetobutylicum* Agr signalling upon *C. roseum*, the experiment suggests that *C. acetobutylicum* can exploit the competing strains signalling. Further exploration of the potential social interaction between *C. acetobutylicum* and *C. roseum* would be on interest, as though closely related these organisms have distinctly different restriction systems, likely limiting DNA transfer between the organisms (based on unpublished genome sequence data; Winzer personal communication), and it has been postulated that social conflict can readily diversify quorum sensing to through driving both the occurrence of cheating and cheating immune signalling mutation (Eldar 2011).

In this context the limited exchange of genetic material between the organisms may be of interest to explore if one of the two organisms had evolved to 'listen in' on signalling of the other or alternatively if the systems enable interspecies communication (Olson et al. 2014) and collaboration (Atkinson & Williams 2009; Hosni et al. 2011). Furthermore, the *C. aurantibutylicum* genome has recently been sequence (Poehlein, University of Gottingen, unpublished) and potentially provides another AgrD ring structure, that similar to that of *C.*

roseum, shows high conservation around the proposed peptide ring structure only varying by a single amino acid at the same position as the putative *C. roseum* AgrD. As with the *C. roseum* Agr system this may be of interest in investigating with respect to cross inhibition/activation with the *C. acetobutylicum* system, and the effect it plays upon sporulation and granulose accumulation, in order to further understand *Clostridium* quorum sensing in an environmental context.

Despite recognition of amid-bond based rings and the *C. roseum* AgrD derived peptide, the Agr quorum sensing of *C. acetobutylicum* was not observed to be as robust as the recent observations in *C. perfringens* may have suggested (Ma et al. 2015) as *C. acetobutylicum agrB::*CT*ermB* showed no capacity to respond to the provision of AgrD derived linear signalling peptides.

In conclusion this work has developed the known and proposed roles and functions of the Agr and RNPP quorum sensing systems of *C. acetobutylicum* with respect to the mechanisms, mode of action and regulatory role beyond that which was previously established, developing the former regulation model from that described in Figure 1.7 to what has now been established in Figure 6.2. This work acts as a basis to further develop the knowledge of structure, function and regulatory hierarchy of these systems in *C. acetobutylicum* to allow for better directed strain manipulation for the production of solvents in the future.

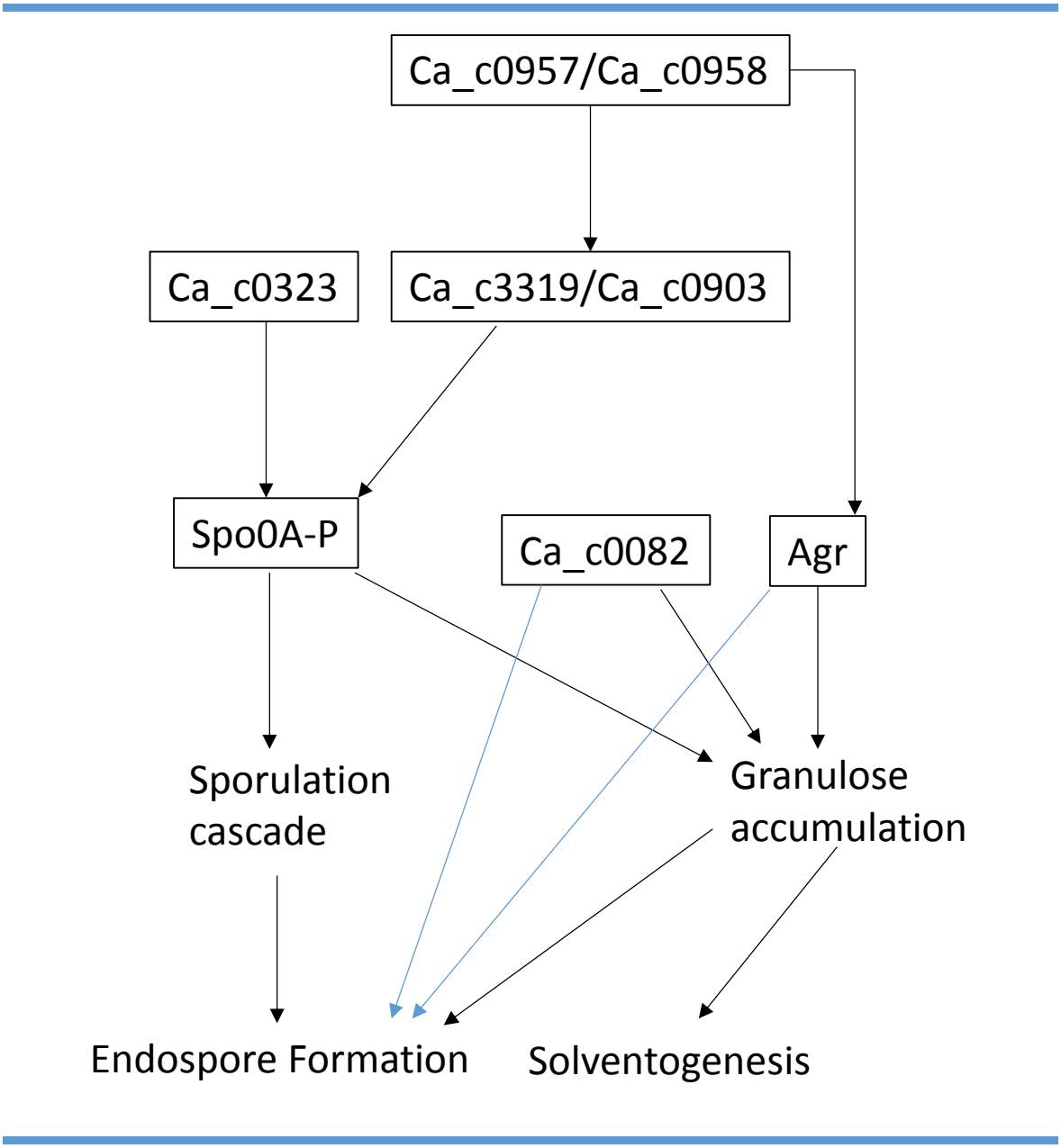


Figure 6. 2 The regulatory control of stationary phase phenotypes in *C. acetobutylicum* following the completion of this study.

Ca_c0957/Ca_0958 act to control the activity of histidine kinases (Ca_c0903/Ca_c3319) that act upon Spo0a to influence solventogenesis and sporulation, however it regulates granulose accumulation through a more direct *agrBD* induction. The Agr system alone does not control granulose accumulation as both Spo0A and Ca_c0082 have been shown to be essential for this process to occur. Granulose accumulation plays a significant role in increasing the level of endospore formation and is the primary means through which the Agr system and Ca_c0082 regulate spore formation, though both show some activity outside of this pathway when carbon/energy is limited. Additionally early solvent formation appears to be influenced negatively when granulose accumulation is allowed to occur, while the proposed early phase regulator of σ^k was shown to have no influence outside of its role in the sporulation cascade.

Bibliography

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

Achari, G.A. & Ramesh, R., 2015. Characterization of bacteria degrading 3-hydroxy palmitic acid methyl ester (3OH-PAME), a quorum sensing molecule of *Ralstonia solanacearum*. *Letters in Applied Microbiology*, 60(5), pp.447–455.

Al-Hinai, M. a, Fast, A.G. & Papoutsakis, E.T., 2012. Novel system for efficient isolation of Clostridium double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. *Applied and environmental microbiology*, 78(22), pp.8112–21.

Al-Hinai, M. a, Jones, S.W. & Papoutsakis, E.T., 2014. σK of *Clostridium acetobutylicum* is the first known sporulation-specific sigma factor with two developmentally separated roles, one early and one late in sporulation. *Journal of bacteriology*, 196(2), pp.287–99.

Allan, E., Mullany, P. & Tabaqchali, S., 1998. Construction and characterization of a *Helicobacter pylori* clpB mutant and role of the gene in the stress response *J Bacteriol*, 180(2), pp.426–429.

Amador-Noguez, D. et al., 2011. Metabolome remodeling during the acidogenic-solventogenic transition in *Clostridium acetobutylicum. Applied and Environmental Microbiology*, 77(22), pp.7984–7997.

Ansaldi, M. et al., 2002. Specific activation of the Bacillus quorum-sensing systems by isoprenylated pheromone variants. *Molecular Microbiology*, 44(6), pp.1561–1573.

Atkinson, S. & Williams, P., 2009. Quorum sensing and social networking in the microbial world. *Journal of the Royal Society Interface*, 6(40), pp.959–978.

Bahl, H. et al., 1986. Nutritional Factors Affecting the ratio of solvents produced by

Clostridium acetobutylicum. *Applied and Environmental Microbiology*, 52(1), pp.169–172.

Bahl, H. & Lutke-Eversloh, T., 2011. Metabolic engineering of *Clostridium acetobutylicum* : recent advances to improve butanol production. *Current Opinion In Biotech.*, 22(5), p.634.

Bainton, N.J. et al., 1992. N-(3-oxohexanoyl)-L-homoserine lactone regulates carbapenem antibiotic production in *Erwinia carotovora*. *Biochemical Journal*, 288(Pt 3), pp.997–1004.

Barber, C.E. et al., 1997. A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. *Molecular Microbiology*, 24(3), pp.555–566.

Bassler, B.L. et al., 1993. Intercellular signalling in *Vibrio harveyi*: sequence and function of genes regulating expression of luminescence. *Molecular Microbiology*, 9(4), pp.773–786.

Bauer, W.D. & Mathesius, U., 2004. Plant responses to bacterial quorum sensing signals. *Current Opinion in Plant Biology*, 7(4), pp.429–433.

Becker, J.J., 1988. Fixed Bed Gasification of Lignocellulosic Biomass: The Cemagref Process BT - Research in Thermochemical Biomass Conversion. In A. V Bridgwater & J. L. Kuester, eds. Dordrecht: Springer Netherlands, pp. 1016–1025.

Bertsch, J. & Müller, V., 2015. Bioenergetic constraints for conversion of syngas to biofuels in acetogenic bacteria. *Biotechnology for Biofuels*, 8(1), pp.1–12.

Biebl, H., 2001. Fermentation of glycerol by *Clostridium pasteurianum* - Batch and continuous culture studies. *J Ind Microbiol Biotechnol*, 27(1), pp.18–26.

Biotechnology, E.R.A.I., 2016. Rational Engineering of Advanced Clostridia for Transformational Improvements in Fermentation. Available at: http://www.era-ib.net/reactif [Accessed June 27, 2016].

Boisset, S. et al., 2007. *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes & Development*, 21(11), pp.1353–1366.

Bongiorni, C. et al., 2006. Rap phosphatase of virulence plasmid pXO1 inhibits *Bacillus anthracis* sporulation. *Journal of Bacteriology*, 188(2), pp.487–498.

Bowles, L. & Ellefson, W., 1985. Effects of Butanol on *Clostridium-acetobutylicum. Applied and Environmental Microbiology*, 50(5), pp.1165–1170.

BP, 2016. BP Statistical Review of World Energy 2016. , (June), pp.1–48.

Brymora, A., Valova, V.A. & Robinson, P.J., 2001. Protein-Protein interactions identified by pull-down experiments and mass spectrometry. In *Current Protocols in Cell Biology*. John Wiley & Sons, Inc.

Bullifent, H., Moir, A. & Titball, R., 1995. The construction of a reporter system and use for the investigation of *Clostridium perfringens* gene expression. *FEMS Microbiol Lett.*, 15(131), pp.99–105.

Burbulys, D., Trach, K. a & Hoch, J. a, 1991. Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. *Cell*, 64(3), pp.545–52.

Busuioc, M. et al., 2009. Role of intracellular polysaccharide in persistence of *Streptococcus mutans*. *Journal of Bacteriology*, 191(23), pp.7315–7322.

Cai, X. & Bennett, G.N., 2011. Improving the *Clostridium acetobutylicum* butanol fermentation by engineering the strain for co-production of riboflavin. *Journal of industrial microbiology & biotechnology*, 38(8), pp.1013–25.

Cao, J. & Meighens, E.A., 1989. Purification and Structural Identification of an Autoinducer for the Luminescence System of *Vibrio harueyi*, *J Biological Chemistry* , 264(36), pp.21670–21676.

Cartman, S.T. et al., 2012. Precise manipulation of the *Clostridium difficile* chromosome

reveals a lack of association between the tcdC genotype and toxin production. *Applied and Environmental Microbiology*, 78(13), pp.4683–4690.

Cartman, S.T. & Minton, N.P., 2010. A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile*. *Applied and environmental microbiology*, 76(4), pp.1103–9.

Chatterjee, C. et al., 2005. Biosynthesis and mode of action of lantibiotics. *Chemical Reviews*, 105(2), pp.633–683.

Chaturvedi, V. & Verma, P., 2013. An overview of key pretreatment processes employed for bioconversion of lignocellulosic biomass into biofuels and value added products. *3 Biotech*, 3(5), pp.415–431.

Chen, F. et al., 2013. Quorum quenching enzymes and their application in degrading signal molecules to block quorum sensing-dependent infection. *International Journal of Molecular Sciences*, 14(9), pp.17477–17500.

Chen, X. et al., 2002. Structural identification of a bacterial quorum-sensing signal containing boron. *Nature*, 415(6871), pp.545–549.

Cheng, Z. et al., 2010. Structural basis of the sensor-synthase interaction in autoinduction of the wuorum sensing signal DSF Biosynthesis. *Structure*, 18(9), pp.1199–1209.

