

**QssB: a pleiotropic RNPP-type
cell-cell communication
system in *C. acetobutylicum*
ATCC 824**

Oliver Jack Severn B.Sc.

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy
September 2017

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.

Oliver Jack Severn

September 2017

Abstract

Clostridium acetobutylicum is renowned for the ability to convert sugars into acids and solvents, including the potential biofuel butanol. However, regulation of its fermentation metabolism remains poorly understood, especially regarding the shift from acid to solvent metabolism. Several RNPP-type cell-cell communication systems have recently been discovered within *C. acetobutylicum* ATCC 824. The aim of this study was to investigate one such system, Quorum Sensing System B (QssB), consisting of a regulator, QsrB, and a signalling peptide of unknown structure derived from a precursor protein, QspB. The primary objectives were to characterise this system in greater detail, its effects on metabolism and endospore formation, and to gain insights into the nature of the signalling process. Overexpression of *qsrB* confirmed a dramatic reduction of solvent production and sporulation, and revealed increased production of acetate and butyrate, with considerably decreased consumption of glucose. These phenotypes were overcome by the addition of synthetically produced peptides matching the C-terminal region of the conjugate signalling peptide QspB. In doing so, the minimum peptide sequence required to relieve the effects of *qsrB* overexpression was shown to be the heptamer AEPTWGW. The system was shown to be highly selective for this sequence, with an alanine scan showing the sequential reduction in biological activity. QssB did not engage in cross talk with the other RNPP-type systems present within the organism, or with a homologous system discovered in *Clostridium roseum*, further evidence of selectivity. Attempts were made to identify the Specific Binding Protein (SBP) required for QspB uptake via ABC transport. ClosTron mutation of two SBP encoding genes, CA_0176 and CA_C3632, conferred resistance to the effects of QsrB overproduction on sporulation. An iTRAQ-based proteomics approach was used to study the wider effects of *qsrB* overexpression, as well as their alleviation via synthetic peptide. This analysis confirmed that the *qsrB* overexpressing strain remained in an acidogenic state which could be overcome through peptide addition. Although glucose consumption was decreased upon QsrB overproduction, key glycolysis enzymes were found to be upregulated. The stationary phase regulator CA_C0957 was found to be downregulated under these conditions, whereas proteins associated with motility and chemotaxis increased in abundance. GusA reporter assays supported the former, whilst swimming and swarming assays remained inconclusive. Enzyme assays

revealed increased acetate and butyrate kinase activity, agreeing with the acidogenic state suggested from the iTRAQ analysis. QsrB was confirmed to be a repressor of solvent production and the production of endospores, with new evidence indicating further complexity.

Acknowledgments

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

First and foremost, I would like thank God without whom I would not have made it this far.

I would like to thank Klaus Winzer for his exemplary supervision during the course of my studies, for all his advice, ideas, and support. I look forward to sharing a drink upon hand in.

Thank you to my second supervisor Alan Cockayne for all of his practical advice and experience.

The SBRC Nottingham have given me a bench and the equipment required, and I would like to thank all members past and present that have given me inspiration and a firm kick when needed.

Thank you to my 'bench supervisor' Zak Bean, and to Jon Humphreys, Ronja Breitkopf, Pawel Piatek and Alex Grosse-Honebrink. Without their contributions, this work would be much lesser. Your practical and social contributions to my PhD experience are not given justice here.

I would like to thank Singer Instruments for providing me with a placement in industry, and a fantastic insight into the wonderful world of robotics.

I thank my wife Emily Severn, for her considerable contribution to the work enclosed; through her constant support, understanding, and encouragement. Without your support, I would have dropped out of secondary school.

Finally, I would like to submit this thesis in memory of the 'Victorious' Victor Tudorica. We should have submitted together my friend.

Chapter 5: Effect of *qsrB* overexpression on the *C.acetobutylicum* ATCC 824 proteome Error! Bookmark not defined.

1. Introduction	16
1.1. The genus <i>Clostridium</i>	16
1.2. The solventogenic clostridia	17
1.2.1. Biphasic nature of ABE metabolism, energy conservation and regulation	19
1.3. Spore formation	21
1.4. A short history of industrial Acetone-Butanol-Ethanol (ABE) fermentation	24
1.5. Tools for the genetic engineering of <i>Clostridium acetobutylicum</i>	26
1.5.1. Insertional mutagenesis using ClosTron	26
1.5.2. In-frame deletion	28
1.5.3. Transposon mutagenesis	31
1.5.4. CRISPR	31
1.6. Bacterial cell-cell communication	32
1.6.1. Cell-cell communication in the Gram-negative organisms	33
1.6.2. Cell-cell communication in Gram-positive organisms	35
1.6.2.1. <i>Agr</i> type signalling	35
1.6.2.2. <i>RNPP</i> type signalling	38
1.7. iTRAQ proteomics	44
1.8. Aims and objectives	45
2 Materials and Methods	48
2.1 List of bacterial strains, plasmids, oligonucleotides and peptides	48
2.2 Microbiological methods	57
2.2.1 Media and growth conditions	57

2.2.2	Storage and revival of strains	59
2.2.3	Measurement of optical density (OD)	60
2.2.4	Initiation of growth experiments	60
2.3	Molecular biological methods	61
2.3.1	Isolation of plasmid DNA from <i>E. coli</i>	61
2.3.2	Isolation of genomic DNA from <i>C. acetobutylicum</i> ATCC 824 via column purification	61
2.3.3	Isolation of genomic DNA from <i>C. acetobutylicum</i> ATCC 824 via phenol chloroform extraction	61
2.3.4	Methods of DNA transfer	62
2.3.4.1	<i>Generation of competent cells and transformation of E. coli</i>	62
2.3.4.2	<i>Generation of competent cells and transformation of C. acetobutylicum ATCC 824.</i>	63
2.3.4.3	<i>Dialysis of DNA</i>	64
2.3.5	DNA manipulation and analysis	64
2.3.5.1	<i>Determination of quality and concentration of DNA samples</i>	64
2.3.5.2	<i>DNA amplification through polymerase chain reaction (PCR)</i>	65
2.3.5.3	<i>Restriction digests and Ligation</i>	65
2.3.5.4	<i>DNA gel electrophoresis and DNA recovery</i>	66
2.3.5.5	<i>Gibson assembly</i>	67
2.3.5.6	<i>DNA sequencing</i>	67
2.3.5.7	<i>Methylation of Plasmid DNA</i>	67
2.3.6	Protein expression and analysis	68
2.3.6.1	<i>Strep-tag purification of QsrB</i>	68
2.3.6.2	<i>SDS-PAGE gel electrophoresis</i>	69
2.4	Synthesis of DNA constructs and peptides	70

2.4.1	Generation of re-targeted ClosTron plasmids	70
2.4.2	Generation of synthetic peptides	71
2.5	<i>Genetic manipulation of C. acetobutylicum ATCC 824</i>	71
2.5.1	Isolation and confirmation of in-frame deletion mutants	71
2.5.2	Isolation and confirmation of ClosTron mutants	71
2.6	<i>Phenotypic analysis of C. acetobutylicum ATCC 824</i>	72
2.6.1	Preparation of <i>C. acetobutylicum</i> ATCC 824 cultures for phenotypic analysis	72
2.6.2	Light microscopy	72
2.6.3	Granulose staining and screening for amylase activity	73
2.6.4	Determination of heat-resistant colony-forming units (CFU)	74
2.6.5	Gas chromatography	74
2.6.6	High performance liquid chromatography of glucose consumption	75
2.6.7	iTRAQ proteomics sample preparation	76
2.6.8	GusA reporter assay	77
2.6.9	Attempted swimming and swarming assays	78
2.7	<i>Determination of enzyme activities</i>	78
2.7.1	Generation of crude lysate via sonication	78
2.7.2	Hydroxamate assay of acetate and butyrate kinase activity	79
2.7.3	Butanol dehydrogenase assay	79
2.7.4	Nitroreductase assay	80
2.8	<i>In silico software and computer tools</i>	80
3	Characterisation of the RNPP type system CA_C0324 (QsrB) and its conjugate peptide (QspB)	83
3.1	Introduction	83
3.2	Aim of the study	86
3.3	Results	87

3.3.1	Creation of a clean in-frame deletion of the <i>qsrB</i> gene	87
3.3.2	Elucidating the minimum QspB peptide length required for complementation of the pMTL85143- <i>qsrB</i> phenotype	89
3.3.3	Investigating the dose response of QspB 7 as the conjugate peptide of QsrB	102
3.3.4	Investigating the ability for QssB of <i>C. acetobutylicum</i> ATCC 824 to engage in interspecies cross talk, using the synthetic QspB 9 peptide from <i>C. roseum</i>	105
3.3.5	Identification of QspB 7 amino acid positions required for biological activity.	107
3.3.6	Investigation of <i>C. acetobutylicum</i> ATCC 824 RNPP-type system cross talk	112
3.3.7	Supplementation of <i>C. acetobutylicum</i> ATCC 824 wild type with synthetic QspB peptides in order to assess its ability to affect solvent yield	114
3.4	Discussion	116
4	The insertional mutation of the specific binding proteins of the oligopeptide permease systems within <i>C. acetobutylicum</i> ATCC 824	123
4.1	Introduction	123
4.2	Objectives	127
4.3	Results	128
4.3.1	Bioinformatic identification of putative oligopeptide permeases in <i>C. acetobutylicum</i> ATCC 824	128
4.3.2	Insertional inactivation of the identified oligopermease SPBs via the ClosTron method	132
4.3.3	Analysis of granulose formation within the SBP insertional mutations	133
4.3.4	Analysis of fermentation products formed by SBP mutants	137
4.3.5	Introduction of the pMTL85143- <i>qsrB</i> overexpression plasmid into the SBP insertional mutation strains	142

4.4	Discussion	150
5	Effect of <i>qsrB</i> overexpression on the <i>C. acetobutylicum</i> ATCC 824 proteome	159
5.1	Introduction	159
5.2	Objectives	162
5.3	Results	163
5.3.1	Generation of proteomics data	163
5.3.1.1	Detection of QsrB within the proteomic data	166
5.3.1.2	Glycolysis enzymes	168
5.3.1.3	ABE fermentation pathway	168
5.3.1.4	Components of the sporulation pathway	169
5.3.1.5	Oligopermease SBPs	170
5.3.1.6	Proteins relating to motility	170
5.3.1.7	Stationary phase regulators	171
5.3.2	Analysis of enzymatic activity	171
5.3.3	GusA reporter analysis of significant promoters distinguished from phenotypic and proteomic analyses	173
5.3.4	Swimming and swarming assays of the <i>C. acetobutylicum</i> ATCC 824 pMTL85143 and <i>C. acetobutylicum</i> ATCC 824 pMTL85143- <i>qsrB</i> strains	177
5.3.5	Attempted tagging and isolation of the QsrB regulator	179
5.4	Discussion	180
6	General discussion and further research	186
6.1	Background	186
6.2	Understanding the QspB signal peptide	186
6.2.1	Key findings of this study	186
6.2.2	Further research and summary	189
6.3	Oligopeptide permease uptake systems and QssB in <i>C. acetobutylicum</i> ATCC 824	191

6.3.1	Key findings of this study	191
6.3.2	Further research and summary	193
6.4	Proteomic analysis of QsrB over production	195
6.4.1	Key findings of this study	195
6.4.2	Further research and summary	197
6.5	General Conclusion	200
7	Bibliography	203
	Appendix	221

List of figures:

Chapter 1: Introduction

- Figure 1.1 A schematic diagram of the ABE pathway. 18
- Figure 1.2 Diagram displaying the molecular action of the ClosTron technique. 27
- Figure 1.3 Diagram showing the genetic mechanism of inframe deletion technique and associated protocol. 29
- Figure 1.4 A schematic diagram of the Agr quorum sensing system in *S. aureus*. 32
- Figure 1.5 General scheme of RNPP-type cell-cell communication. 36
- Figure 1.6 Schematic representation of the 8 RNPP type cell-cell communication systems that have been identified (A), and the alignment of the putative signalling precursor proteins 29

Chapter 3: Characterisation of the RNPP type system CA_C0324 (QsrB) and its conjugate peptide (QspB)

- Figure 3.1 Effects of *qsrB* overexpression on fermentation profile (A) and endospore formation (B). 91
- Figure 3.2 Sequence of QspB and derived synthetic peptides. 92
- Figure 3.3 Butanol production of *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* after 120 hours of batch culture fermentation, with addition of 10 μ M, 1 μ M and 100 nM of QspB 9, respectively. 93
- Figure 3.4 Butanol production and endospore formation by *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* in the presence of synthetic QspB peptides. 96
- Figure 3.5 Biologically active (green) and inactive (red) QspB peptide variants. 99
- Figure 3.6 Formation of solvents (A), organic acids (B) and endospores (C) by *C. acetobutylicum* ATCC 824 pMTL85143*qsrB* supplemented with synthetic QspB-derived peptides of varying lengths. 100
- Figure 3.7 Dose response relationship for the Qsp7 peptide. 103
- Figure 3.8 Comparison of the QspB 9 amino acid sequence from *C. acetobutylicum* ATCC 824 and the homologous sequence from *C. roseum*. 106
- Figure 3.9 Butanol production of *C. acetobutylicum* pMTL85143-*qsrB* in the presence of Qsp9 peptides from *C. roseum* and *C. acetobutylicum* ATCC 824. 107
- Figure 3.10 Biological activity of QspB 7 variants generated by alanine scanning. 110

Figure 3.11 Synthetic peptides derived from other native RNPP type systems did not activate solvent formation by *C. acetobutylicum* ATCC 824 pMTL85143-qsrB. 113

Figure 3.12 Butanol production of *C. acetobutylicum* ATCC 824 wild type upon addition of QspB 9 roseum, QspB 10, QspB 9, QspB 8, QspB 7 and DMSO. 115

Chapter 4: The insertional mutation of the specific binding proteins of the oligopeptide permease systems within *C. acetobutylicum* ATCC 824

Figure 4.1 Composition (A) and operon organisation (B) of a typical Gram positive oligopeptide ABC transporter complex, as exemplified by the Opp system. 124

Figure 4.2 Arrangement of the CA_C0176, CA_C2377, CA_C2667 and CA_C3179 genes within the *C. acetobutylicum* ATCC 824 genome. 130

Figure 4.3 Genetic layout of CA_C3632, CA_C3634 and CA_C3640 within the *C. acetobutylicum* ATCC 824 genome. 131

Figure 4.4 Lack of granule formation and amylase activity within the *C. acetobutylicum* ATCC 824 CA_C2667::CTermB strain. 135

Figure 4.5 Lack of granule formation with retained amylase activity within the *C. acetobutylicum* ATCC 824 CA_C3632::CTermB strain. 136

Figure 4.6 Analysis of solvent formation after over 120 hours by each of the independent SBP mutants. 139

Figure 4.7 Heat resistant CFU/ml of each SBP insertional mutation after 120 hours of fermentation. 141

Figure 4.8 Acetone and Butanol production of SBP insertional mutations, containing either pMTL85143 empty vector control, pMTL85143-qsrB or pMTL85143-qsrB supplemented with 10 μ M QspB. 143

Figure 4.9 Acid of SBP mutants, containing the pMTL85143 control plasmid, the pMTL85143-qsrB plasmid or the pMTL85143-qsrB plasmid with 10 μ M QspB 7 addition.. 135

Figure 4.10 Formation of heat resistant CFU/ml for SBP mutants containing the pMTL85143 empty vector control, pMTL85143-qsrB and pMTL85143-qsrB with 10 μ M QspB 7 addition. 147

Figure 4.11 Glucose consumption by SBP mutants containing empty pMTL85143 vector, pMTL85143-qsrB, and pMTL85143-qsrB with 10 μ M QspB addition. 148

Figure 4.12 Product profile of the reisolated CA_C2377::CTermB mutant strains. 150

Chapter 5: Effect of qsrB overexpression on the C.acetobutylicum ATCC 824 proteome

Figure 5.1 Fermentation products of C. acetobutylicum ATCC 824 pMTL85143 (black), C. acetobutylicum ATCC 824 pMTL85143-qsrB (light grey) and C. acetobutylicum ATCC 824 pMTL85143-qsrB with 10 μ M QspB 7 (dark grey) at the 24 hour time point of preparing samples for proteomic analysis. 164

Figure 5.2 Changes in ABE fermentation pathway enzymes as revealed by iTRAQ analysis for C. acetobutylicum ATCC 824 pMTL85143-qsrB, C. acetobutylicum ATCC 824 pMTL85143 vector control, and C. acetobutylicum ATCC 824 pMTL85143-qsrB with 10 μ M QspB 7 addition. 167

Figure 5.3 Enzymatic activity of Acetate Kinase, Butyrate Kinase, Nitroreductase and Alcohol Dehydrogenase after 24 hours of fermentation in crude extracts of C. acetobutylicum ATCC 824 pMTL85143 (black), C. acetobutylicum ATCC 824 pMTL85143-qsrB (vertical hashed lines) and C. acetobutylicum ATCC 824 pMTL85143-qsrB with 10 μ M QspB 7 addition (angled hashed lines). 172

Figure 5.4 Promoter activity during in C. acetobutylicum ATCC 824 WT (black) and C. acetobutylicum ATCC 824 CA_C0324::CTermB (grey) as determined by transcriptional gusA reporters. 176

Figure 5.5 Swarming assay of the C. acetobutylicum ATCC 824 pMTL85143 and C. acetobutylicum ATCC 824 pMTL85143-qsrB. 178

List of Tables

Chapter 2: Material and Methods

Table 2.1	Bacterial strains used in this study	49
Table 2.2	Plasmids used in this study	52
Table 2.3	Oligonucleotides used in this study	54
Table 2.4	Synthetic peptides used in this study	56
Table 2.5	CGM recipe	57
Table 2.6	CBM-S recipe	58

Chapter 3: Characterisation of the RNPP type system CA_C0324 (QsrB) and its conjugate peptide (QspB)

Table 3.1	QspB 7 peptide variants used in alanine scanning experiments	108
Table 3.2	Biological activity of Qsp 7 alanine scan variants	112

Chapter 4: The insertional mutation of the specific binding proteins of the oligopeptide permease systems within *C. acetobutylicum* ATCC 824

Table 4.1	Putative oligopeptide-specific SNPs identified in <i>C. acetobutylicum</i> ATCC 824	128
Table 4.2	Production of granulose and extracellular amylase activity by SBP mutants	133

1. Introduction

1. Introduction

1.1. The genus *Clostridium*

The work here focuses on the Gram-positive, strictly anaerobic, solventogenic bacterium *Clostridium acetobutylicum* ATCC 824. This type strain was isolated from agricultural soil in 1924 (Nölling et al., 2001b). The taxonomy of this species is as follows: It belongs to the phylum *Firmicutes*, class *Clostridia*, order *Clostridiales*, Family *Clostridiaceae*, and genus *Clostridium*. This genus is not considered to be monophyletic and hence many of its species require reclassification. However, *C. acetobutylicum* ATCC 824 is closely related to the type species of the genus, *Clostridium butyricum*, and thus a member of what has been called the genus ‘*Clostridium sensu stricto*’. The latter, while monophyletic, is phenotypically diverse, with members occupying a wide range of ecological niches and habitats and possessing very different physiological and metabolic traits, many of which are of interest, or concern, for humans. For instance, ancient acetogens such as *C. ljungdahlii* and *C. autoethanogenum*, utilise the Wood-Ljungdahl pathway for converting carbon monoxide, carbon dioxide and hydrogen to acetyl-CoA, which in turn is used for ATP generation through the production of acetate and ethanol (Schuchmann and Müller, 2014). Waste gases, or the production of syngas can be used as a feedstock for these organisms for the production of high value chemicals, and as such they are now widely regarded as organisms of industrial importance (Liew et al., 2016).

Clinically relevant *Clostridium* species include the infamous pathogens *C. perfringens*, the causative agent of gas gangrene wound infections and *C. botulinum*, producer of a powerful secreted neurotoxin which is the causative agent of botulism (Rossetto et al., 2014; Shindo et al., 2015). While *C. perfringens* fits firmly within the proteolytic phenotypic group, *C. botulinum* can exhibit either proteolytic or non-proteolytic (saccharolytic) phenotypes, a possible artefact of its taxonomic definition relying on the presence of the botulinum neurotoxin (Marshall et al., 2014; Rossetto et al., 2014). The solventogenic *Clostridium* spp. such as *C. acetobutylicum* ATCC 824 and *C. beijerinckii* are capable of producing the solvents acetone, butanol and ethanol through the ABE pathway. *C. acetobutylicum* ATCC 824, as a recognised model of the ABE fermentation pathway will form the focus of this study (George et al., 1983;

González-Peñas et al., 2014; Lee et al., 2015). Historically, the ABE pathway has been used to industrially produce acetone and butanol from a variety of feed stocks, which will be covered in further detail in heading 1.4 and 1.5 (Lütke-Eversloh, 2014; Tracy et al., 2012).

1.2. The solventogenic clostridia

The solventogenic clostridia are loosely defined as strictly anaerobic endospore forming bacteria that produce a variety of alcohol and other solvents by anaerobically degrading simple and complex carbohydrates (Papoutsakis, 2008). Solventogenic *Clostridia* are therefore chemoorganoheterotrophs capable of growth on a wide range of organic substrates (Tracy et al., 2012). Generally, the term refers to *Clostridium* spp. that engage in the production of acetone, butanol and ethanol via the ABE fermentation pathway, a schematic of which is shown in Figure 1.1. (Dürre, 2005; Jones and Woods, 1986)

The ABE fermentation pathway begins with pyruvate, the final product of glycolysis, at which point there has been a net generation of ATP and NADH. The fermentation initiates with the production of acetic acid and butyric acid, as well as carbon dioxide and hydrogen, a stage of growth referred to as acidogenesis (Jones, 2001). Under certain conditions lactic acid can be formed, too. The process has previously been linked with exponential phase of the organism, and in performing acidogenesis a maximal amount of approximately 3.4 ATP is generated per molecule of hexose substrate (Jones and Woods, 1986). The generation of these acids causes the external pH to drop, and the organism risks the dissipation of the proton motive force as undissociated acids begin to diffuse back into the cell (Jones and Woods, 1986). Simultaneously, the population density has increased and the bacterium enters stationary phase, and solventogenesis begins (Dürre, 2005; Jones and Woods, 1986).

During batch culture fermentation the observational concurrence of the initiation of stationary phase and the onset of solventogenesis would suggest the two are linked. However, it is now deemed that although the two share regulatory mechanisms, such as control by the master regulator Spo0A, they may not be linked directly (Kumar et al., 2014; Millat and Winzer, 2017). Substrate consumption continues during

solventogenesis, and acids are partially re-assimilated (Jones, 2001; Jones and Woods, 1986). During solventogenesis ethanol is produced from acetyl-CoA, and butanol is generated from butyryl-CoA, whilst acetone is produced from acetoacetyl-CoA in a reaction that is dependent on acid re-assimilation (Dürre, 2005; Jones and Woods, 1986). It should be noted, that also during the phase of solventogenesis the complicated machinery of endospore formation is initiated, as is explained under heading 1.3 (Alsaker et al., 2005).

C. acetobutylicum ATCC 824 is considered the model organism for solventogenic clostridia, and between itself and *C. beijerinckii* represents the majority of published understanding of ABE fermentation (Millat and Winzer, 2017). That is not to say deviations from the scheme given in Figure 1.1 do not exist, as even strains of *C. beijerinckii* have been shown to further reduce acetone to isopropanol (George et al., 1983). *C. aurantibutyricum* has been shown to produce both acetone and isopropanol (George et al., 1983).

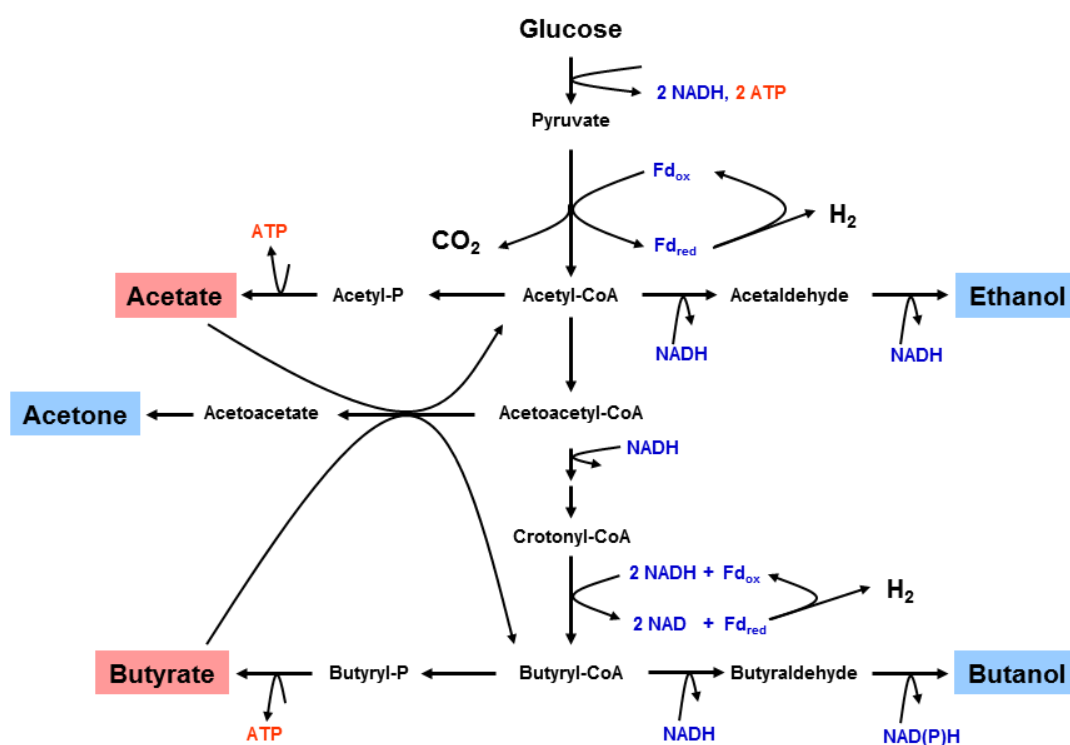


Figure 1.1 **A schematic diagram of the ABE pathway.** The initial step from Glucose to Pyruvate represents glycolysis, 2 pyruvate are generated, but due to downstream branching of the pathway, the figure only displays stoichiometries for individual reactions and not the entire pathway. Reactions producing ATP (red) and using or producing reducing equivalents (blue) are shown, together with final intermediates and final fermentation products, acids (red underlay) and solvents (blue underlay).

Other alterations include *C. tetanomorphum* which does not produce acetone, and *C. puniceum* which produced butanol in a much greater ratio to acetone and also during logarithmic growth (George et al., 1983; Gottwald et al., 1984; Holt et al., 1988; Lund et al., 1981). While the species mentioned here are by no means conclusive of the variations in ABE fermentation, as the solventogenic clostridia appear versatile in nature, it does show the broad spectrum of such organisms. *C. acetobutylicum* ATCC 824 is an appealing industrial strain due to its intrinsic ability to utilize a range of carbon sources (Abd-Alla et al., 2015; H. Li et al., 2014; Madihah et al., 2001; Naik et al., 2010). As expected, catabolite repression plays a role in carbon utilisation, and several catabolic operons within the organism have been studied in response to carbohydrates (Servinsky et al., 2010). Key components of catabolite repression within *C. acetobutylicum* ATCC 824 is the catabolite control protein, CcpA, a pleiotropic regulator involved in glucose repression of alternative carbon source utilization (Ren et al., 2012, 2010). It should be noted that under the batch culture conditions of this study (2.2.1) *C. acetobutylicum* doubled once every hour. The exponential phase of the organism would continue for 12-14 hours to an OD 600 of 4, upon which the culture enters stationary phase. As previously mentioned the master regulator of stationary phase characteristics is the well-studied Spo0A, the state of which will be commented on under the next heading (Harris et al., 2002; Steiner et al., 2011).

1.2.1. Biphasic nature of ABE metabolism, energy conservation and regulation

Batch fermentation utilising the ABE pathway is often understood as a biphasic metabolism (Jang et al., 2012). The mechanism involved in switching the metabolism of the organism from acidogenesis to solventogenesis is still far from understood. During continuous culture of *C. acetobutylicum* ATCC 824, acidogenesis and solventogenesis can be maintained by the alteration of chemostatic conditions (Bahl et al., 1983, 1982; Girbal and Soucaille, 1998). By maintaining a neutral pH and growing on glucose acidogenesis is maintained, and by lowering the pH a solventogenic culture is promoted (Girbal and Soucaille, 1998). However, the picture appears more complex than this, given that accumulation of un-dissociated acids alone appears to effect the occurrence of 'acid crash' and 'acidogenic fermentations' in

batch culture experiments, the latter which can be avoided by simply increasing the starting glucose concentration or lowering concentration of yeast extract in the media (Maddox et al., 2000).

‘Acid crash’ is a phenomenon seen in batch culture or poorly maintained continuous culture, where upon the pH drop of acidogenesis, the organism fails to enter solventogenesis and instead perishes presumably due to the collapse of the proton motive force (Jones and Woods, 1986; Maddox et al., 2000). An ‘acidogenic fermentation’ as the name suggests is the failure of the organism to enter solventogenesis and instead continue to produce acids while maintaining viability, it can occur in batch cultures that are initiated at around neutral pH, and can be maintained in continuous cultures as previously commented (Girbal and Soucaille, 1998; Maddox et al., 2000). Interestingly, the choice of the organism to remain in acidogenesis or begin solventogenesis also has been shown to rely on the availability or need for ATP (Girbal and Soucaille, 1998). As previously discussed, acidogenesis allows the organism to produce the maximal amount of ATP per molecule of substrate (Jones and Woods, 1986). It is not a surprise then that low ATP availability leads to sustained acidogenesis, carbon limitation is an example of this (Girbal and Soucaille, 1998). In contrast, high ATP availability favours the production of solvents, for example during periods of sufficient carbon and lower growth rates (Girbal and Soucaille, 1998).

Energy conservation strategies have recently been shown to differ slightly within the solventogenic *Clostridia*. While all solventogenic *Clostridia* can synthesise ATP by substrate level phosphorylation during glycolysis; *C. beijerinckii*, *C. puniceum*, *C. saccharobutylicum* and *C. sacchroperbutylaceticum* have also been shown to contain *rnf* genes, encoding a membrane-bound complex capable of reducing ferredoxin and producing NADH, and in doing so generating an ion gradient which can be utilised by ATPase (Poehlein et al., 2017). Such a process is not present in the ABE fermentation model organism *C. acetobutylicum* ATCC 824, a concept that should be considered moving forward (Poehlein et al., 2017).

Slow progress is also being made in unravelling the regulatory mechanisms that govern acid and solvent metabolism, as well as sporulation. The master regulator Spo0A is known to be required for high solvent production in solventogenic *Clostridium* spp., although small amounts are still produced in its absence, and is also essential for the initiation of sporulation (Harris et al., 2002; Ravagnani et al., 2000). Other regulators proposed to take part in the regulation of solventogenesis in *C. acetobutylicum* ATCC 824 are the global regulator CodY and the catabolite control protein A (CcpA) (Poehlein et al., 2017; Ren et al., 2012). The latter positively regulates the *sol* operon, encoding key genes required for acetone and butanol formation.

Transcription of several solvent genes is activated by the binding of phosphorylated Spo0A to specific sites, 0A boxes, which are present in the promoter regions of these genes (Harris et al., 2002; Thormann et al., 2002). The phosphorylation state of Spo0A in turn is dependent on the activity of three orphan histidine kinases as well as an intracellular kinase-like protein which acts as a phosphatase (Steiner et al., 2011). Unfortunately, none of the signals that activate or inhibit these proteins is currently known. Thus, while general environmental conditions for acid and solvent formation are well established, and some of the underlying regulatory components have been identified, we are still largely ignorant of the precise chemical and physical nature of the signals which ultimately control the expression of solvent forming pathways.

1.3. Spore formation

The genus *Clostridium* is renowned for its ability to produce endospores, the hardiest form of life on earth, surviving for hundreds if not thousands of years, and even coping with the intense radiation in the vacuum of space (Horneck et al., 1994; Nicholson, 2003). Much of what is known of the endospore formation process was elucidated from *Bacillus* spp. and some have suggested that the research into the endospore formation of *Bacillus anthracis* in the 1870's marks the beginning of the modern age of microbiology (Yung and Ponce, 2008). In order to begin the process of forming an endospore, initially the cell undergoes a form of asymmetric division, in which a small membrane bound section named the forespore, and a larger membrane bound section termed the mother cell are formed (Errington, 2001). This process is

named septation, due to the membrane separating the two compartments being referred to as the septa (Pogliano et al., 1999). The process of sporulation is the focus of intense research, and new insights into the process are fertile ground for publication (Gamba et al., 2016).

What is relatively well founded is the understanding of the sporulation cascade of sigma factors. The formation of an endospore comes about through the sequential expression of sigma factors, allowing for different waves of transcription and hence protein expression (Dürre and Hollergschwandner, 2004). In order, the expressed sigma factors required for the generation of a spore were shown in *B. subtilis* as σ^F , σ^E , σ^G and finally σ^K (Al-Hinai et al., 2015; Dürre, 2014; Dürre and Hollergschwandner, 2004). Studies into the homologous sigma factors of *C. acetobutylicum* ATCC 824 showed the same sequential expression before the production of spore, and hence this is taken as strong evidence that the mechanistic process of spore formation is very similar to that of *B. subtilis*, with only minor differences (Santangelo et al., 1998). Of the four sigma factors; σ^F and σ^G are active in the forespore regulated by anti-sigma factors and anti-anti-sigma factors, whereas σ^E and σ^K are present as a pro-form in the mother cell where they are activated via proteolysis (Al-Hinai et al., 2015; Dürre, 2014; Dürre and Hollergschwandner, 2004).

Following the formation of the forespore and the mother cell, the mother cell engulfs the forespore, and in doing so causes it to be encapsulated by two membranes (Al-Hinai et al., 2015; Dürre, 2014). The chromosome is packed within forespore and differing actions take place in each relating to the difference in sigma factors (Dürre and Hollergschwandner, 2004). Two layers of peptidoglycan form in between the two membranes of the forespore, the cortex and the germ cell wall (Dürre and Hollergschwandner, 2004; Leggett et al., 2012). Multiple layers of protein are layered around the outside of the forespore; this is the spore core and is generated by the action of the mother cell (Dürre and Hollergschwandner, 2004; Leggett et al., 2012). The forespore contains a high concentration of dipicolinic acid, which acts a chelating agent for divalent cations and is implicated in the spores heat resistance, as well as 3-phosphoglycerate which is required for an ATP source during germination (Al-Hinai et al., 2015; Dürre, 2014; Dürre and Hollergschwandner, 2004). Small acid soluble

proteins (SASPs) bind the chromosome into a protective format, leading to resistance to UV, heat and other radiation damage (Raju et al., 2006; Wetzal and Fischer, 2015). The SASPs can also be broken down upon germination in order to provide nutrition to the newly vegetative cell (Dürre and Hollergschwandner, 2004). The endospore is finally released via the lysis of the mother cell.

The difference between spore formation in *B. subtilis* and *C. acetobutylicum* ATCC 824 occurs before the σ -factor cascade, in the phosphorylation of Spo0A (Al-Hinai et al., 2015; Dürre, 2014; Dürre and Hollergschwandner, 2004). Spo0A is often referred to as a master regulator in *C. acetobutylicum* ATCC 824, given its role in initiating sporulation, but also in the production of butanol and acetone (Steiner et al., 2011). In *Bacillus* spp. Spo0A is phosphorylated via a phosphorelay system which is conspicuously absent in *Clostridium* species (Al-Hinai et al., 2015; Dürre, 2014; Dürre and Hollergschwandner, 2004; Steiner et al., 2011). The difference between the two in regards to this may also explain their relative difference in sporulation conditions, in contrast to *Bacillus* spp. starvation is not an affirmative signal of sporulation in *Clostridium* spp., instead solventogenic clostridia may rely on sporulation to withstand the toxic build-up of solvents in the local area (Dürre and Hollergschwandner, 2004). Regardless, Spo0A in *C. acetobutylicum* ATCC 824 was shown to reach a phosphorylated state via several orphan histidine kinase genes, from this it was hypothesised that the organism represented a proto spo0A activation system, and through evolution the phosphorelay components were gathered to represent the system seen in *Bacillus* spp. (Steiner et al., 2011). There are other minor differences in the transcription levels of a few of the sigma factors and transcriptional regulators of the sporulation system, but such detail is beyond the scope of this study (Dürre and Hollergschwandner, 2004).

It is noteworthy that the endospore formation in selected members of the genera *Bacillus* and *Clostridium* has been shown to be subject to regulation by cell-cell communication (Peng et al., 2015; Perchat et al., 2016a, 2011; Rocha-Estrada et al., 2010; Steiner et al., 2012).

1.4. A short history of industrial Acetone-Butanol-Ethanol (ABE) fermentation

The history of ABE fermentation is intertwined with that of Chaim Weizmann, the first president of Israel, who settled in England in 1904, when he took up an appointment at the University of Manchester (Meltzer, 2010). Weizmann was a chemist, but after training himself as a microbiologist he isolated a strain that produced butanol and acetone from many starchy compounds, which he called BY (Jones and Woods, 1986). The organism was later named *Clostridium acetobutylicum*, and Weizmann's interest in it was due to the formation of butanol, that many believed would be a key product in the production of synthetic rubber (Jones and Woods, 1986).

During World War I (WWI), the British munitions industry was in great need of acetone for the manufacture of cordite, a smokeless gun powder (Jones and Woods, 1986). Due to cessation of trading, there was a severe shortage in calcium acetate, imported to Britain for the chemical manufacture of acetone. Production by bacterial fermentation was investigated as an alternative and Weizmann's organism chosen due to superior acetone yields from a wide variety of feedstocks (Weizmann, 1915) thus leading to industrial scale production first in Britain and then, following feedstock supply issues, North America (Jones and Woods, 1986). Following WWI, biological production continued, driven by the booming automobile industry's demand for mainly butanol for the production of fast-drying lacquers. Starch as the feedstock of the fermentation was replaced by the cheaper molasses, requiring the isolation of new solventogenic strains. During WWII, the fermentation served again to support war efforts, for instance for Japan which suffered from vulnerability in terms of access to aviation fuel (Jones, 2014). Following WW2, industrial ABE fermentation went into rapid decline in most parts of the world as solvents could be produced more cheaply from petroleum, with exception of countries such as Russia, China and South Africa, the latter suffering from oil embargos. Over the last decades interest in a revival of the process has been mounting, driven largely by a desire for a sustainable process based on renewable resources. In particular, China has been investing considerably into increasing its capacity (Ni and Sun, 2009) and numerous companies have set out to

develop economical process solutions. However, current efforts stagnate due to low oil prices.

Green Biologics (Oxon, UK) is one such company attempting to produce biobutanol and acetone in a manner competitive with the petrochemical industry (www.greenbiologics.com). They are currently attempting to produce ABE products from a retro fitted ethanol pilot plant in Kansas USA (www.greenbiologics.com). Ethanol has a maximal blending potential of around 10%, above this value the fuel mixture becomes corrosive to unmodified engines, and the combustion potential is reduced (Nigam and Singh, 2011). Butanol does not suffer the same problem, and is able to be used at an 85% blend with petrochemical fuel without engine modification (Nigam and Singh, 2011). It also has the benefit of having a much lower vapour pressure than both conventional petrochemical fuel and ethanol, making it safer to ship and handle (Nigam and Singh, 2011).

Originally, corn was used as the main feedstock for ABE fermentation and is still used for the generation of biofuel, however, use of arable food crops for the generation of fuel has sparked the beginning of what has been called the food vs fuel debate (Tenenbaum, 2008). Allegedly, the increased use of arable land mass for the production of crops intended for fuel use has contributed to an increase in global food prices, having a detrimental effect on the world's poor (Tenenbaum, 2008). While such crops and land are still used for the production of biofuel, much research is devoted to utilisation of non-edible plant matter and preferably those that can grow on previously unused marginal land (Menon and Rao, 2012; Pfomm et al., 2010). Due to the wide range of feedstocks it is able to utilise, *C. acetobutylicum* ATCC 824 and other saccharolytic *Clostridium spp.* are very useful as biofuel producers (H. Li et al. 2014; J. Li et al. 2014; Abd-Alla & Elsadek El-Enany 2012; Abd-Alla et al. 2015; Jones & Woods 1986). In fact, *C. acetobutylicum* ATCC 824 contains a non-functional cellulosome, an apparatus of enzymes required for the breakdown and subsequent fermentation of lignocellulosic material (Parsiegla et al., 2008). Metabolic engineering work has been initiated in an attempt to allow *C. acetobutylicum* ATCC 824 to break down such material (Kovács et al., 2013; Willson et al., 2016). Furthermore, considerable efforts have been made to investigate and optimise

production of butanol from pre-treated lignocellulosic materials (Qureshi et al., 2013). Thus, in the future, butanol and other solvents may be derived in a more sustainable manner from the fermentation of lignocellulosic waste.

1.5. Tools for the genetic engineering of *Clostridium acetobutylicum*

The genus *Clostridium* is filled with species that are recalcitrant to transformation and, thus, genetic modification (Pyne et al., 2014). It is rare for a publication detailing the genetic engineering of a new strain or species of solventogenic clostridia, not to dictate the intricate changes needed in order to increase its transformation efficiency, and also alter each genetic tool to its specific requirements (Schwarz et al., 2017). This includes *C. acetobutylicum* ATCC 824 which contains type II restriction systems (CA_C1502 and CA_C3535), that require specific methylation of transformed DNA to prevent its breakdown (Croux et al., 2016; Papoutsakis and Mermelstein, 1993). The development of genetic tools for *Clostridium spp.* has lagged behind other organisms, hindering genetic engineering endeavours. The genome of *C. acetobutylicum* ATCC 824 was published in 2001 by Nölling and colleagues, where it was discovered the organism contained a 3.94 mbp chromosome and interestingly a 192 kbp megaplasmid (Nölling et al., 2001b). The megaplasmid itself contains the majority of the genes required for the production of solvents, its maintenance is essential for solventogenesis and also the formation of endospores (Nölling et al., 2001b). *C. acetobutylicum* can lose its megaplasmid, which is one known cause of degeneration, the loss of solvent production and spores. A common method for checking megaplasmid retention is iodine staining for the production of granulose, or observing starch degradation, as an integral amylase gene is encoded on the megaplasmid (Nölling et al., 2001a; Steiner et al., 2011). These methods are covered in 2.6.3 and utilised in Chapter 4. This section will shed light on the tools available for the manipulation of the *C. acetobutylicum* ATCC 824 genome.

1.5.1. Insertional mutagenesis using ClosTron

A big step forward towards targeted gene disruption in *Clostridium spp.* was the development of the ClosTron technique, an expedient method that despite alternatives is still widely used today (Liew et al., 2017; Liu et al., 2016). This technique is an

extension of an earlier methodology termed TargeTron, which in turn is based on a derivative of the *Lactococcus lactis* group II mobile intron Ll. LtrB that had been repurposed for defined, targeted gene insertion (Karberg et al., 2001). This intron functions by encoding a ribonucleoprotein (RNP) complex, within which a section of the RNA targets the intron to a specific location within the chromosome (Perutka et al., 2004). At this site, the RNP complex causes insertion of one strand of the RNA intron into the chromosome and then uses its inherent reverse transcriptase activity to catalyse the formation of the opposing strand using a cleaved section of the RNA intron as a 3' primer; this process is termed retrohoming (Zhong et al., 2003). Without selection, the insertional disruption frequency of a target gene using the TargeTron methodology was 0.1-1% in *Escherichia coli*, a process that was improved upon by the addition of a retrotransposon-activated-marker (RAM) (Zhong et al., 2003).

The RAM functions by encoding a selectable marker under control of its own promoter in the reverse orientation within the intron. This selectable marker is itself disrupted by a self-splicing group I intron in the forward orientation. During the retrohoming activity the group I intron splices out of the RNA sequence, which leads to an active selectable marker after DNA integration (Zhong et al., 2003). For the purpose of using this methodology within *C. acetobutylicum* ATCC 824 two major changes were introduced. The retro-transposition activating marker (RAM) system was modified to consist of an *ermB* cassette conferring erythromycin resistance. Furthermore, in its most advanced version the mobile group II intron relies on an IPTG inducible T7 promoter for intron expression, which is not functional in *Clostridium*

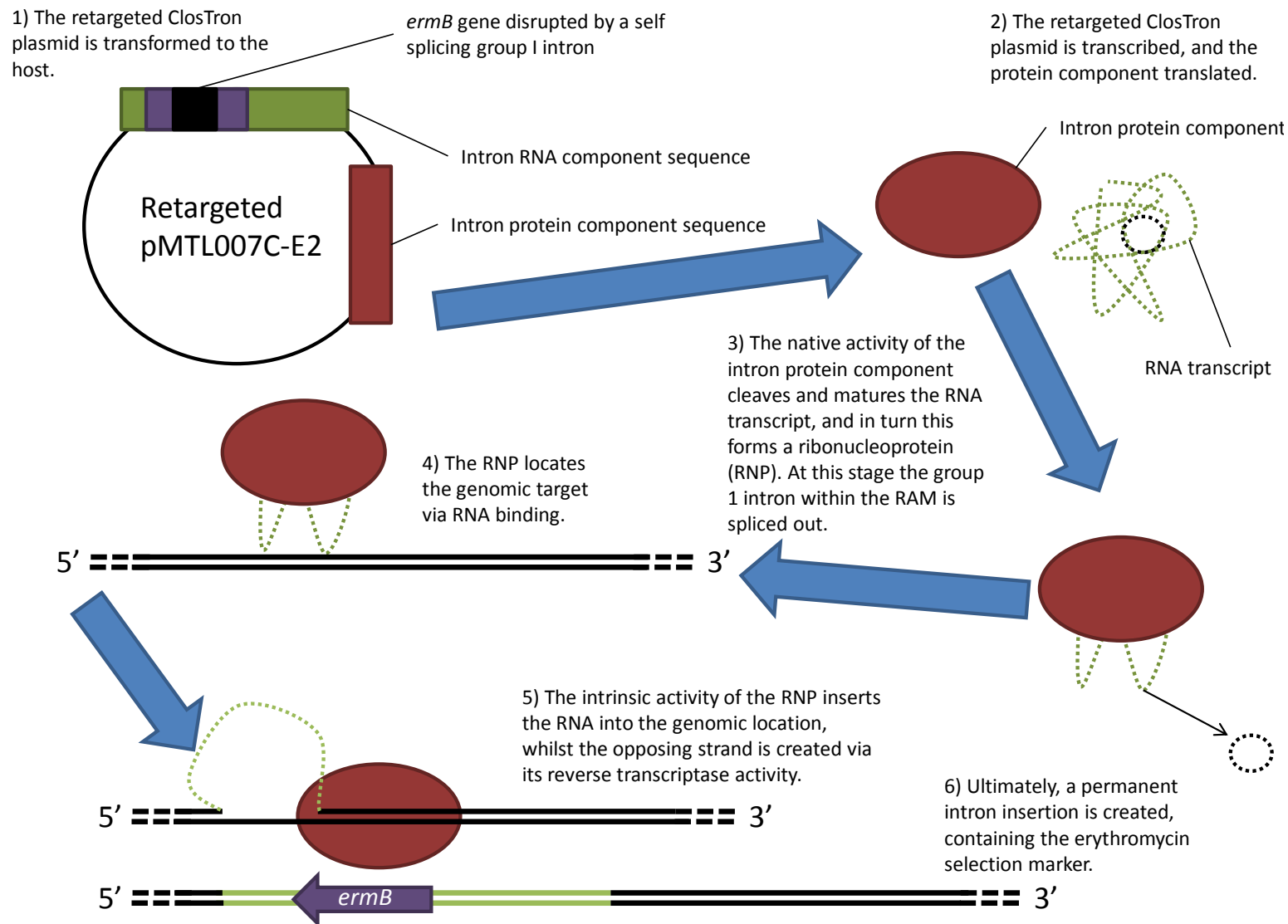


Figure 1.2 Diagram displaying the molecular action of the ClosTron technique. 1) The retargeted ClosTron plasmid is transformed to the host. 2) The protein and RNA components of the intron undergo transcription, and the former undergoes translation. 3) The RNA component of the intron undergoes maturation via the native activity of the protein intron component. The components form the ribonucleoprotein (RNP) complex. The group 1 intron within the retroactive marker (RAM) splices out. 4) The RNP is targeted via the RNA interaction with the host genome. 5) The intrinsic RNP activity leads to the insertion of the RNA into the host genome at the target location, and the opposing strand is produced by the RNPs reverse transcriptase activity. 6) A permanent insertion is created at the target genome location, due to the splicing out of the mobile group 1 intron, an active *ermB* selection marker can be used for the positive insertion of the ClosTron.

spp. Secondly, the available TargeTron plasmids with a RAM encode for kanamycin or trimethoprim resistance, antibiotics to which most clostridia have innate resistance (Heap et al., 2007). Moreover, in its most advanced manifestation, expression of the mobile group II is driven by the constitutive *fdx* promoter of *C. sporogenes* (Heap et al., 2010, 2009). A diagram of this system is shown in Figure 1.2.

The ClosTron system benefits from being expedient, but is not without drawbacks. The intron can cause so called ‘polar effects’, the alteration in the expression of nearby genes by either disruption of their coding operon, or transcription being driven from the promoters used for its function (Pyne et al., 2014; Xue et al., 2017). There is also the potential for ectopic insertions, that will have to be checked for using a Southern blot (Shao et al., 2007). Instability of the ClosTron is often mentioned, but the original source does not seem to link observed instability to the intron itself (Steiner et al., 2011).

1.5.2. In-frame deletion

The in-frame deletion of a gene circumvents the issues associated with ClosTron, i.e. polar effects and ectopic insertions. The *pyrE* gene has also been utilised in the generation of in-frame deletion in clostridia. Using this method for the creation of an in-frame deletion requires a $\Delta pyrE$ background (Ehsaan et al., 2016). The deletion plasmid pMTL-ME3 was generated that requires the cloning of a deletion cassette within a *lacZ* multiple cloning site, that contains homology arms of ~750bp flanking the region to be deleted (Ehsaan et al., 2016). The pMTL-ME3 vector contains the previously mentioned pIM13 replicon, chosen again for its segregationally instability, and also the *pyrE* gene from *C. sporogenes* as well as the *catP* selection marker (Ehsaan et al., 2016). Successful transformants are selected based on Thiamphenicol resistance, and following this are restreaked on media without uracil supplementation (Ehsaan et al., 2016; Ng et al., 2013). Lack of uracil requires the heterologous *pyrE* gene on the pMTL-ME3 plasmid, and again integration via single crossover offers a growth advantage, meaning that larger colonies can be selected to be taken forward (Ehsaan et al., 2016).

