Characterisation of Agr Quorum Sensing in *Clostridium autoethanogenum*

Paweł Piątek, M.Sc.

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

September 2017

Try to learn something about everything and everything about something

- *Thomas Henry Huxley*

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.

Paweł Piątek

September 2017

Abstract

The Gram-positive, anaerobic, acetogenic bacterium *Clostridium autoethanogenum* is regarded as an important biocatalyst in the current advancement of industrial gas fermentation. The ever-widening diversity of industrially-relevant acetogenic bacteria has inspired a rational approach into utilising industrial waste gases as a viable feedstock, with goals of mitigating greenhouse gas emissions, and supporting an alternative means of fuel and high-value chemical production. Combined with this effort, is the expanding repertoire of gene editing tools that have allowed for the improvement of gas fermentation processes and increased spectrum of fermentative products. Despite these advances, there remain many pertinent questions, which addressed, can further the understanding of metabolism and physiology in acetogenic bacteria. This includes cell-to-cell communication and signalling, *Quorum Sensing*.

In this project, these questions are addressed through the study of the Agr QS system in *C. autoethanogenum*. Signalling peptide genes, *agrD1* and *agrD2* were disabled separately and more importantly, in tandem, which effectively abolished Agr signalling. Phenotypic characterisation of the double *agrD* mutants revealed a significant increase in ethanol at the expense of acetate output. Further observations exhibited a complete utilisation of the fructose carbon source, and the inability to fully re-assimilate CO_2. These findings markedly contrasted with the wild type, and both single knock-out *agrD* mutants. Proteomics and enzyme activity analysis of the double AgrD mutant revealed a marked down-regulation of Wood-Ljungdahl pathway genes that included the CO_2-assimilating, carbon monoxide dehydrogenase / acetyl-CoA synthase complex subunits and hydrogenases. An up-regulation of alcohol dehydrogenases was observed explaining ethanol increases, alongside an unexpected upregulation of bacterial micro-compartment clusters.

These findings led to the hypothesis that the *C. autoethanogenum* Agr system influences the ancient Wood-Ljungdahl pathway, primarily as a means of survival by managing carbon-source utilisation and regulation.

Publications

"Whole genome sequence and manual annotation of Clostridium autoethanogenum, an industrially relevant bacterium"

Humphreys, C.M., McLean, S., Schatschneider, S., Millat, T., Henstra, A.M., Annan, F.J., Breitkopf, R., Pander, B., **Piatek, P**., Rowe, P. and Wichlacz, A.T., 2015. *BMC genomics*, *16*(1), p.1085.

Acknowledgements

During my PhD, I never faced a dull moment. This was chiefly due to the tireless help, and guidance of my supervisor, Dr Klaus Winzer, who took me on as a student and changed me into a scientist. I am deeply grateful for his invaluable contributing discussions and ideas in this project. I would also like to thank Prof. Nigel Minton for the opportunity of working on the GASCHEM project, and for his leadership and ambition in developing a successful, world-class research centre.

I would like to extend my appreciation towards the ever-expanding SBRC research group. In particular, Chris Humphreys, who continuously provided me with excellent bench supervision and all-round laboratory knowledge. I am very grateful for the encouragement and advice given by Jon Baker and Pippa Strong, and to my Dutch lab neighbour Bart Pander, whose book-smarts and silliness kept me going through the most trying of times. I would also like to thank Ollie, Zak, Jon, Alex and Dave, who brought good lab advice, laughter, pub antics and pedantic arguing. I would like to thank Tom and Ola (and Maja) Bailey – who became an extended family to me during these 4 years, and showed me irreplaceable friendship, kindness and generosity. Other notable mentions in no particular order; Matt Abbott (five-star technician), Emily Harding, Sophie Vaud, Bashir Rumah, Chris Stead, the KRG massive, James Millard, Rupert Norman, Dr Jamie Gilbert, Dr Gareth Little (thank you for laptop tech support), Dr Alan Burbidge, Louise Dynes, Terry Bilverstone, Dr Kai Schuchmann, Dr Eric Liew, Erik Hildebrandt, and many more that I may have forgotten. I would also like to also extend my gratitude towards the CBS support staff, in particular; Jacky Webb, Evangelos Karvonis, Kate Beaumont and Tullia Randall.

For facilitating my numerous memorable social experiences during my 4 years, I would like to thank the Johnson Arms, Victoria Hotel, The Organ Grinder, Tipoo, Star Inn, Bodega and several other fine Nottingham establishments.

This work was made possible by the BBSRC Strategic Longer and Larger grant (sLoLa) and funding from our industrial partners, LanzaTech USA.

Table of Contents

List of Figures

List of Tables

Abbreviations

1,2-PD	1,2-propanediol
2,3-BD	2,3-butanediol
5-FOA	5-Fluoroorotic acid
5-FU	5-Fluorouracil
ABE	Acetone butanol ethanol
ACE	Allele coupled exchange
ADP	Adenosine diphosphate
Agr	Accessory gene regulator
AHL	Acylated homoserine lactones
AIP	Autoinducing peptides
AMP	Adenosine monophosphate
AOR	Aldehyde:ferredoxin oxidoreductase
ATCC	American type culture collection
ATP	Adenosine triphosphate
BLAST	The Basic Local Alignment Search Tool
BMC	Bacterial micro-compartment
Bp	Base pair
Cas	CRISPR associated proteins
CAT	Chloramphenicol acetyltransferase
CBM	Clostridial Basal Medium
cDNA	Complementary DNA
CFU	Colony forming unit
CGM	Clostridial growth medium
CoA	Coenzyme A
CODH/ACS	Carbon monoxide dehydrogenase / Acetyl-CoA Synthase
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
EDTA	Ethylenediaminetetraacetic acid
FOA	5-fluoroorotic acid
FT	Fischer-Tropsch
FLP	FNR-like protein
FRT	FLP recognition target
GC	Gas chromatography
gDNA	Genomic DNA
GHA	Great homology arm
GHG	Greenhouse gases
HPLC	Higher performance liquid chromatography
HSL	Homoserine lactone
HTH	Helix-turn-helix
IFD	In-frame deletion
InDels	Insertion/deletions

IPTG	Isopropyl β-D-1-thiogalactopyranoside
kPa	Kilopascal
LB	Lysogeny broth
LHA	Left homology arm
mg	Milligram
ml	Millilitre
mm	Millimetre
mM	Millimolar
NADH	Nicotinamide adenine dinucleotide
NCBI	National Centre for Biotechnology Information
NGS	Next generation sequencing
nm	Nanometres
OD	Optical density
ORF	Open reading frame
PBS	Phosphate buffered saline solution
PCR	Polymerase chain reaction
PFOR	Pyruvate:ferredoxin-oxidoreductase
PlcR	Phospholipase C regulator
RAM	Retro-transposition activated marker
Rap	Response regulator aspartate
RHA	Right homology arm
R-M	Restriction methylation systems
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RNPP	Rap, NprR, PlcR, and PrgX
RPM	Rotations per minute
rRNA	Ribosomal ribonucleic acid
SDS-PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
SHA	Small homology arm
SEM	Scanning electron microscopy
SNP / SNV	Single nucleotide polymorphism / variant
SOE	Splicing by overlap extension
TAE	Tris acetate EDTA
TCS	Two-component systems
UV	Ultra violet
V	Volt
YTF	Yeast tryptone fructose
WL	Wood-Ljungdahl
µg	Microgram
µL	Microliter

Chapter 1

Introduction

1.1 Concerns for energy and environment

Increasing global energy demand has been the consequence of growing human activity in the last two centuries, to which it has accelerated in the last 50 years (Solomon *et al.*, 2009; Miller and Sorrell, 2013). This concern has been widely regarded as the main contributing factor to two pressing worldwide problems, climate change, and fossil fuel over-reliance, and opinions argue that the latter issue is largely irreversible (Solomon *et al.*, 2009; Le Quéré *et al.*, 2016).

Before the industrial era, a naturally kept atmospheric CO_2 balance has been maintained primarily photosynthetically through land and ocean life that act as nature's carbon assimilatory sinks in off-setting natural CO_2 emissions (Geider *et al.*, 2001; Hoegh-Guldberg *et al.*, 2007). Excess CO_2 emission has been linked with markedly increased industrial activity intrinsic to a growing global population. As a result, the natural balance has been perturbed by an overabundance of atmospheric CO_2 amongst other greenhouse gases (GHGs) (Karl, 2003). The problem is further compounded by the reduction in natural carbon sinks, which include; deforestation, and the acidification and warming of oceans concomitantly caused by atmospheric carbon oversaturation, (Geider *et al.*, 2001; Hoegh-Guldberg *et al.*, 2007; Le Quéré *et al.*, 2016). Direct anthropogenic activities contributing to this problem include, and are not limited to; fossil fuel usage, industry, transport, commercial and residential activities (Ribeiro *et al.*, 2007; Solomon *et al.*, 2007; Le Quéré *et al.*, 2016).

Therefore, there is a need to further expand development into alternative energy sources while simultaneously mitigating global climate change. These two independent issues can possibly be tackled through a single novel approach relating to anaerobic microbial gas fermentation. Chemicals derived from microbial fermentation are not a new concept and have been established for thousands of years in the form of brewing and food

fermentation (McGovern *et al.*, 2004), but recent advances have explored into harnessing anaerobic bacteria into producing platform chemicals by utilising industrial waste gases as a feedstock (Henstra *et al.*, 2007).

1.2 Towards biofuel development

1.2.1 Overview

The popularity of biofuel has fluctuated over the last half century with accordance to political, economic and environmental factors (Albers, Berklund and Graff, 2016). Moreover, occasional unstable oil prices have also contributed to the interest in biofuels. An example of this was the energy crisis during the 1970s that saw global petroleum shortages primarily due to political instability in the Middle East. The crisis led to global financial and energy turbulence causing oil prices to surge to 121 USD a barrel (with adjustments to inflation) (Bardi, 2009; Brown and Brown, 2014). The resulting panic lead to measures in safeguarding against future occurrences. The crisis was the instigator that led many Western countries into developing renewable energy technology, which meant alternatives to fossil fuels. Alongside this, environmental efforts to specifically curb GHGs and CO_2 emissions have been mandated in policies such as the European Union's Emissions Trading System, which encourages research and development of fossil fuel-replacing technologies (Ellerman and Buchner, 2007; Brown and Brown, 2014; Albers, Berklund and Graff, 2016).

Biofuel has undergone several generational changes, with improving technologies offering new opportunities and approaches. The following section discusses these changes, appropriately considered as first, second and third generation biofuels.

1.2.2 Early efforts and ABE fermentation

The 1970s energy crisis led to a large-scale biofuel production taking place in Brazil where sugarcane feedstocks were employed for ethanol production (Brown and Brown, 2014; Albers, Berklund and Graff, 2016). However, widespread biofuel usage began much earlier in the 20th century (Dürre, 2007), with feedstocks including agricultural plant matter such as starch and vegetable oils which have been refined into bioethanol, biodiesel and other chemicals (Dürre, 2007; Naik *et al.*, 2010).

A historic example of biofuel manufacture at the start of the 20th century, involved acetone-butanol-ethanol (ABE) fermentation. It was promoted by an Israeli-British biochemist, Chaim Weizmann, whose microbially-derived process involved the use of *Clostridium acetobutylicum* (Jones and Woods, 1986; Dürre, 2007). This technique was associated with the pursuit of an alternative means of rubber production. Moreover, during the outbreak of World War 1, acetone, which was initially disregarded as an ABE by-product was found to be much needed in the production of cordite used in ammunitions and ABE fermentation provided a steady supply (Dürre, 2007).

After both World Wars, butanol and acetone's production was repurposed into automobile lacquer and petroleum refinement (Jones and Woods, 1986). The ABE process remained viable in the first half of the 20th century until it was realized that acetone, butanol and ethanol could be produced more cheaply from petroleum. Furthermore, ABE fermentation required cost-inefficient feedstocks such as molasses, maize and other food crops that compete directly with human food sources, as well as the usage of arable land (Jones and Woods, 1986; Dragone *et al.*, 2010; Mohr and Raman, 2013; Brown and Brown, 2014; Albers, Berklund and Graff, 2016). Consequently, ABE fermentation gradually lost favour in the chemical industry.

Feedstock sources therefore needed to avoid the use of food crops, which lead to the advancement of second-generation biofuels. This shift favoured lignocellulose-rich waste, which included sugar cane bagasse, forestry residue, corn cob residue, municipal solid waste and straw (Naik *et al.*, 2010; Sims *et al.*, 2010). These lignocellulosic feedstocks require costly processing through biomechanical and thermochemical means, often using pre-treatment via acid and enzyme hydrolysation into simpler sugars from which ABE fermentations can take place (Qureshi *et al.*, 2010).

As with second generation biofuels, efforts to maintain ABE fermentation's relevance were governed by several important factors. Firstly, the low cost of feedstocks is essential for fermentation products to remain economically viable as ABE products fluctuate with oil price, there may be instances where this is not possible (Green, 2011). Secondly, microbial fermentation is either naturally bottlenecked by the organism's metabolism or because of external feedstock pre-treatment methods (Qureshi *et al.*, 2010; Xiao *et al.*, 2011; Mohr and Raman, 2013; Raut *et al.*, 2016). Thirdly, it has also been noted that increasing the production of biofuels and biochemicals would lower the price of fossil fuel-derived chemicals and reduce ABE's economic competitiveness. Moreover, this

would drive the consumption of fossil fuels, negating goals of reducing GHG emissions (Smeets *et al.*, 2014).

However, within the last 40 years there has been a steady renewed interest in ABE fermentation encouraged by advances addressing *C. acetobutylicum* fermentation challenges (Qureshi *et al.*, 2010; Brown and Brown, 2014). These challenges refer to external and internal bottlenecks faced by solventogenic bacteria, and genetic engineering has proven that these issues can be overcome (Papoutsakis, 2008). Examples in *C. acetobutylicum* include engineering growth on complex cellulosic substrates improving sugar transport and increasing tolerance to fermentative products (Gu *et al.*, 2009; Zheng *et al.*, 2009; Xiao *et al.*, 2011; Mitchell, 2016). Moreover, understanding of sporulation survival mechanisms linked with solventogenesis has led to improvements in product output (Papoutsakis, 2008; Zheng *et al.*, 2009).

1.2.3 Alternative feedstock strategies

Novel approaches in addressing reliable feedstock sources have led to the ambitious effort of exploiting photosynthetic microalgae, whose product repertoire is not limited to biodiesel, biohydrogen, and methane (Dragone *et al.*, 2010; Behera *et al.*, 2014; Albers, Berklund and Graff, 2016). Taking advantage of microalgae's fast-growth, innate CO_2 fixation and phototrophic metabolism, allows for a reliable feedstock source that avoids contention with agricultural crops as these processes do not require arable land and can even take place in a marine environment (Demirbas, 2007; Wei, Quarterman and Jin, 2013). As with solid plant biomass, microalgae can also be converted into synthesis gas, and be processed further similarly to second generation biofuels (Behera *et al.*, 2014). But with these benefits come significant costs in a practical sense; cultivation requires specialised methods that are expensive, require high energy input and a need for large-scale systems to be geographically situated where there is ample sunlight and suitable temperature to be economically viable (Leite, Abdelaziz and Hallenbeck, 2013; Behera *et al.*, 2014).

This issue has led to so-called "fourth generation" biofuel production with the introduction of genetic modification of microalgae into sustainable cell factories (Lu, Sheahan and Fu, 2011). Published works detail engineering of cyanobacteria into increasing biohydrogen, biobutanol and ethanol production, and furthermore there are advances into increasing the repertoire of products such as isoprenoids and terpenoids,

(Desai and Papoutsakis, 1999; Lindblad *et al.*, 2012; Gimpel *et al.*, 2013; Englund *et al.*, 2015). These products go beyond the scope of biofuels and into areas of pharmaceutical pre-cursors, plastics and consumer products (Harun *et al.*, 2010; Wu and Davis, 2016).

1.2.4 Synthesis gas

From this overview, it can be stated that biofuel processes aiming to produce large quantities of sustainable solvents are hindered by a multitude of factors. Primarily the cost of feedstock treatment, product processing and refinement, and the feedstock itself (Dürre, 2017). Moreover, aims to reduce GHG emissions and to avoid using food crops as feedstock have further created difficulties in developing sustainable systems. Approaches with microalgal systems bear the advantage of steady CO_2 fixation coupled with photosynthesis but suffer from underdeveloped and inefficient means of processing. Advances in anaerobic ABE fermentation have shown high productivity potential but are tied to the inflexibility of feedstock sources (Green, 2011).

As an alternative feedstock, synthesis gas (syngas) could hold a possible answer. Syngas can be derived from a vast selection of sources, where high temperature pyrolytic gasification of fossil fuels, plant biomass and municipal solid waste, produces CO, CH_4, N_2, CO_2 and H_2 waste-gases (Henstra *et al.*, 2007; Debabov, 2013). These gases in turn can be directly burned as a fuel, or utilised as a feedstock in chemoautotrophic microbial fermentation (Ragsdale and Pierce, 2008; Naik *et al.*, 2010; Sims *et al.*, 2010; Brown and Brown, 2014; Albers, Berklund and Graff, 2016).

1.3 Microbial gas fermentation

1.3.1 Applications of syngas fermentation

Biofuel development has generally favoured lignocellulosic feedstocks that require conversion into fermentable sugars. However, a large quantity of plant biomass is not viable for efficient biofuel conversion and gasification offers an attractive solution (Henstra *et al.*, 2007). Gasification has been a used for roughly 200 years as a means of producing coal gas and used during WW2 a means of reducing petroleum shortages (Brown and Brown, 2014). Later development of gas to liquid fuel production via metal catalysis was termed the Fischer-Tropsch (FT) process. Within the FT process, the water-

gas shift reaction (WGSR) follows the production of hydrogen from carbon monoxide (Dry, 2002; Demirbas, 2007);

$$CO + H_2O \rightarrow CO_2 + H_2$$

The FT process however requires a strict gas ratio dependent on which type of metal is used in catalysis. Moreover, an added downside is the requirement of high-purity gases and avoidance of certain chemical species, together with the cost of the process which accounts for 60% of a FT process plant's running expenses (Daniell, Köpke and Simpson, 2012). Microbial gas fermentation in contrast does not require stringent gas ratios or high temperatures, and irreversible biocatalytic reactions ensure a steady unidirectional product output. As a result, it is not only possible but economically viable to tap into industrial gas-waste streams (Daniell, Köpke and Simpson, 2012; FungMin Liew *et al.*, 2016). This offers an immediate and attractive solution with industries such as petroleum refineries, municipal waste incinerators, steel mills and chemical plants contributing substantial amounts of readily available CO and CO_2 waste gases. These gases also require minimal pre-treatment and can be directly used as a feedstock for microbial fermentation, effectively removing otherwise toxic pollutants from the environment (FungMin Liew *et al.*, 2016).

Could the development of microbial syngas fermentation answer the question of simultaneously reducing GHG emissions while curbing fossil fuel usage? The technology has already been put into commercial practice by multinational company Lanzatech in cooperation with two Chinese steel mill plants producing 0.1 million gallons of ethanol per year with more plant schemes to follow (Liew 2016). But with these successes, there still are substantial barriers to overcome such as improving product recovery, poor gas-liquid mass transfer, slow cell growth rates and alleviating product inhibition to name a few (Abubackar, Veiga and Kennes, 2011; Latif *et al.*, 2014; FungMin Liew *et al.*, 2016). Besides the engineering challenges, the latter problems might be solved through genetic modification of gas fermenting microbes (Simpson, Koepke and Liew, 2011; Koepke and Al-Sinawi, 2013).

1.3.2 Acetogenic bacteria

Microbial gas fermentation has been predominantly performed by a group of anaerobic, carboxydotrophic and hydrogenotrophic microorganisms that include mesophilic and

thermophilic bacteria (Henstra *et al.*, 2007). These anaerobes are typically referred to as acetogenic bacteria, or "acetogens", that are defined as users of the acetyl-CoA pathway in reducing CO_2 into acetyl-CoA, in order to conserve energy and assimilate of CO_2 into cellular carbon (Drake, Gößner and Daniel, 2008). Acetogens play an integral role in the global carbon-assimilation cycle as they are ubiquitous in nature and can be found in soil sediments, intestinal tracts of humans and mammals, and form symbiotic relationships with insects (Drake, Kusel and Matthies, 2006). Amongst several noteworthy facts, one suggests that the combined annual production and metabolism of acetate ranges from 10^{10} – 10^{13} kg annually, 10 % of which is thought that acetogens are responsible for and likely through the Wood-Ljungdahl pathway (Drake, Kusel and Matthies, 2006). Therefore, the importance of acetogens cannot be understated. Within the known 22 genera of acetogens, *Acetobacterium* and *Clostridium* accommodate the most known acetogenic species (Drake, Gößner and Daniel, 2008). Drake *et al.* (2008) shows a comprehensive list of all known isolated acetogens, Liew *et al.* (2016) further designates current relevant acetogens with growth substrates and product outputs and these data are adapted into Table 1.1.

Table 1.1, Mesophilic acetogens: Optimal temperature and pH ranges are given for each species. Table adapted from Drake *et. al.* (2008) and Liew *et. al.* (2016).

Species	Substrates	Products formed	Temp (°C)	pH	Isolation source	Ref.
Acetobacterium woodii	H_2/CO_2, CO	Acetate	30	6.8	Marine sediment	(Genthner and Bryant, 1987)
Acetonema longum	H_2/CO_2	Acetate, butyrate	30 - 33	7.8	Wood-eating termite gut	(Kane and Breznak, 1991)
Alkalibaculum bacchi	H_2/CO_2, CO	Acetate, ethanol	37	8.0-8.5	Livestock-impacted soil	(Allen *et al.*, 2010)
Butyribacterium methylotrophicum	H_2/CO_2, CO	Acetate, ethanol, butyrate, butanol	37	6	Sewage digester	(Lynd and Zeikus, 1983)
Clostridium aceticum	H_2/CO_2, CO	Acetate	30	8.3	Soil	(Adamse, 1980)
Clostridium autoethanogenum	H_2/CO_2, CO	Acetate, ethanol, 2,3-butanediol, lactate	37	5.8-6.0	Rabbit faeces	(Abrini, Naveau and Nyns, 1994)
Clostridium carboxidivorans / "P7"	H_2/CO_2, CO	Acetate, ethanol, butyrate, butanol, lactate	38	6.2	Lagoon sediment	(Liou *et al.*, 2005; Bruant *et al.*, 2010)
Clostridium coskatii "PS02"	H_2/CO_2, CO	Acetate, ethanol,	37	5.8 - 6.5	Estuary sediment	(Zahn and Saxena, 2011)
Clostridium difficile AA1	H_2/CO_2, CO	Acetate, ethanol, butyrate	35 - 40	6.5 - 7.0	New born lamb rumen	(Rieu-Lesme *et al.*, 1998)
Clostridium drakei	H_2/CO_2, CO	Acetate, ethanol, butyrate	25 - 30	3.6 - 6.8	Coal mine pond sediment	(Küsel *et al.*, 2000; Liou *et al.*, 2005)

Clostridium formicoaceticum	CO	Acetate, formate	37	7.6	Sewer mud	(Andreesen, Gottschalk and Schlegel, 1970; Yang, Tang and Okos, 1987)
Clostridium glycolicum	H_2/CO_2	Acetate	37 - 40	7.0 - 7.5	Sea-grass roots	(Küsel *et al.*, 2001)
Clostridium ljungdahlii	H_2/CO_2, CO	Acetate, ethanol, 2,3-butanediol, lactate	37	6	Chicken waste	(Tanner, Miller and Yang, 1993)
Clostridium magnum	H_2/CO_2	Acetate	30 - 32	7.0	Freshwater sediment	(Schink, 1984)
Clostridium mayombei	H_2/CO_2	Acetate	33	7.3	Soil-feeding termite gut	(Kane, Brauman and Breznak, 1991)
Clostridium methoxybenzovorans	H_2/CO_2	Acetate, formate	37	7.4	Olive oil mill wastewater	(Mechichi *et al.*, 1999)
"Clostridium ragsdalei" or "P11"	H_2/CO_2, CO	Acetate, ethanol, 2,3-butanediol, lactate	37	6.3	Duck pond sediment	(Huhnke, Raymond L.; Lewis, Randy S.; Tanner, 2008)
Clostridium scatologenes	H_2/CO_2, CO	Acetate, ethanol, butyrate	37 - 40	5.4 - 7.5	Soil, coal mine pond sediment	(Weinberg, 1927)
Eubacterium limosum	H_2/CO_2, CO	Acetate	38 - 39	7.0 - 7.2	Sheep rumen fluid	(Rode, Genthner and Bryant, 1981)
Oxobacter pfennigii	H_2/CO_2, CO	Acetate, butyrate	36 - 38	7.3	Steer rumen fluid	(Krumholz and Bryant, 1985)

1.4 Energy conservation in acetogens

1.4.1 The Wood-Ljungdahl pathway

The reductive acetyl-CoA pathway entails the synthesises of acetyl-CoA, step-wise from assimilated and reduced CO_2, this pathway is also more commonly known as the Wood-Ljungdahl pathway (WL-pathway), and was characterized by Harland G. Wood and Lars G. Ljungdahl (Drake, Kusel and Matthies, 2006; Ragsdale and Pierce, 2008). The defining feature of this ancient pathway is by facilitating energy conservation through facultative autotrophic carbon assimilation, whose final metabolic product is acetate produced from 2 molecules of CO_2 (Imkamp and Müller, 2007). Before taking the name, initial observations of this pathway's CO_2 to acetate synthesis were observed by Barker and Kamen in 1945 through $^{14}CO_2$-labelling (Wood, 1952b; Ivey and Ljungdahl, 1986). At the time, experiments were performed comparing glucose conversion into acetate using thermophile, *Moorella thermoacetica* (previously, *Clostridium thermoaceticum*), against *C. aceticum*'s ability of acetate generation from H_2/CO_2, which demonstrated CO_2 uptake and uniform integration into acetate's carboxyl and methyl carbons respectively (Wood,

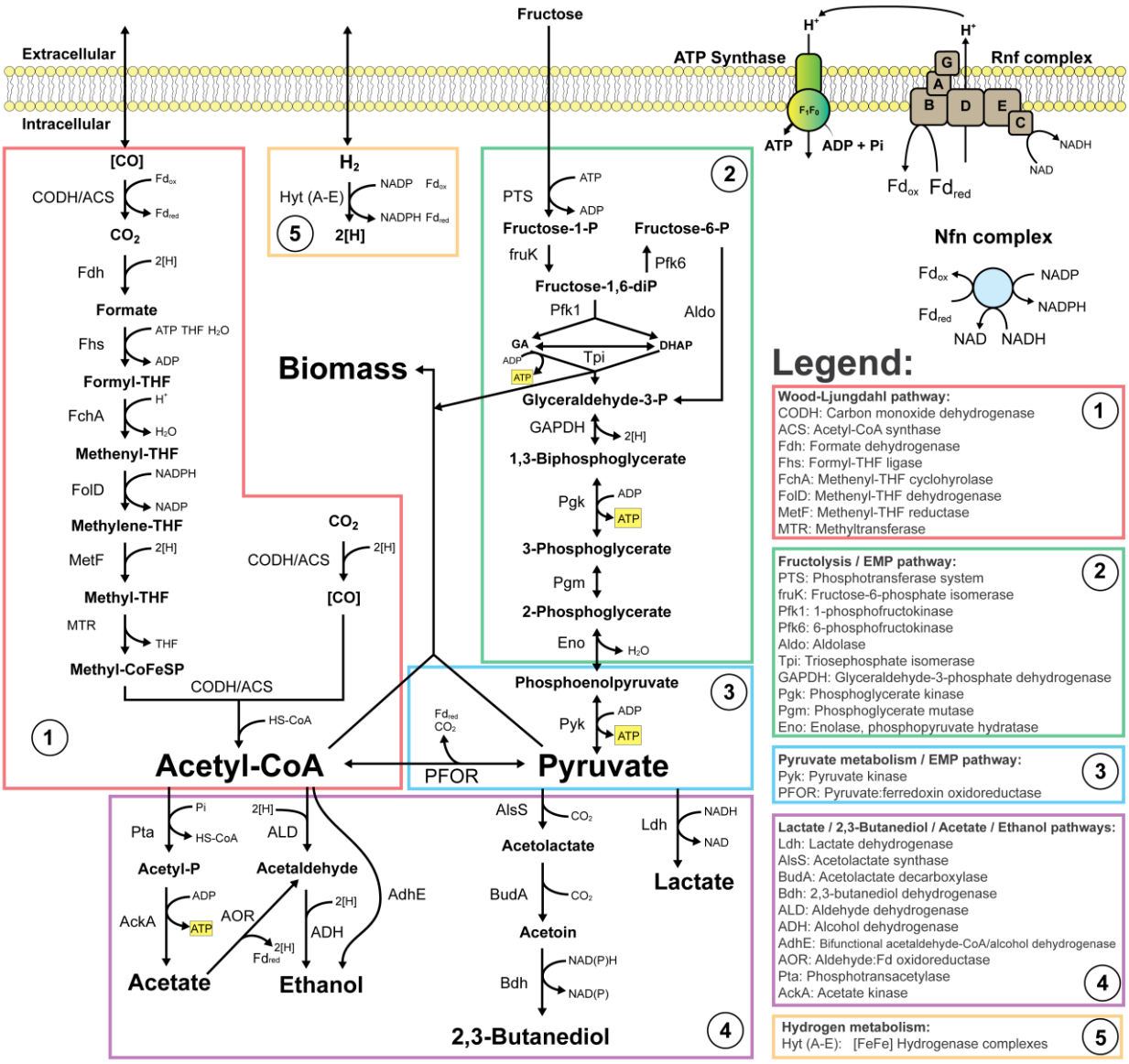


Figure 1.1, Overview of energy metabolism in *C. autoethanogenum*: Pathways are grouped according to specific metabolic roles and show the coupling between the WL-pathway (1) and fructolytic pathway (2). The WL-pathway (1) features the methyl (left side) and carbonyl (right side) branches. (4) Features all pathways leading to fermentation products naturally produced by *C. autoethanogenum*. Not shown; full pathway scheme of hydrogenases.

1952a). Isotope labelling was a relatively new technique at the time and several methods needed to be modified to accommodate characterisation of WL-pathway reactions that require strictly oxygen-free environments (Ragsdale and Pierce, 2008) and hence the next 60 years were spent to fully understand this pathway (Drake, Gößner and Daniel, 2008).

History aside, the WL-pathway can be described as a twin-branched pathway, with branches referred to as the methyl and carbonyl branches or, in some cases, “Eastern” and “Western” branches, respectively. Each branch handles the assimilation of 1

molecule of CO_2 (Ragsdale, 1997). CO and H_2 act as reducing agents for this metabolic process. In the absence of H_2, CO is the sole reducing agent. Required H_2 is produced through the water-gas shift reaction described earlier (Sipma *et al.*, 2006).

The WL-pathway starts by reducing CO_2 to CO in the carbonyl branch and a methyl-group in a methyl branch. The methyl branch first reduces CO_2 into formate, then activates it through formation of formyltetrahydrofolate (formyl-THF), and in a series of enzymatic reactions this is finally reduced to a methyl group. The methyl group is transferred to a subunit of the carbon monoxide dehydrogenase/acetyl-CoA synthase complex (CODH/ACS), with the aid of a methyltransferase (MTF) and corrinoid iron-sulphur protein (CoFeSP). The carbonyl branch pathway follows the reduction of CO_2 into CO that is directly processed through the CODH/ACS complex. The bifunctional CODH/ACS complex fuses the carbonyl-processed CO with the released methyl group to form acetyl-CoA, which in turn is processed by phosphotransacetylase (Pta) and acetate kinase (AckA) into acetate. The last enzymatic reaction produces a 1 molecule of ATP through substrate level phosphorylation (SLP), which is paid back into formate activation and therefore no net ATP is produced via SLP (Imkamp and Müller, 2007; Ragsdale and Pierce, 2008; Köpke, Mihalcea, Bromley, *et al.*, 2011; Daniell, Köpke and Simpson, 2012; Debabov, 2013; Schuchmann and Müller, 2014), Figure 1.1. It is also worth noting that from the WL-pathway reactions the iron-sulfur protein ferredoxin is responsible for much of the electron transport and redox mediation (Schuchmann and Müller, 2014). Its role in acetogen metabolism is described further in the text.

1.4.2 Carbon assimilation

The WL-pathway's performance rests on the vital CODH/ACS complex which is responsible for CO and CO_2 fixation to acetyl-CoA formation (Ragsdale, 2004). As a carbon source, CO provides the dual function of an electron and carbon source, and therefore it is not necessary to supplement acetogenic growth with additional H_2. This is in contrast to growth on CO_2 where H_2 is required (Ragsdale, 2004). Moreover, CO has a roughly 1000-fold reduction potential compared to NADH and forms stable complexes with enzyme metal binding sites, such as Ni, Mo, Fe, Cu -CO complexes in ACS (Chen *et al.*, 2003; Ragsdale, 2004). CODH also functions separately and is capable of biologically performing the water-gas shift reaction, utilising CO and H_2O, and providing H_2 and CO_2 (Ragsdale, 2004). Besides liberating required H_2, energy is captured via reduced

ferredoxin, and CO_2 can be re-assimilated through other WL-pathway reactions (Ragsdale and Pierce, 2008; Köpke, Mihalcea, Bromley, *et al.*, 2011). CODH's close relationship with ACS has been crystallographically shown in *M. thermoacetica* to form a 138 Å long channel that facilitates intermolecular CO transport (Doukov *et al.*, 2008). These findings could infer a similar complex architecture in *Clostridium* species.

1.4.3 Autotrophic energy conservation

As there in no net ATP gain directly from the WL-pathway via SLP, energy conservation must come about in a different form. The membrane-bound, multi subunit ferredoxin-NAD:oxidoreductase complex (Rnf) is responsible for proton translocation which in turn drives the ATP synthase complex (ATPase) into chemiosmotically generating ATP (Biegel *et al.*, 2011; Tremblay *et al.*, 2013). The Rnf complex is noted to work differently dependent on acetogens and as such the *A. woodii* Rnf complex which translocates sodium ions, but *C. ljungdahlii* and *C. autoethanogenum* Rnf complexes transport protons instead (Müller *et al.*, 2008; Biegel *et al.*, 2011; Hess, Schuchmann and Müller, 2013; Tremblay *et al.*, 2013; Schuchmann and Müller, 2014). The energy conservation reaction requires electron transfer via reduced ferredoxin (Fd) to NAD^+, which drives this proton translocation (Biegel *et al.*, 2011; Tremblay *et al.*, 2013). The extracellular proton gradient drives a F_0F_1 ATPase into generating cellular ATP (Ivey and Ljungdahl, 1986; Biegel *et al.*, 2011; Schuchmann and Müller, 2014). In brief, it can be emphasised that a high ferredoxin/NAD^+ ratio plays the vital role of energy conservation in acetogens (Biegel *et al.*, 2011; Schuchmann and Müller, 2014). Other energy conservation systems include electron bifurcation via hydrogenases (Herrmann *et al.*, 2008) and this mechanism has also been shown for other complexes such as the NADH-dependent reduced ferredoxin:NADP oxidoreductase complex (Nfn) (S. Wang *et al.*, 2013). Nfn is proposed to play a role in CO_2 reduction to formate by providing ferredoxin and NADPH.

1.4.4 The EMP pathway

Acetogens can also heterotrophically survive on hexoses using the Embden-Meyerhof-Parnas (EMP) pathway to produce 3 molecules of acetate for every molecule of hexose, via pyruvate synthesis. The oxidation of hexose results in the net synthesis 2 molecules of ATP through SLP (Wood, 1952a; Ivey and Ljungdahl, 1986). Of the 3 molecules of acetate produced, 2 are directly a result of the EMP pathway which oxidises pyruvate to

acetyl-coA. The remaining acetate is produced indirectly through the WL-pathway using reducing equivalents (H^+) that are released by the EMP pathway (Müller, 2003). Moreover, pyruvate that is synthesised by the EMP pathway undergoes a reversible reaction into acetyl-CoA via pyruvate:ferredoxin-oxidoreductase (PFOR) that concomitantly releases CO_2 and reduced ferredoxin in the process (Müller, 2003). Pyruvate also is an important component in the synthesis of 2,3-butanediol and lactate, both highly sought platform chemicals that can be naturally produced by acetogens (Köpke, Mihalcea, Liew, *et al.*, 2011).

1.5 The genus *Clostridium*

The genus *Clostridium*, belongs to the phylum *Firmicutes*, class *Clostridia*, order *Clostridiales*, family *Clostridiaceae*, comprised of organisms with a high degree of phylogenetic variation. Before implementation of advanced 16S ribosomal RNA techniques, this genus was described for their anaerobic metabolism, endospore formation, low GC content (< 50 %), Gram-positive cell wall, and non-sulphate reduction (Lawson *et al.*, 1993). After the advent of 16S rRNA sequencing technology, the genus has reclassified many of its species into other genera including *Eubacterium*, *Moorella*, *Oxobacter*, *Syntrophospora* and *Thermoanaerobacterium* (Dürre, 2005).

The diversity of *Clostridium* organisms span over 300 species listed in culture collections (ATCC, American Type Culture Collections and DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen). These cultures include many important and historic pathogens that are known to produce a potent range of deadly toxins, the latter not being limited to *Clostridium botulinum, Clostridium difficile, Clostridium perfringens* and *Clostridium tetani* (Hatheway, 1990; Dürre, 2005). Well-known examples include botulinum neurotoxin produced by *C. botulinum*, which is responsible for flaccid paralysis (A Münchau, Münchau and Bhatia, 2000) and *C. difficile* that produces Toxins A and B, responsible for enterotoxic and cytotoxic activities (Hatheway, 1990; Di Bella *et al.*, 2016).

Conversely, many *Clostridium* species have been found to be industrially relevant, with historic example being the previously mentioned *C. acetobutylicum* and newly studied *Clostridium* acetogens listed in Table 1.1. Other notable non-acetogenic, solventogenic *Clostridia* include *Clostridium saccharoperbutylacetonicum, Clostridium beijerinckii*, *Clostridium pasteurianum* and *Clostridium saccharobutylicum* (Poehlein *et al.*, 2017). Many of these species are also attractive candidates for further genetic engineering, with

goals of improving product output, aerotolerance, solvent tolerance, high-density fermentations and general stability in continuous bioprocessing (Papoutsakis, 2008).

1.5.1 *Clostridium autoethanogenum*

Clostridium autoethanogenum is a mesophilic, Gram-positive, lithotrophic, motile anaerobic bacterium that resembles a rod-like structure (Abrini, Naveau and Nyns, 1994). The organism was enriched and isolated from rabbit feces and is capable of autotrophically fixing carbon via CO and CO_2 + H_2. As a typical acetogen, it features the WL-pathway as its means of producing acetate from CO but it also can produce appreciable amounts of ethanol, 2,3-butanediol and lactate using this pathway (Abrini, Naveau and Nyns, 1994). The organism can utilise a diverse range of carbon sources for energy conservation, which include pyruvate, xylose, arabinose, fructose, rhamnose, L-glutamate, and several less-used substrates (Abrini, Naveau and Nyns, 1994; Drake, Kusel and Matthies, 2006). With the capability to assimilate CO and CO_2, and transform them into ethanol, *C. autoethanogenum* has become an attractive acetogen in strategies for producing platform chemicals, while reducing GHG emissions. Interestingly, the closely-related *C. ljungdahlii* shares > 98% of its genome sequence with *C. autoethanogenum* and is almost identical in its 16S rRNA sequence, suggesting very high relatedness, but substantial feature differences separate the organisms (Stackebrandt *et al.*, 1999; Humphreys *et al.*, 2015). Studies analysing orthologous protein profiles of both organisms show that a difference of 221 unique proteins exist. A large portion of these proteins are hypothetical, but others are involved in metabolic activity (Brown *et al.*, 2014). Therefore, it is speculated that these highly similar acetogens shared a common ancestor. In terms of research, *C. ljungdahlii* has been studied more than *C. autoethanogenum* with regards to gas fermentation applications and study of the Wood-Ljungdahl pathway, possibly explained through *C. ljungdahlii* earlier isolation and characterisation as an industrially relevant strain (Tanner, Miller and Yang, 1993). Both acetogens can produce similar solvent and fermentative products as listed in Table 1.1, but studies have shown differences in their fermentative product outputs on syngas and sugars, as well as defined medium requirements (Cotter, Chinn and Grunden, 2009; Guo *et al.*, 2010).

1.5.1.1 Engineering *C. autoethanogenum* into an industrial biocatalyst

As *C. autoethanogenum* has a relatively wide product repertoire (Köpke, Mihalcea, Bromley, *et al.*, 2011), it has become the focus of several studies, with patents filed in the past 10 years that have aimed to enhance its bioprocess functionality and to expand the range of fermentative products. To date, there have been patents that claim to enhance *C. autoethanogenum*'s ethanol output, solvent tolerance and nutrient synthesis (Simpson, Koepke and Liew, 2011; Koepke and Al-Sinawi, 2013; Min *et al.*, 2013; Nagaraju *et al.*, 2016). Moreover, further patents and studies have also sought to broaden *C. autoethanogenum*'s product spectrum to include; butanol, 2-butanol, butanone, acetone, isopropanol, 3-hydroxypropionate and isoprene products, with the introduction of plasmid-based, exogenous genes (Köpke and Liew, 2011, 2012; Chen, Liew and Koepke, 2013; Koepke and Chen, 2013; Mueller, Koepke and Nagaraju, 2013).

Further work has also aimed to better understand *C. autoethanogenum's* metabolic pathway enzymes using gene editing tools to disrupt genes for important enzymes such as aldehyde:ferredoxin oxidoreductase (AOR), bifunctional aldehyde/alcohol dehydrogenase (AdhE) and the CODH/ACS complex (Fungmin Liew *et al.*, 2016; Liew *et al.*, 2017). *C. autoethanogenum* has also been the focus of study in acetogenic energy conservation by examining specific co-factors of oxidoreductases (Mock *et al.*, 2015), differences between auto- and heterotrophic growth (Marcellin 2015), and studying flavin-based electron bifurcation complexes (S. Wang *et al.*, 2013). Commercial *C. autoethanogenum* strains LBS1560 (DSM19630) and LBS1561 (DSM23693) have also been patented by Lanzatech through strain "evolution" that have claims to enhanced fermentation products and bioprocess performance (Simpson *et al.*, 2009; Heijstra *et al.*, 2013).

Although there is rapid progress into the better understanding of this industrially important acetogen, there are many areas that still require further enquiry. Modern RNA-sequencing methods have become a standard means in studying an organism's transcriptome. Recent and upcoming publications have examined *C. autoethanogenum*'s transcriptome during auto- and heterotrophic conditions (Marcellin *et al.*, 2016), and pH/gas-shift conditions (Henstra, 2018). Available transcriptomic data from these studies showed genes for expected metabolic, physiological and stress-related proteins to be highly expressed, which include CODH, PFOR, ferredoxins, and flagellar components to name a few (Marcellin *et al.*, 2016; Henstra, 2018). Interestingly, amongst these highly

transcribed RNA, putative quorum sensing peptides are also found to be prominent. Reasons as to why these genes are so highly expressed has not been made clear as the role of quorum sensing has not yet been understood in acetogens; this study aims to do so.

1.6 Quorum sensing

1.6.1 Overview

Quorum Sensing (QS), is defined as a system of stimuli and gene regulation in response

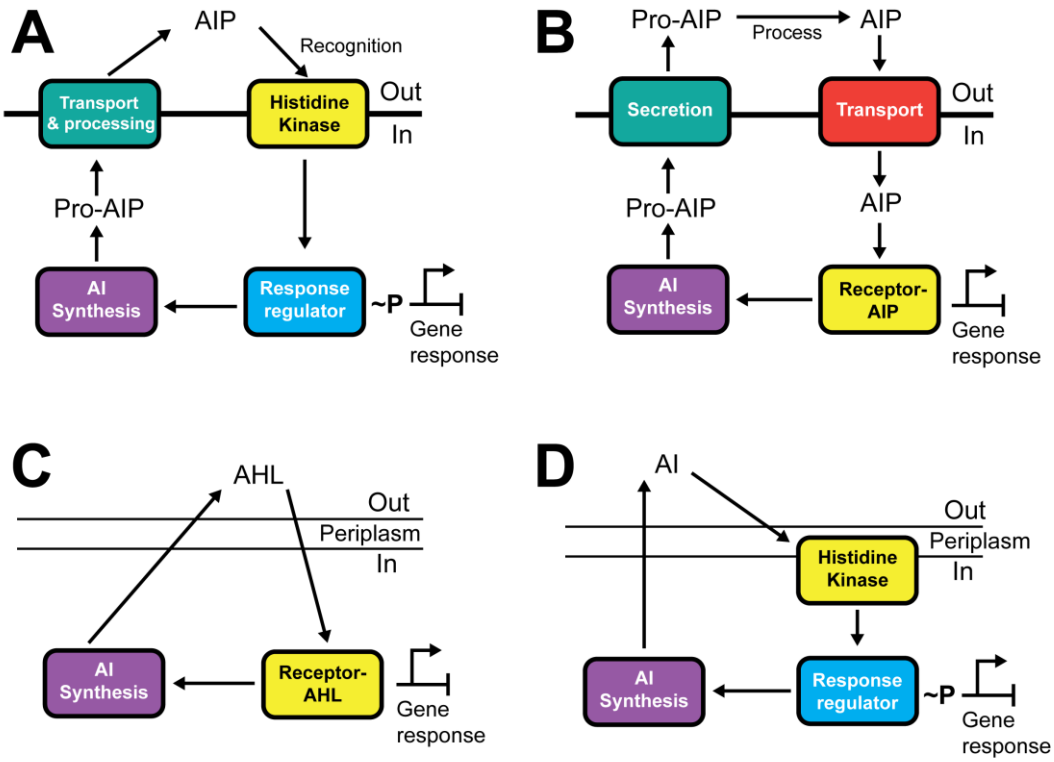


Figure 1.2, Schematic overview of QS circuits; Found in Gram-positive bacteria; (A) Two-component signalling (e.g. Agr), and (B) AIP transport back into the cell for recognition (e.g., RNPP). Found in Gram-negative bacteria; (C) LuxI/LuxR-type system, AIs in the form of AHLs are diffuse through the cell, (D) two-component system, with histidine kinase receptor function similar to (A). Autoinducing peptide (AIP). Figure adapted from Rutherford and Bassler (2012).

to cell population density (Bassler, 1999). QS activity amongst cells can also be modulated and integrated by environmental stimuli such as temperature, pH, osmolarity, nutrient availability and even other species (Whitehead, 2001), which aid organism's environmental survival. The best studied cases of QS have been within bacteria and observed in a wide range of genera that exhibit a diverse variety of physiological actions

that include biofilm formation, virulence, toxicity, sporulation, swarming, motility and antibiotic production to name a few (Miller and Bassler, 2001; Whitehead, 2001). Both Gram-negative and Gram-positive bacteria undertake QS-related behavioural changes and perform the actions listed, but either Gram class produces uniquely different low-molecular mass signalling molecules; Gram-negative species employ the use of fatty-acid derivatives, most notably acylated homoserine lactones, while Gram-positive species use processed oligopeptides (Miller and Bassler, 2001; Whitehead, 2001; Rutherford and Bassler, 2012). The term "two-component systems" (TCS) or two-component regulatory system (TCRS) requires definition. In a broad sense, two-component systems are a means in which bacteria can sense and respond accordingly to environmental stimuli, stresses and growth conditions (Stock, Robinson and Goudreau, 2000; Skerker et al., 2005). Certain QS systems fall within this broad definition and can be considered a TCS sub-class. The core TCS arrangement follows a membrane-bound histidine kinase detecting a defined environmental stimulus, e.g.: antibiotic, pH, chemoattractant or signal molecule, etc., and a cogent response regulator coordinates the cellular response, typically through mediated gene expression (Mascher, Helmann and Unden, 2006). There have been numerous TCSs characterised interacting with a wide variety of stimuli, but the sub-class of QS systems only responds to auto-inducing signal molecules.

1.6.2 QS in Gram-negative bacteria

Many Gram-negative bacteria species possess a QS circuit scheme similar to the originally studied in marine prokaryote *Vibrio fischeri*. In this early studied system, there are two distinct proteins, the autoinducer generating enzyme LuxI and the receptor LuxR (Ng and Bassler, 2009). The former protein is responsible for the synthesis of *N*-acylated homoserine lactone (AHL) molecules which are the signalling molecules of the QS circuit. Depending on the species producing them, AHL molecules can possess acyl chains spanning 4 to 18 carbons in length (Ng and Bassler, 2009). As the produced AHL increases in concentration, they diffuse back into the cell and bind to cognate LuxR transcription factors thus forming a complex which activates transcription of target genes (Engebrecht, Nealson and Silverman, 1983; Rutherford and Bassler, 2012). The LuxI and LuxR signal-receiver arrangement is conserved in other Gram-Negative bacteria and referred to as "LuxI/LuxR-type". Within this arrangement exist several hierarchical combinations (Engebrecht, Nealson and Silverman, 1983; Miller and Bassler, 2001).

V. fischeri's QS system is a classic example of mutualistic behaviour with its host (Ruby, 1996). Its QS system was found to be responsible for regulation of light-producing bioluminescence, and it has been observed that this trait is beneficial in the symbiotic relationship that *V. fischeri* has with fish and squid hosts. Depending on the host, bioluminescence provides a means of either attracting mates, prey or warding off predators (Ruby, 1996; Visick *et al.*, 2000). How this QS system operates is through *V. fischeri*'s coalescence into a dense cell culture confined in the host's "light organ" or storage organ. This increase in *V. fischeri* cell concentration stimulates AHL molecule biosynthesis and after reaching a specific signal threshold, the genetic response is translated into a bioluminescent light signal (Ruby, 1996; Visick *et al.*, 2000).

Further examples can be found in the human pathogens such as *Pseudomonas aeruginosa* (Passador *et al.*, 1993; Nadal Jimenez *et al.*, 2012; Lee and Zhang, 2014). *P. aeruginosa*'s QS system consists of 2 LuxI/LuxR-type homologues, LasI/LasR and RhII/RhlR with each LuxI homolog producing respective AHLs. The pairs act in conjunction, with the LasI/LasR system being the "master" QS regulator, controlling RhlI/RhlR, which in turn mediates a secondary QS response cascade (Lee and Zhang, 2014). This approach allows the pathogen to arrange QS-mediated, cell-density-dependent responses to expedite infection. Specifically, in *P. aeruginosa*'s case, virulence factors are mediated which initiates proteases leading to host tissue damage. The RhlI system is subsequently stimulated into producing the antibiotic pyocyanin, targeting *P. aeruginosa*'s bacterial competition, damaging host's airway cilia and immune system defences (Miller and Bassler, 2001; Ng and Bassler, 2009; Lee and Zhang, 2014). The formation of biofilm is also the result of these systems, which creates an ideal micro-environment for further bacterial growth and protection from medical treatment. Interestingly, one of the symptoms of cystic fibrosis is the proliferation of *P. aeruginosa* within its biofilm by coating the lung canals of patients that further exacerbate cystic fibrosis symptoms (Singh *et al.*, 2000). In summary, Gram-negative QS systems produce specific AHL molecules capable of intra-species communication. Receptors of these systems exist within the cytoplasm and AHLs can typically permeate cell walls and membranes (Taga and Bassler, 2003; Ng and Bassler, 2009). Besides *P. aeruginosa*, this class of QS system is wide-spread amongst many Gram-negative bacteria. Further examples highlight the systems diverse mechanisms and physiological functions, and can be studied in Whitehead *et al.*'s (2001)

excellent review paper and several other (Antunes *et al.*, 2010; Papenfort and Bassler, 2016).

1.6.4 QS in Gram-positive bacteria

In contrast to Gram-negative bacteria, members of the phylum *Firmicutes* produce predominantly short peptides as signalling molecules, which can either be linear or cyclic. These can be defined by circuit receptors that either receive AIPs intra- or extracellularly (Rocha-Estrada *et al.*, 2010; Cook and Federle, 2014; Monnet, Juillard and Gardan, 2014). Within the latter two classes there exist four general families, the 1) RNPP (intra-cellular receptor), 2) Agr-type (extra-cellular), 3) Competence stimulating peptides (CSPs) and bacteriocins (extra-cellular) and 4) Rgg-type (intra-cellular receptors) which are related to the RNPP-type systems (Cook and Federle, 2014; Monnet, Juillard and Gardan, 2014). Further examples in literature detail roles of QS in pathogenic *Clostridium* spp., and how these pathogens use their QS systems for sporulation, cell adhesion, biofilm and toxin production, notably in *C. botulinum*, *C. difficile*, *C. perfringens* and *Clostridium sporogenes* (Carter *et al.*, 2005; Sebaihia *et al.*, 2006; Zhao, Montville and Schaffner, 2006; Cooksley *et al.*, 2010; Li *et al.*, 2011; Ohtani, 2016). Out of the 4 QS groups listed, the RNPP and Agr-type are found to be conserved in several *Clostridium* species.

1.6.4.1 RNPP quorum sensing

The RNPP group refers to the RAP histidine kinases, NprR regulator, PrgX receptor and PlcR transcription regulator, found in *Bacillus subtilis*, *Enteroccocus faecalis* and *Bacillus cereus* respectively, (Kozlowicz, Bae and Dunny, 2004; Rocha-Estrada *et al.*, 2010; Cook and Federle, 2014). The RNPP mechanism is based on regulators that employ linear AIPs, and both RNPP regulator and AIP genes are easily identified through their close chromosomal proximity. As with other QS systems, AIPs are matured intra- or extracellularly via generic enzyme cleavage before being transported back into the cell through membrane-bound oligopermeases to be received by respective regulators (Rocha-Estrada *et al.*, 2010; Kotte, 2013).

In *Clostridium* spp., the RNPP-like system has been identified through sequence and structure similarities to systems found in the *Bacillus* genus (Rocha-Estrada *et al.*, 2010; Kotte *et al.*, 2017). Although there are not many RNPP studies in *Clostridia*, RNPP-like regulator orthologs were studied in *C. difficile* and found to regulate toxin production and

sporulation and have been aptly designated as regulator of sporulation and toxins (Rst), although no signal molecule encoding gene could be distinguished (Edwards, Tamayo and McBride, 2016). Ongoing work has also sought to characterise this system in solventogenic *C. acetobutylicum* (Kotte *et al.*, 2017). A total of 10 putative RNPP-like receptor genes were identified with 8 respective signal peptide genes (Kotte, 2013). The purported function of this system is that specific RNPP signalling peptides are responsible in the regulation in solventogenesis and sporulation (Bean, 2017).

1.6.5 Accessory gene regulator QS

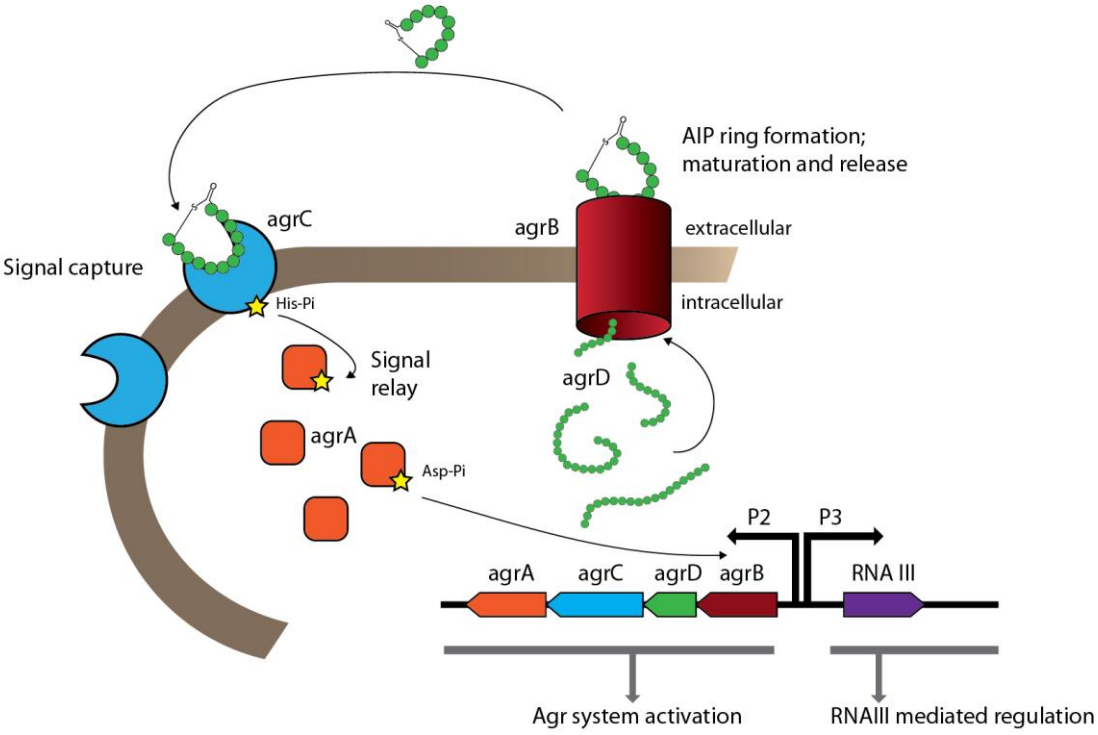


Figure 1.3, Overview of the Agr system in *S. aureus*; This system is divided into four parts, the *agrB* functions as a membrane-bound AIP maturation and exportation unit and *agrD* is encoding a precursor of the AIP signal. *agrA* and *agrC* encode for the two-component receptor/response functions of this system via phosphorylation. AgrC is recognized as a membrane-bound histidine kinase which phosphorylates *agrA* in response to binding the AIP signal. Phosphorylated *agrA* binds to two promoters (P2 and P3), activating the divergently transcribed *agrBDCA* operon and the regulatory RNAIII. This action results in a positive feedback loop for production of more AIPs and production of RNAIII, thus resulting in global gene regulation for the rest of the cell (Ji et. al. 1995; Ji et. al. 1997).

The accessory gene regulator (Agr) system is noted to be evolutionarily conserved across *Firmicutes*, and found in many *Clostridium* pathogens (Wuster and Babu, 2008; Atkinson and Williams, 2009). The Agr system was originally studied in detail in *Staphylococcus*

aureus, to which a wealth of literature explains the mechanism of the Agr QS circuit (Recsei et al., 1986; Peng et al., 1988; Ji, Beavis and Novick, 1997; Novick, 2003; Thoendel et al., 2011; Bronesky et al., 2016). This contrasts with studies in the *Clostridium* species that have been primarily focused in Agr's role in pathogenesis (Cooksley et al., 2010; Darkoh et al., 2015; Ohtani, 2016).