Chevalier, C. et al., 2010. *Staph aureues* RNAIII Binds to two distant regions of coa mRNA to arrest translation andpPromote mRNA degradation. *PLoS Pathog*, 6(3), p.e1000809.

Chhabra, S.R. et al., 2005. Extracellular Communication in Bacteria. In S & Schulz, ed. *The Chemistry of Pheromones and Other Semiochemicals II:* Springer Berlin Heidelberg, pp. 279–315.

Chiao, J.S. & Sun, Z.H., 2007. History of the acetone-butanol-ethanol fermentation industry in China: Development of continuous production technology. *Journal of Molecular Microbiology and Biotechnology*, 13(1–3), pp.12–14.

Choi, S.J. et al., 2012. Effects of nutritional enrichment on the production of acetone-butanol-ethanol (ABE) by *Clostridium acetobutylicum*. *Journal of Microbiology*, 50(6), pp.1063–1066.

Chuang, S.S.C., 2012. Conversion of Syngas to Fuels. In W.-Y. Chen et al., eds. *Handbook of Climate Change Mitigation*. New York, NY: Springer US, pp. 1605–1621.

Chylinski, K., Le Rhun, A. & Charpentier, E., 2013. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA biology*, 10(5), pp.726–37.

Committee, House. of Commons Environmental Audit ., 2008. Are Biofuels Sustainable? , First Report 2007-2008.

Cook, L.C. & Federle, M.J., 2014. Peptide pheromone signaling in *Streptococcus* and *Enterococcus*. *FEMS Microbiology Reviews*, 38(3), pp.473–492.

Cooksley, C.M. et al., 2010. Regulation of neurotoxin production and sporulation by a putative agrBD signaling system in proteolytic *Clostridium botulinum*. *Applied and environmental microbiology*, 76(13), pp.4448–60.

Cooksley, C.M. et al., 2012. Targeted mutagenesis of the *Clostridium acetobutylicum* acetone-butanol-ethanol fermentation pathway. *Metabolic engineering*, 14(6), pp.630–41.

Costerton, J.W., Stewart, P.S. & Greenberg, E.P., 1999. Bacterial Biofilms: A Common Cause of Persistent Infections. *Science*, 284(5418), pp.1318–1322.

Cousineau, B. et al., 1998. Retrohoming of a Bacterial Group II Intron. *Cell*, 94(4), pp.451–462.

Dabrock, B., Bahl, H. & Gottschalk, G., 1992. Parameters affecting solvent production by *Clostridium pasteurianum*. *Appl Environ Microbiol*, 58(4), pp.1233–1239.

Dahlsten, E. et al., 2013. Alternative sigma factor sigK has a role in stress tolerance of group I *Clostridium botulinum* strain ATCC 3502. *Applied and Environmental Microbiology*, 79(12), pp.3867–3869.

Darch, S.E. et al., 2012. Density-dependent fitness benefits in quorum-sensing bacterial populations. *Proceedings of the National Academy of Sciences of the United States of America*, 109(21), pp.8259–8263.

Darkoh, C. et al., 2015. Toxin synthesis by *Clostridium difficile* is regulated through quorum signaling. *mBio*, 6(2), pp.1–10.

Declerck, N. et al., 2007. Structure of PlcR: Insights into virulence regulation and evolution of quorum sensing in Gram-positive bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), pp.18490–5.

Deng, Y. et al., 2009. Differential modulation of *Burkholderia cenocepacia* virulence and energy metabolism by the wuorum-sensing signal BDSF and its synthase. *Journal of Bacteriology*, 191(23), pp.7270–7278.

Diggle, S.P., Griffin, A.S., et al., 2007. Cooperation and conflict in quorum-sensing bacterial populations. *Nature*, 450(7168), pp.411–414.

Diggle, S.P. et al., 2006. Functional genetic analysis reveals a 2-alkyl-4-quinolone signaling system in the human pathogen *Burkholderia pseudomallei* and related bacteria. *Chemistry & Biology*, 13(7), pp.701–710.

Diggle, S.P., Matthijs, S., et al., 2007. The *Pseudomonas aeruginosa* 4-quinolone signal molecules HHQ and PQS play multifunctional roles in quorum sensing and iron entrapment. *Chemistry & Biology*, 14(1), pp.87–96.

Doyle, S.M., Genest, O. & Wickner, S., 2013. Protein rescue from aggregates by powerful molecular chaperone machines. *Nat Rev Mol Cell Biol*, 14(10), pp.617–629.

Drepper, T. et al., 2007. Reporter proteins for in vivo fluorescence without oxygen. *Nat Biotech*, 25(4), pp.443–445.

Driks, A., Setlow, P. & Setlow, P., 2000. Morphogenesis and Properties of the Bacterial Spore. In *Prokaryotic Development*. American Society of Microbiology, pp. 191–218.

Dubnau, D. et al., 1994. Two-component regulators and genetic competence in *Bacillus subtilis*. *Research in Microbiology*, 145(5), pp.403–411.

Dürre, P., 2007. Biobutanol: An attractive biofuel. *Biotechnology Journal*, 2(12), pp.1525–1534.

Dürre, P. & Hollergschwandner, C., 2004. Initiation of endospore formation in *Clostridium acetobutylicum*. *Anaerobe*, 10(2), pp.69–74.

Eberhard, A. et al., 1981. Structural identification of autoinducer of *Photobacterium fischeri* luciferase. *Biochemistry*, 20(9), pp.2444–2449.

Edwards, A.N., Tamayo, R. & Mcbride, S.M., 2016. A novel regulator controls *Clostridium difficile* sporulation, motility and toxin production. *Molecular Microbiology*, 100(March), pp.954–971.

Ehrlich, P.R. & Holdren, J.P., 1971. Impact of Population Growth. *Science*, 171(3977), pp.1212–1217.

Ehsaan, M. et al., 2016. Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. *Biotechnology for Biofuels*, 9(1), p.4.

Eldar, A., 2011. Social conflict drives the evolutionary divergence of quorum sensing. *Proceedings of the National Academy of Sciences*, 108(33), pp.13635–13640.

Engelberg-Kulka, H., Hazan, R. & Amitai, S., 2005. mazEF: a chromosomal toxin-antitoxin module that triggers programmed cell death in bacteria. *Journal of Cell Science*, 118(19), pp.4327–4332.

Engelmoer, D., 2012. The evolution of natural competence in *Streptococcus pneumoniae*. University of Manchester.

Eswaramoorthy, P. et al., 2010. Single-cell measurement of the levels and distributions of the phosphorelay components in a population of sporulating *Bacillus subtilis* cells.

Microbiology, 156(8), pp.2294–2304.

EU, 2012. Proposal for a Directive of the European Parliament and of the Council: amending Directive 98/70/EC relating to the quality of petrol and diesel fuels and amending Directive 2009/28/EC on the promotion of the use of energy from renewable sources. *Official Journal of the European Union*, 288, pp.1–23.

Ezeji, T.C. et al., 2005. Improving performance of a gas stripping-based recovery system to remove butanol from *Clostridium beijerinckii* fermentation. *Bioprocess and Biosystems Engineering*, 27(3), pp.207–214.

Ezeji, T.C., Qureshi, N. & Blaschek, H.P., 2004. Acetone butanol ethanol (ABE) production from concentrated substrate: reduction in substrate inhibition by fed-batch technique and product inhibition by gas stripping. *Applied Microbiology and Biotechnology*, 63(6), pp.653–658.

Fadeev, A.G. et al., 2000. Fouling of poly[-1-(trimethylsilyl)-1-propyne] membranes in pervaporative recovery of butanol from aqueous solutions and {ABE} fermentation broth. *Journal of Membrane Science*, 173(1), pp.133–144.

Fargione, J. et al., 2008. Land Clearing and the Biofuel Carbon Debt. *Science*, 319(5867), pp.1235–1238.

Flavier, A.B., Ganova-Raeva, L.M., et al., 1997. Hierarchical autoinduction in *Ralstonia solanacearum*: control of acyl-homoserine lactone production by a novel autoregulatory system responsive to 3-hydroxypalmitic acid methyl ester. *Journal of Bacteriology*, 179(22), pp.7089–7097.

Flavier, A.B., Clough, S.J., et al., 1997. Identification of 3-hydroxypalmitic acid methyl ester as a novel autoregulator controlling virulence in *Ralstonia solanacearum. Molecular microbiology*, 26(2), pp.251–259.

Fletcher, K., 2016. Ethanol to n-Butanol in the Land of 10 , 000 Lakes. *Green Biologics Press*

Statement.

Fontaine, L. et al., 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. *Journal of Bacteriology*, 184(3), pp.821–830.

Fontes, C.M.G.A. & Gilbert, H.J., 2010. Cellulosomes: Highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annual Review of Biochemistry*, 79(1), pp.655–681.

Fujita, M., González-Pastor, J.E. & Losick, R., 2005. High- and low-threshold genes in the Spo0A regulon of *Bacillus subtilis. Journal of Bacteriology*, 187(4), pp.1357–1368.

Fuqua, C., Parsek, M.R. & Greenberg, E.P., 2001. Regulation of gene expression by cell -to-cell communication by sensing, acyl-homoserine lactone quorum sensing. *Annu. Rev. Genet*, 35(1), pp.439–68.

Fuqua, W.C., Winans, S.C. & Greenberg, E.P., 1994. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *Journal of Bacteriology*, 176(2), pp.269–275.

Furey, T.S., 2012. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat Rev Genet*, 13(12), pp.840–852.

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

Gallagher, L.A. et al., 2002. Functions required for extracellular quinolone signaling by *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 184(23), pp.6472–6480.

Gallego del Sol, F. & Marina, A., 2013. Structural basis of Rap phosphatase inhibition by Phr peptides. *PLoS biology*, 11(3), p.e1001511.

Ghim, S., Nielsen, P. & Neuhard, J., 1994. Molecular characterization of pyrimidine biosynthesis genes from the thermophile Bacillus caldolyticus. *Microbiology*,

73308(140), pp.479–491.

Girbal, L. et al., 2005. Homologous and heterologous overexpression in *Clostridium acetobutylicum* and characterization of purified Clostridial and Algal Fe-only hydrogenaes with high specific activities. *Applied and Environmental Microbiology*, 71(5), pp.2777–2781.

Girbal, L. et al., 1995. Regulation of metabolic shifts in *Clostridium acetobutylicum* ATCC 824. *FEMS Microbiology Reviews*, 17(3), pp.287–297.

Gomez, L.D., Steele-King, C.G. & McQueen-Mason, S.J., 2008. Sustainable liquid biofuels from biomass: the writing's on the walls. *New Phytologist*, 178(3), pp.473–485.

Grandvalet, C., De Crécy-Lagard, V. & Mazodier, P., 1999. The ClpB ATPase of *Streptomyces albus* G belongs to the HspR heat shock regulon. *Molecular Microbiology*, 31(2), pp.521–532.

Green, E.M. et al., 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. *Microbiology*, 142(8), pp.2079–2086.

Green, E.M. & Bennett, G.N., 1996. Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. *Applied Biochemistry and Biotechnology*, 57(1), pp.213–221.

GreenBiologics, 2016. Green Biologics Begins Customer Shipments at First Commercial Plant. *GBL press statement 13 December 2016*.

Grimmler, C. et al., 2011. Genome-wide gene expression analysis of the switch between acidogenesis and solventogenesis in continuous cultures of *Clostridium acetobutylicum. Journal of Molecular Microbiology and Biotechnology*, 20(1), pp.1–15.

Gupta, R.S. & Gao, B., 2009. Phylogenomic analyses of clostridia and identification of novel protein signatures that are specific to the genus Clostridium sensu stricto (cluster I). *International Journal of Systematic and Evolutionary Microbiology*, 59(2), pp.285–294.

Haas, D., Blumer, C. & Keel, C., 2000. Biocontrol ability of fluorescent pseudomonads genetically dissected: Importance of positive feedback regulation. *Current Opinion in Biotechnology*, 11(3), pp.290–297.

Haas, W., Shepard, B.D. & Gilmore, M.S., 2002. Two-component regulator of *Enterococcus faecalis* cytolysin responds to quorum-sensing autoinduction. *Nature*, 415(6867), pp.84–87.

Hall, M.P. et al., 2012. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chemical Biology*, 7(11), pp.1848–1857.

Harris, L.M., Welker, N.E. & Papoutsakis, E.T., 2002. Northern , Morphological , and Fermentation Analysis of spo0A Inactivation and Overexpression in Clostridium acetobutylicum ATCC 824. *Journal of bacteriology*, 184(13), pp.3586–3597.

Heap, J.T. et al., 2009. A modular system for Clostridium shuttle plasmids. *J Microbiol Methods*, 78(1), pp.79–85.

Heap, J.T. et al., 2012. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic acids research*, 40(8), p.e59.

Heap, J.T. et al., 2014. Spores of Clostridium engineered for clinical efficacy and safety cause regression and cure of tumors in vivo. *Oncotarget*, 5(7), pp.1761–1769.

Heap, J.T. et al., 2007. The ClosTron: a universal gene knock-out system for the genus Clostridium. *Journal of microbiological methods*, 70(3), pp.452–64.

Heap, J.T. et al., 2010. The ClosTron: Mutagenesis in Clostridium refined and streamlined. *Journal of microbiological methods*, 80(1), pp.49–55.

Heng, N.C.K. et al., 2007. The Diversity of Bacteriocins in Gram-Positive Bacteria. *Bacteriocins: Ecology and Evolution*, pp.45–92.