Creation of a clean knockout mutation relies on the formation of a double crossover and the loss of the plasmid backbone, this is selected for by plating the previous

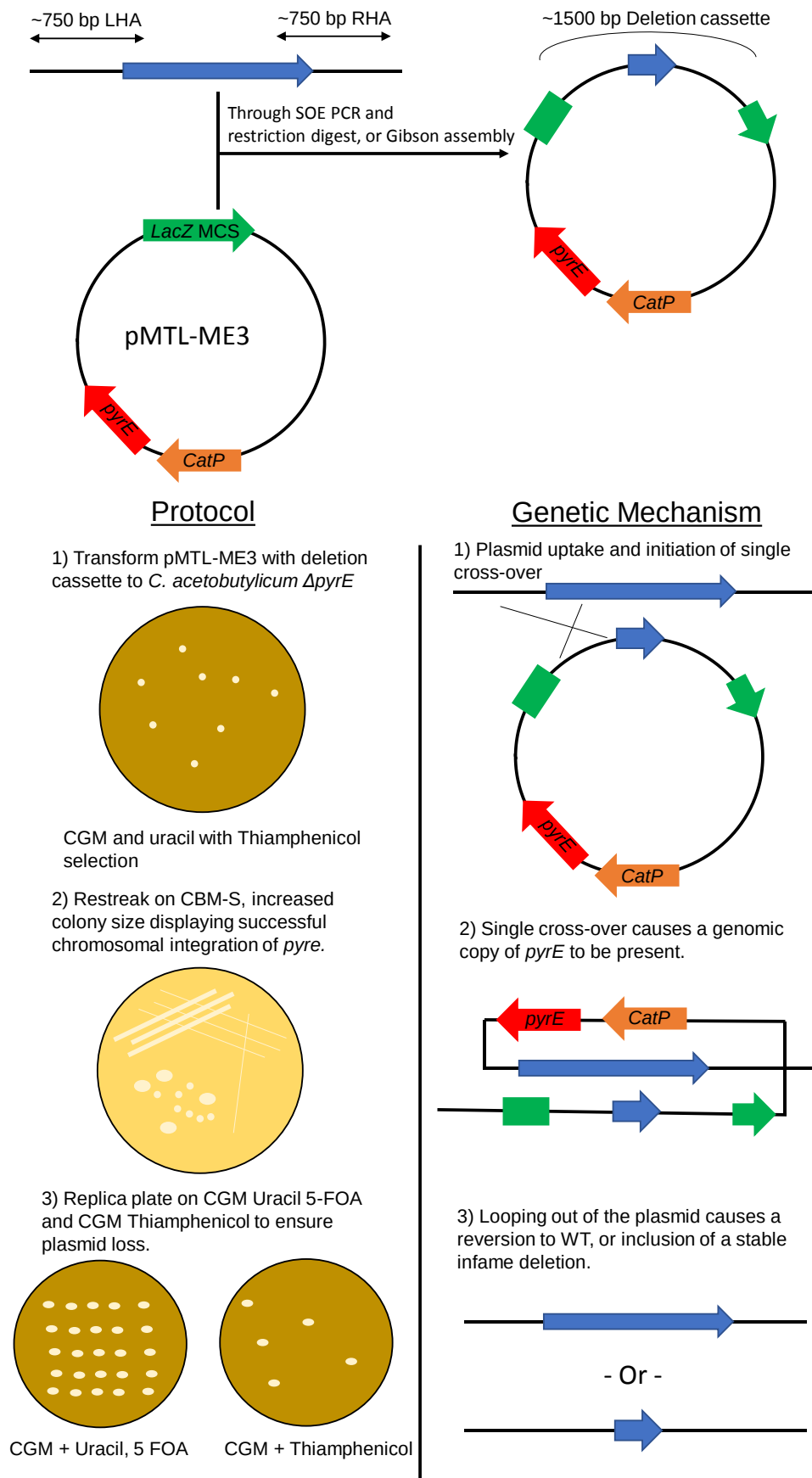


Figure 1.3 Diagram showing the genetic mechanism of the inframe deletion technique and the associated protocol. The upper panel shows the creation of homology arms for the deletion cassette, using primers to clone ~750bp of the upstream, and ~750bp of the downstream regions. Via SOE PCR or Gibson assembly, these fragments are combined within the *LacZ* MCS of the pMTL-ME3 plasmid. Genetic mechanism: 1) Following transformation of the plasmid, a single crossover can occur via the homologous recombination of one of the two homology arms. 2) Single crossover incorporates the plasmid within the genome, including the *pyrE* gene. 3) A secondary homologous recombination event results in either the formation of an inframe deletion mutation, or else reversion to WT. Protocol: 1) pMTL-ME3 with the desired deletion cassette is transformed to *C. acetobutylicum* ATCC 824 $\Delta pyrE$, the plasmid is selected for via *catP*. 2) Transformants are restreaked on CBM-S, where single crossover causes a growth advantage, due to the chromosomal inclusion of the *pyrE* gene. Larger colonies are picked and replica plated on CGM supplemented with uracil and 5-FOA and CGM supplemented with Thiamphenicol. 3) Growth on the former indicates successful double crossover, either creating the inframe deletion or else reverting to WT. Growth of the strains on CGM Thiamphenicol may indicate plasmid retention.

picked colonies onto media supplemented with uracil and 5FOA (Ehsaan et al., 2016). Details of the isolation protocol, and molecular mechanism, are displayed in Figure 1.3. These strains are then analysed via PCR and sequencing to confirm a deletion event. Following single crossover one of two eventualities can occur, that being double crossover or reversion to WT by the reverse of the single crossover, theoretically this should occur 50% of the time but in reality rarely does (Ehsaan et al., 2016). As the final strain includes the in-frame deletion of the target gene, and a *ΔpyrE* phenotype, it allows for chromosomal complementation of the gene to confirm the deletion effects.

Utilisation of the *pyrE* locus is not the only system that would allow the production of in-frame deletion mutants without antibiotic selection, and the *codA* gene has also been used (Cartman et al., 2012). The *codA* gene codes for a cytosine deaminase which catalyses the conversion of cytosine to uracil, but also the harmless 5-fluorocytosine derivative into the highly toxic 5-fluorouracil which again irreversibly inhibits thymidylate synthase, causing DNA replication stalling and cell death (Minton et al., 2016). This conversion occurs through the action of uracil phosphoribosyltransferase (Upp), and hence use of the *codA* selection system requires the organism to contain a functional *upp* gene and no *codA* gene (Cartman et al., 2012). The *codA* gene used is a heterologous gene from *E. coli*, and its presence on a plasmid within *C. difficile* or *C. acetobutylicum* ATCC 824 confers sensitivity 5-fluorocytosine (Cartman et al., 2012; Ehsaan et al., 2016). With respect to the selection methodology, it superficially matches that of the *pyrE* system, but with the use of 5-fluorocytosine instead of 5FOA and without the need for uracil supplementation (Ehsaan et al., 2016; Ng et al., 2013). This system has the benefit of not requiring a mutant starting point, and hence removes the optional mutant repair step following isolation of a double crossover event (Ng et al., 2013).

Other methods of selecting double crossover mutants also exist, e.g. the use of the *E. coli mazF* toxin gene (Al-Hinai et al., 2012). Each technique has its own advantages and disadvantages, and many reviews on the subject of engineering solventogenic

clostridia exist to testify to this (Lütke-Eversloh and Bahl, 2011; Minton et al., 2016; Papoutsakis, 2008; Xue et al., 2017).

1.5.3. Transposon mutagenesis

Another important component within the molecular biology toolset is the ability to create random mutant libraries, allowing screening for certain phenotypes and to analyse the essentiality of genes (Turner et al., 2015; Winterberg and M., 2008). Transposon mutagenesis was developed in *Clostridium spp.* using the mariner transposon system, and was first shown to be effective in *C. difficile* (Cartman and Minton, 2010). However, plasmid retention led to the creation of multiple insertions, an undesirable quality for downstream analysis (Zhang et al., 2015).

In this system, expression of the mariner transposase was under control of the promoter controlled by the unique *C. difficile* sigma factor TcdR. By introducing the encoding *tcdR* gene into a different clostridial host, the host range for the systems could be expanded (Zhang et al., 2015). In addition to this, the updated system included a newly developed conditional replicon that relied on an IPTG inducible promoter, forcing transcription of the pCB102 replication region and interfering with plasmid maintenance, possibly through the overexpression of RepH (Zhang et al., 2015). This in effect, upon addition of IPTG, allows total loss of the transposon plasmid, ensuring single transposon insertions in the creation of transposon libraries (Zhang et al., 2015).

1.5.4. CRISPR

Very recent advances in genetic engineering have led to the utilisation of the prokaryotic adaptive immunity system, clustered regularly interspaced short palindromic repeats (CRISPRs) and CRISPR associated (Cas) proteins, for the fast and efficient generation of defined mutations, and such systems are now available in the *Clostridium* genus (Bruder et al., 2016). CRISPRs allow a prokaryote to recognise viral DNA, or unwanted conjugative plasmids, and through the use of Cas proteins cause a double-strand break in such DNA in order to interfere with them (Marraffini and Sontheimer, 2011). Generally, the CRISPRs are transcribed as one continuous sequence in a unidirectional manner, following this Cas proteins process the RNA

transcript into small crRNAs, which form a complex with Cas proteins (Marraffini and Sontheimer, 2011). The crRNAs allow sequence recognition, and upon successful binding allows the Cas protein to create the targeted double-strand break (Marraffini and Sontheimer, 2011). These systems were first repurposed in mammalian cells, where double strand breaks can be repaired through non-homologous end joining.

Prokaryotes generally do not repair double strand breaks in that manner and hence the repair of Cas9 formed breakages require the inclusion of DNA fragments encoding homology arms (Bruder et al., 2016). This is reminiscent of the homology arms required in order to create a double-crossover mutation, i.e. failure to create a double crossover event following Cas 9 cleavage would mean no double strand break repair, and hence cell death. However, highly active Cas9 systems can result total in the elimination before a successful recombination event, leading to cumbersome calibration of the technique (Xue et al., 2017). For several species of the genus *Clostridium*, including *C. ljungdahlii*, *C. beijerinckii* and *C. acetobutylicum* ATCC 824, CRISPR/Cas9-based gene editing systems have now been described (Huang et al., 2016; Wang et al., 2016; Xue et al., 2017).

1.6. Bacterial cell-cell communication

Many bacteria are known to communicate via diffusible chemical signals (Papenfort and Bassler, 2016; Rocha-Estrada et al., 2010; Sturme et al., 2002; Taga and Bassler, 2003). It is often assumed that the excretion of signal molecules enables bacterial communities to sense the prevailing cell density and to coordinate gene expression accordingly (Papenfort and Bassler, 2016). The signal molecule concentration in a population is thought to reflect its population density and trigger, once a critical threshold has been reached, a population-wide response in the form of phenotypic changes. Whilst this view is probably too simplistic, and whilst there has been criticism of the claim that bacteria are involved in social undertakings, the body of evidence supporting the phenomenon has become too much to refute (Redfield, 2002; West et al., 2012).

Much of the social communication between bacterial organism is referred to as quorum sensing (QS), which specifically refers to the above described process of a

bacterial community sensing its population density and performing actions when sufficient densities are reached (Papenfort and Bassler, 2016) but which is often used in a more general sense to describe systems that rely on diffusible signals.

Communication signals produced by the Gram negative organisms are generally referred to as autoinducers, and those produced by the Gram-positive organisms named autoinducing peptides (AIPs) (Rocha-Estrada et al., 2010; Taga and Bassler, 2003). This is often attributed to the positive feedback regulation observed for many of these systems, leading to rapid upregulation (Papenfort and Bassler, 2016), although the original meaning was that this was a self-produced signal that led to induction of a given phenotype. As the work conducted here largely is comprised of investigating a supposed cell-cell communication system of unknown regulation, the more general terms of pheromone or signalling peptide are used preferentially (Bouillaut et al., 2008; Otto et al., 2001; Perchat et al., 2016b; Shi et al., 2005). While this may seem pedantic, inadequate use of terminology when discussing research into microbial communication has been indicated as an issue within the field (Diggle et al., 2007).

1.6.1. Cell-cell communication in the Gram-negative organisms

Cell-cell communication in Gram-negative bacteria relies on *N*-acylhomoserine lactones (AHLs) and other fatty acid derivatives as signal molecules, which are able to diffuse freely through the bacterial membrane (Papenfort and Bassler, 2016). One of the first bacterial cell-cell communication systems to be discovered was that of *Vibrio fischeri* (now renamed *Aliivibrio fischeri*) which is the main control system behind its characteristic bioluminescent phenotype (Engebrecht et al., 1983). This bacterium possesses a typical LuxR/LuxI-type quorum sensing system where LuxI is an autoinducer synthase which catalysis the formation of AHLs from S-adenosylmethionine and Acyl-ACP, and LuxR the corresponding transcriptional regulator which is activated by AHL binding (Miyashiro and Ruby, 2013). In the case of *V. fischeri* the AHL produced is *N*-(3-oxohexanoyl)-L-homoserine lactone (3-oxo-C6-HSL). Upon binding to LuxR, transcription of the luciferase-encoding *luxCDABEG* operon is initiated, leading to bioluminescence (Miyashiro and Ruby, 2013). As excreted 3-oxo-C6-HSL diffuses away in the seawater, free-living *V.*

fischeri cells are dark. However, when enclosed in the light organs of the squid *Euprymna scolopes*, the autoinducer can accumulate beyond the critical threshold concentration, thus leading to the bioluminescent cells.

More recently it has been established that quorum sensing-mediated regulation of bioluminescence in *V. fischeri* is far more complex than initially thought and in part resembles that found in another marine bacterium, *Vibrio harveyi* (Anetzberger et al., 2012). In addition to 3-oxo-C6-HSL, *V. fischeri* produces two more autoinducers, *N*-octanoyl-homoserine lactone (C8-HSL) and autoinducer-2 (AI-2). C8-HSL is produced by AinS, an enzyme that differs from the typical LuxI synthase, whereas AI-2 is produced by the enzyme LuxS as part of the activated-methyl cycle (Miyashiro and Ruby, 2013). In contrast to the LuxRI system, the receptors for these signals are the histidine kinases located in the membrane. The AinR histidine kinase binds C8-HSL directly, whereas LuxQ interacts with the periplasmic binding protein LuxP, once the latter has bound to AI-2. Activation of bioluminescence as a result of binding is mediated via a complex regulatory network involving a phosphorelay system, a small regulatory RNA, and the transcriptional regulator LitR (Miyashiro and Ruby, 2013). Interestingly, LuxS-catalysed production of AI-2 has been found across the spectrum of Gram negative and Gram positive bacteria, and is believed to represent a universal signal for inter-species communication (Chen et al., 2002; Henke and Bassler, 2004; Papenfort and Bassler, 2016). AI-2 is a signal that is ultimately derived from *S*-adenosylmethionine, through the production of 4,5-dihydroxy-2,3-pentanedione, a highly reactive compound that readily interchanges into cyclic furanones, an argument fielded for its ability to act as a pheromone with inter-species specificity (Papenfort and Bassler, 2016), even though a LuxP/LuxQ-type receptor is missing in most bacteria. It should be noted that AI-2 acting as a universal bacterial pheromone is not without its critics, and it has been suggested that it's wide presence may simply have a nutritional context (Winzer et al., 2002).

LuRIE-type quorum sensing systems exist in many other bacterial species, sometimes in multiple copies and have been well-studied in animal and plant pathogenic bacteria such as *Pseudomonas aeruginosa* and *Erwinia carotovora*, to name only a few (Barnard et al., 2007; Kariminik et al., 2017) where they usually regulate a multitude

of traits including the production of toxins and other virulence factors as well as biofilm formation. In principle they function as described for *V. fischeri* in that the produced AHLs accumulate in the culture and, once a critical threshold concentration is reached, trigger a transcriptional response via binding to their cognate LuxR-type regulator. In some bacteria, such as members of the genera *Pseudomonas*, *Rhizobium*, and *Yersinia*, complex regulatory hierarchies exist with one LuxRI-type system regulating another (Atkinson et al., 2006; Kariminik et al., 2017; LaRock et al., 2013; Sanchez-Contreras et al., 2007).

This is by no means a conclusive overview of cell-cell communication in Gram-negative bacteria as there are several other systems, such as “diffusible signal factor” (DSF)-based systems of plant associated species such as *Xanthomonas*, and also *Burkholderia* spp. which produce cis-11-methyl-2-dodecenoic acid and related derivatives (Papenfort and Bassler, 2016; Ryan et al., 2015). However, the above outlined organisms give a sufficient overview of the Gram negative systems, without straying too far from the crux of this body of work, the cell-cell communication systems of *C. acetobutylicum* ATCC 824.

1.6.2. Cell-cell communication in Gram-positive organisms

In contrast to their Gram-negative relatives, Gram-positive bacteria employ mostly peptides for cell-cell communication. These can either be linear, cyclised or otherwise modified. Several systems have been described, but only the well-studied Agr and RNPP-type systems will be described here.

1.6.2.1. Agr type signalling

The Accessory Gene Regulator (Agr) system is most notable within *Staphylococcus aureus* and *Enterococcus faecalis* where it plays an integral role in the regulation of virulence (Cook and Federle, 2014; Le and Otto, 2015; Novick and Geisinger, 2008; Sturme et al., 2002). Agr systems have been also been shown to be present within the genus *Clostridium* genus, being shown to be involved in the regulation of sporulation and virulence factor production of *C. botulinum* and *C. perfringens*, as well as sporulation and granule production of *C. acetobutylicum* ATCC 824 (Cooksley et al., 2010; Li et al., 2011; Steiner et al., 2011). The Agr system is formed of at least 4

components, AgrA, AgrB, AgrC and AgrD, which are encoded by the *agrBDCA* gene cluster (Le and Otto, 2015; Novick and Geisinger, 2008; Otto, 2004).

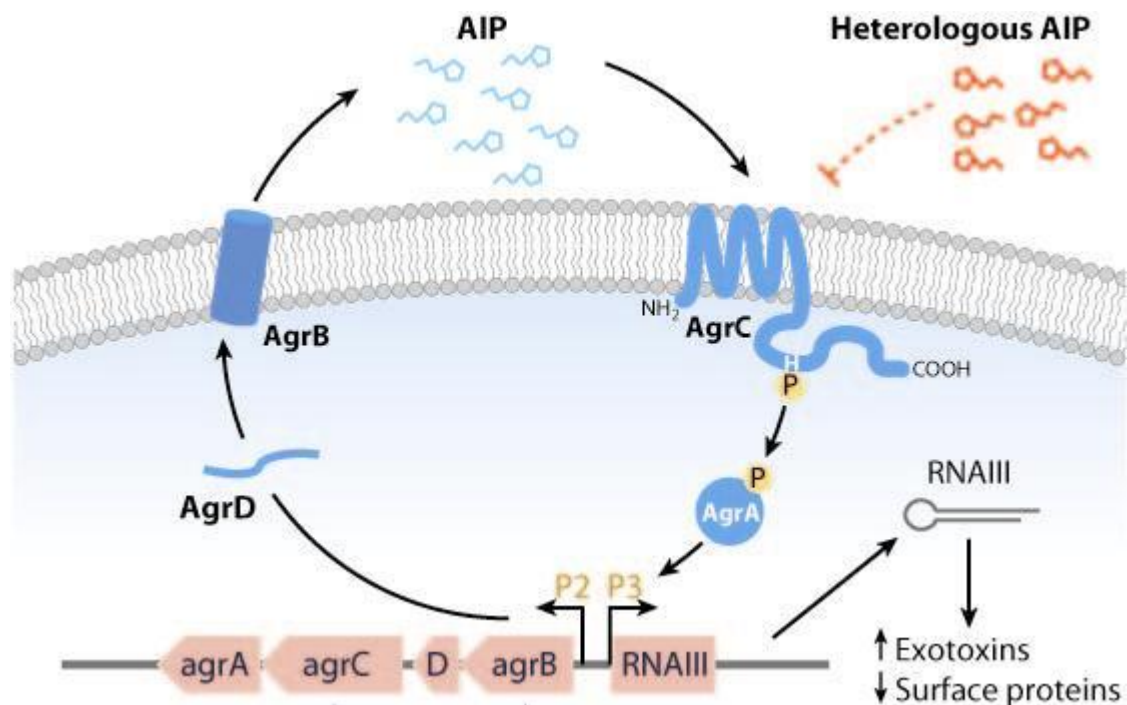


Figure 1.4 A schematic diagram of the Agr quorum sensing system in *S. aureus* (taken and modified from Novick & Geisinger., 2008) The *agrBDCA* operon and the regulatory RNAIII are divergently transcribed. The AgrD precursor peptide is processed and exported by the membrane associated AgrB and accumulates in the extracellular environment. Binding to the histidine kinase AgrC leads to activation of the AgrA response regulator via phosphorylation. This stimulates transcription of RNAIII, a regulatory RNA responsible for the activation of numerous target genes.

A general overview of the staphylococcal Agr system is given in Figure 1.4: the *agrD* encoded signalling peptide precursor is proteolytically processed and then exported by the membrane-associated AgrB protein, with the mature form possessing a five membered ring and a short N-terminal tail of 2-4 amino acids (Le and Otto, 2015; Novick and Geisinger, 2008). Accumulation of secreted AIP in the extracellular environment within a growing population eventually leads to its efficient binding to the transmembrane histidine kinase AgrC (Le and Otto, 2015; Novick and Geisinger, 2008). The activity of AgrC then leads to the phosphorylation of the response regulator AgrA, which in *S. aureus* is shown to activate its own promoter and that of the adjacent regulatory RNAIII (Novick and Geisinger, 2008). RNAIII is the primary effector of the *S. aureus* Agr system and its structure is highly conserved across multiple strains and also related species, even though the primary sequence is not (Novick and Geisinger, 2008).

The structure itself leads to the upregulation of α -haemolysin, and down regulation of multiple genes by binding the translator initiation region of their transcripts (Novick and Geisinger, 2008). In this fashion, RNAIII alters the relative abundance of multiple exoproteins and pleiotropic regulators, leading to a large metabolic burden, which has been given as a possible reason for the occurrence of *agr* mutations in *S. aureus* lab strains (Novick and Geisinger, 2008). Interestingly, AgrA has also been shown to have effects on the cell's metabolism, which are independent of the action of RNAIII (Queck et al., 2008). This is of particular interest when considering that within the Agr system of *C. acetobutylicum* ATCC 824 an RNAIII homolog was not found, and it possibly operates without one (Steiner et al., 2012). Also, while in *S. aureus* *agrBDCA* is transcribed together as a gene cluster, within *C. acetobutylicum* ATCC 824 *agrBD* is predicted to be transcribed separately to *agrCA* (Paredes et al., 2004; Steiner et al., 2011).

The AIPs involved in Agr-signalling systems form thiolactone rings through reaction of an internal cysteine-derived SH group with the C-terminal carboxyl group (Lyon and Novick, 2004; Sturme et al., 2002), although lactone ring structures also form a serine residue between the C-terminal carboxyl group. Between *Staphylococcus* species and even *S. aureus* pheromone subgroups, the amino acid sequence of these cyclic structures can differ greatly (Otto et al., 2001). However, there is some consensus among the central cysteine and the distance to the C-terminus, in order to form a uniform 5-membered ring (Lowy, 1998; Otto et al., 2001). This overall similarity of structure allows *S. aureus* strains to engage in inhibitory cross-talk, in which the pheromone from one subgroup will prevent another from detecting its own signal, a process that likely evolved from the close quarters competition in the organisms' ecological niche (Otto et al., 2001). There are variations of the 5 amino acid ring structure of *Staphylococci*, with the Fsr system of *Enterococcus faecalis* employing a 9 amino acid ring pheromone, and the Agr system of *C. acetobutylicum* ATCC 824 containing a possible 6 amino acid ring structure (Nakayama et al., 2001; Steiner et al., 2012).

In *C. acetobutylicum* ATCC 824, the Agr system was shown to have some control over the level of spores the organism produced, as well as regulate the production of

granulose, an intracellular starch-like storage compound (Steiner et al., 2012). ClosTron insertional mutations of *agrA*, *agrB* and *agrC* led to a 10-fold reduction in spores when grown in liquid culture, whilst sporulation on solid media had a more pronounced effect, and each of the insertional inactivation strains failed to produce the granulose storage compound (Steiner et al., 2012). A mathematical model was employed to gain further insight into the *C. acetobutylicum* ATCC 824 system (Jabbari et al., 2013). The model suggested that the system might act as an energy saving mechanism (Jabbari et al., 2013). Interestingly, the model also suggested that solventogenesis might benefit from being regulated dependent on cell density, showing a possible theoretical link to its control via a cell-cell communication system (Jabbari et al., 2013). Important sections of the *C. acetobutylicum* ATCC 824 life cycle are clearly under cell-cell communication control, and as such it is not a stretch to suggest other systems could be in place to modulate other aspects of the organism's molecular decision making.

1.6.2.2. RNPP type signalling

The RNPP-type QS systems take their name from the number of well-studied systems that have some conformity of mechanism, namely Rap, Npr, PlcR and PrgX (Rocha-Estrada et al., 2010). RNPP-type systems have a unique mechanism relative to other systems as the signal peptide interacts directly with its conjugate regulator. This regulator can then perform its cellular action, which is either as a Rap-type Phosphatase, a transcriptional regulator, or in one particular case both (Rocha-Estrada et al., 2010).

The Rap systems present in the genus *Bacillus* are unique amongst the RNPP family, as they do not contain a HTH domain. They therefore do not engage in regulation through the binding of DNA, but rather through their phosphatase activity on phosphorelay systems or through physical binding properties (Gallego del Sol and Marina, 2013). The Rap proteins have 7 tetratricopeptide repeat (TPR) domains, each of which forms a helix (Gallego del Sol and Marina, 2013). These domains mediate protein-protein or protein-peptide interactions (Rocha-Estrada et al., 2010). Interestingly, the systems conjugate Phr peptides inhibit the activity of their cognate Rap phosphatase by binding to it allosterically and changing the Rap proteins shape

(Gallego del Sol and Marina, 2013). The TPR helices together form a superhelix and Phr binding causes this superhelix to constrict rendering the Rap protein unable to perform its function (Gallego del Sol and Marina, 2013). *B. subtilis*, for example, contains at least 7 Rap systems that have control over sporulation, genetic competence and other gene functions relating to a quorum (Pottathil and Lazazzera, 2003). The Phr peptides are excreted from the cell in a preform through a SecA-dependent transportase, and is processed externally via proteases to its mature form, which is a linear pentapeptide (Pottathil and Lazazzera, 2003).

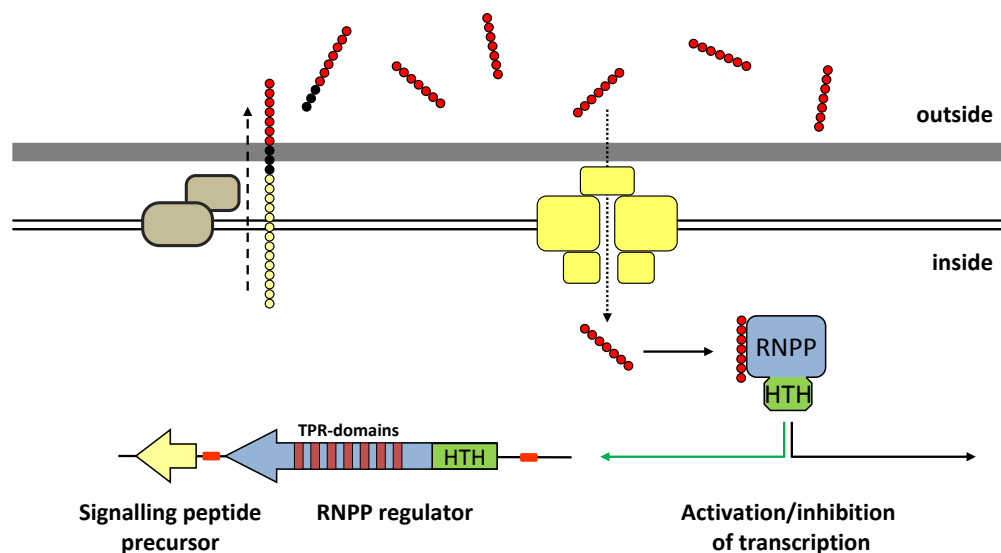


Figure 1.5 **General scheme of RNPP-type cell-cell communication.** The diagram shows a simplified RNPP-type quorum sensing circuit for a regulator (blue) containing a helix-turn-helix motif (green). Regions encoding TPR domains (dark red) are shown for the encoding gene only (large arrow). The red box upstream of the RNPP gene indicates a promoter. The signalling peptide precursor encoding gene (small yellow arrow) is often expressed from an independent promoter. Following transcription and translation, the pre-propeptide is exported via the Sec machinery and processed by peptidases (light brown ovals) into its final mature form. Following accumulation in the extracellular environment, peptide uptake is mediated by oligopeptide permeases (ABC transporters, yellow). Inside the cell, the peptide binds to and activates (or inhibits) its cognate RNPP regulator, resulting in conformational changes and changes in the oligomerisation state (not) shown. For some systems a positive feedback loop has been described (green arrow). Double lines and grey bar represent cell membrane and wall, respectively.

Figure 1.5 displays a general overview of RNPP-type cell-cell communication for those regulators that contain a HTH, which are described below. The Npr system, most notable in the insect pathogen *Bacillus thuringiensis*, is a fascinating cell-cell communication system which has major implications for the life-style of the organism

(Perchat et al., 2016a). The Npr system is unique in the fact that it contains both a HTH domain and two Rap-like TPR domains (Rocha-Estrada et al., 2010). NprR in its free form exists as a dimer, but upon binding of its conjugate peptide it forms an active tetramer (Do and Kumaraswami, 2016). In *B. thuringiensis*, NprR is key to its necrotrophism within the insect larvae cadaver post-infection (Lemy et al., 2012). A metabolic shift must occur from pathogenicity to necrotrophism, and has been shown to be linked to the Npr system, which itself regulates 41 genes (Lemy et al., 2012).

Included in these genes are several degradative enzymes, that were shown as essential for in vitro substrate degradation (Lemy et al., 2012). In addition to this the non-ribosomal peptide kurstakin was also regulated via the Npr system, shown to be essential for survival during a necrotrophic lifestyle due to its role in swarming and biofilm formation (Lemy et al., 2012). More research into the role of NprR in this organism showed the unbound dimer to act as a sporulation inhibitor, and as a sporulation activator when present as a tetramer and in complex with its conjugate peptide NprX (Perchat et al., 2016a). The Rap like TPR domains of NprR allow the free dimer to act as a phosphatase dephosphorylating Spo0F-P, and in doing so inhibiting the initiation of sporulation (Perchat et al., 2016b). The dual action of this regulator allows it to make a key molecular decision in the life cycle of *B. thuringiensis*, from sporulation to necrotrophism (Perchat et al., 2016a). The minimal length of NprX shown to activate the HTH binding activity of NprR was a heptapeptide represented by the sequence SKPDIVG (Perchat et al., 2011). The described dual action of the regulator has led to the suggestion that NprR is an evolutionary intermediate between the Rap phosphatases and RNPP-type regulators that contain a HTH domain (Perchat et al., 2016b)

PlcR homologs have been shown to be the key regulators in *Bacillus cereus* and *B. thuringiensis* pathogenicity (Agaisse et al., 1999). The regulator is found to be functionally equivalent in both *B. cereus* and *B. thuringiensis*, but interestingly but it is truncated in *Bacillus anthracis* and none functional as a transcriptional regulator (Agaisse et al., 1999). Virulence was abolished in various animal models in *plcR*-deficient mutants (Grenha et al., 2012) in agreement with it being a pleiotropic

regulator, with genes under its regulation widely dispersed amongst the genome (Agaisse et al., 1999).

It appears PlcR only binds to DNA via its HTH domain following ordered polymerisation induced by ligand binding (Declerck et al., 2007). It has been suggested that this polymerisation, forming a loop around the DNA bound via its HTH domain, could be a form of regulation, not uncommon in transcriptional regulators (Declerck et al., 2007). The minimal active PapR peptide required for the activation of the regulator was originally thought to be 5 amino acids long and was represented by the sequence LPFEF. However, further study showed the minimal active peptide to be a heptapeptide, represented by the sequence ADLPFEF (Bouillaut et al., 2008).

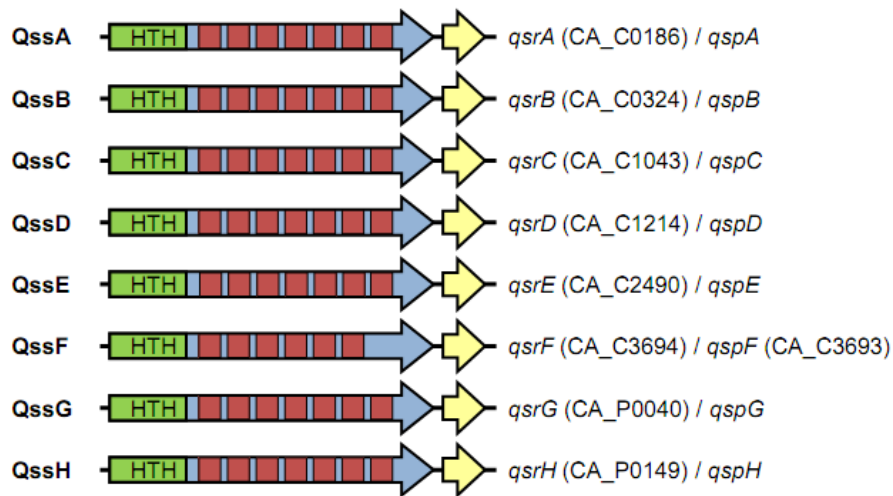
To complete the section on RNPP type cell-cell communication, the PrgX system of *Enterococcus faecalis* will be commented on. The PrgX system regulates conjugation in the organism, and in doing so employs a conjugate peptide signal and an additional inhibitor peptide signal (Shi et al., 2005). The conjugate peptide, cCF10, is encoded on the genome, whereas the regulator PrgX is encoded on the pCF10 plasmid (Shi et al., 2005). The competitive inhibitor iCF10 is also encoded on the plasmid, and is present to prevent the autoactivation of the pCF10 containing cell (Shi et al., 2005). Plasmid free cells, i.e. potential conjugative recipients, produce the conjugate peptide cCF10 and it therefore increases in the local environment. It is taken up by pCF10 containing cells and now outcompetes the inhibitory iCF10 signal (Shi et al., 2005). While bound to iCF10, PrgX forms a tetramer that inhibits the *prgQ* conjugative operon found on the pCF10 plasmid (Kozłowicz et al., 2004; Shi et al., 2005). However, when in the presence of conjugative recipients producing cCF10 signal, PrgX bound to cCF10 becomes a dimer, activating the conjugative operon and promoting conjugation (Shi et al., 2005).

It can be seen here that individual RNPP-type systems are being used to regulate a vast range of bacterial activities, from changing the lifestyle of *B. thuringiensis* to plasmid conjugation in *E. faecalis*. Homologous systems were therefore studied in *C. acetobutylicum* ATCC 824, which was found to contain at least 8 RNPP-type cell-cell

communications systems (QssA-QssH), each consisting of the respective RNPP-type regulator (QsrA-QsrH) and signalling peptide (QspA-QspH). These are displayed within Figure 1.6. They are represented by the gene loci CA_C0186, CA_C0324, CA_C1043, CA_C1214, CA_C2490, CA_C3695/CA_C3694, CA_P0040 and CA_P0149 respectively. In addition, two orphan RNPP-type regulators (QsrI and QsrJ, CA_C0957 and CA_C0958) were also encoded in the genome (Kotte, 2013). These were discovered via the use of KEGG homolog and pblast searches, interrogating each database using the PlcR, Rap, PrgX and NprR sequences (Kotte, 2013; Kotte et al., 2017). All of these systems underwent preliminary characterisation, based on ClosTron insertional gene disruption, and in some cases, plasmid based overexpression, in order to elucidate their function. ClosTron mutations were screened phenotypically with respect to growth, colony morphology, starch degradation, granule formation, sporulation and solvent formation (Kotte, 2013; Kotte et al., 2017). The two orphan RNPP-type regulators QsrI and QsrJ were found to be essential for sporulation and high level solvent production (Bean, 2016; Kotte, 2013). The genes for these regulators are arranged in tandem and thought to act in the same regulatory pathway, as the respective mutant phenotypes were almost identical.

However, the *qsrIJ* locus was found to not contain a putative signalling peptide encoding gene and it is therefore uncertain whether QsrI and QsrJ are part of a quorum sensing circuit (Kotte, 2013). One other regulator, QsrB, was also of particular interest, as disruption of the encoding gene led to an increase in solvent production during early solventogenesis, suggesting that it might act to repress solvent formation at earlier time points. This view was further supported through *qsrB* overexpression, which resulted in a severe reduction of solvent and spore formation, a phenotype that could be overcome through the addition of synthetic signalling peptide (Kotte et al., 2017). It remains to be seen why *C. acetobutylicum* ATCC 824 requires so many RNPP-type systems and it has been hypothesised that it may be a requirement of its complex metabolism and life cycle, including endospore formation (Kotte et al., 2017). Similar RNPP-type regulators have been discovered in other solventogenic *Clostridium* spp., sometimes together with several Agr-type quorum sensing systems and it thus seems that solventogenic *Clostridium* spp. contain an unprecedented number of communication systems (Poehlein et al., 2017).

A)



B)

B)

	positively charged	hydrophobic	mature signalling peptide	
QspG	---MLL KKHKL AFSVL-ALTLLF---	LGVNFLGNNVGLSVSPMTPW-----	39	
QspC	---MFL KRHKLI FSIL-LLGTLI---	LGVTFF KSS-LLAFY PGMPW-----	38	
QspA	-MITVL KKHRFL ISVL-TLGVLF---	LCIVFTNSSYILSGAPCIPW-----	41	
QspE	MRIVFFKRHK FLTLVL-ALGILF---	LGATFFGNS-LLSAF PDLPW -----	41	
QspD	--MFFL KKHKKF LSML-VLTLVIG--	VLVTT KGGISIL SMF PDAPW -----	41	
QspF	---MFF KRYKKLL FIAGILAAVLT--	INLGLNGN--LLSGL PLVWW -----	39	
QspH	---MNS RVKKLL FLV-LCSFII---	IGYLSVNHISVLAS YPGWSW -----	41	
QspB	---MLP KKSNKF ILVLILCAIVFVSAFALNTSGT	BSLLGAEP TWGW NISKLLF 50		
	. . : . : :	:: .	* . * *	

Figure 1.6 Schematic representation of the 8 RNPP type QS systems that have been identified (A) and the alignment of the putative signalling precursor peptides (B). A) The eight identified systems are shown, encoding an RNPP-type regulator (QsrA to QsrH, large arrows) and a signalling peptide precursor (QspA to QspH, short yellow arrows). Locus tags are provided where available. Regions encoding the tetracopeptide repeats are indicated (red), and the helix turn helix motif indicated (green). B) Clustal Omega amino acid sequence alignment of the eight predicted Qsp peptides. Conserved amino acids are indicated (red), positively charged (blue) and negatively charged (green) amino acids are indicated. Identical (*), conserved (:) and semiconserved (.) substitutions are shown. Numbers indicate the length of the putative peptide. Positively charged, hydrophobic and predicted signalling peptides are shown (blue, black and red lines respectively). Figure adapted from (Kotte et al., 2017).

As previously mentioned, RNPP type systems require uptake of their respective signalling peptides, a task which is accomplished with the help of ATP bind cassette

(ABC) dependent oligopeptide permeases (Gominet et al., 2001; Neiditch et al., 2017; Rocha-Estrada et al., 2010; Varahan et al., 2014). ABC transporters are fascinating systems present in all domains of life, in bacteria they often include specific binding proteins (SBPs) that are specific to the permease function (Locher, 2004; Solomon et al., 2003). These systems form the basis of Chapter 4, and will be fully introduced there. However, it is important to note that uptake of signalling peptides via oligopeptide permeases is vital for the signalling process and adds another level of regulatory control. For instance, inactivation of the *B. subtilis opp* system, shown to be required for signalling peptide uptake, shows a severe sporulation defect (Lazazzera et al., 1997; Perego et al., 1991; Rudner et al., 1991).

1.7. iTRAQ proteomics

As previously discussed, the relatively complex bi-phasic nature of *C. acetobutylicum* ATCC 824 metabolism is an area of much study. The nature of the changes during the solventogenic organisms' life style lends itself well to the utilization of numerous – omics approaches (Aristilde et al., 2015; Lee et al., 2015; Mao et al., 2011).

Recently, isobaric tag for relative and absolute quantification (iTRAQ) proteomics has been applied to *C. acetobutylicum* ATCC 824 (Raut et al., 2016). The iTRAQ method allows for the comparative study of proteomes, across differing experimental conditions, by the creation of peptide libraries tagged with an individual isobaric label (Rauniyar and Yates, 2014). These libraries are then combined, and separated based on size via liquid chromatography (Raut et al., 2016). These separated peptides are investigated using tandem mass spectrometry, and the ratio of peptides from each condition calculated via the use of the isobaric labels. These peptides are linked via the use of a protein library, to the protein to which they form a constituent part, and hence a comparative picture is drawn between the experimental conditions (Raut et al., 2016).

There are some key advantages to the use of iTRAQ. As differing constituent peptides are linked to the same protein during the analysis stage, comparative proteomics can be conducted with a very high degree of statistical validation (Ow et al., 2009). This

allows for even small changes in the relative level of a protein of interest to be viewed with a greater level of certainty (Ow et al., 2009). However, recent criticisms of the technique have argued against its level of quantification, displaying that iTRAQ can under-report fold changes between experimental states (Ow et al., 2009). iTRAQ remains a highly popular and influential application of the proteomics field, and one that will be expanded upon and utilized within Chapter 5.

1.8. Aims and objectives

As mentioned above, *C. acetobutylicum* ATCC 824 contains at least 8 RNPP-type cell-cell communication systems (Kotte, 2013). While a detailed investigation of each of these systems would be unsurmountable within the given timeframe, building on the knowledge of one such system promised a favourable outcome. The choice was made to conduct further study into Quorum Sensing System B (QssB), consisting of Quorum Sensing Regulator B (QsrB) and its conjugate peptide coded immediately down stream of its regulator, Quorum Sensing Peptide B (QspB). As described above, QssB was shown to have an effect over solventogenesis and sporulation, arguably two of the most interesting characteristics of *C. acetobutylicum* ATCC 824 (Kotte, 2013; Kotte et al., 2017). The overall aim of this study was therefore to provide in depth insights into workings and role of this system. Specifically, this study aimed to achieve the following key objectives:

- (i) To attempt to inactivate the *qsrB* gene by in-frame deletion, and to phenotypically characterise the resulting mutants
- (ii) To fully characterise the effects of *qsrB* overexpression in terms of acid production (acetate, butyrate), solvent formation (butanol, acetone, ethanol) and sporulation
- (iii) To assess synthetic QspB peptide variants in terms of their biological activity to identify sequences with maximum activity, define the specificity of the QsrB receptor, and assess potential inter-species cross-activation or inhibition
- (iv) To investigate the role of oligopeptide permeases for solventogenesis, sporulation and granule formation and, if possible, to identify the system responsible for QspB uptake

- (v) To apply an iTRAQ-based proteomics approach to define the set of proteins controlled directly or indirectly by QsrB and to confirm key changes by enzyme and reporter gene assays, as appropriate

2. Materials and Methods

2 Materials and Methods

2.1 List of bacterial strains, plasmids, oligonucleotides and peptides

Wild type and genetically modified *E. coli* and *C. acetobutylicum* ATCC 824 strains used in this study are displayed in Table 2.1. The plasmids used, both unmodified and their engineered variants, are listed in Table 2.2. Oligonucleotides used for polymerase chain reactions (PCRs) are shown in Table 2.3, and synthetic peptides used are shown in Table 2.4.

Table 2.1 **Bacterial strains used in this study**

Strain	Relevant information	Source
<i>E. coli</i> TOP10	Genotype: F- <i>mcrA</i> Δ (<i>mrr</i> - <i>hsdRMS-mcrBC</i>) Φ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74 recA1</i> <i>araD139</i> Δ (<i>ara leu</i>)7697 <i>galU</i> <i>galK rpsL</i> (StrR) <i>endA1 nupG</i>	Invitrogen
<i>E. coli</i> TOP10 pAN2	<i>E. coli</i> TOP10 transformed with the pAN2 methylation plasmid	(Heap et al., 2007)
<i>E. coli</i> TOP10 pMTL85143- <i>qsrB</i> -Strep	Modified <i>qsrB</i> gene encoding Strep-tagged version for purification	This study
<i>C. acetobutylicum</i> ATCC 824	<i>C. acetobutylicum</i> ATCC 824 parental wild type; type strain of the species (Accession number: AE001437)	Bahl Laboratory - Rostock University, Germany
<i>C. acetobutylicum</i> ATCC 824 Δ <i>pyrE</i>	<i>C. acetobutylicum</i> ATCC 824 with a truncated <i>pyrE</i> gene	(Ehsaan et al., 2016)
<i>C. acetobutylicum</i> ATCC 824 pMTL85143	Vector control strain used for overexpression experiments.	(Kotte, 2013)
<i>C. acetobutylicum</i> ATCC 824 pMTL85143- <i>qsrB</i>	Plasmid based overexpression of the RNPP regulator gene <i>qsrB</i> .	(Kotte, 2013) – remade this study
<i>C. acetobutylicum</i> ATCC 824 CA_C3632::CTermB	<i>C. acetobutylicum</i> ATCC 824 CA_C3632 ClosTron mutant	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C3632::CTermB pMTL85143	<i>C. acetobutylicum</i> ATCC 824 CA_C3632 ClosTron mutant vector control strain	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C3632::CTermB pMTL85143- <i>qsrB</i>	<i>C. acetobutylicum</i> ATCC 824 CA_C3632 ClosTron <i>qsrB</i> overexpression strain	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C3632::CTermB pMTL85141-CA_C3632	<i>C. acetobutylicum</i> ATCC 824 CA_C3632 ClosTron complementation strain	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C2377::CTermB	<i>C. acetobutylicum</i> ATCC 824 CA_C2377 ClosTron mutant	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C2377::CTermB pMTL85143	<i>C. acetobutylicum</i> ATCC 824 CA_C2377 ClosTron mutant vector control strain	This study

<i>C. acetobutylicum</i> ATCC 824 CA_C2377::CTermB pMTL85143- <i>qsrB</i>	<i>C. acetobutylicum</i> ATCC 824 CA_C2377 ClosTron <i>qsrB</i> overexpression strain	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C2667::CTermB	<i>C. acetobutylicum</i> ATCC 824 CA_C2667 ClosTron mutant	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C2667::CTermB pMTL85143	<i>C. acetobutylicum</i> ATCC 824 CA_C2667 ClosTron mutant vector control strain	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C2667::CTermB pMTL85143- <i>qsrB</i>	<i>C. acetobutylicum</i> ATCC 824 CA_C2667 ClosTron <i>qsrB</i> overexpression strain	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C2667::CTermB pMTL85141- CA_C2667	<i>C. acetobutylicum</i> ATCC 824 CA_C2667 ClosTron complementation strain	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C0176::CTermB	<i>C. acetobutylicum</i> ATCC 824 CA_C0176 ClosTron mutant	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C0176::CTermB pMTL85143	<i>C. acetobutylicum</i> ATCC 824 CA_C0176 ClosTron mutant vector control strain	This study
<i>C. acetobutylicum</i> ATCC 824 CA_C0176::CTermB pMTL85143- <i>qsrB</i>	<i>C. acetobutylicum</i> ATCC 824 CA_C0176 ClosTron <i>qsrB</i> overexpression strain	This study
<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ::CTermB	<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ClosTron mutant	(Kotte, 2013)
<i>C. roseum</i> Wis42-A	<i>C. roseum</i> type strain	DSMZ (7320)
<i>C. acetobutylicum</i> ATCC 824 pMTL85141Pfdx_GusA	<i>C. acetobutylicum</i> ATCC 824 with <i>fdx</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ::CTermB pMTL85141Pfdx_GusA	<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ClosTron mutant with <i>fdx</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 pMTL85141P <i>qsrB</i> _GusA	<i>C. acetobutylicum</i> ATCC 824 with <i>qsrB</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham

<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ::CTermB pMTL85141PqsrB_GusA	<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ClosTron mutant with <i>qsrB</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 pMTL85141P <i>adc</i> _GusA	<i>C. acetobutylicum</i> ATCC 824 with <i>adc</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ::CTermB pMTL85141P <i>adc</i> _GusA	<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ClosTron mutant with <i>adc</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 pMTL85141P <i>adhE</i> _GusA	<i>C. acetobutylicum</i> ATCC 824 with <i>adhE</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ::CTermB pMTL85141P <i>adhE</i> _GusA	<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ClosTron mutant with <i>adhE</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 pMTL85141P <i>ptb</i> _GusA	<i>C. acetobutylicum</i> ATCC 824 with <i>ptb</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ::CTermB pMTL85141P <i>ptb</i> _GusA	<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ClosTron mutant with <i>ptb</i> promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 pMTL85141PCA_C0957_GusA	<i>C. acetobutylicum</i> ATCC 824 with CA_C0957 promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ::CTermB pMTL85141PCA_0957_GusA	<i>C. acetobutylicum</i> ATCC 824 <i>qsrB</i> ClosTron mutant with CA_0957 promoter <i>gusA</i> fusion	Dr Zak Bean and Jessica Locker, University of Nottingham
<i>C. roseum</i> ATCC 17797	(Accession number: LZYV00000000)	NRRL agricultural research service culture collection

Table 2.2 **Plasmids used in this study**

Plasmid	Relevant Information	Source
pAN2	Plasmid containing the ϕ 3TI methyltransferase	(Heap et al., 2007)
pMTL-ME3	<i>C. acetobutylicum</i> ATCC 824 inframe deletion vector	(Ehsaan et al., 2016)
pMTL85141	<i>Clostridium</i> modular plasmid containing the pIM13 Gram positive replicon, <i>catP</i> selection marker, <i>ColE1</i> Gram negative replicon and a <i>lacZ</i> multiple cloning site.	(Heap et al., 2009)
pMTL85143	<i>Clostridium</i> modular overexpression plasmid containing the pIM13 Gram-positive replicon, <i>catP</i> selection marker, <i>ColE1</i> Gram negative replicon and a <i>lacZ</i> multiple cloning site which is driven by the <i>fdx</i> promoter.	(Heap et al., 2009)
pMTL85143- <i>qsrB</i>	<i>qsrB</i> overexpression plasmid	(Kotte, 2013) – remade this study
pMTL85143- <i>qsrB</i> -Strep	Attempted <i>qsrB</i> Strep-tag plasmid	This study
pMTL85141- CA_C2667	CA_C2667 complementation plasmid	This study
pMTL85141- CA_C3632	CA_C3632 complementation plasmid	This study
pMTL007C-E2::CA_C0176-259s	ClosTron plasmid retargeted to CA_C0176	DNA2.0 – This study
pMTL007C-E2::CA_C2376-307s	ClosTron plasmid retargeted to CA_C2376/CA_C2377 possible frameshift	DNA2.0 – This study
pMTL007C-E2::CA_C2667-466s	ClosTron plasmid retargeted to CA_C2667	DNA2.0 – This study
pMTL007C-E2::CA_C3179-439s	ClosTron plasmid retargeted to CA_C3179	DNA2.0 – This study
pMTL007C-E2::CA_C3632-305s	ClosTron plasmid retargeted to CA_C3632	DNA2.0 – This study
pMTL007C-E2::CA_C3640-241s(CA_C3634)	ClosTron plasmid retargeted to CA_C3640 and its duplicate CA_C3634	DNA2.0 – This study

pMTL85141P <i>fdx</i> _GusA	Clostridial modular vector, with the <i>gusA</i> gene from <i>Streptococcus agalactiae</i> under the <i>fdx</i> promoter	Dr Zak Bean and Jessica Locker, University of Nottingham
pMTL85141P <i>qsrB</i> _GusA	Clostridial modular vector, with the <i>gusA</i> gene from <i>Streptococcus agalactiae</i> under the <i>qsrB</i> promoter	Dr Zak Bean and Jessica Locker, University of Nottingham
pMTL85141P <i>adc</i> _GusA	Clostridial modular vector, with the <i>gusA</i> gene from <i>Streptococcus agalactiae</i> under the <i>adc</i> promoter	Dr Zak Bean and Jessica Locker, University of Nottingham
pMTL85141P <i>adhE</i> _GusA	Clostridial modular vector, with the <i>gusA</i> gene from <i>Streptococcus agalactiae</i> under the <i>adhE</i> promoter	Dr Zak Bean and Jessica Locker, University of Nottingham
pMTL85141P <i>ptb</i> _GusA	Clostridial modular vector, with the <i>gusA</i> gene from <i>Streptococcus agalactiae</i> under the <i>ptb</i> promoter	Dr Zak Bean and Jessica Locker – University of Nottingham

Table 2.3 Oligonucleotides used in this study

Oligonucleotide	Sequence
ClosTron mutant screening	
CA_C0176_F	TGGGATGATTCTGGAGAGCC
CA_C0176_R	TCCCGTCAGTACCTTTCTTCC
CA_C2377_F	GCATGTGGACAAATTGCTGTG
CA_C2377_R	TCGTAGTAATCCTTCCAACCCT
CA_C2667_F	ATGGAGCGATGGGGATGATT
CA_C2667_R	TGTTTTGCTCTTCCAGGATATCA
CA_C3179_F	GGTATTGCCGTTGTAGACCC
CA_C3179_R	GTACGTTAGCCACCTCTCCA
CA_C3632_F	GGGTTGGGCAGCTAAGAAAC
CA_C3632_R	TGCTGCGTCCTTATTAACAAGT
CA_C3634_F	TGCTAAGTGGTCAGACGGAA
CA_C3634_R	GGAAAGCCCTTACCGTCTG
CA_C3640_F	GTAGTTCAAGCAGCACGTCA
CA_C3640_R	CCTTTCGGTCCGCAAGTTTT
EBS universal	CGAAATTAGAACTTGCGTTCAGTAAAC
RAM-F	ACGCGTTATATTGATAAAAATAATAATAGTGGG
RAM-R	ACGCGTGCGACTCATAGAATTAT TTCCTCCCG
Cloning and Gibson assembly primers	
QsrB_C_F1	TATATACCTGCAGGCTACATTACTCAAAGCATATAAATACG
QsrB_C_R1	TATATAGCGGCCGCTTACTTAACTTTATTAATAAATTTAATATTTTA TCTATGTC
QsrB_R_Strep	AACATCGGATCCTTTTTCAAAGTGCAGGATGGCTCCATATTTACTTACT TAAAC
F1_QsrB_KO_F	AAACAGCTATGACCGCAAATAATAGGCCTTATAGAAAATATAACC
F1_QsrB_KO_R	TCATATTTGTGTTTTAATCCTTTGCGCAATAG
F2_QsrB_KO_F	GGATTAAAACACAAATATGATATTTAATGTAAATTGAACAATAATG
F2_QsrB_KO_R	GGTCATATGGATACAGCAACTAATGGAAAATATTTGAATAAGTATG
F1_QsrB_HTH_RBS_KO_F	AAACAGCTATGACCGCTCATCTTGAGAACTTGTAGGAAATC
F1_QsrB_HTH_RBS_KO_R	ATAAACTTATAAAGAACTTAAAGGAAATCGC
F2_QsrB_HTH_RBS_KO_F	TTAAGTTCTTTATAAGTTTATTAGCACTAATAATATCCTTAC
F2_QsrB_HTH_RBS_KO_R	GGTCATATGGATACAGCTGTTAAACTAATGGAAAATATTTG
Gib_CA_C2667_F_85141	CTAGAGTCGACGTCACGCGTTCCATATTTCAATAAAAAGTAAAAATTA G
Gib_CA_C2667_R_85141	GTCTGCAGGCCTCGAGATCTATGCTGTTTTTATATTTAACAATATTAC
Gib_CA_C3632_F_85141	CTAGAGTCGACGTCACGCGTAAAAACATATACTGTAGTTTATACAAA ATATATAAG
Gib_CA_C3632_R_85141	GTCTGCAGGCCTCGAGATCTGCCTTAGAATACAGTGAG
Gib_CA_C2667_F_85143	CTAGAGTCGACGTCACGCGTATGAAGAATTTTATAAGAGTATGTATT G

Gib_CA_C2667_R_85143	GTCTGCAGGCCTCGAGATCTTTACCTTAATGAATTTACTATAAGC
Gib_CA_C3632_F_85143	CTAGAGTCGACGTCACGCGTATGAAAAGTAAAAAATTATTGTCAGTA ATAC
Gib_CA_C3632_R_85143	GTCTGCAGGCCTCGAGATCTTTATTTATTTTTTCCTGAAATATAAGCG
Generic sequencing primer	
M13-R	AGGAAACAGCTATGACCAT

Table 2.4 Synthetic peptides used in this study

Peptide Name	Sequence	Purity %	Information
QspB 10	LLGAEPTWGW	93.9	
QspB 10	LLGAEPTWGW	97.3	Independently synthesized
QspB 9	LGAEPTWGW	97.1	
QspB 9	LGAEPTWGW	99.2	Independently synthesized
QspB 8	GAEPTWGW	99.5	
QspB 8	GAEPTWGW	95.0	Independently synthesized
QspB 7	AEPTWGW	98.2	
QspB 7	AEPTWGW	99.6	Independently synthesized
QspB 6	EPTWGW	98.7	
QspB 5	PTWGW	97.6	
QspB 9-1	LGAEPTWG	92.8	
QspB 9-2	LGAEPTW	95.8	
QspB 9-3	LGAEPT	97.0	
QspB 9 <i>roseum</i>	AEPTRPW	97.7	<i>Clostridium roseum</i> homolog of QspB
QspB 7A2	AAPTWGW	97.2	Peptide generated for the alanine scan experiment
QspB 7A3	AEATWGW	97.0	As above
QspB 7A4	AEPAGW	98.8	As above
QspB 7A5	AEPTAGW	99.3	As above
QspB 7A6	AEPTWAW	99.1	As above
QspB 7A7	AEPTWGA	99.5	As above
QspA 9	LSGAPCIPW	94.8	Putative QsrA conjugate peptide
QspC 9	LAFYPGMPW	97.1	Putative QsrC conjugate peptide
QspD 9	LSMFPDAPW	96.5	Putative QsrD conjugate peptide
QspE 9	LSAFPDLPW	98.0	Putative QsrE conjugate peptide
QspF 9	LASYPGWSW	93.3	Putative QsrF conjugate peptide
QspG 9	LSVSPMTPW	92.4	Putative QsrG conjugate peptide
QspH 9	LSGLPLWVW	97.4	Putative QsrH conjugate peptide
QspB 8M	MAEPTWGW	97.5	Peptide generated to test the function of QspB active region with additional start codon for future overexpression work.