The "accessory gene regulator" term was first coined by Recsei et. al., (1986) who initially described the system in *S. aureus* as a "global exo-protein regulator". The outline of the gene locus structure consists of two outwardly facing operons with respective P2 and P3 promoters, Figure 1.3 (Ji, Beavis and Novick, 1997). P2 promotes the expression of the agr operon that encodes four key component genes paired in a TCS-style arrangement; peptide and exporter genes, *agrBD* and receiver and signal relay genes, *agrCA* (Ji, Beavis and Novick, 1997; Novick, 2003). The *agrD* gene is transcribed and translated into an immature AIP, which is active only after maturation and extracellular export by the membrane-bound AgrB transporter. The AgrD peptide is post-translationally modified and consists of seven to nine residues, depending on Agr sub-group. Thus, the peptide consists a 5-membered ring structure and, formed via thiolactone bond formation between the SH group of a cysteine and the C-terminal carboxy group, and an N-terminal tail. As other AIPs, it behaves as an extracellular ligand marker for population density (Ji, Beavis and Novick, 1995). AgrD signal recognition is through the AIP's binding to the transmembrane AgrC signal receptor, a histidine kinase, which relays an intracellular response through phosphorylation of AgrA, the response regulator (Ji, Beavis and Novick, 1997; Novick, 2003). The response is mediated back to promoters P2 and P3, and a positive-feedback effect causes further AgrD pro-peptides to be expressed. The P3 promoter mediates a RNAIII transcription cascade which is responsible for target gene regulation (Novick 2003). The Agr system in *S. aureus* mediates control over numerous genes including food poisoning enterotoxins B, C, D, haemolytic cytotoxins, anti-bactericidal enzymes and nutrient regulating enzymes (Chamberlain and Imanoel, 1996; Vojtov, Ross and Novick, 2002; Novick, 2003; Thoendel et al., 2011). This circuit is reliant on the autocatalytic activity of the AgrD peptide which follows the typical mode of QS systems by needing to cross the signal concentration threshold to initiate an upregulated, positive-feedback.

1.6.5.1 Agr systems in the genus Clostridium

With regards to *Clostridium* spp., the Agr system is not as well understood in comparison to *S. aureus*, but as evidence suggest it is widely conserved (Wuster and Babu, 2008). The pathogenic examples in the literature point to utilisation of the Agr as an accessory TCS, similarly as described in the *S. aureus* counterpart.

Studies in *C. perfringens* have shown evidence of an Agr-like system, termed CpAL, which behaves similarly to *S. aureus* counterparts AgrB and AgrD (Vidal, Shak and Canizalez-Roman, 2015). The CpAL system mediates a host of toxin-producing genes via a two-component VirR/VirS-RNA cascade system that includes alpha-toxin encoding *plc*, perfringolysin O toxin encoding *pfoA*, and collagenase kappa-toxin gene *colA* (Katayama *et al.*, 1993; Lyristis *et al.*, 1994; Rood, 1998; Ohtani, 2016). Recent studies have also shown biofilm formation to be regulated by the CpAL system (Vidal, Shak and Canizalez-Roman, 2015). Moreover, non-pathogenic activity mediated by CpAL includes sporulation through the initiator *spo0A* gene and sigma factors, σ^F and σ^G, (Li *et al.*, 2011).

Agr-like systems have also been studied in the pathogen *C. botulinum* and non-pathogen, *C. sporogenes*. Both possess two copies of *agrB* and *agrD* genes, but these were not located next to *agrC* and *agrA* genes, and thus the precise TCSs mediating the QS response remained unknown (Cooksley *et al.*, 2010). It was speculated that a *C. perfringens* AgrCA-like, VirR/VirS homolog fulfils this role, represented by *C. botulinum* genes CBO0575 and CBO1052 respectively, and in *C. sporogenes*, CLSPOx_03050 and CLSPO_c10570 respectively. Study of the Agr system began with the manipulation of *C. sporogenes* AgrB through antisense RNA modulation, resulting in a 100 to 1000-fold reduction in sporulation. In *C. botulinum*, both *agrD* genes were disrupted with ClosTron mutagenesis. Respective mutants revealed dissimilar phenotypic changes. *agrD1* mutants showed a greatly reduced sporulation, while the *agrD2* mutant phenotype was shown to include reduced neurotoxin production (Cooksley *et al.*, 2010).

C. difficile also possesses Agr-type systems, although its role is not as clear. Initially reported in the *C. difficile* 630 ribotype 12 genome, a partial *agr* locus (termed *agr1*) was found containing AgrBD components that share 30 % and 24 % sequence identity with staphylococcal AgrBD proteins, respectively (Sebaihia *et al.*, 2006; Stabler *et al.*, 2009). In the sequenced *C. difficile* 027 R20291 strain a full *agr* locus was annotated, termed *agr2* alongside the partial *agr1* locus (Stabler *et al.*, 2009; Martin *et al.*, 2013). After targeting

R20291's *agrA* response regulator via ClosTron mutagenesis, an RNA-Seq transcriptomic analysis was performed to discern phenotypic changes that identified 75 differentially expressed genes which included flagellar synthesis, uncharacterized TCSs and above all, the down-regulation of toxin A producing *tcdA* (Martin *et al.*, 2013). This may come as no surprise as there are numerous TCSs found in *C. difficile* which potentially play roles in the expression of toxin genes *tcdA* and *tcdB* (Dineen *et al.*, 2007; Antunes, Martin-Verstraete and Dupuy, 2011; Saujet *et al.*, 2011; Martin *et al.*, 2013; Smits *et al.*, 2016).

These examples show the Agr signalling infrastructure in pathogenic *Clostridium* species which are associated with in toxin production, virulence, spore formation and mediation of infection amongst other behaviours. A large emphasis is placed upon developing potential cures and treatments, which inevitably furthers understanding into pathogenic Agr mechanisms.

Agr QS has not been fully understood in industrially-relevant clostridia except for recent pivotal work in *C. acetobutylicum* (Steiner *et al.*, 2012). *C. acetobutylicum* possess a complete *agrBDCA* cluster with components *agrBD* and *agrCA* expressed through independent promoters. Targeting these clusters with ClosTron mutagenesis resulted in a reduction of heat-resistant spores and loss of starch-like storage compound granulose, (Steiner *et al.*, 2012). Moreover, this study elegantly showed that *agr* mutants could be complemented through the addition of agar-diffusible, AgrD-derived peptides, performed either through proximally streaking of WT *C. acetobutylicum* near mutants, or with the addition of synthetic cyclic AgrD peptides directly into cultures (Steiner *et al.*, 2012). The study gave strong evidence for the utility of the Agr system in a non-pathogenic, solventogen and raised important questions as to why this system is prevalent in many non-pathogenic *Clostridium* spp.

Despite genomic sequencing revealing the presence of putative Agr systems in many acetogenic *Clostridium* spp., no work has been published to characterise them and their role in the organism's life cycle. Specifically, in *C. autoethanogenum* the *agrD1* encoding gene exhibits pronounced transcription levels (Marcellin *et al.*, 2016; Henstra, 2018). What could the plausible role of the Agr system be in *C. autoethanogenum* and other acetogens? Is it possible that it plays a similar role as in *C. acetobutylicum*, or could it potentially act as a stress response regulator as observed in *C. thermocellum* (Verbeke *et al.*, 2017)? Besides the mechanistic means of action, what is the necessity of these

systems existing in the first place and does a functional Agr system deliver a survival advantage to a species?

1.7 Genetic tools for *Clostridium* spp.

Examples used in this chapter have referred to genetic modification of different *Clostridium* species through a variety of techniques. A recent review by Minton *et. al.* (2016) offers a comprehensive roadmap for the implementation of these techniques in *Clostridium* spp., with numerous case examples of successful usage in pathogens and solventogens alike. The importance of these techniques and means of implementation cannot be understated as interest in studying pathogenic and industrially-relevant *Clostridia* spp. has grown correspondingly to the increasing availability of new genetic tools (Bannam and Rood, 1993; Green *et al.*, 1996; Papoutsakis, 2008; Leang *et al.*, 2013; Ng *et al.*, 2013; Pyne *et al.*, 2013, 2014; Zhang, Grosse-Honebrink and Minton, 2015; Huang *et al.*, 2016; Molitor *et al.*, 2016; Nagaraju *et al.*, 2016).

1.7.1 ClosTron mutagenesis

The first novel implementation of genetic disruption with an intron system was reported using *Lactococcus lactis ltrB* (Ll.ltrB) mobile group II intron in *E. coli*, which targeted a specific gene site mediated through a RNA retro-homing mechanism (Karberg *et al.*, 2001). This method was later adapted for *Clostridium* which lead to the creation of the ClosTron (Heap *et al.*, 2007). However, as ClosTron was based on the group II intron systems, also known as the TargeTron (Zhong, Karberg and Lambowitz, 2003), it required further optimisation to improve colony isolation and screening. The strategy adopted a retro-transposition activating marker (RAM) system within the group II intron (Heap *et al.*, 2007) which incorporated an antibiotic cassette *ermB* conferring erythromycin resistance and vastly improved isolation of integrant mutants. Moreover, the ClosTron's utility was further improved with the addition of FLP and FRT (FLP recognition target) features into the ClosTron plasmids (Heap *et al.*, 2010). The FLP enables the introduction of up to 1 kB heterologous cargo into the intron site and FRT allows for the removal of the RAM resistance cassette which enables further insertional mutations to be made (Heap *et al.*, 2010). The ClosTron has been implemented in a range of *Clostridium* hosts, not limited to; *C. acetobutylicum, C. difficile, C. beijerinckii, C. botulinum*, *C. sporogenes, C. autoethanogenum* and *C. ljungdahlii* (Shao *et al.*, 2007; Cooksley *et al.*, 2010; Heap *et*

al., 2010; Köpke, Mihalcea, Liew, *et al.*, 2011; Steiner *et al.*, 2012; Liew *et al.*, 2017). However, there are several disadvantages in this system, as traditionally Southern Blot techniques are necessary for the validation of correct insertion. Moreover, an inherent possibility of unwanted polar effects also exists which may hinder efforts to characterise desired mutations (Kuehne and Minton, 2012). The program to design ClosTron insertions is readily available on the website: http://clostron.com/ (last accessed 26/09/17), and rapid procurement of ClosTron plasmids can be done through a DNA synthesis services.

1.7.2 Allelic-exchange methods

A novel gene manipulating technique termed "Allelic-coupled Exchange" (ACE) allows the possibility of precisely editing genes through use of homologously recombining allelic-exchange vectors able to either cleanly delete or insert genes of varying sizes into target hosts (Ng *et al.*, 2013). Gene deletions are done "in-frame" meaning inclusive of whole gene codon triplets as to prevent interference with downstream open reading frames. Conversely, introduction of heterologous cargo is possible with this method (Heap *et al.*, 2012; Ng *et al.*, 2013) with extreme examples demonstrating the iterative introduction of a 48.5 kB phage lambda genome into *C. acetobutylicum* (Heap *et al.*, 2012).

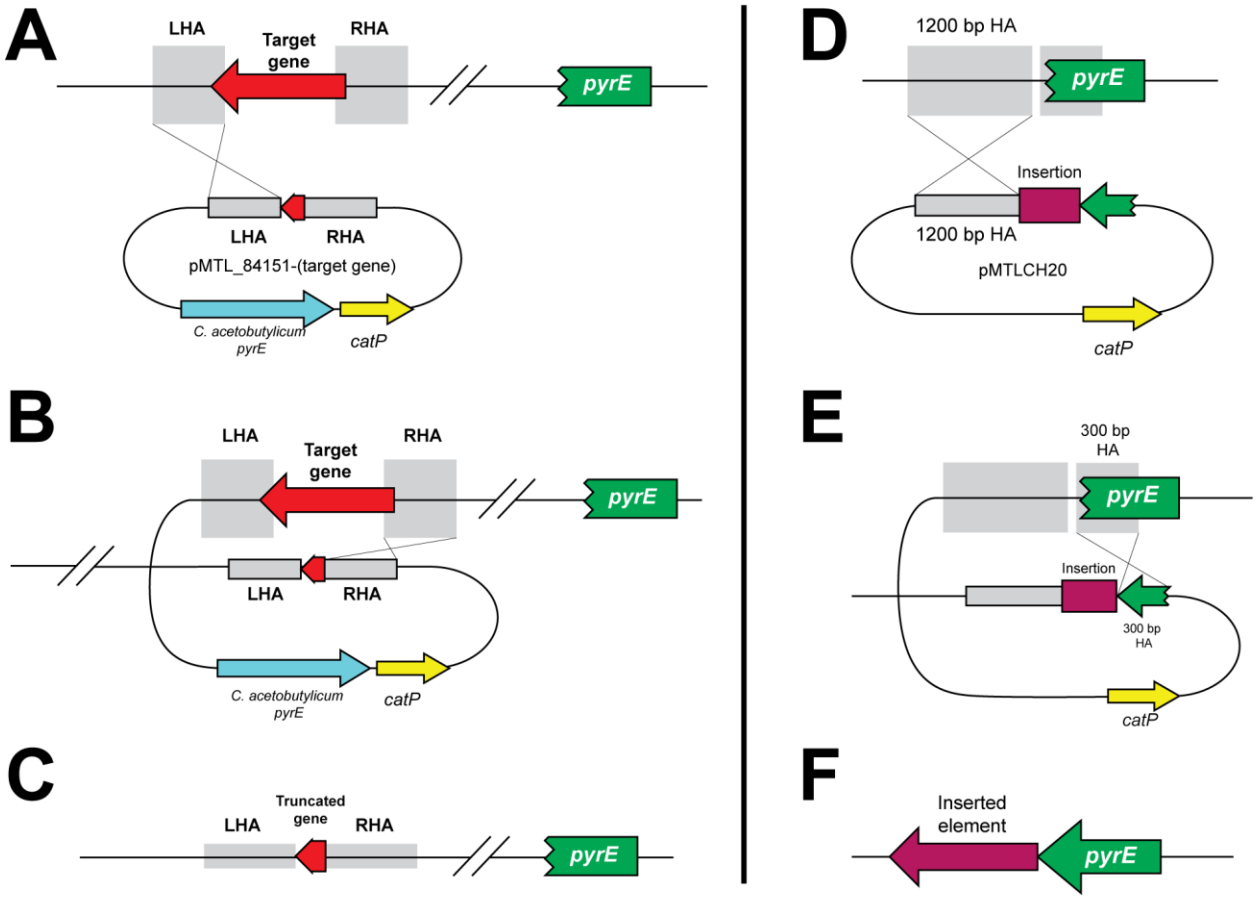
Despite these significant developments, implementation of the ACE techniques comes at a price. In contrast to ClosTron, ACE requires several recombination events for gene deletion/insertion, followed by stringent screening and mutant isolation. Moreover, the process is reliant on negative selection markers and introduction of ACE plasmids into a desired host requires established transformation techniques.

1.7.2.1 *pyrE* as a negative-selection marker

A key component of ACE is the usage of a counter-selectable marker which enables selection for rare recombination events in single or double-cross mutants (Ng *et al.*, 2013). Counter-selectable markers may come in varied forms that target biosynthesis genes essential to growth or may confer toxicity if expressed, in either case these markers have been exploited during the selection processes.

In *Clostridium* spp., the *pyrE* gene encodes orotate phosphoribosyl transferase, responsible for the *de novo* synthesis of the pyrimidine base, uracil (Ghim, Nielsen and Neuhard, 1994). Disruption of the *pyrE* gene results in uracil auxotrophy, which is lethal

if growth medium is not suitably supplemented. However, a lack of a functional *pyrE* gene results in the concomitant resistance to 5-fluoroorotic acid (FOA). FOA is analogous to the pyrimidine intermediate orotic acid but metabolism of FOA leads to the accumulation of toxic 5-fluorouridine monophosphate (5-FUMP) that causes cell death (Boeke *et al.*, 1987). As a result, a *pyrE*-deficient mutant confers uracil auxotrophy and resistance to FOA and is invaluable in ACE selection procedures.



Figure, 1.4, Overview of Allelic-Coupled Exchange (ACE) in *C. autoethanogenum*: (A) In-frame deletion plasmid pMTL_84151 contains left and right homology arms, LHA and RHA, respectively, that encompass a target gene. The plasmid is transformed into a *pyrE*-deficient *C. autoethanogenum* strain. Thiamphenicol resistance conferred by catP selects for transformed colonies. Fast colonies growing are assumed to be single-cross mutants. (B) Single-cross mutants are subjected to FOA treatment and screening for uracil auxotrophy which is attained after the double-cross event. This event also indicates plasmid loss. (C) After counter-selection on FOA and through uracil auxotrophy screening, the double-crossover mutant now possesses an in-frame truncated, non-functional gene. (D) For *pyrE* restoration and/or mutant complementation, pMTLCH20 is transformed in the newly created mutant under thiamphenicol selection. Asymmetric homology arms (HA) of 1200 and 300 bp within the plasmid ensure that the longer HA is recombined first. (E) A second recombination event occurs with the shorter HA and double-cross mutants are screened for uracil auxotrophy. (F) Mutants are selected based on uracil prototrophy and further PCR screened. The inserted element can present a gene module or a non-functional LacZα fragment, alongside a fully restored pyrE locus

It is worth noting that alternative selection markers such as cytosine deaminase *codA* and endoribonuclease toxin *mazF* have been shown to function in *Clostridia* (Al-Hinai, Fast and Papoutsakis, 2012; Cartman *et al.*, 2012; Ehsaan *et al.*, 2016). As *codA* is exceptionally rare in *Clostridia* and *mazF* is only present in *E. coli*, WT *Clostridium* strains with *pyrE* can be used directly. Like *pyrE*, *codA* functions in uracil biosynthesis but through cytosine to uracil conversion. By the same token, *codA* produces a toxic 5-fluorouracil compound when converting from 5-fluorocytosine (Cartman *et al.*, 2012), but relies on uracil phosphoribosyltransferase (*upp*) gene to confer toxicity, which may not always be present in hosts. Interestingly, *codA* counter-selection has been explored in *C. autoethanogenum* and deemed to be an applicable marker (Liew, 2015).

Counter-selectable marker *mazF* encoding an mRNA interferase targets the mRNA 5'-ACA sequence resulting in slow growth leading to cell death (Al-Hinai, Fast and Papoutsakis, 2012). A vector with *mazF* is coupled with a lactose-inducible promoter and a FRT-flanked thiamphenicol resistance cassette sits between target gene homology arms. After vector recombination into the chromosome, the thiamphenicol resistance cassette substitutes the target gene but is later removed using the Flp/FRT system. Screening on media supplemented with lactose and thiamphenicol validates whether plasmid loss occurred and whether the double-cross over event took place (Al-Hinai, Fast and Papoutsakis, 2012).

Overall, it is still preferable to use a *pyrE* selection marker due to the straight-forward gene complementation methods required to validate mutant phenotypes (Minton *et al.*, 2016).

1.7.2.2 Implementation of ACE

Following the route of *pyrE*-based allelic-coupled exchange, a *pyrE*-deficient mutant background is required to perform desired deletion or insertions (Heap *et al.*, 2012). Targeting a gene for deletion first requires the design of a vector that contains a pair of homology arms flanking the target gene. Both arms are spliced together and ligated into a vector that contains an unstable replicon, a resistance cassette and an exogenous copy of *pyrE* (Ng *et al.*, 2013; Ehsaan, Kuehne and Minton, 2016). After plasmid transformation into the host, conferred resistance allows colonies to propagate on selection media. In addition to antibiotic resistance, large fast-growing colonies appearing infer a single-cross plasmid to chromosome integration event. Confirmed single cross mutants are screened

for on FOA/uracil-supplemented media to select for second recombination. This event may either excise the desired deletion fragment out of the chromosome resulting in WT genotype reversion, or preferably the plasmid containing the resistance cassette and exogenous *pyrE* are excised, Figure 1.4.A-C.

Colonies that able to resist FOA are negatively selected against using replica plating techniques with and without uracil, to confirm plasmid loss. Final screening procedures employ PCR and sequencing to determine any WT-revertant contaminants (Heap *et al.*, 2012; Ng *et al.*, 2013; Ehsaan, Kuehne and Minton, 2016).

The final step in mutant completion is restoration of the truncated *pyrE* gene which is performed using a restoration vector with asymmetric homology arms which flank *pyrE*. After transformation, colonies are rapidly screened through streaking onto minimal medium lacking uracil. Propagated colonies suggest successful *pyrE* restoration, (Heap *et al.*, 2012; Ng *et al.*, 2013; Ehsaan, Kuehne and Minton, 2016). Moreover, the introduction of heterologous cargo can be done through ligation onto *pyrE* restoration vectors and following similar procedures described, Figure 1.4.D-F.

1.7.4 Transposon mutagenesis

A new paradigm of *Clostridium* spp. genetic modification entails the usage of "forward genetics" tools which in contrast to above "reverse genetics" techniques, aims to identify mutant genotypes from resultant phenotypes produced through transposable elements, (Zhang, Grosse-Honebrink and Minton, 2015). The Himar1 mariner transposon on which many transposon systems are based on, effectively "jumps" across a host's genome and consequently produces random gene disruptions (Cartman and Minton, 2010). The advantage of this method is the procurement of large random mutant libraries that can be screened for desired phenotypes. Despite this prospect, the system requires fine tuning as to how and when transposition can occur in efforts to avoid multiple insertion events. Solutions have included coupling the expression of transposase with *C. difficile* toxin B promoter (P_{tcdB}) which is exclusively activated through sigma factor TcdR. However, for this system to function in *Clostridium* species lacking TcdR, an exogenous copy must be introduced via ACE complementation. This has been successfully achieved in *C. sporogenes* (Zhang, Grosse-Honebrink and Minton, 2015). Further optimisations of transposon mutagenesis plasmids have been through inducible IPTG promoters placed

upstream of unstable replicons which encourage a high chance of plasmid loss when in contact with IPTG. This strategy reduces the chances of multiple transpositions (Zhang, Grosse-Honebrink and Minton, 2015; Minton *et al.*, 2016).

1.7.5 CRISPR genome editing

CRISPR/Cas9 has been widely regarded as a revolutionary genome engineering tool that is capable of precise gene editing with exceptionally high efficiency in prokaryotic and eukaryotic hosts alike (Hwang *et al.*, 2013; Hsu, Lander and Zhang, 2014). CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats), is the archaeal and bacterial answer to an adaptive immune system which protects primarily against exogenous nucleic acids and bacteriophages (Sorek, Lawrence and Wiedenheft, 2013). The adaptive CRISPR system works in tandem with Cas (CRISPR-associated) proteins and the process begins with the Cas complex acquiring foreign DNA and in incorporating it as a spacer into a CRISPR loci array. Newly incorporated DNA is flanked by repeats, which are continually added into this array as foreign DNA is recognised. The CRISPR array is transcribed into pre-CRISPR RNA / guide RNA and subsequently processed into individual CRISPR RNAs. These crRNA units bind to the complementary foreign DNA sequence that directs Cas proteins to nucleolytically cleave in a binding complex. This results in degraded foreign DNA, (Sorek, Lawrence and Wiedenheft, 2013; Hsu, Lander and Zhang, 2014).

CRISPR/Cas systems are characterised in types I-III, which are defined by the number of Cas effector module subunits and modes of action. Type II was found to be particularly relevant due to the requirement of only a single effector module, Cas9, which streamlines the application of this technique greatly (Hsu, Lander and Zhang, 2014). By exploiting the CRISPR/Cas9 system for gene editing, it is possible to design guide RNA to target genes for knockouts and insertion of heterologous genes (Ran *et al.*, 2013). Moreover, isolation of pure mutants with minimal labour is possible as CRISPR/Cas9 acts as a powerful counter-selection maker leading to high efficiencies of double-cross over events.

The CRISPR/Cas9 system has been so far been successfully applied to *Clostridium* spp. Studies were performed in *C. beijerinckii*, *C. ljungdahlii*, *C. autoethanogenum*, *C. cellulolyticum* and *C. pasteurianum*, with varying levels of success ranging between 25 – 100 % in genes deletion efficiencies (Richter *et al.*, 2012; Brown *et al.*, 2014; Wang *et al.*, 2015; Xu *et al.*, 2015; Nagaraju *et al.*, 2016; Pyne *et al.*, 2016). Further optimisations are

still needed for *Clostridium* spp., but from the onset this tool offers an innovative approach in gene editing.

1.8 Aims of this study

The primary objective of this project was to characterise the Agr quorum sensing system in industrially-relevant acetogen, *Clostridium autoethanogenum*. The seemingly dissimilar topics covered in this chapter have formed the motivation in pursuing this primary goal. Although it has been thoroughly shown in literature that industrially-relevant acetogens hold a possible to answer to solving energy demands in an environmentally conscious way, many questions still lay unanswered.

Quorum sensing research has been chiefly confined to pathogens but is now beginning to show its significance in industrially important solventogens. The Agr QS system which is conserved across many *Firmicutes*, plays an integral role in sporulation, toxin and biofilm production in pathogens. However pivotal work has shown Agr's function in *C. acetobutylicum*, can similar discoveries be made in *C. autoethanogenum*?

This study has three key aims which can be summarised as follows:

I. Establishing a robust method in targeting and disrupting key Agr genes with the application of proven in-frame deletion techniques that have been exemplified in other *Clostridium* spp. Several obstacles are needed to be overcome which include conjugation and transformation inefficiencies, as well as previously unknown restriction systems. Furthermore, in-depth analysis of constructed mutants is necessary to determine the extent of single nucleotide variants, and whether they have any phenotypic bearing on constructed mutants.

II. Phenotypic characterisation of *C. autoethanogenum* Agr mutants through growth experiments and assays aiming to explore every possible aspect it may show in contrast to the WT strain. Findings from these experiments will lay the foundation of further work aiming to verify Agr-dependent changes.

III. The final aim will be to employ the use of advanced proteomic analysis to discern protein-level changes within mutants with respect to the WT. Apparent differences will be individually examined using enzyme activity and other assays, as required.

Each aim will lead to a conclusion which will discuss the plausible reasons as to why the Agr system is not only present, but beneficial to *C. autoethanogenum*.

Chapter 2

Materials and Methods

2.1 Bacterial strains and plasmids

2.1.1 Bacterial strains

All strains used listed in this project are listed in Table 2.1. Strains of independently produced mutants used in this study are also listed.

Table 2.1, **Strains created and used in this study:** C. a. = *C. autoethanogenum*, and E. c. = *E. coli*; DSMZ = Deutsche Sammlung von Mikroorganismen und Zellkulturen (Braunschweig, Germany), ATCC = American Type Culture Collection.

Strains	Description	Source
C. a. JA1-1 DSM 10061	Wild type strain	DSMZ
C. a. ΔpyrE	*pyrE* gene deficient strain for allele coupled exchange	This study
C. a. pyrE^{fix}	Previously *pyrE* gene deficient strain with restored *pyrE* gene	This study
C. a. ΔagrD1	*agrD1* deficient strain	This study
C. a. ΔagrD2	*agrD2* deficient strain	This study
C. a. ΔagrD1D2	*agrD1* and *agrD2* deficient strain	This study
C. a. ΔagrD1D2_16	*agrD1* and *agrD2* deficient strain; independently obtained	This study
C. a. ΔagrD1D2_34	*agrD1* and *agrD2* deficient strain; independently obtained	This study
C. a. ΔagrD1 ΔpyrE	*agrD1* and *pyrE* truncated strain	This study
C. a. ΔagrD2 ΔpyrE	*agrD1* and *pyrE* truncated strain	This study
C. a. ΔagrD1D2 ΔpyrE	*agrD1*, *agrD2* deficient strain and *pyrE* truncated strain	This study
C. a. ΔagrD1D2_16 ΔpyrE	*agrD1*, *agrD2* deficient strain and *pyrE* truncated strain	This study
C. a. ΔagrD1D2_34 ΔpyrE	*agrD1*, *agrD2* deficient strain and *pyrE* truncated strain	This study
C. a. ΔagrD1D2_D1^{comp}	Double knockout mutant with *agrD1* complementation down-stream of *pyrE*	This study
C. a. ΔagrD1D2_D2^{comp}	Double knockout mutant with *agrD2* complementation down-stream of *pyrE*	This study
C. a. agrB::CtermB-234a	QS gene *agrB* disrupted via ClosTron	This study

E. c. One Shot® TOP10	General cloning chemically-competent *E. coli* strain	Thermo-fischer
E. c. CA434	*E. coli* HB101 strain carrying R702 conjugative plasmid, used for all conjugations	Dr Chris Humphreys (Uni. Of Nottingham)
E. c. "Sexpress"	NEB Express *E. coli* carrying R702 conjugative plasmid	Craig Woods (Uni. Of Nottingham)
C. acetobutylicum ATCC 824	Control wild type strain	ATCC / Oliver Severn
C. ljungdahlii DSM 13582	Control wild type strain	DSMZ (Wan Ho)

2.1.2 Plasmids

Plasmids that were constructed or obtained in this study are listed in Table 2.2. ClosTron plasmids were designed through using the "Intron targeting and design webtool" (Heap *et al.*, 2007) and ordered through DNA 2.0 Gene Synthesis service (California, USA), (https://www.atum.bio/).

Table 2.2, Plasmids created and used in this study: Table lists final versions of plasmids.

Plasmid	Description	Source
pMTL84151_ΔpyrE	Used in creation of the *C. autoethanogenum* Δ*pyrE* strain	Anne Henstra (Uni. Of Nottingham)
pMTLCH20	*pyrE* correction cassette and gene knock-in plasmid	Dr Chris Humphreys (Uni. Of Nottingham)
pMTL84151_0816	In-frame deletion plasmid for *agrD1*	This study
pMTL84151_3094	In-frame deletion plasmid for *agrD2*	This study
pMTLCH20-D1	Chromosomal complementation plasmid of *agrD1*	This study
pMTLCH20-D2	Chromosomal complementation plasmid of *agrD2*	This study
pMTL007C-E2::CAUagrB-234a	*agrB* CLOSTRON plasmid anti-sense	This study
pMTL007C-E2::CAUagrB-232s	*agrB* CLOSTRON plasmid sense	This study

2.2 General growth methods

2.2.1 Antibiotics

Antibiotics used in this thesis are detailed in Table 2.3, all references to antibiotics in this thesis refer to the stock and working concentration values detailed in this table**.**

Table 2.3, Antibiotics used in this project: C. a. = *C. autoethanogenum*, and E. c. = *E. coli*

Antibiotic	Abbreviation	Solvent	Stock Conc. (mg/mL)	Working Conc. (µg/mL)	Used in
Ampicillin	Amp	dH_2O	50	100	*E. c.*
Chloramphenicol	Cm	100 % EtOH	25	25	*E. c.*
Clarithromycin	Cla	100 % DMSO	6	6	*C. a.*
D-cycloserine	D-cyc	dH_2O	50	250	*C. a.*
Spectinomycin	Spec	dH_2O	100	100	*C. a./E.c.*
Thiamphenicol	ThPh	100 % DMSO	15	7.5	*C. a.*

2.2.2 Aerobic media and growth conditions

For general *E. coli* work, Lysogeny broth (LB) (Sigma Aldrich), was used in liquid and in 1.5% solid agar (Luria, Adams and Ting, 1960). Solid agar LB plates were incubated at 37 °C, and liquid cultures shaken at 225 RPM at 37 °C.

2.2.3 Anaerobic media and growth conditions

General work with *C. autoethanogenum* was performed in an Don Whitley MG 1000 anaerobic workbench (Don Whitley, Yorkshire, UK) at 37 °C and 55 – 65 % humidity with a 80:10:10 gas atmosphere ratio composed of N_2:H_2:CO_2, respectively. Work was also carried out in an anaerobic chamber, (Coy Laboratory Products), referred to as a "COY tent" or "anaerobic tent" hereinafter. The COY tent had a similar gas atmosphere ratio as Don Whitley anaerobic workbenches but was not humidity or temperature controlled and was operated at room temperature. Research grade gases were used in either anaerobic workstations (CO, CO_2, H_2, and N_2), purchased from BOC (BOC, UK).

Table 2.4, Composition of media used in general *C. autoethanogenum* culturing: All component quantities are added into dH_2O for 1 L of media. Ingredients were added in order as listed below. Agar was added into individual bottles after all components were dissolved.

Medium component	YTF	PETC	SDMM
Yeast Extract	10 g	-	-
Tryptone	16 g	-	-
Fructose	10 g	-	7.2 g
Sodium Chloride	0.2 g	-	-
MES	-	20 g	20 g
Casein Amino Acids	-	1.5 g	10 g
Sodium Formate	-	-	1 g
5 M NaOH	-	-	2 mL
Acid Stock Soln. (1000x)	1 mL	-	1 mL
Base Stock Soln. (1000x)	1 mL	-	1 mL
Wolfe's Vitamin Soln. (1000x)	1 mL	1 mL	1 mL
Macronutrient Soln. (50x)	-	20 mL	20 mL
Trace Metal Soln. (100x)	-	10 mL	-
Resazurin (0.1% Stock)	1 mL	1 mL	1 mL
Agar	15 g	15 g	15 g
Media buffered to pH 5.8			

Culturing of *C. autoethanogenum* employed the use of 3 different media dependent on requirement. Yeast-tryptone-fructose (YTF) medium was used for general culturing and semi-defined minimal medium (SDMM) was employed for *pyrE* deficient and uracil auxotroph mutant screening. Semi-defined PETC 1754 medium (American Type Culture Collection, ATCC) was predominantly used in serum flask experiments and typically the carbon source was added separately as per experiment requirement. Compositions for all 3 media are listed in Table 2.4.

Table 2.4, Composition of media used in general C. autoethanogenum culturing

Supplements to media listed were filter sterilised, and comprised of the following:

Acid stock solution (1000x): For 1 L; 50 mM HCl, 100 mg H_3BO_3, 230 mg $MnCl_2{\cdot}4H_2O$, 780 mg $FeCl_2{\cdot}4H_2O$, 103 mg $CoCl_2{\cdot}6H_2O$, 602 mg $NiCl_2{\cdot}6H_2O$, 78 mg $ZnCl_2$, 50 mg $CuSO_4{\cdot}5H_2O$ and 50 mg $AlK(SO_4)_2{\cdot}12H_2O$.

Base stock solution (1000x): For 1 L; following: 10 mM NaOH, 58 mg Na_2SeO_3, 53 mg Na_2WO4 and 52 mg $Na_2MoO_4{\cdot}2H_2O$.

Wolfe's vitamin solution (1000x): For 1 L; 50 mg p-aminobenzoate, 50 mg riboflavin, 50 mg thiamine HCl, 50 mg nicotinate, 100 mg pyridoxine hydrochloride, 50 mg calcium pantothenate, 1 mg vitamin B_{12}, 20 mg biotin, 20 mg folic acid and 50 mg thiotic acid. The final pH adjusted to 7.8.

Macronutrient solution (50x): For 1 L; a 50x stock solution per 1 L: 2 g nitrilotriacetic acid, 50 g NH_4Cl, 40 g NaCl, 5 g KCl, 10 g KH_2PO_4, 10 g $MgSO_4 \cdot 7H_2O$, 2 g $CaCl_2 \cdot 2H_2O$, 2.4 g Fe $(SO_4)_2(NH_4)_2 \cdot 6H_2O$.

Trace metal solution (100x): For 1 L; 1 g $MnSO_4 \cdot H_2O$, 0.8 g Fe $(SO_4)_2(NH_4)_2 \cdot 6H_2O$, 0.2 g $CoCl_2 \cdot 6H_2O$, 0.2 mg $ZnSO_4 \cdot 7H_2O$, 20 mg $CuCl_2 \cdot 2H_2O$, 20 mg $NaMoO_4 \cdot 2H_2O$, 20 mg Na_2SeO_3, 20 mg $NiCl_2 \cdot 6H_2O$, 20 mg $Na_2WO_4 \cdot 2H_2O$.

Further additions included 200x (4% w/v) stock solution of L-cysteine, prepared with anaerobic distilled water and added to the medium post-autoclaving. A 1000x (0.1 % w/v) of resazurin dye was added to aid in determining acidity and anaerobicity of media.

2.2.3.1 Preparation of culture vessels

Experiments performed outside of the anaerobic cabinet required the use of sterile, 250 mL sealable serum flasks. Experiments predominantly used PETC medium after which it was made, was boiled then sparged with N_2 gas while cooled in an ice-bath to remove oxygen. The sparged medium containing no carbon source, was placed into the COY tent for 3-4 days to allow further oxygen loss. Before dispensing media, serum flasks would be placed into the anaerobic tent overnight to allow residual oxygen to escape. Serum flasks would be filled with PETC medium, sealed and autoclaved. After autoclaving, 1 mL of 4 % w/v L-cysteine was injected into each flask for further medium reduction, characterised by a medium colour change from pink to clear. Flasks were stored at 4 °C until needed. For experiments on sugar, a 3 M concentrated D-fructose stock solution was prepared anaerobically, sterile filtered and added to 90 mL PETC media after autoclaving to a final concentration of 50 mM.

Experiments performed in smaller volumes used 25 mL Hungate tubes with 15 mL PETC medium, containing 50 mM D-fructose. Tubes were prepared similarly as 250 mL serum flasks and used identical reducing agents and sampling methods.

Experiments involving growth on C1 gas substrates used serum flasks containing 50 mL of PETC which were dispensed with either 150 kPa (1.5 bar) of 100 % CO, or 200 kPa (2

bar) of 80% H_2, 10% CO_2 and 10% N_2. This procedure was carried out with the use of an anaerobe gas-exchange system (Anaerobe Gaswisselsysteem), model GW6400 (GR Instruments, Netherlands). Flasks were incubated statically at 37 °C and placed horizontally to maximise gas-liquid surface area. A DPG 110 hand-held pressure gauge (Omega Engineering, USA) was used to determine serum flask gas headspace pressures.

2.2.4 Storage

E. coli competent cell seed stocks were stored in 10 % glycerol in -80 °C. Preservation of plasmid stocks involved transformation into an *E. coli* host and growth on selective LB plates. Colonies were scraped from plates using a 10 µL loop and resuspended in a Microbank™ microbead vial (Pro-Lab Diagnostics, UK) and stored at -80 °C.

Storage of *C. autoethanogenum* was performed through culturing in either YTF or PETC medium for 3 days. 900 µL was pipetted into a 2 mL cryovial containing 100 µL DMSO. Stocks were placed at -80 °C.

2.3 General molecular biology techniques

2.3.1 Genomic DNA extraction and purification

Genomic DNA (gDNA) of high concentration and quality was required for next-generation Illumina sequencing analysis. This gDNA was extracted using adapted methods (Bouillaut, McBride and Sorg, 2011) and commercial kits were employed for all other routine gDNA work.

2.3.1.1 Phenol chloroform extraction

Samples of *C. autoethanogenum* were cultured in 5 mL of YTF for 3 days before centrifugation at 6,000 rpm for 2 min. The cell pellet was resuspended in 180 µL of lysozyme digestion buffer (25 mM Tris-HCl, pH 8.0, 2.5 mM EDTA, 1% Triton X-100), containing 20 mg/mL of lysozyme VI (MP Biomedicals), for 30 min at 37 °C. Following incubation, 25 µL Proteinase K (Sigma Aldrich), 85 µL of dH_2O and 110 µL of 10% SDS (Sigma Aldrich) were added to the lysed samples and mixed by inversion, then incubated at 55 °C for 30 min. After incubation, 400 µL of phenol chloroform (or equal volumes to sample volume), were pipetted into a 2 mL Manual Phase Lock Gel™ Heavy tube (5 Prime, UK), the tube contents were inverted for approx. 15 sec before centrifugation for 3 min

at maximum speed. The clear top layer of the sample was transferred into a fresh 1.5 mL Eppendorf tube and 5 µL of RNAse A (Sigma Aldrich) was added and the sample was incubated at room temperature for 10 min.

The sample was placed into a fresh Phase Lock tube containing 400 µL of phenol chloroform and was mixed by inversion then centrifuged at max speed for 3 min. This step was repeated twice and after the final chloroform wash the top layer was placed into a 1.5 mL tube containing pre-chilled (-80 °C) 3 M sodium acetate and 100 % ethanol. The mixture was inverted and placed back into the -80 °C freezer for 30 min.

Samples were ethanol precipitated by first thawing to room temperature and then centrifuging at max speed for 15 min. The supernatant was removed, leaving a visible pellet. 1 mL of 70 % ethanol was added and mixed by inversion, and centrifuged at max speed for 3 min. The samples were air dried at room temperature for 45 min. The dry gDNA pellet was resuspended in 50 µL of sterile, nuclease-free dH_2O. Samples that underwent Illumina next-generation sequencing (NGS) were required to be eluted in elution buffer (10 mM Tris-Cl, pH 8.5).

2.3.1.2 Conventional gDNA extraction

Routine genomic DNA extraction predominantly used two different commercial kits; GenElute Bacterial Genomic DNA Kit (Sigma Aldrich) and PureLink Genomic DNA extraction (Invitrogen/ThermoFischer). Protocols were followed as per manufacturer's instructions; genomic DNA was eluted in sterile, 50 µL nuclease-free dH_2O and stored at -20 °C.

2.3.1.3 Analysis of genomic DNA

Phenol-chloroform extracted samples were analysed using the Qubit 3.0 fluorometer (Life Technologies, Invitrogen), using 500 µL Qubit assay tubes and Qubit high sensitivity reagents. Procedure followed manufacture protocol. Samples were diluted from a range of 10^0 to 10^{-3} for accurate measurement. DNA was run on a 1 % agarose gel to allow examination of quality and RNA impurities.

Routinely extracted gDNA samples were analysed using a NanoDrop Lite (Thermo Scientific, UK). 1 µL of sample was measured at 260 nM to determine concentration and at the 260/280 nm for sample purity. Methods followed manufacturer's instruction.

2.3.2 Polymerase chain reaction methods

2.3.2.1 PCR Polymerases and Programs

Routine and analytical PCR procedures used DreamTaq Green DNA polymerase (Thermo Scientific, UK), as per manufacturer's instruction in 20 µL reaction volumes, Table 2.5 details the PCR program used.

Table 2.5, DreamTaq PCR program

Step	Temp (°C)	Time	
Initial denaturation	95	2 min	
Denaturation	95	10 sec	
Annealing	X	30 sec	28 cycles
Elongation	72	1 min per kB	
Final elongation	72	10 min	
Storage	15	∞	

For proof-read amplification of DNA fragments, for short-length sequencing and cloning, Phusion High-Fidelity (HF) DNA Polymerase (NEB, UK) was used in 50 µL reaction volumes as per manufacturer's instruction. Table 2.6 details the PCR program used.

Table 2.6, Phusion HF PCR program

Step	Temp (°C)	Time	
Initial denaturation	98	2 min	
Denaturation	98	10 sec	
Annealing	X	30 sec	30 cycles
Elongation	72	30 sec per kB	
Final elongation	72	10 min	
Storage	15	∞	

Annealing temperatures of primers were calculated using Oligo-Evaluator™ (Sigma Aldrich. All PCR reactions were performed using a SensoQuest PCR machine (SensoQuest GmbH, Germany).

2.3.2.2 Primers

Primers were ordered and synthesized from Sigma Aldrich or MWG Eurofins. Synthesis parameters were at a 0.025 µmol scale and purified through desalting. Full primer list can be found in the Appendix section 7.1.

2.3.3 Application of PCR techniques

2.3.3.1 Colony PCR

Colony PCR was performed by picking single colonies and resuspending in 10 µL of distilled water. This resuspension mixture was used as a 1 µL template in a 20 µL reaction mixture using DreamTaq polymerase and PCR program (Table 2.5). This method was used routinely in *E. coli* and *C. autoethanogenum* colony screening.

2.3.3.2 Splicing by Overlap Extension PCR

Construction of in-frame deletion plasmids relied on Splicing by Overlap Extension (SOE) PCR, and used Phusion HF Polymerase in a 50 µL reaction volumes. A set of four primers were designed, two pairs to amplify two separate right and left homology arms flanking the target gene. The right and left homology arms were used as templates and "spliced" together using the forward primer of the left homology arm, and reverse primer of the right homology arm. Primers were designed with an annealing temperature difference of ± 3 °C, thus improving PCR reaction efficiency.

2.3.4 Agarose gel electrophoresis

Routine PCR reactions were analysed using agarose gel electrophoreses. Standard procedures used a 1 % Tris-acetate-EDTA (TAE) agarose gel which when melted, was mixed with 5 - 10 µL of SYBR® Safe DNA stain (Thermo Fisher Scientific, UK). Routine gels were run at a 120 V, 400 A current for 20 – 25 min and visualised using a Syngene UltraBright-LED blue light transilluminator (Syngene, UK) or Licor Odyssey Fc Imaging System (Licor, USA).

2.4 Plasmid construction and cloning

2.4.1 Plasmid purification

Plasmid purification from *E. coli* was performed using GenElute™ Plasmid Miniprep kits (Sigma Aldrich, UK), as per manufacturers instruction, eluted in 50 µL distilled water. For enhanced plasmid concentration yields, elution water was heated to 50 °C and aliquoted into each spin column and allowed to incubate at room temperature for 5-10 min before centrifugation at maximum speed for 3 min.

Plasmid extraction carried out on *C. autoethanogenum* was carried out similarly, but with an added lysis step which followed similar procedures as gDNA extraction; pellet was resuspended in 20 mg/mL of lysozyme VI (MP Biomedicals), and incubated for 30 min at 37 °C. The lysed suspension was treated as per manufacturers instruction using using GenElute™ Plasmid Miniprep kits (Sigma Aldrich, UK). Plasmid were eluted in 50 µL distilled water.

2.4.2 DNA restriction digestion and purification

CutSmart™ Restriction Enzymes (NEB, UK) were used in all DNA endonuclease digestions. All methods followed manufacturer's protocols and were typically performed in 50 µL reaction volumes using 1 µg of DNA and incubated at a longer duration of 1 hour at 37 °C. Digested fragments were separated using agarose gel electrophoresis.

DNA fragments cut out from agarose gels were purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research, USA). DNA fragments obtained from PCR amplification, were purified using QIAquick PCR Purification kit (Qiagen, UK). All purified DNA fragments were eluted in 10 µL dH_2O, and concentrations measured using NanoDrop Lite (Thermo Scientific, UK).

2.4.3 Dephosphorylation and ligation

DNA fragments were dephosphorylated using FastAP Thermosensitive Alkaline Phosphatase (Thermo Fischer Scientific, UK), and ligation was performed in a 20 µL reaction volume using 1 µL of T4 DNA Ligase (NEB, UK), 2 µL of 10x ligation buffer and a ratio of 3:1, insert to plasmid, into a final volume of 20 µL. This reaction mixture was incubated at room temperature for 1 hour.

2.4.4 Construction of in-frame deletion plasmids

Plasmid pMTL84151_ΔpyrE was donated by Dr Anne Henstra (Uni. Of Nottingham) and provided the framework for in-frame deletion (IFD) plasmid creation in this project. pMTL84151_ΔpyrE follows the pMTL80000-series Modular Plasmid System (Heap 2009), and targets the *pyrE* gene locus for in-frame truncation, through an asymmetrical homology arm design, Figure 2.1.A. Typically in IFD plasmids, both homology arms are between 600 – 750 bp in length, but in the case of pMTL84151_ΔpyrE, the short homology arm (SHA) is 303 bp and great homology arm (GHA) is 1.2 kB, with a lacZα fragment separating them which recombines at the 3' end of the WT *pyrE* gene. For allelic-coupled exchange techniques, a non-homologous *C. acetobutylicum pyrE* gene and ribosomal binding site is included in pMTL84151_ΔpyrE that serves as a counter-selection marker when screening on 5-fluoroorotic acid (FOA). The catP cassette confers chloramphenicol and thiamphenicol antibiotic resistance.

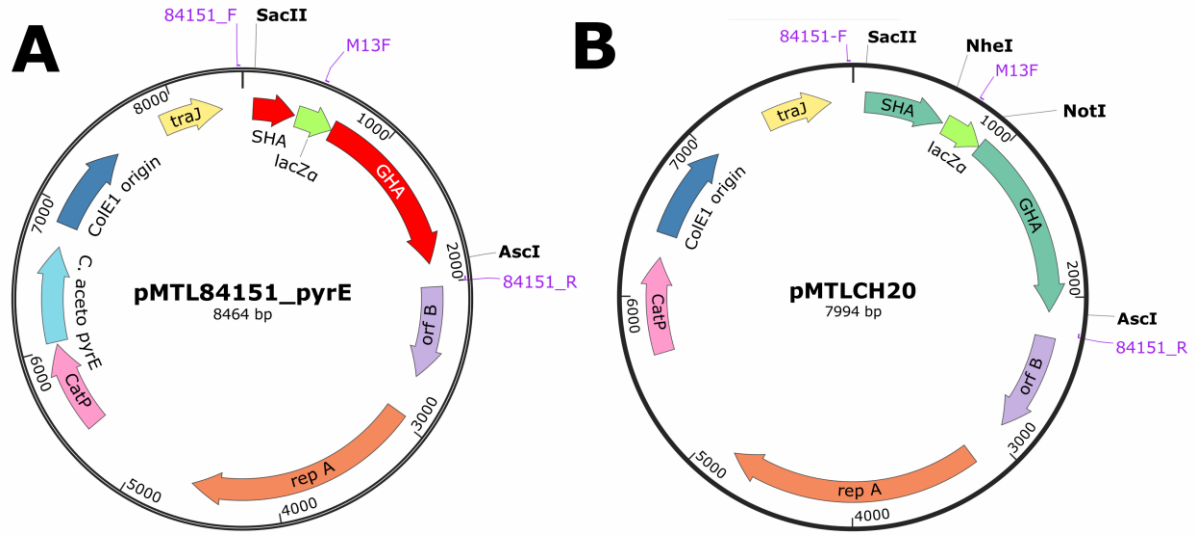


Figure 2.1, Modular pMTL plasmids employed in gene deletion and repair: pMTL84151_ΔpyrE (A), used in *pyrE* deletion and later modified to host IFD fragments. Repair plasmid pMTLCH20 (B), used in *pyrE* restoration, and modified to host *agrD* complementation fragments. Primers 84151_F and R, and M13F, were used in PCR screening and Sanger sequencing procedures.

The multi-cloning site in pMTL84151_ΔpyrE is flanked by restriction sites, SacII and AscI that release a 1.8 kB insert containing the asymmetric *pyrE* homology arms when digested. The pMTL84151 backbone was used in the ligation of IFD inserts that target *agrD1* and *agrD2* genes. Both inserts possessed homology arms of roughly equal size and the IFD regions were designed to truncate up to 80% of each *agr* gene. With this strategy

in mind, plasmids pMTL84151_Δ0816 and pMTL84151_Δ3094 were constructed which followed amplification of the right and left homology arms for *agrD1* and *agrD2*, using the homology arm primers listed in Appendix, Table 7.1. Homology arms were combined by SOE PCR as described in 2.3.3.2, and ligated into plasmid backbone of pMTL84151.

2.4.5 PyrE restoration and complementation plasmids

The *pyrE*-restoration plasmid pMTLCH20 resembles pMTL84151_ΔpyrE but lacks a *C. acetobutylicum pyrE* gene, Figure 2.1.B. Instead it contains homology arms that repair the truncated *pyrE* locus in *C. autoethanogenum_ΔpyrE*, and incorporates a non-functional *LacZα* 264 bp fragment marker downstream of restored *pyrE*. Plasmid pMTLCH20 also allows for the ligation of a cargo gene cassette, which is incorporated in place of the *LacZα* marker using restriction enzymes NheI and NotI.

The creation of *agrD* complementation plasmids pMTLCH20-D1 and pMTLCH20-D2 exploited this feature. Gene fragments, *agrD1* (210 bp) and *agrD2* (211 bp) were amplified by PCR using primer pairs 0816_Comp_F + 0816_Comp_R, and 3094_Comp_F + 3094_Comp_R respectively, and ligated separately into pMTLCH20. Both inserts contained native RBS sites, but lacked their respective native promoters. Complementation was designed to be driven by the *pyrE* promoter.

2.4.6 *E. coli* DNA transformation methods

2.4.6.1 Preparation of chemically-competent *E. coli* cells

One Shot TOP10 Chemically Competent *E. coli* cells (Thermofischer Scientific) were used as a seed stock. This adapted method (Hanahan, Jessee and Bloom, 1991) uses CCMB80 buffer, which 1 L is comprised of: 10 mM KOAc pH 7.0, 80 mM $CaCl_2{\cdot}2H_2O$, 20 mM $MnCl_2{\cdot}4H_2O$, 10 mM $MgCl_2{\cdot}6H_2O$, 10 % glycerol, pH adjusted to 6.4 and sterile filtered. The competent cell procedure follows: (i) seed stock cultures were grown overnight in 5 mL of LB at 37 °C; (ii) 1 mL of the over-night culture was used to inoculate a 250 mL LB in a sterile conical flask and grown in a 37 °C shaker until OD_{600} 0.3-0.4 is reached; (iii) samples were aliquoted into 50 mL falcon tubes and centrifuged at 4 °C at 3000x g for 10 min; (iv) supernatant was removed and the pellets are resuspended with 45 mL ice-cold CCMB80 buffer and incubated on ice for 20 min; (v) cells were centrifuged at 4 °C, supernatant removed and resuspended in 5 mL CCMB80 buffer and incubated on ice for 20 min; (vi) the 50 µL of cells were aliquoted into pre-chilled 1.5 mL Eppendorf tubes and

tubes were labelled and stored in -80 °C. Competency was tested using supplied pUC19 plasmid with heat-shock transformation procedures.

2.4.6.2 Heat-shock transformation

Chemically competent One Shot® TOP10 *E. coli* was used in routine heat-shock transformations. Plasmid DNA volumes varied between newly ligated plasmid and intact plasmids in 5 µL and 1 µL volumes respectively. The procedure was as follows: (i) Aliquots of plasmid DNA were resuspended into 50 µL of thawed competent cells and then incubated on ice for 30 min; (ii) mixtures were heat-shocked at 42 °C for 30 - 45 sec; (iii) 450 µL of recovery SOC medium (Sigma Aldrich) was added and incubated at 37 °C for 1 hour; (iv) samples were plated onto LB agar plates containing appropriate antibiotics and incubated at 37 °C overnight.

2.4.6.3 Electroporation

Electro-competent CA434 *E. coli* was used in the following procedure: (i) Plasmid DNA was resuspended in 50 µL of thawed competent cells; (ii) the mixture was placed into a pre-chilled 0.2 cm wide electroporation cuvette (Biorad, UK); (iii) and electroporated at 2.50 kV using a MicroPulser electroporator (Bio-Rad, UK); (iv) 450 µL of recovery SOC medium was added to the cuvette and the mixture was transferred into an Eppendorf tube to recover at 37 °C for 1 hour in a 225 RPM shaker; (v) the mixture was plated onto LB agar plates containing appropriate antibiotics and incubated at 37 °C overnight.

2.5 Mutant creation in *C. autoethanogenum*

2.5.1 General plasmid conjugation methods

Plasmids were introduced into *C. autoethanogenum* through methods adapted from Purdy *et. al.* (2002), using *E. coli* CA434 as a donor strain. This method was employed in the conjugation of ClosTron plasmids and IFD plasmids used in ACE. Later in this project, an alternative *E. coli* donor, NEB Sexpress, was used to greatly improve conjugation efficiency through different methylation patterns as explained in chapter 3. *E. coli* Sexpress transformations followed heat-shock methods and conjugation procedures continued similarly as for CA434.

First, *C. autoethanogenum* was grown in 3 mL YTF to an OD_{660} of 0.8 – 1.0 that was reached after 3 to 4 days. On the second day of *C. autoethanogenum* growth, donor CA434 was electroporated with the intended plasmid and plated onto LB agar containing chloramphenicol and spectinomycin at concentrations specified in Table 2.3, and incubated overnight at 37 °C. Colonies were picked and cultured overnight at 37 °C in 5 mL of LB medium with chloramphenicol and spectinomycin.

After 20 hours incubation, 1.5 mL of CA434 cultures were pelleted at 4,000x g for 3 min and washed with 500 µL of PBS. The resuspension was centrifuged once more and pellets were placed into the anaerobic cabinet for 30 min to remove residual oxygen. CA434 pellets were resuspended with 200 µL of grown *C. autoethanogenum* culture and the mixture was spotted in 20 µL aliquots onto YTF agar mating plates. Mating plates were incubated for 24 hours.

Conjugated cell growth was scraped from the plates and resuspended in 500 µL of anaerobic PBS. 150 µL aliquots of resuspension was plated onto YTF agar containing D-cycloserine and thiamphenicol, at concentrations specified in Table 2.3. Growth on selective media took place over the course of 5 days, and colonies of *C. autoethanogenum* typically appeared after 4 to 5 days. Larger colonies that were assumed to be successful transformants were picked and restreaked onto fresh selective YTF containing the same antibiotics.

2.5.2 Allelic-coupled exchange (ACE)

Conjugation methods are crucial in the implementation of ACE. This technique required the use of a foundation strain with a truncated *pyrE* gene from which all future mutations take place. Principals of this technique are based on the methods described by *Ng et al.* (2013) and Ehsaan *et. al.* (2016).

2.5.2.1 Creation of *C. autoethanogenum* Δ*pyrE*

Wild-type *C. autoethanogenum* was conjugated with pMTL84151_ΔpyrE. After the second selection round, larger colonies were picked and placed into 3 mL of YTF containing thiamphenicol and grown for 3 days. 1 mL stocks containing 10 % DMSO were aliquoted into 2 mL cryovial and stored at -80 °C. The remaining ~2 mL culture was used in gDNA extractions, after which purified gDNA served as a PCR template to determine successful plasmid integration of selected colonies.

DreamTaq polymerase was used in PCR screening for pMTL84151_ΔpyrE integrants, with an extended elongation time of 1.5 min. Three individual PCR reactions took place using the same program. First, one which consisted of flanking primer pyrE_RSP with internal plasmid primer 84151_F, secondly flanking primer pyrE_FSP with internal plasmid primer 84151_R and finally, both flanking pyrE primers were used together. The first two reactions determined if the plasmid has integrated chromosomally. If so, then the third reaction which encompassed the entire integration region should reveal no bands as the amplicon is too large to form. Observed amplicons in the third reaction suggested plasmid integration is contaminated with the presence of WT *C. autoethanogenum*, and required further selection to obtain a pure single-cross integrant.

Confirmed single-cross integrants were plated onto YTF plates containing 1 mg/mL of 5-fluoroorotic acid (FOA) and 20 µg/mL uracil to select for the rare double crossover event. Plates were kept in the dark by being wrapped in aluminium foil to prevent FOA degradation. Colonies appearing after 3 days were picked and replica plated onto two sets of SDMM agar plates with and without 20 µg/mL uracil. Colonies that grew exclusively on uracil supplemented SDMM plates were re-streaked once more onto (+/-) uracil SDMM plates. Confirmed colonies were cultured in YTF for gDNA extraction and -80 °C stocks.

Extracted gDNA was used as a template to amplify the truncated region using flanking primers pyrE_FSP and pyrE_RSP. Amplicons were purified, and Sanger sequenced, as described in 2.8.5, to determine *pyrE* truncation. The final step involved streaking the newly created *ΔpyrE* strains onto YTF plates containing thiamphenicol. Lack of growth indicated pMTL84151_ΔpyrE plasmid loss.

2.5.2.2 Creation of in-frame deletion (IFD) mutants

IFD plasmids pMTL84151_0816 and pMTL84151_3094 were conjugated separately into *C. autoethanogenum_ΔpyrE*. After the second round of selection, large colonies were verified for single cross integration, and screening procedures were employed as described previously in 2.6.2.1. Right and left homology arm flanking primers for each target gene were paired with primers 84151_F and 84151_R.

After successfully isolating single-cross integrants, integrants were plated on FOA and grown for 3 days to promote the double crossover event. Visible colonies were picked and replica plated onto SDMM uracil +/- plates and grown for 4 days, this step was

repeated. To screen for successful target gene truncation, colonies that grew exclusively on uracil supplemented plates were picked and screened using flanking primers. Sanger sequencing confirmed in-frame truncation of the target gene.

2.5.2.3 Restoration and complementation techniques

Restoration of the *pyrE* locus is required for the completion of mutants and used plasmid pMTLCH20. This plasmid was conjugated into *pyrE* deficient IFD mutant strains and this procedure did not require PCR screening and instead large and fast-growing colonies were streaked directly onto SDMM uracil +/- plates and grown for 4 days. Colonies that could grow without uracil supplementation were picked and this step was repeated. pyrE_FSP and pyrE_RSP flanking primers were used for PCR screening and Sanger sequencing to confirm *pyrE* restoration. A 264 bp *LacZα* marker was incorporated into *pyrE* restored mutants.