Hentzer, M. & Givskov, M., 2003. Pharmacological inhibition of quorum sensing for the treatment of chronic bacterial infections. *Journal of Clinical Investigation*, 112(9),

pp.1300–1307.

Herrmann, G. et al., 2008. Energy conservation via electron-transferring flavoprotein in anaerobic bacteria. *Journal of Bacteriology*, 190(3), pp.784–791.

Hobson & P. Nasr, H., 1951. An amylopectin-type poly- saccharide synthesized from sucrose by C. butyricuit. *Chem. Soc.*, p.1855–1857.

Horng, Y.-T. et al., 2002. The LuxR family protein SpnR functions as a negative regulator of N-acylhomoserine lactone-dependent quorum sensing in *Serratia marcescens*. *Molecular Microbiology*, 45(6), pp.1655–1671.

Hosni, T. et al., 2011. Sharing of quorum-sensing signals and role of interspecies communities in a bacterial plant disease. *Internation society for microbial ecology*, 5(12), pp.1857–1870.

Huber, H. & Prangishvili, D., 2006. *An introduction to the family Clostridiaceae,*

Huedo, P. et al., 2015. Decoding the genetic and functional diversity of the DSF quorum-sensing system in *Stenotrophomonas maltophilia*. *Frontiers in Microbiology*, 6, p.761.

Hughes, L. & Rudolph, J., 2011. Future world oil production: Growth, plateau, or peak? *Current Opinion in Environmental Sustainability*, 3(4), pp.225–234.

Ismaiel, A.A. et al., 1993. Purification and characterization of a primary-secondary alcohol dehydrogenase from two strains of *Clostridium beijerinckii*. *Journal of Bacteriology*, 175(16), pp.5097–5105.

Jabbari, S. et al., 2013. The putative influence of the agr operon upon survival mechanisms used by *Clostridium acetobutylicum*. *Mathematical biosciences*, 243(2), pp.223–39.

Jeon, E. et al., 2009. Production of cellulosic ethanol in *Saccharomyces cerevisiae* heterologous express-ing *Clostridium thermocellum* endoglucanase and *Saccharomycopsis fibuligera* Beta-glucosidase genes. *Mol. Cells*, 28(4), pp.369–373.

Ji, G., Beavis, R.C. & Novick, R.P., 1995. Cell density control of staphylococcal virulence

mediated by an octapeptide pheromone. *Proceedings of the National Academy of Sciences of the United States of America*, 92(26), pp.12055–9.

Jiricny, N. et al., 2014. Loss of social behaviours in populations of *Pseudomonas aeruginosa* infecting lungs of patients with cystic fibrosis. *PLoS ONE*, 9(1), pp.1–10.

Jones, D.T. et al., 1982. Solvent Production and Morphological Changes in *Clostridium acetobutylicum*. *Applied and Environmental Microbiology*, 43(6), pp.1434–1439.

Jones, D.T. & Woods, D.R., 1986. Acetone-Butanol Fermentation Revisited. *Microbiol. and Mol. Bio Review*, 50(4), p.484.

Jönsson, L.J., Alriksson, B. & Nilvebrant, N.-O., 2013. Bioconversion of lignocellulose: inhibitors and detoxification. *Biotechnology for Biofuels*, 6(1), pp.1–10.

Jung, K. et al., 2012. Histidine kinases and response regulators in networks. *Current Opinion in Microbiology*, 15(2), pp.118–124.

Kalkum, M., Lyon, G.J. & Chait, B.T., 2003. Detection of secreted peptides by using hypothesis-driven multistage mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, 100(5), pp.2795–2800.

Karberg, M. et al., 2001. Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nat Biotech*, 19(12), pp.1162–1167.

Kashket, E.R. & Cao, Z.-Y., 1995. Papers presented at the Meeting on Solventogenic Clostridia, Clostridial strain degeneration. *FEMS Microbiology Reviews*, 17(3), pp.307–315.

Katagiri, H., Takeda, I. & Imai, K., 1958. Synthesis of Riboflavin by Microorganisms. *The Journal of Vitaminology*, 4(2), pp.156–162.

Keis, S. et al., 1995. Taxonomy and phylogeny of industrial solvent-producing clostridia. *International journal of systematic bacteriology*, 45, pp.693–705.

Keis, S., Shaheen, R. & Jones, D.T., 2001. Emended descriptions of *Clostridium*

acetobutylicum and *Clostridium beijerinckii*, and descriptions of Clostridium saccharoperbutylacetonicum sp. nov. and *Clostridium saccharobutylicum* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, 51(6), pp.2095–2103.

Kessel, J.C. Van et al., 2013. Analysis of Activator and Repressor Functions Reveals the Requirements for Transcriptional Control b. , 4(4), pp.1–10.

De Kievit, T.R., 2009. Quorum sensing in *Pseudomonas aeruginosa* biofilms. *Environmental Microbiology*, 11(2), pp.279–288.

de Kievit, T.R. & Iglewski, B.H., 2000. Bacterial Quorum Sensing in Pathogenic Relationships D. A. Portnoy, ed. *Infection and Immunity*, 68(9), pp.4839–4849.

Kirk, D.G. et al., 2012. Involvement of *Clostridium botulinum* ATCC 3502 Sigma Factor K in early-stage sporulation. *Applied and Environmental Microbiology*, 78(13), pp.4590–4596.

Kirke, D.F. et al., 2004. The *Aeromonas hydrophila* LuxR homologue AhyR regulates the N-acyl homoserine lactone synthase, AhyI positively and negatively in a growth phase-dependent manner. *FEMS Microbiology Letters*, 241(1), pp.109–117.

Klein, J.R. et al., 2004. A Conjugation-Based System for Genetic Analysis of Group II Intron Splicing in *Lactococcus lactis*. *Journal of Bacteriology*, 186(7), pp.1991–1998.

Koenig, R.L. et al., 2004. *Staphylococcus aureus* AgrA Binding to the RNAIII-agr Regulatory Region. *Journal of Bacteriology*, 186(22), pp.7549–7555.

Kong, K., Vuong, C. & Otto, M., 2006. *Staphylococcus* quorum sensing in biofilm formation and infection. *Journal of Med. Microbiol.*, 296, p.2451.

Konieczny, I. & Liberek, K., 2002. Cooperative action of *Escherichia coli* ClpB protein and DnaK chaperone in the activation of a replication initiation protein. *Journal of Biological Chemistry*, 277(21), pp.18483–18488.

Kool, J. et al., 2011. Studying protein–protein affinity and immobilized ligand–protein affinity interactions using MS-based methods. *Analytical and Bioanalytical Chemistry*, 401(4), pp.1109–1125.

Köpke, M. et al., 2011. 2,3-butanediol production by acetogenic bacteria, an alternative route to chemical synthesis, using industrial waste gas. *Appl Environ Microbiol*, 77(15), pp.5467–5475.

Kosaka, T. et al., 2007. Characterization of the sol operon in butanol-hyperproducing *Clostridium saccharoperbutylacetonicum* strain N1-4 and its degeneration mechanism. *Bioscience, biotechnology, and biochemistry*, 71(1), pp.58–68.

Kotte, A.-K., 2013. RNPP-type quorum sensing in *Clostridium acetobutylicum*. University of Nottingham.

Kozlowicz, B.K., Bae, T. & Dunny, G.M., 2004. *Enterococcus faecalis* pheromone-responsive protein PrgX: genetic separation of positive autoregulatory functions from those involved in negative regulation of conjugative plasmid transfer. *Molecular Microbiology*, 54(2), pp.520–532.

Krupnick, A. et al., 2011. Understanding the costs and benefits of deepwater oil drilling regulation. *Resources for the Future Discussion Papers*, 10(January), pp.10–62.

Kuehne, S. a & Minton, N.P., 2012. ClosTron-mediated engineering of Clostridium. *Bioengineered*, 3(4), pp.247–54.

Kujawska, A. et al., 2015. ABE fermentation products recovery methods - A review. *Renewable and Sustainable Energy Reviews*, 48(August), pp.648–661.

Lambowitz, A.M. & Zimmerly, S., 2011. Group II introns: Mobile ribozymes that invade DNA. *Cold Spring Harbor Perspectives in Biology*, 3(8), pp.1–19.

Lampe, D.J., Churchill, M.E. & Robertson, H.M., 1996. A purified mariner transposase is sufficient to mediate transposition in vitro. *The EMBO Journal*, 15(19), pp.5470–5479.

Lanckriet, A. et al., 2009. Generation of single-copy transposon insertions in *Clostridium perfringens* by electroporation of phage mu dna transposition complexes. *Applied and Environmental Microbiology*, 75(9), pp.2638–2642.

Laue, B.E. et al., 2000. The biocontrol strain *Pseudomonas fluorescens* F113 produces the Rhizobium small bacteriocin, N-(3-hydroxy-7-cis-tetradecenoyl) homoserine lactone, via HdtS, a putative novel N-acylhomoserine lactone synthase. *Microbiology*, 146(10), pp.2469–2480.

Lawson, P.A. et al., 1993. Towards a phylogeny of the clostridia based on 16S rRNA sequences. *FEMS Microbiology Letters*, 113(1), pp.87–92.

Leoneti, A.B., Aragão-Leoneti, V. & de Oliveira, S.V.W.B., 2012. Glycerol as a by-product of biodiesel production in Brazil: Alternatives for the use of unrefined glycerol. *Renewable Energy*, 45, pp.138–145.

Lépine, F. et al., 2004. Electrospray/mass spectrometric identification and analysis of 4-hydroxy-2-alkylquinolines (HAQs) produced by *Pseudomonas aeruginosa*. *Journal of the American Society for Mass Spectrometry*, 15(6), pp.862–869.

Li, J. et al., 2011. The Agr-like quorum-sensing system regulates sporulation and production of enterotoxin and beta2 toxin by *Clostridium perfringens* type A non-food-borne human gastrointestinal disease strain F5603. *Infect. and Immunol*, 79(6), pp.2451–9.

Li, S.Y. et al., 2011. Performance of batch, fed-batch, and continuous A-B-E fermentation with pH-control. *Bioresource Technology*, 102(5), pp.4241–4250.

Lithgow, J.K. et al., 2000. The regulatory locus cinRI in *Rhizobium leguminosarum* controls a network of quorum-sensing loci. *Molecular Microbiology*, 37(1), pp.81–97.

Liu, D. et al., 2014. Biobutanol production in a *Clostridium acetobutylicum* biofilm reactor integrated with simultaneous product recovery by adsorption. *Biotechnology for Biofuels*, 7, p.5.

Longley, D.B., Harkin, D.P. & Johnston, P.G., 2003. 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer*, 3(5), pp.330–338.

Losick, R. et al., 1992. Crisscross regulation of cell-type-specific gene expression during development in *B. subtilis*. *Letters To Nature*, 355(6361), pp.601–604.

LoVetri, K. & Madhyastha, S., 2010. Antimicrobial and antibiofilm activity of quorum sensing peptides and peptide analogues against oral biofilm bacteria. In A. Giuliani & C. A. Rinaldi, eds. *Antimicrobial Peptides: Methods and Protocols*. Totowa, NJ: Humana Press, pp. 383–392.

Ludwig, W., Schleifer, K.-H. & Whitman, W.B., 2009. Revised road map to the phylum Firmicutes. In P. De Vos et al., eds. *Bergey's Manual® of Systematic Bacteriology: Volume Three The Firmicutes*. New York, NY: Springer New York, pp. 1–13.

Lumjiaktase, P. et al., 2017. Quorum sensing regulates dpsA and the oxidative stress response in *Burkholderia pseudomallei*. , (2006), pp.3651–3659.

Ma, M., Li, J. & McClane, B.A., 2015. Structure-function analysis of peptide signaling in the *Clostridium perfringens* Agr-like quorum sensing system. *Journal of Bacteriology*, 197(10), pp.1807–1818.

Mabee, W.E., McFarlane, P.N. & Saddler, J.N., 2011. Biomass availability for lignocellulosic ethanol production. *Biomass and Bioenergy*, 35(11), pp.4519–4529.

Maddox, I.S. et al., 2000. The cause of "acid-crash" and "acidogenic fermentations" during the batch acetone-butanol-ethanol (ABE-) fermentation process. *Journal of molecular microbiology and biotechnology*, 2(1), pp.95–100.

Makarova, K.S. et al., 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Micro*, 13(11), pp.722–736.

Martin, M.J. et al., 2013. The agr Locus Regulates Virulence and Colonization Genes in *Clostridium difficile* 027. *Journal of Bacteriology*, 195(16), pp.3672–3681.

McGlade, C. & Ekins, P., 2015. The geographical distribution of fossil fuels unused when limiting global warming to 2 [deg]C. *Nature*, 517(7533), pp.187–190.

McKnight, S.L., Iglewski, B.H. & Pesci, E.C., 2000. The *Pseudomon*as quinolone signal regulates rhl quorum sensing in *Pseudomonas aeruginosa*. *Journal of bacteriology*, 182(10), pp.2702–2708.

MDowell, P. et al., 2001. Structure, activity and evolution of the group I thiolactone peptide quorum-sensing system of *Staphylococcus aureus*. *Molecular Microbiology*, 41(2), pp.503–512.

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. *Applied and Environmental Microbiology*, 59(4), pp.1077–1081.

Millat, T. et al., 2014. Coenzyme A-transferase-independent butyrate re-assimilation in *Clostridium acetobutylicum*-evidence from a mathematical model. *Applied Microbiology and Biotechnology*, 98(21), pp.9059–9072.