2.2 Microbiological methods

2.2.1 Media and growth conditions

E. coli and *C. acetobutylicum* ATCC 824 were maintained at 37 °C unless noted otherwise. *E. coli* was grown aerobically on Lysogeny Broth (LB, 10 g/l tryptone, 5 g/l yeast extract, 5 g/l sodium chloride, pH 7.2), and shaken at 200 rpm. 15 g/l of agar was used for solidified LB medium. When selection was required for *E. coli* strains the medium was supplemented with: 25 µg/ml chloramphenicol, 10 µg/ml tetracycline, 500 µg/ml erythromycin and 50 µg/ml ampicillin. When performing any prolonged expression in *E. coli* TOP10, the temperature was reduced to 30°C, to lessen the possible toxic effects of overexpression.

Two solid media were used routinely for the growth of *C. acetobutylicum* ATCC 824. Clostridial Growth Media (CGM - Table 2.5) and Clostridial Basal Media (CBM-S – Table 2.6) (O'Brien & Morris 1971, modified). Rich CGM medium was used whenever expedient surplus growth of the organism was required, such as post electroporation recovery, or for the revival of stocks. CBM-S medium was used when controlled growth of the organism was required; the use of casein acid hydrolysate amino acids rather than yeast extract leads to a less rich medium.

Table 2.5 CGM recipe

Component	Weight/Volume
Glucose	50 g
Tryptone	2 g
(NH ₄) ₂ SO ₄	2 g
Yeast extract	1 g
K ₂ HPO ₄	1 g
MgSO ₄ ·7H ₂ O	0.1 g
20g/l FeSO ₄ ·7H ₂ O	750 µl
20 g/l CaCl ₂	500 µl
20 g/l MgSO ₄ ·H ₂ O	500 µl
20 g/l CoCl ₂	100 µl
20 g/l ZnSO ₄	100 µl
Agar	10.2 g
Use reverse osmosis water to bring the volume to 1 litre and then autoclave	

Table 2.6 **CBM-S recipe**

Component	Weight/Volume
20 g/l $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	10 ml
Casamino Acids (acid hydrolysate)	4 g
20 g/l MnSO_4	379 μl
20 g/l $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	500 μl
1 g/l p-amino-benzoic acid	1 ml
1 g/l thiamine-HCl	1 ml
0.1 g/l biotin	20 μl
For solid media, 1.5% (w/v) agar added. Make up to 840 ml with reverse osmosis water and autoclave. Post autoclaving the following was added:	
5 g/l CaCO_3 (autoclaved)	20 ml
50 g/l K_2HPO_4 anhydrous (filter sterilised)	10 ml
50 g/l KH_2PO_4 anhydrous (filter sterilised)	10 ml
50 % glucose (w/v) (autoclaved)	120 ml

Two liquid media were routinely used. CBM-S without the addition of agar, and 2xYTG. 2xYTG was made with 10 g/l yeast extract 16 g/l tryptone 5 g/l sodium chloride and made up to 880 ml in reverse osmosis water and pH adjusted to 5.2, post autoclaving 120 ml of 50% (w/v) glucose was added for a final glucose concentration of 6% (w/v) Following the same premise as above 2xYTG was used whenever healthy, surplus growth of the organism was required, such as transformation recoveries and genome preparations. CBM-S liquid media was more widely used, and was used for all growth experiments.

When required for selection, *C. acetobutylicum* ATCC 824 growth media was supplemented with 15 $\mu\text{g/ml}$ Thiamphenicol or 50 $\mu\text{g/ml}$ erythromycin.

For the growth of strict anaerobes, a Don Whitley A95TG anaerobic workstation was used. An anaerobic environment was maintained by supplying a mixture of the gases hydrogen, carbon dioxide and nitrogen. The internal growth atmosphere was maintained at 10% hydrogen, 10% carbon dioxide and 80% nitrogen, with a temperature of 37°C. Contaminating oxygen was removed with a platinum-based catalyst. For the removal of oxygen, solid and liquid media were placed in the anaerobic workstation for at least 4 h and 16 h, respectively. Plastic ware introduced

into the anaerobic workstation for later use was generally transported via an airlock system, and incubated within the cabinet for a minimum of 4 hours, to allow oxygen diffusion from within the polymer.

2.2.2 Storage and revival of strains

E. coli strains were stored with the use of Microbank tubes (Pro-Lab Diagnostics Inc.). For liquid cultures, the liquid within the Microbank tube was removed and 1 ml of culture pipetted into the tube and gently mixed. Following a brief incubation at room temperature, the liquid culture was removed and discarded. The tube was then stored at -80°C. For cultures on solid media, a small number of colonies were picked with a 10 µl loop, and then resuspended within the Microbank tube. Again, following a brief incubation at room temperature the liquid component was removed and discarded. This was stored as previously stated.

During revival, a single CRYOBANK bead was taken from the tube under sterile conditions, and placed in 10 ml LB supplemented with the appropriate antibiotic. This was incubated at 37°C, shanking at 200 RPM (rounds per minute) for at least 12 hours before further use. If single colonies were required, a bead was removed and placed on the surface of an appropriate agar plate, where it was moved with a 10 µl inoculation loop to distribute attached cells prior to incubation.

C. acetobutylicum ATCC 824 strains were stored at -80°C. For the preparation of a DMSO stock, 900 µl of culture was taken and 100 µl of 100% DMSO added to preserve it from ice crystal damage. For revival, the tubes were placed in a -80°C storage block and transferred into the anaerobic workstation. A 10 µl loop of frozen cells was spread onto a CGM plate containing appropriate antibiotic. If single colonies were required an appropriate streaking pattern was used.

Where possible, *C. acetobutylicum* ATCC 824 strains were stored as spore stocks. The strains were grown in 10 ml 2xYTG for at least 5 days. Following growth, the samples were centrifuged at 5000 x g for 10 minutes and then washed in an equal volume PBS to remove media components and cell debris. After resuspension and a second round

of centrifugation the supernatant was removed and the pellet resuspended in 1 ml of PBS. Spore stocks were stored indefinitely at 4°C. In order to revive these stocks, 1 µl was diluted in 30 µl of PBS, this was heat shocked at 80°C for 10 minutes in a PCR machine or aluminium heating block and then plated on a CGM plate with appropriate antibiotic selection.

It should be noted that *C. acetobutylicum* ATCC 824 was always revived from spore stock on solid medium; as a general rule the organism grows poorly, if at all, when inoculated directly from heat treated spore stock into liquid media.

2.2.3 Measurement of optical density (OD)

Culture growth was measured through OD at a wavelength of 600 nm (OD₆₀₀). All ODs were taken relative to sterile media blanks. Measurement was achieved using a 1 cm gap cuvette, and a BioMate 3 spectrophotometer (BIO-RAD). This spectrophotometer only returned accurate OD measurements below an OD of 1.0. For this reason any sample with a greater measurement then this was diluted by 1:10 in the same batch of sterile medium. Cultures were thoroughly mixed before sampling, by pipetting up and down for a few seconds. Occasionally, undissolved CaCO₃ particles present in some of the media lead to disruption of the measurement and a higher than expected OD was recorded. In these circumstances, cuvettes were left for roughly 1 minute to allow the CaCO₃ to settle, and the OD measurement was repeated.

2.2.4 Initiation of growth experiments

Post revival, *C. acetobutylicum* ATCC 824 strains were grown on CGM agar or CBM-S agar, and in the liquid media CBM-S or 2xYTG. For the initiation of growth experiments *C. acetobutylicum* ATCC 824 cultures were used within the exponential phase of growth. This was done by inoculating 10 ml of CBM-S liquid media with a small number of colonies, then serially diluting to 10⁻⁴. All dilutions were incubated overnight. The following day dilutions that displayed OD₆₀₀ between 0.3 and 0.8 were used to initiate the growth experiment with a starting OD of 0.05. Following the inoculation, OD was measured every hour as previously described, to follow growth of the organism.

2.3 Molecular biological methods

2.3.1 Isolation of plasmid DNA from *E. coli*

Plasmid DNA was isolated using the Sigma-Aldrich GenElute Plasmid Miniprep kit according to the manufacturer's instructions except that DNA was eluted with 50 µl Elution solution (10 mM Tris-HCL, 1 mM EDTA). Plasmid preparations were stored at -20°C. A secondary kit was also used, the New England BioLabs Monarch Miniprep Kit, again the only alteration to the manufacturer's protocol was the reduction in elution buffer to 50 µl.

2.3.2 Isolation of genomic DNA from *C. acetobutylicum* ATCC 824 via column purification

A Sigma-Aldrich GenElute Bacterial Genomic DNA kit was used for the isolation of gDNA from *C. acetobutylicum* ATCC 824. 5 ml of bacterial culture was pre-treated with 100 µg/ml of lysozyme for 30 minutes at 37°C, following this the manufacturers' protocol was followed, except that the final elution in elution solution (10 mM Tris-HCL, 0.5 mM EDTA) used a volume of 100 µl.

2.3.3 Isolation of genomic DNA from *C. acetobutylicum* ATCC 824 via phenol chloroform extraction

For increased yields of gDNA from *C. acetobutylicum* ATCC 824 a phenol chloroform extraction was performed. 10 ml of an overnight culture was centrifuged at 10,000 g for 5 minutes at room temperature. The pellet was resuspended in 180 µl of PBS containing 10 mg/ml lysozyme and incubated at 37°C for 30 minutes. 20 µl of 20-40 mg/ml RNase A solution (Sigma-Aldrich) was added to the lysate and incubated for 10 minutes at 37°C. 25 µl of 20 mg/ml protease K (Sigma-Aldrich) was added with 110 µl 10% (w/v) SDS solution, and the mixture incubated at 65°C for 30 minutes.

An equal volume of phenol-chloroform-isoamyl alcohol (ration 25:24:1, saturated with 10 mM Tris, 1 mM EDTA and pH 8) was added to each sample, which were then

thoroughly mixed via inversion. The samples were loaded into Phase Lock Gel tubes (Quantabio) and centrifuged for 3 minutes at 16,000 g at room temperature. The aqueous top layer was then removed and combined with equal volumes phenol-chloroform-isoamyl alcohol, and following thorough mixing again, placed into a phase lock tube. The samples were then centrifuged as before. This purification process was repeated 3 times in all.

Following the final purification step, the aqueous phase was transferred to a 1.5ml reaction tube and 40 µl 3 M sodium acetate with 800 µl ice-cold 100% ethanol added for DNA precipitation. DNA precipitation was encouraged by incubation at -80°C for 30 minutes. Precipitated DNA was pelleted via centrifugation at room temperature and 16,000 g for 1 minute. The supernatant was removed and the pellet washed with 70% ethanol. Following another centrifugation step as before, the supernatant was carefully removed via pipetting; any additional supernatant was left to air dry for at least 30 minutes. The final DNA pellet was resuspended gently in 50 µl of distilled water. As with other DNA samples, storage temperature was -20°C.

2.3.4 Methods of DNA transfer

2.3.4.1 Generation of competent cells and transformation of *E. coli*

10 ml of LB was inoculated with the *E. coli* strain required for transformation. This was incubated overnight at 37°C. The following day the entire overnight culture was used to inoculate 200 ml of LB in a 500 ml conical flask or, preferably, a baffled flask which allowed head space for sufficient gas exchange. This was incubated at 37°C shaking at 200 RPM, and the OD measured every hour until an OD of 0.6 to 0.8 was reached.

Following this, cells were harvested via centrifugation at 4,000 g and 4°C for 5 minutes. Cell pellets were washed in 40 ml ice-cold distilled water, followed by another centrifugation step. Following this cell pellets were washed again with 20 ml of ice-cold distilled water and the centrifugation was repeated. The final cell pellet was resuspended in 1 ml 10% (v/v) glycerol. Electrocompetent cells were stored in 50 µl aliquots in microcentrifuge tubes at -80°C until required.

When required, an aliquot of competent cells were thawed on ice and 2 µl (150-300 ng) of a plasmid preparation added to the microcentrifuge tube. The mixture was transferred to a 2 mm gap electroporation cuvette (Bio-Rad), which had been pre-cooled on ice, and the mixture briefly incubated on ice for approximately 2 minutes. Electroporation was performed at 200 Ω, 2.5 kV and 25 µF with a Bio-Rad MicroPulser. The electroporated cells were immediately resuspended in 1 ml LB medium and left to recover at 37°C for 1-3 hours, shaking at 200 RPM.

Following recovery, 200 µl of the transformation was plated on LB agar medium containing the required selection antibiotic. A serial dilution was performed if required. The resulting plates were incubated at 37°C overnight.

2.3.4.2 Generation of competent cells and transformation of C. acetobutylicum ATCC 824.

Initially, *C. acetobutylicum* ATCC 824 was revived from a spore stock on CGM agar. Following overnight growth, a small number of colonies were used to inoculate 10 ml 2xYTG medium. The inoculated tube was serially diluted to 10⁻⁴ and all dilutions incubated overnight. The following day, the dilution that had shown the least amount of growth, between OD 0.4-0.8, was used to inoculate a flask of 75 ml 2xYTG medium.

OD measurements were taken every hour until the OD reached 0.6-0.8. At this point the cells were pelleted via centrifugation at 4,000 g and 4°C for 10 minutes. At all times the cells were kept cool by transporting on ice, each step being performed swiftly. The supernatant was poured off inside the anaerobic workstation and the pellet resuspended in 10 ml ice-cold electroporation buffer (EPB).

EPB buffer consisted of 19 ml 284 mM sucrose solution and 1 ml 100 mM sodium phosphate buffer, which in turn was made of 19 mM monobasic sodium phosphate and 81 mM dibasic sodium phosphate.

Resuspended pellets were centrifuged as before and the supernatant poured off once more. The pellet was resuspended in 2 ml of EPB. Aliquots of 540 µl were made and preferably used immediately; if not they were mixed with 60 µl of DMSO and stored at -80°C until required.

During electroporation, 20 µl (1.5-30. µg) of appropriately methylated plasmid (2.3.5.7) was placed within a pre-chilled 4 mm gap electroporation cuvette. These were transported on ice to the anaerobic workstation where aliquots of electrocompetent *C. acetobutylicum* ATCC 824 were added. The mixture was incubated for 2 minutes. Following this, the cuvette was pulsed at 1 Ω, 2 kV and 25 µF, and the electroporation mixture immediately resuspended in 10 ml 2xYTG. The cells were left to recover for 3-4 hours, before centrifugation at 4000 g at room temperature for 10 minutes. The cell pellet was resuspended in 1 ml 2xYTG, and the resuspension plated on CGM plates containing the appropriate selection antibiotic.

2.3.4.3 Dialysis of DNA

Prior to DNA transformation, excess salt from the plasmid preparation could be removed by dialysis. This was performed by floating a 0.0025 µm nitrocellulose membrane (Millipore) on deionised water within a Petri dish, and loading 10 µl of plasmid preparation onto the membrane. After 15 minutes the 10 µl of plasmid preparation was carefully taken up, and subsequently used for transformation.

2.3.5 DNA manipulation and analysis

2.3.5.1 Determination of quality and concentration of DNA samples

Preparations of DNA were analysed using a Nanodrop ND-1000 spectrophotometer (ThermoFisher). Samples were quantified at a wavelength of 260 nm, and purity reflected by the ratio of measurements at 260 nm/280 nm and 260 nm/230 nm (Desjardins and Conklin, 2011). A 260 nm/280 nm ratio of 1.8 is generally accepted as pure for DNA. A 260 nm/230 nm ratio between 1.8 and 2.2 is generally considered pure for DNA (Desjardins and Conklin, 2011). Purity ratios take into account

common contaminants of the sample, typically solvents and reagents used during sample preparation.

2.3.5.2 DNA amplification through polymerase chain reaction (PCR)

Two polymerase sets were used throughout the course of this study, DreamTaq master mix (Thermo Scientific) and Phusion 2x master mix (NEB). All routine screening PCR reactions were performed using DreamTaq master mix in a 20 µl total volume. 100-200 ng of template DNA was added to 10 µl of DreamTaq master mix. 1 µl of 10 µM forward and reverse primers were added, and the final volume made up to 20 µl with nuclease-free water (NEB). Where DNA amplification was needed for cloning, Gibson assembly or sequencing Phusion 2x master mix was used. 100-200 ng of template DNA was added to 10 µl of Phusion 2x master mix, and 1 µl of 10 µM forward and reverse primers added with the final volume being made up to 20 µl with nuclease-free water (NEB).

Thermocycler programs were as indicated in the polymerase master mix instruction, with annealing temperature set to 5°C below the primer melting temperature or 50°C for a general program. PCR reactions were performed in a T3000 Thermocycler (Biometra) or a Sensoquest labcycler (Geneflow). When required, PCR products were purified using a Zymo DNA clean & concentrator kit (Zymo Research) as per the manufacturers' protocol.

2.3.5.3 Restriction digests and Ligation

Restriction enzymes were supplied by ThermoFisher and were of the FastDigest variety. This was due to reduced digest times and reduction in star activity. Digestions were performed according to the manufacturers' instructions using 1 µg of DNA, 1 µl of each required restriction enzyme, 10 µl of FastDigest Green 2x reaction buffer (ThermoFisher) and the final volume made up to 20 µl with deionised water.

T4 DNA ligases used during the course of the study were from NEB, ThermoFisher or Promega. Ligation mixes were combined to an insert to vector ratio of 3:1 according to the manufacturers' instructions and incubated following one of three

regimes independent of manufacturer: A) Components were incubated for an hour at room temperature. B) Components were combined in a microcentrifuge tube and placed on a 50:50 mix of ice and water, the mixture was then incubated overnight at room temperature. C) Components were combined in a microcentrifuge tube and then incubated at 4°C for a minimum of 12 hours.

2.3.5.4 DNA gel electrophoresis and DNA recovery

Agarose gels were made to a percentage (w/v) appropriate for size of DNA fragments to be separated; in the vast majority of cases agarose gels were made to 1% (w/v). Agarose was stored as a solidified 1% (w/v) stock dissolved in TAE (40 mM Tris, 1mM EDTA and 0.1% (v/v) glacial acetic acid) buffer.

This stock was melted in a microwave, and placed in a water bath at 50°C for 15 minutes to reduce handling temperature. Following this, the gel was poured into a casting tray containing the appropriate comb to generate wells, and also containing the appropriate amount of SYBR Safe DNA stain (ThermoFisher), which was 1:10,000 (v/v). The gel cast was mixed on the tray while still molten to ensure equal distribution of SYBR Safe. The gel was allowed to set for 15 minutes. Following this the gel was placed into an electrophoresis tank filled with 1x TAE buffer. Loading dye was mixed with the samples if appropriate (ThermoFisher), and the samples were loaded into the gel wells after comb removal. This was not necessary when DreamTaq master mix (ThermoFisher) or FastDigest enzymes (ThermoFisher) were used which already contained loading dye buffer. For DNA fragment size determination a 2-log DNA ladder was run alongside all samples (ThermoFisher or NEB).

Gels of 100 ml in volume were run at 100 mV for 60 minutes. Gels of 200 ml in volume were run at 120 mV for 60 minutes. A number of different power packs were in operation during the project, all manufactured by Bio-Rad. Gel casting trays and electrophoresis tanks were manufactured by Biometra. Post electrophoresis, gels were visualised using trans-UV light within a Bio-Rad GelDoc system XR+. The images were recorded via the Image Lab software (Bio-Rad).

Should fragments of DNA be required for a further purpose, the bands were cut from the gel using a scalpel. Gel recovery was preformed using a Zymoclean Gel DNA recovery kit (Zymogen), or with a Monarch Gel extraction kit (NEB) following the manufacturer's instructions.

2.3.5.5 Gibson assembly

Fragments that required generation for Gibson assembly were created using Phusion master mix (ThermoFisher) as described (2.3.5.2). The plasmids within which the fragments were to be included were linearised with either NotI or NcoI digestion as described (2.3.5.3). The fragments were combined with the linearised plasmid in a ratio of 1:2, in a 20 µl reaction volume with 10 µl Gibson assembly master mix (NEB) without the total amount of fragments exceeding 0.2 pmol as per the manufacturer's instructions. The reaction volume was made up to 20 µl with nuclease-free water (NEB). The reaction mixture was incubated in a thermocycler at 50°C for 15 minutes, and following this transformed to *E. coli* TOP10 (2.3.4.1).

2.3.5.6 DNA sequencing

Source Bioscience Gene Service (UK, Nottingham) was used for the Sanger sequencing of DNA samples. Samples were prepared as per the service providers instructions and collected in 1.5 ml microcentrifuge tubes. Whole genome sequencing (Illumina sequencing) was performed by MicrobesNG (Birmingham, UK), and samples prepared as per the service providers instruction.

2.3.5.7 Methylation of Plasmid DNA

Prior to transformation of *C. acetobutylicum* ATCC 824, plasmids required appropriate methylation. This was performed by transforming the desired plasmid into *E. coli* TOP10 (pAN2) (2.3.4.1). Following recovery, the cells were spread on LB supplemented with tetracycline and the appropriate antibiotic for selection. Following overnight incubation, confluent growth was harvested and resuspended in 400 µl resuspension buffer from a plasmid isolation kit (2.3.1). The manufacturers' plasmid

isolation procedure was then followed, but due to the increased bacterial content and viscosity the volumes of the lysis and neutralisation steps were doubled.

2.3.6 Protein expression and analysis

2.3.6.1 Strep-tag purification of *QsrB*

E. coli strains carrying the overexpression plasmid pMTL85143-*qsrB*-Strep (Table 2.2) were grown in 200 ml LB medium at 30°C overnight alongside a vector control strain. The total volume was separated into 50 ml centrifuge tubes that were centrifuged at 5,000 g for 10 minutes. Each pellet was resuspended in 5 ml of PBS and washed before being combined into a single 50 ml centrifuge tube and a second centrifugation under the same conditions. After pouring off the supernatant, the resulting pellet was either processed as described below or stored at -20°C overnight when required.

E. coli cells were lysed using BugBuster (Novagen). The 10x concentrate variety (BugBuster 10x protein extraction reagent) was used and diluted to a working concentration in Strep-tag wash buffer (100 mM Tris Base and 150 mM NaCl, pH 8). 10 ml of BugBuster was added to the cell pellet, which was placed on a gently shaking platform for 1 hour at room temperature. Following the lysis step, a 40 µl sample was taken for later SDS-PAGE analysis. The cell debris were removed by centrifugation at $\geq 12,000 \times g$ for 1 minute, and again 40 µl samples taken from both the cell debris (after resuspension in Strep-tag wash buffer) and the clear lysate.

Affinity chromatography was carried out using a Poly-Prep Chromatography column with a 10 ml reservoir (Bio-Rad). After thorough mixing, 1 ml of 50% (w/v) Strep-Tactin Sepharose (IBA Life Sciences) was loaded into the column, and the liquid drained. Following this 10 ml of Strep-tag wash buffer was loaded onto the column for equilibration, and the effluent removed. The total lysate was then passed through the column, and 40 µl samples taken from the eluate as previously described. Following this, 1 ml of Strep-tag wash buffer was run through the column, this was step was repeated 5 times, and each time SDS-PAGE samples were taken as described above. Following the wash steps, the elution steps were performed by running 0.5 ml

of Step-tag elution buffer (as wash buffer but with 2.5 mM desthiobiotin) through the column and storing the effluent at -80°C. The elution step was performed 6 times. Samples were taken at each step as previously described. Following the elution steps, a final wash step was performed by running 5 ml of Strep-tag wash buffer through the column. The final wash was performed twice, and SDS-PAGE samples taken.

In order to regenerate the column for future use, it was washed with 10 ml of HABA buffer. HABA buffer was as described for Strep-Tag wash buffer, but with the addition of 1 mM 4'-hydroxyazobenzene-2-carboxylic acid. A colour change from yellow-orange to red indicated the removal of d-desthiobiotin from the Strep-Tactin Sepharose, which was then washed out by running Strep-Tag wash buffer through the column until it returns to a white colouration.

2.3.6.2 SDS-PAGE gel electrophoresis

Samples for SDS-PAGE gel electrophoresis were obtained as explained in section 2.3.6.1. Premade Bis-Tris 4 to 12% acrylamide gels were used for SDS-PAGE analysis (ThermoFisher – NuPAGE Bis-Tris gels). These were loaded into an XCell SURELock Mini-cell Electrophoresis system (ThermoFisher), and the central gap in connection with the cathode was filled with NuPAGE MOPS SDS running buffer (ThermoFisher). The tank was checked to ensure no running buffer was leaking into the outer chamber connected with the anode, as this would prevent running. Following this the outer chamber was filled with the before mentioned running buffer until it was in contact with the gel.

Each of the samples were mixed with NuPAGE LDS Sample Buffer loading dye (ThermoFisher), and were denatured by heating at 100°C for 10 minutes.

The comb of the gel was removed and 10-20 µl of sample loaded into each well; any empty wells were filled with loading buffer diluted with H₂O to prevent “gel smiling”. For protein size determination a BenchMark Prestained Protein ladder (Invitrogen) was run alongside the samples. The gel tank was connected to a BIO-RAD power pack 300 and run for 60 minutes at 120 mV. The gel was assessed for ladder distribution, and running continued in 10 minute intervals if deemed appropriate.

Following running, the gel cassette was opened and the gel washed briefly in H₂O. The gel was then submerged in Coomassie blue staining buffer and gently shaken for an hour. Coomassie blue staining buffer contained 2.5 g Coomassie R-250, 450 ml methanol, 450 ml H₂O and 100 ml glacial acetic acid. Following this, the gel was briefly washed in H₂O, and then submerged in destaining buffer, and gently shaken for between 4 hours and overnight. Destaining buffer contained 100 ml glacial acetic acid, 500 ml methanol and 400 ml H₂O. The gels were visualised using a GelDoc system (Bio-Rad). The images were recorded via the Image Lab software (Bio-Rad).

2.4 Synthesis of DNA constructs and peptides

2.4.1 Generation of re-targeted ClosTron plasmids

The pMTL007C-E2 ClosTron plasmid was re-targeted using the intron targeting and design tool available on <http://www.clostron.com/clostron2.php> (Heap et al., 2010). Targeting regions were chosen based on the encoded protein's structure, ideally within the coding region of an active site, and the highest score given by the targeting algorithm (www.clostron.com). The before mentioned algorithm compensates for the lack of simple rules regarding intron re-targeting, as the generated ribonucleoprotein complex recognises multiple nucleotide residues within the target site and none are absolutely required (Perutka et al., 2004). The algorithm was trained using 3 databases of re-targeted introns, and effectively supplies multiple possible re-targeted sequences attributing them an arbitrary score. The higher the score, the more likely the re-targeted sequence is to be specific to the desired target region and hence inversely, less likely to produce ectopic insertions (Perutka et al., 2004). Following the choice of targeting sequence, the plasmids were synthesised by DNA2.0, and delivered lyophilised on filter paper. A syringe tip was used to pierce a hole into a 0.2 –ml-reaction tube, and the filter paper placed over the hole. The 0.2 ml reaction tube was placed inside a 1.5 ml microcentrifuge tube, and 50 µl of water loaded on top of it. This was centrifuged briefly, and the resulting elution was used to transform *E. coli* for long term plasmid storage.

2.4.2 Generation of synthetic peptides

Synthetic peptides were synthesised by PeptideSynthetics, Hampshire UK, and made to a minimum purity of 90%. A number of peptides were also reordered to a minimum purity of 95%, as shown in Table 2.4. Peptides were supplied as a lyophilised powder. Delivered tubes were centrifuged briefly to avoid loss of peptide, and made to a stock concentration of 20 mM using pure DMSO (Sigma). To avoid stability issues, peptides were only dissolved prior to first use, and stored at -80°C. The synthetic peptides used in this study are listed in Table 2.4.

2.5 Genetic manipulation of *C. acetobutylicum* ATCC 824

2.5.1 Isolation and confirmation of in-frame deletion mutants

Following construction of the appropriate in-frame deletion cassette via Gibson assembly (2.3.5.5) within pMTL-ME3, the plasmid was methylated (2.3.5.7) and then electroporated into *C. acetobutylicum* ATCC 824 Δ pyrE (2.3.4.2). Putative transformants were selected based on Thiamphenicol selection, and larger colonies suspected to be single crossover events were restreaked on CBM-S agar. Following overnight incubation, colonies were picked and restreaked on CGM agar with uracil (2 µg/ml) and 5-FOA (200 µg/ml) supplementation. Successful overnight growth indicated a double crossover event or reversion to wild type, colonies were picked and used to inoculate 5 ml 2xYTG supplemented with uracil (20 µg/ml) for genomic DNA preparation. Colonies were also screened for plasmid loss by restreaking on CGM agar supplemented with Thiamphenicol (2.2.1). Following genomic DNA isolation of possible double crossover events (2.3.1), a PCR screen was performed using flanking primers.

2.5.2 Isolation and confirmation of ClosTron mutants

ClosTron plasmids were synthesised as described under section 2.4.1. and used to transform *Clostridium acetobutylicum* via electroporation as previously described (2.3.4.2) Putative transformants were selected for by plating on CGM agar plates containing 15 µg/ml Thiamphenicol. After 24 hours of growth, putative transformants were restreaked on CGM agar containing 50 µg/ml erythromycin, selecting for cells expressing the erythromycin retroactive marker (RAM). After 24 hours growth,

colonies were picked and resuspended in 5 ml of 2xYTG medium supplemented with erythromycin. Overnight growth was then used for a genomic DNA preparation, and screened by PCR for a 1.9 kb ClosTron insertion using insertion site flanking primers (Table 2.2). Primers flanking the RAM were used to indicate presence, as well as a junction PCR using the EBS universal primer and the forward flanking primer, as the ClosTrons were inserted in the sense strand (Table 2.2).

2.6 Phenotypic analysis of *C. acetobutylicum* ATCC 824

2.6.1 Preparation of *C. acetobutylicum* ATCC 824 cultures for phenotypic analysis

Following strain revival, (2.2.2) colonies were used to inoculate 10 ml of CBM-S broth, and were serially diluted to 10^{-4} to generate pre-cultures. A pre-culture dilution was selected based on having an OD₆₀₀ of between 0.4-0.9. This was used to inoculate 30 ml of CBM-S broth within a 50 ml centrifuge tube or 75 ml CBM-S within a 200 ml conical flask, to a starting OD₆₀₀ of 0.05, depending on the experimental parameters. Initial samples were taken after inoculation for initial solvent analysis. All phenotypic experiments were undertaken in triplicate. Biological replicates were performed for all generated mutants during phenotypic analysis; these were in triplicate or duplicate as stated in the relevant sections.

A variation of the above method was used for the peptide addition experiments conducted within Chapter 3 and Chapter 4. Pre-cultures were generated as previously stated, and used to inoculate a master culture of between 50 – 300 ml of CBM-S broth to a starting OD of 0.05. Following 4 hours of growth, 10 ml of master culture was aliquoted into a 15 ml centrifuge tube which contained either 5 µl of 20 mM experimental synthetic peptide dissolved in DMSO (Table 2.4) for a final concentration of 10 µM, or else 5 µl of DMSO as a control. This ensured the experimental conditions were identical apart from peptide addition or lack thereof.

2.6.2 Light microscopy

Where light microscopy was used, 10 µl of a sample was taken, placed on a microscope slide and a coverside placed over it, taking care to avoid air bubbles as per

standard microscopy technique. Slides were viewed under an inverted light microscope (Nikon Eclipse TE200) and an appropriate magnification used. A serial dilution was performed in PBS if the culture was too dense. Microscopy was normally used to look for signs of cell stress when trouble shooting problems with the anaerobic workstation, diagnose a potential contamination event or simply as visual confirmation of spore formation. Spore formation was monitored using phase contrast illumination without staining which showed spores as phase bright under the light microscope.

2.6.3 Granulose staining and screening for amylase activity

Strains that required granulose straining and screening for amylase activity, alongside a wild type control for comparison, underwent serial dilution as previously described (2.6.1). Following overnight incubation, a culture with between 0.4-0.9 OD₆₀₀ was selected, and streaked using a 10 µl loop on to CBM-S agar in which the calcium carbonate had been omitted for visual clarity, and CBM-S Starch agar. The latter was generated following the CBM-S recipe, but omitting calcium carbonate and the standard amount of glucose, and instead including 20 g/l of potato starch (Sigma Aldrich) and 2.5 g of glucose pre autoclaving. For the production of CBM-S starch plates, filter-sterilised K₂HPO₄ and KH₂PO₄ solutions (Table 2.6) were added after autoclaving as soon as the handling temperature was reduced and the resulting final starch agar poured into petri dishes. After plate setting and curing, plates were incubated at 4°C for at least 48 hours to allow starch precipitation.

Following inoculation, the CBM-S and CBM-S starch plates were incubated for 36-48 hours before staining. Following this the plates were flooded with Grams iodine solution (1 g of Iodine, 2 g of potassium iodide dissolved in 300 ml of distilled water) with granulose positive colonies changing colour to blue/black. Amylase activity was indicated on CBM-S starch plates as halos are formed around the respective colonies. Post staining this is shown as a transparent halo in a blue/black background. Pictures were taken with a Nikon COOLPIX S3700 digital camera.

2.6.4 Determination of heat-resistant colony-forming units (CFU)

Samples for heat-resistant CFU analysis were taken concurrently with samples for solvent analysis or at the investigation end point, following vigorous vortexing or mixing. 200 µl samples were taken from each culture and placed in a 96 well plate, to be diluted serially to 10^{-4} using PBS, again ensuring vigorous mixing. The plate was heated to 80°C for 10 minutes in a Sensoquest labcycler (Geneflow) as per standard heat shock protocol. The plate was then taken to the anaerobic workstation, where 5 µl of each dilution per sample was spotted onto dry CGM agar simultaneously using a multichannel pipet. Each of these was performed in triplicate. After the spotted samples had dried, the plates were overturned and left to incubate. Following 24 hour incubation, colonies of an appropriate dilution were enumerated, and the heat resistant CFU/ml calculated.

2.6.5 Gas chromatography

Quantification of acetone, butanol, ethanol, butyric acid and acetic acid present in bacterial culture supernatants was carried out using a Focus GC (Thermo Fisher). 1 ml samples were taken of the investigated culture and centrifuged for 3 minutes at $\geq 16,000$ g to remove cells and debris. 900 µl of supernatant was then removed and placed into a screw cap tube to prevent solvent loss, and stored at -20°C until required for analysis.

Samples were prepared by slowly defrosting the supernatants, mixing, removing 500 µl and placing it in a 1.5 ml microcentrifuge tube. This was acidified using 5 µl of 10 M H_2SO_4 to denature remaining proteins for removal, as well as protonating acids for extraction. The mixture was vortexed briefly before the addition of 500 µl of propyl propionate, with an internal standard addition of 50 mM valeric acid. Following this the mixture was intensely vortexed for 10 seconds, and centrifuged at $\geq 21,000$ g for 1 minute at 4°C. The formation of a bilayer allowed the top 300 µl to be removed containing the extracted solvents, which were placed into an appropriate gas chromatography (GC) vial.

The Focus GC was equipped with a 30 m TR-FFAP column with an internal diameter of 0.25 mm. Hydrogen was used as a carrier gas, and components detected via flame ionisation. The hydrogen carrier gas was supplied at 0.8 ml/min, with the internal flame maintained with compressed air (350 ml/min), hydrogen (35 ml/min) and nitrogen (30 ml/min). The injector and detector temperatures were set to 240°C and 270°C respectively. Resolution of peaks was performed with the column at an initial temperature of 40°C, after 2 minutes the temperature was ramped at 80°C/min to 150°C, and a final 45°C/min ramp to 210°C at which the temperature was held for 1 minute. 1 µl of sample was injected, which was performed by the use of an AI/AS 1310 series autosampler (ThermoFisher).

Two independent standard sets were used during each GC run for calibration. These were diluted from a 1 M stock, containing; acetone, butanol, ethanol, acetate and butyrate. This was diluted using ELGA water to concentrations of 200 mM, 150 mM, 100 mM, 50 mM, 30 mM, 10 mM, 5 mM, 3 mM, 1 mM, alongside a CBM-S broth and water blank. Extraction of these standards was performed in the same manner as the above mentioned samples.

2.6.6 High performance liquid chromatography of glucose consumption

High performance liquid chromatography (HPLC) was performed to analyse the glucose concentration of cultures alongside GC analysis of acids and solvents. Samples for glucose consumption were taken at the time of inoculation (time 0) and at the end of the experiment. A detailed account of the HPLC analysis method employed here is given by Schwarz and colleagues (2017). A 1 ml sample of culture was taken and centrifuged (3 minutes, 16,000 g) and 200 µl of cell free supernatant stored at -20°C until use. The samples were thawed and mixed with an equal volume of mobile phase containing internal standard (50 mM valeric acid in 0.005 M H₂SO₄). The mixture was filter sterilised through a 0.22 µm syringe filter (GE Healthcare Life Sciences) into a HPLC vial and capped. Samples were analysed using a Dionex Ultimate 3000 HPLC system (ThermoFisher) which was equipped with a Rezex ROA-Organic Acid H⁺ (8%) (150mm x 4.6mm x 8µm) column, a refractive index (35°C) and diode array detector at UV 210 nm, and a flow rate of 0.5 ml/min of

mobile phase (0.005 M H₂SO₄). The column was maintained at a temperature of 35°C for 30 minutes. A 20 µl injection volume was used. Glucose standards between 5 and 500 mM were used alongside a water blank, which were prepared in an identical manner to the samples.

2.6.7 iTRAQ proteomics sample preparation

After 24 hours of incubation, with cultures inoculated as previously described (2.6.1), the whole cell culture was placed on ice, swiftly transferred to a centrifuge, centrifuged at 5,000 g for 10 minutes at 4°C. The supernatant was removed and the pellet washed in 10 ml ice-cold PBS. This wash step was performed a further 2 times for a total of 3 PBS wash steps. Following this, the pellet was washed with 10 ml 0.5 M triethyl ammonium bicarbonate (Sigma Aldrich, pH 8.5), which was at 4°C. The samples were centrifuged as before and the supernatant poured away, following this the samples were snap frozen in liquid nitrogen and stored at -80°C for a maximum of 7 days. The samples were delivered on dry ice to the collaborators. iTRAQ analysis of the samples was carried out by Dr Mahendra Raut (University of Sheffield) in collaboration with Prof Phillip Wright (University of Newcastle). The protocol that was used for the downstream processing of the samples using an 8-plex isobaric labelling, as was performed within this study, is covered in detail in a recent publication covering iTRAQ in *C. acetobutylicum* ATCC 824 from the Wright Lab (Raut et al., 2016). The cells were disrupted via bead beating and unbroken cells and debris removed via centrifugation. Proteins were concentrated via acetone precipitation, before resuspension and eventual digestion to peptides by trypsin (Promega). Peptides were treated with an 8-plex iTRAQ reagent Multiplex kit (ABSciences) as per the manufacturers' protocol, with each sample receiving an isobaric label with a reporter region that has a differing mass. The samples were mixed and iTRAQ-labelled peptides off-line fractionated using hydrophilic interaction liquid chromatography (HILIC) using an Agilent 1100 HPLC (Berkshire, UK). Resuspension of dried peptides was performed in HILIC buffer A (80% (v/v) acetonitrile (ACN), 10 mM ammonium formate, pH 3). Samples were separated using a polyhydroxyethyl-A column, with a 5 µm pore size, 100 mm length, 4.6 mm ID (Polu LC, MD, USA) at a flow rate of 0.5 mL/min using an UVD170U detector

(Dionex/LC packings, The Netherlands) at 280 nm. HILIC buffer A and B (5% ACN, 10 mM ammonium formate, pH 5) were used to perform a 90 min gradient of 0% B for 10 min, 0-20% B for 15 min, 20-40% B for 30 min, 40-60% B for 15 min, 60-100% B for 5 min and 100% A for 15 min. Fractionation and the chromatogram were monitored using Chromeleon software (Thermo, Hamel Hempstead, UK). Mass spectrometry was performed by solubilising HILIC fractions in reverse phase (RP) buffer A (3% (V/V) acetonitrile, 0.1% (v/v) formic acid (FA)), and subjected to a QstarXL Hybrid ESI Quadrupole Time-of-Flight Tandem mass spectrometer (ABSciex, Farnham, MA) coupled with an online Ultimate 3000 HPLC system (Dionex, Surrey, UK). Reverse-phase peptide separation was performed on a C18 Acclaim PepMap100 column (3 μ m, 100 Å, 15 cm) at a flow rate of 300 nL per min. A gradient was applied of RP buffer A to RP buffer B (97% (v/v) ACN, 0.1 % (v/v) FA) 0-3% B for 5 min, 3-35% B for 90 min, 35-90% B for 0.5 min, 90% B for 6.5 min, with a final 3% buffer A for 18 min. Data were acquired in positive ion mode in the data-dependant acquisition mode. The MS survey scan was set to cover the m/z range of 350-1800 Th, and the MS/MS survey scan set to the m/z range of 100-1600 Th using Analyst QS 2.0 software (ABSciex, Farnham, MA). Peptide of charge +2,+3 and +4 for each TOF-MS scan (400 – 1250 m/z) were dynamically selected and isolated for MS/MS fragment ion scans (100-1600 m/z). Two RP-HPLC_MS runs per HILIC fraction were performed. The iTRAQ reporter region ions allowed for the relative quantification of proteins within the samples, and were calculated using the pipeline described by Ow and colleagues (2009). The protein database was used for *C. acetobutylicum* ATCC 824 (taxon ID: 272562) containing 3825 protein sequences, downloaded from Uni-prot (www.uniprot.org, November 2017).

2.6.8 GusA reporter assay

Investigated *C. acetobutylicum* ATCC 824 cultures, grown as before stated (2.6.1), and containing appropriate GusA reporter plasmids (Table 2.1 and 2.2) were sampled at 8 hours, 24 hours and 48 hours. Putative reporter regions of the investigated genes were used, these consisted of the 200bp upstream region from the first codon. A 1 ml sample was taken, and centrifuged at 16,000 g for 1 minute. The supernatant was removed, and the pellet resuspended in 500 μ l of Z-buffer (60 mM Na₂HPO₄, 45 mM

NaH₂PO₄, 10 mM KCl, 1 mM MgSO₂·7H₂O, adjusted to pH 8.5). 75 µl of each resuspension was placed in an appropriate clear bottomed 96 well plate for plate reading. The reaction was initiated by adding 25 µl of 40 µg/ml 4-methylumbelliferyl-β-D-glucuronide dissolved in Z-buffer, this was performed via multichannel pipet. The plate was immediately placed in a CLARIOstar plate reader (BMG labtech), and the fluorescence at 440-460 nm measured using an excitation profile of 255-275 nm. The CLARIOstar was set to a distance of 7 mm and gain to 750. Each reaction was allowed to take place for 15 – 30 minutes. The relative activities were calculated by taking the linear gradient of the fluorescence, and correcting for the OD₆₀₀ of the culture.

2.6.9 Attempted swimming and swarming assays

Swimming assays were conducted by creating serial dilutions of the strains to be investigated as previously described (2.6.1) and following overnight incubation a culture of between OD₆₀₀ 0.4-0.9 was selected. 2 µl of this growing culture was stabbed in the centre of a semi-solid 2xYTG plate (as with 2xYTG but with addition of 0.3% agar, and made to 0.5% glucose post autoclave). These plates cannot be inverted for incubation. Following inoculation, growth was analysed after 24 hours during which a halo of growth should be measured.

Swarming assays were performed as above with the use of an altered semi-solid 2xYTG. This was made as 2xYTG but with 0.24% Gelzan CM (Sigma Aldrich), the medium was made up to 0.5% glucose post autoclave. Colony growth was analysed every 24 hours for 120 hours, following the deposit of 2 µl of culture on the surface of the agar.

2.7 Determination of enzyme activities

2.7.1 Generation of crude lysate via sonication

Whole cell cultures from 24 hours of incubation, inoculated as described (2.6.1), were centrifuged anaerobically for 10 minutes at 5000 g and 4°C. The supernatant was removed and the cell pellet snap frozen in liquid nitrogen, before storage at -80°C for

no more than 48 hours. Upon use the cell pellets were resuspended in 1 ml of 700 mM Tris-HCL, and anaerobically underwent sonication. Keeping the cell cultures on ice, the samples underwent sonication for a total of 10 minutes, with 30 seconds of sonication followed by 30 seconds of rest. Following this, the samples were centrifuged at 4°C and 5,000 g for 90 minutes to remove cell debris. The protein content of the lysate was calculated using a Bradford protein assay (Bio-Rad) that was conducted as per the manufacturers' instructions. This crude lysate was used for the enzymatic assays of *C. acetobutylicum* ATCC 824.

2.7.2 Hydroxamate assay of acetate and butyrate kinase activity

The reaction mixture (0.6 ml) contained the following components (final concentrations), 200 mM Tris-HCl (pH 7.4), 10 mM ATP, 267 mM potassium acetate or 267 mM of sodium butyrate, 11 mM MgCl₂, and 6% (w/v) hydroxylamine hydrochloride. The hydroxylamine hydrochloride was neutralised with potassium hydroxide before addition (adapted from Winzer et al. 1997). 100 µl of crude lysate was used to initiate the reaction, which was conducted for 20 minutes at 37°C. The reaction was vigorously agitated and debris removed by centrifugation at 5000 g for 1 minute. The reaction was stopped by the addition of 0.3 ml 10% (w/v) trichloroacetic acid. Colour was developed by the addition of 0.6 ml of 1.25% (w/v) FeCl₃ in 1 M HCl. The formation of the chloride-hydroxamide acid complex was measured at 540 nm. The extinction coefficient of hydroxamate (169 M⁻¹cm⁻¹) was used to calculate the enzymatic activity.

2.7.3 Butanol dehydrogenase assay

This assay was performed anaerobically and was used as in Dürre et al. 1987. This was done by combining the reaction components within a quartz cuvette with a silicon stopper (Hellma Analytics). The reaction mixture (1 ml) contained (final concentrations) 77 mM Tris-HCl (pH 7.6) and 0.23 mM NADPH, the final volume was made up to 959 with deionised water. 50 µl of crude lysate was added, and the reaction was initiated by the injection of 1 µl 99% butyraldehyde (Sigma Aldrich). The decreasing absorbance at 340 nm was measured for the consumption of NADPH,

and the subsequent gradient used for the calculation of enzymatic activity. The extinction coefficient of $28 \text{ mM}^{-1}\text{cm}^{-1}$ was used for this calculation.

2.7.4 Nitroreductase assay

The assay was performed as per personal correspondence with Dr Kubiak-Bailey, University of Nottingham. The reaction mixture (1 ml) contained, (final concentrations) 10 mM Tris-HCl (pH 7.5), 1 mM menadione, 10 mM NADH and 700 μM cytochrome C. The reaction was initiated by the addition of 10 μl of crude lysate. The reduction of cytochrome C was measured at 550 nm, and the gradient of this reaction used to calculate the nitroreductase activity using the extinction coefficient of cytochrome C ($28 \text{ mM}^{-1}\text{cm}^{-1}$).

2.8 *In silico* software and computer tools

The following computer tools were used during the course of the current work:

- Basic Logical Alignment Search Tool (BLAST): Used for the comparison of nucleotide sequences, and searching for homologous genes across organisms. BLASTp was used for protein identity comparisons (NCBI Resource Coordinators, 2017).
- Ugene: Bioinformatics software used mainly for amino acid sequence visualisation. However, it is a powerful software with capabilities in multiple sequence alignments, genome sequencing data analysis and more (Okonechnikov et al., 2012).
- GENTle: Also used for amino acid visualisation, due to user friendly alignments and virtual PCR functions (<http://gentle.magnusmanske.de/>).
- FinchTV: For the visualisation of DNA sequences traces, used for troubleshooting sequencing anomalies, and trimming trace files.
- GraphPad Prism: Used for the generation of figures and data visualisation; statistical significance levels were determined via two tailed t-tests performed with this software (Prism version 7 for windows www.graphpad.com).
- Chromeleon 7.2 Chromatography Data System: used for signal analysis of the HPLC data (Thermo Fisher).

- Microsoft office applications: Microsoft Excel 2013 (Microsoft) was used for the storage of raw data, and for the calibration of analytics data, standard error was calculated using this software. The calculation of enzymatic activity was also done within this software package. The body of the thesis was composed within Microsoft Word 2013 (Microsoft).

3. Characterisation of the RNPP type system CA_C0324 (QsrB) and its conjugate peptide (QspB)

3 Characterisation of the RNPP type system CA_C0324 (QsrB) and its conjugate peptide (QspB)

3.1 Introduction

Members of the phylum Firmicutes often engage in cell-cell communication via oligopeptides, that are linear or cyclic depending on the system, and are often referred to as autoinducing peptides (AIPs), signalling peptides or pheromones (Slamti et al., 2014). The signalling systems involved usually conform to one of two general types. First, the peptide can be exported, modified and then externally recognised by a membrane associated histidine kinase. This in turn leads to the phosphorylation of a response regulator protein, which goes on to regulate gene expression (Slamti et al., 2014). The accessory gene regulator (agr) systems conform to this definition, including the recently characterised *Clostridium acetobutylicum* agr system which was shown to regulate spore formation, as well as the formation of the putative storage compound granulose (Jabbari et al., 2013; Steiner et al., 2012). The second family of signalling systems relies on the re-uptake of the secreted and processed peptide via ABC transporters. Following its entry into the cell, binding of the peptide to a regulator leads to the latter's activation or inhibition (Gominet et al., 2001).