Complementation of *agrD* mutants followed similarly but used a modified version of pMTLCH20 containing complementation fragments, with plasmids pMTLCH20-D1 and pMTLCH20-D2. Conjugation took place as described in 2.6.1, colonies successfully transformed were picked and replica plated onto SDMM uracil +/- plates and grown for 4 days. Colonies growing without uracil supplementation SDMM were replica played once more to further confirm that the newly restored *pyrE* locus containing complemented genes was integrated chromosomally. Genomic DNA was isolated and screened using pyrE_FSP and pyrE_RSP flanking primers. PCR amplicon band sizes were expected to be larger with the inclusion of respective desired complemented genes, Sanger sequencing confirmed full complementation.

2.5.2.4 ClosTron plasmid conjugation

The ClosTron method of gene disruption was performed in a *C. autoethanogenum ΔpyrE* recipient, that allowed for later complementation. Conjugation of the ClosTron plasmids pMTL007C-E2::CAUagrB-234a and pMTL007C-E2::CAUagrB-232s used the donor *E. coli* Sexpress strain for its optimised conjugation efficiency, the protocol followed is described in 2.5.1.

After growth on initial selective D-cycloserine/thiamphenicol YTF plates for plasmid conjugation, colonies were streaked onto plates containing clarithromycin to select for intron insertion at the target locus. Clarithromycin resistant colonies were isolated and replicated plated onto YTF agar containing clarithromycin or thiamphenicol. Plasmid loss

occurred when replica colonies only grew on clarithromycin. Colonies were picked and screened for ClosTron site integration using a flanking *agrB* primer and erythromycin cassette primer found within the intron. Amplicon bands showed to be 1.8 kB in size relative to the WT, successful integrants were isolated and Sanger sequenced. Instead of a Southern Blot analysis to determine correct insertion location of the ClosTron, gDNA was phenol-chloroform extracted and sent for Illumina Sequencing, (Microbes NG, Birmingham).

2.6 Phenotypic analysis

2.6.1 Growth curve experiments

Growth experiments comparing *C. autoethanogenum agrD* mutants Δ*agrD*1, Δ*agrD2* and Δ*agrD1D2*, and complementation mutants Δ*agrD1D2_D1*^{comp} and Δ*agrD1D2_D2*^{comp}, against WT *C. autoethanogenum* were performed in 250 mL serum flasks as described in 2.2.4.1. Bacterial stocks were restreaked onto YTF agar plates and single colonies were picked and resuspended in 10 mL PETC with 50 mM D-fructose and grown for 3 days. Cultures were grown to late-exponential phase to an approximate OD_{660} of 0.8. Calculations were made to ensure that the flask starting OD_{660} was approximately 0.01 and exact amounts of pre-cultures were injected accordingly into serum flasks. Initial OD readings were measured marked time-point "T0" and flasks were cultured for up to 10 days with 10-12 time-points. Inoculated bottles were labelled and placed into a New Brunswick Innova 44 shaking incubator (Eppendorf) 225 rpm, at 37 °C. Solvent samples were collected at each time-point as described in 2.6.1.1 and analysed, described in 2.7. Growth curve and analytical data from technical and biological replicates was mapped and scaled with exponential phases.

For growth on C1 gases, serum flasks were prepared and used as described in 2.2.3.1. Bacterial stocks and pre-cultures were prepared similarly as above with cultures inoculated at a starting OD_{660} of 0.01 in 50 mL of PETC. The "T0" time point measurements were performed before gas addition to minimize gas substrate leakage and were cultured for up to 30 days. Samples were taken approximately every 3 – 5 days.

2.6.1.1 Sampling and optical density measurement

For measurement of optical density and solvent content, samples were extracted from serum flasks using 1 mL needled syringes, (Becton Dickinson, USA). Specific timepoints

were measured at an optical density of 660 nM using a Jenway 7300 spectrophotometer (Bibby Scientific, UK). The optical density of 660 nm was chosen to avoid absorption interference from the resazurin dye that absorbs between 590 – 640 nm with a peak at 610 nm (Zrimšek *et al.*, 2004).

During growth experiments, samples entering exponential phase were diluted 1:10 with sterile PETC medium to allow OD measurements to be read within the spectrophotometric range. PETC medium was used as a reference blank. For solvent analysis, 1 mL of undiluted serum flask samples were spun down at maximum speed at 4 °C for 10 min. The supernatant was pipetted into 2 mL cryovials and stored at -80 °C for later chromatographic analysis.

2.6.2 Cell survival assay

Analysis of cell fitness and viability between WT *C. autoethanogenum* and Δ*agrD1D2* mutants begun during the mid-stationary phase of fructose-grown fermentations. The experiment continued over the course of a month and samples were serially diluted and plated onto YTF agar medium. Dilutions ranged from 10^{-3} to 10^{-7} for the first 2 weeks, then to 10^{0} to 10^{-3} as cell viability dropped. Colonies were counted and averaged.

2.6.3 Synthetic peptide complementation

An alternative route to phenotypically complementing Δ*agrD1D2* mutants was through chemical means done by synthetic AgrD1 and AgrD2 peptides that were designed based on their respective putative residue sequences. These peptides were ordered to be synthesized by Peptide Synthetics (Peptide Protein Research Ltd, UK), Table 2.7.

Peptides, AgrD1-T0R5, AgrD1-T1R5, AgrD1-T2R5 and AgrD1-T3R5 were only received, and at inferior purities of 50 – 75 %, in low yields (< 2 mg). Received peptides were dissolved in filter sterilized DMSO to 20 mM stock concentrations and stored at -20 °C. Growth experiments with the use of peptides involved culturing WT *C. autoethanogenum* and *C. autoethanogenum ΔagrD1D2* in triplicate in 15 mL of PETC as described in 2.2.3.1. Peptides added to the medium were added to a final concentration of 20 µM (1 µL stock per 1 ml), with equivalent volumes of pure DMSO serving as negative control. Sampling was limited to 5 time points to minimise loss in culturing medium. Samples were collected and analysed as described in 2.6.1.1.

Table 2.7, Sequences of ordered synthetic peptides: Sequence in brackets denoted lactone ring structure.

Peptide	Residue Sequence	Modification at C-terminus	Protein
AgrD1-T0R5	(STGGA)	Lactone bond between the C-terminal carboxyl group and the OH group of the serine residue.	AgrD1
AgrD1-T1R5	T(STGGA)		
AgrD1-T2R5	PT(STGGA)		
AgrD1-T3R5	VPT(STGGA)		
AgrD2-T0R5	(SAGCY)	Lactone bond between the C-terminal carboxyl group and the OH group of the serine residue.	AgrD2
AgrD2-T1R5	A(SAGCY)		
AgrD2-T2R5	PA(SAGCY)		
AgrD2-T3R5	VPA(SAGCY)		

Peptides ordered to 5 mg, >90 % purity

2.6.4 Heat resistant spore assay

Determining the presence of germinating spores in wild-type *C. autoethanogenum* was performed using heat shock-killed vegetative and non-spore forming cells, allowing heat-resistant spores to survive (Kamiya *et al.*, 1989; Talukdar *et al.*, 2015). Stocks of WT *C. autoethanogenum* were grown in 3 mL YTF in triplicate for up to 5 days. 2 mL from each culture was spun down, washed and resuspended with 400 µL of anaerobic PBS. The 400 µL resuspended mixture was split into two 200-µL aliquots, one acting as a control, whereas the other was heat shocked at 80 °C for 10 min using a SensoQuest (SensoQuest GmbH, Germany) PCR machine.

The experimental and control triplicates were then serially diluted ranging from 10^0 to 10^{-2} and plated onto YTF plates. Colonies were expected to appear in 4-5 days. As a means of technique validation, *C. acetobuytlicum* and *C. ljungdahlii* were used as controls. *C. acetobuytlicum* was grown on clostridial basal medium (CBM) (O'Brien and Morris, 1971) and *C. ljungdahlii* on YTF medium. Both control organisms were subjected to identical heat-shock procedures as *C. autoethanogenum* and plated on either CBM or YTF agar plates respectively.

2.6.5 Granulose staining

Determination of granulose formation followed similar methods as described in Steiner *et. al.* (2012). WT *C. autoethanogenum* was plated onto YTF agar and allowed to grow for four days until colonies were visible. An inverted petri-dish was placed onto a 37 °C

heating block in a fume hood, where Iodine crystals were placed inside and allowed to evaporate. The base of the petri dish contained the iodine vapours. Plates with colonies were removed from the anaerobic cabinet and the agar containing base was placed over the evaporating iodine crystals for approximately 10 min. As the colonies formed a dark yellow/brown colour the colony plate was removed from the heating block to allow excess iodine vapour to evaporate and subsequent colour changes were noted.

2.6.6 Microscopy and spore staining

Bacterial culture samples were observed using a Nikon Eclipse Ci-E microscope (Nikon Instruments Europe B.V., UK) using standard microscopy procedures. Visualising the presence of spores was done using Schaeffer and Fulton Spore Staining Kit (Sigma Aldrich, UK) as per manufacturer's instructions.

2.6.7 Salt tolerance

The determination of salt tolerance between the WT and *ΔagrD1D2* was performed using serum flasks containing 90 mL PETC with 50 mM D-fructose. NaCl was added to flasks to a concentration of 200 mM, and control flask contained no NaCl. Starting inoculum was OD_{660} 0.01 and the experiment took course over a 5-day period in a 225 RPM shaking incubator at 37 °C, with periodic visual assessment. In addition to liquid cultures, stocks of WT and *ΔagrD1D2* were streaked onto 200 mM NaCl +/- YTF agar plates and grown for up to 5 days.

2.7 Analytics

Sample supernatants from growth experiments as described in 2.6.1.1, were stored at -80 °C and were thawed at 4 °C overnight prior to chromatographic analysis.

2.7.1 Gas chromatography (GC)

Analysis of solvents using GC allowed for large quantities of samples to be examined. Acetate and ethanol solvents were quantified using the following method: 500 µL volumes were pipetted from freshly thawed samples into 1.5 mL Eppendorf tubes and acidified with 5 µL of a 10 M H_2SO_4. Samples were vortexed and 500 µL of propyl propionate containing 50 mM valeric acid is added as an internal standard. Sample were vortexed again for 10 sec. The extraction mixture was spun down at 21, 000 x g for 1 min

to remove denatured proteins. Samples separated into a biphasic solution, from which 300 µL of the organic upper phase was removed and pipetted into a 300 µL glass insert that was placed inside a 2 mL brown sampling vial. Vials were sealed with the use of snap-caps. Numbered vials were placed into the auto-loading carousel of a Focus™ gas chromatograph (Thermo Fischer Scientific) which used a 30 m TRACE TR-FFAP column (Thermo Fischer Scientific) with a 0.25 mm internal diameter. The chromatograph used a flame ionisation detector which is maintained with 350 mL/min compressed air, 35 mL/min H_2 and 30 mL/min N_2. The carrier gas is H_2 supplied at 0.8 mL/min. Samples were injected at 1 µL volumes. Injector inlet temperature was at 240 °C and detector temperature was set to 270 °C. For resolution of peaks the following column program was used: after injection, temperature was held at 40 °C for 2 min, then ramped to 150 °C at a rate of 80 °C/min, and finally to 210 °C for 1 min.

Standards for GC analysis contained 1 M ethanol and 1 M sodium acetate. To create a standard curve, dilutions of 0, 1, 3, 5, 10, 30, 50 and 150 mM were diluted using ELGA water and were prepared for the GC in an identical manner as experimental samples.

2.7.2 Gas headspace analysis

Composition of serum flask headspace gas were measured using a TRACE gas chromatograph (Thermo Fischer Scientific). Flasks from growth and solvent experiments were intermittently measured during the early growth phase or at the end of the experiment. Using a 10 mL Hamilton SampleLock syringe (Hamilton, USA), 1 mL of gas was removed from the serum bottle and manually injected into the TRACE GC machine. For the resolution of peaks, the following column program was used: after injection, temperature was held at 40 °C for 3 min, then ramped to 250 °C at a rate of 30 °C/min, the duration of each run took 10.7 minutes, after which the next sample was injected. Standards for gas analysis, involved injecting 1 mL of an anaerobic mix which comprised of 1 % CO, 1 % CO_2, 1 % H_2, 97 % N_2. To measure methane composition, standards consisting of 0.5 % CO, 0.5 % CO_2, 0.5 % H_2, 25 % CH_4, and 73 % N_2 were used. Data obtained was analysed using Chromeleon 7.2 Chromatography Data System (Thermo Fisher Scientific, UK).

2.7.3 High performance liquid chromatography

Measurement of a wider spectrum of fermentation productions required the use of HPLC. Thawed samples were mixed with 1:1 volumes (300 µL) of sample supernatant to internal standard and mobile phase solution, (80 mM valeric acid and 5mM H_2SO_4). Mixed samples were vortexed and syringed filtered into HPLC vials containing 300 µL glass inserts and sealed with snap-caps. Vials were placed inside a 4 °C cooled, auto-loading carousel and measured using a Dionex Ulti-Mate 3000 HPLC system (Thermo Fischer Scientific, UK). Volumes of 20 µL were injected into a Bi-Rad Aminex HPX-87H (Biorad UK) "short" column, which was kept at 35 °C and had a uniform flow rate of 5mM H_2SO_4 mobile phase at 0.5 mL/min. Elution took approximately 30 min per sample. Detection was done through a refractive index (RI) and diode array detector (DAD) at UV 210 nm.

Standards used included fermentation products, ethanol, sodium acetate, 2,3-butanediol, D-lactate and carbon source, D-fructose. For standard curves, dilutions of 0, 1, 3, 5, 10, 35, 50, 75 and 100 mM were diluted using ELGA water and handled in an identical manner as experimental samples. Data obtained was analysed using Chromeleon 7.2 Chromatography Data System (Thermo Fisher Scientific, UK).

2.8 Enzyme assays and proteomics

The measurement of enzyme activity in *C. autoethanogenum was* performed in strict anaerobic conditions with the use of spectrophotometric absorbance.

2.8.1 Strain growth conditions

C. autoethanogenum WT and Δ*agrD1D2* cells were grown in 400 mL PETC with 50 mM D-fructose. Cells were reached an OD_{660} of 0.4 – 0.45, to which cells were spun down in the COY tent and supernatant was fully removed. The pellets were flash frozen in liquid N_2 and stored at -80 °C until further use.

2.8.2 Crude enzyme extraction

Cell pellets were resuspended thoroughly in a combination of 750 µL of BugBuster® protein extraction reagent (Merck Milipore, United States) and 750 µL 50 mM TRIS-HCl and 0.5 mM $MgCl_2$, pH 7.0, both solutions were prepared and mixed anaerobically. Lysozyme, 10 mg/mL, and 1 mg of DNase (Sigma Aldrich) was added to the mixture and vortexed. The lysis mixture was placed into a pre-evacuated 15 mL Hungate tube and

sealed. The tube was placed onto a rocker at room temperature at 50 RPM for one hour. Tubes were placed back into the COY tent and the mixture was pipetted into a 2 mL cryovial (Sarstedt, Germany) and spun at 20,000xg at 4 °C for 20 min. The clear supernatant was injected into a 5 mL serum bottle pre-vacuum-filled with pure N_2 and stored on ice ready for analysis.

2.8.3 Specific activity measurements

Protocols followed adapted methods of Mock *et. al.* (2015), and Wang *et. al.* (2013). Enzyme activities were performed at 37 °C, with the use of 3.5 mL OS glass cuvettes (Hellma Analytics, Germany), measured using a Specord 250 spectrophotometer (AnalytikJena AG, Germany). Cuvettes were sealed with silicone stoppers to ensure gas impermeability, while allowing needles to pass through. Buffers prepared for each assay were done in a strictly anaerobic environment using sealed 100 mL serum bottles. Stock solutions of 1 M methyl viologen, 1 M DTT and 1 M sodium formate were mixed using anaerobic distilled water, and sparged with N_2. Buffers for each assay were made with 100 mM Tris-HCl (pH7.5) using anaerobic dH_2O in 100 mL sealed serum bottles. On the day of the experiment 10 mM methyl viologen and 2mM DTT was added. Separate bottles of buffers for each assay were made and dependant on the assay were sparged with either pure CO, H_2 or N_2 gas.

2.8.3.1 CO dehydrogenase

Cuvette assay mixture contained 100 mM Tris-HCl, 10 mM methyl viologen and 2 mM DTT. Cuvettes were sparged with CO for 5 min to ensure a 100% CO gas phase. Cuvettes were placed into the spectrophotometer and the background was measured for up to 60 sec, subsequently varying amounts of crude enzyme extract were injected into the cuvettes and measured immediately. As a negative control, a set of cuvettes containing N_2-sparged buffer, as well as 100% N_2 gas-phase were used.

2.8.3.2 Hydrogenase

Cuvette assay mixture contained 100 mM Tris-HCl, 10 mM methyl viologen and 2 mM DTT. Cuvettes were sparged with H_2 for 5 min to ensure a 100% H_2 gas phase. Cuvettes were placed into the spectrophotometer and the background was measured for up to 60 sec. Varying amounts of crude enzyme extract were injected into the cuvettes and

measured immediately. As a negative control, a set of cuvettes containing N_2-sparged buffer as well as 100% N_2 gas-phase were used.

2.8.3.3 Alcohol dehydrogenase

Cuvette assay mixture contained 100 mM potassium phosphate (pH 6), 1.1 mM acetaldehyde, 1 mM NADH and 2 mM DTT. Cuvettes were sparged with N_2 for 5 min to ensure a 100% N_2 gas phase. Varying amounts of crude enzyme extract were injected into the cuvettes and measured immediately. As a negative control, cuvettes containing no acetaldehyde were used.

2.8.3.4 Calculations

After the start of each reaction the reduction of methyl viologen with extinction coefficient, ε = 9.8 mM^{-1} cm^{-1} was monitored spectrophotometrically at 578 nm. NADH with extinction coefficient ε = 6.22 $mM^{-1}cm^{-1}$ was monitored spectrophotometrically at 340 nm. Calculation of specific enzyme activity used equation 2.1:

$$U = \frac{V\frac{\Delta A}{\min}}{v\varepsilon d} \quad \text{(eq. 2.1)}$$

Where U is specific enzyme activity, V is total assay volume, v is volume of crude extract added, $\left(\frac{\Delta A}{min}\right)$ is absorbance change per minutes (the slope), d is the path length (1 cm) and ε is the extinction coefficient. Units are first given as µmol $min^{-1}ml^{-1}$, then to obtain a µmol $min^{-1}mg^{-1}$ unit (Specific enzyme activity (U), values were divided by protein concentration.

Negative control samples were deducted from averaged experimental samples for final values. Values were normalised against WT and *ΔagrD1D2* protein concentrations which were determined by Bradford Assay, (Sigma Aldrich, UK).

2.8.4 Proteomics analysis

Cultures of *C. autoethanogenum* WT and *C. autoethanogenum ΔagrD1D2* were grown in triplicate in 250 mL serum flasks containing PETC medium with 50 mM D-fructose, by first inoculating this medium to at an approximate starting OD_{660} of 0.01. Cultures were harvested at OD_{660} 1.0 after 3-4 days and were spun down. The supernatant was removed and the pellet was washed twice with PBS and once with protein extraction buffer, triethyl ammonium bicarbonate (TEAB), 0.5M pH 8.5 (Sigma Aldrich). Samples were flash

frozen in liquid nitrogen and sent to the University of Sheffield to be analysed by Dr Mahendra Raut with the use 4-plex reagents isobaric and tags for relative and absolute quantitation (iTRAQ). Briefly, harvested cells were ruptured using bead-beating and proteins were acetone-precipitated, then treated with 4-plex labelling before being separated by Hypercarb fractionation. Fractions were subjected to LC-MS/MS analysis and data was processed using MaxQuant, see reference for full procedure, (Raut *et al.*, 2016).

2.8.5 DNA sequencing

DNA and plasmid samples that required sequencing verification was performed by Source BioScience Gene Service (Nottingham, United Kingdom) or through Eurofins Genomics (Germany). Samples were prepared as per either services specifications and collected for delivery in barcoded 2 ml Cryotubes.

2.9 Bioinformatic tools

2.9.1 In-silico software and tools

The National Centre for Biotechnology Information (NCBI) primary database was used for the identification of nucleotide sequences, sequencing comparison identification of nucleotide and amino acid sequences and sequence comparisons, using the Basic Local Alignment Search Tool (BLAST) (Altschul 1990). Multiple short sequences were aligned using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) (Edgar, 2004). Online research platform Benchling (Benchling Inc. USA) was used for in-silico construction and visualization of plasmids, primer design and sequence alignment. The Sorting Intolerant From Tolerant (SIFT) tool was used to predict if amino acid substitution can affect protein functions, and if resulting changes are tolerated, (Ng and Henikoff, 2003).

2.9.2 Genome sequencing software

Raw genomic data was processed and assembled using CLC Genomics Workbench (v. 8.0.2), (Qiagen, Denmark) and were mapped against the Humphreys *et. al.* (2015) *C. autoethanogenum* DSM 10061 genome. A CLC Genomics Workbench workflow, curated by Dr Alexander Grosse-Honebrink (Grosse-Honebrink, 2017), included ordered parameters to initially map, then call single nucleotide variants (SNVs), insertion/deletion (InDels) regions and breakpoints.

2.9.3 Data analysis software

Data that was presented graphically used Graphpad Prism 7, (GraphPad Software Inc, USA). Numerical data was stored, calculated and analysed using Microsoft Excel 2016 (Microsoft). Calculation included determination of standard error of the mean, and statistical significance values (P value) between data sets was determined using Student's T-test.

Chapter 3

Design and Construction of *agrD* mutants

3.1 Introduction

3.1.1 Current understanding of the Agr system in industrially-relevant Clostridium spp.

The development of metabolic engineering in *C. autoethanogenum* has acquired steady progress, hastened by companies and academia alike (Nagaraju *et al.*, 2015; Liew *et al.*, 2017). The past decade of research into acetogenic and solventogenic *Clostridia* spp. has generated a wealth of knowledge allowing a basis upon which new cutting-edge developments have taken place (Papoutsakis, 2008). Numerous publications have listed methods of optimizing solvent production, that include reducing unwanted metabolic by-products, improving product tolerance and improving the understanding of energy conservation and enzyme activity (Minton *et al.*, 2016). Although these advances cover vital topics in pursuit of optimising *C. autoethanogenum* into a microbial cell factory, there are substantial areas that have not yet been understood that could further contribute to this goal.

Studies on the industrial applications of *C. acetobutylicum* have uncovered much of the organism's solventogenic mechanisms and strategies for further optimisation (Jones and Woods, 1986; Papoutsakis, 2008). Inevitably this field has crossed into other areas such as sporulation, sugar transport, metabolism, and cell to cell signalling (quorum sensing), to name a few (Jones *et al.*, 1982; Brückner and Titgemeyer, 2002; Tangney and Mitchell, 2007; Steiner *et al.*, 2011). The latter cell signalling study was first elucidated in Steiner and Jabbari *et al.*'s work, that primarily sought to understand the relationship between sporulation and solventogenesis, which in turn lead to the conclusion that Agr QS plays a role in sporulation and granulose formation (Steiner *et al.*, 2011; Jabbari *et al.*, 2013). Furthermore, *C. acetobutylicum*'s RNPP system has also been uncovered to play a part in sporulation and solventogenesis, specifically the organisms mechanism of switching from acid to solvent production (Kotte *et al.*, 2017).

As industrial fermentations are carried out at high cell densities where QS systems are typically active, it is relevant to investigate the roles they play in the regulation of important physiological processes. Apart from the above-mentioned *C. acetobutylicum*, very little is known about these systems in other industrially relevant *Clostridium* species.

Neither Agr or RNPP QS systems have been examined in detail in *C. autoethanogenum*, but the RNPP systems proves to be the least favourable to study as respective signal peptide encoding genes have yet to be identified in *C. autoethanogenum* and *C. ljungdahlii*. This is dissimilar to *C. acetobutylicum* where characteristic signal peptide encoding genes which are typically identified downstream of their cogent regulators (Kotte *et al.*, 2017).

The Agr system initially gained attention through its characterisation in *S. aureus* and is typified by four genes, *agrB, agrD, agrC,* and *agrA*, encoding a protein required for signalling peptide maturation and export (AgrB), the signalling peptide precursor (AgrD), and a two-component signal transduction system (AgrA, AgrC), respectively (Recsei *et al.*, 1986; Novick *et al.*, 1993; Kong, Vuong and Otto, 2006). Moreover, this system is similarly represented in many firmicutes (Wuster and Babu, 2008), and can occasionally can be found in multiple copies and locations within a genome. Annotation of the *C. autoethanogenum* genome suggested the presence of two incomplete *agr* clusters and it was decided that investigations into their physiological role should be the central focus of this study.

Contributions by *Steiner et. al.,* were one of the few prominent motivators in literature in assessing the role of the Agr system in a solventogenic bacterium (Steiner *et al.*, 2011). Other relevant studies explaining Agr's role in Clostridia are also noteworthy, with brief examples including the mediation of toxin and sporulation in *C. botulinum* (Cooksley *et al.*, 2012), toxin induction in *C. difficile* (Darkoh *et al.*, 2015) and in *C. perfringens* (Ohtani and Shimizu, 2015). These prominent studied gave confidence in targeting the Agr system in *C. autoethanogenum* as a means of studying its role in acetogens.

3.1.2 Genetic manipulation of *Clostridia*

Advances in genetic modification of *Clostridia* spp. over the past three decades have been influenced by the focus on medical and industrial applications that involved species not limited to; *C. ljungdahlii, C. acetobutylicum*, *C. beijerinckii*, *C. difficile*, *C. botulinum, C. pasteurianum* and *C. perfringens* (Bannam and Rood, 1993; Green *et al.*, 1996; Leang *et*

al., 2013; Ng *et al.*, 2013; Pyne *et al.*, 2013; Zhang, Grosse-Honebrink and Minton, 2015; Molitor *et al.*, 2016). However, these advances are still hindered by inherent drawbacks unique to Gram-positive bacteria, notably of which involves the transfer of foreign DNA. Electroporation has been a widely used technique in exogenous DNA transferal, however efficiencies are notably higher in Gram-negatives, compared to Gram-positives (Meddeb-Mouelhi, Dulcey and Beauregard, 2012; Trevors *et al.*, 2012), this can be attributed to their innate dense outer peptidoglycan cell walls which hampers efficient DNA transfer (Pyne *et al.*, 2014).

Coupled with this difficulty, there are further hurdles such as restriction-methylation (RM) systems present in some *Clostridia* spp. that actively destroy unmethylated foreign DNA or DNA with a differing methylation signature. This can be circumvented by the introduction of the recipient's methylation patterns onto exogenous DNA via an appropriate *E. coli* donor equipped with the necessary methylases, or in vitro methylation, as demonstrated in *C. acetobutylicum, C. botulinum* and *C. pasteurianum* and several other species (Davis *et al.*, 2000; Berto *et al.*, 2011; Minton *et al.*, 2016). In recent studies, a more up-to-date approach has been to use genomic sequencing to identify RM systems and target them for deletion using a variety of established genetic manipulation techniques (Dong *et al.*, 2010; Croux *et al.*, 2016). These proven techniques include the ClosTron (Heap 2007), allelic-coupled exchange (ACE) (Ng *et al.*, 2013), transposon mutagenesis (Cartman and Minton, 2010) and recently developed, CRISPR/Cas9 (Richter *et al.*, 2012; Huang *et al.*, 2016; Nagaraju *et al.*, 2016), that have formed a comprehensive molecular toolbox for *Clostridia* spp. manipulation.

Focusing on acetogens, methods of overcoming difficulties of introducing foreign DNA varies between each organism. Electroporation in *C. ljungdahlii* has been proven successful (Leang *et al.*, 2013; Molitor *et al.*, 2016), but later efforts replicating this method in *C. autoethanogenum* were largely unfruitful. The attempts that did yield colonies did so with very low efficiencies, suggesting the presence of various active restriction systems elucidated in literature (Brown *et al.*, 2014; Humphreys *et al.*, 2015) and according to experience and opinion of Dr Eric Liew (personal communication, Uni. of Nottingham, 2013). However, electroporation has been shown to work in other acetogens that include; *A. woodii* (Stratz *et al.*, 1994), *C. aceticum* (Schiel-Bengelsdorf and Dürre, 2012), *M. thermoacetica* (Kita *et al.*, 2013), unspecified proprietary strains (Berzin, Kiriukhin and Tyurin, 2012) and recently, *E. limosum* (Mansfield, 2018).

Despite this list, electroporation is still not a fully universal application with regards to several notable acetogens that include *C. autoethanogenum*, therefore requiring a different route to efficient DNA transfer. A proven method was through conjugation donor, *E. coli* CA434 (Liew *et al.*, 2017), which was used throughout this project. CA434 was the preferred means as it could avoid the recipient's high nuclease activity which targets foreign DNA (Stratz *et al.*, 1994). The CA434 strain is comprised of *E. coli* HB101 donor cells which carry the R702 plasmid, that is a non-specific incompatibility group plasmid (IncP), which enables the transfer of non-conjugative plasmids into Gram-positive organisms (Williams, Young and Young, 1990). With this available method, it is possible to introduce a variety of plasmids which need to be selected for during the transformation process and can be introduced repeatedly into the same *C. autoethanogenum* strain. Despite successes with this donor strain, it was noted that *C. autoethanogenum* possesses an active type IV restriction system (Woods, 2018), which required further understanding, and possible circumvention discussed in later in the chapter.

In-frame deletion (IFD) methods based on allelic-coupled exchange (ACE) technique (Ng *et al.*, 2013), were the preferred means of gene disruption and were employed in this study. It is worth noting that ClosTron gene disruption does carry several advantages, but one of its few drawbacks include the inability to target relatively small genes (< 400 bp) reducing the possibilities of targets, and therefore is an unsuitable strategy in small gene disruption (Heap *et al.*, 2010). The ACE-based IFD technique involves the initial creation of an orotate phosphoribosyltransferase (*pyrE*) deficient background strain of *C. autoethanogenum* which acts as a foundational mutant strain from which all mutations are created. With the encoded PyrE enzyme truncated, the organism is strictly uracil auxotrophic, and is unable to metabolise of 5-fluoroorotic acid into cytotoxic 5-fluorouracil (Ng *et al.*, 2013), both features allow a high degree of selection at key stages of plasmid to chromosome recombination. This method first described by Heap et. al., in *C. acetobutylicum* (Heap *et al.*, 2012), and later adapted to *C. autoethanogenum* by Liew et. al. (2017). This technique bears the advantage of straight-forward genetic complementation strategies that are concomitant with *pyrE* repair back to WT genotype (Ng *et al.*, 2013), as shown in Figure 1.4.

3.1.3 Aims of this study

The main aims of this study were to provide a bioinformatical assessment of the annotated *agr* clusters in *C. autoethanogenum* and to create a collection of *agrD*-deficient mutants based on ACE methods (Heap *et al.*, 2012; Ng *et al.*, 2013). For the latter, this would necessitate the creation of a IFD *pyrE* deficient *C. autoethanogenum* mutant, followed by designing IFD plasmids specifically targeting *argD1* and agr*D2* QS genes. After successful conjugation and transformation, genetic confirmation would take place through colony PCR and Sanger-sequencing. Ultimately, all strains would be whole-genome Illumina sequenced to determine single nucleotide variants and small structural variants. Unexpected nucleotide variants would be assessed to determine any deleterious proteins changes, specifically which may affect expected phenotypic changes. The key areas of focus were:

- Production of *agrD1* and *agrD2* deficient mutants

After obtaining *ΔpyrE* strains, subsequent conjugations with Δ*agrD1* and Δ*agrD2* IFD plasmids will take place based on methods previously implemented in *Clostridium* spp. If there are difficulties with these methods, are there straightforward means of circumventing them?

- Complementation of Δ*agrD1 and* Δ*agrD2* deficient mutants

To complement mutants, a means of re-introducing *agrD1* and *agrD2* needs to be performed via plasmid or chromosomal complementation. In either case, both genes must be placed on separate plasmids, with either native or plasmid-encoded promoters. Complemented mutants will also require full *pyrE* restoration.

- Screening mutants via whole-genome Illumina sequencing

After successful deletion of target genes and full restoration through the repair of *pyrE*, does sequencing reveal undesirable single nucleotide variants (SNVs) and insertion deletions (InDels)? Will these nucleotide alterations adversely impact key enzymes and proteins?

Further aims included reproducing phenotypic effects seen in *agrD* mutants through *agrB* ClosTron creation and transformation.

3.2 Results

3.2.1 Bioinformatic analysis of Agr systems in *C. autoethanogenum*

C. autoethanogenum possess two incomplete sets of *agr* genes, separated into two clusters that contain *agrB, agrC1, agrD1* and *agrD2, agrC2*, and can be schematically represented as shown in Figure 3.1. Notable features in these clusters includes the absence of a putative *agrA* gene, and second cluster proteins, AgrD2 and AgrC2 share 70.21 % and 74 % identity with AgrD1 and AgrC1 respectively.

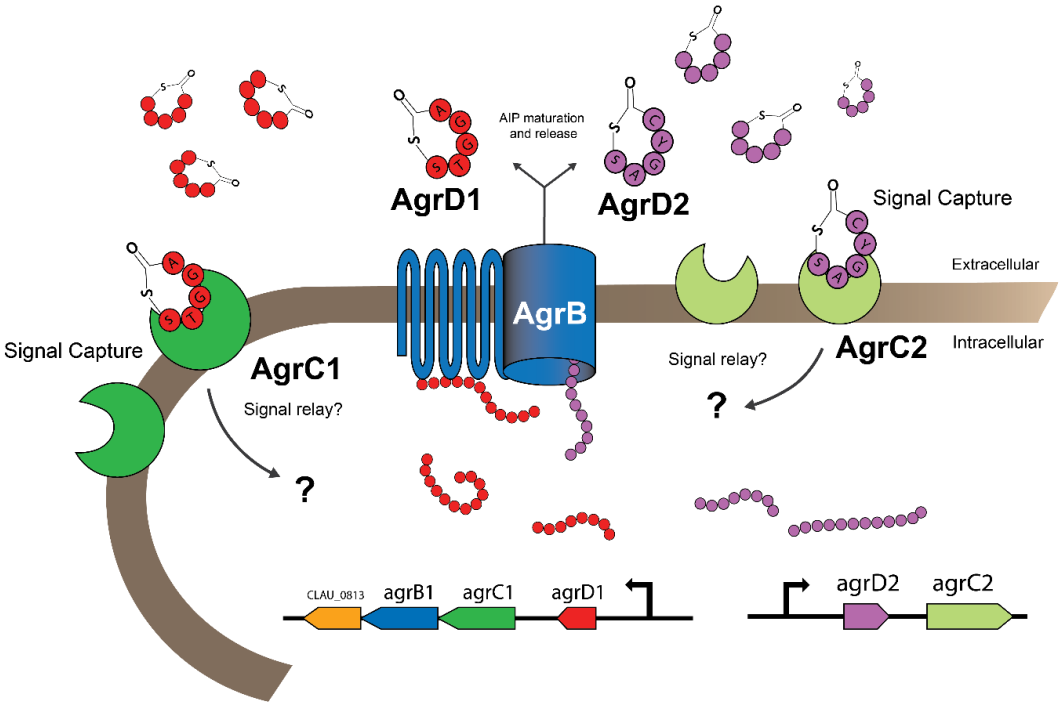


Figure 3.1, Proposed schematic overview of the Agr QS system in *C. autoethanogenum*.

Despite two sets of AgrD and AgrC, there appears to be only one copy of AgrB, which suggests that it may perform the function of maturation and export of both AgrD1 and AgrD2 signalling peptides. This idea is supported by studies in *S. aureus* where sets of chimeric-fusions of two *agrB* homologs resulted in *S. aureus* AgrD1 and AgrD2 signalling activity (Zhang and Ji, 2004).

Comparing *C. acetobutylicum*'s Agr protein identities to *C. autoethanogenum*, it is noted they share a low percentage identity with AgrB; 25.4, AgrD1/D2; 30.4/34.8 and AgrC1/C2; 22.5/23.5, respectively. Using P-BLAST to discern *C. autoethanogenum's* respective Agr

proteins with high-similarity sequences, several acetogens were identified that included, Table 3.1.

Table 3.1, Comparison of Agr proteins from C. autoethanogenum and related acetogens: Absent similarities are represented by a dash "-". * = likely to represent AgrA. *C. scatologenes* only possesses one copy of AgrD.

	Cluster 1				***Cluster 2***		**Genome Sequence**
Species	AgrB1	AgrC1	AgrD1	Spo0E-like*	AgrD2	AgrC2	
C. ljungdahlii	100	82	98	100	83.3	98	(Köpke *et al.*, 2010)
C. ragsdalei	92	79	77	91	79	82	(Huhnke, Raymond L.; Lewis, Randy S.; Tanner, 2008)
C. coskatii	98	79	79	-	77	81	(Bengelsdorf *et al.*, 2016)
C. scatologenes	73	64	51	-	51	63	(Zhu *et al.*, 2015)
C. drakei	73	64	-	-	-	62	(Jeong *et al.*, 2014)

The closest identities are shared with *C. ljungdahlii*; *C. ljungdahlii* has 3 separate gene clusters, two of which appear to be homologous to the ones found in *C. autoethanogenum*, meaning that the respective proteins share high percentage identities, Table 3.1.

Given that the role of these systems was to be investigated in *C. autoethanogenum*, an important question to consider was which of the *agr* genes should targeted for inactivation by mutation. AgrD proteins have conserved lactone ring structure that are a ubiquitous feature in AgrD amongst many species with a putative *agr* system (Wuster and Babu, 2008; Steiner *et al.*, 2012), Table 3.2.

In addition to the conserved lactone ring structure, a Pro-X-X-Pro motif that has been suggested to represent the recognition site in which AgrD interacts with AgrB, (Zhang and Ji, 2004). In the absence of clearly identifiable *agrA* genes, it was decided to inactivate the *agrD* signal peptide genes for the following reasons; first, inactivation of a single *agrD* gene would allow for the study of the roles of these systems individually, in contrast to an *agrB* mutation. Second, this would allow future testing of native or synthetic peptides to chemically complement any observed phenotypic defects, something that would not be possible for *agrC* mutants. Either *C. autoethanogenum agrD* gene is of a relatively small size with distinguishable structural features and both have no obvious problematic

up- and downstream regions that could hinder genetic manipulation. Interestingly, *agrB* also lies upstream of multiple tRNA regions that could hamper efforts in the creation of identity arms amplicons, discussed further in this chapter.

Table 3.2, Conserved AgrD ring structures and putative AgrB recognition sites in related acetogens: Locus tags present the following strains; CLAU: *C. autoethanogenum*, CLJU: *C. ljungdahlii*, CLRAG: *C. ragsdalei*, CSCA: *C. scatologenes*, CLCOS: *C. coskatii* PTA-10522. Ring structure region = yellow highlight, a Pro-X-X-Pro motif = green highlight. Consensus symbols represent; *= fully conserved residue, : = conserved residue with similar properties, .= conserved residue with similar properties

```
AgrD1
CLAU_0816       MKNLKKSLLSKTTKVVGSLSLFLAAIVIVPT STGGA YQPKCPDEFLK
CLJU_c28500     MKNLKKSLLSKTMKVVGSLSLFLAAIVIVPT STGGA YQPKCPDEFLK
CLRAG_03770     MKNLKKSLLSKTMKAVGSLSLFLAAIVITPA SLGMG HQPKCPEDLLK
CSCA_2234       MMFFKKNFSKKVLKTLGSVCLFLGTLAIVPT SYASG QQPKCPDELLK
CLCOS_10790     MKNLKKTLLSKIMKVVGSISLFLAAIVITPT SLGLG YQPKCPKELLK
                *  :**.: .*  *.:**:.***.:::*.*: * . .  *****.::**

AgrD2
CLAU_3094       MKNLKESVLKKSMKVVGCLSLFLEALVIVPA SAGCY HQPKCPDELLK
CLJU_c25570     MKNLKESVLEKSMKVVGKLSLGLAEIVITPA SSGCA YQPKCPDELLK
CLRAG_34580     MKNLKKNFLNKSMKMVGYLSLFLAALVITPA SAGCG HQPQCPDELLK
CSCA_2234       MMFFKKNFSKKVLKTLGSVCLFLGTLAIVPT SYASG QQPKCPDELLK
CLCOS_35050     MKNLKESVLEKSMKVVGKLSLGLAEIVITPA SSGCA YQPKCPDELLK
                *****:. :****:** *** *  :**.** *:*.  :**:*******
```

3.2.2 Development of the Δ*pyrE* strain

The initial step toward *agrD* mutant generation was developing a *pyrE* deficient strain of *C. autoethanogenum* which followed general and in-frame conjugation procedures outlined in 2.5.1 and 2.5.2. Colonies found to be growing exclusively on uracil-supplemented SDMM medium were picked and cultured for further gDNA extraction. This gDNA was used as a template for PCR screening to show the *pyrE* truncation and lacZα fragment integration only occurring in the *pyrE* deficient strain and absent in the WT, Figure 3.2. Primer M13F which anneals to the LacZα fragment that is not present in the WT, created a 634 bp amplicon along with primer pyrE_FSP, Figure 3.2.B. The second PCR reaction produced a band of approximately 2,064 bp for the WT and 2,101 bp for the *pyrE* mutants, Figure 3.2.A. This size discrepancy originates from a 227 bp partial deletion of the *pyrE* gene, replaced with a 264 bp LacZα fragment on the 3' end of *pyrE*, reducing the PCR fragment size differences between WT and Δ*pyrE* strains to 37 bp. These ~2,000 bp fragments were gel purified and sent for Sanger-Sequencing for full confirmation of

pyrE truncation. Whole-genome sequencing was performed later to determine SNVs and InDels.

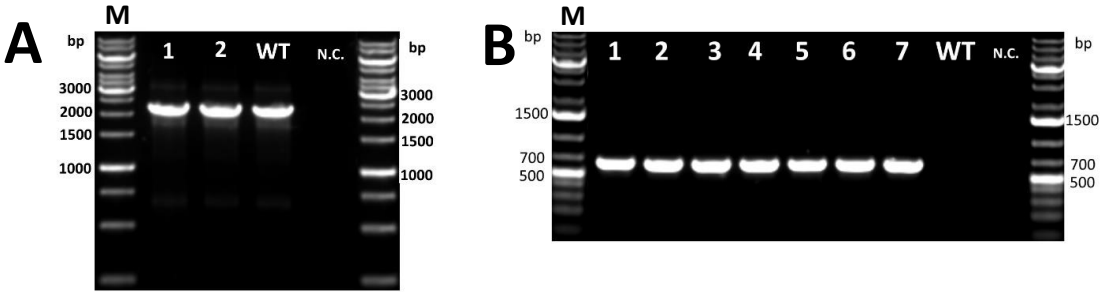


Figure 3.2, PCR screening of putative *ΔpyrE* mutants: (A) ~2 kB fragment amplifying the *pyrE* truncated region containing the *LacZα* fragment, lanes 1-2. (B), 634 bp fragment in *pyrE* mutants (1-7). WT = wild type control, N.C. = negative control (water). M = molecular weight ladder.

As an aside, producing the Δ*pyrE* background strain was undertaken by several members of our lab, to maximise the chances of producing a successful *pyrE*-deficient double-cross mutant with the least number of SNVs. This chapter has demonstrated successful creation of *C. autoethanogenum* Δ*pyrE* but ultimately the strain with the least amount of SNVs was created by Ronja Breitkopf (Uni. of Nottingham), and thus was used for the entirety of this project.

3.2.3 Construction of *agrD* mutants

Construction of *agrD* IFD plasmids followed steps outlined in 2.4.5 and is explained schematically in Figure 3.3. Plasmid pMTL84151 served as a backbone to either Δ*agrD1* or Δ*agrD2* insert, each spanning, 1,270 bp and 1,437 bp respectively. As the genes were of a relatively small, 144 bp size, in-frame deletion would remove approximately 80 % of either, leaving a non-functional 30 bp genetic truncation in the chromosome. Conjugation procedures followed as described in 2.5.1 using the *E. coli* CA434 donor.

Initial attempts proved unsuccessful as conjugation efficiencies were exceptionally low. This problem was circumvented by multiplying conjugation reactions 6-fold, increasing the likelihood of conjugative DNA transfer. Conjugated cell growth from each mating plate was harvested, resuspended and divided onto 3 selective YTF agar plates containing D-cycloserine and thiamphenicol. Increasing the number of selection plates elevated chances of successful plasmid conjugation, but proportionally, yields of *C. autoethanogenum* colonies remained relatively low. With more colonies, the likelihood

of obtaining single-over mutants was improved. Later in the project, *E. coli* Sexpress, was substituted as the donor strain, (Table 2.1) with greatly improved conjugation efficiencies. The same volumes of mating plates were retained to enable a wider selection of independent mutants further into the transformation procedure.

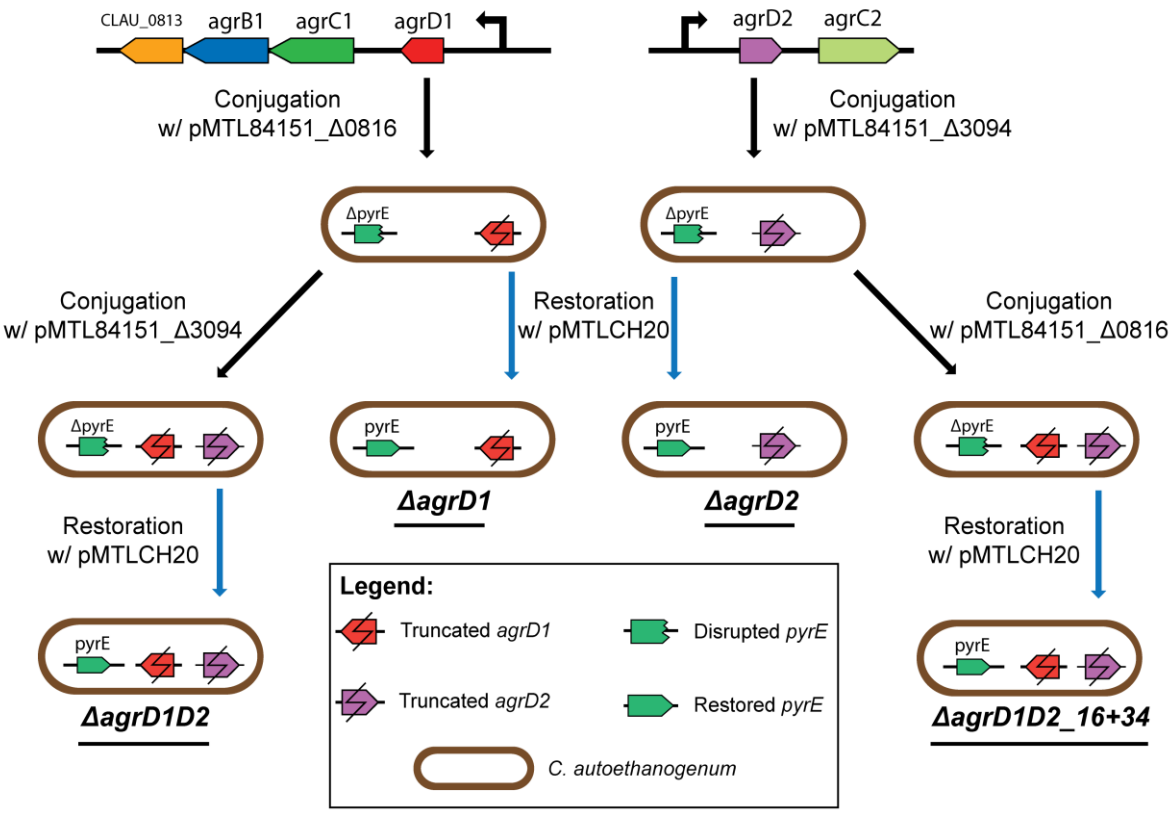


Figure 3.3, Work flow of agrD mutant creation in *C. autoethanogenum*: Black arrows represent transformation via conjugation using plasmids with targeting knock-out cassettes; pMTL84151_Δ0816 (Δ*agrD1*) and pMTL84151Δ3094 (Δ*agrD2*) respectively. Restoration of *pyrE* function was done using plasmid pMTLCH20. The step is represented through blue arrows. Finished, "*pyrE*-corrected" mutant strains are highlighted in bold and underlined.

Analysis of successful single-cross mutants followed procedures outlined in 2.6.2.1, and PCR screening confirmed plasmid integration via left and/or right homology arms, Figure 3.4. A third screening which used gene flanking primers that amplified the whole integration region was also carried out to reveal if there was WT contamination, Figure 3.4. Absence of a band suggested successful plasmid integration as amplicon size was too large, whereas presence of a band implied there was no integration event and the WT genotype persisted. During routine mutant creation of either *agrD* gene, a portion of single-cross integrants contained WT background DNA, which required further isolation and screening. It was observed that even after several iterations of re-isolation and screening, very faint WT gel banding would be still observed, Figure 3.4. Selection for double-cross mutants involved treatment with FOA (to select for plasmid backbone loss via a second crossover), followed by screening on uracil +/- SDMM media which yielded colonies that were PCR amplified and Sanger sequenced. This PCR screening procedure used flanking primers of each *agrD* gene, with predicted amplicons sizes of 1,270 bp and 1,437 bp for *ΔagrD1* and *ΔagrD2* mutants respectively, Figure 3.5.A, lane 1 and 2 left and right panel, respectively. WT amplicon sizes were 1,601 bp and 1,734 bp, for the *agrD1* and *agrD2* locus, as shown in the 'WT' lanes of Figure 3.5.A left and right panel, respectively. Colonies found to revert to WT genotypes were discarded. The final step of mutagenesis was achieved through *pyrE* restoration in both mutants, Figure 3.3.

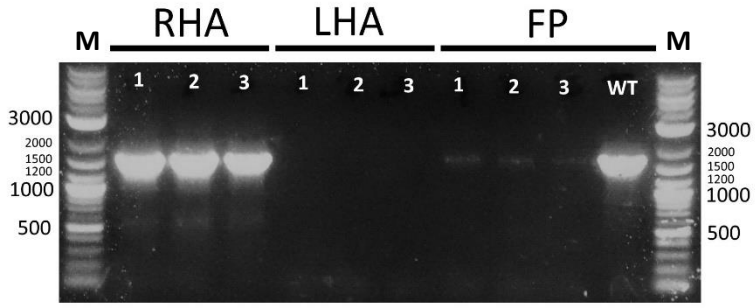


Figure 3.4, Screening for single-cross integration events and detection of WT contamination. RHA, LHA, lanes used primers (0816_SCR_F/3094_SCR_F with 84151_R, or 0816_SCR_R/3094_SCR_R with 84151_F), on the integrant plasmid and within the chromosome of the organism to detect crossover events via right and left homology arms, respectively; FP, used primers only on the chromosome, outside the region of integration (0816_SCR_F + 0816_SCR_R, or 3094_SCR_F + 3094_SCR_R). No band refers to successful integration, a band represents a persistent WT genotype. WT= Wild type control, M = molecular weight ladder.

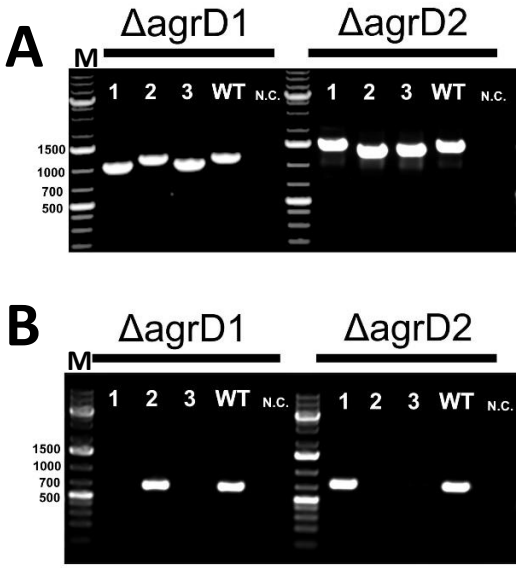


Figure 3.5, PCR screening of *agrD* mutants: (A) Two gels showing *agrD1* (left) and *agrD2* deletions (right) in mutant strains *ΔagrD1*, *ΔagrD2* and *ΔagrD1D2*, represented in lanes 1, 2 and 3, respectively. (B) "Quick-Screening" gel technique for *agrD 1* (left) and *agrD2* deletions (right), showing bands only in strains where either *agrD* gene persists. *ΔagrD1D2* shows neither intact *agrD*. WT = Wild type control, N.C. = Negative Control (water). M = molecular weight ladder.

For the creation of a *ΔagrD1D2* double mutant, stocks of *ΔagrD1_ΔpyrE* and *ΔagrD2_ΔpyrE* mutant strains were used; the strategy is schematically explained in Figure 3.3. Similar screening methods were employed by using flanking primers of each *agrD* gene, to which 1,270 bp and 1,437 bp amplicons were produced for *agrD1* and *agrD2* respectively, Figure 3.5.A, lane 3.

Figure 3.5.A shows small band size differences between *ΔagrD1*, *ΔagrD2*, *ΔagrD1D2* and WT strains, requiring extended gel running times. A more decisive screening method was implemented, Figure 3.5.B, whereby reverse primers, N_R0816 and N_ R3094, annealed at the centre of each *agrD* gene, producing 755 bp and 842 bp amplicons in combination with each gene's forward flanking primers 0816_SCR_F and 3094_SCR_F, respectively. The resulting amplicon would only exist if the respective gene is present in the chromosome. This method was used extensively to detect cross-contamination in growth experiments. The advantage of this method is faster gel running time, and clearer distinction between mutant and WT strains.

3.2.4 Restoration of *pyrE*

Mutant strains that possessed a truncated *pyrE*, required restoration to allow mutants to regain uracil autotrophy. Therefore, similar methods as described in 2.6.3.3 were

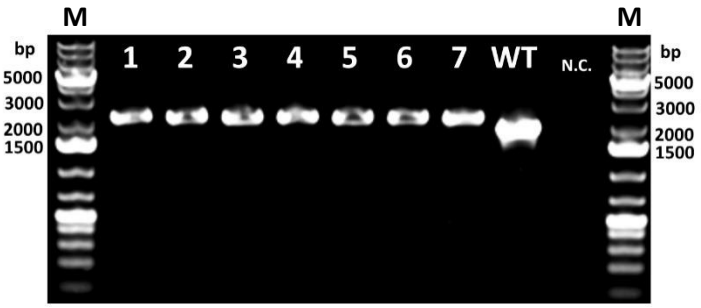


Figure 3.6, PCR screening of *pyrE* restoration: Amplification of restored *pyrE* mutants, the increased size confirming presence of the remaining *lacZα* fragment (lanes 1-7). WT = wild type control, N.C. = negative control (water). M = molecular weight ladder.

employed to conjugate the *pyrE* complementation plasmid, pMTLCH20, into mutants. Successful transformants were screened for uracil autotrophy on uracil-negative SDMM medium. Plasmid pMTLCH20 also carried the *lacZα* fragment placeholder fragment, allowing it to remain downstream of the 3' end of restored *pyrE* gene. Figure 3.6 shows restoration of *pyrE* back into a WT genotype using pMTLCH20, (materials and methods 2.4.5). Size discrepancy is due to *pyrE* restoration, adding 227 bp alongside the remaining *lacZα* fragment (264 bp).

This technique was also implemented in mutant complementation, described in 2.5.2.3, by replacing the pMTLCH20 *lacZα* fragment with the desired complementation gene insert. Successful transformants homologously integrated the complementation fragment chromosomally behind the *pyrE* gene, allowing a mutated genotype reversion back to WT. This strategy was preferred over plasmid complementation, as it eliminated the need for plasmid control and continual antibiotic maintenance. In addition, it provides a more stable transcription of complemented genes, without the hassle of matching parent strain gene transcription levels. Results of complementation are found under section 4.2.2.

3.2.5 Complementation of AgrD mutants

Strategies for complementation considered all routes for returning *ΔagrD1D2* back into *ΔagrD1* or Δ*agrD2* genotype. Phenotype characterization described in detail in chapter 4,

revealed that Δ*agrD1* and Δ*agrD2* both exhibited phenotypically similar profiles to the WT in contrast to the *ΔagrD1D2* strain. Therefore, it was pertinent, to use Δ*agrD1D2_ΔpyrE* and reverted it back to a *ΔagrD1* and Δ*agrD2* genotype, respectively.

Chromosomal complementation was employed over plasmid complementation to exploit the advantage of the ACE, and *pyrE* selection methods. The plan to restore *pyrE* in tandem with either *agrD* gene required the creation of plasmids pMTLCH20-D1 and pMTLCH20-D2 (Table 2.2). The proven strategy was to create both complementation plasmids without native or exogenous promoters and recombine respective correction fragments

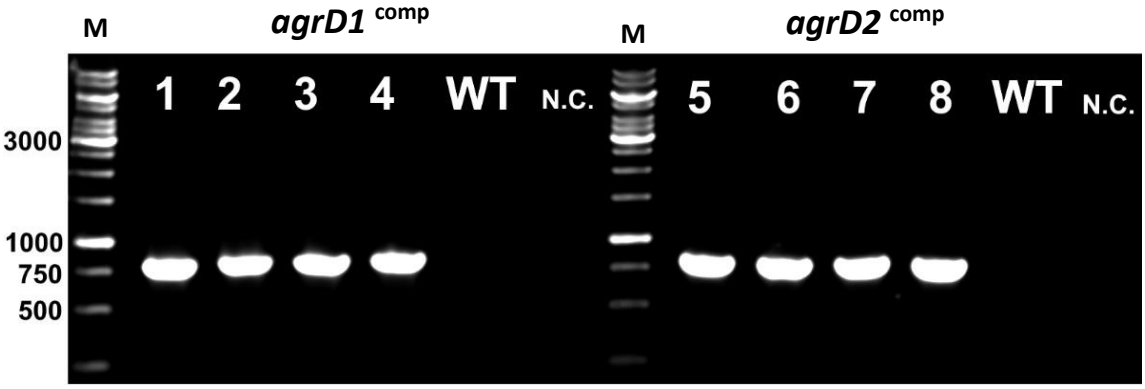


Figure 3.7, PCR screening of *agrD* complementation strains: *ΔagrD1D2* was complemented separately with pMTLCH20-D1 (lanes 1-4) and pMTLCH20-D2 (lanes 5-8), which resulted in 854 bp and 856 bp amplicons respectively. WT = Wild type control, N.C. = Negative Control (water), M = molecular weight ladder.

directly behind the *pyrE* gene, replacing the LacZα fragment. The complemented gene would then be expressed through the *pyrE* promoter. With this approach, either *agrD* gene was cloned into pMTLCH20 and conjugated into *ΔagrD1D2_ΔpyrE* using *E. coli Sexpress* as a donor, thus creating complementation strains *ΔagrD1D2_D1*^{comp} and *ΔagrD1D2_ D2*^{comp}. Screening procedures followed similarly to Figure 3.2; primer M13F was used in conjunction with *agrD* internal screening primers, N_0816 and N_3079, Figure 3.7., confirming successful generation of complementation strains.