Mohimani, H. et al., 2011. Sequencing cyclic peptides by multistage mass spectrometry. *Proteomics*, 11(18), pp.3642–3650.

Mohr, A. & Raman, S., 2013. Lessons from first generation biofuels and implications for the sustainability appraisal of second generation biofuels. *Energy Policy*, 63, pp.114–122.

Mohr, S.C. et al., 1991. Binding of small acid-soluble spore proteins from *Bacillus subtilis* changes the conformation of DNA from B to A. *Proceedings of the National Academy of Sciences of the United States of America*, 88(1), pp.77–81.

Münchau, A. & Bhatia, K.P., 2000. Uses of botulinum toxin injection in medicine today. *BMJ : British Medical Journal*, 320(7228), pp.161–165.

Murphy, R. et al., 2011. Global developments in the competition for land from biofuels. *Food*

Policy, 36(SUPPL. 1), pp.S52–S61.

Musto, H., Romero, H. & Zavala, A., 2003. Translational selection is operative for synonymous codon usage in *Clostridium perfringens* and *Clostridium acetobutylicum*. *Microbiology*, 149(4), pp.855–863.

Naik, S.N. et al., 2010. Production of first and second generation biofuels: a comprehensive review. *Renew Sust Energ Rev*, 14(2), pp.578–597.

Nakayama, J. et al., 2001. Gelatinase biosynthesis-activating pheromone: a peptide lactone that mediates a quorum sensing in *Enterococcus faecalis*. *Molecular Microbiology*, 41(1), pp.145–154.

Nations, U., 2015. World populatiion prospects: the 2015 revision. *Department of Economics and Social affairs, Population Division*.

Ng, W.-L. & Bassler, B.L., 2009. Bacterial quorum-sensing network architectures. *Annual review of genetics*, 43, pp.197–222.

Ng, Y.K. et al., 2013. Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using pyrE alleles. *PloS one*, 8(2), p.e56051.

Niven, R.K., 2005. Ethanol in gasoline: environmental impacts and sustainability review article. *Renewable and Sustainable Energy Reviews*, 9(6), pp.535–555.

O'Brien, R. & Morris, J., 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *Journal of General Microbiology*, 68, pp.307–318.

O'Donovan, G. a & Neuhard, J., 1970. Pyrimidine metabolism in microorganisms. *Bacteriological reviews*, 34(3), pp.278–343.

Ohtani, K. et al., 2009. Virulence gene regulation by the agr system in *Clostridium perfringens*. *Journal of bacteriology*, 191(12), pp.3919–27.

Okada, M. et al., 2005. Structure of the *Bacillus subtilis* quorum-sensing peptide pheromone

ComX. *Nat Chem Biol*, 1(1), pp.23–24.

Olson, M.E. et al., 2014. *Staphylococcus epidermidis* agr quorum-sensing system: Signal identification, cross talk, and importance in colonization. *Journal of Bacteriology*, 196(19), pp.3482–3493.

Otto, M. et al., 1999. Inhibition of virulence factor expression in *Staphylococcus aureus* by the *Staphylococcus epidermidis* agr pheromone and derivatives. *FEBS Letters*, 450(3), pp.257–262.

Otto, M., 2004. Quorum-sensing control in *Staphylococci* – a target for antimicrobial drug therapy? *{FEMS} Microbiology Letters*, 241(2), pp.135–141.

Otto, M., Echner, H. & Voelter, W., 2001. Pheromone cross-inhibition between *Staphylococcus aureus* and *Staphylococcus epidermidis Infection and immunity*, 69(3), pp.1957–1960.

Pan, H. et al., 2012. The *clpB* gene is involved in the stress response of *Myxococcus xanthus* during vegetative growth and development. *Microbiology*, 158(Pt 9), pp.2336–2343.

Papoutsakis, E.T., 2008. Engineering solventogenic clostridia. *Current Opinion In Biotech.*, 19(5), p.420.

Patching, S.G., 2014. Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(1, Part A), pp.43–55.

Pauly, M. & Keegstra, K., 2008. Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant Journal*, 54(4), pp.559–568.

Pennington, O.J., 2006. The development of molecular tools for the expression of prodrug converting enzymes in *Clostridium sporogenes*. University of Nottingham.

Perchat, S. et al., 2011. A cell-cell communication system regulates protease production during sporulation in bacteria of the *Bacillus cereus* group. *Molecular microbiology*, 82(3),

pp.619–33.

Pesci, E.C. et al., 1999. Quinolone signaling in the cell-to-cell communication system of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences*, 96(20), pp.11229–11234.

Pfromm, P.H. et al., 2010. Bio-butanol vs. bio-ethanol: A technical and economic assessment for corn and switchgrass fermented by yeast or *Clostridium acetobutylicum*. *Biomass and Bioenergy*, 34(4), pp.515–524.

Piggot, P & Losick, R., 2002. Sporulation genes and intercompartmental regulation. *American society of microbiology*, pp.483–517.

Piggot, P.J. & Hilbert, D.W., 2004. Sporulation of *Bacillus subtilis*. *Current Opinion in Microbiology*, 7(6), pp.579–586.

Pinkston, K.L. et al., 2011. The Fsr quorum-sensing system of *Enterococcus* faecalis modulates surface display of the collagen binding MSCRAMM Ace through regulation of gelE. *Journal of Bacteriology*, 193(17), pp.4317–4325.

Pothiraj, C., Kanmani, P. & Balaji, P., 2006. Potential bioproducts and their applications : Biomass Production of extracellular enzymes by fungi : Extensive. *Mycobiology*, 34(4), pp.159–165.

Pottathil, M. & Lazazzera, B.A., 2003. The extracellular Phr peptide-Rap phosphatase signaling circuit of *Bacillus subtilis*. *Frontiers in Bioscience*, 8(1), pp.d32-45.

Pyne, M. et al., 2015. Antisense-RNA-mediated gene downregulation in *Clostridium pasteurianum*. *Fermentation*, 1(1), pp.113–126.

Pyne, M.E. et al., 2013. Development of an electrotransformation protocol for genetic manipulation of *Clostridium pasteurianum*. *Biotechnology for Biofuels*, 6(1), pp.1–20.

Pyne, M.E. et al., 2016. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. *Nature Publishing Group*, (April),

pp.1–15.

Qin, X. et al., 2001. Characterization of fsr, a regulator controlling expression of gelatinase and serine protease in *Enterococcus faecalis* OG1RF. *Journal of Bacteriology*, 183(11), pp.3372–3382.

Queck, S.Y. et al., 2008. RNAIII-independent target gene control by the agr quorum-sensing system: Insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Molecular Cell*, 32(1), pp.150–158.

Ragsdale, S.W. & Pierce, E., 2008. Acetogenesis and the Wood-Ljungdahl Pathway of CO(2) Fixation. *Biochimica et biophysica acta*, 1784(12), pp.1873–1898.

Ramsey, M.M., Korgaonkar, A.K. & Whiteley, M., 2009. Quorum-Sensing in Bacteria. In *Encyclopedia of Microbiology (Third Edition)*. p. 357.

Ransom, E.M., Ellermeier, C.D. & Weiss, D.S., 2015. Use of mCherry red fluorescent protein for studies of protein localization and gene expression in *Clostridium difficile*. *Applied and Environmental Microbiology*, 81(5), pp.1652–1660.

Ravagnani, a et al., 2000. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. *Molecular microbiology*, 37(5), pp.1172–85.

Redfield, R.J., 2002. Is quorum sensing a side effect of diffusion sensing? *Trends in Microbiology*, 10(8), pp.365–370.

Reinharz, J., 1985. *Chaim Weizmann: the making of a Zionist leader*,

Ren, C. et al., 2012. Pleiotropic functions of catabolite control protein CcpA in Butanol-producing *Clostridium acetobutylicum*. *BMC Genomics*, 13(1), pp.1–20.

Reysenbach, a L. et al., 1986. Characterization, biosynthesis, and regulation of granulose in *Clostridium acetobutylicum*. *Applied and environmental microbiology*, 52(1), pp.185–90.

Robock, A., 2015. *CLIMATE AND CLIMATE CHANGE | Nuclear Winter* Second Edi., Elsevier.

Robson, R.L., Robson, R.M. & Morris, J.G., 1974. The biosynthesis of granulose by *Clostridium pasteurianum*. *Biochemical Journal*, 144(3), pp.503–511.

Rocha-Estrada, J. et al., 2010. The RNPP family of quorum-sensing proteins in Gram-positive bacteria. *Applied microbiology and biotechnology*, 87(3), pp.913–23.

Ryan, R.P. et al., 2015. The DSF family of cell–cell signals: An expanding class of bacterial virulence regulators C. Dehio, ed. *PLoS Pathogens*, 11(7), p.e1004986.

Saibil, H., 2013. Chaperone machines for protein folding, unfolding and disaggregation. *Nat Rev Mol Cell Biol*, 14(10), pp.630–642.

Saini, J.K., Saini, R. & Tewari, L., 2014. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *3 Biotech*, 5(4), pp.337–353.

Samsonoff, W.A., Hashimoto, T. & Conti, S.F., 1971. Appendage Development in *Clostridium bifermentans*. *J Bacteriol*, 106(I), pp.269–275.

Sandoz, K.M., Mitzimberg, S.M. & Schuster, M., 2007. Social cheating in *Pseudomonas aeruginosa* quorum sensing. *Proceedings of the National Academy of Sciences of the United States of America*, 104(40), pp.15876–81.

Sarchami, T. & Rehmann, L., 2015. Optimizing acid hydrolysis of jerusalem artichoke-derived inulin for fermentative butanol production. *BioEnergy Research*, 8(3), pp.1148–1157.

Sauer, M., 2016. Industrial production of acetone and butanol by fermentation—100 years later R. Boden, ed. *FEMS Microbiology Letters*, 363(13), p.fnw134.

Saujet, L. et al., 2011. The key sigma factor of transition phase, sigh, controls sporulation, metabolism, and virulence factor expression in *clostridium difficile. Journal of Bacteriology*, 193(13), pp.3186–3196.

Schardinger, F., 1905. *Bacillus marcerans*, ein Aceton bildender Rottebacillus. *Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg.*, (14), pp.772–781.

Schmer, M.R. et al., 2014. Energy potential and greenhouse gas emissions from bioenergy cropping systems on marginally productive cropland. *PLoS ONE*, 9(3).

Schultz, D., Wolynes, P.G. & Ben, E., 2009. Deciding fate in adverse times: sporulation and competence in *Bacillus subtilis* PNAS-2009. 106-21027-34 (1). , 106(50).

Schwartz, K. et al., 2014. The AgrD N-terminal leader peptide of *Staphylococcus aureus* has cytolytic and amyloidogenic properties F. C. Fang, ed. *Infection and Immunity*, 82(9), pp.3837–3844.

Searchinger, T. et al., 2008. Emissions from Land-Use Change. *Science*, 423(February), pp.1238–1240.

Sebaihia, M. et al., 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat Genet*, 38(7), pp.779–786.

Sewald, N. & Jakubke, H.-D., 2009. Fundamental chemical and structural principles. In *Peptides: Chemistry and Biology*. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 5–61.

Shafiee, S. & Topal, E., 2010. A long-term view of worldwide fossil fuel prices. *Applied Energy*, 87(3), pp.988–1000.

Shafiee, S. & Topal, E., 2009a. When will fossil fuel reserves be diminished? *Energy Policy*, 37(1), pp.181–189.

Shi, K. et al., 2005. Structure of peptide sex pheromone receptor PrgX and PrgX/pheromone complexes and regulation of conjugation in *Enterococcus faecalis*. *Proceedings of the National Academy of Sciences of the United States of America*, 102(51), pp.18596–601.

Shih, C.J. & Lai, M.C., 2007. Analysis of the AAA+ chaperone *clpB* gene and stress-response expression in the halophilic methanogenic archaeon *Methanohalophilus portucalensis*. *Microbiology*, 153(8), pp.2572–2583.

Sigma-Aldrich, 2008. *TargeTron gene knockout system user guide*, St. Louis, Mo.

Siirola, J.J., 2014. Speculations on global energy demand and supply going forward. *Current*

Opinion in Chemical Engineering, 5, pp.96–100.

Singh, N.N. & Lambowitz, A.M., 2001. Interaction of a group {II} intron ribonucleoprotein endonuclease with its DNA target site investigated by DNA footprinting and modification interference. *Journal of Molecular Biology*, 309(2), pp.361–386.

Sorek, R., Lawrence, C.M. & Wiedenheft, B., 2013. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annual Review of Biochemistry*, 82(1), pp.237–266.

Steiner, E. et al., 2011. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. *Molecular microbiology*, 80(3), pp.641–54.

Steiner, E., Scott, J. & Minton, N.P., 2012. An agr quorum sensing system That regulates granulose formation and sporulation in C*lostridium acetobutylicum*. *Appl. Environ. Microbiol*, 78(4), p.1113.

Steinmoen, H., Knutsen, E. & Håvarstein, L.S., 2002. Induction of natural competence in *Streptococcus pneumoniae* triggers lysis and DNA release from a subfraction of the cell population. *Proceedings of the National Academy of Sciences* , 99(11), pp.7681–7686.

Sturme, M.H.J. et al., 2002. Cell to cell communication by autoinducing peptides in gram-positive bacteria. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 81(1–4), pp.233–243.