The RNPP-type regulators constitute well-studied examples of this group, which are well represented in the genus *Bacillus* but have also been identified in members of the anaerobic genus *Clostridium* (Kotte et al., 2017; Poehlein et al., 2017; Rocha-Estrada et al., 2010). For instance, while the organism of this study, *C. acetobutylicum* ATCC 824, has only a single agr system, it contains at least 8 RNPP type QS systems (Kotte et al., 2017). The body of work described here focuses on the characterisation of one such system, named QssB, comprised of the regulator QsrB and the downstream encoded conjugate signalling peptide QspB. The QsrB protein contains an N-terminal HTH domain for DNA binding (to an unknown QsrB-Box sequence), where it presumably acts as a transcriptional regulator. Therefore, it differs from the rap-phr systems which have been extensively studied in *B. subtilis* (Gallego del Sol and

Marina, 2013; Pottathil and Lazazzera, 2003). Rap systems are unique within the RNPP family due to their lack of an HTH domain for DNA interaction, and instead functioning as a phosphatase through additional TPR domains, of which it contains 7 (Gallego del Sol and Marina, 2013). The Rap system has been shown in *Bacillus subtilis* to aid in the maintenance of the Spo0A to Spo0A-P ratio, the relative increase in the latter being required for the onset of sporulation (Gallego del Sol and Marina, 2013). This occurs through phosphatase activity of Spo0F-P, blocking the phosphorelay system required for the phosphorylation of Spo0B, which in turn leads to Spo0A phosphorylation (Gallego del Sol and Marina, 2013; Slamti et al., 2014). This is not to say that the acquisition of a HTH domain renders a QS regulator incapable of performing phosphatase activity.

One of the RNPP systems has been shown by recent studies to perform both phosphatase and transcription factor roles, this is the NprR regulator (Perchat et al., 2016b). The NprR regulator is the most structurally similar regulator to Rap, containing the same number of TPR domains, but as previously mentioned it differs by containing a HTH domain (Perchat et al., 2016b). When the conjugate peptide NprX is not bound, NprR is in a dimeric form and the TPR domains capable of performing phosphatase activity are exposed. However, binding of the conjugate NprX peptide results in the formation of the NprR-NprX complex, which consists of NprR in a tetrameric formation (Zouhir et al., 2013). This tetrameric formation has pleiotropic activity, leading to the differential regulation of 41 genes, which the authors of numerous studies refer to as the NprR regulon (Lemy et al., 2012; Perchat et al., 2016b). Interestingly the NprR system does not engage in autoregulation, which some researches wrongly suggest to be a necessary hallmark of a QS system (Perchat et al., 2016b). The regulation of genes was determined by the use of proteomic and transcriptomic data. However use of the terminology ‘regulon’ seems to imply the direct regulation of the 41 genes via the NprR-NprX complex transcriptional activity, failing to take into account the systems indirect effects, for instance the regulation of other transcription factors and phosphorelay systems. This will be considered later within this body of work within Chapter 5.

Perchat and colleagues (2016) suggested that the ancestor to the Rap phosphatase QS system may have been a protein with a binding inhibitor on an elongated flexible C-terminus region, which through evolution was detached to become the exportable conjugate peptide Phr. It was further suggested that the Rap regulator could have acquired a HTH domain, leading to the moonlighting regulator NprR (Perchat et al., 2016b). It is, therefore, not a great leap to consider that the loss of the 3 TPR domains required for phosphatase activity would lead to a regulator without moonlighting activity, simply acting as a transcription factor in the presence of its conjugate peptide, such as the PlcR system found in the *B. cereus* group (Perchat et al., 2016b). The *plcR* gene is well studied due to its role as a master regulator of pathogenicity in *B. cereus* and *B. thuringiensis*, PlcR is therefore a pleotropic transcription factor effecting 45 genes across multiple operons (Agaisse et al., 1999). Its cognate signalling peptide is encoded by *papR*. The PlcR protein exists as a dimer, and on the formation of the PlcR-PapR complex, the HTH domains are able to orientate themselves in order to bind a consensus sequence to allow transcription factor activity (Declerck et al., 2007).

The QsrB regulator studied here was predicted to contain 7 TPR domains, one more than the PlcR regulator but lacking the phosphatase TPR domains of Rap and NprR (Kotte et al., 2017; Perchat et al., 2011). An interesting comparison that will form the basis of this chapter, is that the previously studied PapR peptide, the conjugate peptide of the PlcR regulator, is mainly processed as a heptapeptide (Bouillaut et al., 2008). This result is not in isolation as NprX, the conjugate peptide for the NprR system, has been shown to have a minimum active length of 7 amino acids also, although currently data cannot conclude if this is its native form (Perchat et al., 2011).

The system that forms the basis of this study, is QssB within *C. acetobutylicum* ATCC 824 (Kotte, 2013). Using two previously discovered homologs of RNPP-type regulators in *C. acetobutylicum* ATCC 824 (Pottathil and Lazazzera, 2003), and the PlcR and NprR confirmed RNPP-type regulators within *B. subtilis* and the *B. cereus* group, blastp searchers were used to identify a 11 putative RNPP-type regulators (Kotte, 2013; Kotte et al., 2017). The flanking regions of these regulators were

analysed for small open reading frames (ORFs), indicating the presence of putative conjugate signalling peptides precursors (Kotte, 2013). It has been previously shown that these precursors contain a positively charged N-terminus followed by a hydrophobic region, which is required for peptide export (Pottathil and Lazazzera, 2003), with the C-terminus containing the signalling peptide (Kotte, 2013). Of the afore mentioned putative RNPP-type regulators, 8 were found to have a flanking short ORF (Kotte, 2013). The regulators were therefore designated the names Quorum Sensing Regulator A through to H (*qsrA* – *qsrH*), with the short ORFs being named Quorum Sensing Peptide A through to H (*qspA* – *qspH*). Kotte (2013) had shown that insertional mutation, via the ClosTron method, of *qsrB* caused an increase in solvent production at the early time points of *C. acetobutylicum* ATCC 824 growth. It was also shown that ClosTron insertion of *qspB* had led to some reduction of solvents at later time points. However, overexpression of *qsrB* lead to a dramatic reduction of solvents, and dramatic reduction in the formation of spores (Kotte, 2013). One would expect a disruption of conjugate peptide production to present a similar phenotype to the overexpression of the regulator. As stated this did not appear to be the case with regards to QssB, which could suggest the utilisation of an alternative signalling peptide, which will be investigated within this chapter. As an alternative theory, it could also suggest that the initial ClosTron insertion of *qspB* was ultimately unsuccessful in disrupting its production, which would prompt further work as discussed under 3.4. The phenotype was shown to be relieved by the addition of synthetic peptides matching the C-terminus of the QspB sequence, which was used as evidence that QsrB and QspB form a cell-cell communication system which represses solvents and spore production (Kotte, 2013). The original observations were limited to the ClosTron technique. In addition to this the synthetic peptides used to relieve *qsrB* overexpression were not inclusive enough to determine the minimum peptide sequence required for QsrB binding (Kotte, 2013).

3.2 Aim of the study

The overall aim of this study was to further define QssB-mediated signalling in *Clostridium acetobutylicum*, the foundation of which was performed by Kotte (2013).

Previous mutational studies were limited to the use of insertional ClosTron mutants and the use of a few synthetic QspB peptides variants to alleviate the effects of *qsrB* overexpression. This work was to be substantially extended by pursuing the following, specific aims;

- The creation of a clean in-frame deletion of the *qsrB* gene (CA_C0324), for the further characterization of phenotype, and to further support or reject previous ClosTron-based findings.
- To elucidate the minimum conjugate QspB peptide length required for alleviation of QsrB-mediated repression.
- To achieve a greater understanding of the conserved requirements of the conjugate QspB peptide by performing an alanine scan using the complementation phenotype alleviation model.
- To elucidate the specificity of the QsrB regulator, by addition of conjugate signalling peptides from other RNPP type systems present within *Clostridium acetobutylicum* ATCC 824 to the *qsrB*-overexpression strain.
- To analyse the effect of QspB addition to the wild type *C. acetobutylicum* ATCC 824 strain, in order to investigate the ability for the synthetic conjugate peptide of QsrB to allow greater production of solvents during fermentation.
- To investigate the ability for *C. acetobutylicum* ATCC 824 to engage in cross talk with the closely related *C. roseum*, that contains a homologous RNPP type system to QsrB.

3.3 Results

3.3.1 Creation of a clean in-frame deletion of the *qsrB* gene

For the generation of a clean *qsrB* in-frame deletion the method described by Ehsaan *et al.* (2016) was followed. A deletion cassette was created that eliminated the entire structural gene apart from the first 3 and 3 last codons, as follows. Two pairs of primers, F1_QsrB_KO_F, F1_QsrB_KO_R and F2_QsrB_KO_F, F2_QsrB_KO_R (Table 2.3), were used to generate two 750-bp PCR fragments encoding the *qsrB* upstream and downstream region, respectively; these fragments were subsequently joined within the pMTL-ME3 vector via Gibson assembly (Gibson *et al.*, 2009). The

plasmid pairs were constructed using the NEB builder browser tool (<https://nebbuilder.neb.com/>) and the pMTL-ME3 vector was linearized using a NotI restriction digest prior to assembly. Gibson assembly protocol was followed as describe in the material and methods section (2.3.5.5). The final completed cassette thus comprised a total of 1500 bp of upstream and downstream region, yet included only the first and last 3 codons of the QsrB structural gene. Construction of the plasmid was verified via restriction digest and sequencing. The plasmid was appropriately methylated in strain *E. coli* TOP10 (pan2) and transformed to the *C. acetobutylicum* ATCC 824 Δ *pyrE* strain as previously described (2.3.4.2 and 2.3.5.7, Heap et al. 2009), selecting for Thiamphenicol resistant colonies. Thiamphenicol resistant colonies were obtained and restreaked to purity. Following this, the Thiamphenicol resistant colonies were restreaked on CBM-S media, to select for colonies bearing the plasmid borne *pyrE* gene as described within section 2.5.1 of the Materials and Methods. Single cross-over mutations were not obtained.

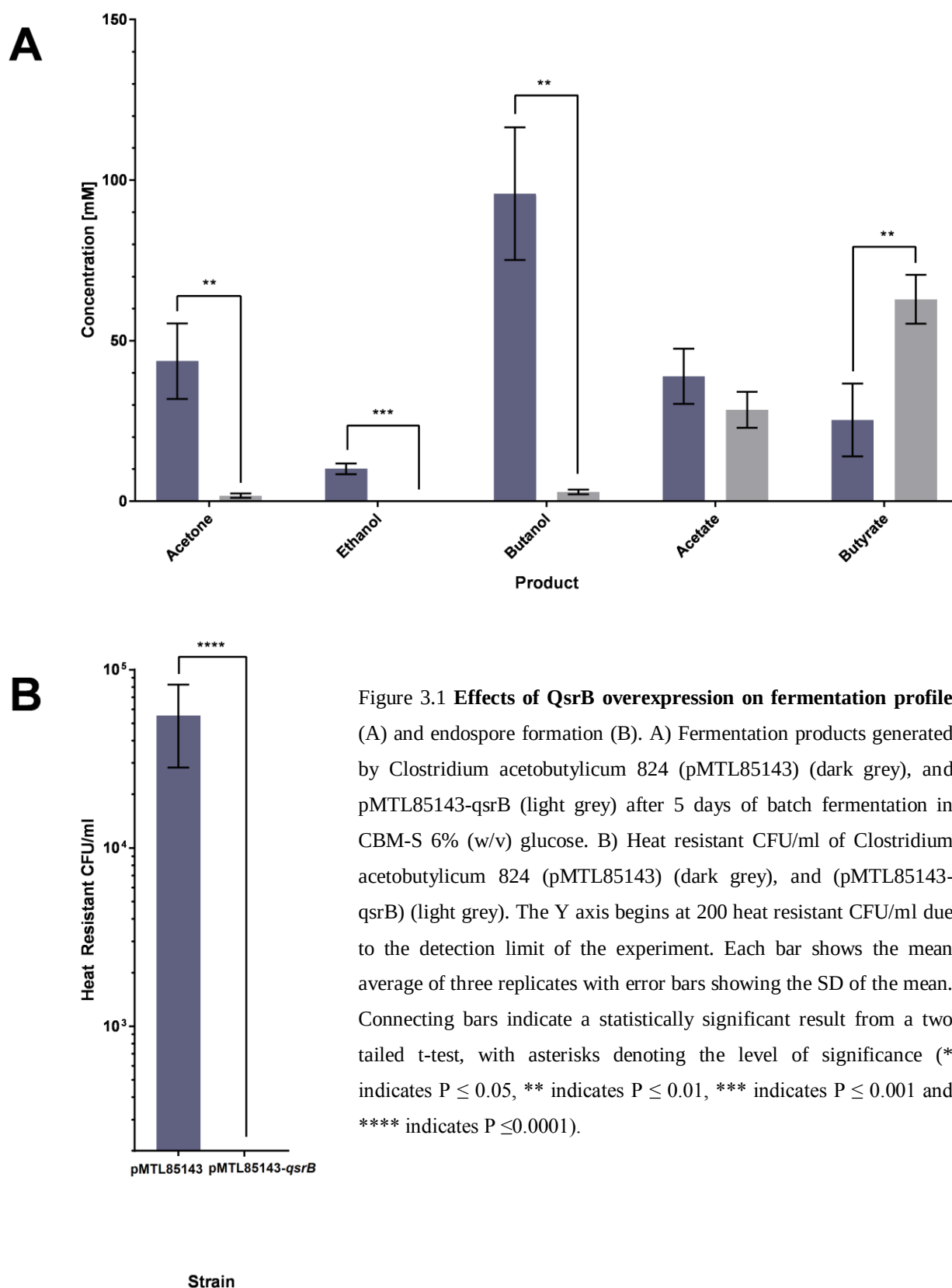
Several transformants containing the vector were obtained and underwent further isolation by restreaking on CBM-S agar to enhance the growth of single crossovers. Small colony growth, reminiscent of the Δ *pyrE* phenotype, were seen. Single crossovers were not obtained. Following numerous unsuccessful attempts to obtain an inframe deletion of *qsrB*, the cassette was redesigned to only remove the RBS, start codon, and HTH domain of the CA_C0324 locus. This plasmid was also constructed using Gibson assembly (Gibson et al., 2009). The two homology arm fragments were constructed using the following plasmid pairs: F1_QsrB_HTH_RBS_KO_F and F1_QsrB_HTH_RBS_KO_R, F2_QsrB_HTH_RBS_KO_F and F2_QsrB_HTH_RBS_KO_R (Table 2.3). As with the former plasmid this was appropriately methylated and transformed to *C. acetobutylicum* ATCC 824 Δ *pyrE*. Again, numerous attempts were made to obtain an inframe deletion of the RBS, start codon and HTH domain of *qsrB*, but were all unsuccessful (data not shown). It was therefore decided to no longer pursue this line of investigation.

3.3.2 Elucidating the minimum QspB peptide length required for complementation of the pMTL85143-*qsrB* phenotype

When the original QsrB overexpressing strain generated by Kotte (2013) was reactivated from culture collection, it was found to be divergent from the recorded results within the body of work, as it did not have a reduced solvent profile. The pMTL85143-*qsrB* overexpression plasmid, which contains the strong constitutive ferredoxin promoter from *C. sporogenes*, was therefore recreated as described within using the primer pair QsrB_C_F1 and QsrB_C_R1. This *fdx* promoter is widely published and used within the *Clostridium* spp. (Heap et al., 2009). The required region was PCR amplified (2.3.5.2), following this the band was restriction digested with *NdeI* and *BamHI* alongside the pMTL85143 vector, and the two subsequently ligated (2.2.5.3). Following this the recreated pMTL85143-*qsrB* plasmid was transformed to *C. acetobutylicum* ATCC 824 to create the overexpression strain (2.3.4.2). Overexpression of the *qsrB* regulator lead to a dramatic reduction in solvent production, shown at the 120h time point in A, compared to a control strain containing the same vector without insert. Combined with this, overexpression of QsrB led to a vast reduction in spore formation as shown in B. Often, spore counts for the *qsrB* overexpression strain were below the detection limits of the experiment (>200 heat resistant CFU/ml). However, spore levels were occasionally seen to reach 1×10^2 to 1×10^3 heat resistant CFU/ml (not shown), making it between a 100 and 1000 fold decrease from the levels of the pMTL85143 vector control. As previously stated; under the batch culture conditions of this study (2.2.1) the exponential phase of the WT organism would continue for 12-14 hours to an OD 600 of 4, upon which the culture enters stationary phase. Overexpression of *qsrB* did not incur a growth defect on the organism.

It has been previously shown by our group that certain synthetically derived QspB variants, when added to cultures of the *qsrB* overexpression strain, could restore butanol production (Kotte, 2013). The previous study only reported this restoration in butanol formation. In addition to this, the previous study lacked the necessary resolution to deduce the minimum peptide length required for QsrB recognition and used peptides of comparatively low purity (Kotte, 2013). A number of different length

QspB fragments were therefore designed and obtained from (PeptideSynthetics, UK) at above 90% purity (see 2.4.2 and table 2.4 in Methods for details).



The full length amino acid sequence of QspB is shown in Figure 3.2, together with the different variations of the QspB synthetic peptides used for complementation comparisons. The peptides were designed to encompass the heptapeptide QspB 7, which in a previous study had been shown to be active in the restoration of butanol production when added to the *qsrB* overexpression strain (Kotte, 2013), in addition to shorter hexapeptides and pentapeptides which acted as additional controls.

It was decided, due to the increase in peptide purity, that a brief pilot study should be

MLPKKSNKFILVLILCAIVFVSAFALNTSGTRSLLGAEPTWGWNI SKLLF	
QspB 9-3	LGAEPT
QspB 9-2	LGAEPTW
QspB 9-1	LGAEPTWG
QspB 9	LGAEPTWGW
QspB 10	LLGAEPTWGW
QspB 8	GAEPTWGW
QspB 7	AEPTWGW
QspB 6	EPTWGW
QspB 5	PTWGW

Figure 3.2 Sequence of QspB and derived synthetic peptides. The amino acid sequence across the top line shows the entire QspB peptide, with the vertical blue lines indicating the maximal synthetic peptide length used for the decapeptide QspB 10. Sequence and position of the other peptides used in this study are shown for comparison.

conducted assessing the complementation effect of different concentrations of QspB 9, a peptide previously shown to restore butanol formation when added to the overexpression strain (Kotte, 2013). This was to establish the approximate concentration ranges to be used in subsequent experiments and performed as stated in 2.6.1 of the Materials and Methods section with the peptides dissolved in dimethyl sulfoxide (DMSO).

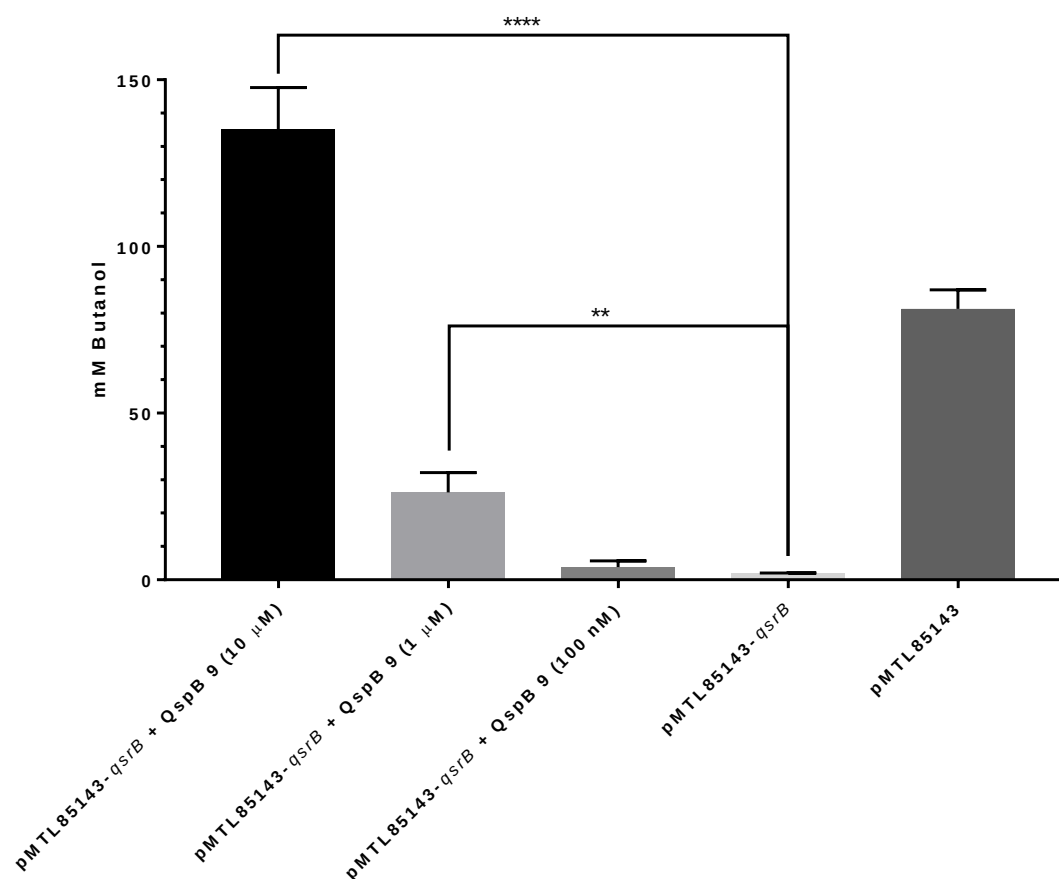


Figure 3.3 Butanol production of *C. acetobutylicum* ATCC 824 (pMTL85143-*qsrB*) after 120 hours of batch culture fermentation, with addition of 10 μ M, 1 μ M and 100 nM of QspB 9, respectively. *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* (no QspB 9 added) and *C. acetobutylicum* ATCC 824 pMTL85143 acted as controls and were treated with equivalent volumes of DMSO. Each bar shows the mean average of 3 replicates with error bars showing the SD of the mean. Connecting bars show a significantly significant result of a two tailed t-test, with asterisks indicating the level of significance (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

The effect of variations in QspB 9 concentration on butanol production by *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* is shown in Figure 3.3. A vector control strain, containing pMTL85143 and supplemented with only DMSO, produced a mean average of 81 ± 2.7 mM of butanol after 120 hours of batch culture fermentation. By comparison, the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strain, supplemented only with DMSO, produced 2 ± 0.04 mM of butanol over the same time period. 10 μ M of QspB 9 was effective at completely restoring butanol production in the *qsrB* overexpression background, showing 135 ± 5.9 mM of butanol produced at 120 hours of batch fermentation. 1 μ M of QspB 9 addition lead to a reduced complementation phenotype with 26 ± 2.8 mM of butanol produced. However, the complementation

using 1 μ M was statistically different from the pMTL85143-*qsrB* DMSO addition (p -value = 0.0022). Addition of 100 nM QspB 9 lead to no significant change in butanol production relative to the overexpression control, producing 4 ± 0.9 mM of butanol. It was decided that 10mM of peptide addition would be sufficient for the complementation analysis of each of the synthetic peptides. Following identification of active peptide variants, a range of concentrations would then be tested to evaluate the dose response.

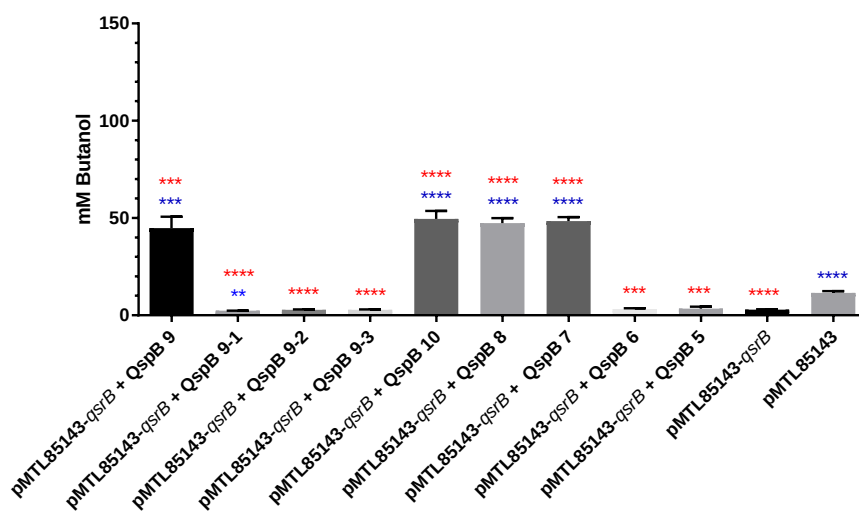
The various synthetic peptide variations shown in Figure 3.2, all synthesised to a purity of $\geq 90\%$, were used to supplement the pMTL85143-*qsrB* strain. Three different time points were considered for this experiment, 24 hours, 48 hours and 120 hours, respectively, with endospore formation only determined for the latter. After 24 hours of fermentation, *C. acetobutylicum* ATCC 824 containing the pMTL85143 plasmid control had produced 12 ± 0.4 mM of butanol compared with the 3 ± 0.1 mM produced by the pMTL85143-*qsrB* strain (p -value = 0.0001). It can be seen that QspB 5, QspB 6, QspB 9-2 and QspB 9-3 all failed to act as a conjugate peptide for the *qsrB* overexpression strain, as butanol levels produced in their presence did not differ from that of the DMSO control. By contrast, QspB 10, QspB 9, QspB 8 and QspB 7 showed butanol production to a level that exceeded that of the empty vector control. Addition of 10 μ M QspB 9 resulted in 45 ± 2.8 mM of butanol (p -value = 0.0007). Supplementation with 10 μ M QspB 10, QspB 8, and QspB 7 showed butanol levels of between 50 ± 1.9 mM and 47 ± 1.2 mM, respectively (All p -values = <0.0001).

This difference increased and remained statistically significant after 48 hours of fermentation, when the vector control had produced 34 ± 1.4 mM of butanol compared to 3 ± 0.1 mM still present within the pMTL85143-*qsrB* strain. Supplementation with 10 μ M QspB 10, QspB 9, QspB 8 and QspB 7 produced between 60 ± 2.1 mM (QspB 8) and 67 ± 2.9 mM (QspB 9) of butanol after 48 hours fermentation.

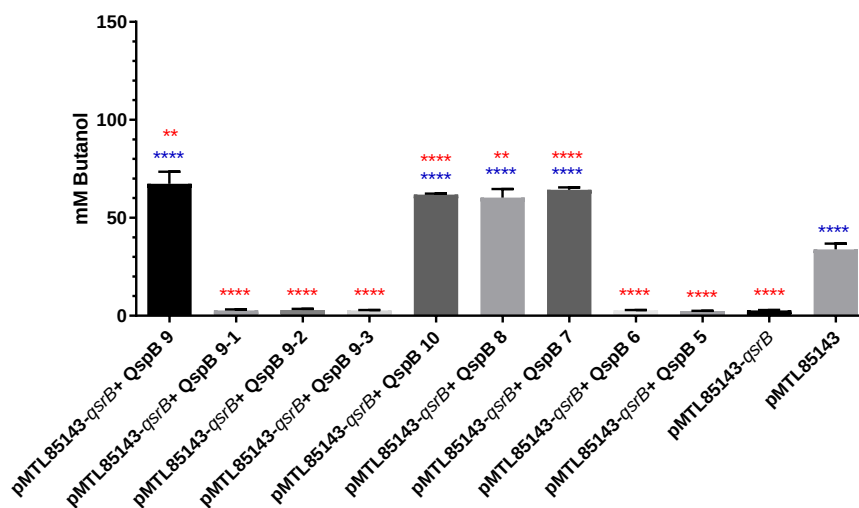
The experiment was run for another 3 days, with an end point sample and a spore sample taken at this point. As shown in section C of Figure 3.4, butanol production

from the *C. acetobutylicum* ATCC 824 pMTL85143 vector control with DMSO was 81 ± 2.7 mM, with 2 ± 0.04 mM being produced by the pMTL85143-*qsrB* DMSO peptide-negative control. Complementation of the pMTL85143-*qsrB* phenotype with the nonapeptide QspB 9 showed the highest production of butanol with 135 ± 5.9 mM (p -value = 0.0024 compared to the empty vector control).

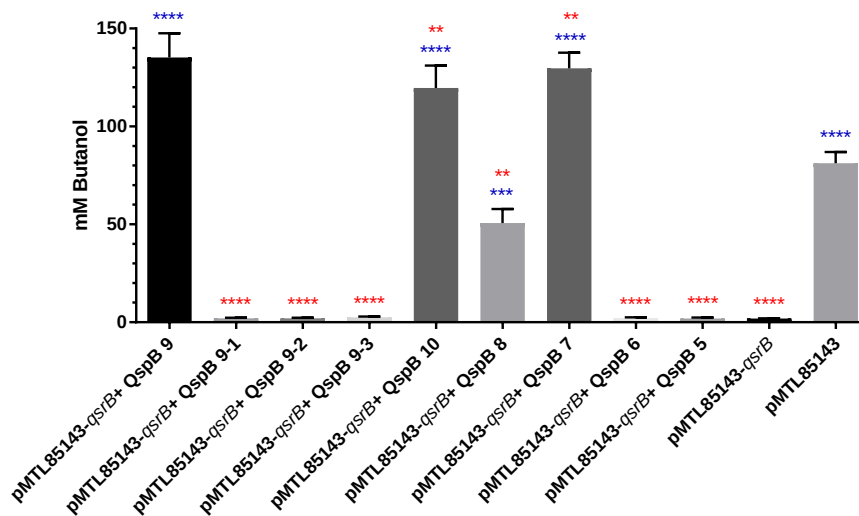
A)



B)



C)



D)

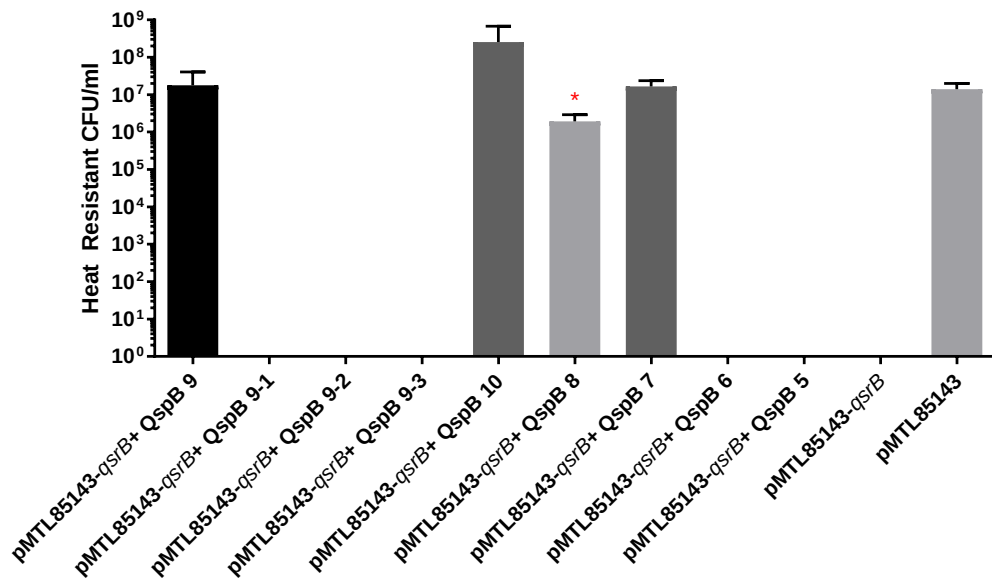


Figure 3.4 Butanol production and endospore formation by *C. acetobutylicum* 824 (pMTL85143-*qsrB*) in the presence of synthetic QspB peptides. Butanol formation by *C. acetobutylicum* ATCC 824 (pMTL85143-*qsrB*) grown on CBM-S medium was assessed after 24 hours (A), 48 hours (B) and 120 hours (C), respectively, in the presence of 10 μ M of the indicated peptides or a DMSO-only control. Endospore formation was measured after 120 hours (D) as heat resistant CFU/ml. *C. acetobutylicum* ATCC 824 (pMTL85143) was included as butanol producing control and supplemented with peptide-free DMSO. Each bar shows the mean average across 3 replicates, with error bars representing the SD of the mean. Two tailed t-tests were performed between each experimental set and the pMTL85143-*qsrB* control (Blue) and the pMTL85143 (Red) control, with asterisks indicating the level of significance (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

Supplementation with 10 μ M QspB 10 and QspB 7 yielded 120 ± 5.4 mM and 130 ± 3.8 mM of butanol respectively, and thus are in the same range as QspB 9 supplementation.

The pH of the cultures was not analysed during the course of this experiment. However, no visual cues of acid crash were observed (Maddox et al., 2000; Wang et al., 2011). In addition, the pMTL85143-*qsrB* containing strain supplemented with DMSO and also those supplemented with biologically inactive conjugate peptides did not seem to dissipate calcium carbonate within the growth media any more than the solvent producing strains.

The ability of the conjugate peptides to complement the spore forming capability of the pMTL85143-*qsrB* phenotype was analysed following 5 days of fermentation. As seen in Figure 3.4D, the sporulation positive *C. acetobutylicum* ATCC 824 pMTL85143 control formed $1.04 \times 10^7 \pm 2.83 \times 10^6$ heat resistant CFU/ml, whereas the pMTL85143-*qsrB* strain failed to produce detectable levels (≥ 200 heat resistant CFU/ml). Furthermore, no detectable levels of spores were seen when the pMTL85143-*qsrB* harbouring strain was supplemented with 10 μ M QspB 9-1, QspB 9-2, QspB 9-3, QspB 6 or QspB 5, respectively. However, when supplemented with QspB 10, QspB 9 and QspB 7 spore levels were shown to reach normal levels, and did not differ significantly from the pMTL85143 DMSO control.

The results shown in Figure 3.4 allowed insight into the minimum peptide length and sequence required for biological activity with QsrB in *C. acetobutylicum* ATCC 824. A summary of active and non-active QspB peptides is shown in Figure 3.5. QspB 7 was the smallest peptide that showed activity. Removal of its N-terminal alanine resulted in no activity, as well as the C-terminal tryptophan.

QspB9-3: LGAEPT
 QspB9-2: LGAEPTW
 QspB9-1: LGAEPTWG
 QspB9: LGAEPTWGW
 QspB10: LLGAEPTWGW
 QspB8: GAEPTWGW
 QspB7: AEPTWGW
 QspB6: EPTWGW
 QspB5: PTWGW

Figure 3.5 Biologically active (green) and inactive (red) QspB peptide variants. The amino acid sequence of each of the QspB variants is shown, with active peptides being able to restore butanol production in *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* cultures.

In order to confirm the results obtained for the synthetic peptides, they were synthesised to a purity of $\geq 95\%$ purity and tested as before. The results for a complete 5 day fermentation are shown in Figure 3.6. As can be seen by Figure 3.6A, the synthetic peptides QspB 5, QspB 6, QspB 9-1, QspB 9-2 and QspB 9-3 again showed no biological activity, and this was the case for every subsequent repeat (data not shown).

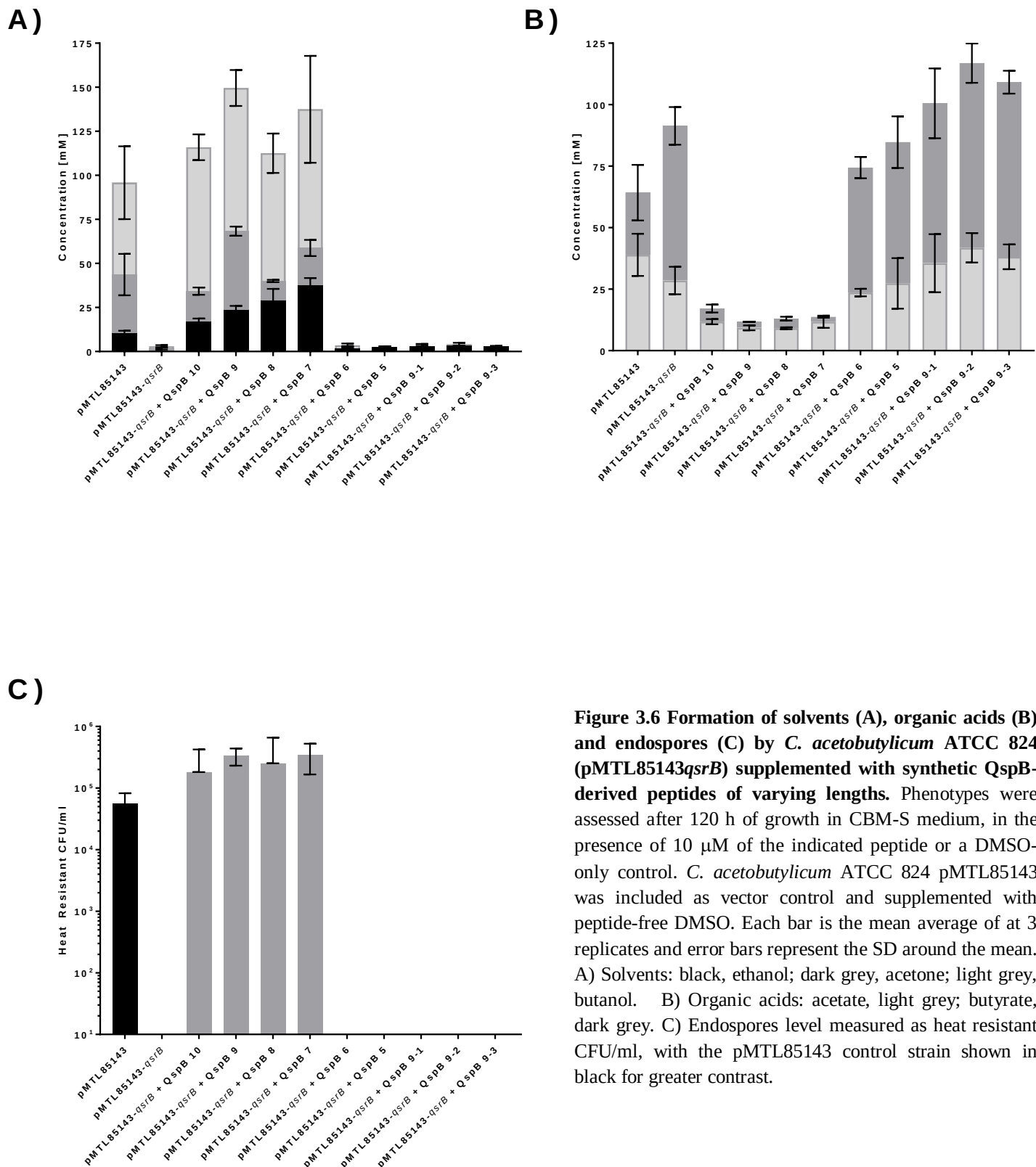


Figure 3.6 Formation of solvents (A), organic acids (B) and endospores (C) by *C. acetobutylicum* ATCC 824 (pMTL85143*qsrB*) supplemented with synthetic QspB-derived peptides of varying lengths. Phenotypes were assessed after 120 h of growth in CBM-S medium, in the presence of 10 μ M of the indicated peptide or a DMSO-only control. *C. acetobutylicum* ATCC 824 pMTL85143 was included as vector control and supplemented with peptide-free DMSO. Each bar is the mean average of at 3 replicates and error bars represent the SD around the mean. A) Solvents: black, ethanol; dark grey, acetone; light grey, butanol. B) Organic acids: acetate, light grey; butyrate, dark grey. C) Endospores level measured as heat resistant CFU/ml, with the pMTL85143 control strain shown in black for greater contrast.

Upon further analysis, all biologically active peptides produced significantly increased ethanol when compared to the pMTL85143 control. The empty vector control produced 10 ± 0.8 mM of ethanol following 5 days of batch fermentation. Supplementation of the pMTL85143-*qsrB* strain with QspB 10, QspB 9, QspB 8 and QspB 7 produced 17 ± 0.7 mM, 24 ± 1.05 mM, 29 ± 3.1 mM and 37 ± 2.0 mM, respectively (*p*-values: QspB 10, 0.006; QspB 9, 0.001; QspB 8, 0.009; QspB 7, 0.0005). The only addition seen to outperform the empty vector control across all solvents in a statistically significant manner was pMTL85143-*qsrB* supplemented with QspB 9. The vector control produced acetone and butanol levels of 44 ± 5.5 mM and 96 ± 9.75 mM respectively, whereas the QsrB overexpression complemented with 10 μ M QspB 9 produced acetone and butanol levels of 68 ± 1.2 mM and 150 ± 4.8 mM respectively (respective *p*-values = 0.024, 0.016).

As can be seen in Figure 3.6B, final levels of acetate and butyrate in QspB 10, QspB 9, QspB 8 and QspB 7 supplemented cultures of the pMTL85143-*qsrB* strain was lower in all respects than the pMTL85143 control, perhaps indicating greater re-assimilation of the acids as opposed to *de novo* generation of solvents. For comparison, the pMTL85143 vector control strain with DMSO addition produced 40 ± 4.04 mM of acetate and 25 ± 5.3 mM of butyrate, whereas the pMTL85143-*qsrB* strain with DMSO addition produced 29 ± 2.6 mM acetate and 63 ± 3.6 mM butyrate.

Figure 3.6C shows the levels of heat resistant CFU/ml, analysed after 5 days fermentation under the before mentioned conditions. The pMTL85143 DMSO addition control produced $5.5 \times 10^4 \pm 7.5 \times 10^3$ heat resistant CFU/ml, whereas the pMTL85143-*qsrB* strain supplemented with only DMSO addition produced no detectable heat resistant CFU/ml (detection limit ≥ 200 heat resistant CFU/ml). The 10 μ M addition of QspB 5, QspB 6, QspB 9-1, QspB 9-2 and QspB 9-3 failed to restore spore formation in the QsrB overexpression phenotype. As previously observed, 10 μ M of QspB 10, QspB 9, QspB 8 and QspB 7, were able to restore the sporulation phenotype of the pMTL85143-*qsrB* harbouring strain. In respect to supplementation with QspB 10 and QspB 8 no significant difference was seen between these experimental replicates and the pMTL85143 with DMSO addition control. However,

addition of QspB 9 and QspB 7 produced significantly more heat resistant CFU/ml (p -value = <0.0001 in both cases).

3.3.3 Investigating the dose response of QspB 7 as the conjugate peptide of QsrB

As the minimum peptide length required for conjugate activity with QsrB was ascertained, a dose response profile was generated to investigate its potency, using butanol production as a phenotypic marker.

Figure 3.7 shows a range of concentrations of QspB 7, from 0.1 μ M to 40 μ M, used to induce butanol production in the pMTL85143-*qsrB* strain. Butanol concentrations were determined after 24 hours of batch fermentation, and again 48 hours and 5 days and compared to those obtained for the controls, strains pMTL85143 and pMTL85143-*qsrB* to which only DMSO had been added.

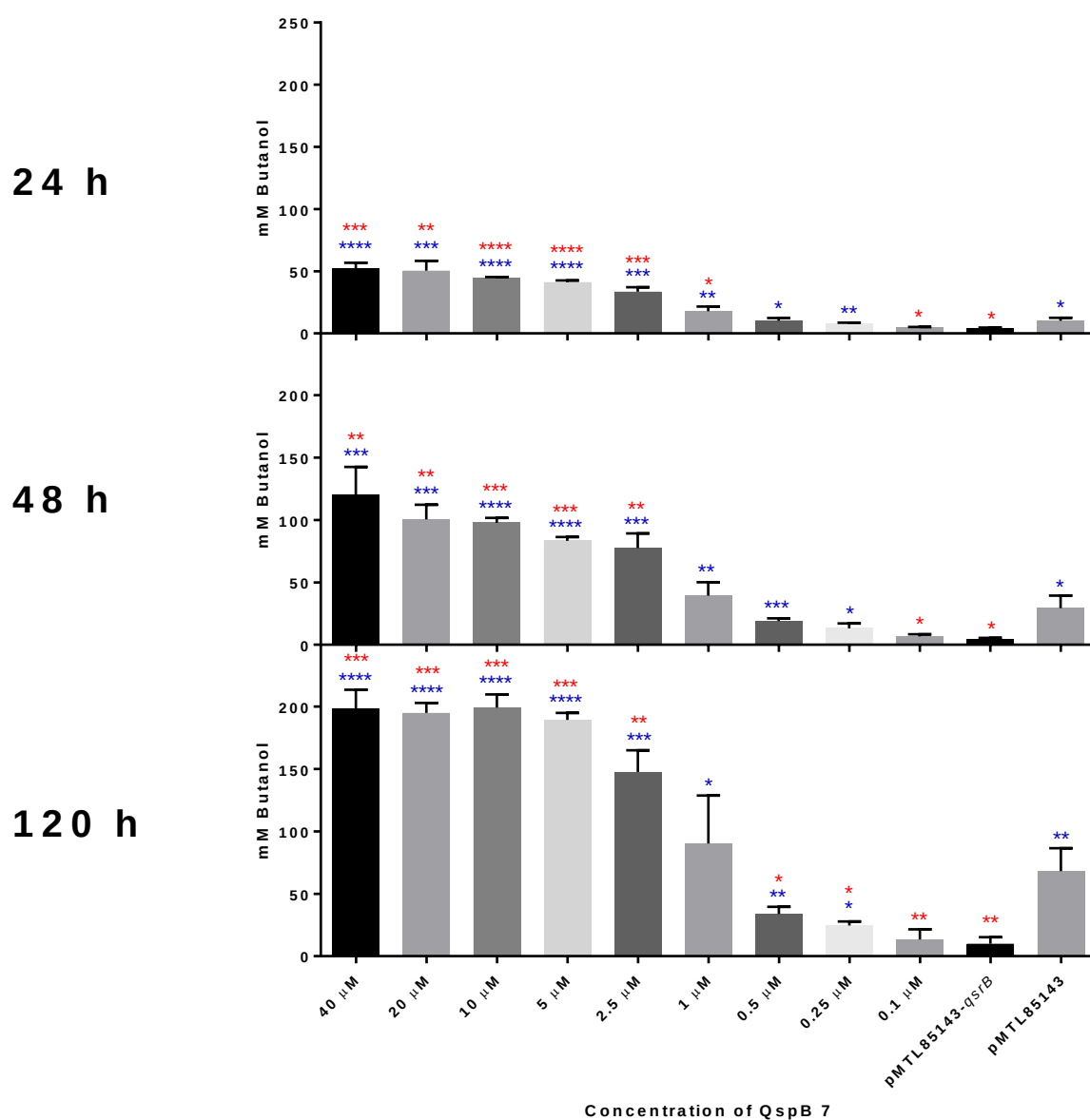


Figure 3.7 Dose response relationship for the Qsp7 peptide. The production of butanol by *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* grown on CBM-S medium was assessed after 24 hours (top) 48 hours (middle) and 5 days (bottom) with the indicated concentrations of QspB 7. This was performed alongside a *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* control supplemented with DMSO only, and a *C. acetobutylicum* ATCC 824 pMTL85143 DMSO control. Each column represents the mean average of 3 replicates, with error bars representing the SD of the mean. Two tailed t-test analysis was performed, showing the statistical difference between the experimental value and pMTL85143-*qsrB* DMSO (blue) and pMTL85143 DMSO (red), the asterisks dictate the level of significance (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

Interestingly, for the 5 day samples, supplementation with 40 μ M, 20 μ M, 10 μ M, 5 μ M and 2.5 μ M of QspB 7 led to dramatically increased butanol levels. These were considerably higher than the 68 ± 8.7 mM obtained for the pMTL85143 DMSO control and statistically significant with p -values ranging from 0.0004 to 0.006. Thus, at the indicated concentrations, addition of QspB 7 to the pMTL85143-*qsrB* strain yielded butanol concentrations well above those normally observed for the empty vector control. Supplementation with 1 μ M QspB 7 resulted in a butanol concentration of 90 ± 18.1 mM which did not differ significantly from the pMTL85143 with DMSO addition (p -value = 0.4). However, it was significantly higher than the 10 ± 2.4 mM produced by the pMTL85143-*qsrB* control showing that alleviation of repression had occurred (p -value = 0.0231). Attempting to overcome QsrB-mediated repression with 0.5 μ M, 0.25 μ M and 0.1 μ M QspB 7 yielded 34 ± 2.8 mM, 25 ± 1.5 mM and 13 ± 3.8 mM of butanol, respectively. For the former two concentrations, this was significantly more than observed for the pMTL85143-*qsrB* DMSO control (p -values of 0.006 and 0.01, respectively). Thus the only QspB 7 concentration that failed to show a significant effect after 5 days was 0.1 μ M.

Similar trends were observed for the earlier time points, which after 24 hours yielded butanol concentrations of 10 ± 1.1 mM and 4 ± 0.1 mM for the pMTL85143 and pMTL85143-*qsrB* DMSO control strains, respectively. These concentrations rose to 29 ± 4.7 mM and 5 ± 0.4 mM after 48 hours.

All supplementations, apart from 0.1 μ M, significantly increased butanol production by the pMTL85143-*qsrB* strain in comparison to the DMSO only control after 24 hours (p -values between 0.0000002 and 0.01). Furthermore, the concentrations achieved with QspB 7 supplementation of 1 μ M and above were significantly higher than those observed for the pMTL85143 DMSO control (p -values between 0.04 and 0.00002). As before, after 48 hours, all butanol concentrations observed for 0.25 μ M QspB 7 and above differed significantly from the pMTL85143-*qsrB* DMSO only control (p -values ranging from 0.000002 to 0.02) and butanol concentrations observed for 2.5 μ M QspB 7 and above were significantly higher than those of the pMTL85143 DMSO only control (p -values ranging from 0.0004 to 0.006).

Based on the above experiment, it was decided that 10 μ M of QspB 7 peptide provided sufficient levels to completely alleviate QsrB mediated repression. This concentration was therefore used in subsequent experiments.

3.3.4 Investigating the ability for QssB of *C. acetobutylicum* ATCC 824 to engage in interspecies cross talk, using the synthetic QspB 9 peptide from *C. roseum*

During the course of this study data was released showing the genome of another industrially relevant species of *Clostridium*, *Clostridium roseum* (Poehlein et al., 2017). *C. roseum* is a saccharolytic, solventogenic *clostridia*, and has many parallels to the closely related *C. acetobutylicum* ATCC 824 (Poehlein et al., 2017). Akin to *C. acetobutylicum* ATCC 824 its ecological niche is poorly understood, however, it differs visibly by producing a pink pigmentation that became its namesake (Lund et al., 1981). Interestingly, the pigment darkens to black upon exposure to oxygen (Lund et al., 1981). The organism contained a homologous system to the QssB studied in *C. acetobutylicum* ATCC 824. Due to the recalcitrant nature of the *C. roseum* organism, genetic manipulation proved to be cumbersome, and beyond the scope of this project. However, personal correspondence, as the manuscript was not yet published, lead to the realisation that the homologous QspB in *C. roseum* had a conserved 9 amino acid stretch that differed in only 3 amino acid residues to the QspB 9 studied in *C. acetobutylicum* ATCC 824 (personal correspondence with Klaus Winzer and Anja Poehlein) Figure 3.8 highlights the comparison between the QspB 9 variants from *C. acetobutylicum* ATCC 824 and *C. roseum*.

C. acetobutylicum QspB 9

LGAEPTWGW

C. roseum QspB 9

LA^AAEPT^RPW

Figure 3.8 Comparison of the QspB 9 amino acid sequence from *C. acetobutylicum* ATCC 824 (Accession number: AE001437) and the homologous sequence from *C. roseum* ATCC 17797 (Accession number: LZYV00000000). Alternative residues in *C. roseum* QspB 9 are highlighted in blue.

The *C. roseum* QspB 9 peptide was synthesised in order to investigate the ability of *C. acetobutylicum* ATCC 824 to engage in interspecies communication with *C. roseum* via the QssB system. The experimental approach was to use 10 μ M of QspB 9 from *C. roseum* in an attempt to alleviate QsrB-mediated repression in the pMTL85143-*qsrB* strain.

As can be seen from Figure 3.9, the QspB 9 synthetic peptide produced from the *C. roseum* sequence failed to restore butanol formation in the pMTL85143-*qsrB* strain. After 120 hours of fermentation the pMTL85143 control strain with DMSO addition had produced 85 ± 1.0 mM of butanol, with the pMTL86143-*qsrB* with DMSO producing only 6 ± 0.03 mM of butanol. As previously shown, and in this case run as a control, the QspB 9 synthetic peptide when added at a 10 μ M concentration was able to restore butanol production when added to the latter strain, which at 120 hours of fermentation had produced 98 ± 2.0 mM of butanol. Addition of QspB 9 from *C. roseum* led to the production of 5 ± 0.7 mM of butanol, which was statistically indistinguishable from the pMTL85143-*qsrB* DMSO control strain. The *C. roseum* QspB 9 was unable to restore the production of the other solvents also (data not shown).