It is worth noting that during complementation procedures, various strategies were met without much success. This included plasmid complementation using a newly characterised, highly stable "P19" replicon isolated from *C. carboxidivorans*, (Millard, 2018), which would avoid the need for antibiotic maintenance. Other attempts tried using native promoters of *agrD1* and *agrD2* as well as plasmid-encoded promoters in order to implement, without favourable results. These failed attempts raised questions about the nature of *agr* promoters and whether recombinantly-expressed immature

AgrD peptides may in fact be toxic to an *E. coli* host. These attempts are explained in greater detail in the discussion.

3.2.6 Whole genome sequencing

Illumina next generation sequencing was subsequently performed after all required mutations and *pyrE* restorations were completed. In-silico work involved assembling raw Illumina data reads against the latest published *C. autoethanogenum* reference sequence, (Humphreys *et al.*, 2015). Sequencing data was analysed using CLC Genomics Workbench. Tables 3.3.A and B list SNVs and InDels in each of the mutant strains. A summary report describing the quality of read-mapping is listed in the appendix, section 7.2.

In Table 3.3.A, a comprehensive overview is given of the conducted sequence analysis which details SNVs which have led to non-synonymous mutations, thus altering amino acid sequence and potentially affecting protein function. The original parental WT Δ*pyrE* strain was *pyrE* restored (designated WT_*pyrE*^{fix}) and included SNVs separately generated during original Δ*pyrE* mutant creation. It could be shown that the WT_*pyrE*^{fix} contained 4 SNVs, which were also present in all derived mutant strains, except for the gene representing amino acid permease (CLAU_0223).

Table 3.3.A, SNV variants found in WT, single, double knock-out and complemented *agrD* mutants: All read-mapping was performed against published *C. autoethanogenum* DSM10061 reference genome, (Humphreys *et. al.,* 2015). SNVs are grouped according to presence in respective strains. "Universal SNVs" refers to SNVs found exclusively in *agrD* mutants. Strain unique SNVs are denoted with i – vii superscript. SNVs that have been carried over to subsequent mutations are denoted in brackets in addition to the strain's own SNVs. Complemented mutant Δ*agrD1D2_D2*^{comp} contains no unique SNVs, and inherits Δ*agrD1D2* SNVs only. WT *pyrE*^{fix} = Wild type with restored *pyrE*, *C. a.* = *C. autoethanogenum*, SNV = Single Nucleotide Variant, Del = Deletion, * = stop codon, F.S. = Frame shift, NCR = non-coding region DS = Downstream, US = Upstream. (Table continues next page).

Pos. in genome	Type	Allele Change	Freq. (%)	Locus tag/ A.A. change	Protein affected
WT *pyrE*^{fix} and *agrD* Mutants					
243106	SNV	C→A	100	**CLAU_0223**: Val167Leu	Amino acid permease **(SNV not present in mutants)**
2570376	SNV	G→T	100	**CLAU_2244 (*pssA*)**: Ala36Asp	Serine O-phosphatidyl-transferase
4317806	SNV	G→A	69	**CLAU_3935**: Ala523Val	Collagen triple helix repeat-containing protein
4318262	SNV	G→A	87	**CLAU_3935**: Thr371Ile	Collagen triple helix repeat-containing protein
Universal SNVs in all *agrD* mutants					
1044274	Del	T→-	97	**CLAU_0940**: Asn15F.S.	NADPH-dependent FMN reductase
2632305	SNV	A→G	100	**CLAU_2395**: Lys131Glu	Glycosyl transferase family 2

3552387	SNV	T→A	100	*NCR*	1 base pair DS of Alanine-tRNA ligase, CLAU_3214 (*alaS*)
C. a. ΔagrD1 [i]					
477182	SNV	C→A	100	**CLAU_0419**: Arg583Met	Molybdopterin-family oxidoreductase alpha subunit
1083237	SNV	A→G	69	**CLAU_0977**: Ile308Thr	Glyoxylate reductase (NADP(+))
2613209	SNV	G→T	100	**CLAU_2378 (*pyk*)**: Gly556*	Pyruvate kinase
3030068	SNV	G→A	89.3	**CLAU_2714**: Ala356Thr	Molybdenum cofactor biosynthesis domain-containing protein
4204139	SNV	G→T	100	**CLAU_3830**: Val252Phe	Na(+) dependent nucleoside transporter domain protein
Universal SNVs in Δ*agrD2* mutants					
1359001	SNV	C→A	100	**CLAU_1232 (*apt*)**: Val164Leu	Adenine phosphoribosyltransferase
3306086	SNV	C→T	100	**CLAU_2960 (*fliM*)**: Gln7*	Flagellar motor switch protein FliM
C. a. ΔagrD2_2 [ii]					
477181	SNV	C→A	100	**CLAU_0419**: Arg583Ser	Molybdopterin-family oxidoreductase alpha subunit
2613206	SNV	G→T	99.2	**CLAU_2378 (*pyk*)**: Ala555Ser	Pyruvate kinase
C. a. ΔagrD2_9 [iii]					
453159	SNV	C→T	100	**CLAU_0398 (*vnfE*)**: Asp392Asn	vnfN-like nitrogenase
483602	Del	A→-	90	**CLAU_0424**: Phe108F.S.	Transcriptional regulator Crp/Fnr family
3423047	SNV	G→T	100	**CLAU_3073**: Val83Phe	Hypothetical protein
C. a. ΔagrD2_27 [iv]					
356088	SNV	C→T	47.83	**CLAU_0322**: Gly387Glu	Cell wall binding repeat 2-containing protein
1744062	SNV	C→A	66.28	**CLAU_1579**: *401Leu	Carbon-monoxide dehydrogenase (acceptor)
2666275	SNV	G→T	100	**CLAU_2422**: Ala441Ser	Amino acid permease-associated protein
3453526	SNV	G→A	48	**CLAU_3110**: No Change	Stage III sporulation protein AA
C. a. ΔagrD1D2 [v]					
1952079	SNV	T→C	40	**CLAU_1765**: Glu4Gly	C_GCAxxG_C_C family protein
1980129	SNV	C→T	100	*NCR*	178 bp US of (CLAU_1794) Alcohol Dehydrogenase
2025343	SNV	G→T	100	**CLAU_1836**: Gln393Lys	PAS modulated sigma54 specific transcriptional regulator Fis family
2041632	SNV	A→C	100	*NCR*	101 bp US of (CLAU_1848) Phosphoenolpyruvate-protein phosphotransferase
2120989	SNV	G→T	100	*NCR*	Between (CLAU_1932) UDP-N-acetylmuramoylalanine-D-glutamate ligase (MurD) and (CLAU_1933) Glycine-tRNA ligase (glyQS)
2167303	SNV	G→A	100	**CLAU_1974**: Ala23Val	Helix-turn-helix domain protein
2214740	SNV	T→C	100	**CLAU_2022**: No Change	Hypothetical protein
2219133	SNV	C→T	99.11	**CLAU_2025**: Gly229Ser	Hypothetical protein
2236445	SNV	G→A	100	**CLAU_2039**: Ser214Leu	Methyl-accepting chemotaxis sensory transducer
2445880	SNV	G→T	100	**CLAU_2235:** Gly602*	Cell wall hydrolase/autolysin
4204139	SNV	G→T	100	**CLAU_3830**: Val252Phe	Na(+) dependent nucleoside transporter domain protein

479613	SNV	A→T	100	**CLAU_0420**: Tyr223*	Major facilitator superfamily MFS_1 (membrane-bound protein)
***C. a. ΔagrD1D2*_16 & 34** vi + (ii)					
477181	SNV	C→A	100	**CLAU_0419**: Arg583Ser	Molybdopterin-family oxidoreductase alpha subunit
813181	SNV	G→A	100	**CLAU_0726:** Asp84Asn	Hypothetical protein
***C. a. ΔagrD1D2_ D1*^{comp}** vii + (v)					
2526258	SNV	G→T	100	**CLAU_2306:** Gly175Trp	Peptide chain release factor 2

Table 3.3.B, Unique InDel regions: InDels present only in double knockouts, *ΔagrD1D2*, *ΔagrD1D2*_16 & 34 and complemented mutants *ΔagrD1D2_D1*^{comp} and *ΔagrD1D2_D2*^{comp}. Expected insertions (LacZα) and complemented *agrD* genes and expected deletions of *agrD1* and *agrD2* are omitted from table.

Position	Deletion	Freq. (%)	Locus	Protein
C. a. ΔagrD1D2				
2848745 - 2848765	GGAGGAGCTTATGGCTATTTA	61	**CLAU_2572**	Type IV pilin N-term methylation site GFxxxE family protein
3389889 - 3389899	TAGGAATACCT	99	**CLAU_3044**	Flagellar biosynthetic protein fliR/flhB

In both *ΔagrD1* and *ΔagrD2*, the number of SNVs vary. This is possibly due to *ΔagrD2* creation requiring multiple rounds of screening for single and double crossover events as described in 3.2.3, that may have led to a greater chance of accumulating spontaneous SNVs, (Miller, 1996; Foster *et al.*, 2015). It was also likely that SNVs were generated during *pyrE* restoration, as for example, the parental *ΔagrD1* strain carries unique SNVs not passed on into *ΔagrD1D2*. A significantly amount number of SNVs were generated in *ΔagrD1D2_16* & *34*, this is potentially due to the substantially faster and more efficient conjugation and transformation methods using NEB *Sexpress*. This was also apparent for the *agrB* ClosTron (*agrB::CtermB-234a*, see 3.2.7), where no newly generated SNVs were detected.

Analysis of InDels in mutants, comprised of expected *agrD* gene deletions and LacZα fragment insertions, were not included in table 3.3.B. Unexpected structural variants in *ΔagrD1D2*, affected "Type IV pilin N-term methylation site GFxxxE family protein", (CLAU_2572), and "flagellar biosynthetic protein FliR/FlhB", (CLAU_3044). In the Type IV pilin N-term methylation site protein, in-frame removal of residues GGAYGYL occurred at 67th to 73rd residue positions within the protein, and no deleterious frameshift occurred. Conversely, FliR/FlhB experienced a frame-shift in form of a 11-nucleotide deletion that occurred at the central 182nd residue, resulting in deleterious changes for the rest of the protein.

Overall, changes to amino acid residues were studied on a case-by-case basis as certain substitutions were synonymous on protein structure and function. Table 3.4 describes SNVs that have substituted residues, which have led to property changes, i.e.: hydrophobic to polar, charged to uncharged, and so on. Evaluation from literature determined if they are potentially detrimental to overall protein structure (Betts and Russell, 2007). In addition to this assessment, the Sorting from Intolerant to Tolerant (SIFT) program (section 2.9.1) (Ng and Henikoff, 2003), was used to determine if residue substitutions will be tolerated in the protein, and if not, do they raise the probability of changing its respective function.

In most of cases, residue changes in listed proteins are predicted to be tolerated and assumed to have not have any further functional changes on these respective proteins. Interestingly, in some cases residues were still predicted as intolerable despite unchanged properties. For example, the change from valine to phenylalanine is regarded this way, and this is likely due to the aliphatic vs. aromatic structural nature of these amino acids, with differences in hydrophobicity that present a significant variance, which lead to incompatible substitutions (Betts and Russell, 2007).

Table 3.4, Determination of tolerated coding substitution changes in non-synonymous residues: Only coding, non-synonymous residue changes were considered. SIFT was used to determine residue change tolerance. S. case = special case amino acids (cysteine, selenocysteine, glycine and proline), * = Occurring in WT *pyrE^{fix}* only. (Table continues next page).

Locus Tag	Protein	A.A. change	Property change	Substitution Tolerated?
WT pyrE^{fix} and Mutants				
CLAU_0223*	Amino acid permease	**Val167Leu**	*No change*	**Yes**
CLAU_2244 (*pssA*)	Radical SAM domain protein / CDP-diacylglycerol/serine O-phosphatidyl-transferase	**Ala36Asp**	*Hydrophobic → Neg. charge (polar)*	**No**
CLAU_3935	Collagen triple helix repeat-containing protein	**Ala523Val**	*No change*	**Yes**
CLAU_3935	Collagen triple helix repeat-containing protein	**Thr371Ile**	*Uncharged Polar → Hydrophobic*	**Yes**
Mutants only				
CLAU_2395	Glycosyl transferase family 2	**Lys131Glu**	*Pos. charge → Neg. charge*	**No**
CLAU_3830	Na(+) dependent nucleoside transporter domain protein	**Val252Phe**	*No Change*	**No**
C. a. ΔagrD1				
CLAU_0419	Molybdopterin-family oxidoreductase alpha subunit	**Arg583Met**	*Pos. charge → Hydrophobic*	**No**
CLAU_0977	Glyoxylate reductase (NADP(+))	**Ile308Thr**	*Hydrophobic → Uncharged Polar*	**Yes**
CLAU_2714	Molybdenum cofactor biosynthesis domain-containing protein	**Ala356Thr**	*Hydrophobic → Uncharged Polar*	**Yes**
C. a. ΔagrD2				

CLAU_1232 (*apt*)	Adenine phosphoribosyltransferase	**Val164Leu**	*No change*	**Yes**
CLAU_0419	Molybdopterin-family oxidoreductase alpha subunit	**Arg583Ser**	*Pos. charge → Uncharged Polar*	**No**
CLAU_2378 (*pyk*)	Pyruvate kinase	**Ala555Ser**	*Hydrophobic → Uncharged Polar*	**No**
CLAU_0398 (*vnfE*)	vnfN-like nitrogenase	**Asp392Asn**	*Neg. charge → Uncharged Polar*	**Yes**
CLAU_3073	hypothetical protein	**Val83Phe**	*No Change*	**Yes**
CLAU_0322	Cell wall binding repeat 2-containing protein	**Gly387Glu**	*S. case → Neg. charge*	**No**
CLAU_2422	Amino acid permease-associated protein	**Ala441Ser**	*Hydrophobic → Uncharged Polar*	**Yes**
C. a. ΔagrD1D2				
CLAU_0726	Hypothetical protein	**Asp84Asn**	*Neg. charge → Uncharged Polar*	**Yes**
CLAU_1765	C_GCAxxG_C_C family protein	**Glu4Gly**	*Neg. charge → S. case*	**No**
CLAU_1836	PAS modulated sigma54 specific transcriptional regulator Fis family	**Gln393Lys**	*Uncharged Polar → Pos. charged*	**No**
CLAU_1974	Helix-turn-helix domain protein	**Ala23Val**	*No change*	**Yes**
CLAU_2025	Hypothetical protein	**Gly229Ser**	*S. case → Uncharged Polar*	**Yes**
CLAU_2039	Methyl-accepting chemotaxis sensory transducer	**Ser214Leu**	*Uncharged Polar → Hydrophobic*	**Yes**

3.2.7 ClosTron mediated inactivation of agrB

Different approaches in validating observed phenotypic changes found in strain *ΔagrD1D2* (see chapter 4), were sought. It was hypothesized that knock-out of *agrB* would prevent the organism from processing and exporting both AgrD peptides, effectively disabling the native Agr QS system. As there was only one putative copy of the *agrB* gene, a single knock-out would be sufficient. In-frame deletion techniques were once more employed, but difficulty in amplifying the left homology arm, potentially due

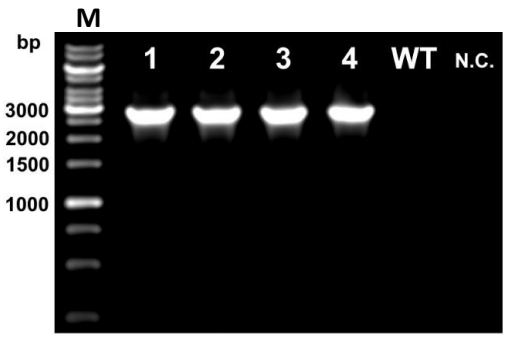


Figure 3.8 PCR screening of *agrB* ClosTron: Amplification of ClosTron insertion site with *agrB* forward flanking primer, (2988 bp). WT = Wild type control, N.C. = Negative Control (water), M = molecular weight ladder.

to the multiple downstream tRNA regions which may have impeded further progress. The alternative strategy was to perform ClosTron disruption of *agrB*.

Two ClosTron plasmids, pMTL007C-E2::CAUagrB-234a and pMTL007C-E2::CAUagrB-232s were synthesized by ATUM (formally DNA 2.0). The encoded introns targeted two different sites within the gene and covered both anti-sense and sense intron orientations. This strategy was employed to minimize the risk of potential polar effects up and downstream of *agrB*. Both plasmids were conjugated into *C. autoethanogenum_ΔpyrE* as described in 2.5.1, and colonies were replica streaked onto thiamphenicol +/- plates. Colonies that did not propagate on thiamphenicol media were deemed to have lost the plasmid and where therefore candidates for further screening. Conjugations using the sense plasmid did not result in ClosTron integrated colonies, as screening showed a WT *agrB* genotype. For putative mutants containing the antisense construct, flanking primer agrB_F1 and internal ClosTron universal primer, EBS2, were used for screening for the insertion, with an expected amplicon size of 2,988 bp. Figure 3.8 shows successful ClosTron insertion, which was further confirmed through Sanger Sequencing. Whole genome sequencing was carried out to ensure a single ClosTron insertion occurred in the desired locus, which was found to be correctly inserted in the desired location. The complementation of ClosTron mutants was not pursued as further analysis discussed in the next chapter revealed no expected phenotypic changes.

3.3 Discussion

The central focus of this chapter was the creation of *agrD* mutants in *C. autoethanogenum* primarily using established ACE techniques (Heap *et al.*, 2012; Ng *et al.*, 2013). The Agr system, first described fully in *S. aureus*, is classically comprised of 4 components that are conserved across firmicutes (Wuster and Babu, 2008). However, in some cases, species possess several systems and/or different permutations of conserved *agr* genes as is the case in *C. autoethanogenum*. This resulted in both *agrD* genes as favourable targets for disruption. Several other acetogens also shared these distinguishing features as outlined in table 3.2. The IFD strategy removed up to 80 % of each gene which included conserved ring structure regions and putative ArgB interaction sites effectively abolishing *C. autoethanogenum*'s Agr signalling capabilities.

Different approaches were considered in undertaking this task and ACE was chosen for its relative ease in targeting multiple gene regions and providing a convenient route to

restoring mutants back to a WT-genotype. Mutant creation began with the construction of the *pyrE* background strain and although work to procure AgrD mutants was ultimately successful, the IFD methods were noted to have several drawbacks. After plasmid introduction via conjugation, faster growing cells were selected and restreaked allowing isolation of single cross-over plasmid integrants. Selection was further determined by PCR. Isolation of pure single cross was made especially difficult by the presence of contaminant WT colonies that reduced the likelihood of isolating a pure single cross mutant. This study found that entirely removing WT background contamination proved difficult, if not impossible, and measures to lower this background included repeatedly re-streaking single-cross integrant stocks onto selective media. This posed a two-fold problem; firstly, re-isolation attempts prolonged the procedure considerably as *C. autoethanogenum* colonies appear on average after 4-5 days and secondly, the probability of acquiring spontaneous SNVs increased as speculated in 3.2.6. Moreover, further steps requiring FOA treatment for double-cross integration selection were noted to produce approximately a 50% likelihood of reversion back to a WT-genotype.

Conversely, ClosTron implementation evidently proved to be effective, stable and rapid. ClosTron plasmids were synthesised and a confirmed ClosTron insertion was achieved in under 2 weeks. Nevertheless, this method also faces several limitations. ClosTron plasmids can be synthesised in pairs, to target sense and antisense DNA of a target gene. This is done to avoid unintended polar effects that may interfere with the expression of genes up- or downstream of the target which is a possibility when using ClosTron, (Heap *et al.*, 2012; Kuehne and Minton, 2012). Other risks include ectopic insertion into the genome (Shao *et al.*, 2007; Pyne *et al.*, 2014), and studies have remarked that this technique merely disrupts a gene rather than deleting it outright which can lead to continued residual activity (Al-Hinai, Fast and Papoutsakis, 2012; Y. Wang *et al.*, 2013; Wang *et al.*, 2015).

Due to time constrains, it was only possible to achieve *agrB* insertion with the plasmid, pMTL007C-E2::CAUagrB-234a. This insertion was verified by PCR screening, Sanger sequencing and later NGS. However, preliminary phenotypic experiments aiming to achieve comparable phenotypes to the *ΔagrD1D2* mutant (as described in chapter 4), showed a vastly different outcome which did not align with an earlier predicted hypothesis, and this strategy was abandoned. The conclusion was a probable promoter read-through effect exerted by the intron's erythromycin RAM marker, causing a polar effect downstream of *agrB*.

Curiously, downstream of this singular *agrB* gene lies a "spo0E-like" gene encoding a putative regulatory phosphorelay protein, that presently, can only be speculated to function similarly to the missing AgrA. As an aside, the Spo0E phosphatase and its corresponding family of phosphatases are documented in the *Bacillaceae* family, and have been shown to be responsible for the inactivation via phosphorelay of the master sporulation factor, Spo0A (Ohlsen, Grimsley and Hoch, 1994; Perego, 2001). Similar phosphorelays have not been observed in *Clostridia* spp., (Dürre and Hollergschwandner, 2004). In *C. acetobutylicum*, Spo0E-equivalent phosphatase have been shown to be represented by kinase CA_C0437, which when deactivated, accelerated sporulation (Steiner *et al.*, 2011). Despite this insight, it is still not clear what the exact purpose of CLAU_0813 in *C. autoethanogenum* is, and whether it is linked with the Agr system.

At the beginning of this study, efforts to introduce plasmids into *C. autoethanogenum* were found troublesome as conjugation efficiencies were remarkably low. It was speculated that this hindrance involved a previously uncharacterized Type IV restriction system in *C. autoethanogenum* which is not present in its close relative *C. ljungdahlii* (Leang *et al.*, 2013). The Type IV restriction system was the probable reason as to why *C. autoethanogenum* is not suitable for electroporation (Liew, 2015). Further insight into this issue meant understanding this restriction system and ways of circumventing it. An earlier curated version of the *C. autoethanogenum* DSM 10061 genome by Utturkar *et. al.* (2015) used PacBio NGS methods which also employed analysis using Single Molecule Sequencing Technology (SMRT) that generated read lengths up to 15,000 bp (Koren *et al.*, 2013). This process enabled recognition for DNA base modifications related to methylation patterns, (Utturkar *et al.*, 2015). This data was processed and made available through the NEB REBASE website (http://tools.neb.com/genomes/, last accessed 12/09/17) (Roberts *et al.*, 2005).

Circumventing this system was through understanding the methylation system in CA434 (*E. coli* HB101 containing pR702). It was noted that commercial HB101 strains have both functioning *dam* and *dcm* methyltransferases, which methylate cytosine (*dcm*) and adenosine (*dam*) in plasmid DNA, leading to recognition by Type IV restriction systems that target methylated DNA with low specificity (Palmer and Marinus, 1994; Loenen and Raleigh, 2014). The next idea was to replace CA434 with a different donor that used an alternative methylation profile, or preferably none. Although commercial dam^-/dcm^- strains are viable alternatives, *dam* is required for modulation and repair of DNA mismatches, a necessary feature when handling plasmid DNA targeting chromosomal

DNA. However it is not yet fully understood what important role *dcm* may play in (Palmer and Marinus, 1994; Marinus and Casadesus, 2009), therefore a compromise in conjugation was met. Work undertaken by Craig Woods (Uni. of Nottingham) selected New England Biolab's *E. coli* NEB express, a BL21 derivative, which was dam^{+}/dcm^{-} and transformed it with plasmid R702, creating *E. coli* Sexpress (Woods, 2018). Experiments were performed to test transformation frequency with commonly used pMTL8000 shuttle vector replicons, (Heap *et al.*, 2009); pCD6 (Purdy *et al.*, 2002), pCB102 (Minton and Morris, 1981), p19 (Millard, 2018), pIP404 (Sloan *et al.*, 1992) that showed a significant increase in conjugation efficiency ranging from a 15 to 750-fold difference in colonies per plate, (Woods, 2018). However, this improvement came about after *pyrE* and *agrD* mutants for this project had already been created. Nevertheless, this solution significantly improved remaining *C. autoethanogenum* conjugations for this project, including the generation of the complementation strains $\Delta agrD1D2\_D1^{comp}$ and $\Delta agrD1D2\_D2^{comp}$.

The ACE technique bears an advantage for straightforward chromosomal complementation procedures. Alterative complementation options included chemical complementation that employed synthetic peptides (section 4.2.2.4) and classic plasmid-based complementation. The latter strategy was attempted with the aid of a newly discovered, highly stable replicon (p19) isolated from *Clostridium Carboxidivorans P7* (Millard, 2018). Replicon p19 was tested in various Gram-positive species; *C. autoethanogenum*, *C. beijerinckii*, *E. limosum* and *B. subtilis*, and was shown to indefinitely retain plasmids without the need for antibiotic markers. This is advantageous as it eliminates the need for antibiotic maintenance in comparative experimental studies (Millard, 2018). The usage of this replicon was successfully implemented in other projects not included in this thesis. However, in this case it was abandoned when attempting to create *agrD* complementation plasmids due to repeated unsuccessful cloning attempts. The cause of these failures possibly indicated to unknown replicon copy number and/or *agrD* promoter toxicity, thus switching strategies to chromosomal complementation via *C. autoethanogenum* with a *pyrE* deficient background was favoured.

The initial design of these complementation plasmids included the respective native promoters of *agrD1* and *agrD2*, but repeated attempts again yielded no positive *E. coli* colonies. The next solution was to replace the native promoter with a high-expression, ferredoxin gene "P_{fdx}" promoter (Takamizawa *et al.*, 2004), that also proved unfruitful. Ultimately, the successful approach was to ligate promoter-less *agrD* amplicons into

pMTLCH20 creating a construct that would simultaneously restore *pyrE* and integrate respective *agrD* complementation fragments in place of the *laZα* fragment. Both plasmids were rapidly cloned, conjugated and transformed into *ΔagrD1D2_ΔpyrE*. A potential downside of this method was that *agrD* gene expression would be modulated via the *pyrE* promoter and may express either gene differently than that of their native promoters. However, phenotype complementation was observed and is described in 4.2.2.

These complementation efforts raised several intriguing questions about *agr* promoters and why they might be useful to study in future projects. Why were both *agrD1* and *agrD2* promoters toxic to a Gram-negative host? Moreover, why was replacing these promoters with P_{fdx} also detrimental to *E. coli* fitness? Both questions imply that transcribed AgrD peptides are toxic to *E. coli*. An example in literature showed heterologous expression of the *S. aureus agrBD* operon in *E. coli* is possible. However, it was concluded no mature peptides could be formed and reasons for this were not clear, (Zhang *et al.*, 2002). Later studies clarified this by showing *E. coli* heterologous expression of functional AgrD AIPs but again only from a tandem *agrBD*-expressing plasmid, (Thoendel and Horswill, 2009). Examples of only heterologously expressing AgrD have not yet been shown. These studies suggested that AgrB presence is essential in AIP maturation and is involved proteolytic processing of AgrD.

With this knowledge, could *C. autoethanogenum*'s AgrD be expressed in a Gram-negative host? If so, it can be hypothesised that high expression may lead to polypeptide mass-aggregation into inclusion bodies that can create a chaotic environment within the host, resulting in cell death (Baneyx, 1999; de Marco *et al.*, 2005).

Promoters of the *agr* operon have been best described in *S. aureus* which possesses two putative *agr*-associated promoters, P2 and P3. The former performs autocatalytic transcription of the *agr* operon, that is regulated by the phosphor-relay of AgrC and AgrA (Novick *et al.*, 1993; Gray, Hall and Gresham, 2013). This function closes the circuit between peptide signal capture and regulation but has not been studied in detail in *Clostridia* spp. Understanding how *agr* promoters operate in acetogenic species may open new possibilities in synthetic biology fields. A recent example demonstrated the transferal of the entire *S. aureus* Agr system into *Bacillus megaterium* which gives example of Gram-positive QS system flexibility (Marchand and Collins, 2016).

Studies that involve genetic manipulation increasingly benefit from undergoing NGS analysis to determine if the proliferation of SNVs and InDels may be contributing to an

unwanted influence on expected phenotypes. NGS has become standard practice in current investigations and arguably raises the standards of research. In this study, Illumina NGS was carried out on all knock-out and complemented mutant strains. As shown in Table 3.3, a combined total of 38 individual SNVs and InDels were generated between *WT_pyrE*^{fix}, Δ*agrD1*, Δ*agrD2_2/9/27*, Δ*agrD1D2*, Δ*agrD1D2_16/34* and Δ*agrD1D2_ D1*^{comp}. SNVs were found to increase in an organism with a progressively higher number of gene knockouts. However, not all SNVs were detected within gene regions and several changes consisted of silent mutations. By subtracting these inconsequential SNVs from the list, the remainder was judged based on respective residue changes. The SIFT program (Ng and Henikoff, 2003) calculated predictions of how substitutions are tolerated and if they may have a deleterious effect on the protein. However, with this data it is difficult to assume to what degree an intolerant substitution can influence its respective protein and to do so would require in-depth analysis of each protein's structure and function which is beyond the scope of this study.

Proteins that sustained frame-shifts and early stop codons needed to be scrutinised individually. For example, an early stop in CLAU_0420; Major facilitator superfamily MFS_1 was observed. In the genome, this is a single copy out of 28 that was affected. Similar examples were noted in other affected genes; CLAU_0424; Transcriptional regulator Crp/Fnr family (1 in 2 copies) and CLAU_0940; NADPH-dependent FMN reductase (1 in 7 copies). These proteins play varied roles in the organism, such as membrane transport of small solutes (Reddy *et al.*, 2012) transcriptional regulation to environmental changes (Körner, Sofia and Zumft, 2003; Zhou *et al.*, 2012) and riboflavin reduction (Duane and Hastings, 1975). More substantive changes were found in double knockout mutants with both deleterious and non-deleterious Indels in pilin and flagellar associated genes but it is not certain if they had any pronounced phenotypic change, as growth experiments described in chapter 4 show no difference when compared with the WT and exhibit a WT-like phenotype when complemented, section 4.2.2. Future work may also entail independent verification of SNVs that may be suspected of changing phenotypes.

Spontaneous mutations can arise and accumulate from a multitude of factors. Either through intrinsic polymerase errors, oxidative stress or exogenous and endogenous agents targeting DNA to name a few examples (Miller, 1996; Foster *et al.*, 2015). Other suggestions as to why *C. autoethanogenum* accumulates a high frequency of mutations was thought to be through FOA treatment during ACE. Although this claim is controversial

as its early described as a non-mutagenic agent (Szybalski, 1958), but later results found in experiments yeast, *Candida albicans*, claimed the opposite (Wellington and Rustchenko, 2005). Thus far, there has been no work done in prokaryotes addressing this issue, and it can be assumed that FOA likely does not play a significant role in mutagenesis. Moreover, later created mutants Δ*agrD1D2_16* & *34* also were exposed to FOA, which showed significantly less SNVs.

Genetic variability is the key component in evolution, and spontaneous mutation is a necessary facilitator in this activity (Foster *et al.*, 2015). It may be concluded that it is very difficult, if not impossible to avoid SNVs when subjecting strains to intensive targeted mutagenic approaches.

Chapter 4

Characterisation of *C. autoethanogenum agrD* mutants

4.1 Introduction

The Agr quorum sensing system has been identified and studied in a range of Gram-positive bacteria, with much focus on pathogens (Recsei *et al.*, 1986; Peng *et al.*, 1988; Ji, Beavis and Novick, 1997; Novick, 2003; Atkinson and Williams, 2009). In the *Clostridium* genus, the Agr system has been experimentally studied in *C. acetobutylicum*, *C. perfringens*, *C. difficile*, and *C. botulinum*, and mutualistic *C. sporogenes*, and found to be present in several others, (Wuster and Babu, 2008). How the Agr system works in each of these species varies, but different combinations of conserved agr genes have been are found in pathogenic clostridia that have been studied to function as regulators relating to virulence, toxin production, sporulation and biofilm formation (Cooksley et al., 2010; Martin et al., 2013; Ohtani and Shimizu, 2015; Vidal, Shak and Canizalez-Roman, 2015). These characteristics contribute to the pathogenicity of each respective species, and have contributed to the wider understanding of the Agr system circuit. However, in non-pathogenic clostridia these systems have not been properly examined.

As described in chapter 3, both genes *agrD1* and *agrD2* were selected for IFD as it was evident that *agrD* mutants would effectively disable the Agr system, and the likelihood of phenotypic change would be observed. References have pointed to *agrD*'s role in respective *Clostridium* species as described above, and this chapter would follow suit in understanding this signalling peptide's activity in *C. autoethanogenum*.

Originally, studies to understand *C. acetobutylicum*'s sporulation initiation lead to the conclusion that its Agr system plays a role in these regulatory networks, (Steiner et al., 2012; Jabbari et al., 2013). Disrupting the putative *agrA*, *agrC* and *agrB* genes resulted in significant changes to heat-resistance, and spore and granulose formation. Transcriptional factor Spo0A, contributing to sporulation initiation and solventogenesis, was hypothesized to be influenced by the response regulator AgrA. Interestingly, there

were no clear changes in acid and solvent production in the *agr* mutants, which lead to the hypothesis that AgrA-dependent modulation may influence the degree in which Spo0A is phosphorylated through orphan histidine kinases (Ravagnani *et al.*, 2000; Steiner *et al.*, 2011, 2012).

As mentioned in the general introduction, RNPP QS systems have also been studied in *C. acetobutylicum*, outlining its role in solventogenesis and sporulation (Kotte *et al.*, 2017), however, besides these two examples, publications concerning the field of Gram-positive QS in solventogenic bacteria are scarce. Hints found in literature pointed to QS being a possible factor in influencing solvent gene expression in *Clostridium saccharoperbutylacetonicum* (Kosaka *et al.*, 2007), but since have not been further investigated. More recent work done in *C. thermocellum* showed an unexpected up-regulation of an AgrD-type signal peptide, when faced with a xylose-induced stress. A synthetically produced version of the signalling peptide was also found to inhibit growth, even in the absence of xylose, which indicated a hypothetical link to sugar stress-response regulation, (Verbeke *et al.*, 2017). These few examples further motivate the need to further the knowledge of QS systems in industrially-relevant bacterial species.

C. autoethanogenum, has been industrially proven for its solvent production capabilities on gas substrates (Köpke and Liew, 2011; Köpke, Mihalcea, Liew, *et al.*, 2011). In addition, there exists a wealth of metabolomic, transcriptomic and enzymatic data (S. Wang *et al.*, 2013; Mock *et al.*, 2015; Marcellin *et al.*, 2016) which have helped in understanding energy conservation with respect to fermentation productions, and in defining metabolomic differences between sugar and gas substrates. When reanalysing Marcellin *et. al.*'s (2015) RNA-seq transcriptomics supplementary data, *agrD1* and *agrD2* exhibited a pattern in their expression with respect to sugar and gas substrate metabolism, and it was found that *agrD1* was 7-fold more upregulated compared to *agrD2* when grown on gas, and 4-fold more when grown on fructose (Marcellin *et al.*, 2016). Intriguingly, this agrD1/D2 disparity corroborates with recent studies from this lab that performed gas and pH-shift experiments on CO-fed continuous stir-tank reactors revealing a similar trend (Henstra, 2018). Interestingly, under these conditions *agrD1* was observed to be in the top 35 of the most expressed genes during these experiments, with transcript levels nearly 150-fold higher than those of *agrD2*.

These independent studies show similar trends of *agrD1* taking a notable lead in transcription levels when compared to *agrD2* and showed a likely role in to gas

metabolism. It has not yet been described how an *agr* QS mechanism might operate in an acetogen, and whether it functions indirectly to help steer the organism towards long-term survival. In this chapter, these questions are put forward and ask how the Agr system might contribute to regulating *C. autoethanogenum's* metabolism, sporulation and overall physiology.

4.1.1 Aims of this study

The purpose of this study was to characterize three newly created *C. autoethanogenum* Δ*agrD* mutants; Δ*agrD1*, Δ*agrD2*, and Δ*agrD1D2* that targeted the peptide encoding genes of the Agr system. Characterisation would examine physiological and the fermentation-product profiles between mutants and the wild type strain to establish any notable differences. Phenotypic experiments performed in *C. acetobutylicum agr* mutants by Steiner *et al.* (2012) would form an initial basis to start testing whether similar phenotypes are present in WT *C. autoethanogenum*, and if so, can they be later examined in *agrD* mutants? The specific goals of this chapter are to address the following key points:

- Determining basic growth characteristics and fermentation product profiles.

This is the first step in establishing whether there are any differences observed amongst Δ*agrD1*, Δ*agrD2* and Δ*agrD1D2* mutants against the WT strain when grown on sugar. If so, does having a tandem knockout of both *agrD1* and *D2* genes produce a more pronounced phenotypic effect? Do subsequent independently generated Δ*agrD1D2* mutants show similar phenotypes?

- Testing mutant growth and behaviour in relation to gas substrates

Gas fixation is mediated via the acetogen's WL pathway, which is also capable of re-assimilating CO_2 from reactions occurring outside of the pathway, such as glycolysis (Furdui and Ragsdale, 2000; Schuchmann and Müller, 2014). Therefore, will mutants behave differently when grown on sugar and C1 gas substrates? Will gas uptake vary, and if so, in what concentrations and partial pressure ratios?

- Testing the presence of discrete physiological differences in *agrD* mutants and WT

Granulose formation, salt tolerance, heat-resistance and spore formation are potential Agr-dependent phenotypes that have been identified in published experiments on

Clostridium species (Steiner *et al.*, 2011; Cooksley *et al.*, 2012; Darkoh *et al.*, 2015; Phillips, Huhnke and Atiyeh, 2017). Can any of these established QS phenotypes be experimentally applied to in WT *C. autoethanogenum*, and if applicable, be later tested in mutants?

4.2 Results

4.2.1 Growth and fermentation product profiles in AgrD deficient mutants

4.2.1.1 Growth profiling on fructose

Fructose fermentations in serum bottles were performed to observe growth and fermentation product differences in Δ*agrD1*, Δ*agrD2* and Δ*agrD1D2* mutants against the WT. Mutant pre-cultures grown in the anaerobic cabinet using PETC medium were used as inoculum when an optical density of OD_{660} 0.8 – 1.0 was reached. The inoculum was injected into 250 mL serum bottles containing 90 mL PETC with 10 g/L fructose, with bottle headspace comprised of anaerobic tent gas mix. The experiments started typically with an OD_{660} of 0.01 – 0.02.

Growth curves are shown in Figure 4.1 and followed an average time course of 175 hours, starting with exponential growth and transitioning to stationary phase typically after 48 hours with peak optical densities that exhibited no statistically significant differences amongst strains. A linear plot of this growth curve can be found in appendix, section 7.3. Fermentation product sampling took place during key time points that covered the start, middle and end of exponential phase, see 4.2.1.2 for results. Later in the project, a further two independent Δ*agrD1D2* mutants (Δ*agrD1D2_16* and *ΔagrD1D2_34*) were generated and included in the growth curve, Figure 4.1. Overall, there was no significant statistical p-value difference in growth between the strains, as determined by student's t-test. However, on occasion, Δ*agrD1D2* remained at a higher OD in stationary phase; this was possibly due to the changes in pH, as explained further in section 4.2.1.2. During the growth experiment, serum bottles were occasionally sampled and samples were examined under a microscope to observe any morphological differences, with no

irregularities noted. Cell motile behaviour, size and shape remained the same amongst all strains (data not shown).

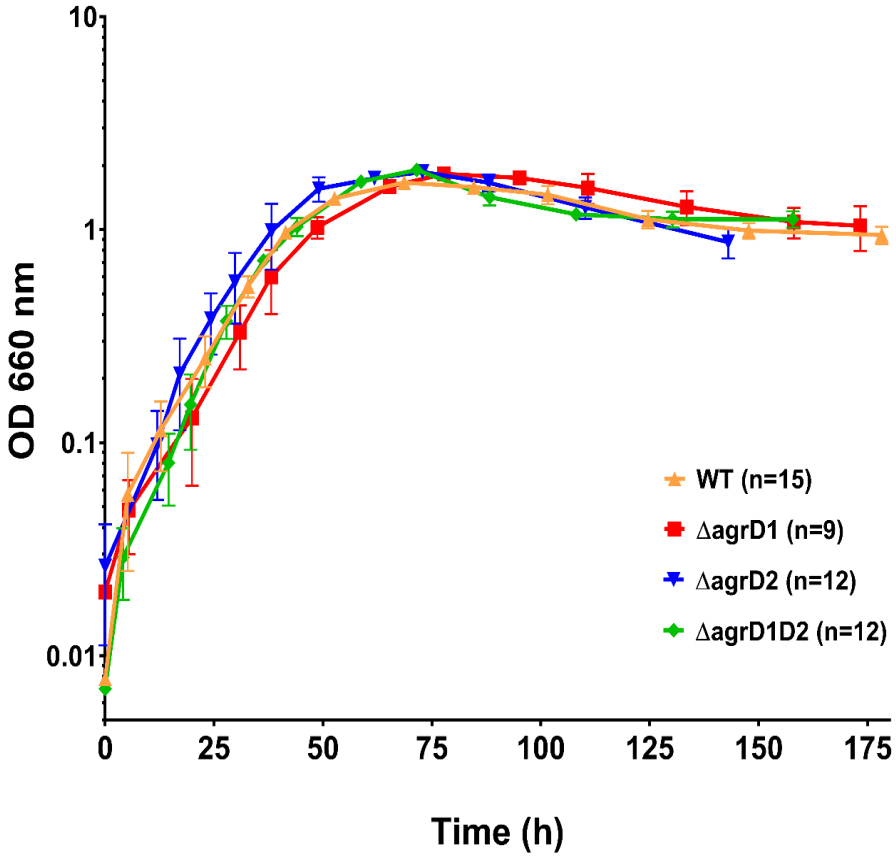


Figure 4.1, Growth of *C. autoethanogenum* WT and *agrD* mutant strains on fructose: Strains were grown in PETC medium containing 10 g/L fructose over an average period of 175 hours. Time points measured OD, and collected samples were saved for later solvent analysis. Error bars indicate standard error of the mean.

4.2.1.2 Fermentation product profiling of AgrD mutants

Assessment of fermentation products took place over the duration of the entire experiment. Gas chromatography was used to detect the main products, ethanol and acetate, Figure 4.2, whereas fructose, lactate and 2,3-butandiol were determined by high pressure liquid chromatography (HPLC), Figure 4.3. The HPLC data also included ethanol and acetate measurements to serve as a means of independently validating GC data. GC results, Figure 4.2 A-D, showed that there was no statistically significant difference between the WT and single-knockout mutants. Acetate was generated earlier and at a higher rate, and reached higher final concentrations of 53.3 ± 4.4 mM, 57.3 ± 7.3 mM and 55.8 ± 3.6 mM, for *ΔagrD1*, *ΔagrD2* and WT respectively. Ethanol, reached relatively low levels with 19.5 ± 3.6 mM, 16.1 ± 3.1 mM and 11.8 ± 0.9, respectively. There was no

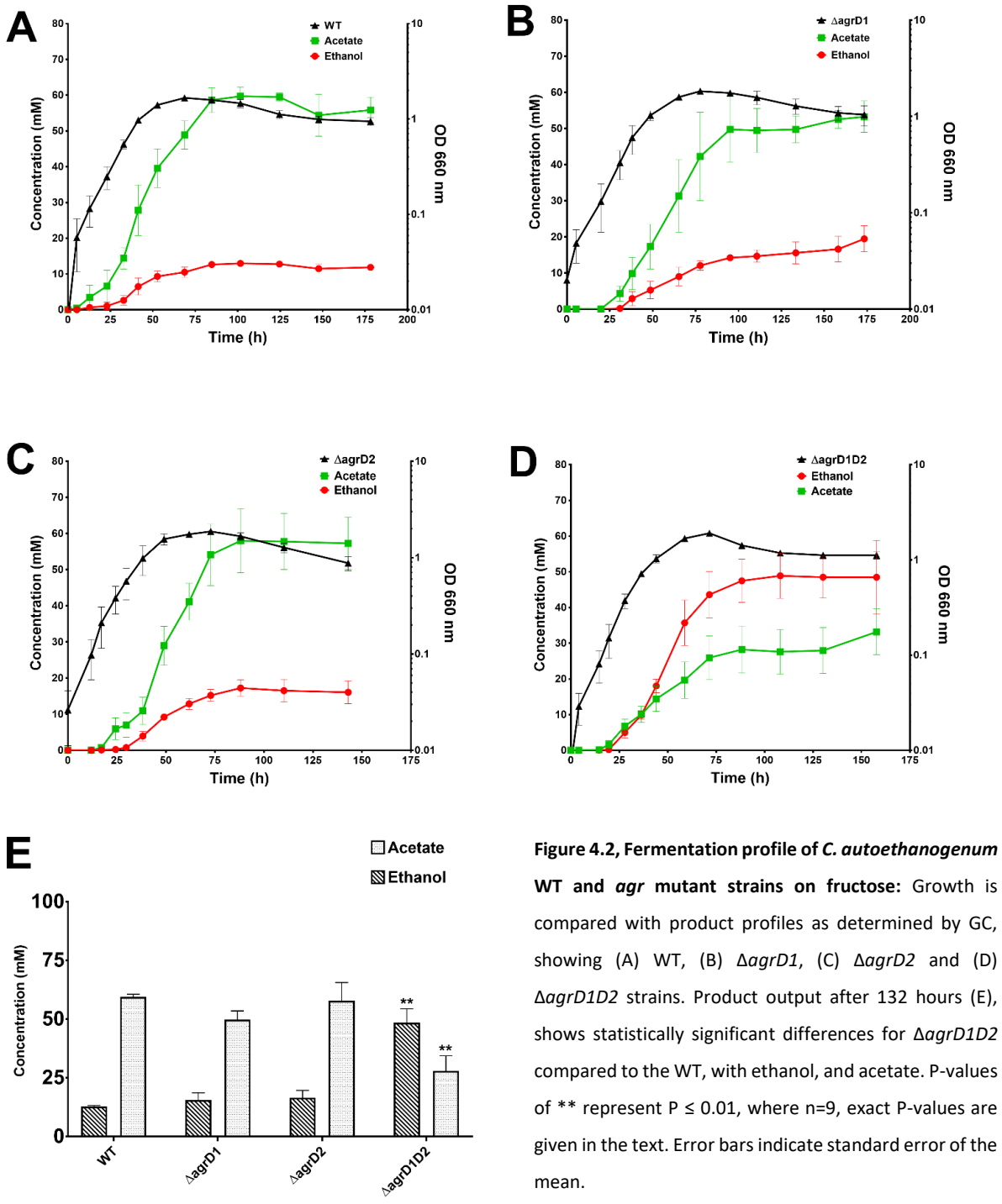


Figure 4.2, Fermentation profile of *C. autoethanogenum* WT and *agr* mutant strains on fructose: Growth is compared with product profiles as determined by GC, showing (A) WT, (B) Δ*agrD1*, (C) Δ*agrD2* and (D) Δ*agrD1D2* strains. Product output after 132 hours (E), shows statistically significant differences for Δ*agrD1D2* compared to the WT, with ethanol, and acetate. P-values of ** represent P ≤ 0.01, where n=9, exact P-values are given in the text. Error bars indicate standard error of the mean.

statistical significance in either product between the WT and single-knockout mutants. However, a striking difference was observed for Δ*agrD1D2* mutants with high statistical significance when compared against the WT, where maximum concentrations of 48.9 ± 10.3 mM (P = 0.0037), and 33.2 ± 6.4 mM, where observed for ethanol and acetate (P = 0.0084), respectively. Thus, for these mutants, ethanol had become the main fermentation product and its production started almost at the same time as that of

acetate. These changes might have accounted for the observed changes in pH in occasional experiments (not shown). Strains Δ*agrD1*, Δ*agrD2* and the WT were observed to have end-point pH values of around 4.4 – 4.6, while for Δ*agrD1D2* occasionally a more basic pH of 4.9 – 5.1 was measured. These pH differences may have allowed to influence growth, with observed higher peak ODs and stationary phases in Δ*agrD1D2* in certain experiments (not shown).

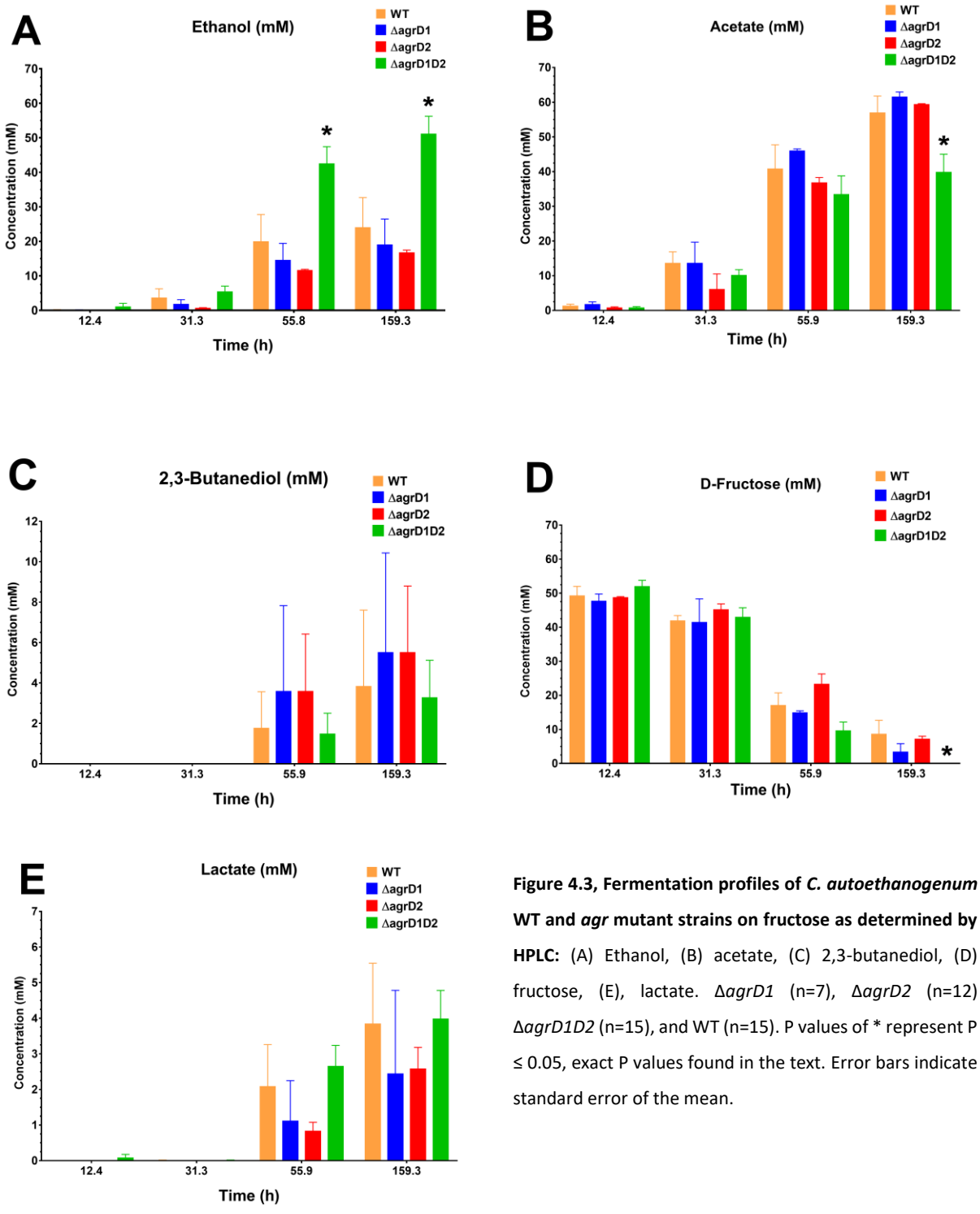


Figure 4.3, Fermentation profiles of *C. autoethanogenum* WT and *agr* mutant strains on fructose as determined by HPLC: (A) Ethanol, (B) acetate, (C) 2,3-butanediol, (D) fructose, (E), lactate. Δ*agrD1* (n=7), Δ*agrD2* (n=12) Δ*agrD1D2* (n=15), and WT (n=15). P values of * represent P ≤ 0.05, exact P values found in the text. Error bars indicate standard error of the mean.

HPLC analysis covered fewer time points but offered a broader product spectrum. Figure 4.3 A and B show that ethanol and acetate values generally agree with the GC data. Key time points that were analysed covered approximate critical stages of the experiment. The first three marked the start (12.4 hours), middle (31.3 hours) and end of exponential phase (55.8 hours), while the final time point (159.3 hours) represented the end of the experiment and final product concentrations. The trends observed for ethanol and acetate production followed that already observed for the GC analysis. For instance, at 55.8 hours, ethanol concentrations for the Δ*agrD1D2* mutants were 42.6 ± 4.8 mM, and thus already significantly differing to the WT's output of 19.9 ± 7.5 mM, (P = 0.03). The Δ*agrD1D2*'s ethanol maximum average output reached 51.2 ± 5 mM (P = 0.018) and acetate reached 39.9 ± 5.1 mM (P = 0.049), which again agreed with the GC data.

However, 2,3-butanediol and lactate production, Figure 4.3 C and E, did not differ significantly between strains. Large error bars represented variation of 2,3-butanediol production in certain experiments, as some failed to yield 2,3-butanediol, but as a general trend, a low mM output was observed at later time points. Fructose consumption, Figure 4.3.D, showed an intriguing difference for Δ*agrD1D2* mutants, with a total absence of fructose after 159.3 hours of fermentation, 0.06 ± 0.02 mM (P = 0.02). This contrasted with Δ*agrD1*, Δ*agrD2* and WT, that all had remaining fructose post fermentation with final concentrations of 3.55 ± 2.3 mM, 7.3 ± 0.7 mM and 8.7 ± 3.9 mM, respectively. This phenomenon was highly reproducible and observed in all growth experiments.

4.2.1.3 Gas headspace analysis

C. autoethanogenum is a gas metabolising organism that utilises H_2 and CO_2, mixotrophically when grown on a sugar carbon source (Abrini, Naveau and Nyns, 1994; Liew *et al.*, 2017). Therefore, CO_2 and H_2 liberated during fructose glycolysis may be re-utilised for the generation of mainly acetate (Tracy *et al.*, 2012). Serum bottles for the experiments described in section 4.2.1.1 and 4.2.1.2 were prepared in an environment containing a standard anaerobic gas mixture that included 10% of H_2 and CO_2 to which the organism is passively exposed to. During fermentation on sugar, additional CO_2 was released as a by-product as described above. During fructose growth experiments, serum bottles were routinely sampled and observed to build up substantial gas pressure in the case of the *ΔagrD1D2* mutants. When measured on this occasion, the pressure was only measured, and *ΔagrD1D2* mutants reached a pressure of 57 ± 3 kPa, whereas pressures remained relatively low for WT and Δ*agrD2* bottles, with 14 and 15 kPa respectively. On

a later occasion, selected bottles of WT, *ΔagrD2* and Δ*agrD1D2* strains were used to precisely measure relative gas pressure ratios and general headspace composition at the end of 175 hours fermentation. Appreciable traces of methane, an observed by-product of the WL-pathway (Fuchs, 1986), was observed for the WT and Δ*agrD2* cultures (not shown) with 8 ± 1 kPa and 14 ± 1 kPa, respectively, but not detected in Δ*agrD1D2* mutant cultures. Conversely, CO_2 was observed at a significantly higher pressure in Δ*agrD1D2* cultures, with 99.02 ± 2.06 kPa, ($P < 0.001$) with respect to the WT, 33.44 ± 0.89 kPa and Δ*agrD2*, 36.29 ± 0.98 kPa, (not shown).

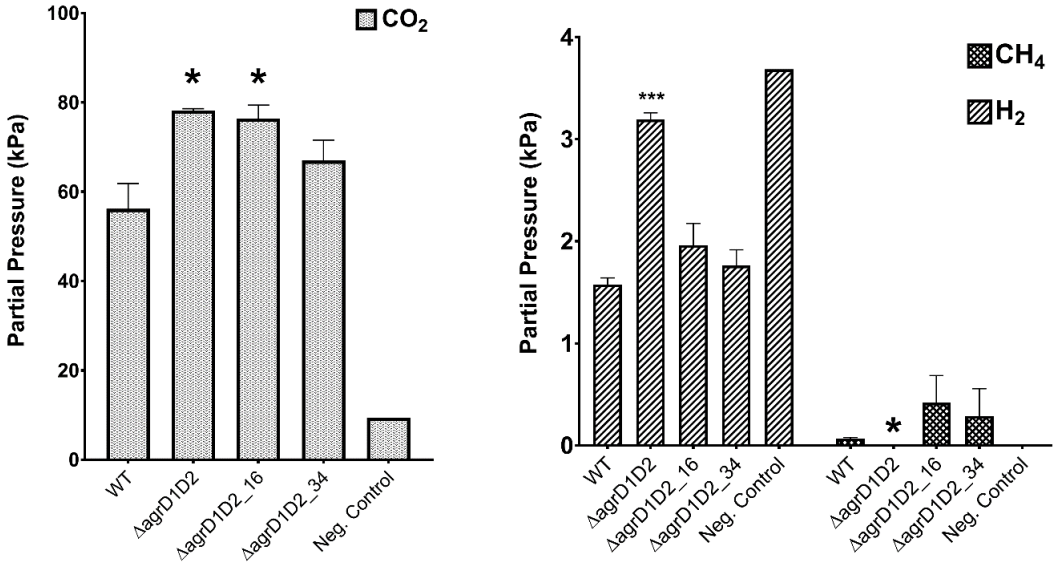


Figure 4.4, Headspace Composition of fructose cultures at 200 hours: (A) CO_2 and (B) methane and hydrogen partial pressures in serum bottles of WT and three independently isolated Δ*agrD1D2* mutants. Δ*agrD1D2* (n=3), Δ*agrD1D2_16* (n=4) and Δ*agrD1D2_34* (n=4), WT (n=4). Negative control represents uninoculated bottles. P values equate to: * ≤0.05, ***≤0.001, exact P values are given in the text. Error bars indicate standard error of the mean.

This experiment was repeated once more, but with the knowledge that there is little effect on single knockout mutants, 3 independently isolated Δ*agrD1D2* mutants (Figure 4.4) and the WT were tested, Figure 4.4. During the experiment, headspace composition in these gas bottles was measured at 200 hours. Once more, a significant increase of CO_2 was observed in the Δ*agrD1D2*, Δ*agrD1D2_16* mutants and less so in Δ*agrD1D2_34*, with 78.17 ± 0.4 kPa ($P = 0.02$), 76.42 ± 2.95 kPa ($P = 0.02$), and 66.99 ± 4.56 kPa, respectively. Hydrogen, starting with a pressure of 3.68 kPa, was consumed in all strains, with Δ*agrD1D2* consuming it at a statistically significant lower rate, remaining at 3.2 kPa ($P = 0.001$). Methane was again not present in Δ*agrD1D2,* but traces were found in

Δ*agrD1D2_16* and Δ*agrD1D2_34* and WT. These results suggest that CO_2 is not readily assimilated in Δ*agrD1D2* mutants when compared to the WT, the methane question remained unanswered as to why it was exclusively not present in one specific Δ*agrD1D2* mutant.