Sturme, M.H.J. et al., 2007. Making sense of quorum sensing in *lactobacilli*: A special focus on *Lactobacillus plantarum* WCFS1. *Microbiology*, 153(12), pp.3939–3947.

Szulczyk, K.R., 2010. Which is a better transportation fuel–butanol or ethanol? *Int J Energy Environ*, 1(3), pp.501–512.

Takamizawa, A. et al., 2004. High-level expression of clostridial sialidase using a ferredoxin gene promoter-based plasmid. *Protein Expression and Purification*, 36(1), pp.70–75.

Talukdar, P.K. et al., 2015. Updates on the sporulation process in *Clostridium* species.

Research in Microbiology, 166(4), pp.225–235.

Tashiro, Y. et al., 2005. High production of acetone–butanol–ethanol with high cell density culture by cell-recycling and bleeding. *Journal of Biotechnology*, 120(2), pp.197–206.

Thoendel, M., 2012. *Synthesis of the accessory gene regulator autoinducing peptide in staphylococcus aureus*. University of Iowa.

Tracy, B.P., Jones, S.W. & Papoutsakis, E.T., 2011. Inactivation of sig E and sig G in *Clostridium acetobutylicum i*lluminates their roles in clostridial-cell-form biogenesis, granulose synthesis, solventogenesis, and spore morphogenesis. *Journal of Bacteriology*, 193(6), pp.1414–1426.

Tye, Y.Y. et al., 2016. The world availability of non-wood lignocellulosic biomass for the production of cellulosic ethanol and potential pretreatments for the enhancement of enzymatic saccharification. *Renewable and Sustainable Energy Reviews*, 60, pp.155–172.

Underwood, S. et al., 2009. Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *Journal of Bacteriology*, 191(23), pp.7296–7305.

Valero, A. & Valero, A., 2012. What are the clean reserves of fossil fuels? *Resources, Conservation and Recycling*, 68, pp.126–131.

Venkataramanan, K.P. et al., 2013. The Clostridium small RNome that responds to stress: the paradigm and importance of toxic metabolite stress in *C. acetobutylicum*. *BMC genomics*, 14, p.849.

Wagner-Döbler, I. et al., 2005. Discovery of complex mixtures of novel long-chain quorum sensing signals in free-living and host-associated marine alphaproteobacteria. *ChemBioChem*, 6(12), pp.2195–2206.

Walker, D.J.F. et al., 2016. A genetic assay for gene essentiality in *Clostridium*. *Anaerobe*, 42, pp.40–43.

Walter, K.A., Bennett, G.N. & Papoutsakis, E.T., 1992. Molecular characterization of two *Clostridium acetobutylicum* ATCC 824 butanol dehydrogenase isozyme genes. *Journal of Bacteriology*, 174(22), pp.7149–7158.

Wang, H., Smith, M.C.M. & Mullany, P., 2006. The conjugative rransposon Tn5397 has a strong preference for integration into its *Clostridium difficile* target site. *Journal of Bacteriology*, 188(13), pp.4871–4878.

Wang, M. et al., 2015. Quorum sensing and policing of *Pseudomonas aeruginosa* social cheaters. *Proceedings of the National Academy of Sciences of the United States of America*, 112(7), pp.2187–2191.

Wang, Y. et al., 2015. Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9 system. *Journal of Biotechnology*, 200, pp.1–5.

Weast, R.C., 1977. *CRC handbook of Chemistry and Physics*,

Wellington, M., Kabir, M.A. & Rustchenko, E., 2006. 5-fluoro-orotic acid induces chromosome alterations in genetically manipulated strains of *Candida albicans*. *Mycologia*, 98(3), pp.393–398.

West, S.A. et al., 2006. Social evolution theory for microorganisms. *Nat Rev Microbiol*, 4(8), pp.597–607.

Widlund, P.O. & Davis, T.N., 2005. A high-efficiency method to replace essential genes with mutant alleles in yeast. *Yeast (Chichester, England)*, 22(10), pp.769–774.

Wiehn, M. et al., 2014. In situ butanol recovery from *Clostridium acetobutylicum* fermentations by expanded bed adsorption. *Biotechnology Progress*, 30(1), pp.68–78.

Wietzke, M. & Bahl, H., 2012. The redox-sensing protein Rex, a transcriptional regulator of solventogenesis in *Clostridium acetobutylicum*. *Applied Microbiology and Biotechnology*, 96(3), pp.749–761.

Wilder, C.N., Diggle, S.P. & Schuster, M., 2011. Cooperation and cheating in *Pseudomonas*

aeruginosa: the roles of the las, rhl and pqs quorum-sensing systems. *The ISME journal*, 5(8), pp.1332–43.

Williams, P. et al., 2007. Look who's talking : communication and quorum sensing in the bacterial world. *Philos Trans R Soc Lond B Biol Sci.* Jul 29;362(1483):1119-34.

Willson, B.J. et al., 2016. Production of a functional cell wall-anchored minicellulosome by recombinant *Clostridium acetobutylicum* ATCC 824. *Biotechnology for Biofuels*, 9(1), p.109.

Winzer, K., Hardie, K.R. & Williams, P., 2002. Bacterial cell-to-cell communication: sorry, can't talk now — gone to lunch! *Current Opinion in Microbiology*, 5(2), pp.216–222.

Woods, D.R. et al., 1987. Physiological Responses of Bacteroides and Clostridium Strains to Environmental Stress Factors. *Adv Microb Physiol*. Academic Press, pp. 1–64.

Wuster, A. & Babu, M.M., 2008. Conservation and evolutionary dynamics of the agr cell-to-cell communication system across firmicutes. *Journal of Bacteriology*, 190(2), pp.743–746.

Yamaguchi, Y. & Inouye, M., 2009. Chapter 12 mRNA Interferases, Sequence- Specific Endoribonucleases from the Toxin–Antitoxin Systems. In *Molecular Biology of RNA Processing and Decay in Prokaryotes*. Progress in Molecular Biology and Translational Science. Academic Press, pp. 467–500.

Yarwood, J.M. & Schlievert, P.M., 2003. Quorum sensing in *Staphylococcus* infections. *The Journal of Clinical Investigation*, 112(11), pp.1620–1625.

Yates, E.A. et al., 2002. N-acylhomoserine lactones undergo lactonolysis in a pH-, temperature-, and acyl chain length-dependent manner during Growth of Y*ersinia pseudotuberculosis* and *Pseudomonas aeruginosa*. *Infection and Immunity*, 70(10), pp.5635–5646.

Yutin, N. & Galperin, M.Y., 2013. A genomic update on clostridial phylogeny: Gram-negative

spore formers and other misplaced clostridia. *Environmental microbiology*, 15, pp.2631–41.

Zeisler, J., Kleinhappl, M. & Hofbauer, H., 2010. Reliable sampling of impurities in product gas and syngas. *Biomass 2020*, pp.1–12.

Zertuche, L. & Zall, R.R., 1982. A study of producing ethanol from cellulose using *Clostridium thermocellum*. *Biotechnology and Bioengineering*, 24(1), pp.57–68.

Zhang, Q. et al., 2012. Evolution of lanthipeptide synthetases. *Proceedings of the National Academy of Sciences of the United States of America*, 109(45), pp.18361–18366.

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(4), p.e0122411.

Zhong, J., 2003. Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Research*, 31(6), pp.1656–1664.

Zhu, L. et al., 2011. Engineering the robustness of *Clostridium acetobutylicum* by introducing glutathione biosynthetic capability. *Metabolic engineering*, 13(4), pp.426–34.

Zimmerly, S. et al., 1995. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell*, 82(4), pp.545–554.

Zouhir, S. et al., 2013a. Peptide-binding dependent conformational changes regulate the transcriptional activity of the quorum-sensor NprR. *Nucleic Acids Research*, 41(16), pp.7920–7933.

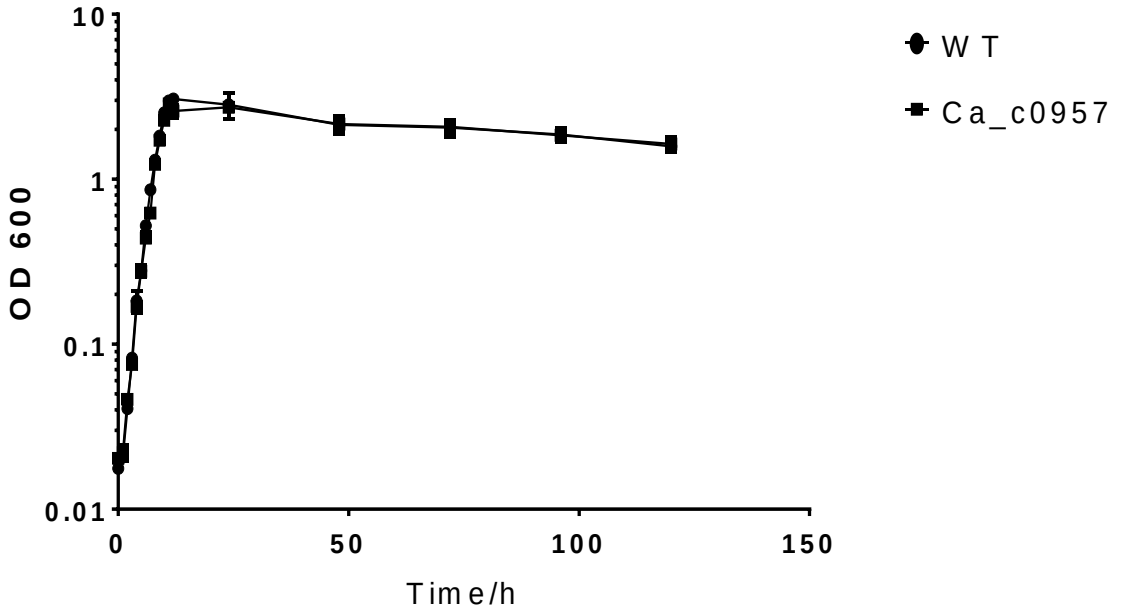
Zouhir, S. et al., 2013b. Peptide-binding dependent conformational changes regulate the transcriptional activity of the quorum-sensor NprR. *Nucleic Acids Research*, 41(16), pp.7920–7933.

Zverlov, V. V. et al., 2006. Bacterial acetone and butanol production by industrial fermentation

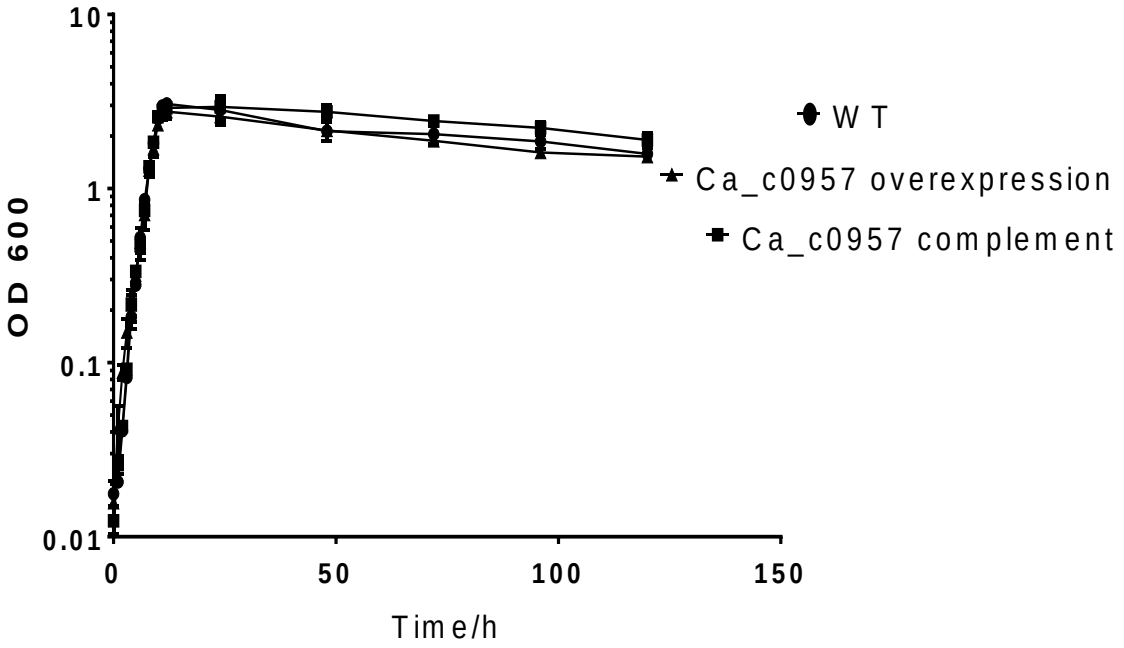
in the Soviet Union: Use of hydrolyzed agricultural waste for biorefinery. *Applied Microbiology and Biotechnology*, 71(5), pp.587–597.