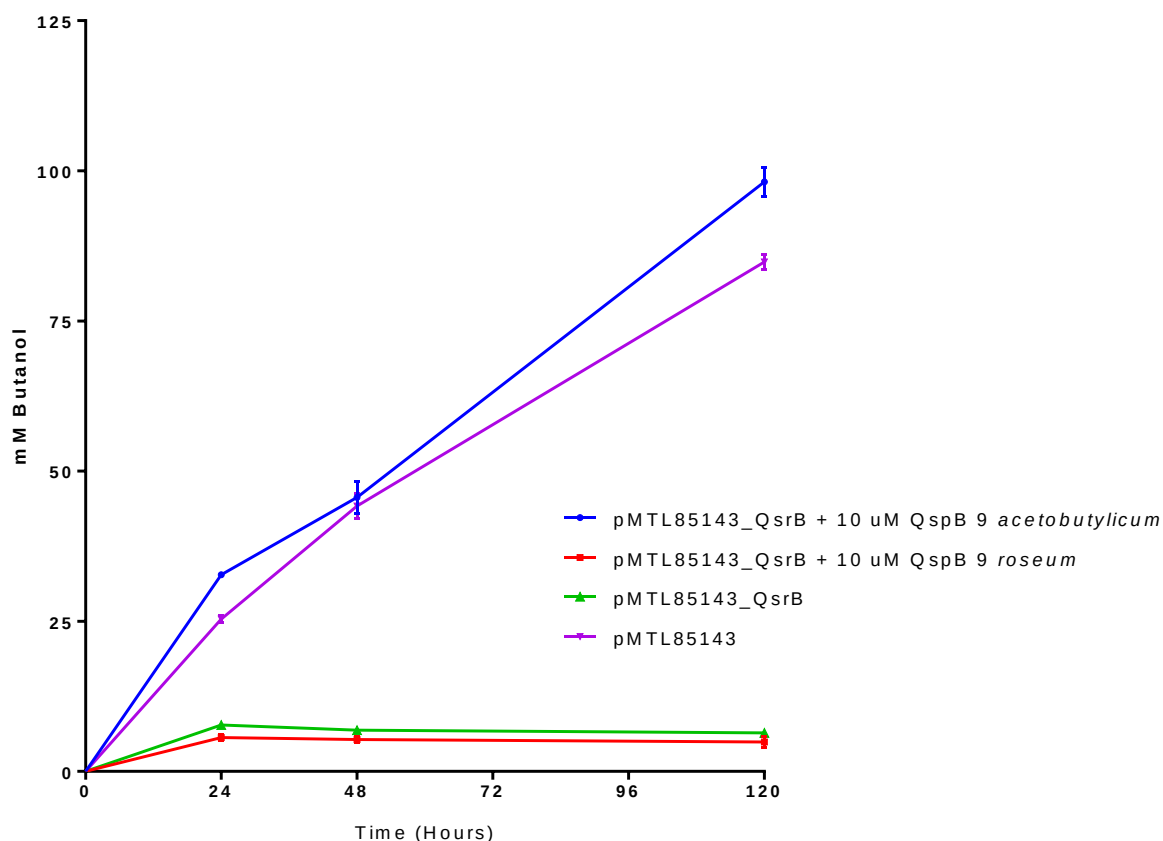


Figure 3.9 Butanol production of *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* in the presence of Qsp9 peptides from *C. roseum* and *C. acetobutylicum* ATCC 824. *C. acetobutylicum* ATCC 824 strains were grown in CBM-S medium supplemented with the indicated peptides or DMSO only. Samples were taken for analytic analysis at 24 hours, 48 hours and 5 days of fermentation. *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* + DMSO (green); *C. acetobutylicum* ATCC 824 pMTL85143 + *C. acetobutylicum* ATCC 824 QspB 9 (blue); *C. acetobutylicum* ATCC 824 pMTL85143 + *C. roseum* QspB 9 (red); alongside a *C. acetobutylicum* ATCC 824 pMTL85143 + DMSO control (purple). Each time point represents the mean of 3 replicates, and error bars show the SD of the mean.

3.3.5 Identification of QspB 7 amino acid positions required for biological activity.

In order to fully discern the contribution of each amino acid to biological activity in the minimal QspB 7 heptapeptide, a number of variants were synthesised at a purity of $\geq 95\%$, replacing the amino acid in each position with an alanine. These are detailed in the Table 3.1, with further details available in the Materials and Methods section (Table 2.4 and 2.4.2).

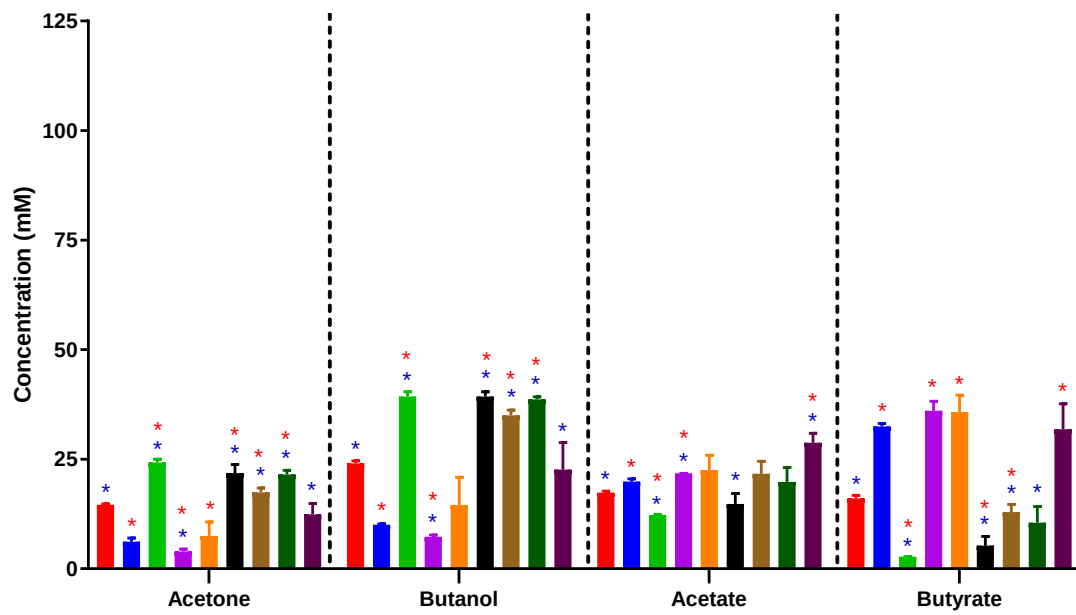
The peptides were used to supplement cultures of the pMTL85143-*qsrB* strain, as described before, alongside appropriate DMSO-only controls. Fermentation profiles are shown in Figure 3.10 for samples obtained after 24 hours, 48 hours and 5 days.

The final fermentation profiles obtained after 5 days reveal a clear pattern, suggesting that all amino acids were required for full biological activity, but some more than others. In terms of acetone and butanol production, substituting the native amino acid in position 3 of the QspB 7 peptide with alanine completely abolished biological activity, as acetone and butanol levels were not significantly different from the pMTL85143-*qsrB* DMSO control. The later produced 3 ± 0.3 mM and 8 ± 0.4 mM of acetone and butanol, respectively, compared to 3 ± 0.9 mM and 12 ± 0.9 mM (p -values of 0.9, 0.4) produced by QspB 7 A3, respectively. The two strains did also not differ in terms of acetic and butyric acid production. QspB 7 A2 also appeared to have no biological activity, having acetone and butanol levels slightly below pMTL85143-*qsrB*, 1 ± 0.2 mM and 6 ± 0.1 mM, respectively (p -values of 0.04, 0.008).

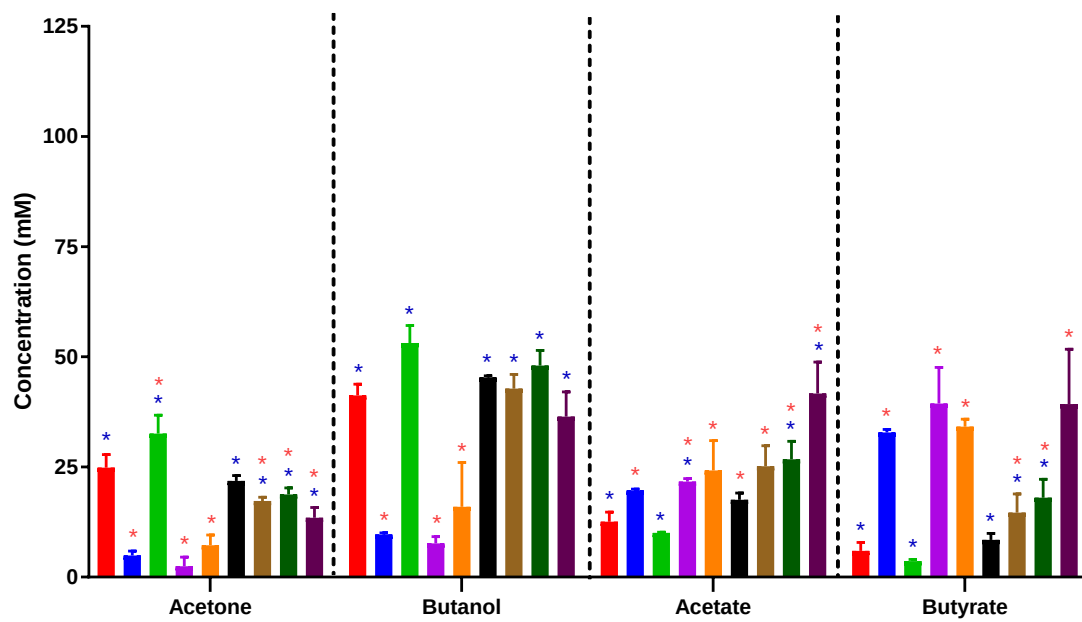
Table 3.1 QspB 7 peptide variants used in alanine scanning experiments

Peptide Name	Amino Acid Sequence
QspB 7	AEPTWGW
QspB 7 A2	AAPTWGW
QspB 7 A3	AEATWGW
QspB 7 A4	AEPAWGW
QspB 7 A5	AEPTAGW
QspB 7 A6	AEPTWAW
QspB 7 A7	AEPTWGA

24 h



48 h



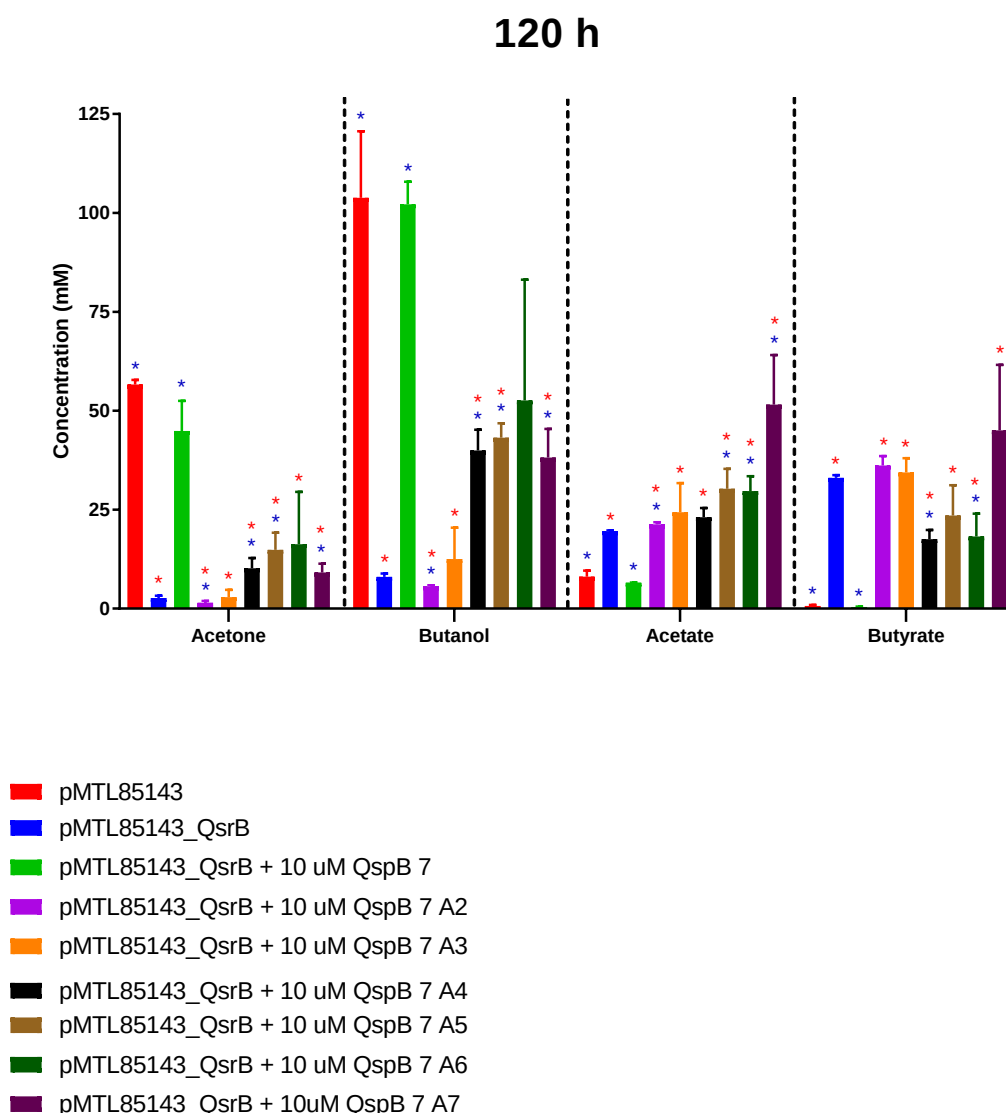


Figure 3.10 Biological activity of QspB 7 variants generated by alanine scanning. *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* was grown in CBM-S medium supplemented with 10 μ M of the indicated QspB 7 peptide variants, with amino acids exchanged for alanine, as shown in Table 3.1. Fermentation products were analysed after 24 hours (top left), 48 hours (top right), and 120 hours (bottom) of fermentation. The bottom right legend shows the coordinated colour for each peptide variant, with their order maintained from left to right for each of the fermentative products. The hashed line indicates when the next group of fermentative products begins. Each column represents the mean average of 3 replicates with error bars representing the SD of the mean. Asterisks indicate a statistically significant result from a two tailed t-test compared with the pMTL85143-*qsrB* DMSO control (blue) or the pMTL85143 DMSO control (red) (* indicates $P \leq 0.05$).

QspB 7 A4 to A7, however, behaved differently, producing significantly more acetone and butanol than the pMTL85143-*qsrB* DMSO control strain, but not as much as the pMTL85143-*qsrB* control to which QspB 7 had been added. The later produced 45 ± 3.6 mM and 102 ± 2.7 mM, acetone and butanol, respectively, whereas QspB 7 variants A4, A5 and A7 produced significantly less with 9 ± 1.0 mM to 15 ± 2.0 mM

acetone and 38 ± 3.4 mM to 43 ± 1.7 mM butanol (p -values between 0.0001 and 0.004 for acetone and 0.0001 and 0.002 for butanol). QspB 7 A6 differed slightly, producing 16 ± 6.2 mM of acetone (p -value 0.03), 53 ± 14.4 mM butanol that was marginally insignificant (p -value 0.0504). However, production of acids was more complicated in QspB 7 A4 to A7 compared to the pMTL85143-*qsrB* DMSO control. Regarding acetate, A4 did not differ significantly whilst A5 to A7 had significantly increased levels of acetate (p -values 0.009 to 0.02). Butyrate levels did not differ significantly for A5 and A7, yet were significantly reduced in A4 and A6 (p -values 0.0004 and 0.01, respectively).

The pMTL85143-*qsrB* DMSO control produced an average of 20 ± 0.1 mM and 33 ± 0.3 mM for acetate and butyrate, respectively, whereas A4 to A6 produced between 23 ± 1.1 mM and 30 ± 2.4 mM of acetate and 18 ± 1.1 mM and 24 ± 3.6 mM of butyrate. Interestingly, QspB 7 A7 behaved differently in terms of acid production, with acetate concentrations of 52 ± 5.9 mM exceeding those observed for the three controls, pMTL85143-*qsrB* DMSO (20 ± 0.1 mM), pMTL85143 DMSO (8 ± 0.7 mM) and pMTL85143-*qsrB* QspB 7 (7 ± 0.03 mM), with p -values of 0.01, 0.004 and 0.003, respectively. Similarly, butyrate levels were higher than the controls, as it produced 45 ± 7.8 mM, differing significantly from pMTL85143 DMSO (1 ± 0.1 mM, p -value 0.01) and pMTL85143-*qsrB* QspB 7 (0.4 ± 0.06 mM, p -value 0.01).

The described trends for acid production in the presence of QspB 7 A7 were also already observed after 24 hours and 48 hours, as was the reduced production of butyric acid for QspB 7 A4 and A6. This was also generally true for the other observations made, i.e. solvent production after addition of QspB 7 A2 and A3 was similar to that of the pMTL85143-*qsrB* DMSO control for both earlier time points, and solvent production in the presence of A4, A5, and A6 reached levels between those pMTL85143-*qsrB* and pMTL85143 DMSO controls.

In summary, as can be seen by heat-mapped Table 3.2, the most obvious changes in biological activity of QspB 7 are associated with the alteration of the glutamic acid and proline residues located in position 2 and 3 from the N-terminus, respectively. The alterations that resulted in the least percentage reduction in butanol formation were the alterations of position 4, 5 and 6 from the N terminus which in turn were

threonine, tryptophan and glycine residues, respectively. Finally, alteration of the C terminal tryptophan lead to a large reduction in activity which was roughly half that seen for positions 2 and 3.

Table 3.2 Biological activity of Qsp 7 alanine scan variants in terms butanol production when *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* is complemented with the corresponding synthetic peptide shown within the left hand column, relative to complementation with the QspB 7 peptide. Values are given for percentage reduction at 24, 48 and 5 days of fermentation indicated across the top row of the table.

Supplemented Peptide Name	Amino Acid replaced	Δ% of Butanol at 24 h	Δ% of Butanol at 48 h	Δ% of Butanol at 120 h
QspB 7 A2	Glutamic Acid	81.4	85.6	94.5
QspB 7 A3	Proline	63.1	70.0	87.8
QspB 7 A4	Threonine	0.1	14.7	60.9
QspB 7 A5	Tryptophan	11.0	19.5	57.7
QspB 7 A6	Glycine	1.6	9.6	48.5
QspB 7 A7	Tryptophan	42.5	31.5	62.7

3.3.6 Investigation of *C. acetobutylicum* ATCC 824 RNPP-type system cross talk

Previous studies have shown that insertional mutation, using the ClosTron technique, of *qspB* did not lead to a discernible reduction in early solvent production (Kotte, 2013). This was unexpected given the phenotype of QsrB overexpression, which suggested that QsrB is acting as a repressor that is inactivated upon binding the QspB peptide. It was hence hypothesised that a possible explanation for this would be the engagement of QssB in intersystem cross talk, allowing a peptide from one of the RNPP type systems present in the organism to also act as a conjugate signal to QsrB. While the specificity of QsrB appears high, given the empirical evidence shown in Figure 3.9 and Figure 3.10, the possibility was investigated to eliminate or confirm this hypothesis. This was performed via the addition of 10 μM of the synthetic conjugate peptides for systems QssA, QssC, QssD, QssE, QssF, QssG and QssH to *C.*

acetobutylicum ATCC 824 pMTL85143-*qsrB*, alongside a pMTL85143 vector control.

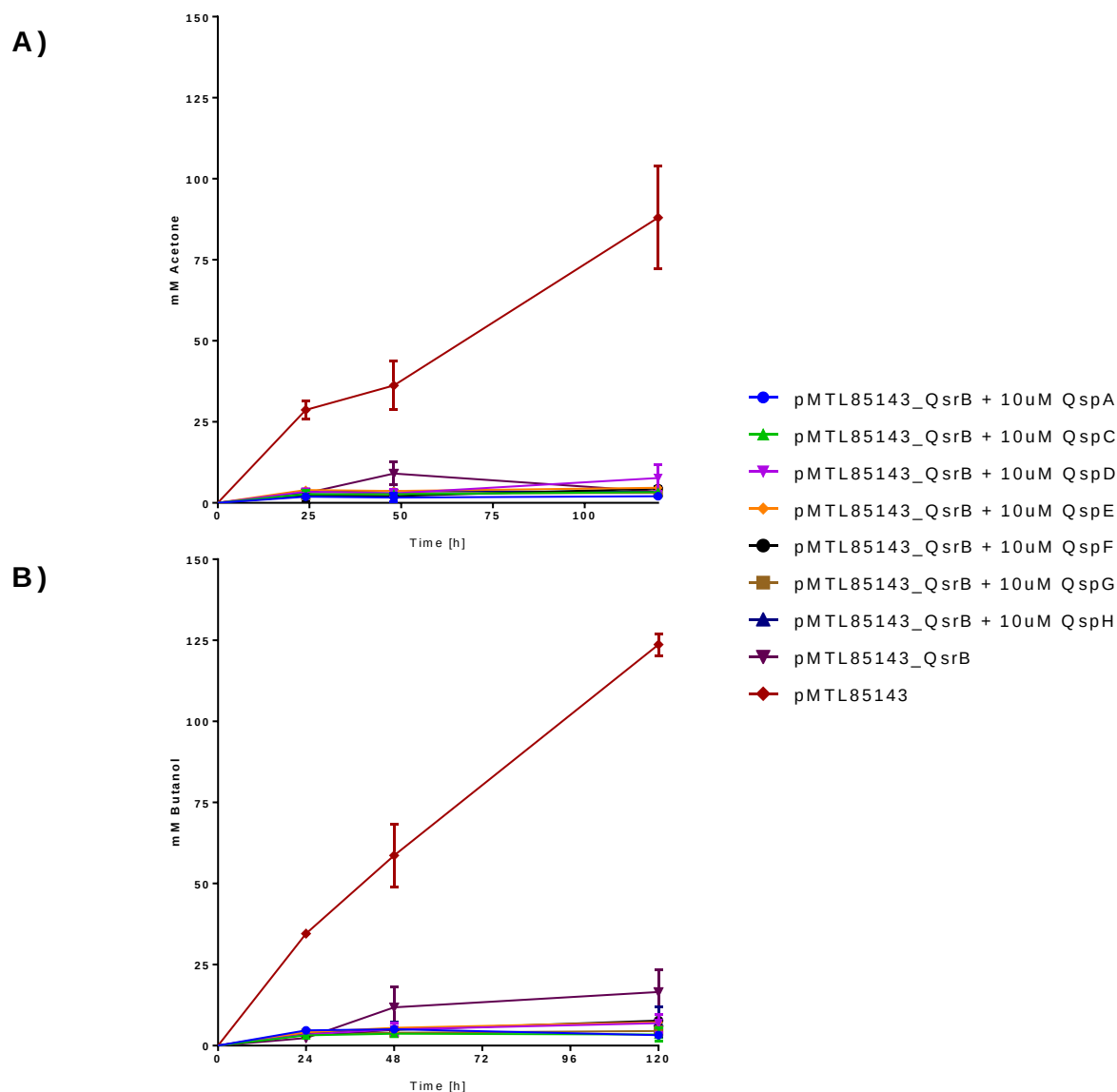


Figure 3.11 Synthetic peptides derived from other native RNPP type systems did not activate solvent formation by *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB*. Seven synthetic peptides representing the predicted signalling peptides for QssA to QssH were used to supplement cultures of the pMTL85143-*qsrB* strain at a final concentration of 10 μ M. Each peptide consisted of a 9 amino acid sequence as previously predicted by (Kotte, 2013; Kotte et al., 2017). The pMTL85143-*qsrB* and pMTL85143 control strains were supplemented with equivalent amounts of DMSO instead. A) and B) show the production of acetone and butanol by these cultures, respectively. Each time point represents the mean of 3 replicates, with the error bars representing the SD of the mean.

Each of the synthetic peptides were generated based on their proposed 9 amino acid sequence (Kotte, 2013; Kotte et al., 2017) (see Table 2.4 for their sequence).

As can be seen from Figure 3.11, each of the provided non-QspB conjugate peptides failed to act in a complimentary fashion upon addition to the pMTL85143-*qsrB* strain. Analysis of the solvent data shown in Figure 3.11 revealed that none of the synthetic peptides from the other RNPP type systems generated acetone or butanol concentrations that differed significantly to the *qsrB* overexpression control.

3.3.7 Supplementation of *C. acetobutylicum* ATCC 824 wild type with synthetic QspB peptides in order to assess its ability to affect solvent yield

The ability for QspB synthetic peptides to increase solvent yields of the pMTL85143-*qsrB* strain above those of the pMTL85143 control, led to experimentation with the synthetic QspB peptides in a wild type *C. acetobutylicum* ATCC 824 background. It was hypothesised that supplementation of the wild type *C. acetobutylicum* ATCC 824 with QspB conjugate peptide could lead to more efficient production of solvents at an earlier time point by binding available QsrB and alleviating any repression it may have on early solventogenesis. Simultaneously, it was hypothesised that the QspB synthetic peptide generated from the *C. roseum* sequence might act as a competitive inhibitor of the QsrB binding, or for interaction with the Specific Binding Protein (SBP) binding during ABC-mediated peptide uptake, leading to increased repression of solvent production. This was analysed by the addition of 10 μ M of QspB 10, QspB 9, QspB 8, QspB 7 and QspB 9 *roseum* to wild type *C. acetobutylicum* ATCC 824 cultures alongside a DMSO addition control.

As can be seen by the top section of Figure 3.12, after 24 hours of fermentation the butanol levels of wild type *C. acetobutylicum* ATCC 824 supplemented with the various QspB peptides did not show any statistically significant increase over that of the DMSO control (all *p*-values > 0.05). After 48 hours of fermentation, indicated by the middle section of Figure 3.12, the *C. acetobutylicum* ATCC 824 wild type DMSO control produced 55 ± 5.5 mM of butanol. None of the peptide supplemented cultures produced any more butanol than that, and the concentrations reached remained statistically indistinguishable from the WT control.

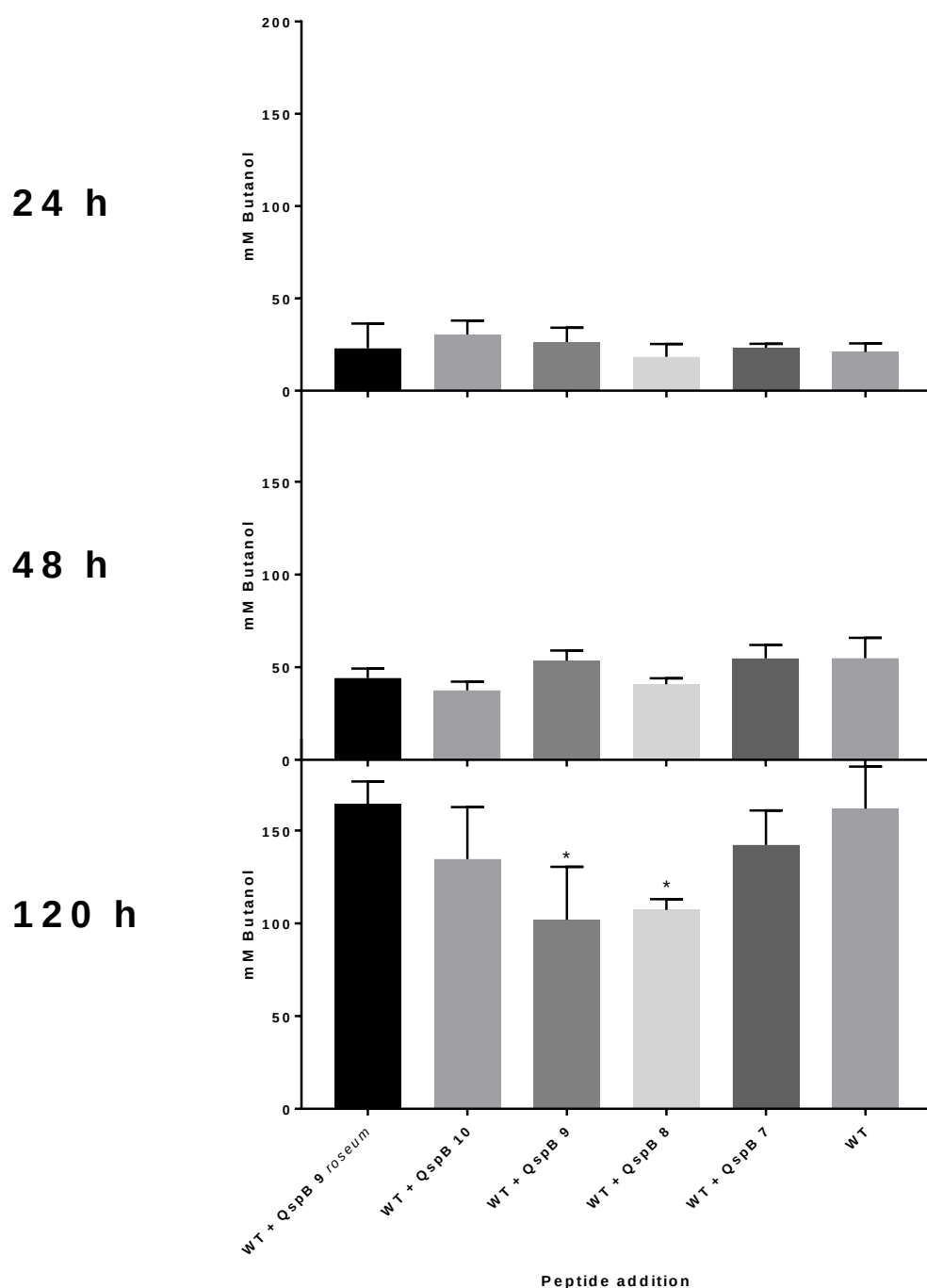


Figure 3.12 Butanol production of *C. acetobutylicum* ATCC 824 wild type upon addition of QspB 9 roseum, QspB 10, QspB 9, QspB 8, QspB 7 and DMSO. The columns represent the mean average of 3 replicates with the error bars representing the SD of the mean. Top, middle and bottom sections show 24 hour, 48 hour and 120 hour time points. Asterisks indicate a statistically significant difference between indicated and the *C. acetobutylicum* ATCC 824 wild type control (* = p -value ≤ 0.05).

However, as can be indicated by the bottom section of Figure 3.12, after 120 hours of fermentation butanol values began to differ. At this time point the *C. acetobutylicum* ATCC 824 wildtype DMSO control produced 162 ± 10.7 mM of butanol. Addition of

10 μ M of *C. roseum* QspB 9, *C. acetobutylicum* ATCC 824 QspB 10 or QspB 7 did not alter the level of butanol produced by the wild type strain, producing 165 ± 5.7 mM, 135 ± 13.2 mM and 142 ± 8.8 mM of butanol respectively. Interestingly, however, supplementation with 10 μ M of QspB 9 or QspB 8 lead to a decrease in butanol produced, 102 ± 13.5 mM and 107 ± 2.7 mM, respectively (p -values 0.047, 0.016).

3.4 Discussion

Overexpression of the RNPP type quorum sensing regulator QsrB was found in this study to cause a vast reduction in solvent production and endospore formation, in line with the previous report (Kotte, 2013). It has previously been shown that inactivation of QsrB via insertional mutation lead to increased solvent production at earlier time points, with the overall solvent production at the end point of the experiment remaining similar to that of the control strain (Kotte, 2013; Kotte et al., 2017). However, insertional mutations such as those created by ClosTron integration can have polar effects and it was thus desirable to independently confirm this observation (Cartman et al., 2012). Unfortunately, however, despite multiple attempts and redesigns of the involved vectors, a clean in-frame deletion of the CA_C0324 regulator could not be obtained. The allelic exchange method which was employed was developed in the CRG group at Nottingham and has been used in multiple applications (Cartman et al., 2012; Ehsaan et al., 2016). Failure to use this method for genes previously inactivated by insertional ClosTron mutagenesis has been observed before on several occasions (Dr Muhammed Ehsaan, personal communication). It is possible that changes to the *qsrB* gene (CA_C0324) region are instable or have polar effects on the essential CA_C0323 flanking gene. The before mentioned gene is one of three histidine kinases that dictate the phosphorylation state of Spo0A, the master regulator of *C. acetobutylicum* ATCC 824 (Steiner et al., 2011). Indeed, in frame deletion of the upstream CA_C0323 gene could not be obtained (Dr Zak Bean, University of Nottingham, personal communication). ClosTron mutation of this locus, encoding a histidine kinase essential for Spo0A phosphorylation was reported by Steiner et al. (2011), although stock cultures of these mutants could not be reactivated and mutant was confirmed to be instable by personal communication with the author.

It could therefore be that the *qsrB* locus is impervious to removal via the in-frame deletion technique. Future work may therefore involve the introduction of much smaller internal deletions or frameshift mutations.

Due to the failure to obtain *qsrB* in-frame deletions, further characterisation of the system made use of the drastic phenotypic effects observed for *qsrB* overexpression. As previously mentioned, addition of active conjugate peptides to the pMTL85143-*qsrB* strain, led on multiple occasions to solvent levels that were higher than those shown by the pMTL85143 control strain. However, this effect was not observed following direct additions of active QspB peptides to a wild type *C. acetobutylicum* ATCC 824 culture. This does not appear to be an anomaly, and likely refers to the restraints of the experiment as it has been observed that inclusion of a plasmid within the *C. acetobutylicum* ATCC 824 can alter its solvent profile, and is the reason special care is taken when selecting controls for growth experiments (Kotte, 2013). The ferredoxin promoter utilised within the pMTL85143 is shown to lead to very strong expression within *C. acetobutylicum* ATCC 824 (Heap et al., 2009). Transcription still takes place within the empty vector, which would still lead to the strong recruitment of RNA polymerase and associated sigma factor, which when considering a multi-copy number plasmid would likely add a considerable metabolic burden on the cell (Mairhofer et al., 2013). This may explain the slight alteration in solvent profile seen in *C. acetobutylicum* ATCC 824 upon presence of a plasmid relative to a wild type *C. acetobutylicum* ATCC 824 strain, as observed here for pMTL85143. The pMTL85143-*qsrB* strain, when supplemented with biologically active QspB, does not appear to suffer accordingly. This could hint that the QssB system has more functions than repressing solvent production, and is possibly involved in other functions.

The observable de-repression of solventogenesis via QspB addition would imply, assuming that QssB complies with the core tenants of QS regulation, that high cell density leads to an onset of solventogenesis. With that in mind, it is interesting that addition of synthetic peptides to the *C. acetobutylicum* ATCC 824 wild type culture had no observable increase in early solvents, and could relate to the QssB system operating at saturated levels during the time frame of investigation. While relatively little is known about the natural niche of *C. acetobutylicum* ATCC 824, it could be

that the in vitro experimentation does not represent the density at which this system is naturally active. This would assume that the expression of QsrB is constitutive, which transcriptomics data has indicated is the case (<http://exploration.weizmann.ac.il/TCOL/>). It could be possible that QssB is implicated within the catabolite repression system of the organism (Ren et al., 2012). Further insight would be gained via considering the glucose consumption of the QsrB overexpression strain, which is investigated within Chapter 4.

With respect to the QspB signalling peptide, we have shown that the minimum sequence required for biological activity is the 7 amino acid peptide AEPTWGW, here referred to as QspB 7. Whilst this does not necessarily represent the length of the native peptide that is active within wild type cultures, it does give us some insight into the requirements of the QssB. The majority of these residues contained within QspB 7 are small and non-polar, with the exception of glutamic acid with a relatively large and acidic side chain, and threonine with a polar side chain. Disruption of the glutamic acid and proline residues of QspB 7 led to the largest reduction in biological activity (Table 3.2). This might be explained if the charged glutamic acid contributes considerably to specific recognition and binding, such as those required in the second position of the Phr proteins of the Rap system (Gallego del Sol and Marina, 2013). The proline residue may have implications in the peptides structure due to angle formation. Another interesting note on the AEPTWGW peptide stretch concerns the tryptophan residues, which while being non-polar do possess relatively large aromatic and hydrophobic side chains. Replacement of both residues caused significantly reduced biological activity, as tryptophan has been shown to have a preference for polar/apolar interfaces, this could indicate the binding site of QsrB (Liu and Caffrey, 2006). Roughly 50 % of the biological activity of QspB 7 was lost when substituting the threonine residue for an alanine. This could be due to the removal of the polar group that threonine has, which could adversely affect its ability to interact with its proposed QsrB binding site. However, in contrast to glutamic acid and proline, its contribution to peptide binding is non-essential. Exchange of the glycine residue for an alanine, examined by the use of QspB 7 A6, lead to the least alteration in activity. This could be due to the relative similarity between alanine and glycine, both being very small and non-polar amino acids. One possibility of decreased biological activity

during the alanine scan is that these alterations effect uptake of peptide, by disrupting SBP binding, and subsequent uptake by the associated ABC transporter. Therefore, experiments are required that study peptide interaction with purified QsrB (See General Discussion).

It was interesting to observe that no biological activity was seen when the QspB 9 *C. roseum* peptide was used to supplement the pMTL85143-*qsrB* strain. The lack of activity of the QspB 9 *C. roseum* peptide could be due to the alteration of amino acid 7 and 8 from the N-terminus, tryptophan and glycine residues, exchanged to an arginine and proline residue, respectively. Arginine contains a large basic side chain, replacing the large aromatic side chain of tryptophan. While both amino acids contain large side chains, which could be required for interaction with conjugate regulator, the QspB sequence tested from *C. acetobutylicum* ATCC 824 contains no basic amino acids, so it is unlikely their presence would have any binding-promoting effects, and instead could substantially hinder the activity of this particular signalling peptide.

As the QspB 9 *C. roseum* peptide is slightly longer the QspB 7 investigated by alanine scanning it is impossible to know from the current evidence if the second residue from the N-terminus contributes to this lack of activity. In QspB 9 *C. roseum* this residue is an alanine, and in QspB 9 it is a glycine. Both are small non-polar amino acids, and so may have similar functions. However, it could be this amino acid alteration affects a cleavage site, preventing the activity of the active peptide post cleavage. It is not clear in *C. acetobutylicum* ATCC 824 if it produces external proteases to activate an external pro-signal for QssB, however such proteases have been shown to be essential for signalling in other organisms (Cabrera et al., 2014). Further experiments are required to fully understand the contribution of the 1st and 2nd amino acid of the QspB 9 peptide to its activity. This could be done by alanine scanning these two positions as per the QspB 7 experiment. The reasoning for lack of QspB 9 *C. roseum* activity could also be further investigated, by singly exchanging amino acids of the *C. acetobutylicum* ATCC 824 QspB 9 peptide for those of the QspB 9 *C. roseum* peptide. Therefore, the contribution of each amino acid to reduction in activity can be viewed. Further detail is given in the General Discussion (6.2.2).

The above gives reasoning for the lack of biological activity shown by QspB 9 *C. roseum*. It could have been expected that *C. roseum* and *C. acetobutylicum* ATCC 824 would engage in cross-activation or inhibition. The organisms likely occupy the same ecological niche, and have a similar physiology, hence are likely in strong competition for resources, which are usual hall marks of cross-talk (Otto et al., 2001). However, that they do not is testament to the exquisite level of specificity operated by *C. acetobutylicum* ATCC 824 in respect to QssB signalling.

This is further shown by the attempt to overcome the effects of QsrB overexpression by adding synthetic conjugate peptides identified for the other 7 RNPP-type systems in this organism (Kotte, 2013; Kotte et al., 2017). The rationale for this experiment was introduced briefly within the results section. In the main, this was to explain the observation that when the conjugate peptide gene *qspB* was disrupted through insertional mutation, only slight changes in solvent production were observed after 72 hours of growth (Kotte, 2013). If QspB acted as the sole peptide signal for this system, one would have expected a dramatic phenotype reminiscent of *qsrB* overexpression, a reduction in solvent production and heat resistant CFU/ml, due to the proposed role of QsrB as a repressor which is inactivated following binding of its cognate signalling peptide (Kotte, 2013). It was therefore hypothesised that another signal was present within the secretome of *C. acetobutylicum* ATCC 824 that might act as a conjugate signal to QsrB. It was shown that none of the proposed peptides from other native systems, QssA and QssC through to QssH, were capable of relieving the QsrB phenotype during supplementation. While it could be argued that there is an alternative signalling peptide within the *C. acetobutylicum* ATCC 824 genome, there is an alternative hypothesis that also could give some insight into the maturation of the QspB signal. The previous study by Kotte (2013) targeted nucleotide position 83/84 of the *qspB* structural gene for insertional mutation via the ClosTron technique (Kotte, 2013). This would leave the coding region for all tested biologically active QspB variants intact. For instance, the QspB 7 coding sequence is located 25 base pairs downstream of the ClosTron insertion site. It is possible that the *erm* promoter of the ClosTron insertion would drive transcription of this area. Interestingly, the leucine residue, the 1st residue in QspB 9, is encoded by TTG which is an alternative start codon. This transcript could be translated and still have activity, as peptides longer

then QspB 10 have previously been shown to have activity (Kotte, 2013). In-frame deletion, if possible, might be used to confirm or discard the offered explanations. Alternatively, further evidence might be gathered through redesign of the ClosTron to target a different site within the QspB 7 sequence.

In summary, QsrB overproduction was confirmed to have effects on the whole metabolic profile of *C. acetobutylicum* ATCC 824, and vastly reduce endospore formation. The minimum active peptide in QssB was shown to be AEPTWGW. The contribution of each amino acid to the peptides biological activity was investigated by alanine scanning, and alteration of glutamic acid and proline shown to have the most drastic effects. The ability of QssB to engage in cross-talk with its other RNPP-type systems, and the homologous QssB system within *C. roseum* was investigated, and it was shown not to, in either an activating or inhibitory sense. The effects of *C. acetobutylicum* ATCC 824 wild type supplementation with QspB were investigated, and no increase was seen in early solvent generation. From the evidence presented here, it is likely that the effect of QssB on solvent production is indirect rather than direct. However, this will be expanded upon towards the end of the next chapter.

4. The insertional
mutation of the specific
binding proteins of the
oligopeptide permease
systems within
C. acetobutylicum ATCC
824

4 The insertional mutation of the specific binding proteins of the oligopeptide permease systems within *C. acetobutylicum* ATCC 824

4.1 Introduction

ABC transporters are a superfamily of membrane import and export systems present in all domains of life (Locher, 2004). In bacteria and archaea, they generally consist of highly conserved ATP binding proteins, two transmembrane proteins, and specific binding proteins (SBPs) (Tomii and Kanehisa, 1998). The SBPs were earlier known as periplasmic binding proteins, and occasionally still are, due to their easy release from the periplasmic extracts of Gram-negative bacteria (Maqbool et al., 2015). The function of an ABC transporter can often be determined by its SBP, and ABC transporters of particular relevance to Gram-positive cell-cell signalling are the oligopeptide permease systems containing an oligopeptide binding SBP (Maqbool et al., 2015). This is because as well as having integral roles in nutrient uptake, for acquiring peptides and hence amino acids, these systems are required for the uptake of peptide pheromones from the surrounding area (Gominet et al., 2001; Rocha-Estrada et al., 2010).

The general layout of an oligopeptide permease is shown in Figure 4.1. It specifically depicts the Opp system with respect to Gram-positive organism, the difference being that the SBP within such organisms is generally considered to be attached to the membrane via a linker region, whilst in Gram-negative organisms they are free to move within the periplasmic space (Perego et al., 1991). Interestingly, in the *B. subtilis* system, mutations in *oppA*, *oppB*, *oppC* and *oppD* genes all caused a failure of the permease to function, whereas a mutation in the *oppF* gene only lead to partial disruption. This has been used to suggest that the OppF subunit within Gram-positive organisms may dissociate from the OppBCDF complex, and perform ATP hydrolysis for another transport system (LeDeaux et al., 1997; Perego et al., 1991).

The SBP, shown in Figure 4.1.A as OppA, binds its cognate ligand, and presents it to the ABC cassette, in this case the OppBCDF complex (Solomon et al., 2003). During ATP binding to the ATP binding subunits, concomitant rearrangement causes the

ligand to be trapped between the integral membrane proteins (here OppB and OppC), and following this ATP hydrolysis drives the opening of the complex to the cytoplasm, allowing the import of the ligand (Beek et al., 2014).

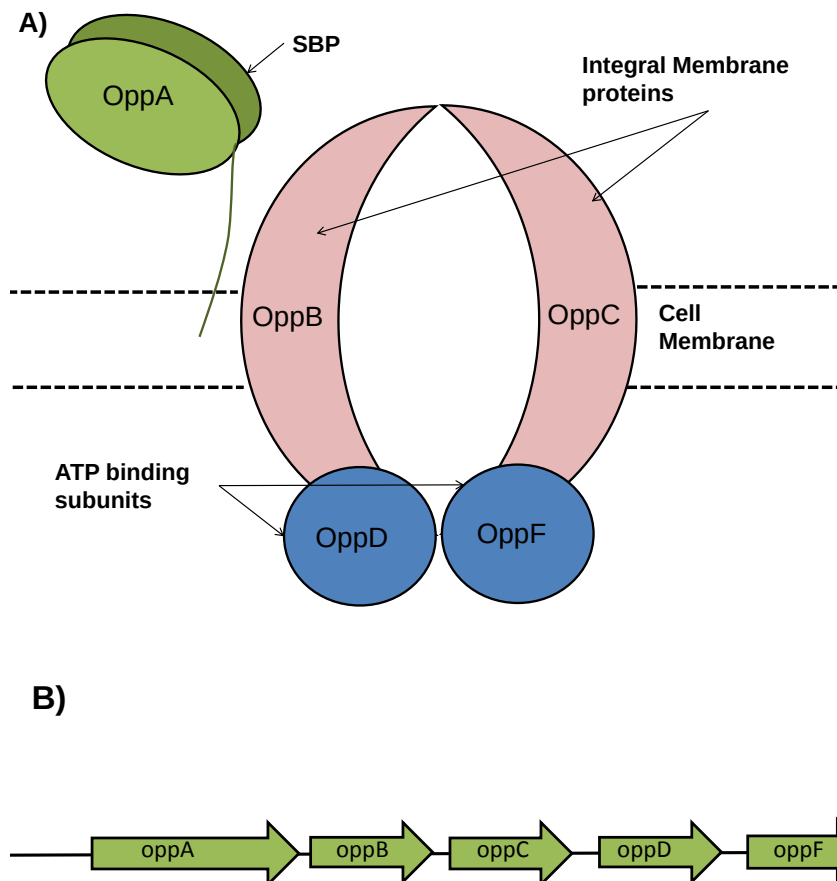


Figure 4.1 Composition (A) and operon organization (B) of a typical Gram positive oligopeptide ABC transporter complex, as exemplified by the Opp system. A) Outline of a typical oligopeptide permease system of Gram-positive bacteria, represented by the OppA and OppBCDF complex. The SBP protein, here OppA (green) is shown attached to the membrane via a linker region, whereas in Gram-negative bacteria it is free to move in the periplasmic space. The two lobes of OppA can be seen, in-between which a peptide can bind in a ‘Venus fly trap’ motion. The integral membrane proteins are shown (pink) here represented by OppB and OppC, together they make up a channel that through ATP hydrolysis allows transport across the membrane. The ATP binding domains (blue) are highly conserved between species and catalyse ATP hydrolysis to allow active, energy-dependent transport across the membrane, here shown as OppD and OppF. B) The encoding *oppABCDF* genes are located in a single operon which is usual for ABC transporter systems.

The relevance of these systems, with respect to Gram-positive cell-cell signalling, arguably became apparent with the identification of the *spo0K* locus of *B. subtilis*. The

locus did not encode a transcriptional regulator, but rather a well-known oligopermease system, the Opp system, present in many organisms and showing similarity to a peptide uptake system present in *Salmonella typhimurium* (Rudner et al., 1991). The SBPs that bind oligopeptides are often considered to bind these ligands generally, and much effort was exerted examining the ability of the SBPs to bind oligopeptides with little selectivity. One such paper demonstrates the ability of the OppA protein of *S. typhimurium* to bind the main chain atoms of the oligopeptide, whilst allowing the amino acid side chains to occupy fluid filled cavities where water molecules act as hydrogen bonding partners, and obscure charges that would hinder interaction (Sleigh et al., 1999). However, more recent studies have challenged this view, for instance, clever use of a mass spectrometer showed the *E. coli* OppA protein to favour positively charged oligopeptides (Klepsch et al., 2011). The selectivity of oligopeptide uptake was a source of debate in the study of *Lactococcus lactis*, it was eventually shown that the specificity of oligopeptide permease systems was wholly due to its SBP. This subject, and that of oligopeptide permeases in *L. lactis* generally, was well reviewed by Doeven and colleagues (2005).

The knockout of *oppA* within *B. subtilis* leads to loss of sporulation and genetic competence (Solomon et al., 2003). A study was conducted selecting for mutants of the *oppA* gene, and subsequently screened for competence (Solomon et al., 2003). It was found that some of the mutations lost endospore formation whilst remaining genetically competent, it was suggested by the authors that the encoded mutant proteins had altered uptake specificity, allowing uptake of some signalling peptides and not others (Solomon et al., 2003). Of course, most species contain multiple oligopeptide permease systems, with *B. subtilis* containing a dipeptide uptake system (Dpp), the before mentioned Opp system and another oligopeptide permease system termed App (Koide et al., 1999; Levnikov et al., 2005). While it has been shown that the Opp system preferentially takes up oligopeptides between 2 and 5 amino acids in length (Solomon et al., 2003), the App system was shown to prefer a 9 amino acid ligand (Levnikov et al., 2005). Further insights into the uptake of cell-cell signalling peptides in Gram-positive bacteria can be gained from *Bacillus thuringiensis*. Well known for its entomopathogenic properties, the virulence factors and life cycle of *B. thuringiensis* are well studied, and have shed considerable light on the quorum sensing

system that regulate them (Lemy et al., 2012; Lereclus et al., 2000; Slamti et al., 2014). An elegant study into insertional mutations which caused a pathogenically deficient phenotype in *B. thuringiensis* led to the discovery that an Opp-type system, homologous to that of *B. subtilis*, was required for pathogenicity (Gominet et al., 2001). It is therefore possible that a homologous system to either the Opp or the App ABC transporters would allow uptake of QspB in the *C. acetobutylicum* ATCC 824 system studied here.

Although a relatively rare occurrence, it is possible for multiple SBPs to function together with the same ABC transporter complex (Beek et al., 2014; Thomas, 2010). The oligopeptide permease systems are by no means an exception to this, and some fascinating examples have been shown of SBPs with different conditional requirements and specificities for oligopeptide binding, using the same ABC complex (Wang et al., 2004). However, the understanding of multiple SBPs using the same ABC cassette is far from complete, with even the *E. coli* K-12 lab strain containing multiple orphan SBPs with unstudied functions (Thomas, 2010).

Oligopeptide permeases are fascinating structures, and their purpose takes on new meaning when viewed through the lens of RNPP type cell-cell signalling. Oligopeptide permeases are required for the internalisation of RNPP type peptide signals (Gominet et al., 2001; Rocha-Estrada et al., 2010; Solomon et al., 2003). Certain studies have shown that oligopeptide SBPs can have a level of selectivity for what peptides they bind, and mutation of the encoding genes was hence shown to selectively disrupt peptide signalling in *B. subtilis* (Klepsch et al., 2011; Solomon et al., 2003; Wang et al., 2004). Further to this, as mentioned above, in certain circumstances multiple SBPs have been shown to interact with the same ABC transport complex, and hence change its specificity (Thomas, 2010; Wang et al., 2004). Very little literature surrounds ABC transporters within *C. acetobutylicum* ATCC 824, and even less so regarding oligopeptide permeases. This allows fertile ground for novel understanding of the oligopeptide permease systems within the organism, and more specifically, into the uptake of the RNPP type oligopeptide signals.

4.2 Objectives

The previous chapter investigated the conjugate signalling component of the QssB system. Within this chapter the focus changes to another important component of the RNPP type quorum sensing systems, the uptake of the conjugate signalling peptide from the culture supernatant into the cell. As previously outlined in the general introduction under heading 1.6.2 and the introduction of this chapter, this is performed via an oligopeptide permease system, that is an ABC transport systems containing SBPs capable of binding short peptides. In order to gain further insights into the uptake of the QspB signalling peptide, and possibly the uptake of signalling peptides of the other seven RNPP type QS systems within *C. acetobutylicum* ATCC 824, a number of objectives were devised:

- The bioinformatic identification of SBPs possibly required for the uptake of RNPP type QS conjugate signalling peptides into *C. acetobutylicum* ATCC 824.
- The insertional inactivation of each SBP via the use of the ClosTron technique, avoiding the isolation of degenerate strains.
- Screening of SBP mutants for phenotypes, particularly those known from previous studies to be linked to cell-cell communication: granulose formation, sporulation, colony size, and solvent formation.
- Identification of SBP mutants affected in uptake of QspB derived signals.

The ultimate aim of the above objectives were to identify the (potentially single) oligopermease uptake system required for the uptake of QspB and hence begin to further characterise the expression and uptake cycle of the RNPP-type QS system referred to as QssB.