4.2.1.4 Growth on a C1 gas substrate

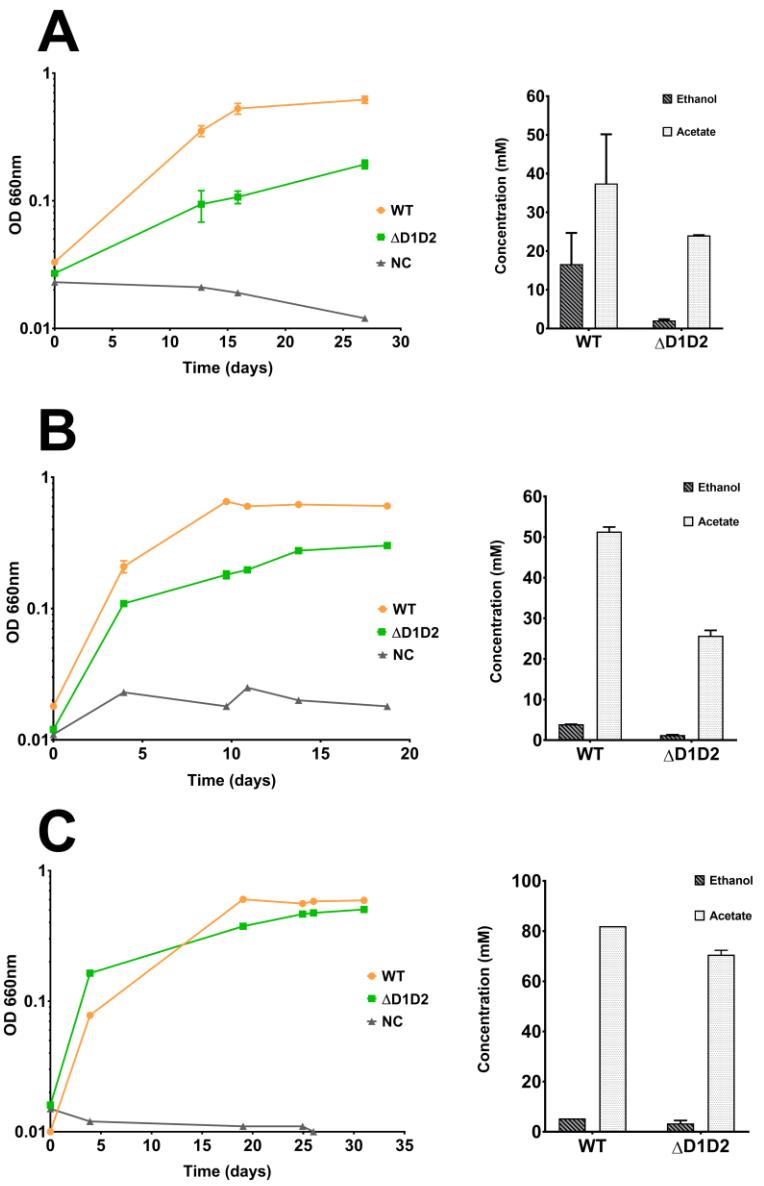


Figure 4.5, Adaptive growth and fermentation product profiles of WT and Δ*agrD1D2* on CO gas: Profiles represent first (A), second (B), and third (C) subcultures of WT (n=2) and *ΔagrD1D2* (n=2) cultures, growing in PETC with 150 kPa of 100 % CO, over a period of 29, 19 and 31 days respectively. Samples analysed using HPLC. NC = Negative control, contained no CO. Error bars indicate standard error of the mean.

Following the discovery of excess CO_2 in the headspace of the Δ*agrD1D2* strain, investigation into whether this double knockout strain can grow on gas was tested. Methods into this test are outlined in section 2.2.3.1. WT and Δ*agrD1D2* were inoculated into serum flasks containing 50 mL of PETC, with 1.5 bar CO. To accommodate for CO metabolism adaptation, both strains were cultured for extended periods of time, after which fermentation endpoints were collected for HPLC analysis, Figure 4.5.

After the initial subculture, the WT showed a 3-fold growth difference to *ΔagrD1D2*, with a peak OD_{660} of 0.62 to 0.19 respectively. The second subculture displayed an improved adaptation of *ΔagrD1D2* with OD_{660} 0.3 to the WT OD_{660} of 0.6. The final subculture showed *ΔagrD1D2* closely matching the WT OD_{660} with 0.5 and 0.59 respectively.

The general fermentation product trend for the WT showed increasing acetate and diminishing ethanol output from the first to the last subculture. The *ΔagrD1D2* strain showed that acetate and ethanol concentrations are both lower than the WT, but as with growth, acetate titres appeared to adapt to the WT level. Ethanol, however remained at a lower concentration throughout, when compared to the WT, Figure 4.5. From these 3 subcultures, the WT adapts to CO metabolism more readily than *ΔagrD1D2*, but *ΔagrD1D2* is able to eventually match a similar growth profile in the final subculture, with fermentative products matching closely.

4.2.2 Complementation of *ΔagrD1D2* mutants

The *ΔagrD1D2* mutant was chromosomally complemented independently with either *agrD1* or *agrD2*. Complementation strategies followed those described in section 3.2.4, whereby successful chromosomal recombination for each gene occurred downstream of the *pyrE* locus. Each complement lacked its native promoter but was driven by the *pyrE* operon promoter instead. Complemented mutants were verified via screening PCR, Sanger sequencing and Illumina sequencing (section 3.2.5 and 3.2.6), and complemented mutant strains were designated *ΔagrD1D2_ D1*^{comp} and *ΔagrD1D2_ D2*^{comp}, respectively.

4.2.2.1 Growth profile of complemented Δ*agrD1D2* mutants

Complemented mutants were inoculated and grown in small-scale fermentations identically as described in 4.2.1.1. Growth followed with no significant differences across strains over a period of 137 hours, with 12-time points at which samples were collected.

However, on this occasion, one of the complemented strains, *ΔagrD1D2_ D2^{comp}*, and the original mutant, *ΔagrD1D2,* grew with an extended lag phase.

4.2.2.2 Fermentation product analysis of complemented Δ*agrD1D2* mutants

Having successfully achieved chromosomal integration of the complementing *agrD1* and *agrD2* genes, it was examined whether this influenced the strains' fermentation product profiles. By following the same workflow as in 4.2.1.1, GC, HPLC and GC Trace techniques were employed to examine samples and bottles for clear signs of complementation.

GC profiles for acetate and ethanol showed clear effects for Δ*agrD1D2* complemented with *agrD1* but not for *agrD2*, Figure 4.6 B - C. The WT acting as a control, displayed acetate and ethanol concentration as previously seen in Figure 4.2.A, with peak concentrations of 69 ± 1.5 mM and 19.2 ± 0.2 mM respectively, Figure 4.6.A. The *ΔagrD1D2_ D1^{comp}* strain displayed a reversion back to a WT-like fermentation profile with final acetate and ethanol concentrations reaching 64.7 ± 2.5 mM and 28.2 ± 1.2 mM respectively Figure 4.6.B. However, *ΔagrD1D2_ D2^{comp}*'s fermentation profile matched closer to *ΔagrD1D2* double mutant and differed considerably from the WT, with 43.9 ± 2.0 mM ($P < 0.0001$) and 32.9 ± 0.1 mM ($P < 0.0001$) for acetate and ethanol respectively, Figure 4.6.E-F. As a comparison, the double mutant produced final concentrations of 39.7 ± 5 mM ($P< 0.001$) and 38.8 ± 9.3 mM ($P < 0.05$), for acetate and ethanol, respectively (Fig 4.6.E-F)

HPLC analysis revealed a similar trend to the GC data, Figure 4.7. Strain *ΔagrD1D2_ D1^{comp}* was similar to the WT in terms of final acetate and ethanol concentrations, and fructose was no longer consumed, although consumption remained somewhat different, with slightly higher fructose concentrations remaining in the WT, 8.4 ± 2.3 mM, compared 4.3 ± 1.4 mM for *ΔagrD1D2_ D1^{comp}*. Concentrations of fructose in *ΔagrD1D2_ D2^{comp}* (1.1 mM) again matched *ΔagrD1D2*'s near complete utilization of fructose, and were statistically significant when compared to the WT. Moreover, 2,3-butanediol production was not observed in this experiment. In previous characterization experiments, 2,3-butanediol was regularly observed at low concentrations with seldom occasions of negligible, to no amounts detected. Lactate was observed in all strains, with the WT, *ΔagrD1D2_ D1^{comp}* and *ΔagrD1D2* all reaching similar maximum averages of 5.3 ± 0.24 mM, 4.9 ± 0.45 mM and 6.16 ± 0.82 mM respectively. Strain *ΔagrD1D2_ D1^{comp}* ended

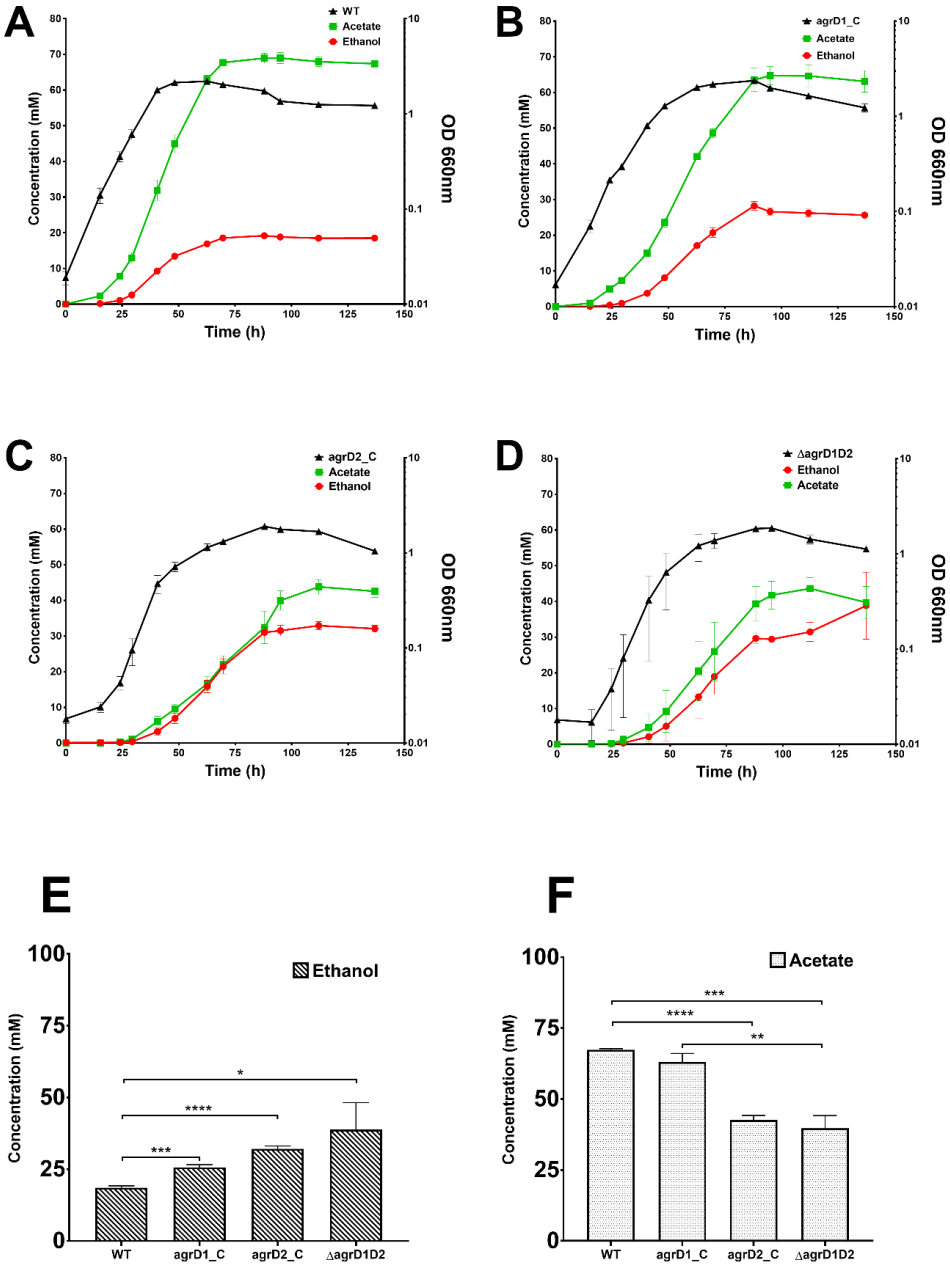


Figure 4.6, Fermentation profiles of chromosomally complemented *C. autoethanogenum agr* mutants on fructose as determined by GC: Growth in PETC 10 g/L fructose, showing (A) WT (n=4), (B) *agrD1_C* (n=4), (C) *agrD2*_C (n=4) and (D) Δ*agrD1D2* (n=3). Ethanol and acetate output after 136 hours (E and F). P values equate to: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001, exact P values are given in the text. Error bars indicate standard error of the mean.

fermentation with the highest significant average of 8.8 ± 0.6 mM. This contrasts with previous fermentations, where lactate output was low, especially in Δ*agrD1* and Δ*agrD2.*

From these findings, *ΔagrD1D2_D1*^{comp} showed a striking similarity to WT, suggesting reasonably successful complementation. *ΔagrD1D2_ D2*^{comp} had a less pronounced complementation effect and resembled a Δ*agrD1D2* phenotype across most product profiles.

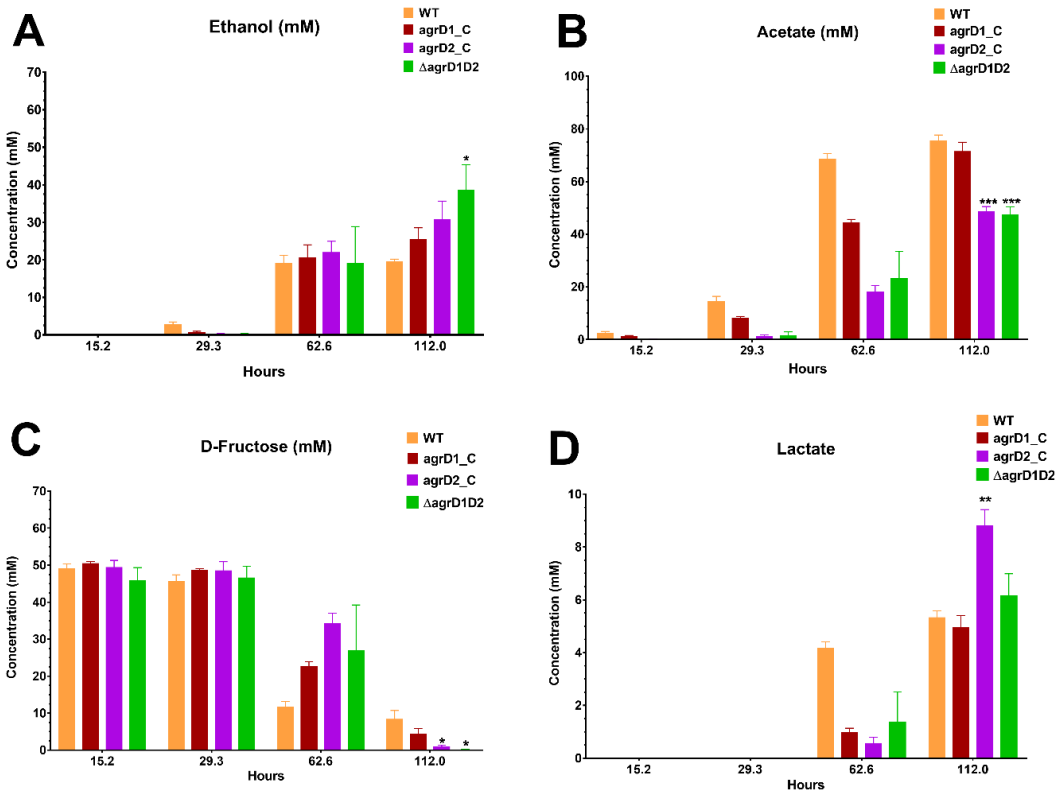


Figure 4.7, Fermentation profiles of chromosomally complemented *C. autoethanogenum agr* mutants on fructose as determined by HPLC: (A) Ethanol and (B) acetate, (C) D-fructose and (D) lactate profiles following 136-hour growth in PETC with 10g/L fructose. WT (n=4), agrD1_C (n=4), agrD2_C (n=4), Δ*agrD1D2* (n=3). Error bars indicate standard error of the mean. P values equate to: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001, exact P values are given in the text.

4.2.2.3 Gas analysis profile

Gas analysis took place at the end of the fermentation experiment described in 4.2.2.2 and showed similar CO_2 and H_2 gas trends between the WT and *ΔagrD1D2* as observed in Figure 4.4. Cultures of the complemented *ΔagrD1D2_ D1*^{comp} strain showed a headspace gas composition which, overall, resembled that of the WT, although CO_2 and H_2 partial pressures remained slightly higher. For *ΔagrD1D2_ D2*^{comp} cultures, CO_2 partial pressures were like those observed for the double mutants, while H_2 partial pressures were similar

to the wildtype, Figure 4.8. Interestingly, along with CO_2 consumption, methane partial pressures were restored to WT levels for *ΔagrD1D2_ D1^{comp}* but not *ΔagrD1D2_ D2^{comp}* cultures.

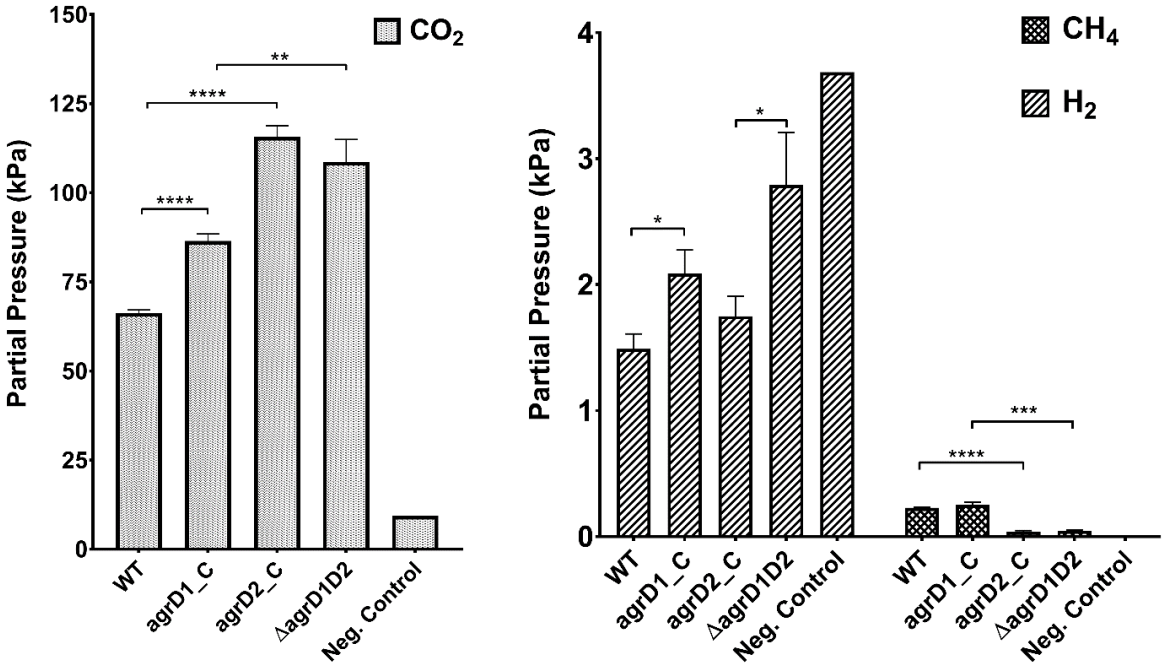


Figure 4.8, Headspace composition of chromosomally complemented *agrD* mutant cultures at 136 hours: Complemented *agrD1^{comp}* and *agrD2^{comp}* strains, WT and Δ*agrD1D2* double mutant were grown in PETC with 10g/L fructose. Negative control represents uninoculated bottles. Error bars indicate the SEM. P values equate to: * ≤0.05, **≤0.01, ***≤0.001 and **** ≤ 0.0001. WT (n=4), *agrD1^{comp}* (n=4), *agrD2^{comp}* (n=4), Δ*agrD1D2* (n=3).

4.2.2.4 Chemical complementation

Another route to complementing *ΔagrD1D2* mutants was undertaken through the addition of synthetic peptides, based on the design of the predicted mature ring structures of AgrD1 and AgrD2. As outlined in 2.6.3, 4 separate iterations of each AgrD peptide were designed, each possessing their respective ring structures with tails of lengths varying between 0 to 3 residues long, as described in table 2.7. From the outsourced synthesis process, only the AgrD1 peptide set was received, with low purity ranging from 50 – 70 %, and yields of 2.1 – 5 mg. Typical quality and yields would consist of 90 – 100 %, and up to 10 mg respectively.

Growth experiments took place in 15 mL PETC containing 10 g/L fructose, with 20 mM concentrations of each respective peptide. No growth differences were observed between the WT, the untreated *ΔagrD1D2* control and each of the 4 treated *ΔagrD1D2*'s cultures with each peptide variant. Fermentation products were measured after 32 hours of growth. A generally similar trend was observed with small statistically significant differences in final product titres across all *ΔagrD1D2* mutants treated with peptides, when compared to untreated *ΔagrD1D2*, Figure 4.9. However, none of the added peptides were capable of reversing the effect of the *ΔagrD1D2* double mutant's phenotype.

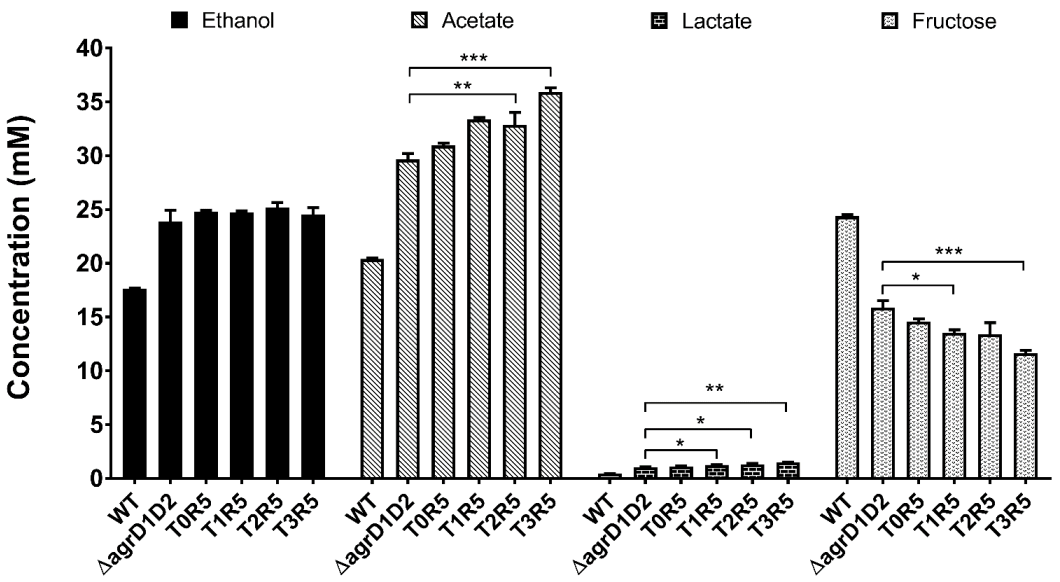


Figure 4.9, Fermentation productions of chemically complemented *ΔagrD1D2* mutant cultures at 32 hours: WT, untreated Δ*agrD1D2* and Δ*agrD1D2* double mutants treated each with 20 mM of AgrD1 peptides; T0R5, T1R5, T2R5, and T3R5, were grown in PETC with 10g/L fructose. Error bars indicate standard error of the mean. P values equate to: * ≤0.05, **≤0.01 and ***≤0.001. WT (n=3), Δ*agrD1D2* (n=3), Δ*agrD1D2* and peptides (T0R5 – T3R5) (n=3)

4.2.3 Physiological characterization

At the start of the project, it was unknown as to what phenotypes were to be expected from generated *agrD* mutants. From the literature, *C. botulinum*, *C. sporogenes* and *C. acetobutylicum* had *agrB* and *agrD* genes targeted for disruption with clear phenotypes showing that QS is linked to heat resistant endospore formation, storage polymers (granulose) generation and neurotoxin production, (Cooksley *et al.*, 2012; Steiner *et al.*, 2012). As toxins are not produced by *C. autoethanogenum*, only the former two

phenotypes could be tested for their presence, as they may also be linked with *agr* QS systems in *C. autoethanogenum*.

4.2.3.1 Testing for granulose accumulation

In Steiner *et al.*'s study (2012), disruption of the *agr genes* via ClosTron resulted in loss of granulose accumulation. Granulose, a starch like storage compound, has been documented in *C. acetobutylicum* (Reysenbach, Ravenschroft and Long, 1986) and the respective genes, notably *glgA,* identified (Nölling *et al.*, 2001). A simple iodine staining experiment can be conducted to determine its presence. No documented *glgA* homologs could be detected in *C. autoethanogenum*'s genome (Humphreys *et al.*, 2015), which genetically suggested that *C. autoethanogenum* does not possess capabilities to accumulate glycogen or starch-like storage compounds under excess carbon conditions. To test empirically whether this was true, *C. autoethanogenum* was grown for 4 days on

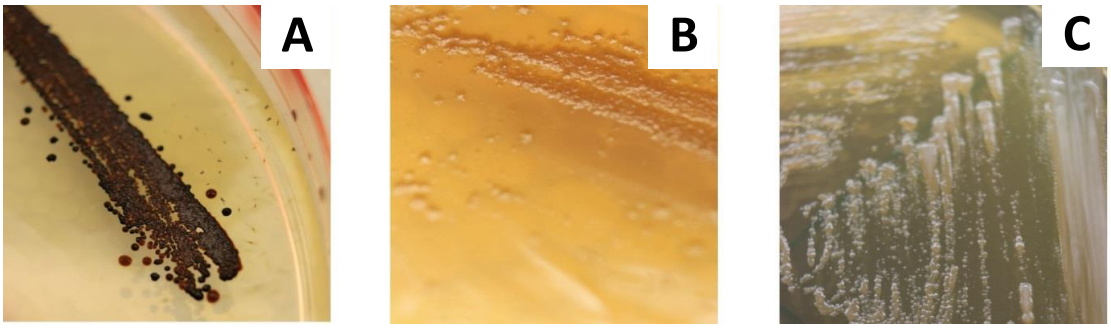


Figure 4.10, Demonstration for the lack of granulose accumulation in *C. autoethanogenum*, as suggested by iodine staining: Strains were grown on agar plates and subjected to iodine staining. (A) Positive control, *C. acetobutylicum* streaked onto CBM medium with 10 g/L glucose. (B) *C. autoethanogenum* grown on YTF with 10 g/L fructose. (C) *C. autoethanogenum* on YTF with 10 g/L fructose, untreated.

YTF agar plates with 10 g/L fructose until visible colonies were noticed. Staining for starch-like polymers was carried out as described in section 2.6.5.

After plates were exposed to iodine vapours, they ventilated for a few minutes to remove excess iodine. The *C. autoethanogenum* plates, initially stained with excess iodine, reverted to a colourless state, while the control *C. acetobutylicum* remained a distinct brown/black colour, Figure 4.10. This experiment conclusively showed that starch-like polymers are not accumulating in *C. autoethanogenum*, at least under the employed conditions, and therefore can be ruled out as a potential phenotype to look for in QS studies.

4.2.3.2 Heat resistant endospore Formation

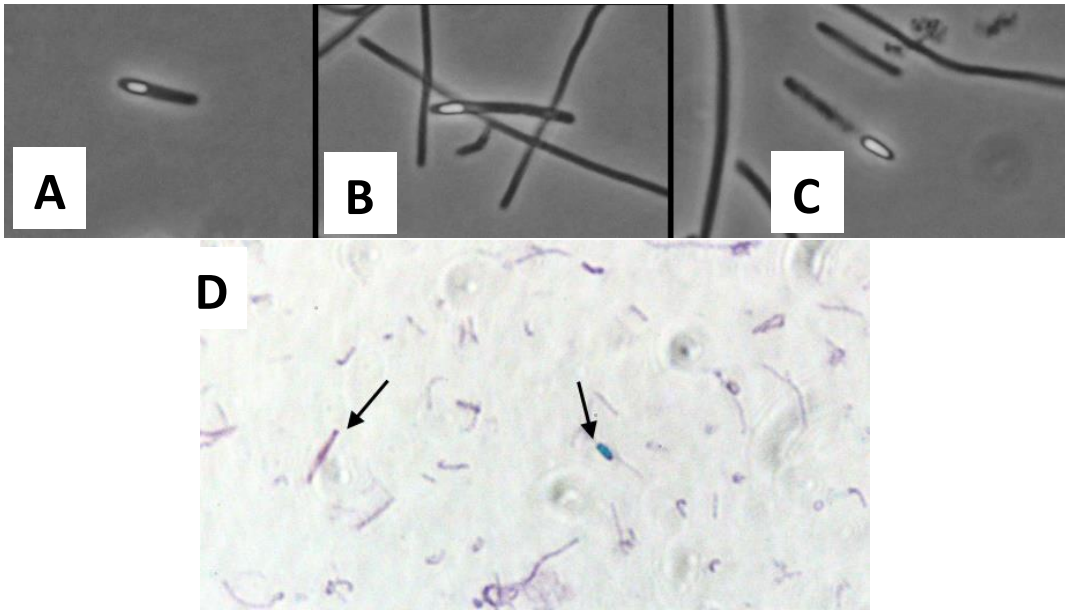


Figure 4.11, Spore-like particles present in CO-grown continuous cultures of *C. autoethanogenum*: Microscope pictures showing distinct endospore-like structures within cells (A, B and C). (D) shows spore staining according to Schaeffer & Fulton (Schaeffer 1990) with arrows pointing to vegetative cells (pink) and free spores (blue).

Where studied, sporulation in some *Clostridia* spp. has been shown to be closely linked with QS activity (Cooksley *et al.*, 2010; Li *et al.*, 2011; Jabbari *et al.*, 2013; Edwards, Tamayo and McBride, 2016; Kotte *et al.*, 2017). *C. autoethanogenum* when isolated and characterized by Abrini et al. (1994), was described as a spore-former, but with endospore formation occurring infrequently. Similarly, the closely related *C. ljungdahlii* has been described as a tentative spore-former, with little to no forespores produced (Tanner, Miller and Yang, 1993).

Detection of heat resistant endospores follows a simple procedure of heat-shocking serially-diluted cell suspensions at 80 °C for 10-20 min and plating them on appropriate nutrient agar plates, section 2.6.4. Resultant colonies are counted and correlated with heat-resistant spore formation. In experiments with *C. autoethanogenum,* this procedure was adopted on numerous occasions in the current study with no colonies forming in any instance. Various alternative methods were implemented to induce spore formation, which included oxygen stress, excess and scarcity of a carbon source and methods to promote spore germination borrowed from other *Clostridia* spp. The motivation behind these approaches are fully outlined in the discussion. Other experiments included a more

holistic approach whereby rabbit faeces were collected and transformed into a sterile supplement added to medium.

The only instance when endospore-like structures were microscopically observed, occurred during a continuous fermentation in a colleague's 2 L bioreactor, which operated under minimal medium conditions with CO gas as a sole carbon source and supplemented with Na_2SO_4 as the sole sulphur source. Microscope imaging and Schaeffer & Fulton's spore-staining (Schaeffer and Fulton, 1990) was performed on that occasion and seemed to suggest that endospores were indeed present, Figure 4.11. However, no viable cells could be recovered following heat shock treatment and plating of the bioreactor samples. Therefore, it is not clear whether the structures observed in the bioreactor represented viable spores or, indeed, whether they were formed by the organism rather than a contaminant.

After all avenues of enquiry were thoroughly exhausted, it was concluded that sporulation is either an exceptionally rare occurrence in *C. autoethanogenum*, or that it required yet unidentified conditions or signals. As a result, plans to use sporulation as a phenotypic marker when screening potential QS mutants had to be put aside.

4.2.3.3 Salt tolerance and biofilm formation

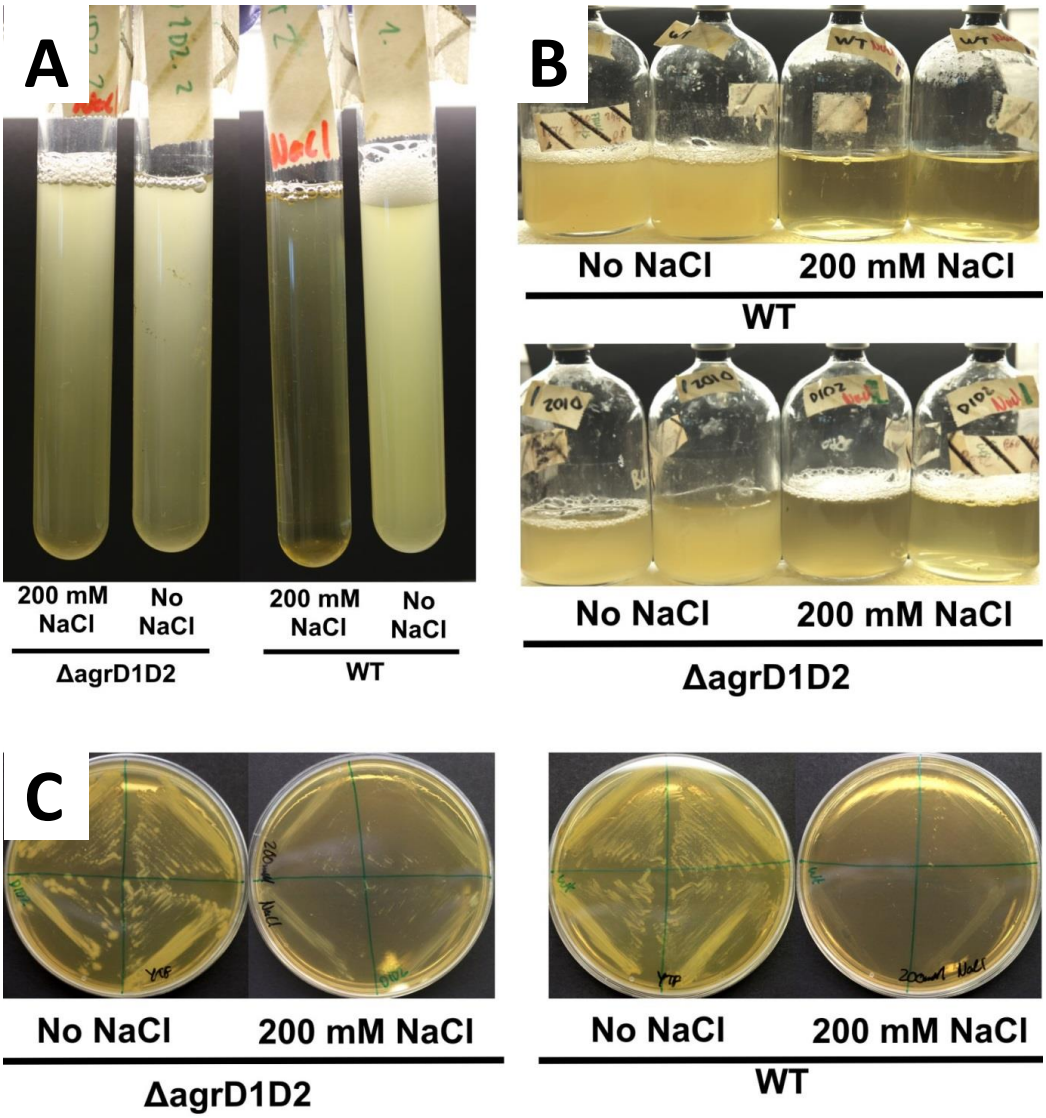


Figure 4.12, Salt tolerance of *C. autoethanogenum* WT and *ΔagrD1D2* mutants: (A) Hungate tubes and (B) 90 mL serum bottle containing WT failing and Δ*agrD1D2* mutants grown in PETC medium the presence and absence of 200 mM NaCl (C) YTF agar plates for WT and Δa*grD1D2* mutant with and without 200 mM NaCl.

Biofilm formation is observed for many bacterial species in their natural habitats and in the laboratory, and has often been linked with QS activity (Singh *et al.*, 2000; Miller and Bassler, 2001; Vidal, Shak and Canizalez-Roman, 2015). However, it is not yet fully understood in many *Clostridia*, besides *C. difficile* and *C. perfringens* whose activity are not representative for the whole genus (Pantaléon *et al.*, 2014). A study done by Philips et. al. (2017) examined biofilm formation in *C. ljungdahlii,* by inducing it through stress, which included adding a variety of salts, which resulted in up-regulation of salt-stress response genes (Philips *et al.*, 2017). Moreover, as *C. ljungdahlii* is closely related to *C.*

autoethanogenum, it was reasonable to assume that applying similar conditions to *C. autoethanogenum* would lead to similar results. Observing biofilm formation under these conditions would then allow for testing whether the phenomenon was dependent on *agr* QS.

In a preliminary experiment, WT *C. autoethanogenum* (n=2) and Δ*agrD1D2* (n=2) were grown in 15 mL PETC with 10 g/L fructose, with and without 200 mM NaCl, with as starting growth OD_{660} of 0.02. After 5 days, cultures were examined, and planktonic growth was observed in Δ*agrD1D2*, in both NaCl-free and NaCl containing Hungate tubes. The WT had only shown growth in the absence of NaCl. Only after a further 3 days a small aggregation of WT cells could be observed at the bottom of the tube, in the presence of NaCl, Figure 4.12 A.

This experiment was repeated once more in 90 mL PETC with 10 g/L fructose, with and without 200 mM NaCl. Similar results were obtained showing that Δ*agrD1D2* did not aggregate at the bottom of the serum bottle, but produced a planktonic culture, although this was less in the presence of NaCl. WT bottles without NaCl showed typical turbid cultures with high densities, but NaCl addition resulted in clear medium with a very light aggregation of cells, Figure 4.12 B. Microscopy revealed atypical short contorted cells. The experiment was repeated on solid YTF agar, with and without 200 mM NaCl. Results showed substantially less WT colonies when streaked on 200 mM NaCl compared to Δ*agrD1D2*, although both strains exhibited less growth under these conditions, Figure 4.12.C.

4.2.3.4 Cell fitness

As reported in section 4.2.1.2 it was observed that Δ*agrD1D2* mutants had completely consumed the provided fructose carbon source by the end of fermentation (159.3 hours), and that more CO_2 gas accumulated in the headspace in comparison to WT. When subcultured multiple times on a C1 gas substrate, as described in 4.2.1.4, *ΔagrD1D2* was eventually able to adapt and match the growth to that of WT, with a similar fermentation productive profile. This lead to the question; does a functional QS system give the WT an advantage in switching from a diminishing carbon source (sugar), to more abundant forms (gaseous), well in advance to ensure survival?

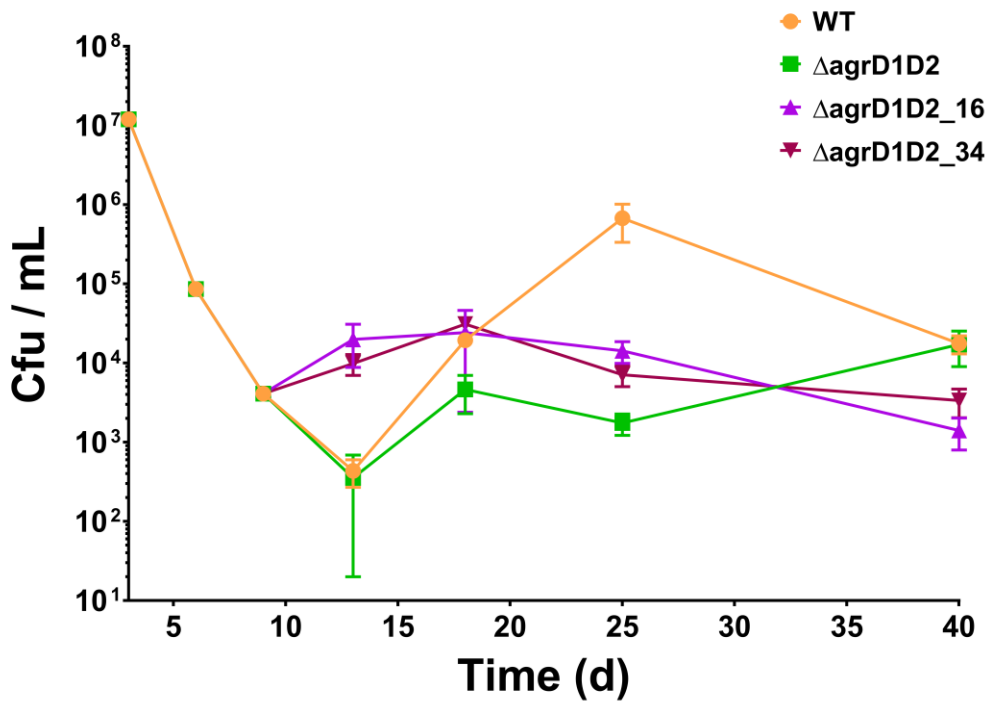


Figure 4.13 Long-term viability of cells present in C. autoethanogenum WT and ΔagrD1D2 cultures grown on fructose: Cfu measurement took place 3 days after inoculation. WT (orange circles), independent ΔagrD1D2 mutants (purple, burgundy, green). Error bars indicate standard error of the mean, n=3.

To test this hypothesis, colony forming unit (Cfu) assays were performed during a growth experiment for WT *C. autoethanogenum* and 3 separate *ΔagrD1D2* mutants to determine the number of viable cells present over time, which would be indicative of a given strain's fitness, (Figure 4.13). During this experiment, cell survival assays were performed as outlined in 2.6.2. During the growth experiment, after reaching stationary phase at 74 hours (3 days) strains reached their maximum OD_{660} values and began to decline. At this time point gas pressure was measured for the WT and Δ*agrD1D2*, *ΔagrD1D2_16* and *ΔagrD1D2_34* mutant strains with 25 ± 1 kPa, 35 ± 1 kPa, 27 ± 2 kPa, and 25 ± 1 kPa respectively.

After 13 days, it was observed that WT steadily increased in terms of cfu count, while mutants remained at a relatively constant level. The WT was observed to reach a maximum of 6.67×10^6 cfu/mL at day 25, while Δ*agrD1D2*, *ΔagrD1D2_16* and Δ*agrD1D2_34* reached 1.35×10^3 cfu/mL, 1.43×10^4 cfu/mL and 7.12×10^3 cfu/mL, respectively. Gas pressures measured on the last day showed, 2 kPa, 12 ± 1 kPa, 11 ± 2 kPa and 9 ± 1 kPa for WT, Δ*agrD1D2*, Δ*agrD1D2_16* and Δ*agrD1D2_34* respectively.

This data suggests that WT *C. autoethanogenum* may have gradually switched to CO_2 + H_2 metabolism in agreement with the data in Figure 4.4 showing that Δ*agrD1D2* mutants retained relatively high CO_2 pressures after 200 hours of growth.

4.3 Discussion

The primary aim of this chapter was to characterise *C. autoethanogenum* mutant's defective in Agr QS in order to establish the role of cell-to-cell communication in metabolic, sporulation and physiological activities. This was carried out through the characterisation of basic growth and fermentation product profiles of strains Δ*agrD1*, Δ*agrD2* and Δ*agrD1D2*, with respect to the WT.

While there was no obvious difference in growth, the initial results of fermentation product profiling showed a clear difference for the Δa*grD1D2* mutant, with a remarkably high ethanol output in comparison to the WT and both single-knockout mutants. Both Δ*agrD1* and Δ*agrD2* mutants produced acetate and ethanol at levels comparable to the WT. From Figure 4.2, the Δ*agrD1D2* strain had a near complete reversal in product output, with an average nearing 50 mM ethanol and 33 mM acetate, whereas the WT and single knockout mutants produced around 15 mM and 56 mM, respectively. A further discrepancy was the complete utilisation of fructose by Δ*agrD1D2* mutants, contrary to WT and both single knockout mutants. Finally, analysis of gas pressure and composition in bottle headspace suggested that a substantially higher amount of CO_2 accumulated exclusively in Δ*agrD1D2* headspace, compared to the WT. From these experimental observations, could there be a unifying explanation?

As an acetogen, hypothetically *C. autoethanogenum* can use the EMP pathway to generate pyruvate from 1 hexose, which is further metabolised into 2 acetate, 2 CO_2, and 8 reducing equivalents [H], this is also coupled to the synthesis of 4 ATP (eq. 1) (Wood, 1952a; Ivey and Ljungdahl, 1986; Müller, 2003).

$$C_6H_{12}O_6 + 2H_2O + 4\,ADP + 4\,P_i \rightarrow 2CH_3COOH + 2CO_2 + 8[H] + 4ATP \qquad \text{(eq. 1)}$$

Provided the liberated CO_2 is then completely re-assimilated, using the reducing equivalents gained during glucose oxidation, a third molecule of acetate can be formed via the WL pathway., (eq. 2) (Wood, 1952a; Ivey and Ljungdahl, 1986; Müller, 2003; Drake, Kusel and Matthies, 2006).

$$2CO_2 + 8[H] + nADP + nP_i \rightarrow CH_3COOH + 2H_2O + nATP \qquad \text{(eq. 2)}$$

As acetyl-CoA largely contributes to acetate formation in most species, this results in what used to be called a homoacetogenic fermentation, defined simply as the generation of 3 acetate from a single hexose molecule in an ideal fermentation (Diekert and Wohlfarth, 1994). With this concept in mind, the following hypothesis can be put forward: WT *C. autoethanogenum* metabolises fructose and re-fixes a portion of the generated CO_2 to produce mostly acetate alongside small amounts of ethanol. The phenotypic changes seen for Δ*agrD1D2* suggest that this mutant shows a greatly reduced degree of carbon fixation via the WL pathway. This has been shown with higher CO_2 headspace pressures, Figure 4.4, but also the increased production of ethanol, and notably reduced formation of acetate, Figure 4.2. Thus, CO_2 is no longer available as an electron acceptor and hence the generated reducing equivalents in the form of H^+ reduce acetyl-CoA instead, leading to an increased ethanol formation in place of acetate. The observed differences in the trace amounts of methane present in the headspace of fructose-grown cultures further support this explanation. Methane is liberated in trace amounts by acetogenic bacteria, as it is generated in a side reaction catalysed by key WL pathway enzymes, CO dehydrogenase and acetyl-CoA synthase (Fuchs, 1986; Schauder *et al.*, 1986). Therefore, the absence of methane in the Δ*agrD1D2* agrees with the presented hypothesis. However, as can be seen from Figure 4.4, the situation is not as clear as previously thought, as two-other independently obtained *agrD* double mutants, Δ*agrD1D2*_16 and Δ*agrD1D2*_34, did not show reduced methane levels under the same conditions, despite methane production being restored in the Δ*agrD1D2* mutant through genetic complementation, Figures, 4.4 and 4.8. There is currently no obvious explanation for this, but as it has been noted in Table 3.3.A in section 3.2.6, both *ΔagrD1D2_16* and *ΔagrD1D2_34* carry different, although significantly less, SNVs, which may contribute to this disparity. It should be noted that strains undergoing intensive genetic manipulation, often accumulate mutations in the form of single nucleotide variants and small structural variants (Miller, 1996; Foster *et al.*, 2015; Humphreys *et al.*, 2015), section 3.2.6. However, it was shown that chromosomal complementation of Δ*agrD1D2* with the *agrD1* gene restored not only acetate and ethanol production to WT levels, but also the ability to produce methane, Figure 4.8.

The question then arises as to how and why an Δ*agrD1D2* mutant should notably affect a central metabolic pathway in this organism? Quorum sensing-dependent regulation of metabolic genes has been observed in different bacteria and allows the cells to prepare for, and handle the stresses and limitations associated with high population density

(Williams *et al.*, 2007; Dandekar, Chugani and Greenberg, 2012; Cornforth and Foster, 2013). With respect to *C. autoethanogenum*, it could be argued that the sensing of population density through peptide signals might allow cells to upregulate the necessary enzymes and systems required for growth on C1 gases before more favourable carbon sources such as fructose are exhausted by a quickly growing population. Results described in sections 4.2.1.4 and 4.2.3.4 hint at this hypothesis, initially showing *ΔagrD1D2*'s slow adaptation to CO compared to WT, Figure 4.5. In later cell fitness experiments, Figure 4.13, the WT showed a steady recovery in growth over time, in comparison to Δ*agrD1D2* mutants, which suggested the result of an adaptive switch to CO_2 metabolism.

However, the question for double AgrD signals persists; why is there such a pronounced phenotype in the double-knock out, and why are two different AgrD peptide signals required? Detailed in Chapter 3, the residues of AgrD1 and AgrD2 share an overall identity of 70.21%, but with distinct conserved ring structures. One possible explanation for the requirement of these similar, but perhaps not identically acting signal peptides, could be to do with a combinatorial mode of function, which is defined as two signals theoretically functioning simultaneously in cultures that are dense and/or have high mass transfer, effectively improving their social and physical environmental resolution (Cornforth *et al.*, 2014). In the context of *C. autoethanogenum*, this concept can be supported by two independent sets of transcriptomic data, which have both observed AgrD1 to be significantly more expressed than that of AgrD2, (Marcellin *et al.*, 2016; Henstra, 2018). Having both signals operate at two specific transcript levels, *C. autoethanogenum* would be theoretically capable of recognising vital AgrD signals on a much wider population density and mass transfer spectrum. Disabling only either one of these signals merely reduces this spectrum, and the Agr system remains largely functional. However, this scenario would require converging signalling pathways for AgrD1 and AgrD2, for instance by interacting with their proposed cognate histidine kinases, AgrC1 and AgrC2, would in turn act on the same phosphorelay protein. There are examples for such as systems in Gram-negative bacteria, for instance the well-studied *N*-acyl homoserine lactone-dependent QS system of *Vibrio harveyi* (Waters and Bassler, 2006; Papenfort and Bassler, 2016). This explanation would be compatible with what was observed for the single knock-out mutants, with no significant difference in either and phenotypes principally matching the WT.

However, complement mutants, *ΔagrD1D2_D1^{comp}* and *ΔagrD1D2_D2^{comp}*, were expected to return to a single-knockout phenotype, yet only *ΔagrD1D2_D1^{comp}* appeared to do so, and *ΔagrD1D2_D2^{comp}* showed similarities resembling a *ΔagrD1D2* phenotype. A possible explanation to the latter would be the nature of the targeted recombination at the *pyrE* locus, as both *agrD* complements made use of the *pyrE* promoter in place of their own as, explained in 3.2.5. A possible theory could be that this promoter is attenuated differently leading to different transcription of either *agrD* gene, which can include the level at which they are natively transcribed. The mechanism of the *agr* operon promoter remains elusive in Clostridia, but in *S. aureus*, it has been made clear that AgrA, along with other transcription factors, activates the highly auto-catalytic *agr* promoter, P2, (Ji, Beavis and Novick, 1995; Novick, 2003). Of course, this mechanism might work entirely differently in *C. autoethanogenum*, and other Clostridia, but it gives an idea to the potential complexity of *agr* expression and mediation. Furthermore, production of the AgrD proteins is perhaps not only dependent on transcription rates and transcript stability, but also translation efficiency, which together may not have resulted in sufficient AgrD2 production by the complemented double mutant.

To address physiological changes in mutants, several experiments were undertaken in a systematic manner by drawing parallels to techniques performed in *C. acetobutylicum* and *C. ljungdahlii*. Initially sporulation and granulose formation was examined, as studies in *C. acetobutylicum* have demonstrated these two activities have been associated with its *agr* system (Steiner *et al.*, 2011, 2012). However, it was observed that *C. autoethanogenum* does not display either of the two phenotypes under the conditions tested.

It was uncertain if *C. autoethanogenum* could synthesise granulose-like compounds as it does not possess any putative *glgA* homologs in its genome (Humphreys *et al.*, 2015). Iodine-staining experiments further supported this, and it was concluded that granulose cannot be used as a marker in *agr* mutant screening. With regards to sporulation, *C. autoethanogenum*'s initial characterization by Abrini *et al.* (1994), reported its spore-forming capabilities, however no spores have ever been documented in our lab besides on occasion as described in Figure 4.11, and in specific bioreactor experiments. In Abrini *et. al.*'s paper, spore-forming rods were reportedly observed after repeated incubation in a $CO/N_2/CO_2$ atmosphere, with cells containing spore-like phase-bright, round-shaped particles in terminal or sub-terminal position. This conclusion was perhaps premature as there were no follow-up heat-shock tests done to confirm if these so-called spores were

viable. During this project, similar but more oval cellular structures were observed in terminal position in instances were gas was used as a carbon source. These hypothetical spores were heat-shocked to induce germination and subsequent growth as performed with other *Clostridia* species (Jones *et al.*, 1982; Ramachandriya, DeLorme and Wilkins, 2010; Pyne *et al.*, 2013), and tests showed no positive outcome. Thus, without passing a heat-shock test, it could not be asserted that the observed structure was indeed a typical endospore.

In *Clostridia*, sporulation is a well-documented event (Talukdar *et al.*, 2015), but it has not been fully understood in acetogens such as *C. autoethanogenum* and *C. ljungdahlii*. Interestingly, *C. ljungdahlii*, when characterized by Tanner *et al.* was described as having highly infrequent "tentative forespores-like structures" (Tanner, Miller and Yang, 1993), but no work has been done to clarify these findings. Tanner et. al's (1993) observations agree with the findings made in this study which support the view that both species do not readily sporulate. This leaves the question as to why *C. autoethanogenum* does not sporulate in the same manner as other members of its genus. Although not fully understood, typically, environmental stress factors such exposure to oxygen, or low pH from fermentation, contribute to triggering sporulation in *Clostridium* spp. (Long, S., Jones, D.T. and Woods, 1984; Sauer *et al.*, 1995; Al-Hinai, Jones and Papoutsakis, 2015). In this genus, the sporulation cascade typically begins with phosphorylation of the master regulator, Spo0A and is mediated through factors σ^F and σ^E (Dürre and Hollergschwandner, 2004; Paredes, Alsaker and Papoutsakis, 2005; Papoutsakis, 2008; Al-Hinai, Jones and Papoutsakis, 2015), both present in *C. autoethanogenum* (Humphreys *et al.*, 2015). Either a key component in this sporulation mechanism must be missing, inactive, or require a very stringent set of conditions for it to be triggered. Thus, regulation of sporulation via Agr QS in *C. autoethanogenum* could not be tested in this study.

Lastly, it was recently described in Philips *et al.* (2017), that sodium chloride stress induced biofilm formation in *C. ljungdahlii*, and a similar approach was implemented in *C. autoethanogenum*. This investigation questioned the possible link that *agr* QS might play in biofilm formation, as this is documented in other Clostridia and various Gram-positive bacteria (Kong, Vuong and Otto, 2006; Li *et al.*, 2011; Dapa *et al.*, 2013; Pantaléon *et al.*, 2014). Results showed a clear difference between *ΔagrD1D2* strains and the WT, with the WT samples failing to show planktonic and turbid cultures when compared to the mutant and tended to aggregate at the bottom of the vessel, Figure 4.12. Experiments

performed on solid YTF agar containing NaCl also showed discrete differences. Whether this outcome was the consequence of a disabled *agr* system is still unclear and further investigation is required. Transcriptomic data from salt-stress experiments on *C. ljungdahlii* did not provide indication that its *agr* homologs are up/down regulated (Phillips, Huhnke and Atiyeh, 2017).

Interestingly, the results found in *ΔagrD1D2* mutants could hint at a possible salt stress-response, as a lack of an AgrD signal prevents the inhibition of growth. This concept has been observed in *C. thermocellum*, where pentose-induced stressed created a significant upregulation of AgrD peptide homologs that were expressed independently from expected high cell densities (Verbeke *et al.*, 2017). The addition of synthetic AgrD also greatly inhibited growth, even when not in the presence of pentose sugars. Could a similar function be attributed to the salt-stressed WT *C. autoethanogenum*? It is still early to conclude, and further experimentation is necessary to explore this concept.

In summary, this chapter elucidated the phenotypic changes in AgrD-deficient *C. autoethanogenum* mutants, and lead to the discovery of a substantially different fermentation product profile as well behaviours towards C1 gas substrates when compared to the WT. The next chapter will focus on the plausible link these phenotypes may have with the ancient WL-pathway.

Chapter 5

Relationship of Agr quorum sensing and the Wood-Ljungdahl pathway

5.1 Introduction

Following results in Chapter 4, a plausible connection between the Agr system and *C. autoethanogenum*'s WL-pathway was established. This followed Δ*agrD1D2* mutants exhibiting a pronounced phenotypic difference in fermentative products, gas headspace composition and complete fructose utilisation during fermentation, relative to the WT. These differences formed the hypothesis that mutant Δ*agrD1D2* shows a significantly reduced degree of carbon fixation through the WL-pathway, and instead diverts its reducing equivalents into reducing acetyl-CoA, thereby increasing ethanol output in place of acetate.

This evidence suggests that *C. autoethanogenum*'s WL and/or EMP pathways (Drake, Kusel and Matthies, 2006; Drake, Gößner and Daniel, 2008) are possibly influenced by the Agr system, but the regulatory point in these pathways where Agr plays this possible role is not clear, and therefore the biochemistry of both pathways must be scrutinised. An overview of the WL pathway, introduced in section 1.4.1, describes the carbonyl and methyl CO_2/CO-assimilating branches, both converge to form acetyl-CoA via the critical bi-functional CODH/ACS complex (Doukov, 2002; Ragsdale and Pierce, 2008). This complex has been noted for condensing CO, a methyl group and CoA into acetyl-CoA, allowing the formation of 1 ATP per acetyl-CoA through subsequent conversion to acetate (Ragsdale, 2004). In heterotrophic conditions, *C. autoethanogenum* chiefly uses the EMP pathway, which was described to produce 8 reducing equivalents per molecule hexose that take part in the WL-pathway's further acetate production (Wood, 1952a; Ivey and Ljungdahl, 1986; Müller, 2003; Ragsdale and Pierce, 2008). Both pathways share a common acetyl-CoA junction; the EMP produces pyruvate, which can be interconverted via pyruvate:ferredoxin oxidoreductase (PFOR) into acetyl-CoA and vice versa (Ragsdale, 2003), described in eq. 5.1:

$$2pyruvate + CoA + Fd_{ox} \leftrightarrow 2CO_2 + Acetyl - CoA + Fd_{red} \quad \text{(eq. 5.1)}$$

From this reaction, a concomitant release of 2 CO_2 occurs as 2 molecules of pyruvate are formed during breakdown of fructose. These can be re-assimilated by *C. autoethanogenum* via the WL-pathway using the 8 reducing equivalents (H^+) produced from the EMP pathway (Drake, Kusel and Matthies, 2006). This CO_2 generation which cannot be re-assimilated to the same extent in the *ΔagrD1D2* strain likely explains the higher CO_2 partial pressure on fructose-grown cultures as seen in Chapter 4.

In summary, during heterotrophic growth, the WL-pathway takes advantage of the EMP pathway's electron donors, CO_2, and reducing equivalents in the form of H_2 (Müller, 2003), but during autotrophic conditions, CO_2 and CO take place of the carbon source, with reducing equivalents being provided by molecular hydrogen and CO (Müller, 2003; Schuchmann and Müller, 2014). Although dissimilar, both pathways play an integral role in energy metabolism in *C. autoethanogenum*. Moreover, as suggested by transcriptomic data, it comes as no surprise that CODH/ACS complexes and hydrogenases are highly upregulated in autotrophic conditions (Marcellin *et al.*, 2016), but exact regulation of these WL-pathway associated enzymes has not yet been clarified in literature.

The observed increase in CO_2 partial pressure and a shift towards higher ethanol production in *agrD* mutants offered the tempting hypothesis that the ACS/CODH complex is likely involved, and a recent study has shown *acsA* inactivated mutants alluding to a set of similar phenotypes (Fungmin Liew *et al.*, 2016). In Liew *et al.*'s study, ClosTron-disrupted *acsA* mutants showed a 2-fold increased ethanol output, very low acetate production, approximately 4-fold CO_2 headspace pressure, and near absent autotrophic growth on CO and CO_2 + H_2, relative to the WT. These finding promoted the compelling hypothesis that the Agr system affects the WL-pathway either directly or indirectly, via the CODH/ACS complex.