Appendix

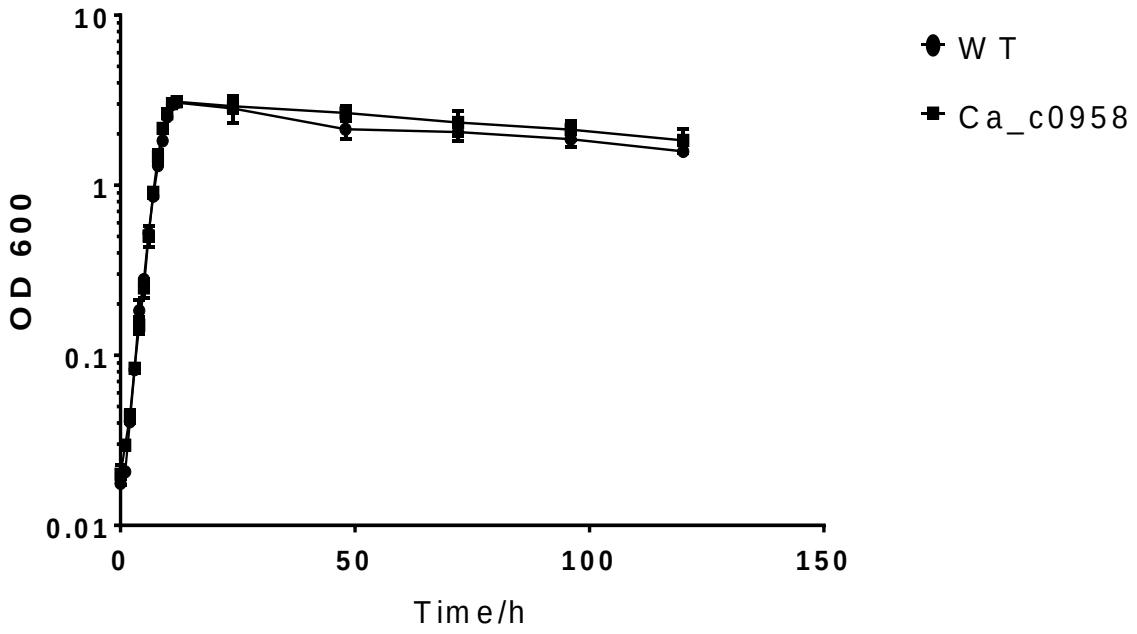
A


10
1
0.1
0.01
OD 600
0
50
100
150
Time/h
WT
Ca_c0957

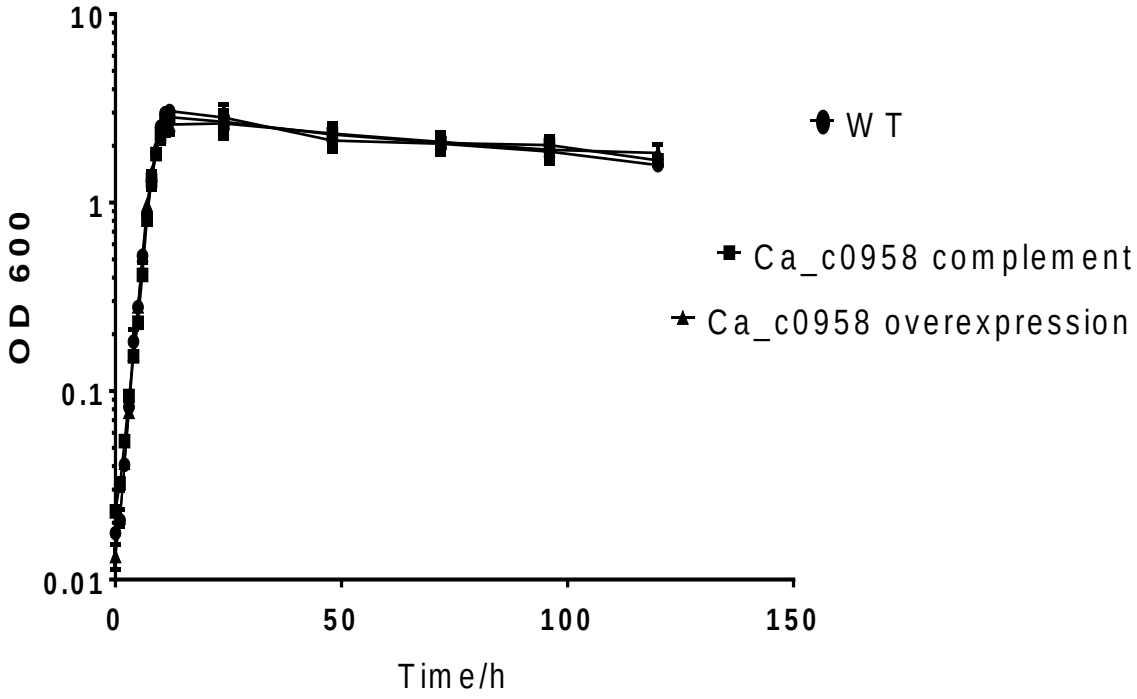
B


10
1
0.1
0.01
OD 600
0
50
100
150
Time/h
WT
Ca_c0957 overexpression
Ca_c0957 complement

C

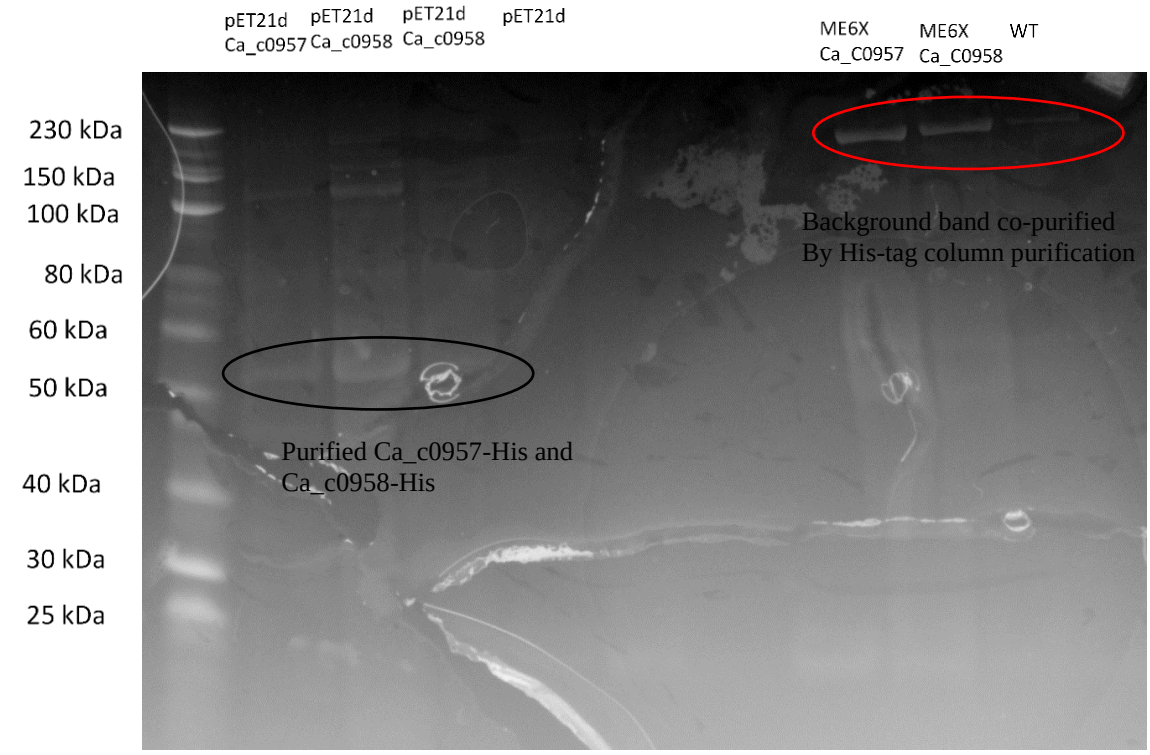

10
1
0.1
0.01
OD 600
0
50
100
150
Time/h
WT
Ca_c0958

D


10
1
0.1
0.01
OD 600
0
50
100
150
Time/h
WT
Ca_c0958 complement
Ca_c0958 overexpression

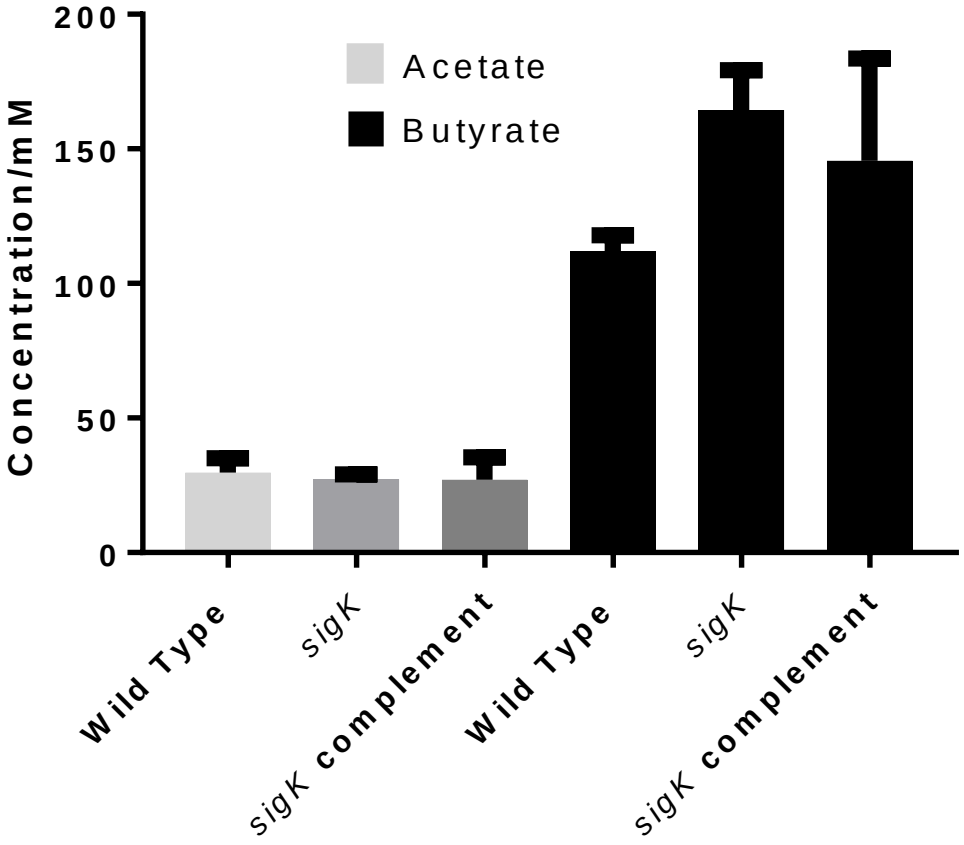
Appendix Figure 1 Growth of Ca_c0957 and Ca_c0958 mutant strains

Growth profiles of (A) a ΔCa_c0957 mutant, (B) Ca_c0957 complementation and overexpression strains (ΔCa_c0957 ME6c Ca_c0957 and ΔCa_c0957 ME6x Ca_c0957 respectively), (C) a ΔCa_c0958 mutant, (D) Ca_c0958 complementation and overexpression strains (ΔCa_c0958 ME6c Ca_c0958 and ΔCa_c0958 ME6x Ca_c0958 respectively) compared to the ATCC 824 wild type. All strains were inoculated to an approximate starting OD of 0.01 and OD_{600} measurements were taken every hour until the cultures appeared to enter stationary phase, after which measurements were made every 24 h. n=9



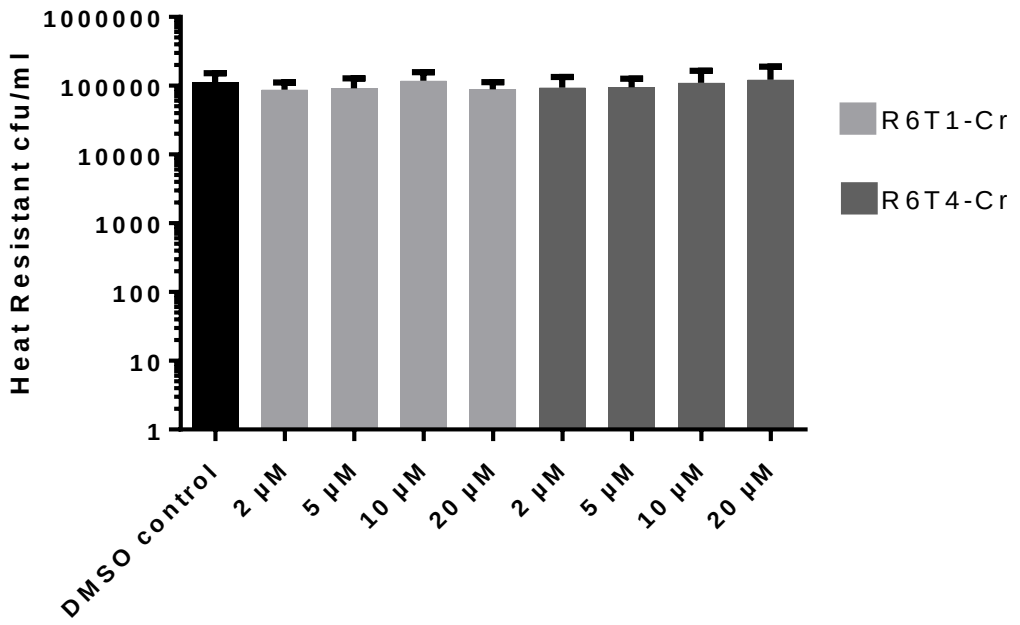
Appendix Figure 2 SDS-PAGE gel of histidine tagged Ca_c0957 and Ca_c0958 protein expressed in *E. coli* and *C. acetobutylicum* following nickel affinity chromatography

Tris SDS PAGE gel stained with Coomassie blue. Lane 1, NEB colour plus 10-230 kDa marker; lane 2, His-tagged Ca_c0957 purified from *E. coli*; lane 3 and 4, His-tagged Ca_c0958 from *E. coli*; lane 5, eluate obtained from the *E. coli* BL21 (DE3) pET21d empty vector control; lane 6, His-tagged Ca_c0957 purified from *C. acetobutylicum* overexpression strain (ΔCa_c0957 ME6X Ca_c0957); lane 7, His-tagged Ca_c0958 purified from *C. acetobutylicum* overexpression strain lane 8 (ΔCa_c0958 ME6X Ca_c0958);, eluate obtained from *C. acetobutylicum wildtype control*. The expected molecular weights were 53.68 kDa and 54.2 kDa for Ca_C0957 and Ca_c0958, respectively.