4.3 Results

4.3.1 Bioinformatic identification of putative oligopeptide permeases in *C. acetobutylicum* ATCC 824

Given that the specificity of bacterial ABC transporter systems is generally defined by their associated SBPs, bioinformatic identification of the oligopeptide permease systems present in *C. acetobutylicum* ATCC 824 was investigated via pBLAST searches using established SBPs as a query sequence. Specifically, the protein SBP sequences for AppA (GenBank accession number: AAA62358.1) and OppA (GenBank accession number: CAB13000.1) from the Gram positive sporulation model organism *B. subtilis*, were used to identify putative homologs within the *C. acetobutylicum* ATCC 824 genome, which is inclusive of the native megaplasmid.

Table 4.1 Putative oligopeptide-specific SNPs identified in *C. acetobutylicum* ATCC 824

Gene Locus	Information	Putative Homologs	OppA Similarity	AppA Similarity
CA_C0176	In a predicted operon with <i>murQ</i> upstream of putative <i>appBCDF</i> operon	<i>appA</i>	264 of 564	264 of 564
CA_C2377	Genome sequence predicts frameshift	<i>appA</i>	78 of 157	63 of 145
CA_C2667	In a predicted monocistronic operon – orphan SBP.	<i>dppE</i>	240 of 561	-
CA_C3179	In a predicted operon with ABC transporter system components.	<i>appA</i>	239 of 563	272 of 563
CA_C3632	Annotated <i>oppA</i> and upstream of a predicted <i>oppBCDF</i> operon.	<i>oppA</i>	285 of 565	150 of 350
CA_C3634	In a predicted operon with ABC transporter system components. See below.	<i>oppA</i>	290 of 519	199 of 498
CA_C3640	CA_C3634 and CA_C3640 share an 89% nucleotide identity. Downstream of a predicted ABC transporter operon lacking a SBP gene.	<i>oppA</i>	312 of 560	230 of 546

The resultant identified genes and their predicted protein sequences were then used to ‘back-search’ against the well-annotated and curated genomes of *E. coli* and *B. subtilis*, to confirm their proposed role as peptide-specific SBPs. Using this method, 7 putative oligopeptide SBPs were identified, further details of which are shown in Table 4.1.

After identification of SBP encoding genes, the *C. acetobutylicum* ATCC 824 genome was analysed to identify their chromosomal location and the structure of their associated operons. A summary of this analysis is given in Figure 4.2 and Figure 4.3.

First, the more widely dispersed genes CA_C0176, CA_C2377, CA_C2667 and CA_C3179 are shown in Figure 4.2. The CA_C0176 gene is downstream of a *murQ* gene, and predicted to be within the same operon as it (Paredes et al., 2004). The *murQ* gene codes for a MurNAc-6-phosphate etherase, necessary for cell wall recycling in Gram-negative organisms, but recently shown to be essential for peptidoglycan recycling during the transition and stationary phases in Gram-positive bacteria (Borisova et al., 2016). Within the *C. acetobutylicum* ATCC 824 genome, CA_C0176 has been annotated as *appA*, which agrees with the above blast analysis, an oligopeptide-binding SBP shown in *B. subtilis* to bind a nonapeptide (Levdikov et al., 2005).

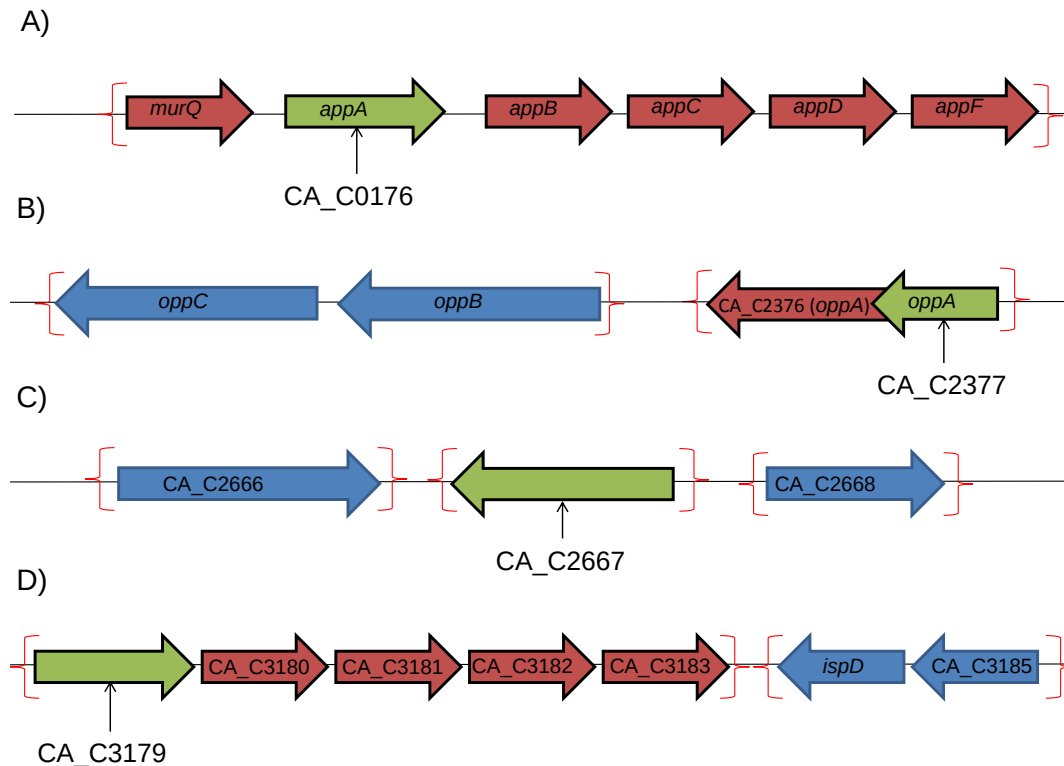


Figure 4.2 Arrangement of the CA_C0176, CA_C2377, CA_C2667 and CA_C3179 genes within the *C. acetobutylicum* ATCC 824 genome. Each arrow represents a gene, the identified SBP genes are shown (green). Red brackets indicate a predicted operon from the following sources; the DOOR² operon database, the Sorek lab transcriptomics database (<http://exploration.weizmann.ac.il/TCOL/>), and published *C. acetobutylicum* ATCC 824 transcriptional units (Paredes et al., 2004). Genes within the same predicted operon as an identified SBP are indicated (red) with all other genes distinguished (blue). CA_C2377 and CA_C2376 appear separated due to a frame shift mutation, causing a stop codon within CA_C2377.

CA_C2377 was annotated as *oppA* (provisionally), disagreeing with the above Blast analysis which returns the AppA protein of *B. subtilis*. The gene is located upstream of a bicistronic operon encoding further ABC transporter components, annotated as *oppB* and *oppC*. There are numerous genes upstream of this which does not appear to be related to oligopermease function. CA_C2373 is a predicted membrane protein of unknown function, whilst CA_C2372 is annotated as *sspA* and is a small acid soluble spore protein. Continuing upstream, CA_C2371 encodes an S-adenosylmethionine-dependent methyltransferase, whilst CA_C2370 encodes an uncharacterised protein. Interestingly, according to the published genome sequence CA_C2377 appears to have been disrupted via a frameshift mutation, causing a premature stop codon to appear 567 bp into the gene, eliminating the final 1,167 bp (which are annotated as separate orf, CA_C2666).

The frameshift mutation present within the published genome sequence (AE001437) was confirmed within the two independent parental genome sequences used within this group (Poehlein et al., 2017).

The third SBP gene, CA_C2667, was predicted to be an orphan located in a monocistronic operon, and flanked by unannotated genes coding in the opposing orientation, and hence little insight could be gained via bioinformatics and empirical data are thus required.

Another SBP homolog, encoded by CA_C3179, is predicted to form an operon together with 4 other genes, predicted by Blast search to encode the remaining components of an ABC transporter. It is likely this operon represents a complete oligopermease transporter.

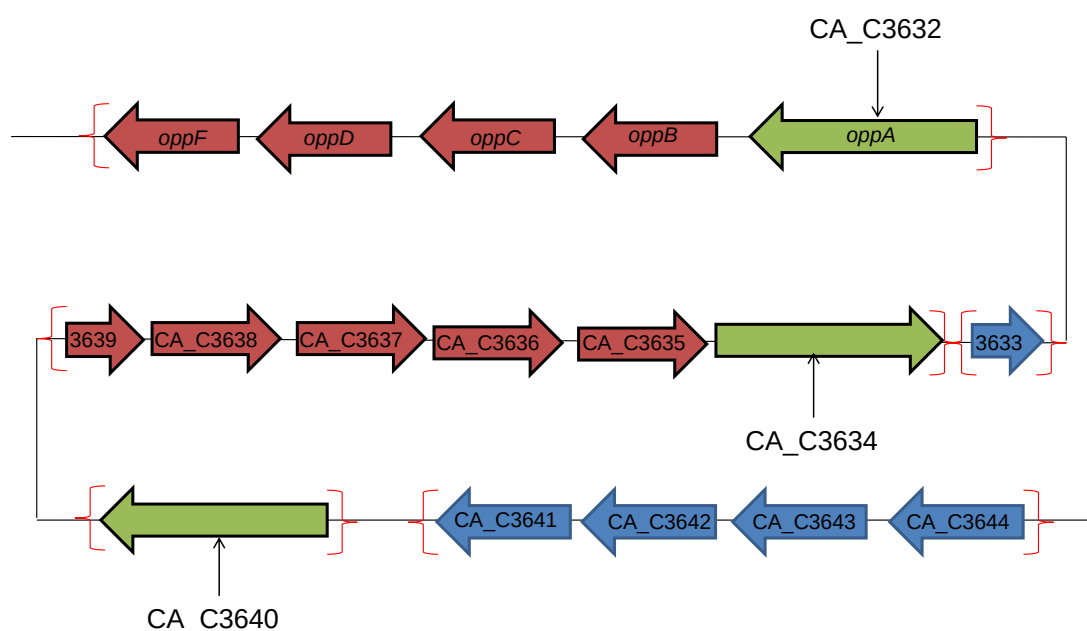


Figure 4.3 Genetic layout of CA_C3632, CA_C3634 and CA_C3640 within the *C. acetobutylicum* ATCC 824 genome. Each arrow represents a gene; the investigated SBP genes are indicated in green and other genes within the same predicted operon are shown as red, all other genes are shown in blue. Red brackets indicate a predicted operon from the following sources; the DOOR² operon database, the Sorek lab transcriptomics database (<http://exploration.weizmann.ac.il/TCOL/>), and published *C. acetobutylicum* ATCC 824 transcriptional units (Paredes et al., 2004).

Three additional putative SBPs, CA_C3632, CA_C3634 and CA_C3640, are encoded in very close vicinity to each other as shown in Figure 4.3. The CA_C3632 gene has been annotated *oppA* and from transcriptomic data is within a polycistronic operon coding the remaining components of a functioning ABC transporter, annotated as *oppBCDF* (<http://exploration.weizmann.ac.il/TCOL/>).

The CA_C3634 gene appeared in a predicted operon of multiple genes, the layout of which suggested a complete ABC transporter. CA_C3639, the first gene of this operon, had similarities to genes in multiple genomes, including multicellular eukaryotic organisms and numerous *C. difficile* strains, and was predicted to encode a CRO repressor like DNA binding protein. The other genes of the operon were predicted by Blast to encode the remaining components of an ABC transporter. It is likely then the CA_C3634 operon includes a complete oligopermease ABC transporter.

When subject to a Blast comparison, the CA_C3634 and CA_C3640 locus proteins were found to share an 89% identity, and therefore are likely duplicates.

The gene for the third SBP encoded in this region, CA_C3640, is predicted to form a monocistronic operon downstream of an operon containing 4 genes predicted to encode oligopeptide ABC transporter components. Blast searches indicated that each of the components had similar sequences within other *Clostridium spp.* all of which, where annotated, related to the functions of an oligopeptide ABC transporter system. It is possible that the two operons together encode all the functions required for a functional ABC transporter.

4.3.2 Insertional inactivation of the identified oligopermease SPBs via the ClosTron method

Following their identification, all seven putative peptide-specific SBPs, were subjected to ClosTron mutagenesis as described in section 2.5.2 (Heap et al., 2010, 2007). It was attempted to obtain two independent insertions in each SBP, to provide early evidence that observed phenotypes were not due to the before mentioned process of degeneration (Cornillot et al., 1997). This led to a large body of work obtaining and

confirming each insertional inactivation, and within the time frame of this project at least two completely independent mutations were obtained for the CA_C0176, CA_C2377, CA_C2667 and CA_C3632. This left insertional mutations for the potential duplicate genes CA_C3634 and CA_C3640, and CA_C3179 which could not be obtained within the time frame available. The insertional mutations were confirmed via PCR, using flanking primers of the insertional region, one flanking primer and the internal EBS universal primer to confirm insertion within the correct location (see Table 2.3 for primer sequences). For further trouble shooting a PCR was performed of the retroactive marker (RAM) using the RAM-F and RAM-R primers. Each insertional mutation was confirmed in this manner. Following this each of the strains were sent for genome sequencing to confirm a single insertion of the intron (awaiting data) and to obtain evidence for potential additional and unrelated mutations that may have accumulated during the process of mutant generation, selection and purification. SNP analysis can be performed to look for unintentional mutation, and ectopic insertions can be identified within the genomic data by locating the ClosTron native, reverse transcribed, sequences.

3 independent ClosTron mutations were obtained for the CA_C3632 locus, CA_C3632::C*TermB* 1A2, 2A1 and 3A1 respectively. 2 were obtained for the CA_C2377 locus, CA_C2377::C*TermB* A1 and 4D respectively. 2 were obtained for the CA_C2667 locus, CA_C2667::C*TermB* 1 and 3 respectively. Finally 3 were obtained for CA_C0176, CA_C0176::C*TermB* 1, 2 and A1 respectively.

4.3.3 Analysis of granulose formation within the SBP insertional mutations

Once confirmed the insertional mutations were analysed for granulose formation as per section 2.6.3 of the materials and methods.

Table 4.2 Production of granulose and extracellular amylase activity by SBP mutants

Strain	Granulose Formation	Amylase Activity
Wild Type	+ve	+ve
CA_C0176:: <i>CTermB</i>	+ve	+ve
CA_C2377:: <i>CTermB</i>	+ve	+ve
CA_C2667:: <i>CTermB</i>	-ve	-ve
CA_C3632:: <i>CTermB</i>	-ve	+ve

During this analysis the mutants were also analysed for amylase activity as an indicator of megaplasmid loss (discussed in 4.4) and hence degeneration. As can be seen from Table 4.2 the CA_C2667::*CTermB* and CA_C3632::*CTermB* strains both failed to produce granulose under the conditions of experimentation. Interestingly the CA_C3632::*CTermB* retained a high level of amylase activity. Each of the independently obtained insertional mutations behaved the same, and hence are displayed in Table 4.2 together. The lack of granulose formation and amylase activity of the CA_C2667::*CTermB* strain is highlighted in Figure 4.4. It is shown here alongside the *C. acetobutylicum* ATCC 824 WT strain for ease of comparison. The strains were grown on CBM-S starch plates, and stained with iodine after 48 hours of growth. Amylase activity can be seen by the zones of clearing around the bacterial growth indicating starch breakdown, whilst granulose accumulation is indicated by bacterial growth changing to a blue-black coloration post iodine stain. It is clear to see from Figure 4.4 that the *C. acetobutylicum* ATCC 824 CA_C2667::*CTermB* strain has done neither, which was shown in both of the independently obtained insertional mutants (data not shown).

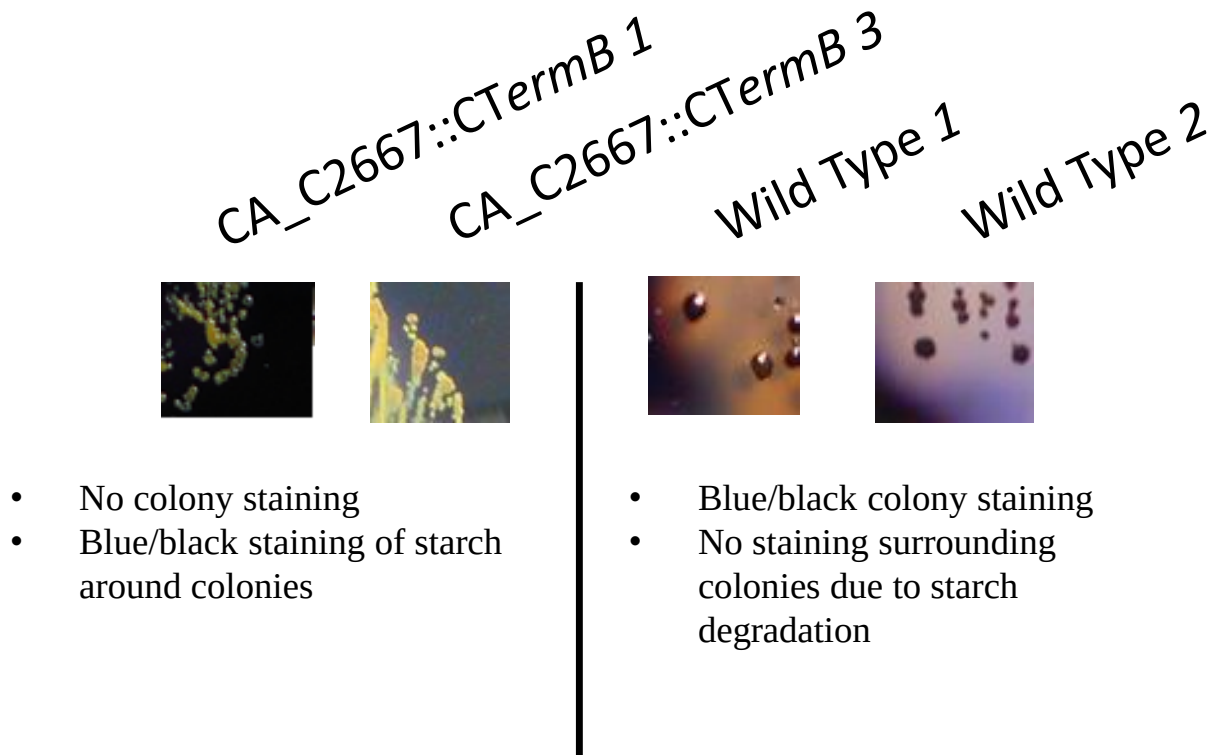


Figure 4.4 Lack of granulose formation and amylase activity within the *C. acetobutylicum* ATCC 824 CA_C2667::CTermB mutants. As the labelling indicates the left hand streak is the before mentioned insertional mutation strain and the right hand shows the *C. acetobutylicum* ATCC 824 WT strain. The strains were grown on CBM-S starch agar plates as outlined within the materials and methods section, and stained with iodine. The above experiment was performed using two independently obtained triplicates, and one triplicate shown for convenience. The image was cropped for more efficient annotation, other than this the image is original and unedited.

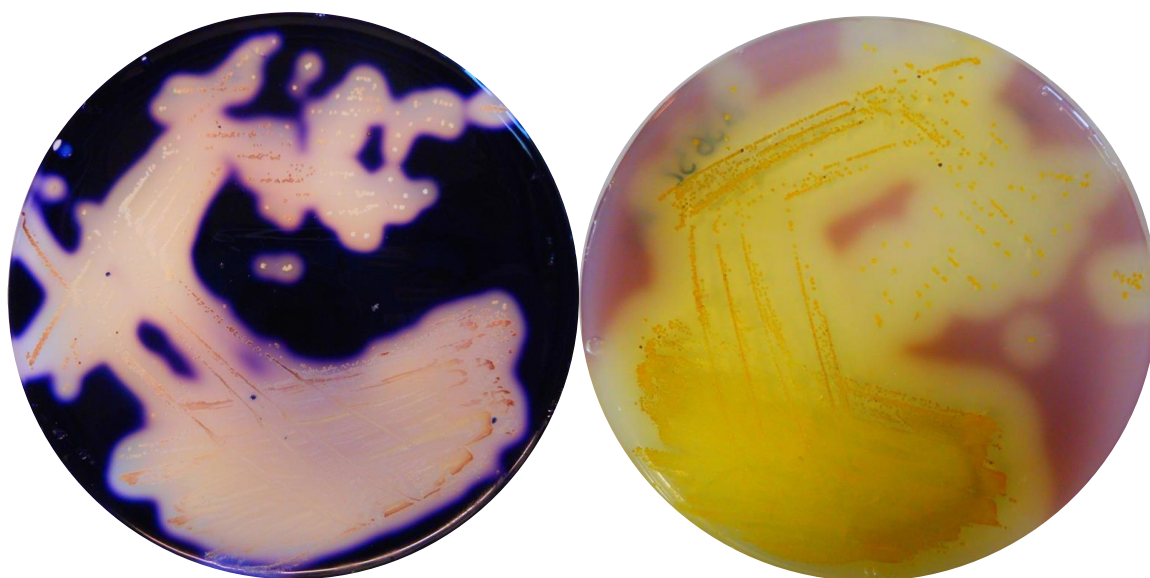


Figure 4.5 Lack of granulose formation with retained amylase activity within the *C. acetobutylicum* ATCC 824 CA_C3632::CTermB strain. The *C. acetobutylicum* ATCC 824 CA_C3632::CTermB strain is shown post iodine stain, on CBM-S starch medium (left) and CBM-S medium with calcium carbonate omitted for improved picture quality (right). The above experiment was performed in triplicate with three independently obtained mutants, but the results for only one replicate is shown for convenience. The images were cropped in a circular fashion to remove the background but where otherwise original and unedited.

The phenotype of the *C. acetobutylicum* ATCC 824 CA_C3632::CTermB mutant is shown in Figure 4.5. As can be seen from the iodine-stained agar plates displayed in the figure, amylase activity has resulted in large zones of clearing surrounding bacterial growth when grown on CBM-S starch plates for 48 hours. It can also be seen on both plates that the mutation of CA_C3632 has led to a lack of granulose accumulation, due to the failure of colonies of the strain to turn blue-black on iodine addition.

4.3.4 Analysis of fermentation products formed by SBP mutants

The solvent production and acid assimilation profile of each of the independently acquired SBP mutant strains was analysed. WT control and mutant strains were grown in CBM-S medium supplemented with 6% glucose and samples taken after 24 h, 48 h and 120 h. Determination of fermentation products was achieved through gas chromatography as previously described within the material and methods section (2.6.5). Unexpectedly, all mutants strains generated showed a reduction in solvent production. As can be seen in the upper section of Figure 4.6 the *C. acetobutylicum* ATCC 824 WT control had produced 20 ± 1.2 mM of acetone, 39 ± 2.2 mM of butanol and 1 ± 0.1 mM of ethanol after the first 24 hours of fermentation. It is interesting to note that the majority of the insertional mutations produced significantly less butanol, after the first 24 hours, then the wild type control. The *p*-values for this ranged from 0.0006 to 0.0066. The exceptions to the before mentioned trend were both isolates of the CA_C2667::CTermB, that did not differ significantly from the wild type control with respect to butanol at this time point, as well as a single variant of the CA_C0176 locus insertion, named CA_C0176::CTermB 1. This trend does not completely hold with respect to the production of acetone after the first 24 hours of fermentation. The majority of insertional mutations still produced significantly less acetone than the WT control, with *p*-values ranging from 0.0006 and 0.0362. However, this also included one of the independent isolations of the CA_C2667 locus ClosTron insertions, with only CA_C2667::CTermB 3 and the previous CA_C0176::CTermB 1 producing levels of acetone statistically indistinguishable from the WT control. It should be noted as previously stated that the WT produced a normal level of ethanol within the 24 hours, a number of the insertional mutations did not produced levels of ethanol detectable by the outlined gas chromatography protocol at this time.

The greatest levels of difference, and statistical significance, between the solvent profiles of the WT control and each of the SBP insertional mutations, occurred at the 48 hour time point. The whole solvent profile was reduced in each of the SBP mutants at the 48 hour time point which can be seen from Figure 4.6. The WT control produced 49 ± 0.2 mM of acetone, 97 ± 3.4 mM of butanol and 12 ± 0.7 mM of ethanol at this stage of the fermentation. Acetone seemed the most effected at this

stage with levels much lower than the WT control, this was corroborated by statistical analysis showing a much greater level of confidence for this observation. The lowest level of acetone produced at this stage was 7 ± 0.7 mM by the CA_C3632::*CTermB* 3A1 strain (p -value = 8.88×10^{-7}). The highest level of acetone produced at this stage, apart from the control, was 27 ± 2.34 mM by the CA_C0176::*CTermB* 1 strain (p -value = 0.0072).

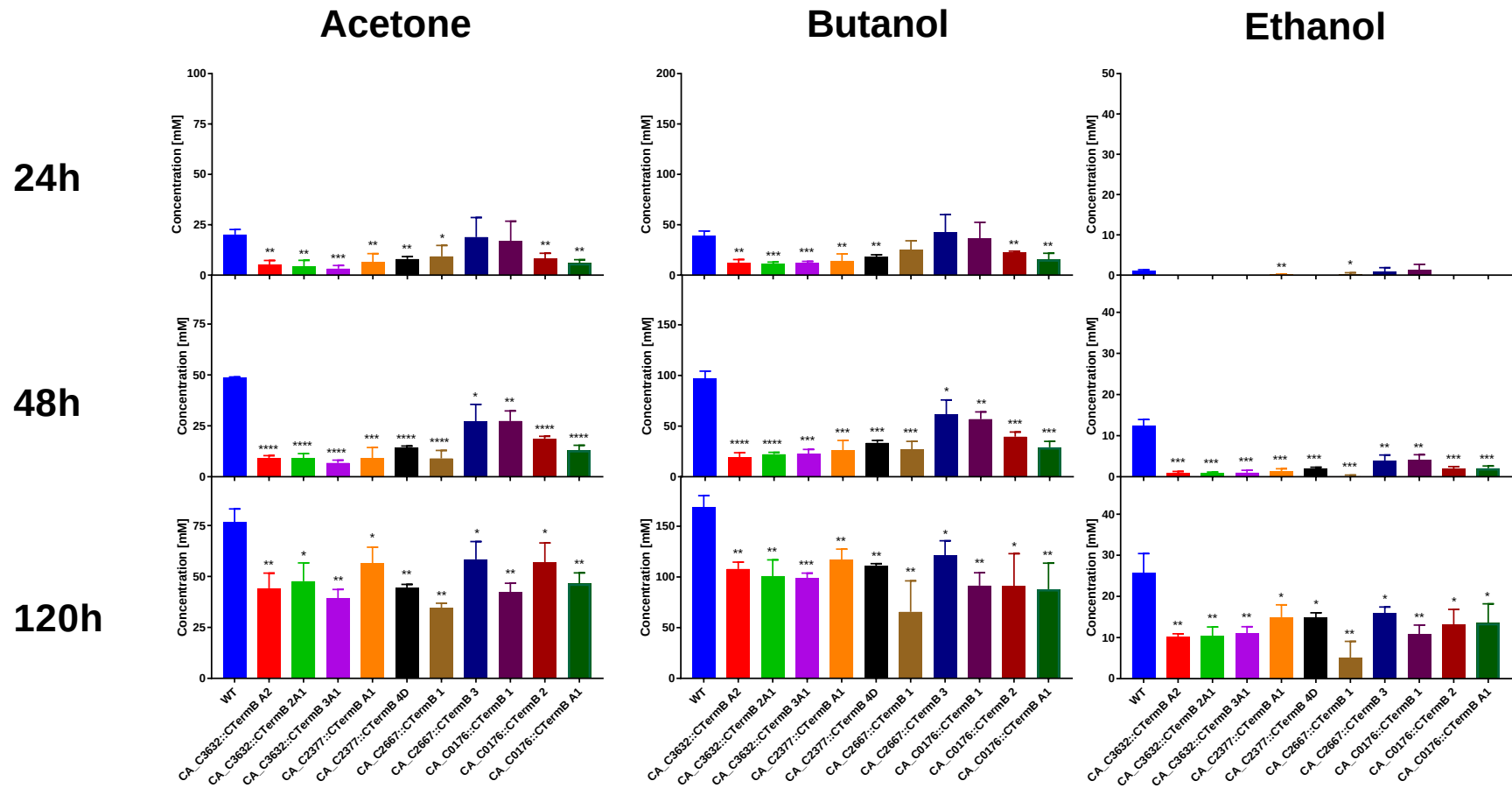


Figure 4.6 Analysis of solvent formation over 120 hours by each of the independent SBP mutants. Samples were taken after 24 h (top row), 48 h (middle row) and 120 h (bottom row). Levels of acetone (left) butanol (middle) and ethanol (right) are shown for each time point. The colour coding of each strain is maintained across the figure, during the course of this experimentation three independent isolations of ClosTron mutations were used for the CA_C3632::CTermB and CA_C0176::CTermB strains, and duplicate independent isolations used for CA_C2377::CTermB and CA_C2667::CTermB strains. Each of the columns represent the mean average of three replicates in order to mitigate biological variation with the error bars representing the standard deviation of the mean. Statistical significance is indicated by asterisks and show the result of a two tailed T-test between the individual strain and the WT control, where * indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$.

Levels of butanol were also severely affected after 48 h, with the lowest and highest level of 19 ± 2.2 mM and 61 ± 6.8 mM being produced by the CA_C3632::CTermB 1A2 and the CA_C2667::CTermB 3 strain respectively (p -values = 9.5×10^{-5} and 0.0187). The minimum level of ethanol seen at this time point was 0.15 ± 0.1 mM, produced from the CA_C2667::CTermB 1 (p -value = 1.78×10^{-4}) the highest level of ethanol produced by an insertional mutant was 4 ± 0.6 mM from the CA_C0176::CTermB 1 strain (p -value = 0.0019).

The trend holds true, but to lesser degree, even after 120 hours of fermentation. That is all of the tested individual oligopermease SBP mutants had a reduced solvent profile relative to the wild type control. After 120 hours of fermentation the WT control had produced 77 ± 3.1 mM acetone, 168 ± 5.6 mM butanol and 26 ± 2.2 mM ethanol. Each of the insertional mutations of the oligopermease SBP produced statistically significant less acetone ranging from 27 ± 6.6 mM produced by the CA_C2667::CTermB 1 strain, to 58 ± 4.1 mM produced by the CA_C2667::CTermB 3 strain (p -values = 0.0051 and 0.0456 respectively).

All of the insertional mutations of SBPs also produced less butanol than the WT control after 120 hours, again ranging from 65 ± 14.4 mM to 121 ± 6.9 mM, produced from the previously mentioned strains (p -values = 0.0055 and 0.0120 respectively).

The same trend is seen across the ethanol production within each of the insertional mutations. The lowest value recorded was 5 ± 1.9 mM and the highest recorded was 16 ± 0.7 mM, these values were again attributed to the before mentioned strains (p -values = 0.0042 and 0.0247 respectively). Despite these strains having insertional mutations within the same locus, CA_C2667, they produced significantly different levels of acetone (p -value = 0.0367), butanol (p -value = 0.0468) and ethanol (0.0117). To further analyse the phenotype of the obtained SBP insertional mutants, levels of heat resistant CFU/ml were determined following 120 hours of fermentation as a measure of endospore formation. This information is displayed within Figure 4.7. At this time the wild type *C. acetobutylicum* ATCC 824 strain produced 2.38×10^5 ($\pm$

3.01×10^4) heat resistant CFU/ml. As indicated in Figure 4.7, 3 strains produced significantly more spores than the WT control and 3 produced significantly less. The CA_C2667 strains are worthy of highlight, given that both independently obtained ClosTrons produced significantly less heat resistant CFU/ml than the WT control.

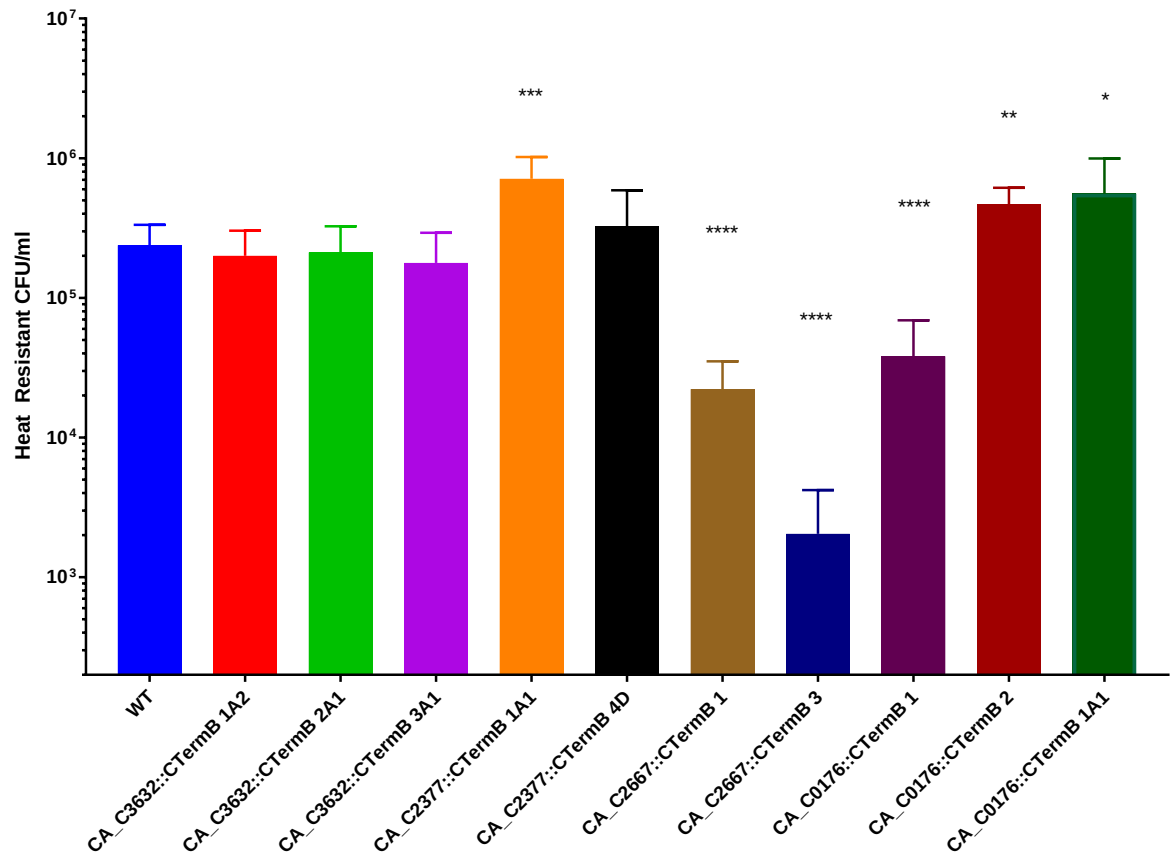


Figure 4.7 Heat resistant CFU/ml of each SBP insertional mutant after 120 hours of fermentation. Heat resistant CFU/ml was measured as indicated in the appropriate section of the materials and methods. Error bars indicate the standard error of the mean, with asterisks indicating the level of statistical significance when each mean is measured against the WT control. (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$). The Y axis initiates at 200 CFU/ml due to the detection limit of the experiment. Each of the columns shows the average of 9 replicates.

4.3.5 Introduction of the pMTL85143-*qsrB* overexpression plasmid into the SBP insertional mutation strains

Having isolated and characterised a number of SBP mutants, it was of interest to observe whether QsrB-mediated repression of solvent formation and sporulation in the individual strains could still be overcome by the addition of synthetic QspB peptide.

A single independent insertional mutant of each of the SBP was chosen to be transformed with the pMTL8153-*qsrB* overexpression plasmid. To minimize the risk that the chosen strain was degenerated due to the presence of additional, undesired mutations, derivatives were chosen that still stably produced solvents and that did not deviate in appearance from other independent mutants of the same SPB locus. The mutant chosen for each strain were CA_0176::CTermB 2, CA_C2377::CTermB 4D, CA_C2667::CTermB 3 and CA_C3632::CTermB 1A2. The ClosTron insertional mutants were transformed as per the protocol outlined within the material and methods section 2.3.4.2. Transformation efficiencies were as expected for the vector, with many colonies present, which were restreaked to purity maintaining Thiamphenicol selection and also erythromycin selection for the ClosTron RAM.

In order to form a fair control for the solvent analysis experiments, each of the SBP insertional mutants were also transformed with the pMTL85143 empty vector in the same manner as described above (2.3.4.2). Each of the transformed strains, as well as appropriate wild type controls, underwent fermentation for 120 hours, either in the presence of 10 mM synthetic peptide QspB7 (*qsrB* overexpression strains) or with equivalent amounts of DMSO added (all strains). At the end of the experiment, the wild type pMTL85143 DMSO control had produced 35 ± 4.2 mM of acetone, 72 ± 12.3 mM of butanol and 8 ± 2.0 mM of ethanol. By contrast, the pMTL85143-*qsrB* DMSO control had produced 3 ± 0.6 mM of acetone, 4 ± 0.6 mM of acetone and 0.6 ± 0.2 mM of ethanol in the same time frame.

The results with respect to acetone and butanol for all mutants and conditions are shown in Figure 4.8. As expected, introduction of the pMTL85143-*qsrB* plasmid into the SBP mutants resulted in a severe reduction of solvents during fermentation which generally could be overcome by the addition of QspB7 peptide.

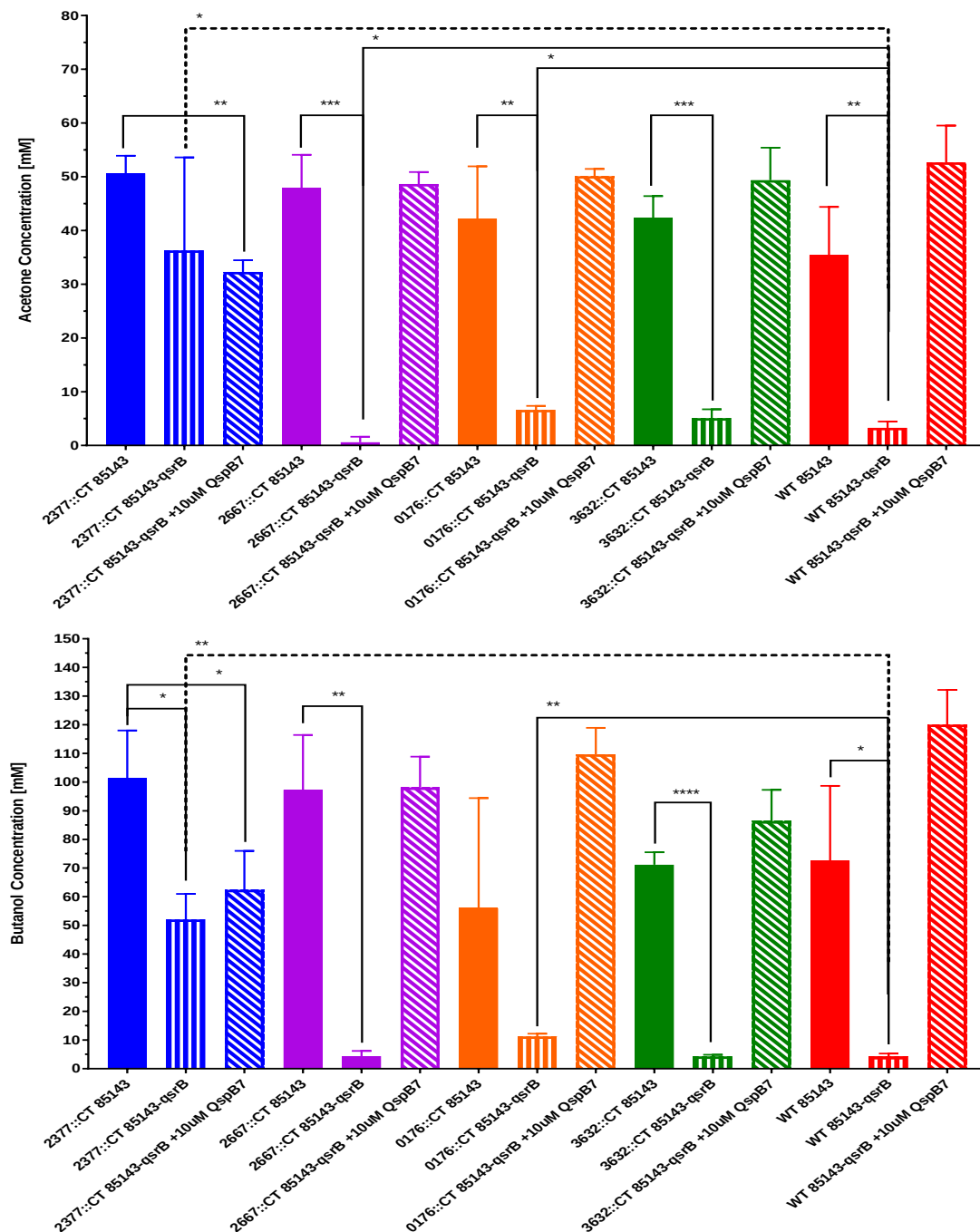


Figure 4.8 Acetone and Butanol production of SBP insertional mutants, containing either pMTL85143 empty vector control, pMTL85143-*qsrB* or pMTL85143-*qsrB* supplemented with 10 μM QspB. Samples were analysed after 120 hours of fermentation for acetone (upper panel) and butanol (lower panel). Each strain is represented by three bars, based on colour, showing the pMTL85143 control (solid colour), pMTL85143-*qsrB* (vertical hashed lines) and pMTL85143-*qsrB* with 10 μM QspB addition (angled hashed lines). Derivatives of CA_C2377::CTermB (blue); CA_C2667::CTermB (purple); CA_C0176::CTermB (orange); CA_C3632::CTermB (green) and *C. acetobutylicum* ATCC 824 control strain (red). Each of the bars show the average of 3 replicates, with error bars showing standard deviation of the mean. Connecting lines between two columns show the application of a T test, the *p*-values are indicated using asterisks (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

However, there was one notable exception: the reduction of solvents was far less dramatic in CA_C2377::CTermB containing the *qsrB* overexpression vector pMTL85143-*qsrB*. This strain produced higher levels of acetone and butanol when compared with the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* control (p -values = 0.03 and 0.0021, respectively). Furthermore, addition of synthetic QspB7 peptide failed to have any discernible effect on this strain with solvent remaining at the reduced levels seen for the DMSO control. The observed difference between CA_C2377::CTermB pMTL85143 vector control and the pMTL85143-*qsrB* containing overexpression strain was statistically significant for butanol but not for acetone. The former produced 101 ± 7.9 mM butanol, and the latter only 52 ± 3.6 mM (p -value = 0.0346), whereas acetone production was 51 ± 1.5 mM and 36 ± 8.2 mM (p -value = 0.231), respectively.

The CA_C0176::CTermB pMTL85143-*qsrB* strain also produced slightly higher levels of acetone and butanol compared to same control, this difference was statistically significant (p -value = 0.0136 and 0.0013 respectively). The levels of the conjugate acids acetate and butyrate after 120 hours of fermentation are shown in Figure 4.9. For the most part the acid data reflect the inverse of the above solvent data, which given the metabolism of *C. acetobutylicum* ATCC 824, is to be expected. At this time point the *C. acetobutylicum* ATCC 824 pMTL85143 supernatant contained 28 ± 1.2 mM of acetate and 10 ± 4.4 mM of butyrate, compared to 58 ± 1.1 mM of acetate and 49 ± 0.94 mM of butyrate present within the supernatant of the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* (p -values = 1.00×10^{-3} and 2.15×10^{-3} respectively). Furthermore, addition of QspB7 peptide reversed the effect of *qsrB* overexpression as already reported in the previous chapter (3.3.2). A similar behaviour was observed for all tested mutant strains with the exception, again, of the CA_C2377 mutant and its plasmid-bearing derivatives. Consistent with what had been observed for solvent production, the CA_C2377::CTermB pMTL85143-*qsrB* strain, assimilated lower levels of acids compared to the wild type, with the supernatant containing 35 ± 4.1 mM of acetate and 6 ± 1.3 mM of butyrate (p -values = 0.011 and 2.87×10^{-5} respectively). Furthermore, addition of QspB7 had little effect and the

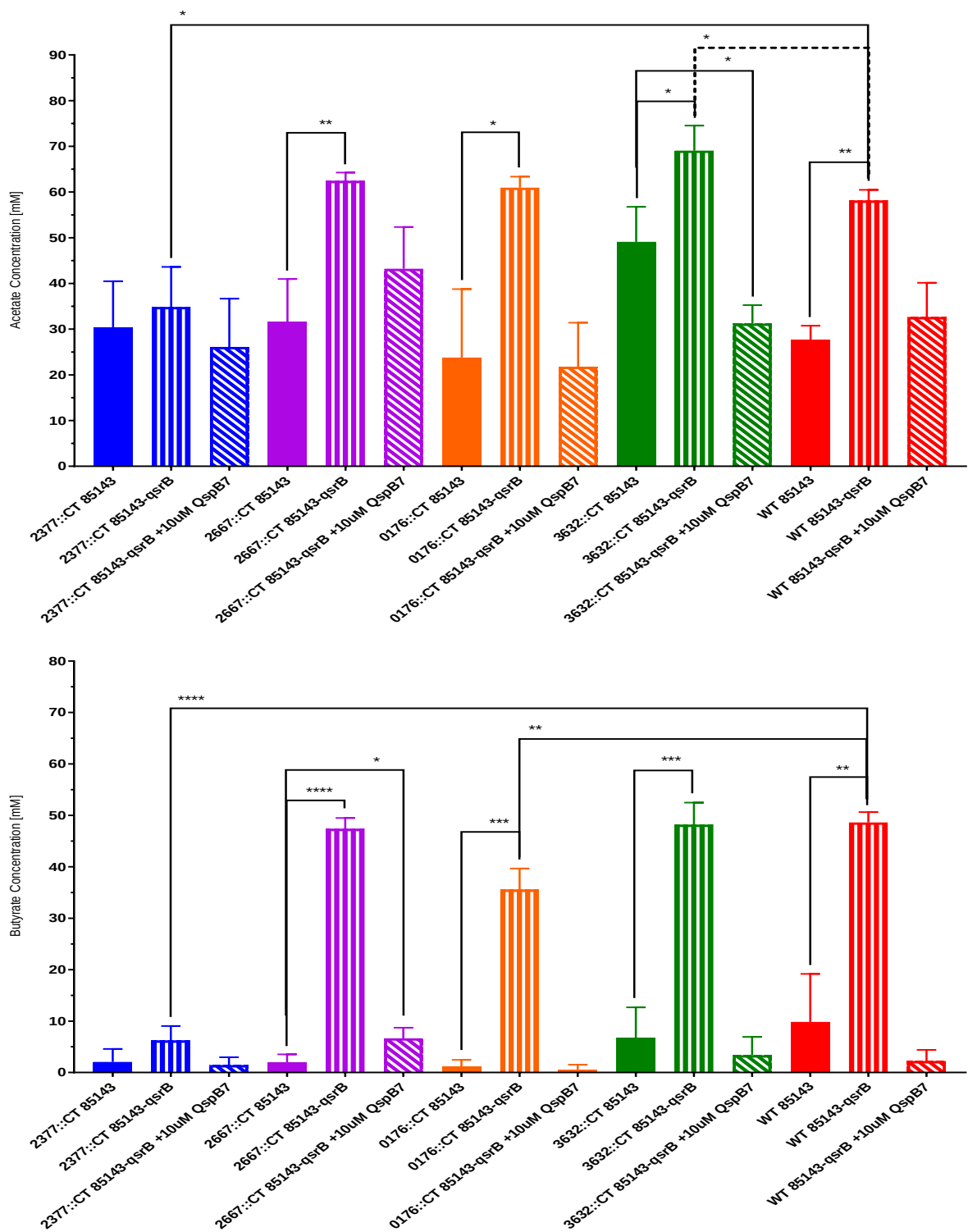


Figure 4.9 Acid of SBP mutants, containing the pMTL85143 control plasmid, the pMTL85143-*qsrB* plasmid or the pMTL85143-*qsrB* plasmid with 10 μ M QspB 7 addition. Each strain is indicated by colour: CA_C2377::CTermB (blue), CA_C2667::CTermB (purple), CA_C0176::CTermB (orange), CA_C3632::CTermB (green) and the *C. acetobutylicum* ATCC 824 WT strain (red). The colour filling of each bar indicates the experimental variable: pMTL85143 control (solid colour), pMTL85143-*qsrB* (vertical hashed lines) and pMTL85143-*qsrB* with 10 μ M QspB 7 addition. Each column represents the mean average of 3 replicates, with error bars representing SD of the mean. Connecting lines show a statistically significant t test result and asterisks indicate the resulting *p*-value (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

Small differences seen for butyrate were shown to be not statistically significant (p -value = 0.061). Interestingly then, the result seen for the level of heat resistant CFU/ml seemed to differ from this general trend, as shown in Figure 4.10. Whilst plasmid bearing mutant control strains produced less spores than the *C. acetobutylicum* ATCC 824 plasmid control, overexpression of *qsrB* did not result in a further reduction of spore counts, with the exception of the CA_C2667::CTermB pMTL85143-*qsrB* strain. Also, addition of synthetic QspB7 to *qsrB* overexpressing mutants had either no or little effect, a clear exception being the corresponding CA_C0176 derivative, which showed a 6 fold increase in final spore levels.

The observed reduction in mutant spore levels for the plasmid control strains was statistically significant for CA_C2667::CTermB pMTL85143 and CA_C3632::CTermB pMTL85143 (p -values of 0.0089 and 0.012, respectively), but not the CA_C0176::CTermB pMTL85143 and CA_C2377::CTermB pMTL85143. Also, the reduced (undetectable) level of heat resistant colony forming units for CA_C2667::CTermB pMTL85143-*qsrB* was statistically significant compared to the 2.12×10^3 ($\pm 5.39 \times 10^2$) produced by the corresponding vector control strain (p -value = 0.0011). Interestingly, for CA_C2667::CTermB pMTL85143-*qsrB* a weak response to 10 mM QspB7 was observed, but spore level remained significantly lower than those of the vector control (p -value = 0.0048). The response of CA_C0176::CTermB pMTL85143-*qsrB* to QspB7 was also found to be significant (p -value = 0.0026). Importantly, the lack of QsrB mediated repression of sporulation was also significant for the CA_C2377, CA_C0176, and CA_C3632 mutants with p -values ≤ 0.0001 .

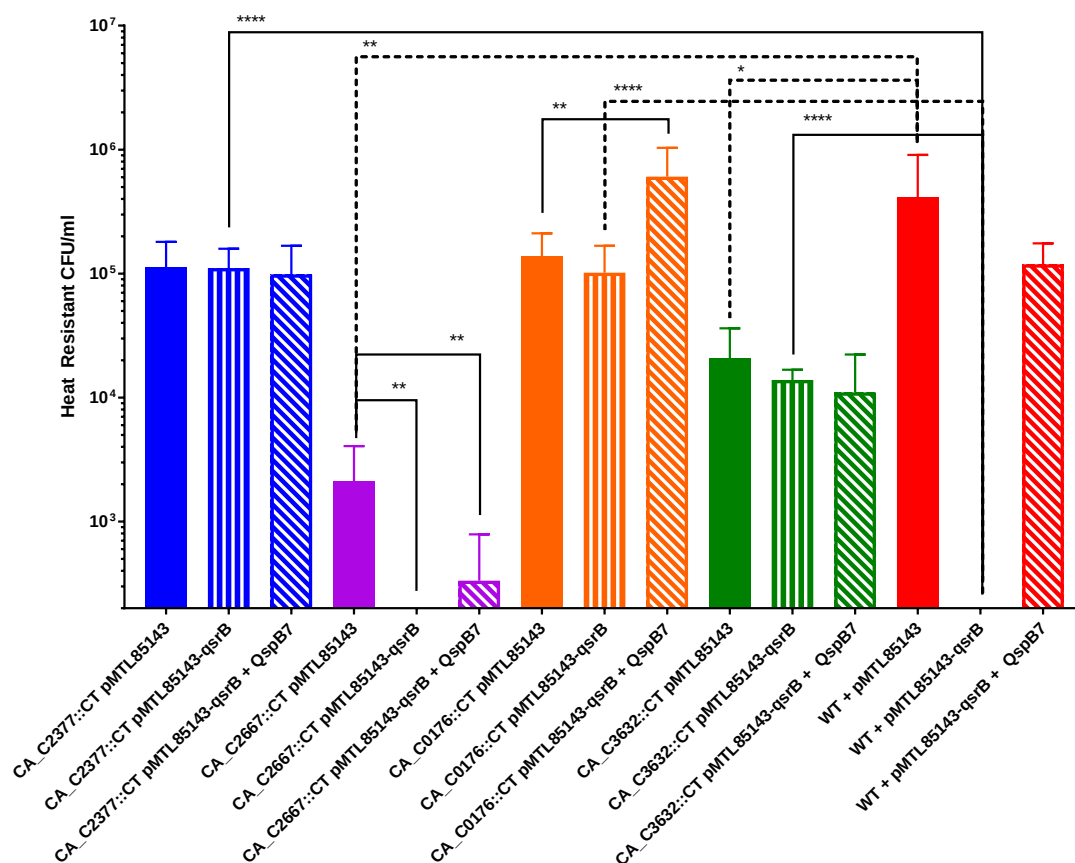


Figure 4.10 Formation of heat resistant CFU/ml for SBP mutants containing the pMTL85143 empty vector control, pMTL85143-*qsrB* and pMTL85143-*qsrB* with 10 μ M QspB 7 addition. Each strain is indicated by colour: CA_C2377::CTermB (blue), CA_C2667::CTermB (purple), CA_0176::CTermB (orange) CA_C3632::CTermB (green) and the *C. acetobutylicum* ATCC 824 WT strain (red). The filling pattern then denotes the experimental variable: pMTL85143 control (solid colour), pMTL85143-*qsrB* strain (vertical hashed lines) and the pMTL85143-*qsrB* strain with 10 μ M QspB 7 addition (diagonal hashed lines). Each bar represents the mean of 9 replicates, with error bars representing the standard deviation of the mean. The y axis begins at 200 heat resistant CFU/ml, in order to take into account the detection limits of the experiment. Statistical analysis was a two tailed t test with *p*-values represented by asterisks as before, the columns that underwent analysis are connected by a line, the line is dashed when it crosses another solid line (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

Over the course of the 120 h fermentation, the level of glucose was measured via high performance liquid chromatography, as described within section 2.6.6 of the materials and methods. This gave new insight, not only into the SBP ClosTron insertion mutants, but also into the physiological effects of *qsrB* overexpression in general.