5.1.1 Aims of this study

After determining phenotypic differences in *ΔagrD1D2* mutants, it was thought that the Agr system may influence the WL-pathway, and to investigate this claim, proteomic analysis and enzymatic assays focusing on key WL-pathway enzymes were carried out. The aims for this chapter can be outlined as follows:

- To compare WT and *ΔagrD1D2* strains through proteomic analysis to quantify and distinguish differences in each strain's respective protein expression.

When both strains are subjected to identical growth conditions and grown to similar densities, what conspicuous differences can be detected in the proteome? How do these differences relate to the observed *ΔagrD1D2* phenotypes described in chapter 4, and are there differences not related to these phenotypes?

- Verifying proteomics data with independent enzyme assays aimed at key WL-pathway enzymes and associated enzymes.

From proteomics data, which enzymes are most up- and down-regulated, and do independent assays verify this data, if so, is there a clear difference between the WT and *ΔagrD1D2* strain?

5.2 Results

5.2.1 Proteomic analysis

Strain culturing procedures prior to analysis were as described in section 2.8.4. Cultures of WT *C. autoethanogenum* and the *ΔagrD1D2* mutant were grown in triplicate and harvested at OD_{660} 1.0. Fermentative profiles from each strain's culture matched trends described in Chapter 4, with high ethanol, lower acetate, and complete fructose utilisation that indicated an expected growth pattern for each respective strain, (fermentative data not shown).

Cell pellets were washed as described and analysed at Sheffield University by Dr Mahendra Raut and Prof. Phillip Wright. Extracted peptides labelled using iTRAQ 4-plex were analysed with MaxQuant, and a total of 282 proteins were detected, with high significant statistical differences ($P < 0.0001$) in relation to peptide expressions levels between the WT and *ΔagrD1D2* strains. A full list of these proteins, grouped according to cluster, can be found in the appendix, section 7.4. Of these 282 proteins, 39 proteins relevant to *C. autoethanogenum's* metabolism were separated into categories which included, the WL-pathway, pyruvate metabolism, glycolysis, energy conservation and acetate, ethanol, 2,3-butanediol formation, Table 5.1. As an aside, when interpreting the relative changes in protein abundance it was important to consider that iTRAQ has been

reported to often underestimate the occurring changes (Ow *et al.*, 2009), this discrepancy is explained further in the discussion.

In Table 5.1 the most down regulated proteins detected for the *ΔagrD1D2* mutant were key WL-pathway enzymes in *ΔagrD1D2*, with CODH/ACS sub-unit complexes (CLAU_1578 and CLAU_1579) showing a -2.94, -2.26-fold difference respectively. Other notably downregulated WL-pathway proteins included further components of the CODH/ACS complex (CLAU_2924 and CLAU_1566), and downregulation of formate metabolising enzymes (CLAU_2713 and CLAU_1567) and the largest group, hydrogenases involved in oxidative reactions outside the WL-pathway during autotrophic growth (CLAU_2717, CLAU_2713, CLAU_1535, CLAU_2722 and CLAU_2719). Interestingly, enzymes involved in ethanol fermentation were split into 2 separate groups on opposite sides of the spectrum, some being significantly up- and others down-regulated. Moreover, recent transcriptomic studies in *C. autoethanogenum* have revealed that individual copies of alcohol and acetaldehyde dehydrogenase genes are expressed differently, dependent on autotrophic or heterotrophic conditions (Marcellin *et al.*, 2016).

Table 5.1, Fold Changes in relative protein abundance for WT and *ΔagrD1D2* mutant: representing enzymes in the WL-pathway (black), pyruvate metabolism (purple), glycolysis (yellow), energy conservation (green, grey) and acetate (blue), ethanol (red), 2,3-butanediol (orange) formation. Fold changes are linear, with positive and negative changes indicating upregulation and down-regulation in the mutant, respectively. Table continues next page.

UniProt ID	Locus Tag	Protein	Fold Change
U5RTV6	CLAU_1766	Alcohol dehydrogenase (Adh)	2.59
U5RXF4	CLAU_3204	Acetaldehyde dehydrogenase (acetylating)	2.51
U5RWW6	CLAU_1783	Acetaldehyde dehydrogenase (acetylating)	2.48
U5RXF4	CLAU_1772	Acetaldehyde dehydrogenase (acetylating)	1.96
U5RTQ9	CLAU_1848	Phosphotransferase system, enzyme I (PtsI)	1.72
U5RPH4	CLAU_0388	Acetolactate synthase (AlsS)	1.69
U5RXJ6	CLAU_3144	Rnf complex (ABCDGE type, C subunit)	1.65
U5RSS2	CLAU_0370	2,3-butanediol dehydrogenase (Bdh)	1.57
U5RXA5	CLAU_2947	Pyruvate:ferredoxin oxidoreductase (PFOR)	1.51
U5RXZ2	CLAU_3146	Rnf complex (ABCDGE type, G subunit)	1.50
U5RXK0	CLAU_3149	Rnf complex (ABCDGE type, B subunit)	1.38
U5RZJ2	CLAU_3861	Alcohol dehydrogenase	1.33
U5RQB1	CLAU_0532	Alcohol dehydrogenase zinc-binding domain protein	1.30
U5RVL5	CLAU_2378	Pyruvate kinase (Pyk)	1.26
U5RT85	CLAU_1553	Pyruvate carboxylase (Pyc)	1.24
U5RYX1	CLAU_2289	ATP synthase subunit alpha (AtpA)	1.15
U5RWQ9	CLAU_1714	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	1.14

U5RTQ5	CLAU_1712	Triose-phosphate isomerase (Tpi)	1.11
U5RT18	CLAU_1572	Methylene-THF reductase (MetF)	1.11
U5RTD3	CLAU_1710	Enolase phosphopyruvate hydratase (Eno)	1.09
U5RX94	CLAU_1713	Phosphoglycerate kinase (Pgk)	1.09
U5RV03	CLAU_2291	ATP synthase subunit beta (atpD)	1.07
U5RWC6	CLAU_1567	Formate-THF ligase (Fhs)	-1.08
U5RXR0	CLAU_3274	Phosphate acetyltransferase (Pta)	-1.09
U5RZ09	CLAU_2322	Fructose-1,6-bisphosphate aldolase (Aldo)	-1.14
U5RP02	CLAU_0089	Aldehyde ferredoxin oxidoreductase (AOR)	-1.24
S5Z7X0	CLAU_2719	4Fe-4S ferredoxin (hytD)	-1.27
U5RWA6	CLAU_2721	Hydrogenase, Fe-only	-1.29
U5RW72	CLAU_2722	4Fe-4S ferredoxin (HydN)	-1.30
U5RW58	CLAU_1535	Hydrogenase, Fe-only	-1.36
U5RWP5	CLAU_2713	Formate dehydrogenase, alpha subunit (MoeA)	-1.38
U5RWA4	CLAU_1566	CODH/ACS complex, AcsB subunit	-1.40
S5ZJ59	CLAU_2717	FeFe-hydrogenase (HytC)	-1.46
U5S0R8	CLAU_2924	Carbon monoxide dehydrogenase (CODH)	-1.72
U5RP39	CLAU_0099	Aldehyde ferredoxin oxidoreductase (AOR)	-1.73
U5RP39	CLAU_0099	Aldehyde ferredoxin oxidoreductase (AOR)	-2.01
U5S2V9	CLAU_3656	Bifunctional acetaldehyde-CoA/alcohol dehydrogenase (AdhE2)	-2.03
U5RTK2	CLAU_1794	Alcohol dehydrogenase (Adh)	-2.13
U5RTE2	CLAU_1579	Carbon-monoxide dehydrogenase, Acceptor (AcsA)	-2.26
U5S2V9	CLAU_3655	Bifunctional acetaldehyde-CoA/alcohol dehydrogenase (AdhE1)	-2.33
U5RSY1	CLAU_1578	Carbon-monoxide dehydrogenase, Acceptor (AcsA)	-2.96

Further expected changes consisted of a down-regulation of phosphate acetyltransferase (Pta) (CLAU_3274) in the *ΔagrD1D2* mutant, that is involved in acetate fermentation, but interestingly acetate kinase did not appear to be significantly different between either strain.

Hydrogenase subunit CLAU_1535 has been studied to be part of the second-highest expressed hydrogenase in *C. autoethanogenum* (Mock *et al.*, 2015) but is not part of the other downregulated hydrogenase subunits (CLAU_2717, CLAU_2719, CLAU_2721, CLAU_2722), which are found in a greater 7-gene cluster that represents NADP-specific electron bifurcating hydrogenase complexes that work in tandem with formate dehydrogenase, (MoeA) (S. Wang *et al.*, 2013).

Enzymes outside of the WL-pathway showed a marginally higher upregulation in the mutant, which included EMP pathway enzyme clusters, (CLAU_1848, CLAU_1714, CLAU_1712, CLAU_1710, CLAU_1713) and pyruvate metabolism enzymes (CLAU_2947, CLAU_2378, CLAU_1553), both might be attributed to Δ*agrD1D2*'s high fructose utilisation phenotype. Electron-transporting Rnf complexes (CLAU_3144, CLAU_3146, CLAU_3149), also experienced a higher upregulation, along with two ATP synthase

subunits, alpha and beta, (CLAU_2289 and CLAU_2291) belonging to the F_0F_1 ATP-synthase complex. Both the Rnf and ATPase complex are coupled, Figure 1.1, the former reduces NAD+ to NADH from ferredoxin generated from CODH, hydrogenase and sugar metabolism, which in turn is concomitant with the translocation of protons (H^+) outside of the cell, the ATPase uses this ion gradient to drive ATP synthesis (Drake, Gößner and Daniel, 2008; Schuchmann and Müller, 2014). Intriguingly, methylene-THF reductase (MetF) (CLAU_1572), is the only upregulated WL-pathway enzyme which drives its methylene-H_4F reduction to methyl-H_4F presumably through NADH reduction, as assumed in *C. ljungdahlii* (Schuchmann and Müller, 2014).

Of note, were non-WL-pathway related enzymes with a high gene upregulation trend, which included several bacterial microcompartment (BMC) proteins, 1,2-propanediol and ethanolamine-utilising enzymes, and uncharacterized proteins, Table 5.2. This was a surprising finding as 1,2-propanediol and ethanolamine traces were noted not to be present in the medium by HPLC analysis, as described in chapter 4. This absence may be due to this observed upregulation of BMC genes, whose purpose is to sequester and metabolise toxic compounds and by-products from the rest of the cell (Cannon *et al.*, 2001).

The first BMC gene cluster, labelled blue, Table 5.2, shows several BMC proteins upregulated (CLAU_1769, CLAU_1772, CLAU_1773, CLAU_1775, CLAU_1785), along with 1,2-propanediol-utilising enzymes (PduL) (CLAU_1771), and ethanolamine-utilising enzymes, (EutQ) (CLAU_1774, CLAU_1779). The gene cluster encoding these proteins is also home to previously mentioned upregulated alcohol and aldehyde dehydrogenases, (CLAU_1766, CLAU_1772, CLAU_1783), Table 5.1, which also can perhaps explain the BMCs concomitant expression, as ethanol utilisation is noted to coincide with BMCs in solventogenic and acetogenic clostridia (Heldt *et al.*, 2009; Schuchmann *et al.*, 2015). Uncharacterized proteins 1, 2 and 3 (CLAU_1786, CLAU_1780 and CLAU_1788) were noted for their high upregulation, and proximity of their encoding genes to BMC clusters. The first and third of these proteins relate to PadR-type transcriptional regulators, involved in phenolic-acid stress (Thi *et al.*, 2011; Isom *et al.*, 2016), the second protein shares high similarity with *C. ljungdahlii's* ethanolamine utilization cobalamin adenosyl transferase (EutT), involved in ethanolamine metabolism (Garsin, 2010).

The second cluster labelled orange in Table 5.2, represents another BMC associated protein (CLAU_3206), together with respective enzymes for 1,2-propanediol

(CLAU_3202) and ethanolamine (CLAU_3205) utilisation. Within this cluster, a second copy of an aldehyde-dehydrogenase (CLAU_3204) is found, Table 5.1. Unrelated to these clusters, glycerol dehydrogenase (GlyDH) (CLAU_0711) is also upregulated and is known to function in 1,2-propanediol metabolism (Altaras and Cameron, 1999).

Table 5.2, Clusters of upregulated BMC components and associated proteins in *ΔagrD1D2*: Showing the first cluster (light blue), and second cluster (orange) of 1,2-propanediol and ethanolamine-utilising enzymes.

UniProt ID	Locus Tag	Protein	Fold Change
U5RTX4	CLAU_1786	Uncharacterized protein 1 (PadR, hypothetical)	3.93
U5RTI0	CLAU_1769	BMC protein	3.77
U5RTW1	CLAU_1771	Phosphate propanoyltransferase (PduL)	3.36
U5RWV7	CLAU_1773	BMC protein	3.22
U5RTF0	CLAU_1772	BMC protein	2.93
U5RTI4	CLAU_1774	Ethanolamine utilization (EutQ family)	2.63
U5RTF0	CLAU_1770	BMC protein	2.59
U5RTJ3	CLAU_3206	BMC protein	2.56
U5RXL2	CLAU_3202	Phosphate propanoyltransferase (PduL)	2.53
U5RTF7	CLAU_1780	Uncharacterized protein 2 (ethanolamine utilization)	2.46
U5S1K9	CLAU_3205	Ethanolamine utilization (EutQ family)	2.42
U5RWX1	CLAU_1788	Uncharacterized protein 3 (PadR, hypothetical)	2.41
U5RXP0	CLAU_1779	Ethanolamine utilization protein (EutJ family)	2.25
U5RTF4	CLAU_1775	BMC protein	2.18
U5RQX8	CLAU_0711	Glycerol dehydrogenase (GlyDH)	2.16
U5RTG1	CLAU_1785	BMC protein	2.13

5.2.1 Enzyme Assays

Proteome differences were established between *ΔagrD1D2* and the WT, particularly showing the differences in WL-pathway and BMC enzymes. This work required independent experimental validation through selecting key WL-pathway enzymes that showed pronounced up and down-regulation. Studies focusing on WL-pathway enzyme assays in *C. autoethanogenum* have provided the experimental methodology (S. Wang *et*

al., 2013; Mock *et al.*, 2015), whereby WT and Δ*agrD1D2* mutant cultures were grown in PETC medium containing fructose until an OD_{660} of 0.45 was reached. Cells were harvested and cell lysates (crude extracts) were prepared by chemical lysis as described in 2.8.2. Crude extracts typically matched between WT and *ΔagrD1D2* samples and had an average concentration of 10 mg/mL. Extracts were then used immediately for the different enzyme assays. Table 5.1 showed that alcohol dehydrogenases, hydrogenases and carbon monoxide dehydrogenases, exhibited marked regulation differences, and were therefore picked to perform this validation study.

Assays as shown in Figure 5.1, tested activity was performed with increasing amounts of crude extract to ensure enough was added to observe activity and also to ensure that the assay was linear over a range of enzyme activities. Assays involving hydrogenase and CODH used the property of methyl viologen (MV) to act as an electron acceptor due to its favourable redox potential, (-444 mV) (Erbes DL, 1978; Ragsdale *et al.*, 1983), and can be described in equations, 5.2 and 5.3:

$$H_2 + 2MV_{oxi}^{2+} \xrightarrow{Hydrogenase} 2H^+ + 2MV_{red} \qquad \text{(eq. 5.2)}$$

and

$$CO + H_2O + 2MV_{oxi}^{2+} \xrightarrow{CODH} CO_2 + 2MV_{red} + 2H^+ \qquad \text{(eq. 5.3)}$$

Reduced MV product formation creates a distinct purple/blue colour change which absorbs at 578 nm, Figure 5.1, A-D. Hydrogenases, Figure 5.1, A-B, are shown to be highly expressed and active in the WT, where rapid product formation corresponded proportionally to amounts of crude extract added. In *ΔagrD1D2*, hydrogenase activity was not as pronounced, and similar crude extract protein concentrations prompted a lower activity than the WT. A similar trend was observed in CODH activity, where again the WT showed a higher product turnover rate than in *ΔagrD1D2*, with similar crude extract protein concentrations used, Figure 5.1, C-D.

In the case of measuring the combined activity of the alcohol dehydrogenases, NADH was used here as an oxidative coenzyme in the reaction of acetaldehyde to ethanol formation (Walker, 1992; S. Wang *et al.*, 2013), described in equation 5.4:

$$CH_3CHO + NADH + H^+ \xrightarrow{adh} CH_3CH_2OH + NAD^+ \qquad \text{(eq. 5.4)}$$

Measuring NADH oxidation is done at an absorbance of 340 nm, which decreases with respect to time, shown in Figure 5.1, E-F. Mutant *ΔagrD1D2* showed a small, but noticeable difference in activity, with steeper linear portions when compared to the WT between 0 and 2 minutes, indicating higher NADH consumption relating to higher activity.

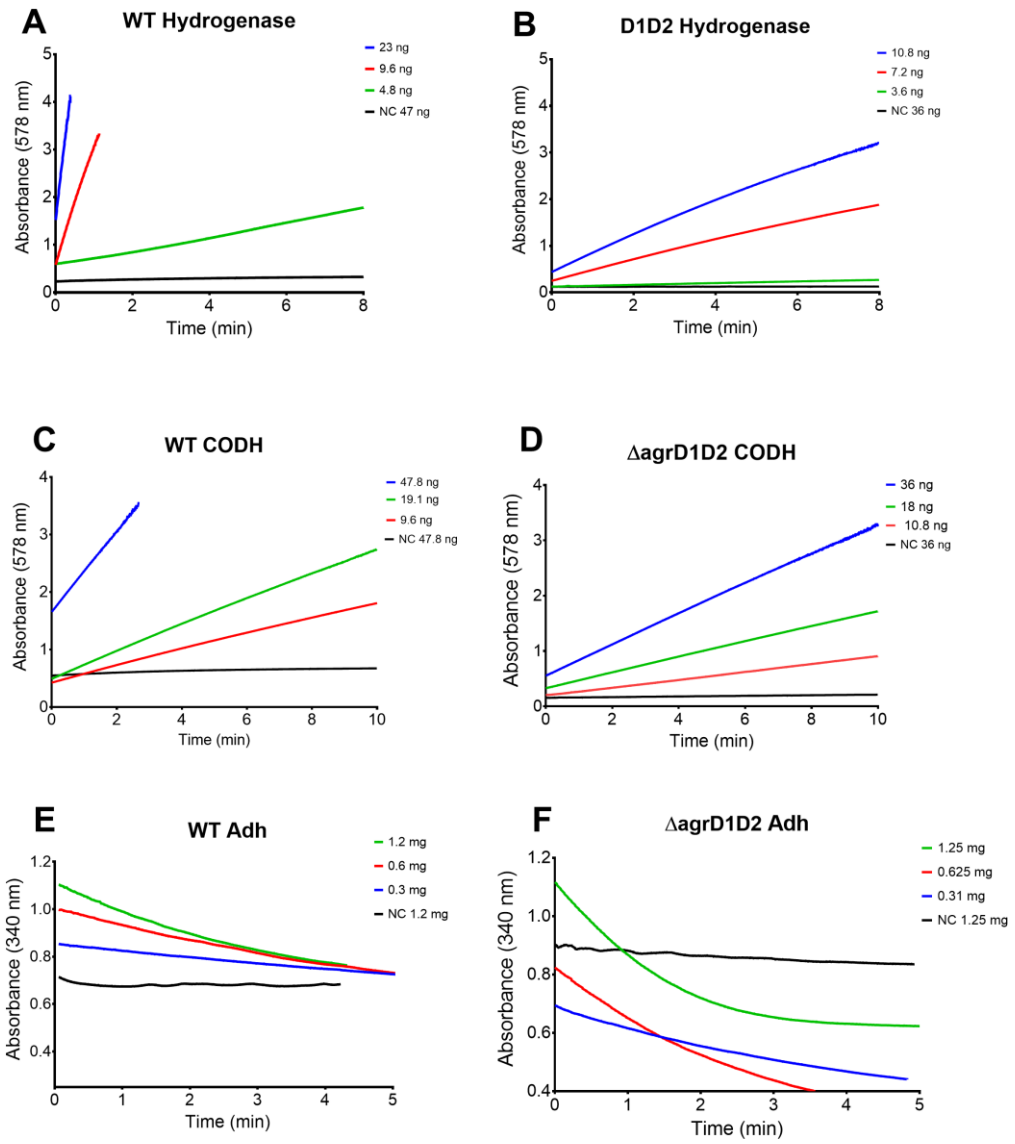


Figure 5.1. Enzyme assays with linear product formation with respect to added cell extract protein amounts over time: Activities of WT and *ΔagrD1D2* showing, hydrogenase, (A – B), carbon monoxide dehydrogenase (CODH), (C – D), and alcohol dehydrogenase, (E – F). Crude cell protein amounts are found in figure legends, NC refers to negative control consisting of enzyme assay reaction mixtures without respective substrates, (H_2, CO and acetaldehyde).

Specific enzyme activities for all three enzymes was calculated using extinction coefficients (ε) of NADH and methyl viologen, 6.2 $mM^{-1}cm^{-1}$ and 9.8 $mM^{-1}cm^{-1}$,

respectively, and the crude extract protein concentrations, with data presented in Table 5.3.

Total hydrogenase and CODH activity shows an 8-fold and a near 2-fold reduction for the *ΔagrD1D2* mutant respectively. Whereas alcohol dehydrogenase was found to be 2-fold increased, however high statistical noise resulted in no statistical significance in this reading. The trend in specific activity values (U/mg) for hydrogenase, CODH and Adh activity in WT, Table 5.3, generally agrees with those found in literature (S. Wang *et al.*, 2013; Mock *et al.*, 2015), with hydrogenase being most active, followed by CODH, and lastly Adh. Crude cell extracts in this study also possessed a short life-span which particularly limited the number of assays possible.

Table 5.3, Specific activities of selected enzymes from cultures grown on fructose.

Enzyme	Substrates	Specific activity (U/mg)		P-value
		WT	***ΔagrD1D2***	
Hydrogenase	H_2 + MV	28.6 ± 2	3.5 ± 0.7	P <0.0001
Carbon monoxide dehydrogenase (CODH)	CO + MV	1.32 ± 0.1	0.71 ± 0.3	P <0.01
Alcohol dehydrogenase	Acetaldehyde + NADH	0.007	0.015	P = 0.07

One unit of activity is equivalent to 2 mol of electrons transferred per minute.

5.3 Discussion

The aim of this chapter was demonstrating the plausible connection between *C. autoethanogenum*'s Agr system and the WL-pathway, and whether QS plays a role in influencing the production of integral WL-pathway enzymes, such as the CODH/ACS-complex. Proteomics analysis and enzyme assays were employed to elucidate this question and in doing so, not only WL-pathway proteins exhibited differences, but a collection of unexpected proteins were also differently produced in the Δ*agrD1D2* mutant.

Regarding proteomic analysis methods, the iTRAQ method offers a robust and high-throughput quantitative analysis and comparison of proteomes, but as with many analytical systems its limitation requires mentioning. In complex mixtures such as bacterial lysates, concerns of accuracy were put into question as it was suggested that iTRAQ's isobaric tags may underestimate protein abundance (Ow *et al.*, 2009; Perkel,

2009). Two short-comings were experimentally identified; isotopic impurity in iTRAQ labels (manufacturing defect), and more importantly, highly complex mixtures can contribute to the dampening down of peptide intensity differences (Ow *et al.*, 2009). Ultimately, the *direction* of the up- and down-regulations remain unchanged. With regards to this chapter's work, absolute fold changes presented in Tables 5.1 and 5.2 may not seem overly large, but with the above concept in mind it can be argued that the trend of regulations is what matters importantly.

The hypothesis of an Agr-influenced production of CODH/ACS complex can be compared with Liew *et. al.*'s (2016) work that explicitly disabled *acsA* (CLAU_1578-1759), encoding a core component of this complex which resulted in surprisingly similar phenotypes to those found in *ΔagrD1D2*. Liew *et. al.*'s study gave two vital clues in the form of elevated CO_2 pressure in bottle headspace, and an increased ethanol output which was proposed as an alternative electron sink. However, several important differences must be highlighted between the Δ*acsA* and *ΔagrD1D2* mutants. Unlike *ΔacsA*, *ΔagrD1D2* grew at a generally similar rate to WT when on fructose, and eventually adapted to growth on CO gas after several repeated passages, which nearly matched OD levels to that of the WT. Moreover, Liew *et. al.* (2016), suggested that the Δ*acsA* mutant relies on higher aldehyde:ferredoxin oxidoreductase (AOR) activity in order to reduce acetate into acetaldehyde, which is converted into ethanol via AdhE with the use of surplus reducing equivalents from the EMP pathway. Contrary to this suggestion, Δ*agrD1D2* proteomics analysis revealed all its AOR copies to be down-regulated, and therefore similar reasoning as to why more overall ethanol is produced in Δ*agrD1D2*, cannot be fully supported.

Apparent similarities between Δ*acsA* and Δ*agrD1D2* phenotypes allow the assumption that the CODH/ACS complex is perhaps affected to some extent, as with Δ*agrD1D2*'s lack of peptide signalling capability, the mutant is may be unable to properly mediate *when* CODH/ACS should be expressed, and instead remains expressed at heterotrophic levels. This hypothesis can be supported as shown in Table 5.1, which describes the CODH/ACS complex, AcsA (CLAU_1578-79), and the constituent subunit, AcsB (CLAU_1566), to be some of the most down-regulated enzymes in the proteomic comparison, in addition, the third and final CODH enzyme (CLAU_2924), is also noted to be down-regulated. Furthermore, enzyme assays also showed a 2-fold less of total CODH activity in Δ*agrD1D2*.

Reactions within WL-pathway have been thought to be the main mechanism for energy conservation in acetogens (Ragsdale and Pierce, 2008) however recent progress revealed that the system to be more complex; During the formation of 1 acetate, 1 ATP is gained through substrate level phosphorylation via acetate kinase, but this ATP is used as compensation during the reduction of CO_2 to formate (Schuchmann 2014). With this in mind the WL-pathway is considered to be coupled to supporting energy conservation mechanisms, such as the reduced ferredoxin:NAD^+ oxidoreductase (Rnf) complex (Schuchmann and Müller, 2014; Mock *et al.*, 2015). This 6-subunit, trans-membrane system plays an important role in net ATP formation by catalysing electron transfer from reduced ferredoxin to NAD^+, thereby creating a chemiosmotic gradient by pumping intracellular H^+ ions across the cell membrane. This extracellular H^+ ion moiety in turn drives the F_0F_1 ATP-synthase complex, which pumps the ions back into the cell and conserves part of the resulting free energy to phosphorylate ADP to ATP (Köpke *et al.*, 2010; Tremblay *et al.*, 2013).

It was found in this proteomic comparison, Rnf subunits, G, B and C, along with ATP-synthase subunit beta, were upregulated, which is interesting for two reasons. Firstly, the Rnf-complex is thought to be essential for autotrophic growth (Tremblay *et al.*, 2013) and thus is accordingly upregulated (Marcellin *et al.*, 2016). Secondly, considering that Δ*agrD1D2* may be generating less ATP due to decreased acetate formation (this study, chapter 4), is it possible that the cells have to employ other means of conserving energy in the form of ATP? During the breakdown of fructose one ferredoxin is reduced for each molecule of acetyl-CoA formed. Considering this fact with the down-regulation of the WL-pathway, less reduced ferredoxin is needed for the reduction of CO_2 to CO, this surplus might feed into the Rnf complex and results in NAD+ reduction alongside energy conservation as described above. This connection maybe tenuous, as ATP-utilising formate dehydrogenase, (MoeA) was downregulated, therefore this ATP surplus could be offsetting a different metabolic activity. Furthermore, given the reduced expression of the AOR enzyme (Table 5.1), this additional NADH may then play a role in furthering ethanol production in the Δ*agrD1D2* mutant via acetaldehyde and alcohol dehydrogenase, as observed in chapter 4. While this hypothesis is in conflict with data presented by Mock *et. al.* (2015), suggesting that enzymes appear to prefer NADPH to NADH under heterotrophic conditions. Therefore, alcohol enzymes that are classically considered to play a role in ethanol production in *C. autoethanogenum*, appear to be downregulated in the proteomics study, while others are upregulated, in particular

enzymes associated with BMC clusters. It is notable to point out that in other bacteria, the latter BMC-associated alcohol enzymes are often typically NADH-dependent (Cheng *et al.*, 2012; Kerfeld and Erbilgin, 2015).

Related to this fact, proteomics data showed two distinct groups of acetaldehyde and alcohol dehydrogenases, both separated on either end of this spectrum, Table 5.1. The down-regulated group included bifunctional alcohol and acetaldehyde dehydrogenases, alongside mentioned AOR enzymes, both groups are studied to be upregulated during autotrophic conditions, specifically in the formation of ethanol via acetate reduction while growing on CO_2/H_2, (S. Wang *et al.*, 2013; Mock *et al.*, 2015; Marcellin *et al.*, 2016). The up-regulated group, has several alcohol and acetaldehyde dehydrogenases, one of which is a primary:secondary alcohol dehydrogenases (CaADH), (CLAU_0532), involved in acetoin to 2,3-butanediol reduction (Köpke, Mihalcea, Liew, *et al.*, 2011). The upregulated collection of alcohol and acetaldehyde dehydrogenases were found to be encoded in a cluster of genes also encoding up-regulated BMC components as well as 1,2-propanediol and ethanolamine utilisation enzymes, Table 5.2, which suggests that the entire operon is co-transcribed, and coincidentally upregulated during Δ*agrD1D2'*s heterotrophic growth. This upregulation perhaps concurs with the higher presence of ethanol in the organism, as a means of sequestering toxic aldehyde intermediate metabolites. Studies in the acetogen, *A. woodii* have strongly suggested that there may be an unspecific means for sensing alcohols, and/or aldehyde intermediates which regulates this cluster (Schuchmann *et al.*, 2015).

An observed upregulation of pyruvate-metabolic enzymes was also noted, specifically pyruvate:ferredoxin oxidoreductase (PFOR) which facilitates acetyl-CoA formation (Ragsdale, 2003; Mock *et al.*, 2015). This reaction is described in the introduction using equation 5.1, which explained the abundant CO_2 partial pressure inside *ΔagrD1D2* serum bottles. Looking beyond the pyruvate metabolic pathway, several EMP pathway enzymes have been noted to be upregulated, presumably because the Δ*agrD1D2* mutant was grown primarily in a heterotrophic state. It was also noted that EMP pathway enzymes were generally expressed marginally higher, which could be owed to the increased fructose consumption as described in chapter 4.

From the evidence that has been presented in this chapter, it may be tenuous to suggest yet that CODH/ACS-complex expression can be correlated with population density, but

the social benefits of engaging in a finely-tuned, population-based activity can be realised through 2 important points;

- Mediating the WL-pathway genes during specific growth conditions may spare valuable ATP from unnecessary WL-pathway protein synthesis, as 50 % of total cellular ATP is expended in cellular biosynthesis (Stouthamer, 1973). Moreover, recent findings have revealed that *C. autoethanogenum* autotrophic growth indeed requires more ATP production maintenance costs at higher biomass (Valgepea, de Souza Pinto Lemgruber, *et al.*, 2017; Valgepea, Loi, *et al.*, 2017).
- The ability to switch to autotrophic conditions in due time, may safeguards long-term cell survival, as described experimentally in section 4.2.3.4.

The CODH/ACS-complex as well as other CODH enzymes have a central role in the WL-pathway (Ragsdalae, 2004), and from their disruption, a "knock-on" effecting can be seen in other genes (Fungmin Liew *et al.*, 2016). A reduced production of acetate, due to a reduced ability to assimilate and reduce CO_2, has likely directed *ΔagrD1D2* towards complete fructose utilisation, whereas WT strains begin to switch WL-pathway enzymes for gas metabolism. The reducing equivalents produced from GADPH can no longer serve for carbon fixation via the WL pathway and are instead redistributed into alcohol and acetaldehyde dehydrogenases that exploit acetyl-CoA and acetaldehyde as an electron sinks, explaining increase ethanol titres (this study, chapter 4), this hypothesis was also shown in *acsA*-deficient mutants, (Fungmin Liew *et al.*, 2016). Similarly, pyruvate metabolism appears to be enhanced in Δ*agrD1D2* mutants, and in a catabolic state, facilitates acetyl-CoA synthesis through oxidation of pyruvate via PFOR. This provides additional CO_2 that can no longer be completely re-assimilated, contributing to higher gas pressures. Acetate production via the WL-pathway is not abolished but instead, downregulated, to which AOR is not observed to be utilised as reducing equivalents are assimilated elsewhere. From these assertions, a new WL-pathway scheme can be drawn that is affected by the absence of an active Agr-system, Figure 5.2.

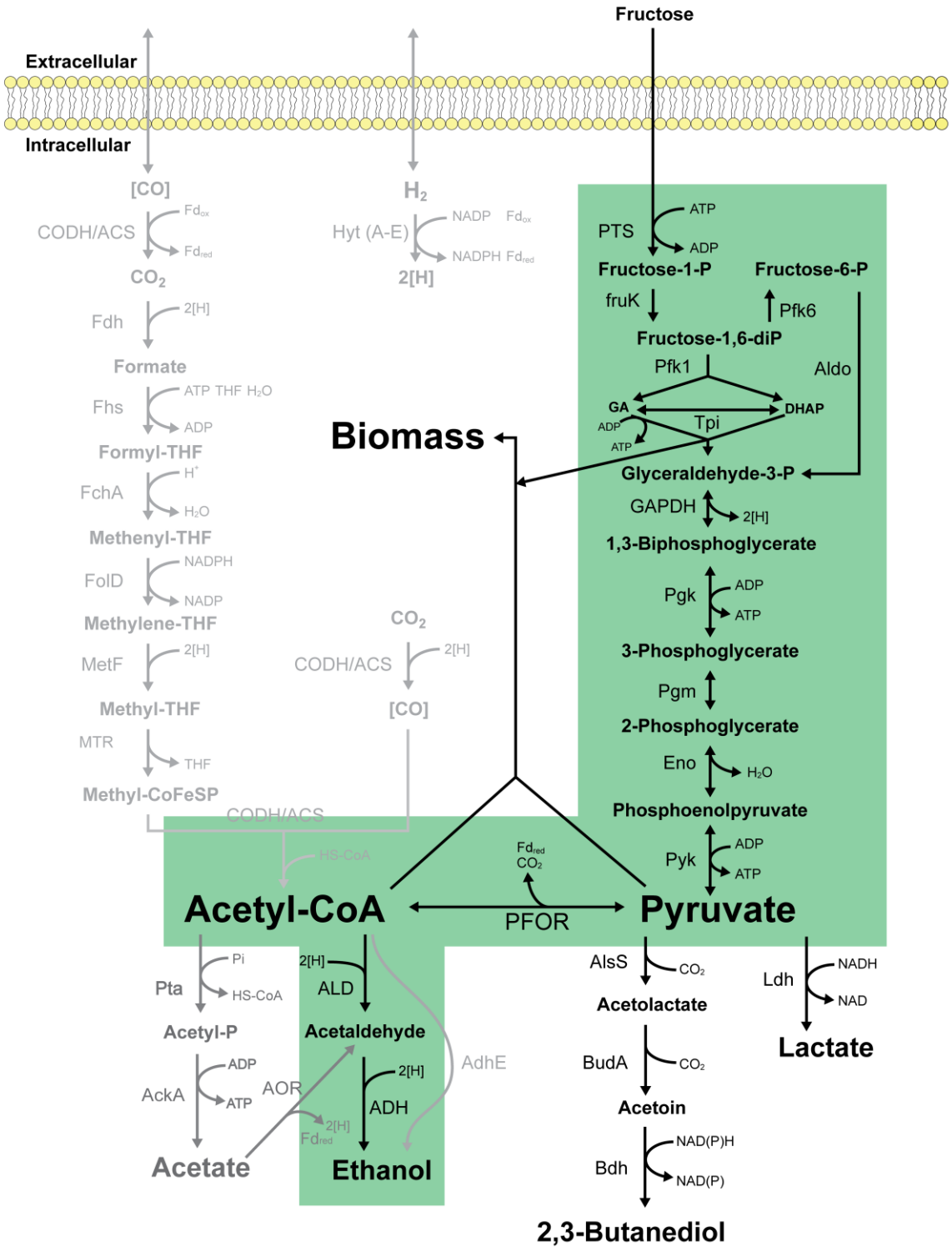


Figure 5.2, Proposed overview of the WL-pathway and main fermentation pathways active in *C. autoethanogenum* in the absence of a functional Agr system: Greyed-out regions represent under-expressed enzymes, which includes the WL-pathway. Green highlighted region represents consequential route of metabolic flux in Δ*agrD1D2* mutants.

Chapter 6

General Discussion

6.1 Background

The Gram-positive, anaerobic, acetogenic bacterium *Clostridium autoethanogenum* has been noted for its important biotechnological potential in the process of industrialised gas fermentation (Köpke, Mihalcea, Bromley, *et al.*, 2011). *C. autoethanogenum* along with several other notable acetogenic bacteria has formed a rational approach to producing platform chemicals from capturing industrial waste-gases and simultaneously mitigating GHG emissions (Dürre, 2017). Moreover, the growing repertoire of gene editing tools in *Clostridium* spp. has allowed for significant improvements in enhancing fermentation processes and expanding the spectrum of products (Fungmin Liew *et al.*, 2016). Despite these accomplishments there are still many unanswered questions which can further advance the understanding of gas-fermenting bacteria.

Amongst these questions is the topic of QS which is seldom examined despite the considerable efforts already invested into QS research of pathogenic bacteria. The Agr QS system that has been identified in *C. autoethanogenum* is widely conserved across Firmicutes, both pathogenic and non-pathogenic species alike (Wuster and Babu, 2008). The system's role in pathogens has been that associated with toxin production, sporulation and biofilm synthesis (Novick *et al.*, 1993; Ji, Beavis and Novick, 1997; Cooksley *et al.*, 2012; Darkoh *et al.*, 2015; Ohtani and Shimizu, 2015; Vidal, Shak and Canizalez-Roman, 2015). Despite these findings, Agr systems were rarely studied in non-pathogens, and it was not until work in *C. acetobutylicum* highlighting its function in sporulation and granulose accumulation (Steiner *et al.*, 2012). Steiner *et. al.*'s work was one of several motivators that led this project into asking similar questions with regards to *C. autoethanogenum*.

6.2 Key findings and future work

6.2.1 Manipulating *C. autoethanogenum*

6.2.1.1 Barriers to conjugation

As of starting this project in 2013, tools and methods to manipulate *Clostridium* spp. were readily available but not optimised for *C. autoethanogenum*. Consequently, creation of *agrD* IFD mutants through ACE was hindered by exceptionally poor conjugation frequencies that required improvement.

Although these issues were later resolved using *E. coli* Sexpress (Woods, 2018), the majority of *agrD* mutants were already produced through using *E. coli* CA434 as conjugative host. The consequence of CA434's low conjugation frequencies was that the few transformed colonies obtained resulted in inflexibility when selecting for single- and double-cross mutants, as either event occurs randomly and spontaneously (Ng *et al.*, 2013). Moreover, isolating a pure mutant through repeatedly passaging single colonies raised the chances of SNVs and Indels (Miller, 1996; Foster *et al.*, 2015). This was observed in Table 3.3, where mutants created with CA434's low conjugation frequencies were found to possess a large amount of SNVs. In contrast, later mutants *ΔagrD1D2_16*, *ΔagrD1D2_34* and complemented mutants, *ΔagrD1D2_D1^{comp}* and *ΔagrD1D2_D2^{comp}* used Sexpress as a donor and experienced relatively low number of SNVs and Indels. This resulted in higher conjugation efficiencies leading to a substantially larger pool of single-colony plasmid integrant mutants to screen. However, this observation requires further examination through systematic testing before final conclusions can be drawn. In conclusion, difficulties faced when transforming uncharacterised organisms should be addressed first by understanding inherent restriction methylation systems, for instance by employing methods such as PacBio-sequencing (Roberts *et al.*, 2015) that reveal prevailing methylation patterns. If this data is unavailable, then systematic testing of different *E. coli* donors with different sets of methyltransferases should be performed.

Future gene editing in *C. autoethanogenum* may eventually reduce the usage of IFD plasmids and instead rely on CRISPR for targeted gene disruptions, (Nagaraju *et al.*, 2016). CRISPR has already been established as a valid tool in a variety of *Clostridium* spp. (Richter *et al.*, 2012; Brown *et al.*, 2014; Wang *et al.*, 2015; Xu *et al.*, 2015; Huang *et al.*, 2016; Nagaraju *et al.*, 2016; Pyne *et al.*, 2016), including *C. ljungdahlii* and *C. autoethanogenum*.

After initial attempts in *C. autoethanogenum*, the method still requires optimisation as native CRISPR systems are present, this contrasts with *C. ljungdahlii*'s absence of a CRISPR system (Brown *et al.*, 2014) where gene deletions showed 50 – 100 % efficiency, (Huang *et al.*, 2016). Possible solutions in improving efficiencies include using a non-constitutive promoter and a nickase-only Cas9 variant (Nagaraju *et al.*, 2016).

As a part continuation of this project, efforts to use CRISPR in order to knock-out AgrB with aims of replicating *ΔagrD1D2* phenotypes are being undertaken by Dr Christopher Humphreys, (Uni. of Nottingham). It is anticipated that precise targeting would produce a "clean" mutant without potential polar effects that have so far been the probable cause of the unexpected phenotypes observed from insertion of ClosTron plasmid pMTL007C-E2::CAUagrB-234a.

6.2.1.2 Usage of a*grD* promoters in synthetic biology

Efforts to complement *ΔagrD1D2* were only possible through creating complementation plasmids without native promoters of either *agrD* gene. It was speculated that high expression exerted through each respective promoter created a toxic environment for *E. coli* hosts, but it is not yet known what induces *agrD* promoter activity in their native background. Therefore, it is suitable to ask whether isolating and characterising these promoters would be beneficial in further synthetic biology studies.

Several reviews have observed the possibilities that QS circuitry can offer in synthetic applications, amongst those digital switches, timers and logic gates (Atkinson *et al.*, 2003; Danino *et al.*, 2010; B. Wang *et al.*, 2011; Carter, Valdes and Bentley, 2012; Bacchus and Fussenegger, 2013; LaSarre and Federle, 2013). Studies furthering the understanding of QS systems in either Gram-type allow contributions towards further novel applications (Carter, Valdes and Bentley, 2012; Bacchus and Fussenegger, 2013). Examples over the past decade have shown exogenous QS systems employed within *E. coli* hosts which include demonstrating "plug and play" synthetic circuitry (Kobayashi *et al.*, 2004), *E. coli* that are able to sense and kill *P. aeruginosa* cells (Saeidi *et al.*, 2014), spatiotemporal gene expression (Basu *et al.*, 2004) and modulation of cell density and mobility into stripe patterns (Liu *et al.*, 2011) to name a few. Most of these studies have exploited the well-understood *V. fischeri* LuxI/LuxR QS system, however Gram-positive QS systems have not yet attracted as much attention towards synthetic circuits. The best example to date, has been through introducing an *S. aureus*-based synthetic *agr* QS system into *Bacillus*

megaterium, (Marchand and Collins, 2016). The system was split into two *B. megaterium* strains; a signal producing *agrBD* strain and a signal receiving *agrCA* strain. A GFP fluorescence signal output used to successfully indicate communication between sender and receiver (Marchand and Collins, 2016).

This noteworthy example highlighted the possibility of introducing a Gram-positive QS system component, into a synthetic circuit scenario. Could a similar arrangement be adopted within *Clostridium* spp.? It could be readily imagined that a synthetic quorum sensing system within a industrially-relevant acetogen to be indispensable in a large fermenter system, especially with working examples already mentioned in *E. coli*. Despite these ambitions, numerous barriers still require overcoming such as creating stable QS changes in the genome, fine-tuning synthetic systems, balancing the metabolic load on organism, avoiding signal, cross-talk and above all, better understanding of individual QS components in *Clostridium* spp. (Brenner, You and Arnold, 2008; Daniel *et al.*, 2013; Marchand and Collins, 2016).

6.2.1.3 Synthetic peptide complementation

The attempt to chemically complement *ΔagrD1D2* with predicted synthetic AgrD peptides was not successful, as limited quantities of inferior quality peptides may have contributed to this short-coming. It would be therefore appropriate to repeat this endeavour with a higher-quality procurement of peptides. Moreover, future work may consist of isolating and identifying *C. autoethanogenum*'s AgrD peptides that could subsequently be accurately synthetically reproduced and tested.

6.2.2 AgrD signalling

Results in chapter 4 described the remarkable differences that mutant *ΔagrD1D2* exhibited in comparison to *ΔagrD1*, *ΔagrD2* and the WT strain. These findings imply that either AgrD signal is sufficient for maintaining regular cell function, as there were no significant differences between single knockout mutants and the WT. Moreover, it may also imply that either AgrD signal received through respective AgrC receptors stimulates the same response regulator. This regulator has not been yet identified in *C. autoethanogenum* but as previously speculated in chapter 3, the Spo0E-like kinase (CLAU_0813) may in fact behave as an AgrA equivalent signal relay mechanism which is

also shares high similarity with homologous protein in related acetogens e.g. *C. ljungdahlii* and *C. ragsdalei*.

Attempts at characterising the ClosTron-disrupted *agrB* mutant revealed unexpected phenotypes, with a markedly higher acetate and lower ethanol output with respect to the WT, (data not shown). As the anti-sense insertion contained a sense direction RAM cassette, it would be reasonable to speculate that a down-stream polar effect may have prematurely triggered, or negatively interfered with, *spo0E* transcription (Kuehne and Minton, 2012). Future studies could aim to characterise this protein through gene deletion or transcriptional knock-down, via antisense RNA expression (Desai and Papoutsakis, 1999).

The observed phenotypes of *ΔagrD1D2* in chapters 4 and 5 strongly suggested that the Agr system influences WL-pathway, specifically the CODH/ACS complex. This assertion is further corroborated with recent studies that directly targeted the CODH/ACS complex, resulting in similar phenotypes to those found in this project (Fungmin Liew *et al.*, 2016). The mechanism of Agr system influencing this complex can only be speculated at this point in time, but the result may be a "knock-on" effect onto the rest of the WL-pathway.

Reducing equivalents generated from the glycolytic EMP pathway are thought to be the alternative means in which acetyl-CoA is reduced, leading to higher formation of ethanol observed in the *ΔagrD1D2* mutant. How these increased reactions rates are mediated remains uncertain, but it was speculated that increased availability of reduced ferredoxin, and up-regulated Rnf-complex sub-units, C, G, B and ATPase lead to a concomitant increase of NADH production. This additional NADH, in the absence of CO_2 would increase the use of acetyl-CoA as an electron acceptor. This may explain the higher ethanol yields through acetaldehyde and alcohol dehydrogenase. NADH usage in BMC-associated alcohol enzymes in other bacteria has also been observed (Cheng *et al.*, 2012; Kerfeld and Erbilgin, 2015). However, a counter-argument to this hypothesis suggests that heterotrophic counterparts to these ethanol pathway enzymes require NADPH (Mock *et al.*, 2015). Interestingly, an example shows that addition of reducing agents into *C. ljungdahlii* cultures did marginally increase ethanol titres (Klasson *et al.*, 1991), which still leaves this topic open to discussion.

It is still not exactly clear how *C. autoethanogenum*'s Agr system influences the WL-pathway. Literature gives some general examples of Agr operation in *Clostridium* spp., but mechanisms of the Agr regulatory networks differ between species or remain elusive.

Sporulation has been found to be linked with the Agr system in several species; *C. botulinum, C. sporogenes, C. acetobutylicum and C. perfringens* (Cooksley *et al.*, 2010; Li *et al.*, 2011; Steiner *et al.*, 2012). Evidence of this relationship could not be established in *C. autoethanogenum* as it was observed not to sporulate readily. In specific cases, the *C. perfringens* Agr (CpAL) system is noted to interact indirectly with other two-component systems (TCSs) involved in virulence and toxin production (Vidal, Shak and Canizalez-Roman, 2015). Similarly, Agr-dependent TCSs have been speculated to function in *C. difficile* (Darkoh *et al.*, 2015). It could be then speculated that similar Agr-dependent TCS arrangements possibly function in *C. autoethanogenum* which would require identification in future studies.

6.2.2.1 Future assays and whole-transcriptome analysis

Due to time constraints, it was possible to only assay a select number of WL-pathway enzymes as described in chapter 5. This range was based on proteomics data showing relative protein expression differences between *ΔagrD1D2* and the WT, of which included CODH, hydrogenase and alcohol dehydrogenase to show the biggest differences. Future studies involving *agrD* mutants could broaden the scope of assays to include further WL-pathway enzymes such as formate dehydrogenase, acetaldehyde dehydrogenase and pyruvate: ferredoxin oxidoreductase. Proven methods would allow further study of these enzymes in terms of their activity (S. Wang *et al.*, 2013; Mock *et al.*, 2015), but a more appropriate option would be to examine transcriptome-wide changes.

To date, studies in *Clostridium* spp. have taken advantages of RNA sequencing technology to better understand genome-wide transcriptional changes (Shimizu *et al.*, 2002; Alsaker *et al.*, 2004; Scaria *et al.*, 2011; Y. Wang *et al.*, 2011; Martin *et al.*, 2013; Tan *et al.*, 2013; Wilson *et al.*, 2013; Phillips, Huhnke and Atiyeh, 2017), that include *C. autoethanogenum* (Marcellin *et al.*, 2016; Henstra, 2018). Published findings typically carry a wealth of RNA sequencing data that exemplify and compare an entire strain's transcriptome which can be used to reference genes of interest. Thus, *agrD* mutants could be subjected similarly to both heterotrophic and autotrophic conditions as described in Marcellin *et. al*.'s (2015) study to determine differences in transcription levels. It would be interesting to observe any obvious changes in WL-pathway enzymes, but also in uncharacterised and hypothetical proteins, which could be used to identify possible TCSs interactions.

6.2.3 The Agr system and cell survival

In this study, questions addressed the utility of the Agr system in *C. autoethanogenum* as figures 4.5 and 4.13 described the *ΔagrD1D2* strains gradual adaptation on CO gas and long-term cell viability respectively. It was therefore partly reasoned that the presence of a functional Agr system equips the organism to adapt and prepare for stresses that are associated with a rapidly growing population, which includes dynamic adjustment to different carbon sources (Williams *et al.*, 2007; Dandekar, Chugani and Greenberg, 2012; Cornforth *et al.*, 2014). This facultative adaptation could come in the form of upregulating necessary metabolic pathway enzymes such as the WL-pathway, which otherwise would be metabolically taxing if expressed constitutively (Stouthamer, 1973).

Moreover, it was speculated as to why *C. autoethanogenum* possesses two copies of differently expressed AgrD peptides. As noted in transcriptomic data from continuous cultures (Marcellin *et al.*, 2016; Henstra, 2018), it was suggested that *agrD1* is significantly more expressed in relation to *agrD2* which can be correlated with working in combination to resolve the organism's environment, and thereby discern between population density and environmental mass transfer states (Cornforth *et al.*, 2014). This is intriguing as neither single-knock out showed any appreciable phenotypic differences which furthers the arguments that response regulators are activated irrespective of which AgrD signal is recognised. Future studies could include focus on these single-knock out mutants, and perhaps include them in RNA-seq experiments, and/or proteomics experiments if indeed there are mutant strain-specific changes.

Other possible uses of the Agr system include responses to environmental stressors. Several Gram-positive bacteria including *S. aureus* and *Listeria monocytogenes* engage in biofilm formation through QS systems as a stress response (Kong, Vuong and Otto, 2006; Riedel *et al.*, 2009; Li *et al.*, 2011; Pantaléon *et al.*, 2014). In *C. perfringens*, biofilm formation is mediated through the Agr system (Vidal, Shak and Canizalez-Roman, 2015), but this activity has not been thoroughly studied in *Clostridium* spp. Known stressors such as salt concentration induces biofilm formation as shown in *C. ljungdahlii*, but transcriptomics data did not fully reveal Agr activity (Philips *et al.*, 2017). When NaCl was supplemented into cultures of *C. autoethanogenum*, samples showed significantly reduced growth and aggregation of the WT, in contrast *ΔagrD1D2*'s planktonic growth. The potential benefits of a stress-activated Agr system could be a means of coping in adverse environments as found in several pathogens (Dawson *et al.*, 2012; Dapa *et al.*,

2013). Interestingly, salt-induced biofilm is noted to have benefits in industrial applications, where biofilms in bio-electrochemical fermenter electrodes can aid in electron transfer rates (Philips *et al.*, 2017).

6.2.3.1 Survival in co-cultures

Future work exploring the utility of the Agr system may be exemplified through co-culture experiments. It can be imagined that *C. autoethanogenum*'s natural environment is subjected to constant competition with other microbiota over shared carbon sources. A functional Agr system would help in directing a growing *C. autoethanogenum* population into adapting to alternative carbon sources that are generally unavailable to the rest of the microbiome before more accessible nutrients run out. For example, experimentation of this premise would use combined cultures of *C. acetobutylicum* and *C. autoethanogenum*, where rapid sugar uptake by *C. acetobutylicum* would prompt *C. autoethanogenum* to switch to gas metabolism and enable the strain to survive on surplus CO_2 and H_2 produced through *C. acetobutylicum* respiration. Mutant *ΔagrD1D2* would be similarly tested, with expected impaired survival. Colony counting and identification could be performed through dilution plating and granulose staining. Numerous publications describe co-culture experiments between various *Clostridium* spp., and other genera which demonstrated synergistic activities that improved product titres particularly in hydrogen and ethanol production (Ng and Zeikus, 1981; Fang, Zhu and Zhang, 2006; Masset *et al.*, 2012; Nasr *et al.*, 2017).

6.3 Concluding remarks

This study has laid the foundation in understanding the Agr QS system in the acetogen, *C. autoethanogenum* through employing optimised targeted gene mutagenesis and robust mutant characterisation approaches. The conclusions heavily suggest that the ancient, carbon-fixing WL-pathway is influenced by the organism's Agr system. The Agr system is likely used as a means of evaluating population density with respect to environmental conditions that include carbon source availability.

Prior to this work, studies describing QS systems in *Clostridium* spp. have been primarily focused on pathogenic species, with notable exceptions in *C. acetobutylicum* (Steiner *et al.*, 2012; Kotte *et al.*, 2017). Presently, studies of QS systems in non-pathogenic acetogenic bacteria are scarce but should be considered a valid field of interest. Acetogens are proven biocatalysts in industrial chemical and fuel processes that undergo

continual optimisation through genetic engineering. Furthering the understanding of inherent QS systems can contribute in this future pursuit.

Chapter 7

Appendix

7.1 Oligonucleotide list

Table 7.1, Oligonucleotides used in this project: Underlined sequences represent restriction sites preceded by ATAT nonsense sequences. Table continues next page.

Oligonucleotides	Sequence 5'-3'	Description
pMTL_84151 screening		
84151_F	GAGCAAGGCAAGACC	IFD plasmid screening
84151_R	GCTAAAAATCTATTTATATTTCACC	
IFD plasmid construction primers		
0816_LHA_F	ATATATCCGCGGGGTACTTTAAAAAAATCATCC	SOE-PCR primers for IFD *agrD1* plasmids
0816_LHA_R	TTCATCAGGATTTTTCATTTCATACACTCC	
0816_RHA_F	AATGAAAAATCCTGATGAATTCTTAAAGTAA	
0816_RHA_R	ATATATGGCGCGCCAGATAAGACCTTCTTTTTCG	
3094_LHA_F	ATATATCCGCGGCCAGATAATTTGTTTTTAGC	SOE-PCR primers for IFD *agrD2* plasmids
3094_LHA_R	TTCATCAGGATTTTTCATTTTTCATATACC	
3094_RHA_F	AATGAAAAATCCTGATGAACTTTTGAAATAA	
3094_RHA_R	ATATATGGCGCGCCGCAAAATACTAAGATAGTAACAC	
IFD mutant screening		
0816_SCR_F	CCTTATTTCTAAAATTATGG	Flanking primers screening *agrD1*
0816_SCR_R	GGTCTGTTTTTCCTTTAAGC	
3094_SCR_F	GGCTATTTACAACATTGTCC	Flanking primers screening *agrD2*
3094_SCR_R	CATTTACAAGTGTTTGTTTTTCC	
N_R0816	GGCAGCAATAGTTATAGTACC	Internal *agrD* screening primer for deletion mutants
N_R3094	CATTTACAAGTGTTTGTTTTTCC	
Complementation plasmid construction primers		
0816_Comp_F	ATATGCGGCCGCGCTATATAATTACAAAATTTTTACTAAAGG	Complementation fragment construction primers
0816_Comp_R	ATATGCTAGCGCAGTATTTGTAAAAAAACTAAG	
3094_Comp_F	ATATGCGGCCGCGAATTTATAAATTAAAACATTATAGGAGG	
3094_Comp_R	ATATGCTAGCACCTAAGTATTTGTAAAAGAC	
Universal screening primers		
M13F	TGTAAAACGACGGCCAGT	*pyrE* restoration and deletion screening primer
EBS2	GCAAGTTTCTAATTTCGGTTCTCATCCGATAGAGGAAAGTGTCT	ClosTron insertion screening primers
agrB_Scr_F	GCAAAGTTTAGAAGCTTAAAGG	
pyrE_FSP	CATCAAAGCTATACTATTTTCCGTATTTACATTTGGG	

pyrE_RSP	GTTTGCAGTACCTACTTGAATTGCATAAGCTC	Flanking primers screening *pyrE*

7.2 Read-mapping summary reports

Table 7.2, Summary reports of Illumina read-mapping: Reference genome used *C. autoethanogenum* DSM10061 (Humphreys *et al.*, 2015). Table continues next page.