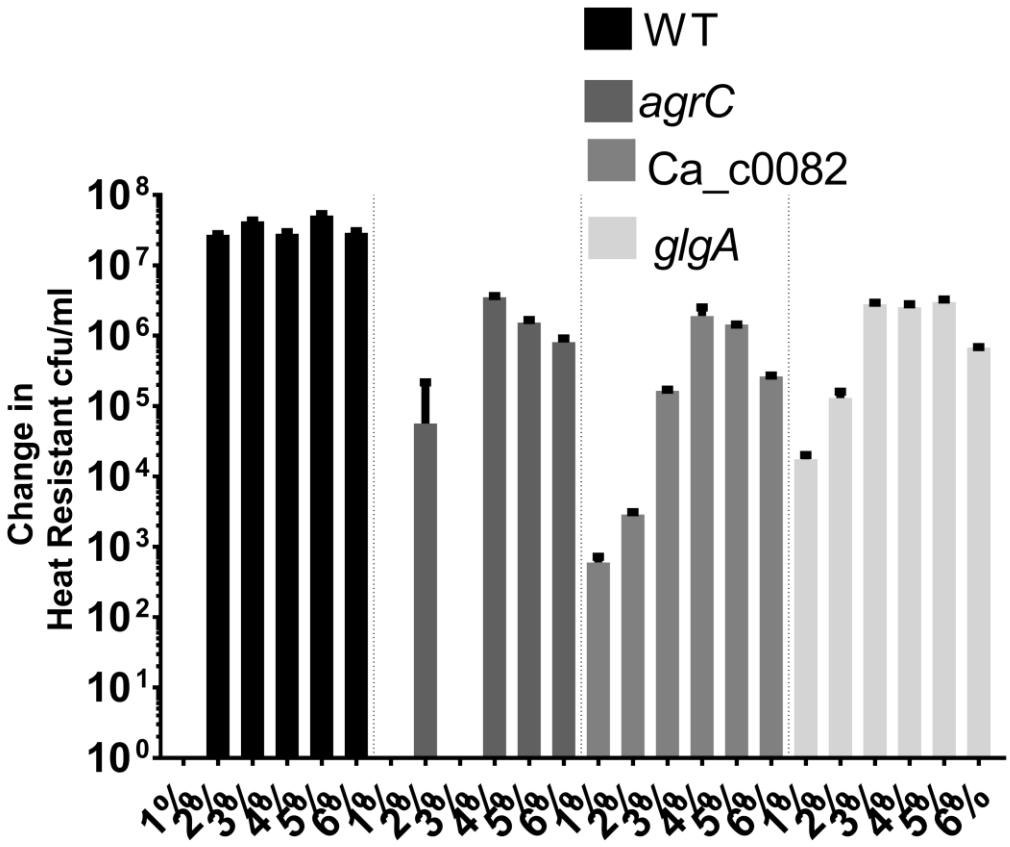
Appendix Figure 3 Acetate and Butyrate production by σ^k mutants

Acetone and Butanol levels of strains inoculated from overnight cultures of CBM-S 6% glucose and grown for 120 h in CMB-S broth 6% (w/v) Included were the ATCC 824 parent strain (WT), the $\Delta\sigma^k$ (*sigK*) mutant and the complemented mutant (*sigK* complement). No significant differences were found from the wild type control in either of the strains. n=9



Appendix Figure 4 Effects of *C. roseum* derived AgrD peptides containing amino acid tails on sporulation of *C. acetobutylicum*

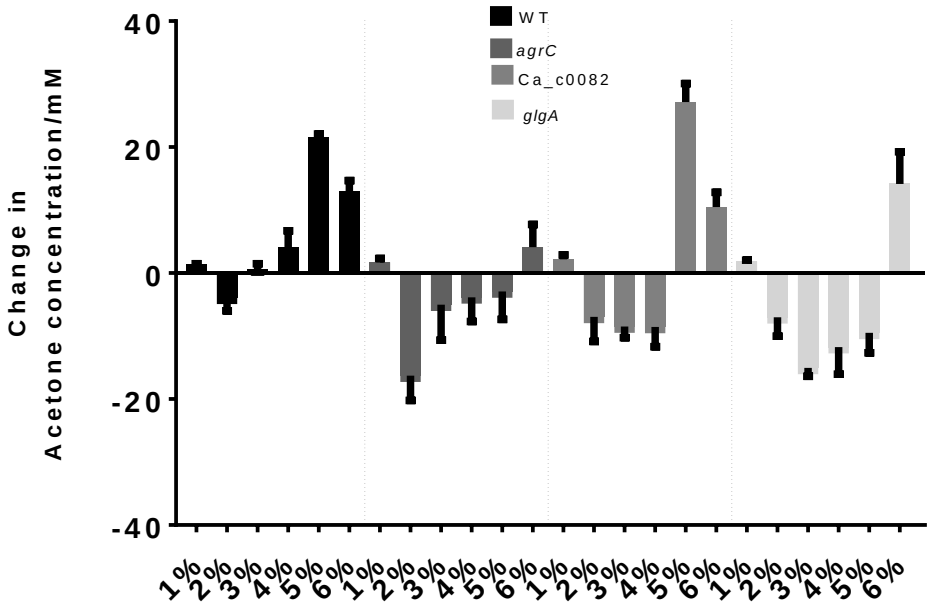
Sporulation of the ATCC wild type was measured as heat resistant cfu/ml after growth in 10 ml CBM-S 6% (w/v) glucose for 120 h in the presence of *C. roseum* derived AgrD peptides containing a 1 and 4-amino acid tails, (R6T1-Cr) and (R6T4-Cr) respectively. Results were compared to a DMSO control by T-test analysis and no statistical significant differences were detected, n=9.



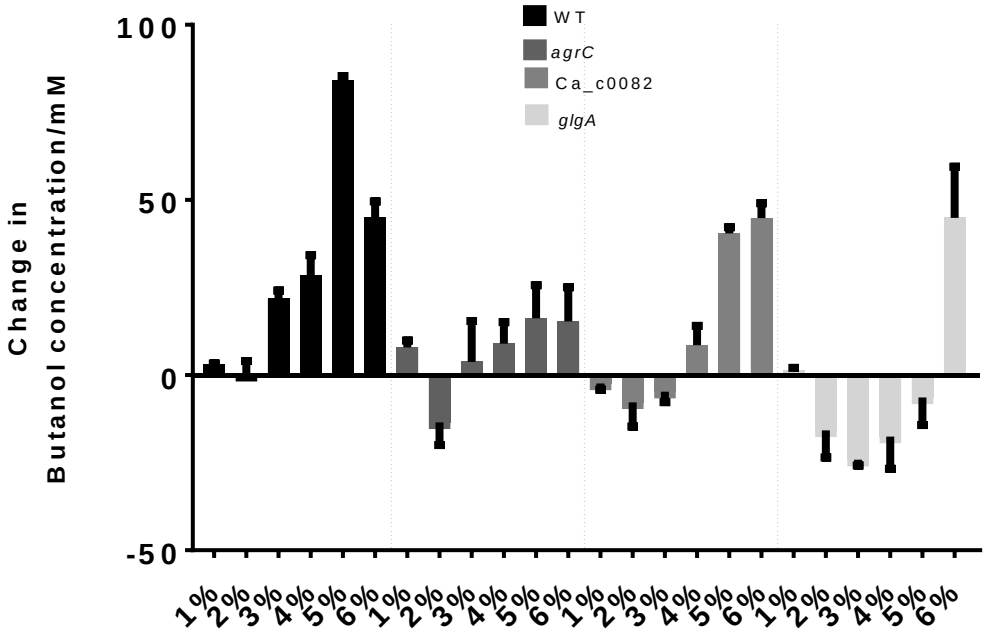
Appendix Figure 5 Change in sporulation of *C. acetobutylicum* wild type, *agrC*::CT*ermB*, Ca_0082::CT*ermB* and Δ*glgA* mutants between 48 h and 120 h at different glucose concentrations

Following growth in 10 ml of CBM-S at the indicated glucose concentrations (1%-6%) samples were taken at 48 h and 120 h time points and analysed for a change in sporulation. n=9

A



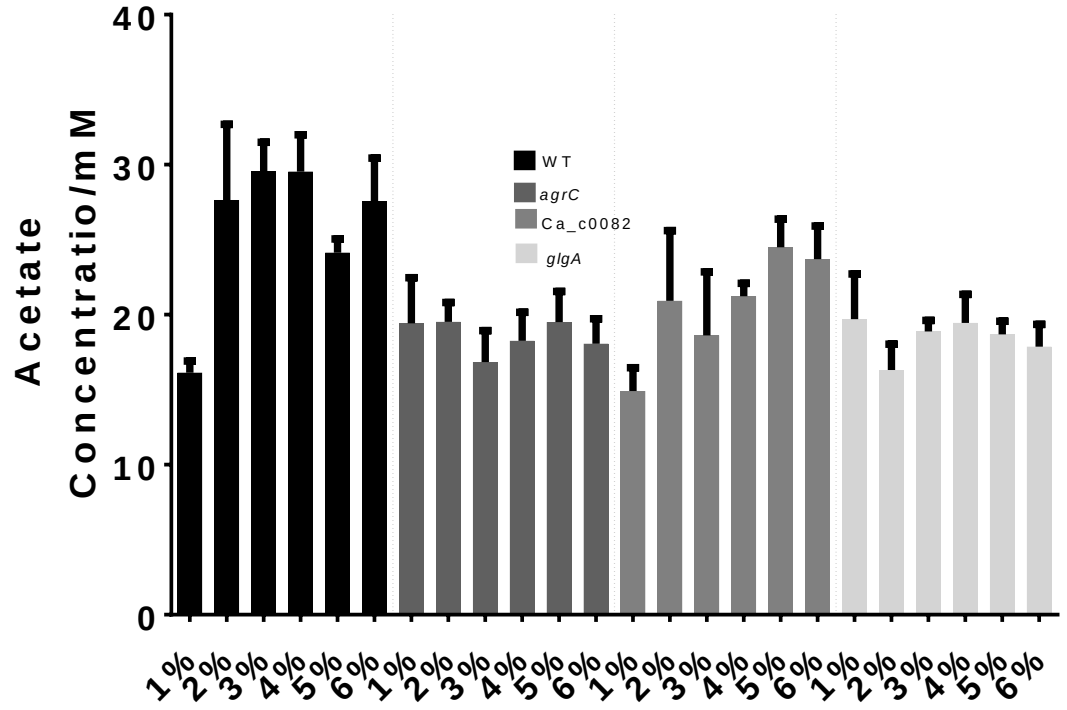
B



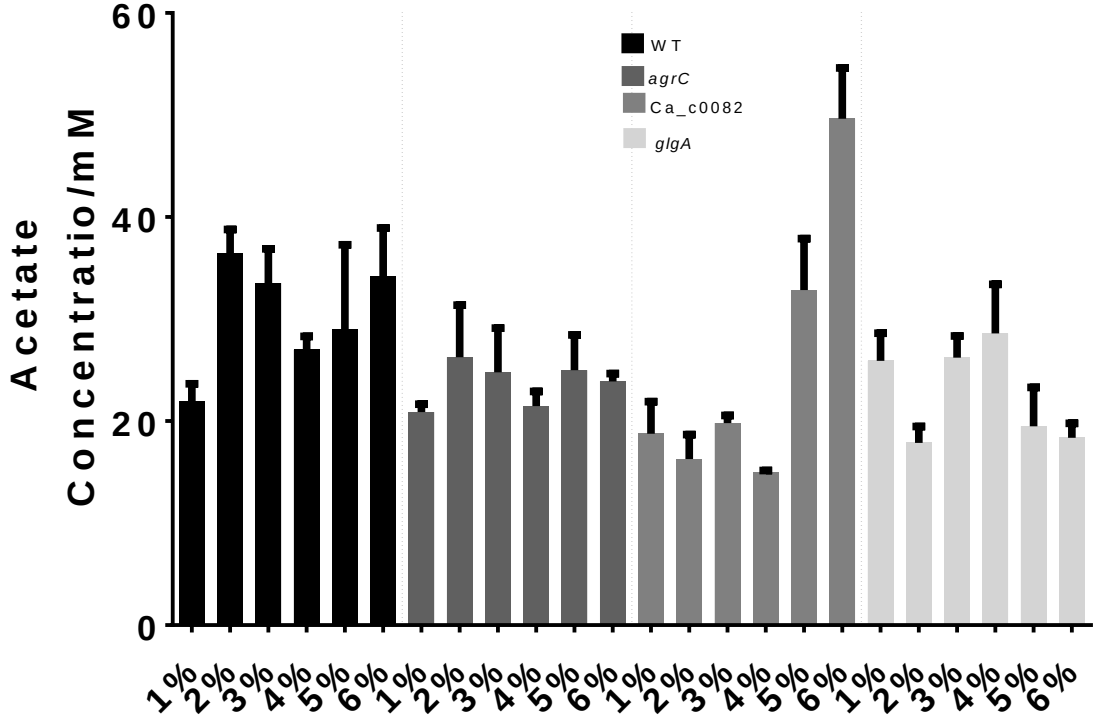
Appendix Figure 6 Change in acetone and butanol concentrations of *C. acetobutylicum* wild type, agrC::CTermB, Ca_0082::CTermB and ΔglgA mutants between 48 h and 120 h at different glucose concentrations