Firstly, it would be necessary to understand the *C. acetobutylicum* ATCC 824 WT

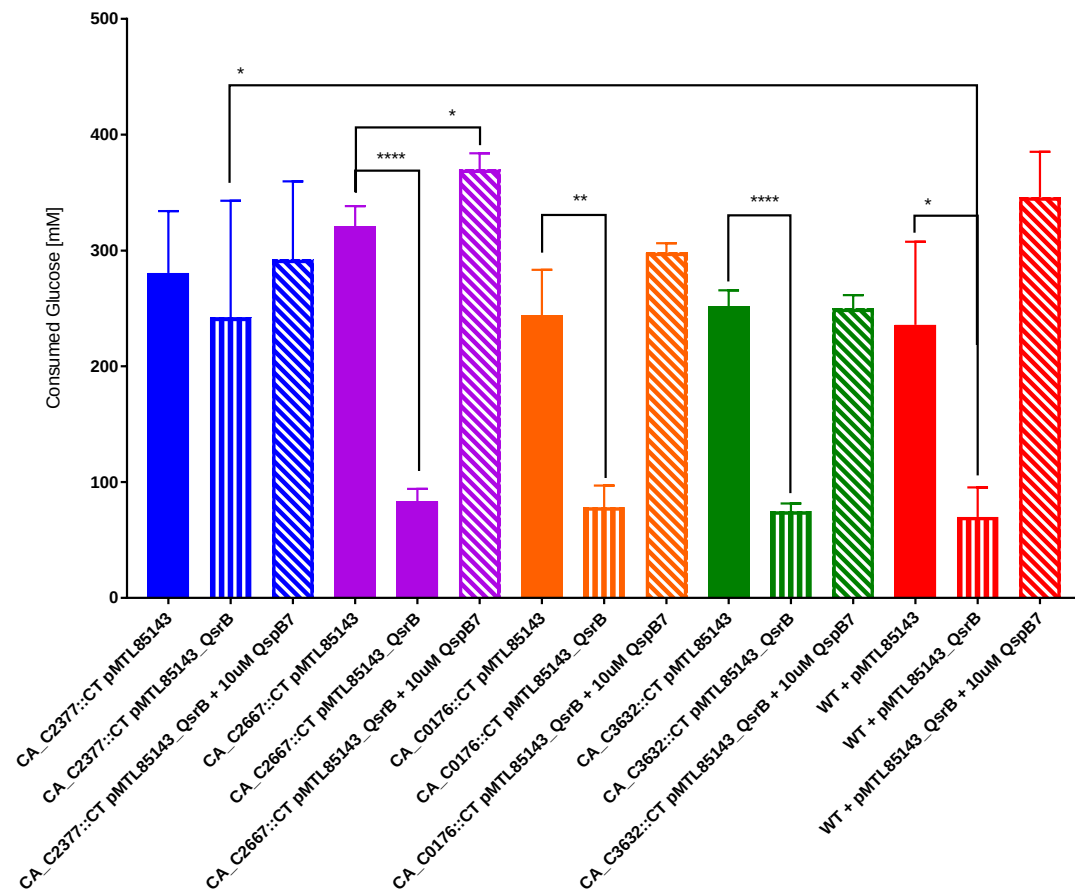


Figure 4.11 Glucose consumption by SBP mutants containing empty pMTL85143 vector, pMTL85143-*qsrB*, and pMTL85143-*qsrB* with 10 μ M QspB7 addition. The above data was generated by deducting the final glucose concentration for each replicate from its initial glucose concentration at initiation of fermentation, and giving a mean average across all replicates represented by each column. The error bars represent the standard deviation from this mean. Each insertional mutant is represented by a colour: CA_C2377::CTermB (blue), CA_C2667::CTermB (purple), CA_C0176::CTermB (orange), CA_C3632::CTermB (green) and the *C. acetobutylicum* ATCC 824 WT strain (red). The filling pattern then denotes the experimental variable: pMTL85143 (solid colour), pMTL85143-*qsrB* (vertical hashed lines) or the pMTL85143-*qsrB* with 10 μ M QspB addition (angled hashed lines). Connecting lines between columns represent a significant result from a two-tailed t test, with asterisks indicating the resulting *p*-value (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

strain glucose consumption data show in Figure 4.11. The *C. acetobutylicum* ATCC 824 pMTL85143 control strain consumed 236 ± 34 mM of glucose after 120 hours of fermentation, while the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strain consumed only 70 ± 12 mM of glucose (p -value = 0.02). When the latter strain was supplemented with 10 μ M of QspB, its consumption of glucose dramatically increased to 346 ± 19 mM, however, this was not sufficient to differ significantly to the WT strain containing the pMTL85143 control plasmid (p -value = 0.08).

The remaining data shown in Figure 4.11 gives further novel insight into the effect of SBP mutation. As expected from the solvent data, the only *qsrB*-overexpressing strain to consume statistically more glucose than the WT pMTL85143-*qsrB* control was the CA_C2377::CTermB pMTL85143-*qsrB* strain. This strain consumed 242 ± 48 mM of glucose (p -value = 0.046). As expected, this level did not differ significantly from the corresponding pMTL85143 plasmid control (p -value = 0.598). Each of the other insertional mutation strains transformed with pMTL85143-*qsrB* did consume significantly less glucose than their own pMTL85143 empty vector control, which is shown in Figure 4.11.

As pointed out in Chapter 3 (section 3.3.2), supplementation of pMTL85143-*qsrB* bearing strains with 10 μ M of QspB 7 often produced increased solvents in previous experiments. Only one strain showed a statistical difference in this respect, which was the CA_C2667CTermB strain. The CA_C2667CTermB pMTL85143 strain consumed 321 ± 8 mM of glucose, and the CA_C2667CTermB pMTL85143-*qsrB* strain consumed 370 ± 7 mM (p -value = 0.02).

Due to the deviating solvent profile of the CA_C2377::CTermB mutant (Figure 4.8), it was subject to further scrutiny. The CA_C2377::CTermB mutant strains were restreaked to single colonies twice to confirm erythromycin resistant WT contamination had not taken place. Following 120 hours of fermentation, the acid and solvent profiles were re-examined. As can be seen in Figure 4.12 the newly created CA_C2377::CTermB pMTL85143-*qsrB* strain now displayed reduced solvent formation.

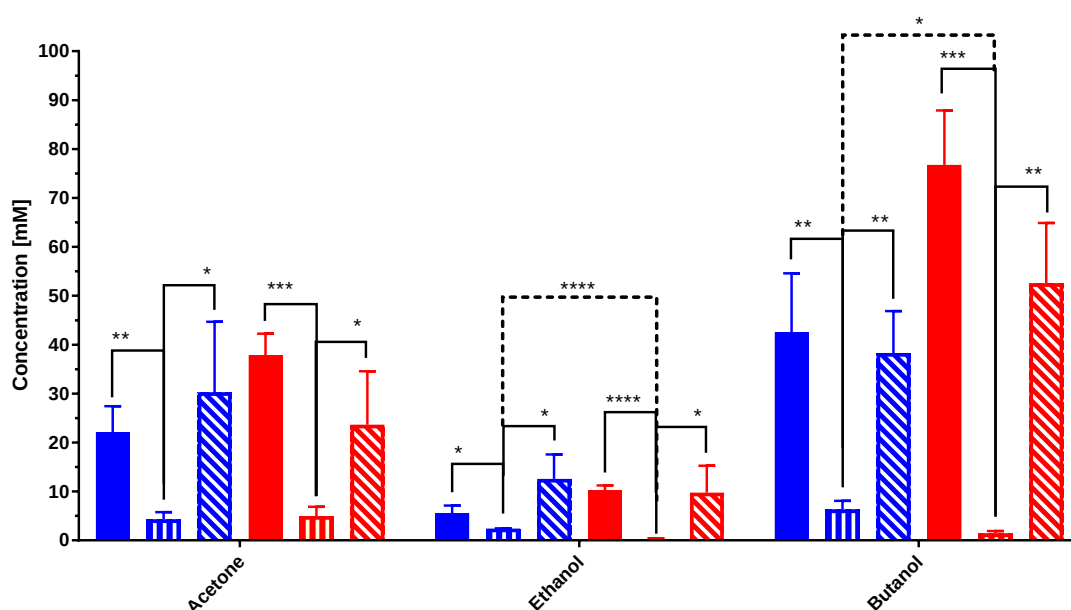


Figure 4.12 Product profile of the reisolated CA_C2377::CTermB mutant strains containing; the pMTL85143 empty vector control, the pMTL85143-*qsrB* vector and the pMTL85143-*qsrB* vector with 10 µM QspB 7 addition. The coloured bars represent the strain: CA_C2377::CTermB (blue) and *C. acetobutylicum* ATCC 824 WT (red). The column filing pattern indicates the experimental conditions; pMTL85143 empty vector control (solid colour), pMTL85143-*qsrB* (vertical hashed lines) and pMTL85143-*qsrB* with 10 µM QspB 7 addition (angled hashed lines). The columns represent the mean of 3 replicates with error bars representing the standard deviation of the mean. Joining lines between columns represent a statistically significant result from a two tailed t test, and are solid or hashed for ease of viewing, the resulting *p*-values are represented by asterisks (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

4.4 Discussion

The ultimate aim of the work described in this chapter was to identify the oligopermease system(s) required for QspB uptake. This required their bioinformatics identification, specifically of the respective substrate binding proteins, mutation of the encoding genes, and characterisation of the generated mutants in terms of their ability to alleviate the effects of QsrB-mediated repression in the presence of synthetic QspB peptide.

When investigating the *C. acetobutylicum* ATCC 824 genome for genes encoding SBPs capable of binding oligopeptides, four of them (CA_C0176, CA_C3632, CA_C3179 and CA_3634) were contained within a predicted operon with the other constituent parts of a putative ABC transporter. As suggested by the literature, it is likely these loci represent fully functioning transporters in their own right (Beek et al.,

2014; LeDeaux et al., 1997; Rudner et al., 1991). An additional SBP encoding gene (CA_C3640) was positioned alongside a predicted operon that contained all the other constituent parts of an ABC transporter, two integral membrane proteins and two ATP binding subunits. It is therefore likely that, despite being coded in different predicted operons, these genes also encode a fully functional transport system as the operon structure is not universally conserved for bacterial ABC transporter systems (Tomii and Kanehisa, 1998). Interestingly, the CA_C0176 gene was predicted to be transcribed together with the *murQ* gene (Dam et al., 2007). The *murQ* gene is essential in the process of peptidoglycan recycling, shown to be crucial for survival in multiple Gram-positive organisms (Borisova et al., 2016).

While this in itself does not give much of a clue as to the function of the CA_C0176 encoded SBP and its associated ABC complex, it hints that at least the gene is transcribed actively during the stationary phase. As *C. acetobutylicum* ATCC 824 will be at its highest cell density during the stationary phase, CA_C0176 cannot be ruled out as participating in the uptake of cell-cell communication peptides. The before mentioned oligopeptide permease systems would give *C. acetobutylicum* ATCC 824 an impressively large number, at least 5 compared to the 3 of *B. subtilis* (Levdikov et al., 2005). While currently unsupported, it could be possible that the higher number of oligopeptide permease systems within the *C. acetobutylicum* ATCC 824 genome is reflective of the high number of RNPP type cell-cell communication systems it employs (Kotte, 2013; Kotte et al., 2017). This is an attractive hypothesis given that the organism is saccharolytic and not known to be able to use peptides as the sole sources of carbon and energy, although of course they may be utilised to obtain amino acids for incorporation into proteins.

Two other putative SBP encoding genes were identified that had some interesting traits. The CA_C2377 gene contained a frameshift mutation, leading to only one third of the gene, 567 bp, to be translated. Effectively eliminating 1,167 bp of coding sequence, it is not known what the downstream effect of this frame shift will be, and if the CA_C2377 gene retains any function at all. Interestingly, two components of an ABC transporter are encoded downstream of the CA_C2377/CA_C2376 locus, and are predicted to be transcribed within their own operon. It is possible that these two components represent a broken ABC transporter, i.e. an artefact left over from an

earlier mutation, an apparently not uncommon occurrence (personal correspondence with Dr. Gavin Thomas, University of York). The CA_C2667 gene encoded for a putative peptide binding SBP that was not located with any components of an ABC cassette, and was predicted to be within its own operon, and was therefore characterised as an orphan SBP (Thomas, 2010). It is possible that this orphan SBP interacts with any of the other ABC transporters, presumably with different specificity to the other, native SBPs (Maqbool et al., 2015; Thomas, 2010; Wang et al., 2004).

Of the aforementioned SBP genes, four could be obtained as ClosTron insertional mutations; CA_C0179, CA_C3632, CA_C2377 and CA_C2667. The insertional mutation of CA_C0176 and CA_C3632 was exciting, given that they were part of a complete ABC transporter gene cluster and likely candidates to encode their cognate SBPs. Furthermore, it is interesting that two of the unobtainable ClosTron insertional mutations were those of CA_C3179 and CA_C3634, both of which were predicted to be within the same operon as the remaining components of an ABC cassette (Dam et al., 2007). As these genes are predicted within the operon of other ABC cassette components, it is more likely they form a fully functional oligopermease system (Thomas, 2010). Therefore, inability to target them with ClosTron could indicate their essentiality within the genome. Whilst it was true that CA_C3640 could not be inactivated via ClosTron either, the sequence shares an 89% identity to CA_C3634. For this reason, it was difficult to design a ClosTron that specifically targeted only one of the two genes, as the mobile group II intron would be likely to integrate into either. Therefore, the inability to obtain a CA_C3634 insertional mutant could be due to the potential lethality of a CA_C3640 insertional mutation, or theoretically vice versa.

The first set of phenotypes to be observed were the accumulation of granulose and amylase activity of the ClosTron insertional mutants. These easily screenable phenotypes are often associated with degeneration in *C. acetobutylicum* ATCC 824, as one of the two amylase genes resides within the megaplasmid of the organism which in degenerate strains is lost (Cornillot et al., 1997) and the other is controlled by the key regulator Spo0A (Jang et al., 2014). Although both CA_C2667::CTermB strains isolated here had lost granulose formation and amylase activity, they were not considered degenerates given that they still produced solvents and spores (albeit the latter at a reduced level) and the same phenotype had been obtained for independent

isolates. The CA_C3632::CTermB strain whilst granulose negative, had retained amylase function. Plasmid complementation was attempted in both of these strains. For both strains a plasmid was constructed that contained the gene of interest with 250 bp of upstream region in an attempt to include the native promoter. These plasmids were successfully transformed to the respective ClosTron insertion strains, however, both failed to show complementation (data not shown). There are numerous reasons for lack of complementation, the simplest being that the native promoter was not sufficiently present in the 250 bp region used. Thus, complementation should be repeated with a strong promoter, such as the *C. sporogenes* ferredoxin promoter used in the pMTL85143 plasmid, to eliminate this possibility. Another possibility is the insertion of a second copy of the mobile group II intron into another site in the genome. However, double insertions are rare events, and it is unlikely a double insertion mutant like this would have occurred in each of the independently obtained mutants (Cooksley et al., 2012).

Nevertheless, the mutant strains were sent for whole genome sequencing to eliminate this possibility (awaiting data). Given that the generated mutants were generated by ClosTron insertion, it is also possible that the observed phenotypes were caused by polar effects on the flanking genes, an issue that might be addressed through the more tedious introduction of clean in-frame deletions (Heap et al., 2010, 2007). These effects can also be investigated via the introduction of plasmid, or genomic, complementation. Reversion back to a WT phenotype upon complementation would suggest the altered phenotype was induced via the ClosTron mutation, and not polar effects. However, it should be kept in mind that the overall aim of the undertaken large scale peptide permease inactivation programme was not the detailed characterisation of the resulting mutant phenotypes and possible associated artefacts. Rather, it was meant to be a fast screen for the identification of permease mutants no longer responding to QspB peptide addition, then to be followed by a more detailed investigation. Thus, it was sufficient to establish that all strains had retained the ability to form solvents and spores, as these were markers for the *qsrB* overexpression test to be employed.

Indeed, each of the insertional mutant strains showed a reduced solvent profile relative to a WT control after 48 and 120 hours of fermentation and was able to produce

spores, suggesting that the results seen were not the result of degeneration. As mentioned above, both of the isolated CA_C2667::CTermB strains produced significantly reduced spores relative to the WT control, the only insertional mutation to do so consistently across all independent isolations.

The SBP required for the uptake of QspB could be determined by transforming the insertional mutations with the pMTL85143-*qsrB* plasmid, and showing that the QsrB overproduction phenotype is maintained in each of the insertional mutations. Following this each of the pMTL85143-*qsrB* containing mutants could be dosed with 10 μ M QspB 7, and the responses relative to solvent and spore formation measured. If an insertional mutation, containing the pMTL85143-*qsrB* plasmid, could not be rescued to solvent and spore formation by the addition of QspB 7, it must be responsible for its uptake.

Eventually, it was concluded that each of the SBP insertional mutations that contained pMTL85143-*qsrB* had a reduced solvent profile compared to its own vector control, which was to be expected from previous results (Kotte et al., 2017). The CA_C2377::CTermB mutation produced results concluded to be anomalous, and was subsequently shown to conform to the above. The CA_C0176::CTermB pMTL85143-*qsrB* strain performed better than the others, with statistically increased acetone and butanol relative to the WT pmtl85143-*qsrB* control (Figure 4.8). However, it was also seen that upon addition of 10 μ M of QspB 7 to the fermentation, the formation of solvents could be restored in each of the SBP insertional mutations harbouring the pMTL85143-*qsrB* plasmid. Taking this result in isolation, this could mean one of several things. Firstly, it may be that none of the SBPs that were successfully mutated via the ClosTron technique are the SBP required for QspB uptake, and hence QspB could be taken up via CA_C3179 and its operon, or the very similar CA_C3634 and CA_C3640 genes coding for SBPs. This assumes that the reason that *C. acetobutylicum* ATCC 824 contains multiple oligopeptide SBP is because they each possess a different ligand preference, and hence are required for different peptide recognition or in different conditions (Doeven et al., 2005; Wang et al., 2004). Secondly, it could show that the SBPs are more likely general uptake, and that QspB can be bound by multiple SBPs and hence taken up through differing ABC cassettes (Sleigh et al., 1999). Either way, more empirical information would be required.

Perhaps a more streamlined approach would be to target an integral membrane protein from each oligopeptide permease system for either KO through inframe deletion, ClosTron insertional mutation, or even the newly adapted CRSPR technique (Bruder et al., 2016; Xue et al., 2017). A disruptive mutation in an integral membrane protein was found to render the entire oligopeptide ABC transport system ineffective in the Gram-positive model *Bacillus subtilis* (LeDeaux et al., 1997). However, failure to create ClosTron mutations of the CA_C3179 and CA_C3634, may indicate that the gene and by extension the operon they reside in, are essential for the organisms survival in the growth conditions. Despite its industrial importance, and the availability of a universal *mariner* transposon, no study has yet been conducted showing the essential genome of *C. acetobutylicum* ATCC 824, and hence it is impossible to predict the essentiality of such genes (Zhang et al., 2015). A more time intensive, but perhaps more conclusive analysis of the SBP would involve the tagging of the SBP followed by purification, as has been done before in other organisms (Levdikov et al., 2005). It would then be possible to perform a ligand-fishing experiment, by dosing the purified protein with QspB peptides, washing, and identifying those that have successfully bound via HPLC/MS (Gogliettino et al., 2010). This has the benefit of identifying multiple SBPs that can bind QspB, in the event that the selectivity of the SBPs overlap, and could also be used to identify the SBPs of other RNPP type systems within the organism by dosing with synthetic peptides QspA through to QspH.

Something interesting did occur with respect to sporulation in the SBP insertional mutations transformed with pMTL85143-qsrB. The CA_C0176::CTermB pMTL85143-qsrB and CA_C3632::CTermB pMTL85143-qsrB strains formed spores of levels comparable to their own WT control. Previous studies have only linked loss of an oligopeptide permease with sporulation defects (Perego et al., 1991; Rudner et al., 1991). There is also little in the wider literature to suggest why loss of an oligopeptide permease may lead to recalcitrance towards what is believed to be a transcriptional repressor (Kotte et al., 2017). It would be a first hypothesis that the named SBP are required for export of the QspB peptide, therefore loss would lead to an increased cytosolic level of QspB and hence reduce the activity of QsrB enough to initiate the sporulation cascade, similar to the marginal increase in sporulation seen when overexpressing QspB in previous studies (Kotte et al., 2017). However, this is

unlikely given that in other organisms RNPP type signalling peptides are exported via SecA dependant transporters, and that oligopeptide permease ABC transporters are very rarely seen as exporters in prokaryotes (Beek et al., 2014; Locher, 2004; Rocha-Estrada et al., 2010; Tomii and Kanehisa, 1998). Another possible suggestion is that these insertional mutations are hindering another uptake of another signalling peptide, to which the conjugate regulator is initiating a sporulation response that supersedes QsrB. There are two orphan regulators within the *C. acetobutylicum* ATCC 824 genome that have been shown to be essential for solventogenesis and sporulation, CA_C0957/CA_C0958 (Kotte, Bean, Winzer, unpublished data). Overexpression of the CA_C0957 gene in a Δ CA_C0957 background, caused an increase in spore formation over the WT control level, whilst increasing solvents to be indistinguishable from a WT control (Bean, 2016). It could be then that lack of a conjugate peptide for this system is causing a sporulation signal, activating the sporulation cascade and superseding QsrB regulation. Whilst simultaneously, the repression of solvent formation by QsrB is performed in such a way to supersede CA_C0957 regulation. An interesting notion but lacking substantiation.

Perhaps the most interesting data to be obtained during the course of this section of inquiry was that of the glucose levels both before and after fermentation shown in Figure 4.11. The data here conclusively shows that overproduction of QsrB, causes a reduction in the amount of glucose consumed after 120 hours of fermentation. This could possibly suggest that the reduction in solvents seen in strains containing the pMTL85143-*qsrB* is not necessarily due to a repression of solvent production, but could be due to a reduction in metabolised carbon source. The higher levels of glucose consumption seen, for example *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* supplemented with QspB 7, represents glucose limitation. It was decided proteomics data could be used to elucidate further the effects of QsrB overproduction, and will form the basis of the next chapter.

To conclude, it was set out in this chapter to identify the possible SBPs required for the uptake of QspB, and other RNPP type signalling peptides of *C. acetobutylicum* ATCC 824. Other than a possible hypothesis relating to the conjugate peptide for the CA_C0957, it cannot be said that this was achieved. This may show that the QspB peptide is taken up through the systems for which a ClosTron insertional mutation

could not be obtained, CA_C3179, CA_C3634 and CA_C3640. Or it could indicate that the QspB peptide is taken up as general uptake by any one of the oligopermease ABC transporters (Sleigh et al., 1999).

However, four insertional mutants of SBP in *C. acetobutylicum* ATCC 824 were obtained as independent replicates. To our knowledge this is the first time such mutations in this organism have been conducted. Two were shown to be granulose negative (CA_C2667::CTermB and CA_C3632::CTermB) and of these CA_C2667::CTermB was shown to be amylase negative also. Though there were complications with the CA_C2377::CTermB strain, eventually it was shown that each of the SBP insertional mutants had significantly reduced solvents and spores when transformed with the pMTL85143-*qsrB* plasmid, and that solvent formation could be revived by the addition of 10 μ M QspB 7. Interestingly, it was seen that the CA_C0176::CTermB pMTL85143-*qsrB* and CA_C3632::CTermB pMTL85143-*qsrB* strain still produced spores, despite retaining the vastly reduced solvent phenotype of QsrB overproduction. Finally, glucose consumption data generated through use of the HPLC indicated that QsrB overproduction lead to decreased glucose uptake, indicating that reduced spore and solvent phenotypes are likely due to restricted glucose usage as opposed to direct repression. The insertional mutations of the oligopeptide SBPs conferred this reduced glucose consumption phenotype, with the exception of CA_C2377::CTermB, which through further experimentation was shown to conform to the phenotype of the other insertional mutations.

**5. Effect of qsrB
overexpression on the
C. acetobutylicum
ATCC 824 proteome**

5 Effect of *qsrB* overexpression on the *C. acetobutylicum* ATCC 824 proteome

5.1 Introduction

C. acetobutylicum ATCC 824 has a complex metabolism. The production of the acids acetate and butyrate, and their subsequent reuptake in order to produce the conjugate solvents acetone and butanol (Jones and Woods, 1986). This transition from acid production to solvent production is often termed the acid solvent switch, and conclusive theory is yet to be produced on what governs it (Sauer and Durre, 1995; Strickmann et al., 2002). It was traditionally seen that the initiation of stationary phase in batch culture lead to the production of solvents, and hence for a while these two were often linked (Jones and Woods, 1986). Conflicting evidence has been given as to whether or not solventogenic metabolism supports growth of the organism at all (Alsaker et al., 2005; Grimmmler et al., 2011). However, for decades the organism has been known to grow and produce solvents when maintained in continuous culture (Bahl et al., 1983). Furthermore, investigation of the acid solvent switch has been conducted on the system and modelling level, which has introduced some interesting challenges to these views (Kumar et al., 2014, 2013). Kumar and colleagues (2014) suggested through modelling that growth can be supported during solventogenesis but only in response to external stress, which they define as a stress not produced by the organism's own acid production, an idea which has been contested by empirical data (Millat and Winzer, 2017).

Many studies have been conducted to more clearly define changes in gene expression and protein composition when cultures make the transition from the acid to solvent production (Bao et al., 2014; Janssen et al., 2010). Understanding the complex nature of the organisms metabolism requires such approaches, as industrially the feedstock available for solvent production may be inherently variable and hence differ strongly from the uniform conditions of the laboratory (Abd-Alla et al., 2015; Alonso et al., 2010; Menon and Rao, 2012; Naik et al., 2010). Use of various proteomic techniques has shed light on intracellular and membrane protein composition of *C. acetobutylicum* ATCC 824 under a range of different conditions (Chen et al., 2014; Janssen et al., 2010; Raut et al., 2016). Proteomic techniques are also increasingly

applied to study adaptation and response of *C. acetobutylicum* ATCC 824 metabolism under different fermentation conditions, in a line of enquiry that will be essential if ABE fermentation is to return to industrial viability (Raut et al., 2016). For instance, Raut and colleagues (2016) investigated through the use of comparative and quantitative proteomics how lignin, a component of lignocellulose treatment, lead to decreased solvent yields in *C. acetobutylicum* ATCC 824. It was therefore concluded that the study of QsrB overproduction and its effects on solvent production and spore formation might benefit from a similar approach (Kotte et al., 2017; Raut et al., 2016).

As for all -omics approaches, the results obtained in proteomic analyses need to be independently verified to ensure that the vast amount of quantitative data generated are reliable and reflect de facto changes rather than artefacts. Often, a representative set of proteins is chosen and their levels under the different employed conditions quantified by different means, for instance Western blot (if specific antibodies exist) or by measuring the biological (e.g. enzymatic) activity.

Enzyme assays have also been used independently of proteomic approaches to extensively characterise and assess physiology and metabolism of *C. acetobutylicum* ATCC 824 in different phases of growth (Andersch et al., 1983; Dürre et al., 1987; Yoo et al., 2015). The nature of *C. acetobutylicum* ATCC 824 being an obligate anaerobe means some of the enzymes essential for its metabolism are poisoned by oxygen, and for this reason can occasionally give negative results when using standard published assays (Dürre et al., 1987). The acidogenic phase of growth lends itself greatly to the assaying of those enzymes involved in the process, such as the acetate and butyrate kinases (Winzer et al., 1997). Studies into the enzymes required for butanol formation led to the discovery that *C. acetobutylicum* ATCC 824 possesses several alcohol and aldehyde dehydrogenases. It is now generally accepted that the organism uses them to various degrees at different stages of growth depending on conditions and some of them (AdhE1 and AdhE2) for the production of both ethanol and butanol (Yoo et al., 2015). Hence, their activity in protein extracts can serve as a measure for the extent of solventogenesis.

Another method for the corroboration of a proteomics data set would be the use of a reporter system in order to indicate the relative activity of key genes. Fluorescent

reporters, while common in other organisms, have been difficult to develop in *C. acetobutylicum* ATCC 824. Two reasons for this are the reliance of well-developed reporters such as green fluorescent protein (GFP), on oxygen and the intrinsic autofluorescence of *C. acetobutylicum* ATCC 824 due to its high riboflavin content (Tummala et al., 1999; Willson et al., 2016). The low G-C content of the *C. acetobutylicum* ATCC 824 genome can also create difficulties for commercially available plasmids for constructing fluorescent fusion proteins. There have been successes with anaerobic reporters, although fluorescence in *Clostridium spp.* was usually very close to the background level, perhaps one of the reasons they did not have the impact anticipated when they were first described 10 years ago (Drepper et al., 2007). Several reporter systems have been adapted for this organism and used for the generation of transcriptional and translational fusions, and a number of which will be summarized here.

The *lacZ* gene from *Thermoanaerobacterium thermosulfurigenes* has been used with varying success. It possesses a GC content similar to that of the *C. acetobutylicum* ATCC 824 host, and the well documented nature of the β -galactosidase assay makes it a good choice for a reporter gene (Feustel et al., 2004; Tummala et al., 1999). Luciferase from the fire fly *Photinus pyralis* has also been utilised as a reporter in *C. acetobutylicum* ATCC 824 (Feustel et al., 2004). An advantage of this system is that it is not subject to background activity, whereas the *lacZ* system would require checking for background activity on a strain to strain basis (Feustel et al., 2004). The *E. coli* β -glucuronidase gene has also been used in reporters in a number of *Clostridium spp.* including *C. acetobutylicum* ATCC 824 (Girbal et al., 2003). *GusA* was shown to be functional and non-toxic in the organism, and native β -glucuronidase activity was low enough not to cause significant background activity (Girbal et al., 2003). During the course of this study, a *gusA* reporter was developed using the gene from *Streptococcus agalactiae* (Bean et al., data not published).

A number of insights can be gained through the tagging and purification of an RNPP-type regulator. Such approaches have been the basis of multiple studies (Do and Kumaraswami, 2016; Huillet et al., 2012). Protein purification has been key to characterising and discovering the mode of action of a number of regulators allowing

deepening understanding of conformational changes brought on by peptide binding (Bouillaut et al., 2008; Perchat et al., 2016a).

It was decided that proteomic data sets might give much needed insight into the action of QssB, in order to fill the current gaps in understanding the system. This data set, obtained for the QsrB overexpressing strain and relevant control, could be corroborated by the use of enzyme assays, and further investigated by the use of a reporter system. In order to fully understand the action of QsrB the protein should also be tagged and purified for further experimentation. This line of enquiry led to the following objectives for this part of the study.

5.2 Objectives

In order to develop a more encompassing understanding of QssB, it was decided to perform a proteomics experiment. The generation of a large data set necessitates the need for further empirical investigation in order to confirm its findings, and as such the following more specific objectives were outlined;

- To generate proteomics data for the strains *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB*, *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* with 10 μ M QspB7 addition and the *C. acetobutylicum* ATCC 824 pMTL85143 'empty vector' control.
- To use this data to better understand the phenotypic effects of QsrB overproduction, by analysing the relative abundances of proteins known to have key effects in the *C. acetobutylicum* ATCC 824 life cycle.
- To empirically investigate the proteomics data, by choosing a number of enzymes that differ significantly in their relative abundance, and assaying their activity. Also, to attempt an assay on any other phenotypic change that may be implied by the proteomics data.
- To utilise a GusA reporter system, in order to generate promoter activity data in a *C. acetobutylicum* ATCC 824 CA_C0324::CTermB (i.e. *qsrB* mutant) background for a number of key genes, including *qsrB* itself, shown to be important either from previous studies or the proteomics data.
- To be able to tag and purify the QsrB regulator protein, in order to create a starting point for protein crystallisation and other downstream experiments.

5.3 Results

5.3.1 Generation of proteomics data

The effects of *qsrB* overexpression in the presence and absence of synthetic QspB 7 peptide were to be investigated to identify proteins directly and indirectly dependent on the QssB system. To generate cells for subsequent iTRAQ-based proteomic analysis, *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB*, *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* with 10 μ M QspB7 and the *C. acetobutylicum* ATCC 824 pMTL85143 control were grown in CBM-S medium in shaking flasks for 24 hours. At this time point the whole cell culture was harvested, and processed as per 2.6.7 (Ow et al., 2009; Raut et al., 2016). As the cell cultures were taken and processed to be sent for proteomics analysis, a 1 ml supernatant sample was also taken of each for gas chromatography analysis (2.6.5), the result of which is shown in Figure 5.1. This time point was chosen as previous experiments had established that the applied conditions led to the start of statistically significant differences between solvents. A large number of changes were expected due to the drastic shift caused by overexpression, it was reasoned that an early time point would reduce the number of proteins whose levels would change indirectly as a consequence of QsrB mediated effects.

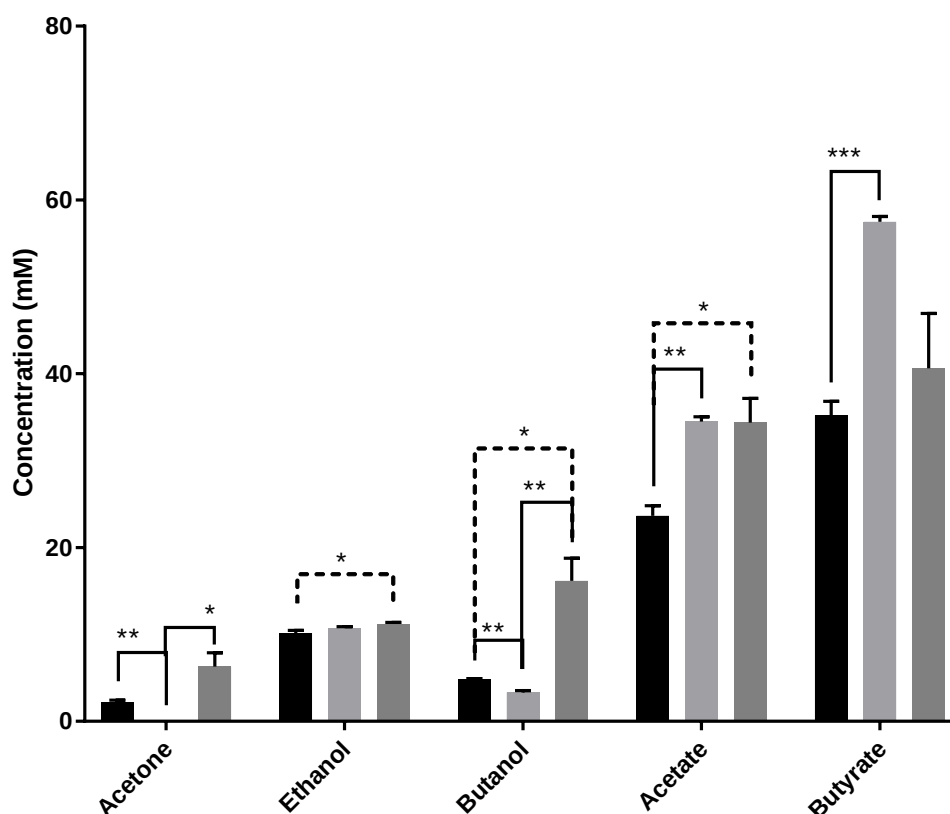


Figure 5.1 Fermentation products of *C. acetobutylicum* ATCC 824 pMTL85143 (black), *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* (light grey) and *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* with 10 µM QspB 7 (dark grey) at the 24 hour time point of preparing samples for proteomic analysis. Each column represents the average of 3 replicates, and error bars show the standard deviation of the mean. Connecting bars indicate a statistically significant result from a two-tailed T test, and can be solid or dashed lines for visual clarity. The *p*-value is represented by asterisks (*p*-value ≤ 0.05 is *, *p*-value ≤ 0.01 is **, *p*-value ≤ 0.001 is ***)

The trend of the data shown in Figure 5.1 was as expected for this time point for all strains (see Figure 3.1 and Figure 3.7 in section 3.3.2 and 3.3.3, respectively for comparison): For the *qsrB* overexpressing strain, butanol and acetone levels were reduced and absent, respectively, whereas acid levels were increased in comparison to the empty vector control, with differences being statistically significant. As seen before, the effects of *qsrB* overexpression could be overcome by the addition of synthetic QspB7 peptide. Thus, the obtained protein samples were deemed fit for further proteomic analysis.

The extent of changes observed for individual proteins ranged from 2.11-fold upregulation to 2.54-fold down regulation when comparing the *qsrB* overexpressing strain with the plasmid control. Given the clear effects of *qsrB* overexpression on

phenotypes such as solvent formation and spore generation, this range was comparatively narrow. iTRAQ based analyses often tend to underestimate the fold changes occurring (Ow et al., 2009). Hence when considering the data presented here, it would seem appropriate to place emphasis on the observed direction of change rather than the actual fold value determined

Proteomic iTRAQ data was produced by Dr. Raut via collaboration with The University of Sheffield and Prof Phillip Wright (now at Newcastle University). The complete set of proteomics data is included in section Appendix Table 1- 3. Three comparisons were made; Firstly, *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* was compared against a *C. acetobutylicum* ATCC 824 pMTL85143 control to investigate the proteomic changes induced by QsrB overproduction. Secondly, *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* with 10 μ M of QspB 7 addition against the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB*, to indicate if the supplementation with conjugate peptide leads to reversion of the phenotype at a proteomic level. Finally *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* with 10 μ M of QspB 7 addition against the *C. acetobutylicum* ATCC 824 pMTL85143 control, to compare QsrB overproduction that is supplemented with its conjugate peptide QspB 7, to the WT strain.

Following protein digestion to a peptide library, iTRAQ analysis was conducted. When comparing the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* with the *C. acetobutylicum* ATCC 824 pMTL85143 control strain, 233 proteins were found to differ significantly, which were represented by a total of 3787 individual peptides tags. Similarly, for the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strain grown with 10 μ M QspB 7, 197 proteins were detected that differed significantly from the QsrB overproduction strain, and of these 3,145 individual peptide tags were detected. Finally, in the data set comparing the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strain with 10 μ M QspB 7 and the empty vector control strain, 59 proteins were detected that differed significantly, and of these 4921 peptides were detected. The entire data set obtained can be seen in Appendix Tables 1- 3.

5.3.1.1 Detection of QsrB within the proteomic data

As an interesting first note, the QsrB protein was detected within the proteomics data set. A 1.25 fold increase was seen in the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strain relative to the pMTL85143 control strain. This level appeared to be more upon addition of the QspB 7 conjugate peptide, with the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strain with 10 μ M QspB 7 addition showing a 2.14 fold increase over the before mentioned control. It therefore follows that the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strain with 10 μ M QspB 7 addition strain produced a 1.75 fold increase over the overproduction strain with no conjugate peptide addition.

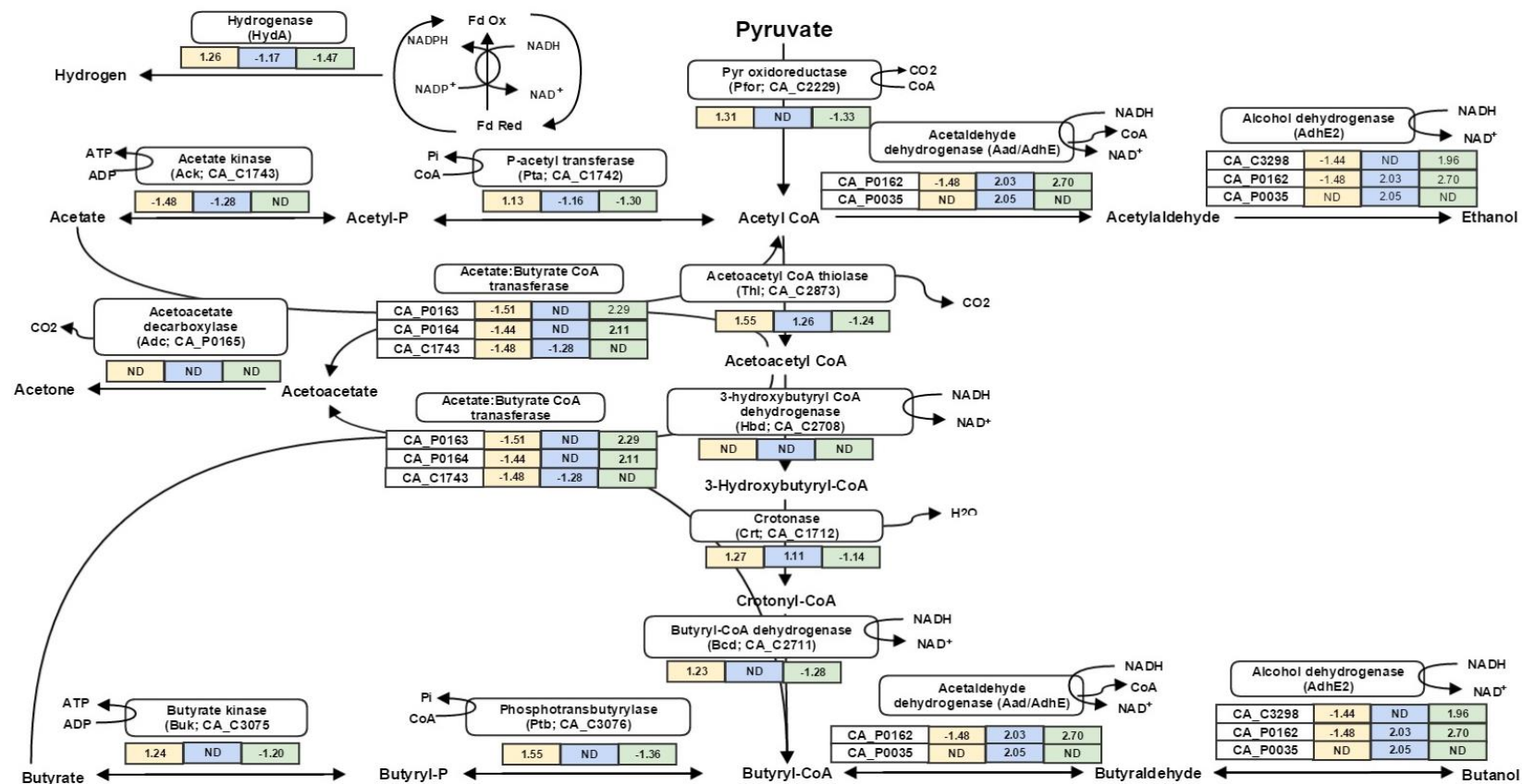


Figure 5.2 Changes in ABE fermentation pathway enzymes as revealed by iTRAQ analysis for *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB*, *C. acetobutylicum* ATCC 824 pMTL85143 vector control, and *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* with 10 μ M QspB 7 addition. The fold changes are shown beneath each heading. Three comparisons are shown for each, *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* against *C. acetobutylicum* ATCC 824 pMTL85143 (yellow), *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* + 10 μ M QspB 7 against *C. acetobutylicum* ATCC 824 pMTL85143 (blue) and finally *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* + 10 μ M QspB 7 against *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* (green). Where multiple gene loci represent each enzyme, they are listed on the left hand side of the comparison. ND denotes that peptides of this protein were not detected in the peptide library. This figure was compiled by Dr. Raut of The University of Sheffield.

5.3.1.2 Glycolysis enzymes

In order to understand more fully the metabolic processes that were taking place during QsrB overproduction, the enzymes involved in glycolysis were analysed from within the proteomics data set. Of all the enzymes that were differentially regulated during glycolysis, each of them were upregulated during QsrB overproduction. Glucose-6-phosphate isomerase was upregulated by 1.18 fold (p -value = 1.60×10^{-6}). Triosephosphate isomerase was upregulated 1.31 fold (p -value = 8.51×10^{-15}). In addition to this, both putative pyruvate kinases (CA_C0518 and CA_C1036) were upregulated 1.25 and 1.23 fold (p -values = 1.80×10^{-18} and 1.08×10^{-10} respectively). Of the 5 proteins mentioned, upon addition of 10 μ M QspB 7 the inverse regulation was seen for 3 of them: triosephosphate isomerase and both pyruvate kinases were down regulated -1.34 fold, -1.27 fold and -1.31 fold respectively (p -values = 2.67×10^{-8} , 2.69×10^{-9} and 5.90×10^{-16}). Adding to this, the phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS), involved in transporting sugars across the membrane, was upregulated 1.7 fold in the QsrB overproduction background, and the converse seen upon QspB 7 addition, -1.48 fold (p -values = 1.45×10^{-5} , 7.4×10^{-4} respectively).

5.3.1.3 ABE fermentation pathway

To continue the analysis of induced changes of *C. acetobutylicum* ATCC 824 metabolism during QsrB overproduction, proteomic data relevant to the ABE fermentation pathway was condensed into Figure 5.2, produced by Dr. Raut (University of Sheffield). It can be seen from this figure that in a pMTL85143-*qsrB* background many of the enzymes leading to the production of butyrate were upregulated. The converse can also be seen when comparing the pMTL85143-*qsrB* strains grown in the presence and absence of 10 μ M QspB 7, indicating the effect can be relieved in a specific manner via conjugate peptide addition.

Interestingly, the acetate kinase enzyme (CA_C1743) was shown to be down regulated in the QsrB overproduction strain by a factor of 1.48 (p -value = 2.85×10^{-11}). This enzyme was not detected in the QspB 7 addition experiment. However, both phosphotransacetylase and hydrogenase were in greater abundance in the QsrB

overproduction strain, being 1.13 and 1.26 fold increased (p -value = 5.75×10^{-6} and 5.68×10^{-12}). The reverse is also seen upon QspB 7 addition, with the proteins in a lesser abundance (factor 1.30 and -1.47 fold) (p -values = 3.61×10^{-13} and 5.56×10^{-22}).

It can also be seen that many of the enzymes leading to the formation of the solvents acetone, butanol and ethanol were downregulated during QsrB overproduction. The two acetate/butyrate CoA transferases subunits CtfA and CtfB (CA_P0163, CA_P0164) conferring the production of acetoacetate, were downregulated by factor 1.51 and 1.44, respectively, in the *qsrB* overexpressing strain in comparison to the control (p -values = 1.97×10^{-7} , 1.03×10^{-3}). The bifunctional acetaldehyde/dehydrogenase AdhE1 (CA_P0162), known to be essential for butanol production under the employed batch culture conditions and utilised in the production of acetylaldehyde and butylraldehyde from acetyl-CoA and butyryl-CoA, respectively, was downregulated -1.48 fold (p -value = 3.65×10^{-16}) (Cooksley et al., 2012). AdhE1 also catalyses the subsequent reduction of acetylaldehyde and butylraldehyde to ethanol and butanol, respectively. There is another bifunctional acetaldehyde/alcohol dehydrogenase in *C. acetobutylicum* ATCC 824, AhdE2 (CA_P0035), although this protein was not detectable within the QsrB overproduction compared to the vector control, and the QsrB overproduction with 10 μ M QspB 7 compared to the QsrB overproduction data sets. The butanol dehydrogenase BdhB (CA_C3298), involved in the final stages of butanol production, was also shown to be downregulated by a factor of 1.44 (p -value = 8.40×10^{-5}). BdhA was not detected within the peptide libraries, but BdhC was shown to be downregulated by a factor of 1.46, and the inverse seen upon QspB 7 addition (1.48) (Yoo et al., 2015). As can be seen from Figure 5.2, the above mentioned protein abundances in the *qsrB* overexpressing strain are reversed upon addition of the conjugate peptide QspB 7.

5.3.1.4 Components of the sporulation pathway

At the early time point analysed, it was not expected to observe proteins relating to sporulation. This was true for the pMTL85143 vs pMTL85143-*qsrB* comparison; however upon addition of QspB 7 to the QsrB overproduction some sporulation related proteins were seen to differ in relative abundance. The SpoIVA (CA_C1713) sporulation protein, playing a role in spore coat morphogenesis, was upregulated by a

factor of 2.48, also SpoIIIAH (CA_C2086) was shown to be upregulated 2.10 fold. The spore cortex lytic enzyme (CA_C3244) was shown to be upregulated by 2.31 fold, despite its role in germination.

5.3.1.5 Oligopermease SBPs

Only one oligopeptide permease SBP was detected in the proteomics screen, CA_C3632. The protein abundance was increased 1.34 fold during QsrB overproduction (p -value = 2.25×10^{-16}). The converse was seen when analysing the pMTL85143-*qsrB* with 10 μ M QspB7 addition strain, that downregulated the protein by -1.25 fold (p -value = 5.03×10^{-10}). On the topic of SBPs, two that bind sugars were detected within the proteomics screen, CA_C0664 and CA_C0662. During QsrB overproduction these proteins were downregulated relative to the vector control, -1.77 and -2.23 fold (p -values = 1.22×10^{-4} and 7.80×10^{-4} respectively). Upon QspB 7 addition these proteins were increased in abundance by 1.96 fold and 2.52 fold (p -values = 5.60×10^{-5} and 1.24×10^{-4}).

5.3.1.6 Proteins relating to motility

It can be difficult to conclude the present activity of an organism from a proteomics data set. However, an interesting set of gene products, related by function were detected and merit highlighting. The Flagellin protein Hag (CA_C2203), the flagellar basal body FlgE (CA_C2154), and MotB (CA_C0305) an important component of the flagella system, were shown to be upregulated when QsrB was overproduced. They were upregulated by 1.51 fold, 1.38 fold and 1.33 fold (p -values = 6.28×10^{-17} , 6.95×10^{-7} and 2.47×10^{-4} respectively). Interestingly, each of these gene products were decreased in abundance upon addition of 10 μ M of QspB 7 (-1.41 fold, -1.39 fold and -1.30 fold p -values = 1.52×10^{-6} , 5.16×10^{-5} and 1.70×10^{-3} respectively). In addition to this, 5 gene products relating to chemotaxis were detected (CA_C1600, CA_C1579, CA_C2217, CA_C0909 and CA_C0433). All of the aforementioned proteins were upregulated between 1.29 – 1.44 fold (p -values ranged from 8.77×10^{-12} to 1.52×10^{-3}). Although it should be noted that 3 other proteins relating to chemotaxis were identified in the latter, CA_P0048, CA_C3545 and CA_C0432. All of these genes

products were down regulated -1.36 fold, -1.42 fold and -1.56 fold respectively (p -values 7.20×10^{-4} , 1.09×10^{-3} and 1.11×10^{-3}).

5.3.1.7 Stationary phase regulators

Interestingly, the gene product of CA_C0957, mentioned in the previous chapter and the basis of a body of work performed by Bean (2016), was downregulated upon QsrB overproduction. CA_C0957 has been shown to be a master regulator of stationary phase characteristics. The associated protein was decreased in abundance by -1.55 fold (p -value = 7.85×10^{-6}). The inverse is seen upon addition of 10 μ M QspB 7 (1.25 fold increase, p -value = 1.72×10^{-3}).

5.3.2 Analysis of enzymatic activity

For the verification of the obtained proteomic data, four enzymatic reactions were chosen based on their relative abundance in the obtained data sets, phenotypic relevance and ease of measurement. These were; acetate kinase (potassium acetate), butyrate kinase (sodium butyrate), butanol dehydrogenase (butaraldehyde and NADPH) and nitroreductase activity (cytochrome C). As highlighted under section 2.7 of the materials and methods, cell samples were taken at the same time point and under the same conditions as for the proteomics analysis and lysates (crude extracts) prepared. Each of the enzymatic assays were performed using crude lysate, as highlighted in section 2.7 of the materials and methods, and the enzymatic activity calculated as U mg^{-1} .