***C. autoethanogenum* strain:**	**Count**	**Read [%]**	**Average length**	**Base count**	**Bases [%]**
WT *pyrE*^{fix}					
References	1	-	4,352,446	4,352,446	-
Mapped reads	4,832,053	98.26	217.59	1,051,428,855	98.36
Not mapped reads	85,615	1.74	204.25	17,486,590	1.64
Reads in pairs	4,440,254	90.29	570.81	999,545,810	93.51
Broken paired reads	391,799	7.97	132.42	51,883,045	4.85
Total reads	4,917,668	100	217.36	1,068,915,445	100
ΔagrD1					
References	1	-	4,352,446	4,352,446	-
Mapped reads	3,194,497	98.45	219.04	699,708,953	98.57
Not mapped reads	50,309	1.55	201.93	10,158,665	1.43
Reads in pairs	2,930,978	90.33	552.68	659,919,415	92.96
Broken paired reads	263,519	8.12	150.99	39,789,538	5.61
Total reads	3,244,806	100	218.77	709,867,618	100
ΔagrD2					
References	1	-	4,352,446	4,352,446	-
Mapped reads	2,709,749	97.93	215.72	584,541,320	98.11
Not mapped reads	57,327	2.07	196.19	11,247,050	1.89
Reads in pairs	2,487,960	89.91	621.6	555,987,844	93.32
Broken paired reads	221,789	8.02	128.74	28,553,476	4.79
Total reads	2,767,076	100	215.31	595,788,370	100
ΔagrD2_9					
References	1	-	4,352,446	4,352,446	-
Mapped reads	916,657	97.49	209.66	192,181,997	97.70
Not mapped reads	23,571	2.51	191.73	4,519,315	2.30
Reads in pairs	893,710	95.05	570.43	187,432,957	95.29
Broken paired reads	22,947	2.44	206.96	4,749,040	2.41
Total reads	940,228	100	209.21	196,701,312	100
ΔagrD2_27					
References	1	-	4,352,446	4,352,446	-
Mapped reads	2,016,382	98.26	212.46	428,401,057	98.43
Not mapped reads	35,656	1.74	191.87	6,841,199	1.57
Reads in pairs	2,006,312	97.77	566.03	426,303,420	97.95
Broken paired reads	10,070	0.49	208.31	2,097,637	0.48
Total reads	2,052,038	100	212.1	435,242,256	100

ΔagrD1D2					
References	1	-	4,352,446	4,352,446	-
Mapped reads	5,455,397	98.37	216.62	1,181,763,978	98.36
Not mapped reads	90,461	1.63	217.63	19,686,799	1.64
Reads in pairs	5,452,630	98.32	500.08	1,181,258,056	98.32
Broken paired reads	2,767	0.05	182.84	505,922	0.04
Total reads	5,545,858	100	216.64	1,201,450,777	100
ΔagrD1D2_16					
References	1	-	4,352,446	4,352,446	-
Mapped reads	1,343,536	97.60	226.33	304,083,726	97.81
Not mapped reads	33,104	2.40	205.55	6,804,536	2.19
Reads in pairs	1,266,866	92.03	655.38	294,815,326	94.83
Broken paired reads	76,670	5.57	120.89	9,268,400	2.98
Total reads	1,376,640	100	225.83	310,888,262	100
ΔagrD1D2_34					
References	1	-	4,352,446	4,352,446	-
Mapped reads	816,708	92.58	226.7	185,144,571	92.76
Not mapped reads	65,412	7.42	220.88	14,447,958	7.24
Reads in pairs	692,916	78.55	723.61	162,380,582	81.36
Broken paired reads	123,792	14.03	183.89	22,763,989	11.41
Total reads	882,120	100	226.26	199,592,529	100
ΔagrD1D2_D1^{comp}					
References	1	-	4,352,446	4,352,446	-
Mapped reads	1,683,390	96.76	227.69	383,287,363	96.91
Not mapped reads	56,308	3.24	216.96	12,216,753	3.09
Reads in pairs	1,503,368	86.42	643.59	356,784,368	90.21
Broken paired reads	180,022	10.35	147.22	26,502,995	6.70
Total reads	1,739,698	100	227.34	395,504,116	100
ΔagrD1D2_D2^{comp}					
References	1	-	4,352,446	4,352,446	-
Mapped reads	609,204	94.96	220.75	134,484,469	94.90
Not mapped reads	32,362	5.04	223.29	7,226,082	5.10
Reads in pairs	604,546	94.23	475.43	133,488,991	94.20
Broken paired reads	4,658	0.73	213.71	995,478	0.70
Total reads	641,566	100	220.88	141,710,551	100
agrB::CtermB-234a					
References	1	-	4,352,446	4,352,446	-
Mapped reads	1,744,919	96.43	228.48	398,678,295	96.64
Not mapped reads	64,633	3.57	214.34	13,853,296	3.36
Reads in pairs	1,533,520	84.75	675.15	361,906,587	87.73
Broken paired reads	211,399	11.68	173.94	36,771,708	8.91
Total reads	1,809,552	100	227.97	412,531,591	100

7.3 Supplementary growth experiment data

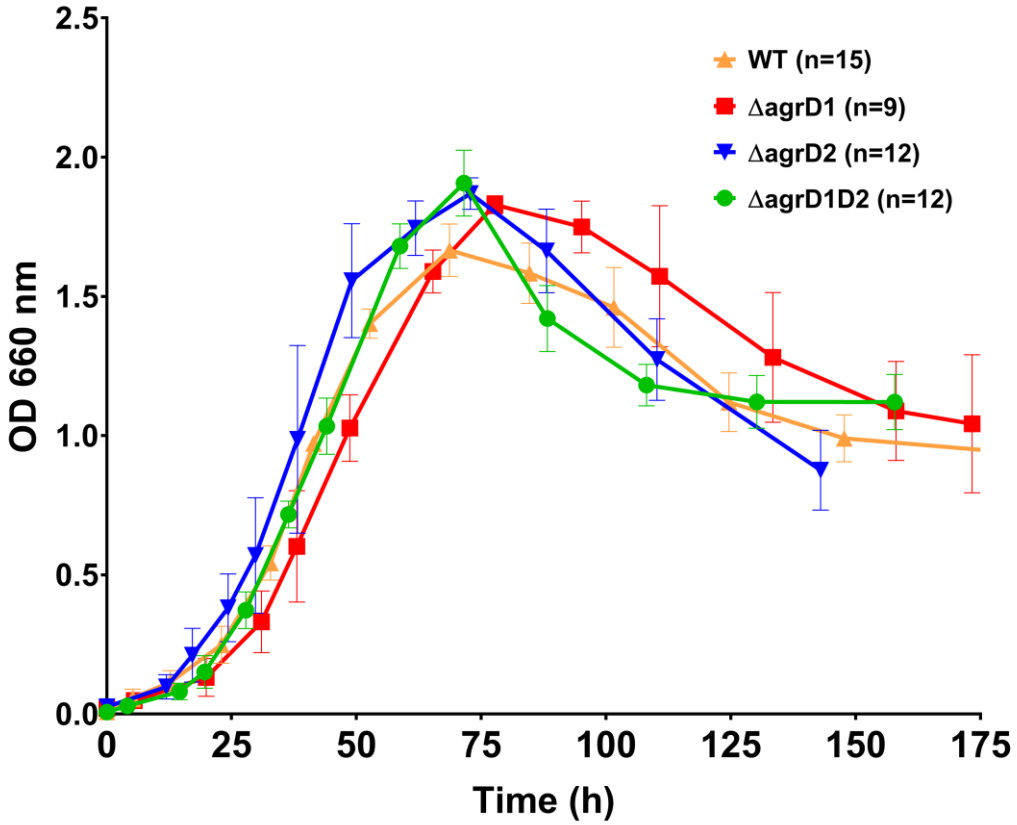


Figure 7.1, Growth of *C. autoethanogenum* WT and *agr* mutant strains on fructose, linear scale: Strains were grown in PETC medium containing 10 g/L fructose over an average period of 175 hours. Time points measured OD, and collected samples were saved for later solvent analysis. Error bars indicate standard error of the mean.

7.4 Full proteomics list

Table 7.3, Full list of protein fold changes in order of locus tags: Changes measured using ITRAQ methods described in Raut *et. al.* (2016). P value refers to an average P value. Table continues over next pages.

Uniprot ID	Locus Tag	Protein	P value	Fold change
U5RRQ8	CLAU_0016	Transcriptional regulator, MarR family	3.78E-07	-1.39
U5RNS7	CLAU_0024	Rubredoxin domain containing protein	1.59E-02	-1.37
U5RNF1	CLAU_0027	NAD(+) diphosphatase (EC 3.6.1.22)	1.46E-03	1.19
U5RSB8	CLAU_0035	Putative transcriptional regulator, GntR family (EC 2.6.1.57)	9.89E-03	1.18
U5RNW7	CLAU_0054	Chaperone protein HtpG (Heat shock protein HtpG) (High temperature protein G)	1.43E-02	-1.17
U5RP02	CLAU_0089	Aldehyde ferredoxin oxidoreductase (EC 1.2.7.5)	2.58E-07	-1.24
U5RP39	CLAU_0089	Aldehyde ferredoxin oxidoreductase (EC 1.2.7.5)	8.78E-05	-1.73

U5RP39	CLAU_0099	Aldehyde ferredoxin oxidoreductase	3.33E-19	-2.01
U5RNZ7	CLAU_0101	UBA/THIF-type NAD/FAD binding protein	2.06E-02	-1.52
U5RP61	CLAU_0117	Ketol-acid reductoisomerase (NADP(+)) (KARI) (EC 1.1.1.86) (Acetohydroxy-acid isomeroreductase) (AHIR) (Alpha-keto-beta-hydroxylacyl reductoisomerase)	9.84E-08	1.54
U5RSQ7	CLAU_0118	Dihydroxy-acid dehydratase (DAD) (EC 4.2.1.9)	5.07E-04	1.15
U5RST6	CLAU_0140	Dihydropteroate synthase DHPS	1.19E-02	-1.21
U5RS85	CLAU_0145	Cobalamin B12-binding domain protein	1.04E-02	-1.3
U5RNV1	CLAU_0205	Putative transcriptional regulator, GntR family (EC 2.6.1.57)	2.31E-04	1.34
U5RSY0	CLAU_0207	Aminotransferase (EC 2.6.1.-)	6.88E-06	1.17
U5RSE1	CLAU_0213	Transcriptional regulator, RpiR family	4.55E-03	-1.16
U5RP20	CLAU_0224	Histidine ammonia-lyase (Histidase) (EC 4.3.1.3)	6.41E-04	-1.37
U5RPF8	CLAU_0226	Urocanate hydratase (Urocanase) (EC 4.2.1.49) (Imidazolonepropionate hydrolase)	6.32E-05	-1.36
U5RP25	CLAU_0228	Glutamate formiminotransferase (EC 2.1.2.5)	2.74E-03	1.31
U5RP41	CLAU_0304	Molybdenum ABC transporter, periplasmic molybdate-binding protein	1.07E-02	1.12
U5RP91	CLAU_0308	Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3)	8.11E-03	-1.15
U5RSS2	CLAU_0370	(R,R)-butanediol dehydrogenase (EC 1.1.1.4)	2.67E-34	1.57
U5RSS7	CLAU_0375	Efflux transporter, RND family, MFP subunit	1.19E-02	1.19
U5RPW4	CLAU_0385	Cysteine desulfurase IscS (EC 2.8.1.7)	3.23E-02	-1.2
U5RPH4	CLAU_0388	Acetolactate synthase (EC 2.2.1.6)	8.64E-13	1.69
U5RPD9	CLAU_0399	Nitrogenase iron protein (EC 1.18.6.1) (Nitrogenase Fe protein) (Nitrogenase component II) (Nitrogenase reductase)	8.29E-03	1.43
U5RPK3	CLAU_0458	Glutamate synthase (NADPH) (EC 1.4.1.13)	1.08E-13	1.88
U5RPK9	CLAU_0463	NADPH dehydrogenase (EC 1.6.99.1)	8.52E-05	1.21
U5RTL9	CLAU_0465	Cyclase family protein	1.02E-09	1.75
U5RQB1	CLAU_0532	Alcohol dehydrogenase zinc-binding domain protein	1.44E-02	1.3
U5RTC4	CLAU_0568	Carbamoyl-phosphate synthase small chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase glutamine chain)	1.52E-02	1.46
U5RPZ5	CLAU_0569	Ornithine carbamoyltransferase (OTCase) (EC 2.1.3.3)	1.99E-08	1.54
U5RQX8	CLAU_0711	Glycerol dehydrogenase (EC 1.1.1.6)	6.99E-20	2.16
U5RQH8	CLAU_0725	Uncharacterized protein	2.36E-02	1.36
U5RTY0	CLAU_0729	Glutamine--tRNA ligase (EC 6.1.1.18) (Glutaminyl-tRNA synthetase) (GlnRS)	2.10E-06	1.24
U5RUF8	CLAU_0733	DegT/DnrJ/EryC1/StrS aminotransferase	1.50E-05	-1.37
U5RTY5	CLAU_0734	Glutamine--scyllo-inositol transaminase (EC 2.6.1.50)	2.81E-02	-1.16
U5RR17	CLAU_0752	Uncharacterized protein	1.23E-07	1.55
U5RU29	CLAU_0789	Glucarate dehydratase (EC 4.2.1.40)	3.06E-09	-1.42
U5RUL7	CLAU_0793	Methylglyoxal reductase (NADPH-dependent) (EC 1.1.1.283)	3.97E-11	-1.75
U5RQN3	CLAU_0795	Dihydrodipicolinate synthase (EC 4.3.3.7)	6.07E-09	-1.57
U5RUT7	CLAU_0865	Cardiolipin synthase (CL synthase) (EC 2.7.8.-)	1.08E-03	1.32
U5RQT1	CLAU_0873	Shikimate dehydrogenase (NADP(+)) (SDH) (EC 1.1.1.25)	4.60E-03	1.21
U5RRE0	CLAU_0879	Phospho-2-dehydro-3-deoxyheptonate aldolase (EC 2.5.1.54)	4.04E-08	-1.21
U5RUV6	CLAU_0886	UPF0210 protein CAETHG_0918	1.11E-06	1.21
U5RUG4	CLAU_0940	NADPH-dependent FMN reductase	3.06E-04	-1.91
U5RUQ1	CLAU_1022	NADP-dependent oxidoreductase domain containing protein	1.45E-02	1.52
U5RRT3	CLAU_1024	Putative methyl-accepting chemotaxis sensory transducer	2.48E-02	1.2
U5RRK2	CLAU_1071	Cob(I)alamin adenosyltransferase	1.56E-02	-1.16
U5RS07	CLAU_1078	Precorrin-2 C20-methyltransferase	1.27E-04	-1.23
U5RS12	CLAU_1083	Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase (NN:DBI PRT) (EC 2.4.2.21) (N(1)-alpha-phosphoribosyltransferase)	1.12E-04	-1.23
U5RVJ6	CLAU_1084	Cobyrinate a,c-diamide synthase (EC 6.3.5.11) (Cobyrinic acid a,c-diamide synthetase)	8.04E-03	-1.21
U5RVL0	CLAU_1099	Lipoprotein	1.26E-05	1.12
U5RRK7	CLAU_1107	Nitroreductase	2.68E-04	1.24

U5RS69	CLAU_1154	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92) (Endopeptidase Clp)	1.81E-02	1.15
U5RS79	CLAU_1164	Phosphomethylpyrimidine kinase (EC 2.7.4.7)	2.31E-02	1.77
U5RRR0	CLAU_1173	ABC-type transporter, periplasmic subunit family 3	2.32E-08	1.23
U5RS92	CLAU_1179	Threonine synthase (EC 4.2.3.1)	1.78E-02	1.07
U5RVS7	CLAU_1185	Cold-shock DNA-binding domain protein	1.99E-09	-1.25
U5RVD2	CLAU_1226	Aspartyl-tRNA synthetase (EC 6.1.1.12)	1.30E-05	-1.17
U5RS67	CLAU_1305	50S ribosomal protein L35	5.95E-03	1.48
U5RSK7	CLAU_1307	Threonine--tRNA ligase (EC 6.1.1.3) (Threonyl-tRNA synthetase) (ThrRS)	8.77E-09	-1.29
U5RW61	CLAU_1332	Phosphoglycerate dehydrogenase (EC 1.1.1.95)	5.72E-09	2.64
U5RVM6	CLAU_1333	Dimethylmenaquinone methyltransferase	8.86E-12	2.68
U5RS58	CLAU_1335	Amidohydrolase	4.02E-06	1.8
U5RVQ0	CLAU_1358	CRISPR-associated protein	3.39E-03	1.37
U5RWD0	CLAU_1413	Transcription elongation factor GreA (Transcript cleavage factor GreA)	4.15E-03	-1.2
U5RSV4	CLAU_1417	Uncharacterized protein	4.89E-08	1.42
U5RSW5	CLAU_1432	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92) (Endopeptidase Clp)	4.34E-04	1.19
U5RSK1	CLAU_1435	Uncharacterized protein	8.05E-03	1.1
U5RSG1	CLAU_1436	Orotate phosphoribosyltransferase (OPRT) (OPRTase) (EC 2.4.2.10)	3.15E-04	-1.28
U5RSK5	CLAU_1440	Aspartate carbamoyltransferase regulatory subunit	1.07E-03	-1.53
U5RSG6	CLAU_1441	Aspartate carbamoyltransferase (EC 2.1.3.2) (Aspartate transcarbamylase) (ATCase)	2.55E-06	-1.41
U5RSK0	CLAU_1477	Hemerythrin-like metal-binding protein	1.01E-02	-1.56
U5RWQ1	CLAU_1529	GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2) (GMP synthetase) (Glutamine amidotransferase)	4.42E-04	-1.13
U5RW54	CLAU_1530	Inosine-5'-monophosphate dehydrogenase (IMP dehydrogenase) (IMPD) (IMPDH) (EC 1.1.1.205)	1.95E-13	-1.16
U5RSU4	CLAU_1531	60 kDa chaperonin (GroEL protein) (Protein Cpn60)	1.41E-10	-1.13
U5RW58	CLAU_1535	Hydrogenase, Fe-only (EC 1.12.7.2)	6.67E-04	-1.36
U5RSQ6	CLAU_1537	NADH dehydrogenase (Ubiquinone) 24 kDa subunit	8.08E-05	-1.64
U5RW67	CLAU_1545	3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)	1.45E-02	-1.36
U5RT85	CLAU_1553	Pyruvate carboxylase (EC 6.4.1.1)	8.04E-07	1.24
U5RW82	CLAU_1561	Uncharacterized protein	6.65E-03	-1.41
U5RWW1	CLAU_1565	Glycine cleavage system H protein	1.23E-03	-1.18
U5RWA4	CLAU_1566	CO dehydrogenase/acetyl-CoA synthase complex, beta subunit (EC 2.3.1.169)	2.48E-47	-1.4
U5RWC1	CLAU_1571	Dihydrolipoyl dehydrogenase (EC 1.8.1.4)	1.02E-07	-1.16
U5RT18	CLAU_1572	Methylenetetrahydrofolate reductase (EC 1.5.1.20)	2.57E-04	1.11
U5RWC6	CLAU_1576	Formate--tetrahydrofolate ligase (EC 6.3.4.3) (Formyltetrahydrofolate synthetase) (FHS) (FTHFS)	1.09E-07	-1.08
U5RSY1	CLAU_1578	Carbon-monoxide dehydrogenase (Acceptor) (EC 1.2.99.2)	1.04E-13	-2.96
U5RTE2	CLAU_1579	Carbon-monoxide dehydrogenase (Acceptor) (EC 1.2.99.2)	1.34E-30	-2.26
U5RT31	CLAU_1587	Nitrite and sulfite reductase 4Fe-4S region	4.68E-04	1.4
U5RT38	CLAU_1647	Aspartokinase (EC 2.7.2.4)	1.53E-02	1.09
U5RTD3	CLAU_1710	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase)	6.94E-06	1.09
U5RTA1	CLAU_1711	2,3-bisphosphoglycerate-independent phosphoglycerate mutase (BPG-independent PGAM) (Phosphoglyceromutase) (iPGM) (EC 5.4.2.12)	2.52E-04	1.1
U5RTQ5	CLAU_1712	Triosephosphate isomerase (TIM) (TPI) (EC 5.3.1.1) (Triose-phosphate isomerase)	1.02E-04	1.11
U5RX94	CLAU_1713	Phosphoglycerate kinase (EC 2.7.2.3)	1.19E-04	1.09
U5RWQ9	CLAU_1714	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.-)	7.44E-18	1.14
U5RTE9	CLAU_1729	Cysteine synthase (EC 2.5.1.47)	4.20E-09	2
U5RTS5	CLAU_1731	Ig domain protein	2.42E-02	-1.24
U5RXB0	CLAU_1732	Rubrerythrin	1.04E-10	-1.65
U5RXD6	CLAU_1757	Pyridoxal 5'-phosphate synthase subunit PdxS (PLP synthase subunit PdxS) (EC 4.3.3.6) (Pdx1)	3.36E-10	1.19

U5RWU8	CLAU_1758	Pyridoxal 5'-phosphate synthase subunit PdxT (EC 4.3.3.6) (Pdx2) (Pyridoxal 5'-phosphate synthase glutaminase subunit) (EC 3.5.1.2)	5.81E-04	1.38
U5RTV6	CLAU_1766	Alcohol dehydrogenase (EC 1.1.1.1)	8.71E-21	2.59
U5RTI0	CLAU_1769	Microcompartments protein	1.41E-15	3.77
U5RTI0	CLAU_1769	Microcompartments protein	2.64E-16	2.59
U5RTW1	CLAU_1771	Phosphate propanoyltransferase (EC 2.3.1.222)	3.85E-05	3.36
U5RXL2	CLAU_1771	Phosphate propanoyltransferase (EC 2.3.1.222)	5.29E-09	2.53
U5RTF0	CLAU_1772	Microcompartments protein	3.31E-16	2.93
U5RXF4	CLAU_1772	Acetaldehyde dehydrogenase (Acetylating)	7.36E-04	1.96
U5RWV7	CLAU_1773	Microcompartments protein	1.57E-05	3.22
U5RTI4	CLAU_1774	Ethanolamine utilization EutQ family protein	6.79E-07	2.63
U5RTF4	CLAU_1775	Microcompartments protein	1.83E-09	2.18
U5RTW4	CLAU_1776	Respiratory-chain NADH dehydrogenase domain 51 kDa subunit	9.75E-16	2.88
U5RXP0	CLAU_1779	Ethanolamine utilization protein EutJ family protein	3.09E-06	2.25
U5RTF7	CLAU_1780	Uncharacterized protein	1.67E-02	2.46
U5RTW9	CLAU_1781	Choline trimethylamine-lyase activating enzyme (EC 1.97.1.-) (Choline utilization protein D) (GRE activase CutD) (Glycyl-radical enzyme activating enzyme CutD) (GRE activating enzyme CutD)	4.78E-03	1.67
U5RXG3	CLAU_1782	Choline trimethylamine-lyase (Choline TMA-lyase) (EC 4.3.99.4) (Choline utilization protein C)	1.29E-43	2.86
U5RWW6	CLAU_1783	Acetaldehyde dehydrogenase (Acetylating) (EC 1.2.1.10)	1.80E-13	2.48
U5RTJ3	CLAU_1784	Microcompartments protein	1.14E-08	2.56
U5RTG1	CLAU_1785	Microcompartments protein	1.17E-08	2.13
U5RTX4	CLAU_1786	Uncharacterized protein	3.46E-05	3.93
U5RWX1	CLAU_1788	Uncharacterized protein	3.21E-05	2.41
U5RTK2	CLAU_1794	Alcohol dehydrogenase (EC 1.1.1.1)	1.59E-10	-2.13
U5RU23	CLAU_1845	Thioredoxin	7.96E-03	-1.13
U5RTQ9	CLAU_1848	Phosphoenolpyruvate-protein phosphotransferase (EC 2.7.3.9) (Phosphotransferase system, enzyme I)	2.18E-04	1.72
U5RXN5	CLAU_1856	Methylaspartate ammonia-lyase (EC 4.3.1.2)	1.13E-05	-1.18
U5RX36	CLAU_1857	Glutamate mutase epsilon subunit (EC 5.4.99.1) (Glutamate mutase E chain) (Glutamate mutase large subunit) (Methylaspartate mutase)	2.06E-03	-1.13
U5RU43	CLAU_1865	30S ribosomal protein S9	4.54E-04	-1.09
U5RXP1	CLAU_1866	50S ribosomal protein L13	3.98E-07	-1.21
U5RXP5	CLAU_1871	50S ribosomal protein L17	9.09E-03	-1.14
U5RX48	CLAU_1872	DNA-directed RNA polymerase subunit alpha (RNAP subunit alpha) (EC 2.7.7.6) (RNA polymerase subunit alpha) (Transcriptase subunit alpha)	1.68E-04	-1.1
U5RTT2	CLAU_1873	30S ribosomal protein S4	1.52E-02	-1.11
U5RU54	CLAU_1875	30S ribosomal protein S13	2.71E-03	-1.16
U5RTT7	CLAU_1878	Adenylate kinase (AK) (EC 2.7.4.3) (ATP-AMP transphosphorylase) (ATP:AMP phosphotransferase) (Adenylate monophosphate kinase)	2.31E-03	-1.15
U5RU58	CLAU_1880	50S ribosomal protein L15	4.72E-05	-1.27
U5RX57	CLAU_1882	30S ribosomal protein S5	3.82E-04	-1.15
U5RTR1	CLAU_1884	50S ribosomal protein L6	1.39E-08	-1.24
U5RX62	CLAU_1887	50S ribosomal protein L5	6.20E-05	-1.13
U5RTU8	CLAU_1888	50S ribosomal protein L24	7.27E-03	-1.2
U5RTR7	CLAU_1889	50S ribosomal protein L14	9.37E-03	-1.22
U5RU69	CLAU_1890	30S ribosomal protein S17	6.46E-03	-1.27
U5RU73	CLAU_1895	30S ribosomal protein S19	5.34E-05	-1.17
U5RX69	CLAU_1897	50S ribosomal protein L23	1.53E-02	-1.14
U5RU77	CLAU_1900	30S ribosomal protein S10	2.92E-04	-1.18
U5RU93	CLAU_1901	Elongation factor Tu (EF-Tu)	3.75E-03	-1.09
U5RX73	CLAU_1902	Elongation factor G (EF-G)	4.16E-14	-1.24
U5RTW7	CLAU_1908	50S ribosomal protein L7/L12	1.44E-10	-1.23

U5RTT9	CLAU_1909	50S ribosomal protein L10	8.37E-07	-1.27
U5RU88	CLAU_1910	50S ribosomal protein L1	1.55E-03	-1.2
U5RXU4	CLAU_1926	ATPase AAA-2 domain protein	3.42E-14	-1.34
U5RX93	CLAU_1927	Protein-arginine kinase (EC 2.7.14.1)	2.97E-04	-1.58
U5RTY8	CLAU_1928	UvrB/UvrC protein	7.07E-07	-1.52
U5RX97	CLAU_1932	UDP-N-acetylmuramoylalanine--D-glutamate ligase (EC 6.3.2.9) (D-glutamic acid-adding enzyme) (UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase)	2.32E-02	1.1
U5RTZ3	CLAU_1933	Glycine--tRNA ligase (EC 6.1.1.14) (Glycyl-tRNA synthetase) (GlyRS)	3.72E-05	1.26
U5RTW6	CLAU_1934	Lysine--tRNA ligase (EC 6.1.1.6) (Lysyl-tRNA synthetase) (LysRS)	8.06E-03	-1.09
U5RUB0	CLAU_1935	Transcription elongation factor GreA (Transcript cleavage factor GreA)	1.66E-02	-1.13
U5RUB5	CLAU_1940	Hypoxanthine phosphoribosyltransferase (EC 2.4.2.8)	2.35E-02	1.3
U5RU02	CLAU_1943	RNA binding S1 domain protein	1.23E-02	-1.23
U5RXB8	CLAU_1952	Foldase protein PrsA (EC 5.2.1.8)	2.93E-04	1.13
U5RU17	CLAU_1958	Ribose-phosphate pyrophosphokinase (RPPK) (EC 2.7.6.1) (5-phospho-D-ribosyl alpha-1-diphosphate) (Phosphoribosyl diphosphate synthase) (Phosphoribosyl pyrophosphate synthase) (P-Rib-PP synthase) (PRPP synthase) (PRPPase)	1.52E-07	-1.29
U5RU03	CLAU_1974	Helix-turn-helix domain protein	1.65E-02	-1.5
U5RY03	CLAU_1976	Glutamine synthetase catalytic region	1.06E-14	1.52
U5RU44	CLAU_1982	SEC-C motif domain protein	1.28E-03	-1.25
U5RY25	CLAU_1995	Acetyl-CoA carboxylase, biotin carboxyl carrier protein	2.01E-04	-1.23
U5RY43	CLAU_2009	Adenylosuccinate synthetase (AMPSase) (AdSS) (EC 6.3.4.4) (IMP--aspartate ligase)	1.08E-02	-1.1
U5RUI7	CLAU_2018	Cupin 2 conserved barrel domain protein	1.18E-05	1.43
U5RY61	CLAU_2024	Putative agmatine deiminase (EC 3.5.3.12) (Agmatine iminohydrolase)	1.80E-04	-2.05
U5RU95	CLAU_2030	Carbamate kinase	2.53E-02	-2.15
U5RU62	CLAU_2031	Putrescine carbamoyltransferase (PTC) (PTCase) (EC 2.1.3.6) (Putrescine transcarbamoylase) (Putrescine transcarbamylase)	2.02E-03	-2.02
U5RY88	CLAU_2051	30S ribosomal protein S18	1.20E-04	-1.38
U5RUB3	CLAU_2053	30S ribosomal protein S6	1.13E-02	-1.15
U5RU85	CLAU_2054	Uncharacterized protein	1.43E-02	1.34
U5RYF0	CLAU_2081	UPF0182 protein CAETHG_2133	8.62E-04	1.17
U5RYK0	CLAU_2147	Nucleoid-associated protein CAETHG_2200	2.00E-03	1.26
U5RYK5	CLAU_2157	Alanine--glyoxylate transaminase (EC 2.6.1.44)	2.03E-04	-1.22
U5RY08	CLAU_2158	Phosphoglycerate dehydrogenase (EC 1.1.1.95)	1.06E-11	-2.8
U5RV15	CLAU_2172	Peptidylprolyl isomerase (EC 5.2.1.8)	9.60E-04	-1.26
U5RYN7	CLAU_2202	Transcriptional regulator, AbrB family	3.40E-05	-1.24
U5RUS5	CLAU_2204	Uncharacterized protein	1.49E-02	1.51
U5RUP7	CLAU_2205	Phage shock protein A, PspA	2.76E-14	1.3
U5RUQ8	CLAU_2220	Methionine--tRNA ligase (EC 6.1.1.10) (Methionyl-tRNA synthetase) (MetRS)	7.56E-03	-1.15
U5RYR3	CLAU_2231	Anaerobic ribonucleoside-triphosphate reductase (EC 1.17.4.2)	1.75E-02	-1.2
U5RUU9	CLAU_2267	CTP synthase (EC 6.3.4.2) (Cytidine 5'-triphosphate synthase) (Cytidine triphosphate synthetase) (CTP synthetase) (CTPS) (UTP--ammonia ligase)	5.04E-03	-1.16
U5RYB0	CLAU_2270	50S ribosomal protein L31	8.30E-03	-1.21
U5RYX1	CLAU_2289	ATP synthase subunit alpha (EC 3.6.3.14) (ATP synthase F1 sector subunit alpha) (F-ATPase subunit alpha)	1.36E-06	1.15
U5RV03	CLAU_2291	ATP synthase subunit beta (EC 3.6.3.14) (ATP synthase F1 sector subunit beta) (F-ATPase subunit beta)	2.52E-03	1.07
U5RYD5	CLAU_2300	S-adenosylmethionine synthase (AdoMet synthase) (EC 2.5.1.6) (MAT) (Methionine adenosyltransferase)	2.74E-07	1.43
U5RV19	CLAU_2305	Protein translocase subunit SecA	1.65E-03	-1.15
U5RYZ2	CLAU_2308	Glutamate dehydrogenase	1.21E-07	-1.27
U5RZ09	CLAU_2322	Fructose-1,6-bisphosphate aldolase, class II (EC 4.1.2.13)	2.33E-05	-1.14
U5RZ41	CLAU_2360	Cell division ATP-binding protein FtsE (EC 3.6.3.28)	2.81E-02	1.24

U5RVL5	CLAU_2378	Pyruvate kinase (EC 2.7.1.40)	3.27E-16	1.26
U5RZ70	CLAU_2394	Methyl-accepting chemotaxis sensory transducer	1.09E-03	1.15
U5RVP0	CLAU_2404	Nitrogen regulatory protein P-II	8.36E-04	-1.47
U5RYS5	CLAU_2445	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain)	2.17E-02	1.16
U5RVG8	CLAU_2492	Cell wall binding repeat 2-containing protein	8.50E-113	1.36
U5RVI8	CLAU_2512	Cell wall binding repeat 2-containing protein	2.27E-03	1.58
U5RZK5	CLAU_2519	Outer membrane efflux protein	7.94E-03	1.18
U5RZ08	CLAU_2520	ResB family protein	1.11E-02	1.22
U5RW09	CLAU_2523	Cell wall binding repeat 2-containing protein	5.52E-49	1.48
U5RVZ8	CLAU_2590	Tubulin-like protein	3.07E-06	1.36
U5RW20	CLAU_2613	D-lactate dehydrogenase (Cytochrome) (EC 1.1.2.4)	1.90E-02	1.36
U5RW30	CLAU_2623	Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde dehydrogenase) (GSA dehydrogenase)	1.20E-20	1.33
U5RZI2	CLAU_2641	Chaperone protein ClpB	1.57E-04	1.17
U5RW10	CLAU_2653	Dihydropteroate synthase (DHPS) (EC 2.5.1.15) (Dihydropteroate pyrophosphorylase)	3.74E-05	-1.4
U5RW34	CLAU_2677	O-acetylhomoserine/O-acetylserine sulfhydrylase (EC 2.5.1.47)	1.38E-26	1.29
U5RZL7	CLAU_2680	ABC-type transporter, periplasmic subunit family 3	7.95E-06	1.16
U5RWL5	CLAU_2683	Argininosuccinate synthase	4.86E-14	1.63
U5S045	CLAU_2683	Argininosuccinate synthase (EC 6.3.4.5)	1.01E-02	1.27
U5RW42	CLAU_2686	Transcriptional regulator, AsnC family	1.19E-02	-1.83
U5RW59	CLAU_2707	Band 7 protein	3.77E-03	1.2
U5RWP5	CLAU_2713	Formate dehydrogenase, alpha subunit (EC 1.7.99.4)	6.46E-04	-1.38
S5ZJ59	CLAU_2717	Electron bifurcating FeFe-hydrogenase dependent on TPN subunit C (NADH-quinone oxidoreductase, E subunit) (EC 1.6.99.5)	1.02E-05	-1.46
S5Z7X0	CLAU_2719	4Fe-4S ferredoxin, iron-sulpur binding domain-containing protein (Electron bifurcating FeFe-hydrogenase dependent on TPN subunit D)	6.38E-04	-1.27
U5RWA6	CLAU_2721	Hydrogenase, Fe-only (EC 1.12.7.2)	1.16E-08	-1.29
U5RW72	CLAU_2722	4Fe-4S ferredoxin, iron-sulpur binding domain-containing protein	4.02E-03	-1.3
U5S087	CLAU_2728	UPF0735 ACT domain-containing protein CAETHG_2806	1.38E-04	1.33
U5RW84	CLAU_2736	Cell shape determining protein, MreB/Mrl family	2.91E-03	1.16
U5RWD3	CLAU_2750	50S ribosomal protein L21	3.62E-05	-1.39
U5RWT5	CLAU_2752	50S ribosomal protein L27	4.63E-03	-1.12
U5RW98	CLAU_2756	Endoribonuclease L-PSP	5.86E-05	1.25
U5RWF3	CLAU_2764	Dihydropteroate synthase DHPS	2.43E-03	-1.44
U5RWA7	CLAU_2765	Uncharacterized protein	4.84E-03	-1.92
U5S0E6	CLAU_2788	Hydroxylamine reductase (EC 1.7.99.1) (Hybrid-cluster protein) (HCP) (Prismane protein)	9.27E-09	1.23
U5S0G1	CLAU_2804	30S ribosomal protein S20	4.89E-03	-1.11
U5RWJ5	CLAU_2811	Protein GrpE (HSP-70 cofactor)	5.98E-04	-1.24
U5RZY4	CLAU_2819	Uncharacterized protein	3.51E-03	-1.13
U5RWJ7	CLAU_2855	PDZ/DHR/GLGF domain protein	6.29E-04	1.42
U5RWK6	CLAU_2865	Hydrolase	1.21E-03	1.09
U5S0L4	CLAU_2867	N5-carboxyaminoimidazole ribonucleotide mutase (N5-CAIR mutase) (EC 5.4.99.18) (5-(carboxyamino)imidazole ribonucleotide mutase)	1.16E-02	-1.39
U5RWQ3	CLAU_2869	Amidophosphoribosyltransferase (ATase) (EC 2.4.2.14) (Glutamine phosphoribosylpyrophosphate amidotransferase) (GPATase)	7.41E-03	-1.17
U5RWS4	CLAU_2889	Uncharacterized protein	4.80E-03	1.23
U5S0R8	CLAU_2924	Carbon monoxide dehydrogenase (EC 1.2.7.4)	2.50E-05	-1.72
U5RXA5	CLAU_2947	Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)	1.15E-39	1.51
U5RXB1	CLAU_2952	CheW protein	1.32E-02	-1.17
U5S0V6	CLAU_2958	Response regulator receiver protein	1.49E-02	1.16
U5RXI1	CLAU_3017	Homoserine dehydrogenase (EC 1.1.1.3)	9.87E-04	1.21

U5RXB4	CLAU_3055	Putative signal transduction protein with CBS domain containing protein (EC 3.6.1.1)	2.55E-03	1.14
U5S1B3	CLAU_3078	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)	7.86E-03	1.36
U5RXJ6	CLAU_3144	Electron transport complex subunit C	9.19E-13	1.65
U5RXZ2	CLAU_3146	Electron transport complex subunit G	1.98E-04	1.5
U5RXK0	CLAU_3149	Electron transport complex, RnfABCDGE type, B subunit	1.96E-05	1.38
U5S1J7	CLAU_3187	Uncharacterized protein	3.41E-05	1.38
U5RXN3	CLAU_3189	Histidinol phosphate phosphatase HisJ family (EC 3.1.3.15)	1.13E-02	1.2
U5S1K9	CLAU_3202	Ethanolamine utilization EutQ family protein	1.57E-05	2.42
U5RXF4	CLAU_3204	Acetaldehyde dehydrogenase (Acetylating) (EC 1.2.1.10)	8.98E-14	2.51
U5RXL6	CLAU_3210	Cysteine desulfurase IscS (EC 2.8.1.7)	1.67E-03	-1.2
U5RY52	CLAU_3211	FeS cluster assembly scaffold protein NifU	1.28E-03	-1.23
U5RXL9	CLAU_3215	UPF0297 protein CAETHG_3298	7.22E-05	-2.04
U5RY95	CLAU_3256	Methionyl-tRNA formyltransferase (EC 2.1.2.9)	6.44E-03	1.39
U5RXP7	CLAU_3260	Protein serine/threonine phosphatase	9.09E-04	1.31
U5RXQ1	CLAU_3265	50S ribosomal protein L28	9.73E-03	-1.17
U5RXR0	CLAU_3274	Phosphate acetyltransferase (EC 2.3.1.8)	3.01E-03	-1.09
U5RYD2	CLAU_3300	GTP-sensing transcriptional pleiotropic repressor CodY	6.07E-03	-1.17
U5S1B8	CLAU_3302	Elongation factor Ts (EF-Ts)	2.14E-04	-1.12
U5S1W2	CLAU_3311	Transcription termination/antitermination protein NusA	1.16E-02	-1.14
U5RXV2	CLAU_3319	30S ribosomal protein S15	6.54E-04	-1.21
U5RYE8	CLAU_3320	Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) (Polynucleotide phosphorylase) (PNPase)	8.05E-07	-1.24
U5RY13	CLAU_3333	Aminotransferase (EC 2.6.1.-)	1.09E-03	-1.11
U5RXW4	CLAU_3334	Phosphotransferase system, phosphocarrier protein HPr	4.90E-08	1.5
U5RYK8	CLAU_3384	Methyl-accepting chemotaxis sensory transducer	3.75E-05	-1.25
U5RYA1	CLAU_3416	Cupin 2 conserved barrel domain protein	6.61E-03	1.28
U5S1Y5	CLAU_3521	Threonine dehydratase	1.08E-08	1.51
U5RYX9	CLAU_3612	Glutamine amidotransferase of anthranilate synthase (EC 4.1.3.27)	8.98E-04	-1.23
U5RYS2	CLAU_3613	Anthranilate phosphoribosyltransferase (EC 2.4.2.18)	1.98E-04	-1.25
U5RZB6	CLAU_3614	Indole-3-glycerol phosphate synthase (IGPS) (EC 4.1.1.48)	6.99E-04	-1.37
U5S2S9	CLAU_3615	N-(5'-phosphoribosyl)anthranilate isomerase (PRAI) (EC 5.3.1.24)	5.10E-03	-1.26
U5S2B5	CLAU_3616	Tryptophan synthase beta chain (EC 4.2.1.20)	8.84E-07	-1.3
U5S2V9	CLAU_3655	Aldehyde-alcohol dehydrogenase	1.13E-12	-2.03
U5S2V9	CLAU_3655	Aldehyde-alcohol dehydrogenase	4.17E-20	-2.33
U5S2J6	CLAU_3721	Rubrerythrin	1.17E-04	1.28
U5RZ45	CLAU_3758	Glutamate synthase (Ferredoxin) (EC 1.4.7.1)	1.03E-02	-1.09
U5S336	CLAU_3760	Dipeptidase	9.73E-11	1.37
U5RZT6	CLAU_3809	Microcompartments protein	3.19E-04	1.35
U5RZF0	CLAU_3821	4-hydroxy-tetrahydrodipicolinate reductase (HTPA reductase) (EC 1.17.1.8)	1.63E-04	1.77
U5RZV4	CLAU_3828	Cytidine deaminase (EC 3.5.4.5)	1.42E-09	-1.18
U5S394	CLAU_3829	Deoxyribose-phosphate aldolase (DERA) (EC 4.1.2.4) (2-deoxy-D-ribose 5-phosphate aldolase) (Phosphodeoxyriboaldolase) (Deoxyriboaldolase)	6.94E-08	-1.21
U5RZA4	CLAU_3832	Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)	1.27E-56	-1.3
U5RZH4	CLAU_3846	Taurine-transporting ATPase (EC 3.6.3.36)	2.72E-05	1.45
U5RZB9	CLAU_3847	ABC-type transporter, periplasmic subunit family 3	3.32E-11	1.77
U5RZX4	CLAU_3848	Uncharacterized protein	7.35E-12	2.17
U5RZJ2	CLAU_3861	Alcohol dehydrogenase (EC 1.1.1.1)	7.72E-15	1.33

Bibliography

"Manuscript in preparation" refers to manuscripts in the drafting/reviewing stages at the time of writing this project.

A Münchau, K. P. B., Münchau, A. and Bhatia, K. P. (2000) 'Uses of botulinum toxin injection in medicine today', *BMJ : British Medical Journal*. British Medical Journal, 320(7228), pp. 161–165.

Abrini, J., Naveau, H. and Nyns, E. (1994) 'Clostridium autoethanogenumt, sp. nov., an anaerobic bacterium that produces ethanol from carbon monoxide', *Archives of Microbiology*, 161(4), pp. 345–351. doi: 10.1007/bf00303591.

Abubackar, H. N., Veiga, M. C. and Kennes, C. (2011) 'Biological conversion of carbon monoxide: rich syngas or waste gases to bioethanol', *Biofuels, Bioproducts and Biorefining*, 5(1), pp. 93–114. doi: 10.1002/bbb.256.

Adamse, A. D. (1980) 'New isolation of Clostridium aceticum (Wieringa)', *Antonie van Leeuwenhoek*. Springer, 46(6), pp. 523–531.

Al-Hinai, M. A., Fast, A. G. and 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. 8112–8121. doi: 10.1128/AEM.02214-12.

Al-Hinai, M. A., Jones, S. W. and Papoutsakis, E. T. (2015) 'The Clostridium sporulation programs: diversity and preservation of endospore differentiation.', *Microbiology and molecular biology reviews : MMBR*, 79(1), pp. 19–37. doi: 10.1128/MMBR.00025-14.

Albers, S. C., Berklund, A. M. and Graff, G. D. (2016) 'The rise and fall of innovation in biofuels', *Nature Biotechnology*. Nature Publishing Group, 34(8), pp. 814–821. doi: 10.1038/nbt.3644.

Allen, T. D. *et al.* (2010) 'Alkalibaculum bacchi gen. nov., sp. nov., a CO-oxidizing, ethanol-producing acetogen isolated from livestock-impacted soil', *International journal of systematic and evolutionary microbiology*. Microbiology Society, 60(10), pp. 2483–2489.

Alsaker, K. V *et al.* (2004) 'Transcriptional Analysis of spo0A Overexpression in Clostridium acetobutylicum and Its Effect on the Cell â€™ s Response to Butanol Stress Transcriptional Analysis of spo0A Overexpression in Clostridium acetobutylicum and Its Effect on the Cell ' s Response', *Journal of bacteriology*, 186(7), pp. 1959–1971. doi: 10.1128/JB.186.7.1959.

Altaras, N. E. and Cameron, D. C. (1999) 'Metabolic Engineering of a 1, 2-Propanediol Pathway in Escherichia coli Metabolic Engineering of a 1,2-Propanediol Pathway in Escherichia coli', *Appl. Environ. Microbiol.*, 65(3), pp. 1180–1185.

Andreesen, J. R., Gottschalk, G. and Schlegel, H. G. (1970) 'Clostridium formicoaceticum nov. spec. Isolation, description and distinction from C. aceticum and C. thermoaceticum', *Archiv für Mikrobiologie*, 72(2), pp. 154–174. doi: 10.1007/BF00409521.

Antunes, A., Martin-Verstraete, I. and Dupuy, B. (2011) 'CcpA-mediated repression of Clostridium difficile toxin gene expression', *Molecular Microbiology*, 79(4), pp. 882–899. doi: 10.1111/j.1365-2958.2010.07495.x.

Antunes, L. C. M. *et al.* (2010) 'Quorum sensing in bacterial virulence', *Microbiology*, 156(8), pp. 2271–2282. doi: 10.1099/mic.0.038794-0.

Atkinson, M. R. *et al.* (2003) 'Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli', *Cell*, 113(5), pp. 597–607. doi: 10.1016/S0092-8674(03)00346-5.

Atkinson, S. and Williams, P. (2009) 'Quorum sensing and social networking in the microbial world.', *Journal of the Royal Society, Interface / the Royal Society*, 6(40), pp. 959–78. doi: 10.1098/rsif.2009.0203.

Bacchus, W. and Fussenegger, M. (2013) 'Engineering of synthetic intercellular communication systems', *Metabolic Engineering*. Elsevier, 16(1), pp. 33–41. doi: 10.1016/j.ymben.2012.12.001.

Baneyx, F. (1999) 'Recombinant protein expression in Escherichia coli', *Current Opinion in Biotechnology*, pp. 411–421. doi: 10.1016/S0958-1669(99)00003-8.

Bannam, T. L. and Rood, J. I. (1993) 'Clostridium perfringens-Escherichia coli Shuttle Vectors That Carry Single Antibiotic Resistance Determinants', *Plasmid*, pp. 233–235. doi: 10.1006/plas.1993.1025.

Bardi, U. (2009) 'Peak oil: The four stages of a new idea', *Energy*, 34(3), pp. 323–326. doi: 10.1016/j.energy.2008.08.015.

Bassler, B. L. (1999) 'How bacteria talk to each other: regulation of gene expression by quorum sensing.', *Current opinion in microbiology*, 2(6), pp. 582–7. doi: 10.1016/S1369-5274(99)00025-9.

Basu, S. *et al.* (2004) 'Spatiotemporal control of gene expression with pulse-generating networks.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(17), pp. 6355–60. doi: 10.1073/pnas.0307571101.

Bean, Z. (2017) 'Role of RNPP and Agr quorum sensing in the regulation of stationary phase phenotypes in Clostridium acetobutylicum'.

Behera, S. *et al.* (2014) 'Scope of algae as third generation biofuels.', *Frontiers in bioengineering and biotechnology*, 2(February), p. 90. doi: 10.3389/fbioe.2014.00090.

Di Bella, S. *et al.* (2016) 'Clostridium difficile toxins A and B: Insights into pathogenic properties and extraintestinal effects', *Toxins*, 8(5), pp. 1–25. doi: 10.3390/toxins8050134.

Bengelsdorf, F. R. *et al.* (2016) 'Industrial acetogenic biocatalysts: A comparative metabolic and genomic analysis', *Frontiers in Microbiology*, 7(JUL), pp. 1–15. doi: 10.3389/fmicb.2016.01036.

Berto, P. *et al.* (2011) 'The cyanobacterium Synechocystis sp. PCC 6803 is able to express an active [FeFe]-hydrogenase without additional maturation proteins', *Biochemical and Biophysical Research Communications*, 405(4), pp. 678–683. doi: 10.1016/j.bbrc.2011.01.095.

Berzin, V., Kiriukhin, M. and Tyurin, M. (2012) 'Elimination of Acetate Production to Improve Ethanol Yield During Continuous Synthesis Gas Fermentation by Engineered Biocatalyst <i>Clostridium sp. MTEtOH550', *Applied Biochemistry and Biotechnology*, 167, pp. 338–347. doi: 10.1007/s12010-012-9697-5.

Betts, M. J. and Russell, R. B. (2007) 'Amino-Acid Properties and Consequences of Substitutions', *Bioinformatics for Geneticists: A Bioinformatics Primer for the Analysis of Genetic Data: Second Edition*, 4, pp. 311–342. doi: 10.1002/9780470059180.ch13.

Biegel, E. *et al.* (2011) 'Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes', *Cellular and Molecular Life Sciences*, 68(4), pp. 613–634. doi: 10.1007/s00018-010-0555-8.

Boeke, J. D. *et al.* (1987) '[10] 5-Fluoroorotic acid as a selective agent in yeast molecular genetics', *Methods in Enzymology*, 154(C), pp. 164–175. doi: 10.1016/0076-6879(87)54076-9.

Bouillaut, L., McBride, S. M. and Sorg, J. A. (2011) 'Genetic Manipulation of Clostridium difficile', in *Current Protocols in Microbiology*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 1–20. doi: 10.1002/9780471729259.mc09a02s20.

Brenner, K., You, L. and Arnold, F. H. (2008) 'Engineering microbial consortia: a new frontier in synthetic biology', *Trends in Biotechnology*, 26(9), pp. 483–489. doi: 10.1016/j.tibtech.2008.05.004.

Bronesky, D. *et al.* (2016) 'Staphylococcus aureus RNAIII and Its Regulon Link Quorum Sensing, Stress Responses, Metabolic Adaptation, and Regulation of Virulence Gene Expression', *Annual Review of Microbiology*, 70(1), pp. 299–316. doi: 10.1146/annurev-micro-102215-095708.

Brown, R. C. and Brown, T. R. (2014) *Biorenewable Resources: Engineering New Products from Agriculture: Second Edition*, *Wiley*. doi: 10.1002/9781118524985.

Brown, S. D. *et al.* (2014) 'Comparison of single-molecule sequencing and hybrid approaches for finishing the genome of Clostridium autoethanogenum and analysis of CRISPR systems in industrial relevant Clostridia.', *Biotechnology for biofuels*, 7, p. 40. doi: 10.1186/1754-6834-7-40.

Bruant, G. *et al.* (2010) 'Genomic analysis of carbon monoxide utilization and butanol production by clostridium carboxidivorans strain P7T', *PLoS ONE*, 5(9), pp. 1–12. doi: 10.1371/journal.pone.0013033.

Brückner, R. and Titgemeyer, F. (2002) 'Carbon catabolite repression in bacteria: Choice of the carbon source and autoregulatory limitation of sugar utilization', *FEMS Microbiology Letters*, 209(2), pp. 141–148. doi: 10.1016/S0378-1097(02)00559-1.

Cannon, G. C. *et al.* (2001) 'Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra Microcompartments in Prokaryotes: Carboxysomes a', *Applied and Environmental Microbiology*, 67(12), pp. 5351–5361. doi: 10.1128/AEM.67.12.5351.

Carter, G. P. *et al.* (2005) 'Quorum sensing in Clostridium difficile: Analysis of a luxS-type signalling system', *Journal of Medical Microbiology*, 54(2), pp. 119–127. doi: 10.1099/jmm.0.45817-0.

Carter, K. K., Valdes, J. J. and Bentley, W. E. (2012) 'Pathway engineering via quorum sensing and sRNA riboregulators-Interconnected networks and controllers', *Metabolic Engineering*. Elsevier, 14(3), pp. 281–288. doi: 10.1016/j.ymben.2011.11.006.

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. doi: 10.1128/AEM.00249-12.

Cartman, S. T. and Minton, N. P. (2010) 'A mariner-Based transposon system for in vivo random mutagenesis of clostridium difficile', *Applied and Environmental Microbiology*, 76(4), pp. 1103–1109. doi: 10.1128/AEM.02525-09.

Chamberlain, N. R. and Imanoel, B. (1996) 'Genetic regulation of fatty acid modifying enzyme from Staphylococcus aureus', *Journal of Medical Microbiology*, 44(2), pp. 125–129.

Chen, J. *et al.* (2003) 'Infrared Studies of Carbon Monoxide Binding to Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase from Moorella thermoacetica', *Biochemistry*, 42(50), pp. 14822–14830. doi: 10.1021/bi0349470.

Chen, W. Y., Liew, F. and Koepke, M. (2013) 'Recombinant microorganisms and uses therefor (US20130323820 A1)'. Google Patents.

Cheng, S. *et al.* (2012) 'The PduQ Enzyme Is an Alcohol Dehydrogenase Used to Recycle NAD+ Internally within the Pdu Microcompartment of Salmonella enterica', *PLoS ONE*, 7(10), pp. 1–11. doi: 10.1371/journal.pone.0047144.

Cook, L. C. and Federle, M. J. (2014) 'Peptide pheromone signaling in Streptococcus and Enterococcus', *FEMS Microbiology Reviews*, 38(3), pp. 473–492. doi: 10.1111/1574-6976.12046.

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. 4448–4460. doi: 10.1128/AEM.03038-09.

Cooksley, C. M. *et al.* (2012) 'Targeted mutagenesis of the Clostridium acetobutylicum acetone–butanol–ethanol fermentation pathway', *Metabolic Engineering*. Elsevier, 14(6), pp. 630–641. doi: 10.1016/j.ymben.2012.09.001.

Cornforth, D. M. *et al.* (2014) 'Combinatorial quorum sensing allows bacteria to resolve their social and physical environment.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(11), pp. 4280–4284. doi: 10.1073/pnas.1319175111.

Cornforth, D. M. and Foster, K. R. (2013) 'Competition sensing: the social side of bacterial stress responses.', *Nature reviews. Microbiology*. Nature Publishing Group, 11(4), pp. 285–93. doi: 10.1038/nrmicro2977.

Cotter, J. L., Chinn, M. S. and Grunden, A. M. (2009) 'Ethanol and acetate production by Clostridium ljungdahlii and Clostridium autoethanogenum using resting cells', *Bioprocess and Biosystems Engineering*, 32(3), pp. 369–380. doi: 10.1007/s00449-008-0256-y.

Croux, C. *et al.* (2016) 'Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer Clostridium acetobutylicum.', *Biotechnology for biofuels*. BioMed Central, 9, p. 23. doi: 10.1186/s13068-016-0432-2.

Dandekar, A. A., Chugani, S. and Greenberg, E. P. (2012) 'Bacterial Quorum Sensing and Metabolic Incentives to Cooperate', *Science*, 338(6104), pp. 264–266. doi: 10.1126/science.1227289.

Daniel, R. *et al.* (2013) 'Synthetic analog computation in living cells', *Nature*. Nature Publishing Group, 497(7451), pp. 619–623. doi: 10.1038/nature12148.

Daniell, J., Köpke, M. and Simpson, S. D. (2012) *Commercial biomass syngas fermentation*, *Energies*. doi: 10.3390/en5125372.

Danino, T. *et al.* (2010) 'A synchronized quorum of genetic clocks', *Nature*, 463(7279), pp. 326–330. doi: 10.1038/nature08753.

Dapa, T. *et al.* (2013) 'Multiple factors modulate biofilm formation by the anaerobic pathogen Clostridium difficile', *Journal of Bacteriology*, 195(3), pp. 545–555. doi: 10.1128/JB.01980-12.

Darkoh, C. *et al.* (2015) 'Toxin synthesis by Clostridium difficile is regulated through quorum signaling', *mBio*, 6(2), pp. 1–10. doi: 10.1128/mBio.02569-14.

Davis, T. O. *et al.* (2000) 'Development of a transformation and gene reporter system for group II, non-proteolytic Clostridium botulinum type B strains.', *Journal of molecular microbiology and biotechnology*, 2(1), pp. 59–69.

Dawson, L. F. *et al.* (2012) 'Characterisation of Clostridium difficile Biofilm Formation, a Role for Spo0A', *PLoS ONE*, 7(12). doi: 10.1371/journal.pone.0050527.

Debabov, V. G. (2013) 'Bioethanol from synthesis gas', *Applied Biochemistry and Microbiology*, 49(7), pp. 619–628. doi: 10.1134/S000368381307003X.

Demirbas, A. (2007) 'Progress and recent trends in biofuels', *Progress in Energy and Combustion Science*, 33(1), pp. 1–18. doi: 10.1016/j.pecs.2006.06.001.

Desai, R. P. and Papoutsakis, E. T. (1999) 'Antisense RNA strategies for metabolic engineering of Clostridium acetobutylicum', *Applied and Environmental Microbiology*, 65(3), pp. 936–945.

Diekert, G. and Wohlfarth, G. (1994) 'Metabolism of homoacetogens', *Antonie van Leeuwenhoek*, 66(1–3), pp. 209–221. doi: 10.1007/BF00871640.

Dineen, S. S. *et al.* (2007) 'Repression of Clostridium difficile toxin gene expression by CodY', *Molecular Microbiology*, 66(1), pp. 206–219. doi: 10.1111/j.1365-2958.2007.05906.x.

Dong, H. *et al.* (2010) 'Engineering Clostridium strain to accept unmethylated DNA', *PLoS ONE*, 5(2), pp. 1–8. doi: 10.1371/journal.pone.0009038.

Doukov, T. I. (2002) 'A Ni-Fe-Cu Center in a Bifunctional Carbon Monoxide Dehydrogenase/ Acetyl-CoA Synthase', *Science*, 298(5593), pp. 567–572. doi: 10.1126/science.1075843.

Doukov, T. I. *et al.* (2008) 'Xenon in and at the end of the tunnel of bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase', *Biochemistry*, 47(11), pp. 3474–3483. doi: 10.1021/bi702386t.

Dragone, G. *et al.* (2010) 'Third generation biofuels from microalgae', *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, pp. 1355–1366. doi: 10.1016/j.apenergy.2011.03.012.

Drake, H. L., Gößner, A. S. and Daniel, S. L. (2008) 'Old Acetogens, New Light', *Annals of the New York Academy of Sciences*, 1125(1), pp. 100–128. doi: 10.1196/annals.1419.016.

Drake, H. L., Kusel, K. and Matthies, C. (2006) *Acetogenic Prokaryotes, Vol. 2*, *The Prokaryotes*.

Dry, M. E. (2002) 'The Fischer-Tropsch process: 1950-2000', *Catalysis Today*, 71(3–4), pp. 227–241. doi: 10.1016/S0920-5861(01)00453-9.

Duane, W. and Hastings, J. W. (1975) 'Flavin mononucleotide reductase of luminous bacteria', *Molecular and Cellular Biochemistry*, 6(1), pp. 53–64. doi: 10.1007/BF01731866.

Dürre, P. (2005) *Handbook on clostridia*. CRC press.

Dürre, P. (2007) 'Biobutanol: An attractive biofuel', *Biotechnology Journal*, 2(12), pp. 1525–1534. doi: 10.1002/biot.200700168.

Dürre, P. (2017) 'Gas fermentation - a biotechnological solution for today's challenges', *Microbial Biotechnology*, 10(1), pp. 14–16. doi: 10.1111/1751-7915.12431.

Dürre, P. and Hollergschwandner, C. (2004) 'Initiation of endospore formation in Clostridium acetobutylicum', *Anaerobe*, 10(2), pp. 69–74. doi: 10.1016/j.anaerobe.2003.11.001.

Edgar, R. C. (2004) 'MUSCLE: Multiple sequence alignment with high accuracy and high throughput', *Nucleic Acids Research*, 32(5), pp. 1792–1797. doi: 10.1093/nar/gkh340.