Following growth in 10 ml of CBM-S at the indicated glucose concentrations (1%-6%) samples were taken at 48 h and 120 h time points and analysed for a change in (A) acetone and (B) butanol concentrations. The negative values represent a net evaporative loss of the respective solvent. n=9

A


Acetate Concentratio/mM
40
30
20
10
0
WT
agrC
Ca_c0082
glgA
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

B


Acetate Concentratio/mM
60
40
20
0
WT
agrC
Ca_c0082
glgA
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

Appendix Figure 7 Acetate concentrations of *C. acetobutylicum* wild type, agrC::CTermB, Ca_c0082::CTermB and ΔglgA mutants at different glucose concentrations

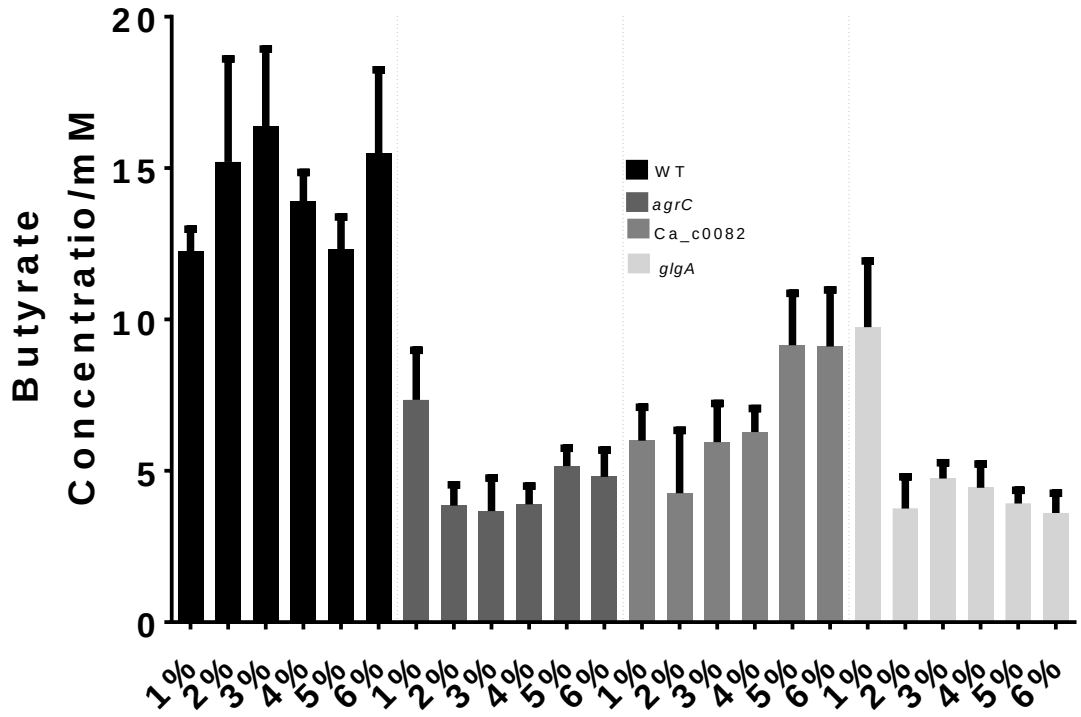
Following growth in 10 ml of CBM-S at the indicated glucose concentrations (1%-6%) samples were taken at 48 h and 120 h time points and acetate concentration determined. Statistical analysis through T-tests summarised in Appendix Table 1.

Appendix Table 1 Statistical significance of acetate concentrations of *C. acetobutylicum* wild type, agrC::CT*ermB*, Ca_c0082::CT*ermB* and Δ*glgA* at different glucose concentrations

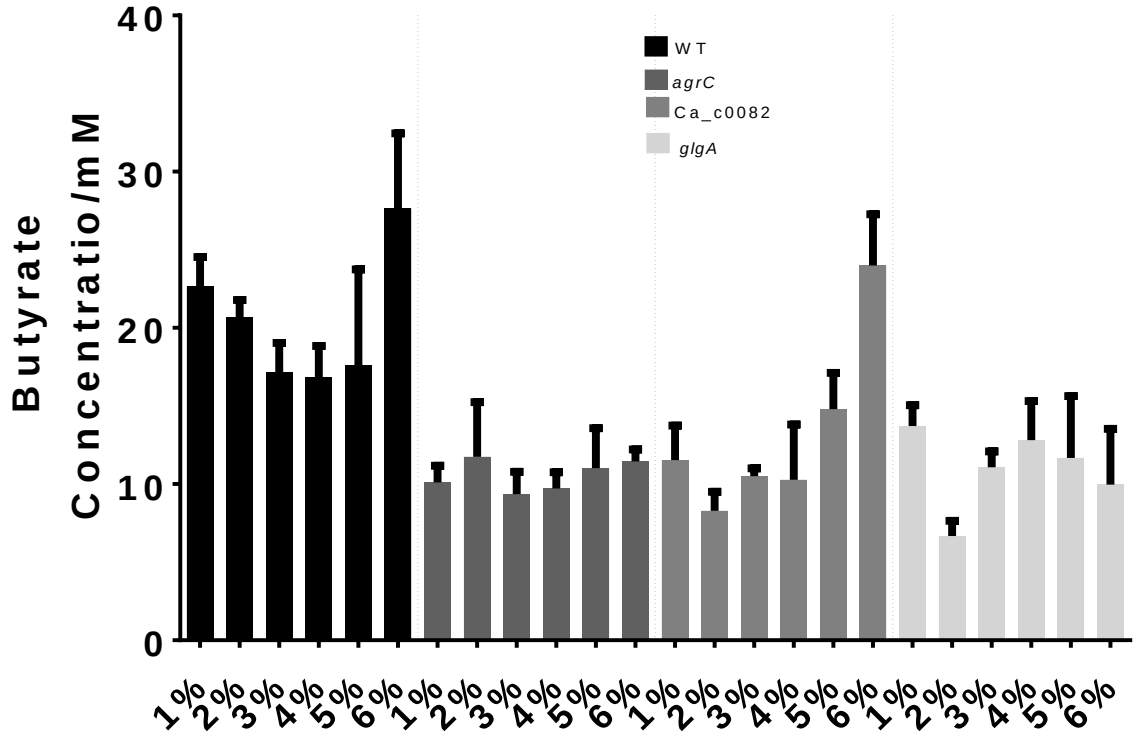
	48h						120h					
	1%	2%	3%	4%	5%	6%	1%	2%	3%	4%	5%	6%
Δ*glgA* vs												
WT	ns	ns	ns	ns	ns	ns	ns	**	ns	ns	ns	*
***agrC* vs**:												
WT	ns	ns	**	ns	ns	*	ns	ns	ns	*	ns	*
Δ*glgA*	ns	ns	ns	ns	ns	Ns	ns	ns	ns	ns	ns	**
Ca_c0082 vs:												
WT	ns	ns	**	*	**	*	ns	**	**	***	ns	*
Δ*glgA*	ns	ns	ns	ns	*	*	ns	ns	*	*	ns	*
agrC	ns	ns	ns	ns	ns	ns	ns	ns	ns	**	ns	*

Statistical analysis through T-tests with P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.000, where n=9

A


Butyrate Concentratio/mM
20
15
10
5
0
WT
agrC
Ca_c0082
glgA
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

B


Butyrate Concentratio/mM
40
30
20
10
0
WT
agrC
Ca_c0082
glgA
1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6% 1% 2% 3% 4% 5% 6%

Appendix Figure 8 Butyrate concentrations of *C. acetobutylicum* wild type, agrC::CTermB, Ca_c0082::CTermB and ΔglgA mutants at different glucose concentrations

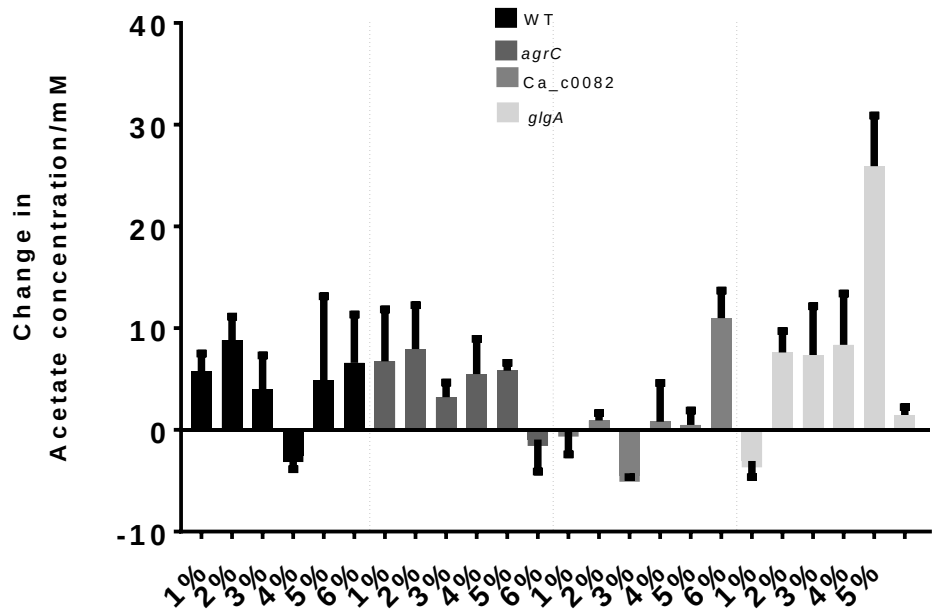
Following growth in 10 ml of CBM-S at the indicated glucose concentrations (1%-6%) samples were taken at 48 h and 120 h time points and butyrate concentration determined. Statistical analysis through T-tests summarised in Appendix Table 2.

Appendix Table 2 Statistical significance of butyrate concentrations of *C. acetobutylicum* wild type, agrC::CT*ermB*, Ca_c0082::CT*ermB* and Δ*glgA* at different glucose concentrations

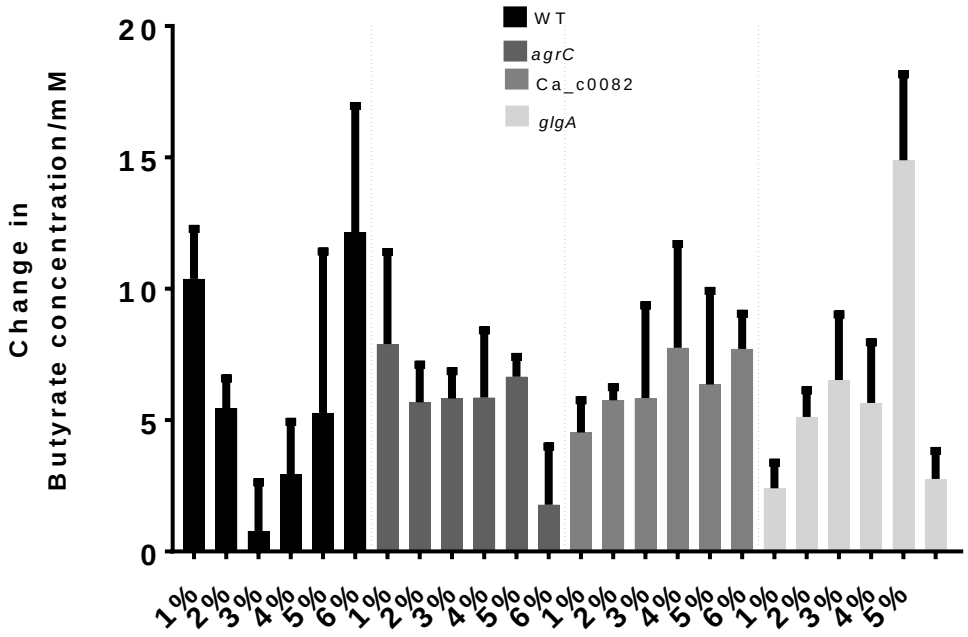
	48h						120h					
	1%	2%	3%	4%	5%	6%	1%	2%	3%	4%	5%	6%
Δ*glgA* vs												
WT	**	*	*	*	ns	ns	**	***	*	ns	ns	ns
***agrC* vs**:												
WT	*	*	**	***	**	*	**	*	*	*	ns	*
Δ*glgA*	ns	ns	ns	*	*	ns						
Ca_c0082 vs:												
WT	ns	*	**	***	***	**	**	***	*	ns	ns	*
Δ*glgA*	ns	ns	ns	ns	*	*	ns	ns	ns	ns	ns	*
agrC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Statistical analysis through T-tests with P values of: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.000, where n=9

A



B



Appendix Figure 9 Change in acetate and butyrate concentrations of *C. acetobutylicum* wild type, agrC::CTermB, Ca_0082::CTermB and ΔglgA mutants between 48 h and 120 h at different glucose concentrations

Following growth in 10 ml of CBM-S at the indicated glucose concentrations (1%-6%) samples were taken at 48 h and 120 h time points and analysed for the change in (A) acetate and (B) butyrate concentrations. The negative values represent a net evaporative loss of the respective solvent. n=9