Edwards, A. N., Tamayo, R. and McBride, S. M. (2016) 'A novel regulator controls Clostridium difficile sporulation, motility and toxin production', *Molecular Microbiology*, 100(6), pp. 954–971. doi: 10.1111/mmi.13361.

Ehsaan, M. *et al.* (2016) 'Mutant generation by allelic exchange and genome resequencing of the biobutanol organism Clostridium acetobutylicum ATCC 824', *Biotechnology for Biofuels*. BioMed Central, 9(1), p. 4. doi: 10.1186/s13068-015-0410-0.

Ehsaan, M., Kuehne, S. A. and Minton, N. P. (2016) 'Clostridium difficile genome editing using pyrE alleles', in *Methods in Molecular Biology*, pp. 35–52. doi: 10.1007/978-1-4939-6361-4_4.

Ellerman, A. D. and Buchner, B. K. (2007) 'The European Union Emissions Trading Scheme: Origins, Allocation, and Early Results', *Review of Environmental Economics and Policy*, 1(1), pp. 66–87. doi: 10.1093/reep/rem003.

Engebrecht, J., Nealson, K. and Silverman, M. (1983) 'Bacterial bioluminescence: Isolation and genetic analysis of functions from Vibrio fischeri', *Cell*, 32(3), pp. 773–781. doi: 10.1016/0092-8674(83)90063-6.

Englund, E. *et al.* (2015) 'Metabolic Engineering of Synechocystis sp. PCC 6803 for Production of the Plant Diterpenoid Manoyl Oxide', *ACS Synthetic Biology*, 4(12), pp. 1270–1278. doi: 10.1021/acssynbio.5b00070.

Erbes DL, B. R. (1978) 'The kinetics of methyl viologen oxidation and reduction by the hydrogenase from Clostridium pasteurianum.', *Biochim Biophys Acta.*, 525, pp. 45–54.

Fang, H. H. P., Zhu, H. and Zhang, T. (2006) 'Phototrophic hydrogen production from glucose by pure and co-cultures of Clostridium butyricum and Rhodobacter sphaeroides', *International Journal of Hydrogen Energy*, 31(15), pp. 2223–2230. doi: 10.1016/j.ijhydene.2006.03.005.

Foster, P. L. *et al.* (2015) 'Determinants of spontaneous mutation in the bacterium Escherichia coli as revealed by whole-genome sequencing.', *Proceedings of the National Academy of Sciences of the United States of America*, 112(44), pp. E5990-9. doi: 10.1073/pnas.1512136112.

Fuchs, G. (1986) 'CO2 fixation in acetogenic bacteria: Variations on a theme', *FEMS Microbiology Letters*, 39(3), pp. 181–213. doi: 10.1016/0378-1097(86)90446-5.

Furdui, C. and Ragsdale, S. W. (2000) 'The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood-Ljungdahl pathway', *Journal of Biological Chemistry*, 275(37), pp. 28494–28499. doi: 10.1074/jbc.M003291200.

Garsin, D. (2010) 'Ethanolamine Utilisation in Bacterial Pathogens: Roles and Regulation', *Nature Reviews Microbiology*, 8(4), pp. 290–295. doi: 10.1038/nrmicro2334.Ethanolamine.

Geider, R. J. *et al.* (2001) 'Primary productivity of planet earth: biological determinants and physical constraints in terrestrial and aquatic habitats', *Global Change Biology*, 7(8), pp. 849–882. doi: 10.1046/j.1365-2486.2001.00448.x.

Genthner, B. R. S. and Bryant, M. P. (1987) 'Additional characteristics of one-carbon-compound utilization by Eubacterium limosum and Acetobacterium woodii', *Applied and Environmental Microbiology*, 53(3), pp. 471–476.

Ghim, S. Y., Nielsen, P. and Neuhard, J. (1994) 'Molecular characterization of pyrimidine biosynthesis genes from the thermophile Bacillus caldolyticus', *Microbiology*, 140(3), pp. 479–491.

Gimpel, J. A. *et al.* (2013) 'Advances in microalgae engineering and synthetic biology applications for biofuel production', *Current Opinion in Chemical Biology*. Elsevier Ltd, 17(3), pp. 489–495. doi: 10.1016/j.cbpa.2013.03.038.

Gray, B., Hall, P. and Gresham, H. (2013) 'Targeting agr- and agr-Like quorum sensing systems for

development of common therapeutics to treat multiple gram-positive bacterial infections.', *Sensors (Basel, Switzerland)*, 13(4), pp. 5130–5166. doi: 10.3390/s130405130.

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. doi: 10.1099/13500872-142-8-2079.

Green, E. M. (2011) 'Fermentative production of butanol-the industrial perspective', *Current Opinion in Biotechnology*, 22(3), pp. 337–343. doi: 10.1016/j.copbio.2011.02.004.

Grosse-Honebrink, A. (2017) 'Forward and Reverse Genetics in Industrially Important Clostridia'.

Gu, Y. *et al.* (2009) 'Improvement of xylose utilization in Clostridium acetobutylicum via expression of the talA gene encoding transaldolase from Escherichia coli', *Journal of Biotechnology*, 143(4), pp. 284–287. doi: 10.1016/j.jbiotec.2009.08.009.

Guo, Y. *et al.* (2010) 'Medium optimization for ethanol production with Clostridium autoethanogenum with carbon monoxide as sole carbon source', *Bioresource Technology*. Elsevier Ltd, 101(22), pp. 8784–8789. doi: 10.1016/j.biortech.2010.06.072.

Hanahan, D., Jessee, J. and Bloom, F. R. (1991) 'Bacterial Genetic Systems', *Methods in Enzymology*, 204(1970), pp. 63–113. doi: 10.1016/0076-6879(91)04006-A.

Harun, R. *et al.* (2010) 'Bioprocess engineering of microalgae to produce a variety of consumer products', *Renewable and Sustainable Energy Reviews*, 14(3), pp. 1037–1047. doi: 10.1016/j.rser.2009.11.004.

Hatheway, C. L. (1990) 'Toxigenic clostridia', *Clinical Microbiology Reviews*, 3(1), pp. 66–98. doi: 320(7228): 161–165.

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. 452–464. doi: 10.1016/j.mimet.2007.05.021.

Heap, J. T. *et al.* (2009) 'A modular system for Clostridium shuttle plasmids', *Journal of Microbiological Methods*. Elsevier B.V., 78(1), pp. 79–85. doi: 10.1016/j.mimet.2009.05.004.

Heap, J. T. *et al.* (2010) 'The ClosTron: Mutagenesis in Clostridium refined and streamlined', *Journal of Microbiological Methods*, 80(1), pp. 49–55. doi: 10.1016/j.mimet.2009.10.018.

Heap, J. T. *et al.* (2012) 'Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker', *Nucleic Acids Research*, 40(8). doi: 10.1093/nar/gkr1321.

Heijstra, B. D. *et al.* (2013) 'Novel bacteria and methods of use thereof'. Google Patents.

Heldt, D. *et al.* (2009) 'Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol utilization in Clostridium kluyveri', *Biochemical Journal*, 423(2), pp. 199–207. doi: 10.1042/BJ20090780.

Henstra, A. M. *et al.* (2007) 'Microbiology of synthesis gas fermentation for biofuel production', *Current Opinion in Biotechnology*, 18(3), pp. 200–206. doi: 10.1016/j.copbio.2007.03.008.

Henstra, A. M. A. M. (2018) 'Manuscript in preparation'.

Herrmann, G. *et al.* (2008) 'Energy conservation via electron-transferring flavoprotein in anaerobic bacteria', *Journal of Bacteriology*, 190(3), pp. 784–791. doi: 10.1128/JB.01422-07.

Hess, V., Schuchmann, K. and Müller, V. (2013) 'The ferredoxin: NAD+ Oxidoreductase (Rnf) from the acetogen acetobacterium woodii requires na+ and is reversibly coupled to the membrane potential', *Journal of Biological Chemistry*, 288(44), pp. 31496–31502. doi: 10.1074/jbc.M113.510255.

Hoegh-Guldberg, O. *et al.* (2007) 'Coral reefs under rapid climate change and ocean acidification', *Science*, 318(5857), pp. 1737–1742. doi: 10.1126/science.1152509.

Hsu, P. D., Lander, E. S. and Zhang, F. (2014) 'Development and applications of CRISPR-Cas9 for genome engineering', *Cell*. Elsevier, 157(6), pp. 1262–1278. doi: 10.1016/j.cell.2014.05.010.

Huang, H. *et al.* (2016) 'CRISPR/Cas9-Based Efficient Genome Editing in Clostridium ljungdahlii, an Autotrophic Gas-Fermenting Bacterium', *ACS Synthetic Biology*, 5(12), pp. 1355–1361. doi: 10.1021/acssynbio.6b00044.

Huhnke, Raymond L.; Lewis, Randy S.; Tanner, R. S. (2008) 'Isolation and Characterization of novel Clostridial Species', *Patent*.

Humphreys, C. M. *et al.* (2015) 'Whole genome sequence and manual annotation of Clostridium autoethanogenum, an industrially relevant bacterium', *BMC Genomics*. BMC Genomics, 16(1), p. 1085. doi: 10.1186/s12864-015-2287-5.

Hwang, W. Y. *et al.* (2013) 'Efficient genome editing in zebrafish using a CRISPR-Cas system', *Nature Biotechnology*. Nature Publishing Group, 31(3), pp. 227–229. doi: 10.1038/nbt.2501.

Imkamp, F. and Müller, V. (2007) 'Acetogenic Bacteria', in *Encyclopedia of Life Sciences*. Chichester, UK: John Wiley & Sons, Ltd, pp. 1–8. doi: 10.1002/9780470015902.a0020086.

Isom, C. E. *et al.* (2016) 'Crystal structure and DNA binding activity of a PadR family transcription regulator from hypervirulent Clostridium difficile R20291', *BMC Microbiology*. BMC Microbiology, 16, pp. 1–12. doi: 10.1186/s12866-016-0850-0.

Ivey, D. M. and Ljungdahl, L. G. (1986) 'Purification and Characterization of the Fl-ATPase from Clostridium thermoaceticum', 165(1), pp. 252–257.

Jabbari, S. *et al.* (2013) 'The putative influence of the agr operon upon survival mechanisms used by Clostridium acetobutylicum', *Mathematical Biosciences*. Elsevier Inc., 243(2), pp. 223–239. doi: 10.1016/j.mbs.2013.03.005.

Jeong, Y. *et al.* (2014) 'Draft genome sequence of acid-tolerant Clostridium drakei SL1T, a potential chemical producer through syngas fermentation', *Genome announcements*, 2(3), pp. 4–5. doi: 10.1128/genomeA.00387-14.Copyright.

Ji, G., Beavis, R. C. and 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. doi: 10.1073/pnas.92.26.12055.

Ji, G., Beavis, R. and Novick, R. P. (1997) 'Bacterial interference caused by autoinducing peptide variants.', *Science (New York, N.Y.)*, 276(5321), pp. 2027–2030. doi: 10.1126/science.276.5321.2027.

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. and Woods, D. R. (1986) 'Acetone-butanol fermentation revisited.', *Microbiological reviews*, 50(4), pp. 484–524. doi: 3540574.

Kamiya, S. *et al.* (1989) 'Heat or Alkali', 28(1989), pp. 217–221.

Kane, M. D., Brauman, A. and Breznak, J. A. (1991) 'Clostridium mayombei sp. nov., an H2/CO2 acetogenic bacterium from the gut of the African soil-feeding termite, Cubitermes speciosus', *Archives of Microbiology*, 156(2), pp. 99–104. doi: 10.1007/BF00290980.

Kane, M. D. and Breznak, J. A. (1991) 'Acetonema longum gen. nov. sp. nov., an H 2/CO 2 acetogenic bacterium from the termite, Pterotermes occidentis', *Archives of microbiology*. Springer, 156(2), pp. 91–98.

Karberg, M. *et al.* (2001) 'Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria', *Nature Biotechnology*, 19(12), pp. 1162–1167. doi: 10.1038/nbt1201-1162.

Karl, T. R. (2003) 'Modern Global Climate Change', *Science*, 302(5651), pp. 1719–1723. doi: 10.1126/science.1090228.

Katayama, S. I. *et al.* (1993) 'Comparison of the alpha-toxin genes of Clostridium perfringens type A and C strains: Evidence for extragenic regulation of transcription', *Infection and Immunity*, 61(2), pp. 457–463.

Kerfeld, C. A. and Erbilgin, O. (2015) 'Bacterial microcompartments and the modular construction of microbial metabolism', *Trends in Microbiology*. Elsevier Ltd, 23(1), pp. 22–34. doi: 10.1016/j.tim.2014.10.003.

Kita, A. *et al.* (2013) 'Development of genetic transformation and heterologous expression system in carboxydotrophic thermophilic acetogen *Moorella thermoacetica*', *Journal of bioscience and bioengineering*, 115(4), pp. 347–52. doi: 10.1016/j.jbiosc.2012.10.013.

Klasson, K. T. *et al.* (1991) 'Bioreactor design for synthesis fermentations', *Fuel*, 70, pp. 605–614.

Kobayashi, H. *et al.* (2004) 'Programmable cells: interfacing natural and engineered gene networks.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(22), pp. 8414–9. doi: 10.1073/pnas.0402940101.

Koepke, M. and Al-Sinawi, B. (2013) 'Selection Method and Recombinant Microorganisms and uses Therefor (US20130316364 A1)'. Google Patents.

Koepke, M. and Chen, W. Y. (2013) 'Recombinant microorganisms and uses therefor (US20130323806 A1)'. Google Patents.

Kong, K. F., Vuong, C. and Otto, M. (2006) 'Staphylococcus quorum sensing in biofilm formation and infection', *International Journal of Medical Microbiology*, 296(2–3), pp. 133–139. doi: 10.1016/j.ijmm.2006.01.042.

Köpke, M. *et al.* (2010) 'Clostridium ljungdahlii represents a microbial production platform based on syngas.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(29), pp. 13087–92. doi: 10.1073/pnas.1004716107.

Köpke, M., Mihalcea, C., Liew, F., *et al.* (2011) '2,3-Butanediol Production By Acetogenic Bacteria, an Alternative Route To Chemical Synthesis, Using Industrial Waste Gas.', *Applied and environmental microbiology*, 77(15), pp. 5467–5475. doi: 10.1128/AEM.00355-11.

Köpke, M., Mihalcea, C., Bromley, J. C., *et al.* (2011) 'Fermentative production of ethanol from carbon monoxide', *Current Opinion in Biotechnology*, 22(3), pp. 320–325. doi: 10.1016/j.copbio.2011.01.005.

Köpke, M. and Liew, F. (2011) 'Recombinant microorganisms and methods for production thereof', *Patent*.

Köpke, M. and Liew, F. (2012) 'Production of butanol from carbon monoxide by a recombinant microorganism', *Patent*.

Koren, S. *et al.* (2013) 'Reducing assembly complexity of microbial genomes with single-molecule sequencing.', *Genome biology*, 14(9), p. R101. doi: 10.1186/gb-2013-14-9-r101.

Körner, H., Sofia, H. J. and Zumft, W. G. (2003) 'Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs', *FEMS Microbiology Reviews*, 27(5), pp. 559–592. doi: 10.1016/S0168-6445(03)00066-4.

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. doi: 10.1271/bbb.60370.

Kotte, A.-K. (2013) *RNPP-type quorum sensing in Clostridium acetobutylicum*. University of Nottingham.

Kotte, A.-K. *et al.* (2017) 'RNPP-type quorum sensing regulates solvent formation and sporulation in Clostridium acetobutylicum', *bioRxiv*.

Kozlowicz, B. K., Bae, T. and 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. doi: 10.1111/j.1365-2958.2004.04286.x.

Krumholz, L. R. and Bryant, M. P. (1985) 'Clostridium pfennigii sp. nov. uses methoxyl groups of monobenzenoids and produces butyrate', *International Journal of Systematic and Evolutionary Microbiology*. Microbiology Society, 35(4), pp. 454–456.

Kuehne, S. A. and Minton, N. P. (2012) 'ClosTron-mediated engineering of Clostridium', *Bioengineered*, 3(4), pp. 247–254. doi: 10.4161/bioe.21004.

Küsel, K. *et al.* (2000) 'Clostridium scatologenes strain SL1 isolated as an acetogenic bacterium from acidic sediments.', *International journal of systematic and evolutionary microbiology*. Microbiology Society, 50(2), pp. 537–546.

Küsel, K. *et al.* (2001) 'Physiological Ecology of Clostridium glycolicum RD-1 , an Aerotolerant Acetogen Isolated from Sea Grass Roots Physiological Ecology of Clostridium glycolicum RD-1 , an Aerotolerant Acetogen Isolated from Sea Grass Roots', *Applied and Environmental Microbiology*, 67(10), pp. 4734–4741. doi: 10.1128/AEM.67.10.4734.

LaSarre, B. and Federle, M. J. (2013) 'Exploiting Quorum Sensing To Confuse Bacterial Pathogens',

Microbiology and Molecular Biology Reviews, 77(1), pp. 73–111. doi: 10.1128/MMBR.00046-12.

Latif, H. *et al.* (2014) 'Trash to treasure: Production of biofuels and commodity chemicals via syngas fermenting microorganisms', *Current Opinion in Biotechnology*. Elsevier Ltd, 27, pp. 79–87. doi: 10.1016/j.copbio.2013.12.001.

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. doi: http://dx.doi.org/.

Leang, C. *et al.* (2013) 'A Genetic system for Clostridium ljungdahlii: A chassis for autotrophic production of biocommodities and a model homoacetogen', *Applied and Environmental Microbiology*, 79(4), pp. 1102–1109. doi: 10.1128/AEM.02891-12.

Lee, J. and Zhang, L. (2014) 'The hierarchy quorum sensing network in Pseudomonas aeruginosa', *Protein and Cell*, 6(1), pp. 26–41. doi: 10.1007/s13238-014-0100-x.

Leite, G. B., Abdelaziz, A. E. M. and Hallenbeck, P. C. (2013) 'Algal biofuels: Challenges and opportunities', *Bioresource Technology*. Elsevier Ltd, 145, pp. 134–141. doi: 10.1016/j.biortech.2013.02.007.

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', *Infection and Immunity*, 79(6), pp. 2451–2459. doi: 10.1128/IAI.00169-11.

Liew, F. *et al.* (2016) 'Gas Fermentation—A Flexible Platform for Commercial Scale Production of Low-Carbon-Fuels and Chemicals from Waste and Renewable Feedstocks', *Frontiers in Microbiology*, 7(MAY). doi: 10.3389/fmicb.2016.00694.

Liew, F. *et al.* (2016) 'Insights into CO 2 Fixation Pathway of Clostridium autoethanogenum by Targeted Mutagenesis', *mBio*, 7(3), pp. e00427-16. doi: 10.1128/mBio.00427-16.

Liew, F. *et al.* (2017) 'Metabolic engineering of Clostridium autoethanogenum for selective alcohol production', *Metabolic Engineering*, 40(January), pp. 104–114. doi: 10.1016/j.ymben.2017.01.007.

Liew, F. M. (2015) *Eric Liew - Thesis 2015*. University of Nottingham.

Lindblad, P. *et al.* (2012) 'Design, engineering, and construction of photosynthetic microbial cell factories for renewable solar fuel production', *Ambio*, 41(SUPPL.2), pp. 163–168. doi: 10.1007/s13280-012-0274-5.

Liou, J. S.-C. C. *et al.* (2005) 'Clostridium carboxidivorans sp. nov., a solvent-producing clostridium isolated from an agricultural settling lagoon, and reclassification of the acetogen Clostridium scatologenes strain SL1 as Clostridium drakei sp. nov.', *International journal of systematic and evolutionary microbiology*. Microbiology Society, 55(5), pp. 2085–2091. doi: 10.1099/ijs.0.63482-0.

Liu, C. *et al.* (2011) 'Sequential Establishment of Stripe Patterns in an Expanding Cell Population', *Science*, 334(6053), pp. 238–241. doi: 10.1126/science.1209042.

Loenen, W. A. M. and Raleigh, E. A. (2014) 'The other face of restriction: Modification-dependent enzymes', *Nucleic Acids Research*, 42(1), pp. 56–69. doi: 10.1093/nar/gkt747.

Long, S., Jones, D.T. and Woods, D. R. (1984) 'The relationship between sporulation and solvent production in Clostridium acetobutylicum P262', *Biotechnology Letters Vol*, 6(8), pp. 1–5. doi: 10.1007/s13398-014-0173-7.2.

Lu, J., Sheahan, C. and Fu, P. (2011) 'Metabolic engineering of algae for fourth generation biofuels production', *Energy Environ. Sci.*, 4(7), pp. 2451–2466. doi: 10.1039/C0EE00593B.

Luria, S. E., Adams, J. N. and Ting, R. C. (1960) 'Transduction of lactose-utilizing ability among strains of E. coli and S. dysenteriae and the properties of the transducing phage particles', *Virology*, 12(3), pp. 348–390. doi: 10.1016/0042-6822(60)90161-6.

Lynd, L. H. and Zeikus, J. G. (1983) 'Metabolism of H2-CO2, methanol, and glucose by Butyribacterium methylotrophicum.', *Journal of bacteriology*, 153(3), pp. 1415–23.

Lyristis, M. *et al.* (1994) 'Identification and molecular analysis of a locus that regulates extracellular toxin production in Clostridium perfringens.', *Molecular microbiology*, 12(5), pp. 761–777. doi: 10.1111/j.1365-2958.1994.tb01063.x.

Mansfield, R. (2018) 'Manuscript in preparation'.

Marcellin, E. *et al.* (2016) 'Low carbon fuels and commodity chemicals from waste gases – Systematic approach to understand energy metabolism in a model acetogen', *Green Chem.*, pp. 3020–3028. doi: 10.1039/C5GC02708J.

Marchand, N. and Collins, C. H. (2016) 'Synthetic Quorum Sensing and Cell-Cell Communication in Gram-Positive Bacillus megaterium', *ACS Synthetic Biology*, 5(7), pp. 597–606. doi: 10.1021/acssynbio.5b00099.

de Marco, A. *et al.* (2005) 'Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones.', *Cell stress & chaperones*, 10(4), pp. 329–339. doi: 10.1379/CSC-139R.1.

Marinus, M. G. and Casadesus, J. (2009) 'Roles of DNA adenine methylation in host-pathogen interactions: Mismatch repair, transcriptional regulation, and more', *FEMS Microbiology Reviews*, 33(3), pp. 488–503. doi: 10.1111/j.1574-6976.2008.00159.x.

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. doi: 10.1128/JB.00473-13.

Mascher, T., Helmann, J. D. and Unden, G. (2006) 'Stimulus perception in bacterial signal-transducing histidine kinases.', *Microbiology and molecular biology reviews : MMBR*, 70(4), pp. 910–938. doi: 10.1128/MMBR.00020-06.

Masset, J. *et al.* (2012) 'Fermentative hydrogen production from glucose and starch using pure strains and artificial co-cultures of Clostridium spp.', ???, 5, p. 1. doi: 10.1186/1754-6834-5-35.

McGovern, P. E. *et al.* (2004) 'Fermented beverages of pre- and proto-historic China.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(51), pp. 17593–8. doi: 10.1073/pnas.0407921102.

Mechichi, T. *et al.* (1999) 'Clostridium methoxybenzovorans sp. nov., a new aromatic o-demethylating homoacetogen from an olive mill wastewater treatment digester', *International Journal of Systematic and Evolutionary Microbiology*. Microbiology Society, 49(3), pp. 1201–1209.

Meddeb-Mouelhi, F., Dulcey, C. and Beauregard, M. (2012) 'High transformation efficiency of Bacillus subtilis with integrative DNA using glycine betaine as osmoprotectant', *Analytical Biochemistry*. Elsevier Inc., 424(2), pp. 127–129. doi: 10.1016/j.ab.2012.01.032.

Millard, J. (2018) 'Manuscript in preparation'.

Miller, J. H. (1996) 'SPONTANEOUS MUTATORS IN BACTERIA: Insights into Pathways of Mutagenesis and Repair', *Annu. Rev. Microbiol*, 50, pp. 625–43. doi: 10.1146/annurev.micro.50.1.625.

Miller, M. B. and Bassler, B. L. (2001) 'Quorum sensing in bacteria.', *Annual review of microbiology*, 55, pp. 165–99. doi: 10.1146/annurev.micro.55.1.165.

Miller, R. G. and Sorrell, S. R. (2013) 'The future of oil supply', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2006), pp. 20130179–20130179. doi: 10.1098/rsta.2013.0179.

Min, F. *et al.* (2013) 'Gas Fermentation for Commercial Biofuels Production', in *Liquid, Gaseous and Solid Biofuels - Conversion Techniques*, pp. 125–173. doi: 10.5772/52164.

Minton, N. and Morris, J. G. (1981) 'Isolation and Partial Characterization of Three Cryptic Plasmids from Strains of [i]Clostridium butyricum[/i]', *J Gen Microbiol*, 127(2), pp. 325–331. doi: 10.1099/00221287-127-2-325.

Minton, N. P. *et al.* (2016) 'A roadmap for gene system development in Clostridium', *Anaerobe*. Elsevier Ltd, 41, pp. 104–112. doi: 10.1016/j.anaerobe.2016.05.011.

Mitchell, W. J. (2016) 'Sugar uptake by the solventogenic clostridia', *World Journal of Microbiology and Biotechnology*. Springer Netherlands, 32(2), pp. 1–10. doi: 10.1007/s11274-015-1981-4.

Mock, J. *et al.* (2015) 'Energy Conservation Associated with Ethanol Formation from H 2 and CO 2 in Clostridium autoethanogenum Involving Electron Bifurcation', *Journal of Bacteriology*. Edited by W. W. Metcalf, 197(18), pp. 2965–2980. doi: 10.1128/JB.00399-15.

Mohr, A. and Raman, S. (2013) 'Lessons from first generation biofuels and implications for the sustainability appraisal of second generation biofuels', *Energy Policy*. Elsevier, 63, pp. 114–122. doi:

10.1016/j.enpol.2013.08.033.

Molitor, B. *et al.* (2016) 'Expanding the molecular toolkit for the homoacetogen Clostridium ljungdahlii.', *Scientific reports*. Nature Publishing Group, 6, p. 31518. doi: 10.1038/srep31518.

Monnet, V., Juillard, V. and Gardan, R. (2014) 'Peptide conversations in Gram-positive bacteria.', *Critical reviews in microbiology*, 7828(January), pp. 1–13. doi: 10.3109/1040841X.2014.948804.

Mueller, A. P., Koepke, M. and Nagaraju, S. (2013) 'Recombinant microorganisms and uses therefor'. Google Patents.

Müller, V. (2003) 'Energy Conservation in Acetogenic', *Applied and Environmental Microbiology*, 69(11), pp. 6345–6353. doi: 10.1128/AEM.69.11.6345.

Müller, V. *et al.* (2008) 'Discovery of a ferredoxin:NAD+-oxidoreductase (Rnf) in Acetobacterium woodii: A novel potential coupling site in acetogens', *Annals of the New York Academy of Sciences*, 1125, pp. 137–146. doi: 10.1196/annals.1419.011.

Nadal Jimenez, P. *et al.* (2012) 'The Multiple Signaling Systems Regulating Virulence in Pseudomonas aeruginosa', *Microbiology and Molecular Biology Reviews*, 76(1), pp. 46–65. doi: 10.1128/MMBR.05007-11.

Nagaraju, S. *et al.* (2015) 'Recombinant microorganisms and methods of use thereof'. Google Patents.

Nagaraju, S. *et al.* (2016) 'Genome editing of Clostridium autoethanogenum using CRISPR/Cas9', *Biotechnology for Biofuels*. BioMed Central, 9(1), p. 219. doi: 10.1186/s13068-016-0638-3.

Naik, S. N. *et al.* (2010) 'Production of first and second generation biofuels: A comprehensive review', *Renewable and Sustainable Energy Reviews*, 14(2), pp. 578–597. doi: 10.1016/j.rser.2009.10.003.

Nasr, N. *et al.* (2017) 'Mono- and co-substrate utilization kinetics using mono- and co-culture of Clostridium beijerinckii and Clostridium saccharoperbutylacetonicum', *Bioresource Technology*. Elsevier Ltd, 241, pp. 152–160. doi: 10.1016/j.biortech.2017.05.086.

Ng, P. C. and Henikoff, S. (2003) 'SIFT: Predicting amino acid changes that affect protein function', *Nucleic Acids Research*, 31(13), pp. 3812–3814. doi: 10.1093/nar/gkg509.

Ng, T. K. and Zeikus, J. G. (1981) 'Ethanol Production by Thermophilic Bacteria Fermentation of cellulose substrate by cocultures of Clostridium termocellum and clostridium thermohydrosulfuricum', 41(6), pp. 1337–1343.

Ng, W.-L. and Bassler, B. L. (2009) 'Bacterial Quorum-Sensing Network Architectures', *Annual Review of Genetics*, 43(1), pp. 197–222. doi: 10.1146/annurev-genet-102108-134304.

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). doi: 10.1371/journal.pone.0056051.

Nölling, J. *et al.* (2001) 'Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum', *Journal of Bacteriology*, 183(16), pp. a823–4838. doi: 10.1128/JB.183.16.4823.

Novick, R. P. *et al.* (1993) 'Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule', *The EMBO Journal*, 12(10), pp. 3967–3975. doi: 10.1016/j.jnutbio.2011.09.003.

Novick, R. P. (2003) 'Autoinduction and signal transduction in the regulation of staphylococcal virulence', *Molecular Microbiology*, 48(6), pp. 1429–1449. doi: 10.1046/j.1365-2958.2003.03526.x.

O'Brien, R. W. and Morris, J. G. (1971) 'Oxygen and the Growth and Metabolism of Clostridium acetobutylicum', *Journal of General Microbiology*, 68(3), pp. 307–318. doi: 10.1099/00221287-68-3-307.

Ohlsen, K. L., Grimsley, J. K. and Hoch, J. A. (1994) 'Deactivation of the sporulation transcription factor SpoOA by the SpoOE protein phosphatase', *Developmental Biology*, 91(March), pp. 1756–1760. doi: 10.1073/PNAS.91.5.1756.

Ohtani, K. (2016) 'Gene regulation by the VirS/VirR system in Clostridium perfringens', *Anaerobe*. Elsevier Ltd, 41, pp. 5–9. doi: 10.1016/j.anaerobe.2016.06.003.

Ohtani, K. and Shimizu, T. (2015) 'Regulation of toxin gene expression in Clostridium perfringens', *Research*

in Microbiology. Elsevier Masson SAS, 166(4), pp. 280–289. doi: 10.1016/j.resmic.2014.09.010.

Ow, S. Y. *et al.* (2009) 'iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly".', *Journal of proteome research*, 8(11), pp. 5347–55. doi: 10.1021/pr900634c.

Palmer, B. R. and Marinus, M. G. (1994) 'The dam and dcm strains of Escherichia coli--a review.', *Gene*, 143, pp. 1–12. doi: 0378-1119(94)90597-5 [pii].

Pantaléon, V. *et al.* (2014) 'Biofilms of Clostridium species', *Anaerobe*, 30, pp. 193–198. doi: 10.1016/j.anaerobe.2014.09.010.

Papenfort, K. and Bassler, B. L. (2016) 'Quorum sensing signal–response systems in Gram-negative bacteria', *Nature Reviews Microbiology*. Nature Publishing Group, 14(9), pp. 576–588. doi: 10.1038/nrmicro.2016.89.

Papoutsakis, E. T. (2008) 'Engineering solventogenic clostridia', *Current Opinion in Biotechnology*, 19(5), pp. 420–429. doi: 10.1016/j.copbio.2008.08.003.

Paredes, C. J., Alsaker, K. V and Papoutsakis, E. T. (2005) 'A comparative genomic view of clostridial sporulation and physiology', *Nature Reviews Microbiology*, 3(12), pp. 969–978. doi: 10.1038/nrmicro1288.

Passador, L. *et al.* (1993) 'Expression of Pseudomonas aeruginosa virulence genes requires cell-to-cell communication', *Science*, 260(5111), pp. 1127–1130. doi: 10.1126/science.8493556.

Peng, H. L. *et al.* (1988) 'Cloning, Characterization, and Sequensing of an Accessory Gene Regulation (agr) in Staphylococcus aureus', *J Bacteriol*, 170(9), pp. 4365–4372.

Perego, M. (2001) 'A new family of aspartyl phosphate phosphatases targeting the sporulation transcription factor Spo0A of Bacillus subtilis', *Molecular Microbiology*, 42(1), pp. 133–143. doi: 10.1046/j.1365-2958.2001.02611.x.

Perkel, J. M. (2009) 'iTRAQ gets put to the test', *Journal of Proteome Research*, 8(11), pp. 4885–4885. doi: 10.1021/pr900897d.

Philips, J. *et al.* (2017) 'Biofilm Formation by Clostridium ljungdahlii Is Induced by Sodium Chloride Stress: Experimental Evaluation and Transcriptome Analysis', *Plos One*, 12(1), p. e0170406. doi: 10.1371/journal.pone.0170406.

Phillips, J. R., Huhnke, R. L. and Atiyeh, H. K. (2017) 'Syngas Fermentation: A Microbial Conversion Process of Gaseous Substrates to Various Products', *Fermentation*, 3(2), p. 28. doi: 10.3390/fermentation3020028.

Poehlein, A. *et al.* (2017) 'Microbial solvent formation revisited by comparative genome analysis', *Biotechnology for Biofuels*. BioMed Central, 10(1), p. 58. doi: 10.1186/s13068-017-0742-z.

Purdy, D. *et al.* (2002) 'Conjugative transfer of clostridial shuttle vectors from [i]Escherichia coli[/i] to [i]Clostridium difficile[/i] through circumvention of the restriction barrier', *Mol Microbiol*, 46(2), pp. 439–452. doi: 3134 [pii].

Pyne, M. E. *et al.* (2013) 'Development of an electrotransformation protocol for genetic manipulation of Clostridium pasteurianum.', *Biotechnology for biofuels*, 6(1), p. 50. doi: 10.1186/1754-6834-6-50.

Pyne, M. E. *et al.* (2014) 'Technical guide for genetic advancement of underdeveloped and intractable Clostridium', *Biotechnology Advances*. Elsevier Inc., 32(3), pp. 623–641. doi: 10.1016/j.biotechadv.2014.04.003.

Pyne, M. E. *et al.* (2016) 'Genome-directed analysis of prophage excision, host defence systems, and central fermentative metabolism in Clostridium pasteurianum', *Scientific Reports*, 6(1), p. 26228. doi: 10.1038/srep26228.

Le Quéré, C. *et al.* (2016) 'Global Carbon Budget 2016', *Earth System Science Data*, 8(2), pp. 605–649. doi: 10.5194/essd-8-605-2016.

Qureshi, N. *et al.* (2010) 'Production of butanol (a biofuel) from agricultural residues: Part II - Use of corn stover and switchgrass hydrolysates', *Biomass and Bioenergy*. Elsevier Ltd, 34(4), pp. 566–571. doi: 10.1016/j.biombioe.2009.12.023.

Ragsdale, S. W. *et al.* (1983) 'Properties of purified carbon monoxide dehydrogenase from Clostridium thermoaceticum, a nickel, iron-sulfur protein.', *The Journal of biological chemistry*, 258(4), pp. 2364–9.

Ragsdale, S. W. (1997) 'The eastern and western branches of the Wood/Ljungdahl pathway: how the east and west were won', *BioFactors (Oxford, England)*, 6(1), pp. 3–11. doi: 10.1002/biof.5520060102.

Ragsdale, S. W. (2003) 'Pyruvate ferredoxin oxidoreductase and its radical intermediate', *Chemical Reviews*, 103(6), pp. 2333–2346. doi: 10.1021/cr020423e.

Ragsdale, S. W. (2004) 'Life with Carbon Monoxide', *Critical Reviews in Biochemistry and Molecular Biology*, 39, pp. 165–195. doi: 10.1080/10409230490496577.

Ragsdale, S. W. and Pierce, E. (2008) 'Acetogenesis and the Wood-Ljungdahl pathway of CO2 fixation', *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1784(12), pp. 1873–1898. doi: 10.1016/j.bbapap.2008.08.012.

Ramachandriya, K. D., DeLorme, M. J. and Wilkins, M. R. (2010) 'Heat shocking of Clostridium strain P11 to promote sporulation and ethanol production', *Biological Engineering*, 2(2), pp. 115–131.

Ran, F. A. *et al.* (2013) 'Genome engineering using the CRISPR-Cas9 system', *Nature Protocols*, 8(11), pp. 2281–2308. doi: 10.1038/nprot.2013.143.

Raut, M. P. *et al.* (2016) 'Quantitative proteomic analysis of the influence of lignin on biofuel production by Clostridium acetobutylicum ATCC 824.', *Biotechnology for biofuels*. BioMed Central, 9(1), p. 113. doi: 10.1186/s13068-016-0523-0.

Ravagnani, A. *et al.* (2000) 'Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia', *Molecular Microbiology*, 37(5), pp. 1172–1185. doi: 10.1046/j.1365-2958.2000.02071.x.

Recsei, P. *et al.* (1986) 'Regulation of exoprotein gene expression in Staphylococcusaureus by agr', *Mol Gen Genet*, 202, pp. 58–61.

Reddy, V. S. *et al.* (2012) 'The major facilitator superfamily (MFS) revisited', *FEBS Journal*, 279(11), pp. 2022–2035. doi: 10.1111/j.1742-4658.2012.08588.x.

Reysenbach, A. L., Ravenschroft, N. and Long, S. (1986) 'Characterization, biosynthesis, and regulation of granulose in Clostridium acetobutylicum', *Applied and Environmental Microbiology*, 52(1), pp. 185–190.

Ribeiro, K. S. *et al.* (2007) 'Transport and its infrastructure. In Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.', in. Cambridge University Press, Cambridge, p. 328.

Richter, H. *et al.* (2012) 'Characterization of CRISPR RNA processing in Clostridium thermocellum and Methanococcus maripaludis', *Nucleic Acids Research*, 40(19), pp. 9887–9896. doi: 10.1093/nar/gks737.

Riedel, C. U. *et al.* (2009) 'AgrD-dependent quorum sensing affects biofilm formation, invasion, virulence and global gene expression profiles in Listeria monocytogenes', *Molecular Microbiology*, 71(5), pp. 1177–1189. doi: 10.1111/j.1365-2958.2008.06589.x.

Rieu-Lesme, F. *et al.* (1998) 'Isolation from the Rumen of a New Acetogenic Bacterium Phylogenetically Closely Related toClostridium difficile', *Anaerobe*. Elsevier, 4(2), pp. 89–94.

Roberts, R. J. *et al.* (2005) 'REBASE - Restriction enzymes and DNA methyltransferases', *Nucleic Acids Research*, 33(DATABASE ISS.), pp. 230–232. doi: 10.1093/nar/gki029.

Roberts, R. J. *et al.* (2015) 'REBASE-a database for DNA restriction and modification: Enzymes, genes and genomes', *Nucleic Acids Research*, 43(D1), pp. D298–D299. doi: 10.1093/nar/gku1046.

Rocha-Estrada, J. *et al.* (2010) 'The RNPP family of quorum-sensing proteins in Gram-positive bacteria', *Applied Microbiology and Biotechnology*, 87(3), pp. 913–923. doi: 10.1007/s00253-010-2651-y.

Rode, L. M., Genthner, B. R. S. and Bryant, M. P. (1981) 'Syntrophic association by cocultures of the methanol-and CO2-H2-utilizing species Eubacterium limosum and pectin-fermenting Lachnospira multiparus during growth in a pectin medium', *Applied and environmental microbiology*. Am Soc Microbiol, 42(1), pp. 20–22.

Rood, J. I. (1998) 'VIRULENCE GENES OF Clostridial Perfringens', *Annual review of Microbiology*, 52, pp. 333–360.

Ruby, E. G. (1996) 'Lessons from a cooperative, bacterial-animal association: the Vibrio fischeri-Euprymna scolopes light organ symbiosis.', *Annual review of microbiology*, 50, pp. 591–624. doi:

10.1146/annurev.micro.50.1.591.

Rutherford, S. T. and Bassler, B. L. (2012) 'Bacterial quorum sensing: its role in virulence and possibilities for its control.', *Cold Spring Harbor perspectives in medicine*, 2(11). doi: 10.1101/cshperspect.a012427.

Saeidi, N. *et al.* (2014) 'Engineering microbes to sense and eradicate Pseudomonas aeruginosa, a human pathogen', *Molecular Systems Biology*. Nature Publishing Group, 7(1), pp. 521–521. doi: 10.1038/msb.2011.55.

Sauer, U. *et al.* (1995) 'Sigma factor and sporulation genes in Clostridium', *FEMS Microbiology Reviews*, 17(3), pp. 331–340. doi: 10.1111/j.1574-6976.1995.tb00216.x.

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. doi: 10.1128/JB.00272-11.

Scaria, J. *et al.* (2011) 'Clostridium difficile transcriptome analysis using pig ligated loop model reveals modulation of pathways not modulated in vitro', *Journal of Infectious Diseases*, 203(11), pp. 1613–1620. doi: 10.1093/infdis/jir112.

Schaeffer, A. B. and Fulton, M. D. (1990) 'A Simplified Method of Staining Endospores', *Science*, p. 194. doi: 10.1126/science.77.1990.194.

Schauder, R. *et al.* (1986) 'Acetate oxidation to CO2 in anaerobic bacteria via a novel pathway not involving reactions of the citric acid cycle', *Archives of Microbiology*, 145(2), pp. 162–172. doi: 10.1007/BF00446775.

Schiel-Bengelsdorf, B. and Dürre, P. (2012) 'Pathway engineering and synthetic biology using acetogens', *FEBS Letters*. Federation of European Biochemical Societies, 586(15), pp. 2191–2198. doi: 10.1016/j.febslet.2012.04.043.

Schink, B. (1984) 'Clostridium magnum sp. nov., a non-autotrophic homoacetogenic bacterium', *Archives of Microbiology*, 137(3), pp. 250–255. doi: 10.1007/BF00414553.

Schuchmann, K. *et al.* (2015) 'Nonacetogenic growth of the acetogen Acetobacterium woodii on 1,2-propanediol', *Journal of Bacteriology*, 197(2), pp. 382–391. doi: 10.1128/JB.02383-14.

Schuchmann, K. and Müller, V. (2014) 'Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria', *Nature Reviews Microbiology*. Nature Publishing Group, 12(12), pp. 809–821. doi: 10.1038/nrmicro3365.

Sebaihia, M. *et al.* (2006) 'The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome.', *Nature genetics*, 38(7), pp. 779–786. doi: 10.1038/ng1830.

Shao, L. *et al.* (2007) 'Targeted gene disruption by use of a group II intron (targetron) vector in Clostridium beijerinckii.', *Cell research*, 17(11), pp. 963–965. doi: 10.1038/cr.2007.91.

Shimizu, T. *et al.* (2002) 'Proteome and Transcriptome Analysis of the Virulence Genes Regulated by the VirR / VirS System in Clostridium perfringens', 184(10), pp. 2587–2594. doi: 10.1128/JB.184.10.2587.

Simpson, S. D. *et al.* (2009) 'Clostridium autoethanogenum strain and methods of use thereof to produce ethanol and acetate'. Google Patents.

Simpson, S. D., Koepke, M. and Liew, F. (2011) 'Recombinant Microorganisms and Methods of Use Thereof'. Google Patents.

Sims, R. E. H. *et al.* (2010) 'An overview of second generation biofuel technologies', *Bioresource Technology*. Elsevier Ltd, 101(6), pp. 1570–1580. doi: 10.1016/j.biortech.2009.11.046.

Singh, P. K. *et al.* (2000) 'Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms.', *Nature*, 407(6805), pp. 762–764. doi: 10.1038/35037627.

Sipma, J. *et al.* (2006) 'Microbial CO conversions with applications in synthesis gas purification and bio-desulfurization.', *Critical reviews in biotechnology*, 26(1), pp. 41–65. doi: 10.1080/07388550500513974.

Skerker, J. M. *et al.* (2005) 'Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: A system-level analysis', *PLoS Biology*, 3(10). doi: 10.1371/journal.pbio.0030334.

Sloan, J. *et al.* (1992) 'Construction of a sequencedClostridium perfringens-Escherichia coli shuttle plasmid',

Plasmid, 27(3), pp. 207–219. doi: 10.1016/0147-619X(92)90023-4.

Smeets, E. *et al.* (2014) 'The impact of the rebound effect of the use of first generation biofuels in the EU on greenhouse gas emissions: A critical review', *Renewable and Sustainable Energy Reviews*. Elsevier, 38, pp. 393–403. doi: 10.1016/j.rser.2014.05.035.

Smits, W. K. *et al.* (2016) 'Clostridium difficile infection', *Nature Reviews Disease Primers*, 2(7), p. 16020. doi: 10.1038/nrdp.2016.20.

Solomon, S. *et al.* (2007) 'Climate Change 2007. The Physical Science basis: Summary for policymakers. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change', *Climate Change 2007: The Physical Science basis: Summary for policymakers. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*.

Solomon, S. *et al.* (2009) 'Irreversible climate change due to carbon dioxide emissions.', *Proceedings of the National Academy of Sciences of the United States of America*, 106(6), pp. 1704–9. doi: 10.1073/pnas.0812721106.

Sorek, R., Lawrence, C. M. and Wiedenheft, B. (2013) 'CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea', *Annual Review of Biochemistry*, 82(1), pp. 237–266. doi: 10.1146/annurev-biochem-072911-172315.

Stabler, R. *et al.* (2009) 'Comparative genome and phenotypic analysis of Clostridium difficile 027 strains provides insight into the evolution of a hypervirulent bacterium', *Genome Biology*, 10(9), p. R102. doi: 10.1186/gb-2009-10-9-r102.

Stackebrandt, E. *et al.* (1999) 'Phylogenetic basis for a taxonomic dissection of the genus Clostridium', *FEMS Immunology and Medical Microbiology*, 24(3), pp. 253–258. doi: 10.1016/S0928-8244(99)00039-5.

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. 641–654. doi: 10.1111/j.1365-2958.2011.07608.x.

Steiner, E. *et al.* (2012) 'An agr quorum sensing system that regulates granulose formation and sporulation in Clostridium acetobutylicum', *Applied and Environmental Microbiology*, 78(4), pp. 1113–1122. doi: 10.1128/AEM.06376-11.

Stock, A. M., Robinson, V. L. and Goudreau, P. N. (2000) 'Two-component signal transduction', *Annual Review Of Biochemistry*, 69, pp. 183–215. doi: 69/1/183 [pii]\r10.1146/annurev.biochem.69.1.183.

Stouthamer, a H. (1973) 'A theoretical study on the amount of ATP required for synthesis of microbial cell material', *Antonie van Leeuwenhoek*, 39(1), pp. 545–565. doi: 10.1007/BF02578899.

Stratz, M. *et al.* (1994) 'Plasmid Transfer Into the Homoacetogen Acetobacterium-Woodii By Electroporation and Conjugation', *Applied and Environmental Microbiology*, 60(3), pp. 1033–1037.

Szybalski, W. (1958) 'Special Microbiological Systems. Ii. Observations on Chemical Mutagenesis in Microorganisms', *Annals of the New York Academy of Sciences*, 76(3), pp. 475–489. doi: 10.1111/j.1749-6632.1958.tb57106.x.

Taga, M. E. and Bassler, B. L. (2003) 'Chemical communication among bacteria.', *Proceedings of the National Academy of Sciences of the United States of America*, 100 Suppl, pp. 14549–14554. doi: 10.1073/pnas.1934514100.

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. doi: 10.1016/j.pep.2004.03.004.

Talukdar, P. K. *et al.* (2015) 'Updates on the sporulation process in Clostridium species', *Research in Microbiology*, 166(4), pp. 225–235. doi: 10.1016/j.resmic.2014.12.001.

Tan, Y. *et al.* (2013) 'RNA-seq-based comparative transcriptome analysis of the syngas-utilizing bacterium *Clostridium ljungdahlii* DSM 13528 grown autotrophically and heterotrophically', *Mol. Biosyst.*, 9, pp. 2775–2784. doi: 10.1039/c3mb70232d.

Tangney, M. and Mitchell, W. J. (2007) 'Characterisation of a glucose phosphotransferase system in Clostridium acetobutylicum ATCC 824', *Applied Microbiology and Biotechnology*, 74(2), pp. 398–405. doi: 10.1007/s00253-006-0679-9.

Tanner, R. S., Miller, L. M. and Yang, D. (1993) 'Clostridium ljungdahlii sp. nov., an Acetogenic Species in Clostridial rRNA Homology Group I', *International Journal of Systematic Bacteriology*, 43(2), pp. 232–236. doi: 10.1099/00207713-43-2-232.

Thi *et al.* (2011) 'Genetic and Biochemical Analysis of PadR-padC Promoter Interactions during the Phenolic Acid Stress Response in Bacillus subtilis 168', *Journal of Bacteriology*, 193(16), pp. 4180–4191. doi: 10.1128/JB.00385-11.

Thoendel, M. *et al.* (2011) 'Peptide Signaling in the Staphylococci, ChemRev 2010', pp. 117–151.

Thoendel, M. and Horswill, A. R. (2009) 'Identification of Staphylococcus aureus AgrD residues required for autoinducing peptide biosynthesis', *Journal of Biological Chemistry*, 284(33), pp. 21828–21838. doi: 10.1074/jbc.M109.031757.

Tracy, B. P. *et al.* (2012) 'Clostridia: The importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications', *Current Opinion in Biotechnology*. Elsevier Ltd, 23(3), pp. 364–381. doi: 10.1016/j.copbio.2011.10.008.

Tremblay, P. *et al.* (2013) 'The Rnf Complex of Clostridium ljungdahlii Is a Proton-Translocating Ferredoxin : NAD+ Oxidoreductase Essential for Autotrophic Growth', 4(1), pp. 1–8. doi: 10.1128/mBio.00406-12.Editor.

Trevors, J. T. *et al.* (2012) 'Electrotransformation of Bacteria by Plasmid DNA', in *Guide to Electroporation and Electrofusion*, pp. 265–290. doi: 10.1016/B978-0-08-091727-6.50020-4.

Utturkar, S. M. *et al.* (2015) 'Sequence data for Clostridium autoethanogenum using three generations of sequencing technologies.', *Scientific data*, 2, p. 150014. doi: 10.1038/sdata.2015.14.

Valgepea, K., Loi, K. Q., *et al.* (2017) 'Arginine deiminase pathway provides ATP and boosts growth of the gas-fermenting acetogen Clostridium autoethanogenum.', *Metabolic engineering*, 41(April), pp. 202–211. doi: 10.1016/j.ymben.2017.04.007.

Valgepea, K., de Souza Pinto Lemgruber, R., *et al.* (2017) 'Maintenance of ATP Homeostasis Triggers Metabolic Shifts in Gas-Fermenting Acetogens.', *Cell systems*, 4(5), p. 505–515.e5. doi: 10.1016/j.cels.2017.04.008.

Verbeke, T. J. *et al.* (2017) 'Pentose sugars inhibit metabolism and increase expression of an AgrD-type cyclic pentapeptide in Clostridium thermocellum', *Scientific Reports*. Nature Publishing Group, (February), p. In-press. doi: 10.1038/srep43355.

Vidal, J. E., Shak, J. R. and Canizalez-Roman, A. (2015) 'The CpAL quorum sensing system regulates production of hemolysins CPA and PFO to build Clostridium perfringens biofilms', *Infection and Immunity*, 83(6), pp. 2430–2442. doi: 10.1128/IAI.00240-15.

Visick, K. L. *et al.* (2000) 'Vibrio fischeri lux genes play an important role in colonization and development of the host light organ', *Journal of Bacteriology*, 182(16), pp. 4578–4586. doi: 10.1128/JB.182.16.4578-4586.2000.

Vojtov, N., Ross, H. F. and Novick, R. P. (2002) 'Global repression of exotoxin synthesis by staphylococcal superantigens', *Proc.Natl.Acad.Sci.U.S.A*, 99(0027–8424 SB–IM), pp. 10102–10107. doi: 10.1073/pnas.152152499.

Walker, J. (1992) 'Spectrophotometric determination of enzyme activity: alcohol dehydrogenase (ADH)', *Biochemical Education*. Headington Hill Hall, 20(1), pp. 42–43. doi: 10.1016/0307-4412(92)90021-D.

Wang, B. *et al.* (2011) 'Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology', *Nature Communications*. Nature Publishing Group, 2, p. 508. doi: 10.1038/ncomms1516.

Wang, S. *et al.* (2013) 'NADP-Specific Electron-Bifurcating [FeFe]-Hydrogenase in a Functional Complex with Formate Dehydrogenase in Clostridium autoethanogenum Grown on CO', *Journal of Bacteriology*, 195(19), pp. 4373–4386. doi: 10.1128/JB.00678-13.

Wang, Y. *et al.* (2011) 'Single-nucleotide resolution analysis of the transcriptome structure of Clostridium beijerinckii NCIMB 8052 using RNA-Seq.', *BMC genomics*, 12(1), p. 479. doi: 10.1186/1471-2164-12-479.

Wang, Y. *et al.* (2013) 'Development of a gene knockout system using mobile group II introns (Targetron) and genetic disruption of acid production pathways in Clostridium beijerinckii', *Applied and Environmental Microbiology*, 79(19), pp. 5853–5863. doi: 10.1128/AEM.00971-13.

Wang, Y. *et al.* (2015) 'Markerless chromosomal gene deletion in Clostridium beijerinckii using CRISPR/Cas9 system', *Journal of Biotechnology*. Elsevier B.V., 200, pp. 1–5. doi: 10.1016/j.jbiotec.2015.02.005.

Waters, C. M. and Bassler, B. L. (2006) 'Vibrio harveyi', *Genes Dev.*, pp. 2754–2767. doi: 10.1101/gad.1466506.Other.

Wei, N., Quarterman, J. and Jin, Y. S. (2013) 'Marine macroalgae: An untapped resource for producing fuels and chemicals', *Trends in Biotechnology*. Elsevier Ltd, 31(2), pp. 70–77. doi: 10.1016/j.tibtech.2012.10.009.

Weinberg, M. (1927) 'Données récentes sur les microbes anaérobies et leur rôle en pathologie, par M. Weinberg, et B. Ginsbourg', *Paris: Masson et Cie.* Paris: Masson et Cie., pp. 1–291.

Wellington, M. and Rustchenko, E. (2005) '5-Fluoro-orotic acid induces chromosome alterations in Candida albicans', *Yeast*, 22(1), pp. 57–70. doi: 10.1002/yea.1191.

Whitehead, N. (2001) *Quorum-sensing in Gram-negative bacteria*, *FEMS Microbiology Reviews*. doi: 10.1016/S0168-6445(01)00059-6.

Williams, D. R., Young, D. I. and Young, M. (1990) 'Conjugative plasmid transfer from Escherichia coli to Clostridium acetobutylicum', *Journal of General Microbiology*, 136(5), pp. 819–826. doi: 10.1099/00221287-136-5-819.

Williams, P. *et al.* (2007) 'Look who's talking: communication and quorum sensing in the bacterial world.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(1483), pp. 1119–1134. doi: 10.1098/rstb.2007.2039.

Wilson, C. M. *et al.* (2013) 'Global transcriptome analysis of Clostridium thermocellum ATCC 27405 during growth on dilute acid pretreated Populus and switchgrass.', *Biotechnology for biofuels*, 6(1), p. 179. doi: 10.1186/1754-6834-6-179.

Wood, H. G. (1952a) 'A study of carbon dioxide fixation by mass determination of the types of C13-acetate.', *The Journal of biological chemistry*, 194(2), pp. 905–31. doi: 10.1088/1751-8113/44/8/085201.

Wood, H. G. (1952b) 'Fermentation of 3, 4-C14-and 1-C14-labeled glucose by Clostridium thermoaceticum.', *Journal of Biological Chemistry*, (199.2), p. 579–583.

Woods, C. (2018) 'Manuscript in preparation'.

Wu, W. and Davis, R. W. (2016) 'One-pot bioconversion of algae biomass into terpenes for advanced biofuels and bioproducts', *Algal Research*. Elsevier B.V., 17, pp. 316–320. doi: 10.1016/j.algal.2016.05.005.

Wuster, A. and 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. doi: 10.1128/JB.01135-07.

Xiao, H. *et al.* (2011) 'Confirmation and elimination of xylose metabolism bottlenecks in glucose phosphoenolpyruvate-dependent phosphotransferase system-deficient clostridium acetobutylicum for simultaneous utilization of glucose, xylose, and arabinose', *Applied and Environmental Microbiology*, 77(22), pp. 7886–7895. doi: 10.1128/AEM.00644-11.

Xu, T. *et al.* (2015) 'Efficient genome editing in clostridium cellulolyticum via CRISPR-Cas9 nickase', *Applied and Environmental Microbiology*, 81(13), pp. 4423–4431. doi: 10.1128/AEM.00873-15.

Yang, S., Tang, I. and Okos, M. R. (1987) 'Kinetics of Homoacetic Fermentation of Lactate by Clostridium formicoaceticum Kinetics of Homoacetic Fermentation of Lactate by Clostridium formicoaceticum', 53(4), pp. 823–827.

Zahn, J. A. and Saxena, J. (2011) 'Novel Ethanologenic Clostridium species, Clostridium coskatii'. Google Patents.

Zhang, L. *et al.* (2002) 'Transmembrane topology of AgrB, the protein involved in the post-translational modification of AgrD in Staphylococcus aureus', *Journal of Biological Chemistry*, 277(38), pp. 34736–34742. doi: 10.1074/jbc.M205367200.

Zhang, L. and Ji, G. (2004) 'Identification of a staphylococcal AgrB segment(s) responsible for group-specific processing of AgrD by gene swapping', *Journal of Bacteriology*, 186(20), pp. 6706–6713. doi: 10.1128/JB.186.20.6706-6713.2004.

Zhang, Y., Grosse-Honebrink, A. and Minton, N. P. (2015) 'A universal mariner transposon system for forward genetic studies in the genus clostridium', *PLoS ONE*, 10(4), pp. 1–21. doi: 10.1371/journal.pone.0122411.

Zhao, L., Montville, T. J. and Schaffner, D. W. (2006) 'Evidence for quorum sensing in Clostridium botulinum 56A', *Letters in Applied Microbiology*, 42(1), pp. 54–58. doi: 10.1111/j.1472-765X.2005.01807.x.

Zheng, Y. N. *et al.* (2009) 'Problems with the microbial production of butanol', *Journal of Industrial Microbiology and Biotechnology*, 36(9), pp. 1127–1138. doi: 10.1007/s10295-009-0609-9.

Zhong, J., Karberg, M. and Lambowitz, A. M. (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. doi: 10.1093/nar/gkg248.

Zhou, A. *et al.* (2012) 'Functional characterization of Crp/Fnr-type global transcriptional regulators in Desulfovibrio vulgaris hildenborough', *Applied and Environmental Microbiology*, 78(4), pp. 1168–1177. doi: 10.1128/AEM.05666-11.

Zhu, Z. *et al.* (2015) 'Complete genome sequence of a malodorant-producing acetogen, Clostridium scatologenes ATCC 25775^{T}', *Journal of Biotechnology*, 212, pp. 19–20. doi: 10.1016/j.jbiotec.2015.07.013.

Zrimšek, P. *et al.* (2004) 'Spectrophotometric application of resazurin reduction assay to evaluate boar semen quality', *International Journal of Andrology*, 27(1), pp. 57–62. doi: 10.1046/j.0105-6263.2003.00448.x.