Acetate kinase and butyrate kinase were measured in the unphysiological direction using potassium acetate and sodium butyrate as substrates, respectively. Combined butanol dehydrogenase activity was determined in the physiological direction using butyraldehyde and NADPH as substrates. Nitroreductase activity was measured via the reduction of cytochrome C.

As can be seen by Figure 5.3 the only enzymatic activities to generate results that differed significantly between conditions were the acetate and butyrate kinase assays. In both cases the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* extracts had a higher activity than those of the pMTL85143 control, and the addition of QspB 7..

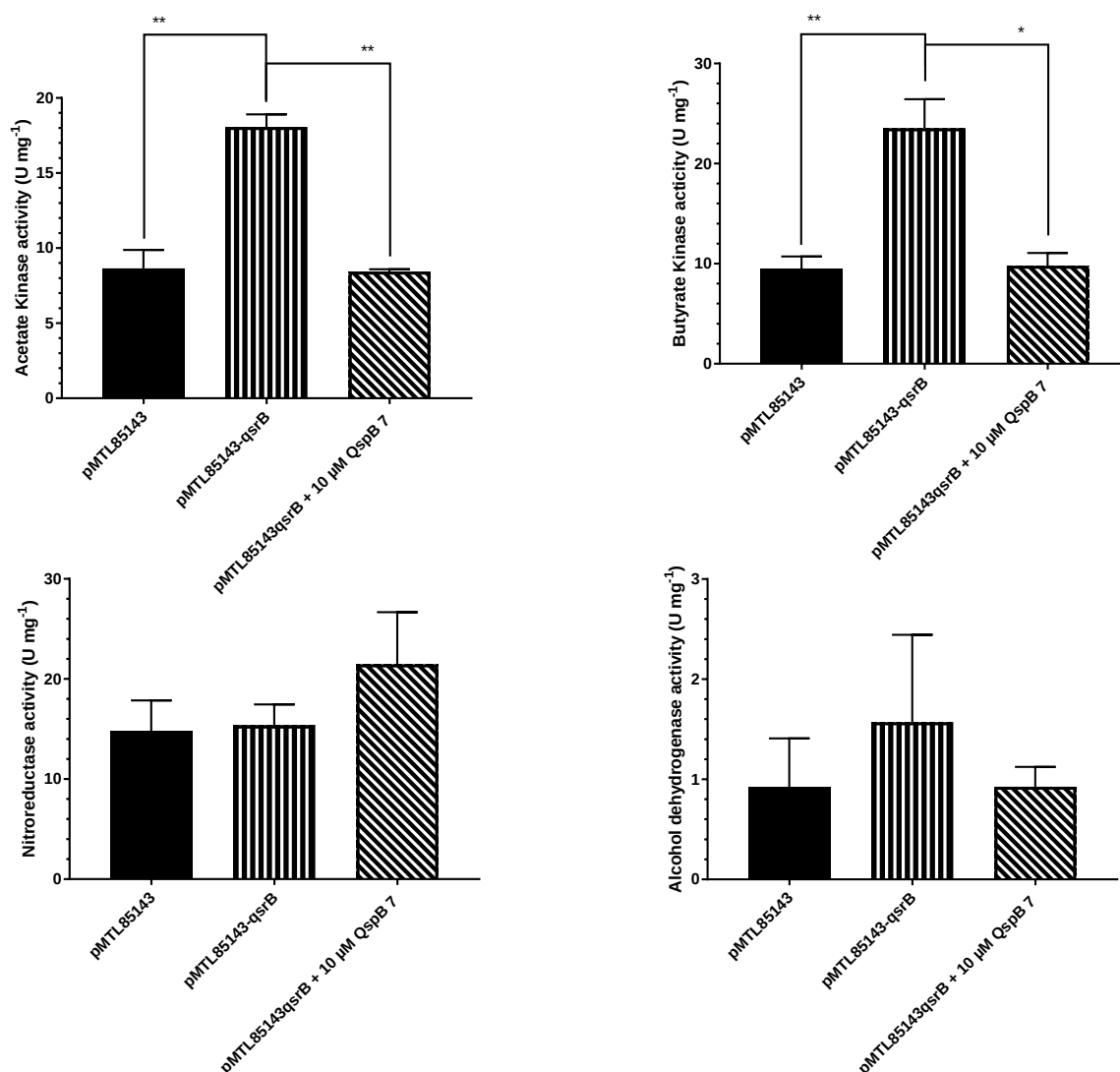


Figure 5.3 Enzymatic activity of Acetate Kinase, Butyrate Kinase, Nitroreductase and Alcohol Dehydrogenase after 24 hours of fermentation in crude extracts of *C. acetobutylicum* ATCC 824 pMTL85143 (black), *C. acetobutylicum* ATCC 824 pMTL85143-qsrB (vertical hashed lines) and *C. acetobutylicum* ATCC 824 pMTL85143-qsrB with 10 μM QspB 7 addition (angled hashed lines). The four graphs show the mean average enzymatic activity of the four enzymes displayed on the Y axis. Error bars represent the standard deviation of the mean. Statistically significant differences were calculated between two sets of measurements, indicated by solid bars with asterisks denoting the level of significance (* indicates $P \leq 0.05$, ** indicates $P \leq 0.01$, *** indicates $P \leq 0.001$ and **** indicates $P \leq 0.0001$).

With regards to acetate kinase the QsrB overproduction strain had an statistically increased activity of $18.0 \pm 0.6 \text{ U mg}^{-1}$, which upon addition of 10 μM of QspB 7 dropped to $8.5 \pm 0.1 \text{ U mg}^{-1}$, almost identical to the vector control level of $8.7 \pm 0.7 \text{ U mg}^{-1}$ (p -values = 3.86×10^{-3} and 2.65×10^{-3} , respectively). The butyrate kinase activity was also statistically increased in the QsrB overproduction strain at 23.0 ± 2.0

U mg⁻¹, compared to the 9.5 ± 0.7 of the vector control and 9.8 ± 0.9 U mg⁻¹ with 10 µM QspB 7 addition (p -values = 4.36×10^{-3} and 2.56×10^{-2} , respectively). These results were in agreement with the increased acid production observed in *qsrB* overexpressing strains, and for butyrate kinase agreed with the proteomics analysis, but acetate kinase results disagreed with the latter.

Nitroreductase and alcohol dehydrogenase activities were not as clear cut as far as statistical significance was concerned. The activities present in crude extracts representing the different experimental conditions did not differ significantly from each other. The QsrB overproduction with 10 µM QspB 7 addition strain had a slightly higher mean activity of 21.5 ± 3.7 U mg⁻¹, but this did not differ significantly from the vector control (p -value = 0.16). The butanol dehydrogenase activity overall was higher than recorded crude extracts (Dürre et al., 1987), and roughly a third of more sophisticated techniques (Yoo et al., 2015). The activity for crude extracts from the pMTL85143-*qsrB* strain was somewhat higher than the others, with an average activity of 1.6 ± 0.6 U mg⁻¹, but this difference was not statistically significant (p -value = 0.34 and 0.41 respectively).

5.3.3 GusA reporter analysis of significant promoters distinguished from phenotypic and proteomic analyses

Following failed attempts to label the QsrB protein with C-terminal SNAP and CLIP-tags (data not shown), it was decided to use a more classical approach and generate transcriptional reporters based on monitoring β-glucuronidase activity conferred by expression of a *gusA* reporter gene. The GusA reporter plasmid was constructed by a DTP rotation student in the lab, Jessica Locker, using the GusA gene from *Streptococcus agalactiae* DSMZ 2134. This gene was chosen because of its GC content and codon usage, which were reasonably close to those of *C. acetobutylicum* ATCC 824. The encoded β-glucuronidase had been shown to possess favorable kinetic properties (Wallace et al., 2015) similar to *E. coli* GusA, the latter having been used for a previously published clostridial reporter construct (Girbal et al., 2003). Briefly, the *S. agalactiae* structural gene was amplified by PCR, using primers with appropriate restrictions sites, and cloned into plasmid pMTL85141, to which the investigated promoters could be introduced using the RBS from the *fdx* gene.

Restriction sites in this construct allow for convenient swapping of promoter regions. Following confirmation of activity in both *E. coli* and *C. acetobutylicum* ATCC 824, the vector was made available for the work described below. The chosen promoters were; *fdx*, *qsrB*, *adc*, *adhE*, *ptb* and CA_C0957. The *fdx* promoter was chosen as a constitutively active positive control. The *adc*, *adhE* and *ptb* promoters were chosen based on their ability to give an insight into the acidogenic or solventogenic state of the organism. The *adc* gene being involved in the production of acetone, the *adhE* gene being involved in the production of butanol and ethanol, and finally the *ptb* gene involved in the production of butyrate. The fold change in the before mentioned genes can be seen in Figure 5.2. The promoter for the CA_C0957 gene was chosen based on its decrease in abundance by -1.55 fold ($p\text{-value} = 7.85 \times 10^{-6}$) during *qsrB* overexpression, and its importance as a stationary phase regulator.

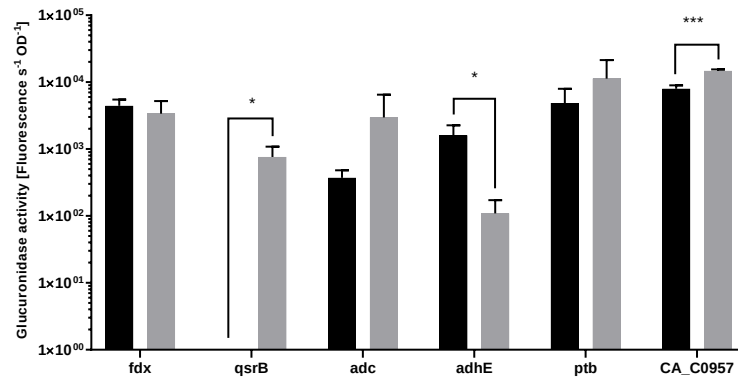
The putative promoter regions, the 200bp region immediately upstream of the first codon, of these genes were amplified via PCR with inclusion of appropriate restriction sites for subsequent cloning into the GusA reporter plasmid. The strains chosen for the study were *C. acetobutylicum* ATCC 824 as a WT control and its derived *qsrB* insertional mutation, *C. acetobutylicum* ATCC 824 CA_C0324::CTermB, generated during the work of Kotte et al. (2017). This was a ClosTron insertional mutation before mentioned in Chapter 3. Both strains were transformed with the aforementioned *gusA* reporter constructs. The *gusA* gene expression was thus driven by each of the investigated promoters, and the resulting enzymatic activity tested following the protocol outlined in section 2.6.8 of the materials and methods chapter. Some difficulty occurred during growth in liquid medium due to the strong tendency of the *C. acetobutylicum* ATCC 824 *qsrB* strains carrying the aforementioned reporter constructs to flocculate. Planktonic cells were present in the flasks, however flocculation of some of the biomass was always seen (data not shown) which made it difficult to ensure that cultures were showing equivalent growth and to take samples containing the same amount of biomass.

As can be seen from Figure 5.4 after only 8 hours of growth, the *fdx-gusA* fusion strains were already showing their highest level of activity. The *C. acetobutylicum* ATCC 824 pMTL85141_P*fdx-gusA* strain had a relative activity of 4430 ± 607

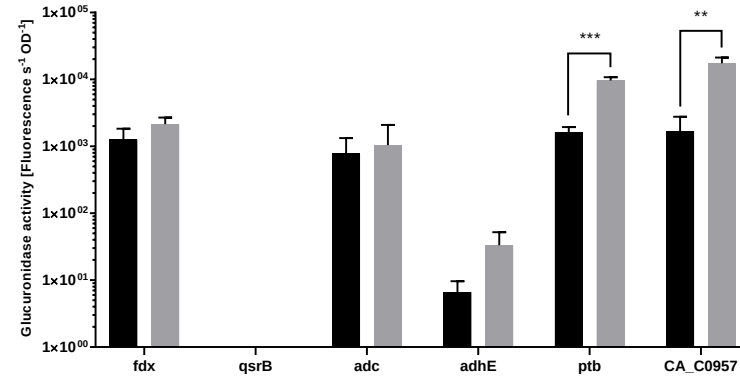
Fluorescence $\text{s}^{-1} \text{OD}^{-1}$, and the respective *qsrB* mutant strain produced a comparable 3372 ± 1065 . There were three statistically significant differences of promoter activity between the various reporter construct harbored by WT and *qsrB* mutant strains after 8 hours of growth. The first was the surprisingly greater activity of the *adhE* promoter in the WT strain, with a measurement of 1606 ± 375 , compared to 109 ± 37 Fluorescence $\text{s}^{-1} \text{OD}^{-1}$ of the *qsrB* mutant strain (p -value = 0.0165). The second was with regards to the CA_C0957 promoter, which was statistically higher in the *qsrB* mutant, with a measurement of 14712 ± 467 relative to the 7935 ± 587 measured for the WT strain (p -value = 8.31×10^{-4}). The third was the *qsrB* promoter, in which the WT strain had promoter measurements suggesting that it was below the experimental detection limit, having a value of -0.086 ± 0.03 Fluorescence $\text{s}^{-1} \text{OD}^{-1}$. The *qsrB* mutant had a much greater level of 761 ± 228 (p -value = 0.02). The *ptb* and *adc* promoters also appear to have a greater activity within the *qsrB* mutant. However, these were not sufficient enough to confer statistical significance.

After 24 hours, the observed differences for CA_C0957 promoter activity were considerably increased due to an overall drop in activity in the WT, producing

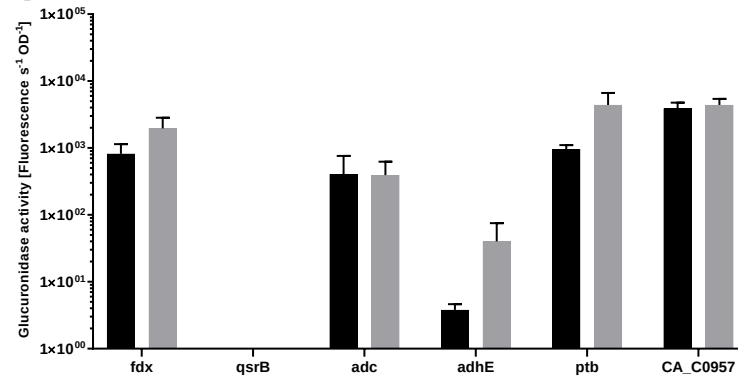
8h)



24h)



48h)



■ *C. acetobutylicum* WT
 ■ *C. acetobutylicum* CA_C0324::CTermB

Figure 5.4 Promoter activity in *C. acetobutylicum* ATCC 824 WT (black) and *C. acetobutylicum* ATCC 824 CA_C0324::CTermB (grey) as determined by transcriptional *gusA* reporters. GusA (β -Glucuronidase) activity was measured at 8, 24 and 48 hours. Each of the columns shows the mean of 3 replicates, with error bars showing the standard deviation of the mean. Statistical significance was determined for each fusion. Where a significant difference was observed, this is indicated by a connecting bar, with asterisks displaying the p -value (p -value ≤ 0.05 = *, ≤ 0.01 = **, ≤ 0.001 = ***).

measurements of 17386 ± 2197 and 1678 ± 632 Fluorescence s^{-1} OD $^{-1}$, respectively (p -value = 0.00235). Similarly, *ptb* promoter activity was seen to decrease in the WT strain producing an activity of 1618 ± 185 relative to the 9618 ± 664 Fluorescence s^{-1} OD $^{-1}$ produced by the *qsrB* mutant (p -value = 3.15×10^{-4}). Interestingly, *qsrB* promoter activity as close to the detection limit with the WT strain producing relative activity measurements of -0.051 ± 0.02 , and the *qsrB* strain producing measurements of 0.14 ± 0.29 Fluorescence s^{-1} OD $^{-1}$.

Levels of the *qsrB* promoter activity remained almost undetectable within the 48 hour samples of the WT and *qsrB* mutant, with measured activities of 0.17 ± 0.21 and 0.098 ± 0.013 Fluorescence s^{-1} OD $^{-1}$, respectively. Following analysis, promoter activities for any of the fusions did not differ significantly between the two strains after 48 hours of growth. The CA_C0957 promoter showed similar activity in the *qsrB* mutant when compared to the wild type, the latter producing 3886 ± 502 and the mutant 4401 ± 591 Fluorescence s^{-1} OD $^{-1}$. Following the increase seen at 8 hours, relative activity of the *fdx* promoter stayed almost constant at the 24 and 48 hour time points. Activity of *ptb* was greater in the *qsrB* mutant strain, measured at 4393 ± 1297 , compared to 957 ± 86 in the WT control, but this difference was not statistically significant.

5.3.4 Swimming and swarming assays of the *C. acetobutylicum* ATCC 824 pMTL85143 and *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strains

Due to the previously mentioned alteration in proteins associated with motility in the proteomics data set, it was decided to attempt to analyse the swimming and swarming phenotypes of the *C. acetobutylicum* ATCC 824 pMTL85143 and *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strains. Plate-based swimming assays require the organism of study to be stabbed into a semi-solid agar, in which it will start swimming, thus producing a halo the size of which can be measured (Cai et al., 2017). However, such assays are notoriously difficult to perform when working with *C. acetobutylicum* ATCC 824 due to the volume of gas it produces. It was no surprise then, that when the swimming assay protocol was followed as per section 2.6.9 of the materials and methods section, within 24 hours the semi-solid media had already

begun to inflate, and hence the assay was void. After several attempts to validate the swimming assay protocol within the strains, the assay was abandoned. A swarming assay protocol was also followed with the pMTL85143 and pMTL85143-*qsrB* strains (Baban et al., 2013). This required a small amount of cells to be placed on the surface of 0.24% Gelzan (Sigma Aldrich) plate, thus permitting the spread of bacteria on the surface without concomitant gas build-up. However, success was limited as shown in Figure 5.5.

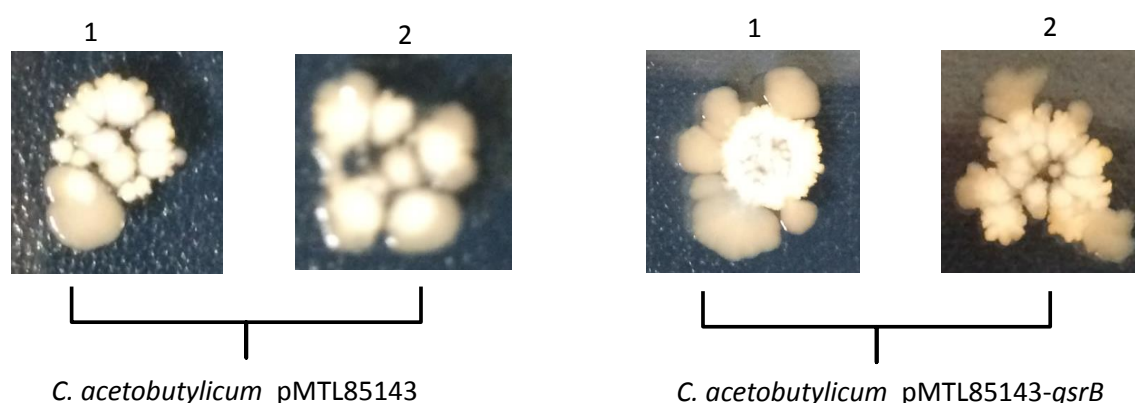


Figure 5.5 Swarming assay of the *C. acetobutylicum* ATCC 824 pMTL85143 and *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB*. 2 μ l aliquots of cells suspensions derived from 24-h-old liquid cultures were deposited on the surface of 0.24% Gelzan 2xYTG agar plates and incubated under anaerobic conditions for 120 hours.

It is perhaps possible to argue that the growth of the *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* strain showed a slightly greater ability to produce outgrowths of swarming than its vector control. However, it is clear from Figure 5.5 that the limited swarming-like outgrowth observed was atypical and differed from the larger structures seen for other bacteria (Baban et al., 2013; Cai et al., 2017). Several attempts were undertaken to optimise conditions, including changing the inoculum from 2 μ l to 5 μ l, varying the incubation time and concentration of Gelzan, but without success. Therefore, the analysis of possible effects of QsrB overproduction on swarming motility was abandoned.

5.3.5 Attempted tagging and isolation of the QsrB regulator

Attempts were made to purify the QsrB regulator to enable further characterisation. The *qsrB* gene was amplified by PCR using primer set QsrB_ C_F2 / QsrB_ C_R2_Strep and cloned into plasmid pMTL85143, thus generating a C-terminal Strep-tag fusion, with tagged gene being under control of the *fdx* promoter. Correct creation of the plasmid was confirmed via restriction digest and Sanger sequencing. Furthermore, the resulting construct would also allow expression in the native *C. acetobutylicum* ATCC 824 background if required. Cells were harvested after another 24 hours of incubation and cells then lysed in accordance to section 2.3.6.1 of the materials and methods chapter. The cell lysate, representing the soluble protein fraction, together with that of an empty plasmid control, were analysed by SDS PAGE for the presence of a protein of the predicted size of approximately 52 kDa (including the strep tag addition). However, no distinct band representing a protein of this size could be detected. Similarly, there appeared to be no protein of the expected size in the insoluble fraction. Nevertheless, considering that only relatively small amounts of QsrB may have been produced, the cell lysates were subjected to Strep-tag affinity purification as described in 2.3.6.1 and the eluted protein fraction analysed by SDS PAGE. However, no protein was detectable in the eluted fractions (data not shown). The experiment was repeated several times without success. The plasmid was re-isolated from the *E. coli* strain, and again analysed via restriction digest and sequencing to confirm that inactivation had not occurred via a mobile element, as sometimes seen for genes encoding toxic proteins.

Following failed production of correctly sized bands from the *E. coli* TOP10 pMTL85143-*qsrB*-Strep combination, an experimental redesign led to the change of strain, vector and tag. A new plasmid was constructed using the ProEXHT vector (Life Technologies), for the creation of his-tag fusions. The *qsrB* gene had the XhoI and NcoI restriction sites added via PCR, but returned two fragments ~1.5kb and ~2kb in length. Sequencing data showed these fragments to be artefacts. Due to time constraints the project was halted at this point to focus on other endeavors.

5.4 Discussion

Comparative proteomics on *C. acetobutylicum* ATCC 824 in order to identify changes in protein composition is not a new approach. 2-dimensional gel electrophoresis (2DGE) based approaches for this organism date back more than 25 years and were first employed to analyse its response to heat shock (Pich et al., 1990). In this early work, proteins of interest were identified following 2DGE separation through their isolation from the gels, fragmentation, and Edman degradation (Bahl et al., 1995; Pich et al., 1990). Subsequent studies for *C. acetobutylicum* ATCC 824 used similar approaches to, for instance, compare different strains (Sullivan and Bennett, 2006), establish the organisms response to oxygen (May et al., 2004), its membrane proteome (Mao et al., 2011) and changes associated with acid and solvent metabolism (Janssen et al., 2010), although in some of this work 2DGE was often coupled to increasingly sophisticated mass spectrometry based procedures for protein identification. More recently, entirely mass spectrometry based methods have been applied to establish the organism's protein interaction network during butanol fermentation (Sivagnanam et al., 2013), and the influence of lignin on solvent production (Raut et al., 2016). In the latter study, an iTRAQ-based approach was used as this allows comparative multiplexing, with increased statistical validation for gaining a comparative view of temporal expression patterns of proteins (Rauniyar and Yates, 2014).

The method is based on creating peptide libraries of each experimental sample through treatment of the proteins with proteases, and then tagging each of the peptides with differing isobaric labels that bind the N terminus, or ϵ -amino group of lysine (Rauniyar and Yates, 2014). The libraries are combined and analysed using an LC-MS/MS. The iTRAQ labels have a reporter region that has a specific mass, which is easily detached during mass spectrometry analysis. The ratios of these reporter regions are what show the relative abundance of the associated protein, which in turn, allows the production of comparative proteomic data. In the work presented here, the iTRAQ approach established by Raut and co-workers was employed to identify key changes in protein production due to overexpression of the quorum sensing regulator QsrB in presence and absence of its cognate signal. Whilst iTRAQ has the benefit of increased statistical validation, and increased throughput, it also has limitations as changes in protein production are often underestimated (Ow et al., 2009). This was

also observed here, as key acid and solvent genes were only identified by this method as being up or downregulated up to 1.55 and 1.51 fold, respectively, although the observed phenotypes suggested far more drastic changes.

Overall, however, the observed trends agree with the phenotypic findings in that enzymes required for solvent formation are generally down-regulated in the *qsrB* overexpressing strain, and acid pathways maintained, an effect that is largely reversed upon addition of synthetic QspB 7 signal. However, with 233 proteins changed upon *qsrB* overexpression, which often will lead to further downstream effects, it is difficult to extract what might constitute a possible core regulon for this regulator.

Despite a reduction in overall consumed glucose seen during *qsrB* overexpression in the previous chapter, multiple enzymes involved in the glycolysis pathway are shown to be in greater abundance during QsrB overproduction relative to the vector control. 5 enzymes were shown to be upregulated, namely glucose-6-phosphate isomerase, triose phosphate isomerase, enolase and both pyruvate kinases (CA_C0518 and CA_C1036). While previous data takes into account the total amount of glucose consumed over a 120 hour time period, the proteomics data is from a single time point at 24 hours of growth. Whilst *C. acetobutylicum* ATCC 824 overproducing QsrB consumes less glucose overall, based on this data it may at this time point be undergoing a greater level of glycolysis. It cannot be said from the current evidence at what time in growth, if ever, generation of acids from glucose ceases. The QsrB overproduction strain is capable of producing all the ATP it requires for growth through acidogenesis, potentially producing an average of 3.25 mol of ATP per 1 mol of glucose fermented, whilst using hydrogen to dispose of excess hydrogen and electrons (Jones and Woods, 1986).

There is however a limit to the production of these acids, given that their increase in the surrounding media could allow for flow of protons into the cell (via undissociated acids), and hence destruction of the promotive force (Jones and Woods, 1986). There is sufficient evidence from the proteomic data comparing QsrB overproduction with the vector control, that *qsrB* overexpression has caused the organism to remain in the acidogenesis. This is suggested by the increase in relative abundance of a number of

acidogenic enzymes displayed in Figure 5.2 and corroborated by other proteomic studies (Jang et al., 2014). Interestingly, the relative abundance of acetate kinase appeared to be reduced during QsrB overproduction, and hence its activity and that of the increasingly abundant butyrate kinase, was measured via hydroxamate assay of the crude lysate. Increased activity, in a statistically significant manner, of both enzymes was recorded during QsrB overproduction, which was shown to be reduced when in the presence of sufficient conjugate peptide. This further corroborates the fact that the organism remains in an acidogenic state even when tested at this 24 hour time point. It is not clear why the observed acetate kinase activity is at odds with the proteome data. There is little evidence in the literature for regulation of acetate kinase activity via posttranslational modification (Venkataramanan et al., 2015), but this is one possibility why in vitro activity and relative abundance did not correlate here. Crude lysate was used for all the enzymatic reactions; it's possible therefore that a less specific enzyme may also utilise the substrate.

It would not be expected to observe proteins associated with sporulation at this time point, and that is true of the pMTL85143 vs pMTL85143-*qsrB* comparison (Alsaker and Papoutsakis, 2005; Jones et al., 2008). Interestingly the SpoIVA and SpoIIAH components of sporulation were shown to be in greater abundance upon addition of QspB 7, perhaps suggesting a direct or indirect activating effect on sporulation genes. With no other indication however, this remains to be seen.

The increased abundance of proteins associated with motility during *qsrB* overexpression could be taken as further evidence of its acidogenic state. Motility has been shown to be essential for efficient *C. acetobutylicum* ATCC 824 metabolism, and increased motility during glucose metabolism has even been linked to its preferential use as a carbon source (Gutierrez and Maddox, 1987; Sivagnanam et al., 2011). Indeed other proteomics studies have linked glucose consumption to increased motility proteins (Sivagnanam et al., 2011). Increased motility 'runs', that is the organism moving in a straight line, are associated with acidogenesis, whilst increased 'tumbling', a form of motion in which the organism revolves around to change direction, is associated with solventogenesis (Gutierrez and Maddox, 1987). Whilst from the current data it is not known if the organism is predominantly performing

‘runs’ or ‘tumbles’, it is known that the organism is in an acidogenic state and fermenting glucose, and hence an increased motility in which the organism is continually performing ‘runs’ is expected (Gutierrez and Maddox, 1987; Sivagnanam et al., 2011).

Intriguingly, upon QsrB overproduction, the oligopeptide permease SBP, CA_C3632 (OppA) was present in statistically increased abundance relative to the vector control. This gene product has been shown to be the most abundant ABC transporter in *C. acetobutylicum* ATCC 824 through proteomic study (Mao et al., 2011). It could be argued that this upregulation was due solely to nutritional purposes, which is possible given that the predicted tripeptide aminopeptidase PepT was also present in increased abundance. However, each of the aminopeptidases CA_C0214, CA_C215 and CA_C216 which are involved in the degradation and utilisation of proteins, are not present within the data. Combined with this the abundance of CAS amino acids available within the growth media, it seems unlikely that the organism would rely on oligopeptide uptake as a nitrogen source. As such, up regulation of CA_C3632 could be, either directly or indirectly, driven by the overproduction of QsrB for another purpose. It could imply that it is used by the organism for uptake of RNPP-type signal peptides, specifically QspB.

Another interesting observation comes from the GusA reporter assays. From this (Figure 5.4) we can see that *qsrB* promoter activity is undetectable throughout each of the analysed time points for the WT. There are several possible explanations for this. Some regulator genes are known to be expressed at low level, as only a low number of regulator proteins is required (Venkataramanan et al., 2015). It is also possible that the 200-250 bp upstream gene region which was fused to the GusA gene did not contain the entire promoter region required for expression. It is also possible that the regulator is produced at a specific time point that was not analysed, or the expression of the *qsrB* gene is reliant on conditions not met by the experiment.

However, the *qsrB* gene has been previously suggested to be involved in early solventogenesis, this is due to its inactivation leading to a slight increase in the production of early solvents (Kotte, 2013). If this were the function of the regulator

we would have expected to see it produced within the first 8 hours of growth, and possibly remain active in the later time points covered by the GusA assay. However, *qsrB* promoter activity was observed at the 8 hour time point in the *qsrB* mutant. This may suggest that, in the wild type, QsrB either directly or indirectly repressed its own regulation together with other factors that come into play at the later time points. Also, the activity of the recently analysed CA_C0957 regulator increased in activity within the same insertional mutation relative to the WT control at the 8 hour and 24 hour time points (Bean, 2016). This is corroborated by the proteomics data, showing the CA_C0957 gene product is decreased during *qsrB* overexpression, which is reversible upon QspB 7 addition. CA_C0957 has been shown to be a master regulator of stationary phase characteristics, and hence it is possible that interaction with this gene or operon could be a mechanism in which QssB effects solvent and spore formation (Bean, 2016).

To conclude, the proteomics screen revealed that the QsrB regulator when overproduced led to the down regulation of a number of solventogenesis proteins, whilst maintaining enzymes required for glycolysis and acidogenesis. It also revealed that a number of other proteins were potentially regulated by the QssB system, directly or indirectly, and those involved in motility have been highlighted. Acetate and butyrate kinase activity corroborated the acidogenic phenotype, even if acetate kinase appeared to disagree with the proteomic data. Whilst other tested enzymes were somewhat inconclusive. Several promoters were analysed in the organism using a GusA reporter, and it lead to the possibility of QsrB being involved in its own repression. So far, attempts to purify the QsrB regulator have been unsuccessful.

6. General discussion and further research

6 General discussion and further research

6.1 Background

RNPP type cell-cell communication, and changes in bacterial lifestyle are not an unusual pairing, and has been shown before (Agaisse et al., 1999; Cabrera et al., 2014). It is aptly demonstrated by the NprR regulator of *Bacillus thuringiensis* (Dubois et al., 2013; Perchat et al., 2016a). NprR is responsible for the change from the pathogenic phase of *B. thuringiensis*'s lifestyle within insect larvae, to its necrotrophic phase within the cadaver (Perchat et al., 2016a). The work conducted here applied this concept to the RNPP-type cell-cell communication system QssB within *C. acetobutylicum* ATCC 824. By working on the foundation set by Kotte (2013), the action of QssB was investigated, and through this work its effect on the entire solvent profile, as well as its effect on sporulation was studied and reported in detail. This will add to the body of work on the cell-cell communication systems of *C. acetobutylicum* ATCC 824 (Bean, 2016; Kotte, 2013; Kotte et al., 2017; Steiner et al., 2012). Previous work on the QsrB regulator began to add merit to the link between the switch from acidogenesis and solventogenesis with cell-cell communication, a topic that urged more research into the area for better understanding of this industrially relevant strain. However, while viewed as a repressor of solventogenesis and sporulation, the work here has perhaps begun to show the action of QssB is more complex. The research questions involved in understanding the QssB system of *C. acetobutylicum* ATCC 824 now benefit from a much greater range of data which are summarised under separate headings below.

6.2 Understanding the QspB signal peptide

6.2.1 Key findings of this study

The study that laid the foundation of this work, after the discovery of the RNPP-type regulators of *C. acetobutylicum* ATCC 824, had shown that overexpression of *qsrB* lead to reduction in the concentration of produced solvents (Kotte, 2013). This phenotype was confirmed and further analysed during the course of this study. The entire range of metabolic products were analysed repetitively, and the effect of regulator overexpression on spore formation also confirmed. Previously, overproduction of QsrB was performed under its native promoter using the multicopy

number plasmid pMTL85141, over the use of the pMTL85143-*qsrB* construct (Kotte, 2013). During the body of this work, overexpression of the regulator was achieved through the strong ferredoxin promoter, through the pMTL85143 plasmid used to recreate pMTL85143-*qsrB* (Kotte et al., 2017).

Initially the project centered on discovering the minimum length peptide required for overcoming the action of QsrB overproduction. A series of high purity peptides were tested based on a pilot study conducted during previous work, the sequence of which was based on a C-terminal region of the precursor QspB protein predicted to encode the mature signaling peptide (Kotte, 2013). We now know with certainty that the minimum length peptide capable of overcoming the action of QsrB is coded by the amino acids AEPTWGW. This is in keeping with published literature on the conjugate peptide of RNPP type regulators, as on more than one occasion has it been shown that a heptapeptide is responsible for regulator interactions (Bouillaut et al., 2008; Perchat et al., 2011).

Further information about the binding properties of the heptapeptide AEPTWGW were divulged through alanine scanning, a process shown to be useful in analysing cell-cell communication signals and also other important regulators (Jiang et al., 1999; Tal-Gan et al., 2013). The process showed the position 2 glutamic acid to be the most essential of all within the heptapeptide sequence, but not by a large margin. Exchange of any of the amino acids for an alanine caused a reduction of the ability of the heptapeptide to overcome the effects of QsrB overproduction by over 50%, apart from the position 6 glycine, that instead reduced it by 48.5%. This was calculated via the production of butanol, and it should be noted that it is unknown what the effect of exchanging the position 1 alanine would be, given that it was not included in the alanine scan.

The effect of ClosTron insertional mutation of *qspB* was shown to be subtle in previous work (Kotte, 2013). This lead to the concept that QssB could be engaging in cross talk with the other RNPP type conjugate signal peptides encoded within the *C. acetobutylicum* ATCC 824 genome. Given the before mentioned alanine scan data, the likelihood seemed low, but the hypothesis was tested nevertheless in order to confirm

the case. Synthetic nonapeptides of each conjugate peptide were used to supplement the QsrB overproduction strain (3.3.6). None were shown to be able to overcome the overproduction of QsrB phenotype.

Further interest in crosstalk was generated on the discovery that the closely related *Clostridium roseum* contained a homologous system to QssB (Poehlein et al., 2017). The *C. roseum* conjugate peptide contained the nonamer LAAEPTRPW, differing in only 3 positions to the *C. acetobutylicum* ATCC 824 nonamer LGAEPTWGW. It was shown that the *C. roseum* nonamer was also unable to alter the QsrB overproduction phenotype, and had no activity (3.3.6). It is not clear why this was the case, given that the alanine scan had shown that the tryptophan and glycine residues of this position were the least conserved with regards to QspB activity. It could be that the *C. roseum* nonamer second residue, alanine in place of a glycine, prevents cleavage of the oligopeptide to a heptamer, and hence no activity is seen. It is also possible that the 7th and 8th residues of the QspB nonamer, tryptophan and glycine, are both required simultaneously for ligand binding and subsequent changes to the regulator which affect its activity. This experiment alone has suggested multiple possible avenues for future experimentation, which could lead to greater understanding of the ligand regulator interaction.

Despite this, experimentation into the other RNPP-type conjugate peptides and that of the structurally similar *C. roseum* QspB indicated the exquisite specificity of the RNPP-type system QssB. One might be forgiven to think that cross-talk and inhibition of Gram-positive cell-cell communication systems happens often, given the body of literature that surrounds the topic, however these generally concern the Agr-type systems (Otto et al., 2001; Tal-Gan et al., 2013). It has been suggested that the increased cross-talk inhibition of the Agr systems is due to similarity in structure (Olson et al., 2014), due to each of the signal peptides being cyclic, yet amino acid composition was also shown to be integral to inhibiting properties (Sturme et al., 2002). Cross-talk inhibition or activation is rare in RNPP-type cell-cell communication, and to the author's knowledge is yet to be shown in-between species or even strains. Perchat and colleagues (2011) have shown the possible cross-talk of the NprR-NprX system in *Bacillus cereus* between pherogroups, but that seems to be

the extent of such data. The implication of this is that the specificity of the signaling peptides, and the binding of regulators make these systems very specific. This makes the multitude of RNPP-type cell-cell communications systems within *C. acetobutylicum* ATCC 824 even more impressive.

Further experimentation into the QspB peptide had shown that the addition of the peptide to *C. acetobutylicum* ATCC 824 WT lead to no observable change in solvent production (3.3.7). This would imply that QssB operates on a level of saturation during the conditions observed within the experiment. Alternatively, it could be that the system was not active under the employed conditions. However, previously published transcriptomic studies indicate expression of all *C. acetobutylicum* ATCC 824 RNPP-type systems (Nguyen et al., 2016). Given the complete lack of a WT response to the addition of QspB signal, it was surprising that on many occasions the *qsrB* overexpressing strain produced more solvents than the WT plasmid control. This suggested that QsrB's role in the cell may not be satisfactorily explained by assuming a simple repressor function.

6.2.2 Further research and summary

The length of the native peptide, i.e. the length of mature, processed QspB used naturally within the system, is not known and is a source of further investigation. Initial experimentation was carried out, attempting to detect the RNPP type peptides via the uses of LC-MS-MS. Initial spiking experiments displayed that the small oligopeptides could be extracted and detected from crude supernatant. However, attempts at detecting possible native peptides were unsuccessful (data not shown). It is likely that with further calibration such peptides could be detected. This would not only aid in the identification of the native peptide, but could also help identify the SBP required for binding of QspB. Previous examples of such studies have involved the purification tagged protein, exposure to peptides, isolate complexes and identify bound peptides via LC-MS/MS (Gogliettino et al., 2010).

It was previously commented on that the conjugate peptides from the other RNPP-type quorum sensing systems were incapable of alleviating the effects of QsrB overproduction, a processes that could have explained the mild phenotype observed

during insertional mutation of QspB (Kotte, 2013). In addition, the *C. roseum* QspB nonamer was also unable to overcome the effects of QsrB overproduction, despite its close residue similarity to QspB. This may give a novel opportunity for further analysing the selectivity of the QssB system. A number of novel synthetic peptides could be designed and their biological activity tested in conjunction with the QsrB overproduction strain.

A number of novel synthetic peptides could be designed and their biological activity tested in conjunction with the QsrB overproduction strain. These may involve amino acid changes in larger peptide precursors to identify residues required for processing, i.e. potential peptidase cleavage sites. Receptor specificity may be further explored by systematically replacing each position in the heptamer by closely related amino acids (i.e. an amino acid of similar size, charge or polarity) where possible. For instance, the glutamine present in position 2 of the heptamer may be replaced by an aspartate residue, the threonine by a serine, and the tryptophans by phenylalanines. Furthermore, based on the outcome of this approach, it may be possible to design “hybrid peptides”, for example *C. acetobutylicum* ATCC 824 QspB 7 derivatives in which single amino acids have been changed to their respective *C. roseum* counterparts. This way, it would be possible to identify the minimum number of (mutational) changes required for the conversion of the latter peptide into a sequence capable of stimulating the *C. acetobutylicum* ATCC 824 system.

However, when undertaking the above experiments it ought to be kept in mind that specific interactions are likely to occur at two different levels, the extracellular binding of the peptide to one or several SBPS and its intracellular binding to the QsrB receptor. Thus, ultimately, purification of both receptors and in vitro characterisation of their respective interactions with the different peptide derivatives will be required to fully elucidate the constraints placed on biological activity.

6.3 Oligopeptide permease uptake systems and QssB in *C. acetobutylicum* ATCC 824

6.3.1 Key findings of this study

C. acetobutylicum ATCC 824 was found, during the course of this study, to be in possession of an arsenal of oligopeptide SBPs. A possible 7 SBPs were identified, with perhaps one inactivated by a frame shift mutation (CA_C2377). Even when taking into account the different SBP loci, there appear to be at least 5 complete oligopeptide transporter systems. Compared to the 2 oligopeptide permease systems Opp and App, and one dipeptide permease system Dpp, of the Gram-positive spore forming model organism *B. subtilis*, *C. acetobutylicum* ATCC 824 would appear to contain an disproportionally large number (Levdikov et al., 2005), especially as it is known as saccharolytic, not a peptolytic, organism (Jones and Woods, 1986). The exact reasons for this are unknown, it could be that each oligopeptide permease is expressed within differing physiological conditions, and hence each covers a separate ecological niche (Doeven et al., 2005). However, an attractive working hypothesis is that it could be due to the relatively large number of RNPP-type cell-cell communication systems that are used by the organism (Kotte et al., 2017). It was therefore of interest to establish which of these systems were contributing to QspB 7 uptake.

Each of the oligopeptide permease SBPs were targeted for insertional mutation via the ClosTron method (Heap et al., 2010) to see whether this would affect phenotypes associated with QsrB regulation and, more importantly, whether the effects of *qsrB* overexpression in these mutants could still be alleviated through addition of synthetic peptide. Interestingly, of the three mutations that could not be obtained, the genes were located within operons or gene clusters including the components of an oligopeptide permease system, and in the case of CA_C3634 and CA_C3640 shared a very high sequence similarity to one another. It is possible that the insertional mutation would disrupt the function of the entire operon, and hence the inability for these genes to be targeted could indicate essentiality for any of the genes involved in the system.

However, despite insertional mutation of 4 oligopeptide permease SBPs within the repertoire of *C. acetobutylicum* ATCC 824, a single SBP specific to the uptake of the QspB conjugate peptide could not be found. Unfortunately, this neither confirms or denies the prospect that a single specific SBP is responsible for the uptake of QspB, or that the conjugate peptide is taken up by several less specific transporter systems (Gominet et al., 2001; Klepsch et al., 2011).

An interesting side effect of these experiments was that at least two of the SBP insertional mutations appeared to confer a resistance to the reduction in spore formation usually seen during QsrB overproduction. These were the CA_C0176::CTermB and the CA_C3632::CTermB strains. The reason why this took place is not sufficiently understood through the data at hand and also awaits final confirmation through gene complementation. However, if confirmed it would indicate a key finding from this body of work that was previously not considered: the regulation of solvent production and spore formation via QssB may be performed independently of one another. Based on the data available it is not possible to determine if the effects on solvent profile or sporulation are direct consequences of the increased relative abundance of QsrB, or rather due to its indirect downstream effect. However, the results for both CA_C0176 and CA_C3632 mutants show that whilst QsrB no longer repressed sporulation it did so for solvents and this effect could still be relieved through peptide addition. The cause of the resistance to the low sporulation phenotype cannot truly be speculated on, but it does imply that the regulation of QssB with regards to sporulation can be superseded. The fact that this was only visible upon the interruption of oligopeptide uptake could also imply that the observation is signal based. Perhaps the action of QsrB operates in tandem with another regulator, even another RNPP-type regulator. Without further study, it would be impossible to know with any degree of certainty.

Another interesting insight developed by this body of work was the reduction of glucose consumption by *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB*. It was perhaps this sole acquisition of data that began to reformulate the previous theory that QsrB was a repressor of solvent production (Kotte, 2013; Kotte et al., 2017). While it certainly can be said that *qsrB* overexpression leads to the vast reduction of solvents,

the observation that the organism has reduced glucose consumption was an intriguing novel observation (Figure 4.11 section 4.3.5). If the reduction of solvent profile was due to a decrease in glucose consumption, QsrB can hardly be credited with direct solvent repression, but could therefore be a repressor of glucose consumption. Indeed, in contrast to many other bacteria which sporulate in response to nutrient starvation (Dürre, 2014; Dürre and Hollergschwandner, 2004), *C. acetobutylicum* ATCC 824 shows reduced sporulation and solvent formation when supply of glucose is limiting (Bean, 2016). However, it is not possible to know at this stage why *C. acetobutylicum* ATCC 824 overproducing QsrB consumed less glucose, as again the process could either be a direct or indirect effect of its action.

6.3.2 Further research and summary

The evidence gained here led to some interesting questions that merit further study. The first experiment to be performed with respect to the SBP ClosTron mutants would be the independent verification of the results via plasmid complementation with the respective genes cloned into the pMTL85143 vector (Table 2.2 Materials and Methods). It should also be independently verified by the production of in-frame deletion mutants of each gene (Ehsaan et al., 2016).

It was not the subject of this body work to analyse why *C. acetobutylicum* ATCC 824 contains so many oligopeptide uptake systems compared to other organisms (Levdikov et al., 2005). The study of this question can be paired with further investigation of which oligopeptide permease system is required for the uptake of QsrB. In order to more expediently pinpoint the uptake system responsible for QspB uptake, one of the integral membrane proteins of each oligopeptide permease can be targeted (Figure 4.1 for system diagram). This would also lead to a null mutation of the system (Gominet et al., 2001). As it has been shown SBPs can act with multiple systems (Thomas, 2010) targeting an integral membrane protein may simplify the question. While ClosTron insertional mutation would be adequate, inframe deletion methods may even remove the whole coding operon (Ehsaan et al., 2016). Upon obtaining mutations of each of the oligopeptide permease systems, phenotypic screening could be undertaken in order to understand the effects of system interruption which could give indication as to their function. These mutations could

then undergo screening as was performed with the SBP insertional mutations of this study, by transforming them with pMTL85143-*qsrB* and supplementing with QspB 7 (4.3.5). Failure of the conjugate peptide to overcome the overproduction of its regulator in the presence of a specific integral membrane protein mutation would indicate that the targeted system is responsible for QspB uptake. However, this line of inquiry does make the assumption that there is no redundancy in the oligopeptide uptake systems with regard to RNPP-type conjugate peptides, which may be the case, and in this case in frame deletion techniques would allow the generation of double or triple mutants, if required

The insertional mutation of CA_C0176 and CA_C3632 allowed spore formation in the presence of QsrB overproduction whilst the CA_C2667 insertional mutation did not allow production of spores during *qsrB* overexpression (4.3.5). The data outlined here may suggest that the action of solvent repression and sporulation repression by QsrB could be independent of each other. There are multiple regulators implicated in the control of sporulation and solventogenesis, it is therefore possible that this may override the action of QsrB. Although, it is intriguing why this would occur in an SBP mutant.

C. acetobutylicum ATCC 824 is often cited as an interesting industrial strain relative to *E. coli* and *Saccharomyces cerevisiae* due to its wide range of carbon source utilisation (Abd-Alla et al., 2015; H. Li et al., 2014; Madihah et al., 2001; Naik et al., 2010). It should therefore be investigated whether or not the same effect of QsrB overproduction is observed when the organism is grown on a differing carbon source such as xylose, fructose, arabinose or starch. The possibility of glucose limitation, and the mixed sugars of the putative ecological niche, could be more reminiscent of the natural action of QssB. This could give an indication not only if the effects are glucose specific, but also give some idea as to whether the phenotype is related to physical uptake of certain molecules, or if catabolite repression is integral to the natural action of QssB (Ren et al., 2010). This relates specifically to the utilisation of the glycolysis pathway relative to the pentose phosphate pathway and phosphoketolase pathway, utilised by pentose sugars with minimal flux into glycolysis (Aristilde et al., 2015; Jin et al., 2014).

In summary, a number of interesting observations were made during this section of the work, even if they were not the initial objectives of the experimentation. It was found that *C. acetobutylicum* ATCC 824 harbors many more oligopeptide permease systems, assuming that they are functional, than was expected from the literature (Levdikov et al., 2005). Despite the targeting and interruption of 4 oligopeptide SBPs, none of these were shown to be the sole protein responsible for the uptake of QspB. This could be due to QspB uptake being general, or that the SBP responsible for uptake could not be inactivated via insertional disruption using the ClosTron method. Perhaps the most interesting section of evidence to arise from this area of work was the realisation that at least two of the ClosTron inactivated SBPs conferred a resistance to the reduction in sporulation seen during QsrB overproduction. However, this phenotype awaits confirmation through gene complementation. For the first time it was observed that QsrB overproduction lead to a reduction in glucose consumption. This introduced the idea that QsrB may not act as a direct repressor of solvent production and sporulation but rather a regulator of glucose consumption.

6.4 Proteomic analysis of QsrB over production

6.4.1 Key findings of this study

The RNPP type regulators have often been shown to be pleiotropic, and from the proteomics data it is clear to see that QsrB is no exception to this (Agaisse et al., 1999; Lereclus et al., 2000). To our knowledge, this is the first time proteomic analysis has been used in *C. acetobutylicum* ATCC 824 to analyse the effects of cell to cell communication, despite various proteomics approaches being well established for the organism (Janssen et al., 2010; Raut et al., 2016; Sivagnanam et al., 2011). That being said, RNPP type cell-cell communication with regards to *C. acetobutylicum* ATCC 824 is not present in the body literature and hence the work here is relatively novel within the field (Bean, 2016; Kotte, 2013; Kotte et al., 2017). Despite this, studies into the RNPP type cell-cell communication are appearing in *Clostridium spp.* and hence the work embodied here could represent an emerging field in these organisms (Martin-Verstraete et al., 2016; Peng et al., 2015).

During overexpression of *qsrB*, 233 proteins were shown to be differentially abundant. With addition of QspB 7 140 of these proteins returned to their original level supporting the idea that the peptide was required for QsrB inactivation. 142 proteins were shown to be in greater abundance during QsrB overproduction, and 91 shown to be in a lower abundance. Of the 91 proteins in lower abundance during QsrB overproduction, 48 of them were shown to be increased upon QspB 7 addition. If QsrB is indeed a repressor, and binds to gene promoters, these 48 proteins could possibly be investigated as direct targets. The proteomics dataset largely confirmed the *qsrB* overexpression observed phenotypes. Enzymes leading to the production of butanol and ethanol, namely AdhE1 and AdhE2 were down regulated. The butyrate CoA transferases, leading to butyrate and acetate assimilation and acetone production, were also shown to be in decreased abundance (Figure 5.2, section 5.3.1). Enzymes involved in the production of butyrate, such as Ptb and Buk, were in greater abundance, despite those involved in acetate production seemingly downregulated. Interestingly, proteins responsible for glucose metabolism seemed to be in greater abundance, despite the observed reduced glucose consumption of the strain. Multiple motility genes were shown to be upregulated, further supporting the strains acidogenic phase (Gutierrez and Maddox, 1987; Jones and Woods, 1986).

The *gusA* assay, coupling glucuronidase activity to a number of promoters, was beneficial in further understanding the action of QsrB as well as reinforcing the information obtained during the proteomics data. It should be noted that the experiments performed here were done so in the *C. acetobutylicum* ATCC 824 WT strain and the *C. acetobutylicum* ATCC 824 *qsrB::CTermB* insertional mutant. Activity of the *qsrB* promoter was only seen after 8 hours in the insertional mutation strain. It is possible, from this indication that the QsrB-QspB complex represses its own expression. This investigation also showed that during the 8 hour and 24 hour time points, promoter activity of the CA_C0957 regulator was significantly increased in the insertional mutation strain. In agreement with this finding, CA_C0957 protein levels were reduced upon *qsrB* overexpression but increased following QspB 7 peptide addition. This indicates that the activity of QsrB could be linked to this key regulator, which was previously proposed to be above the master regulator Spo0A in the regulatory hierarchy (Bean, 2016) thus providing a possible explanation for the

observed changes in solvent and spore formation in *qsrB*-deficient, showing increased solvents at earlier time points, and overexpressing strains, showing vastly reduced solvents and spores (Kotte et al., 2017).

Enzyme assays confirmed that the QsrB overproduction strain had greater acetate and butyrate kinase activity than the empty vector control, and in addition this activity was reduced upon supplementation with the conjugate peptide QspB. This is in agreement with the obtained physiological data but disagrees with the proteomics data indicating the QsrB overproduction strain had a lower relative abundance of acetate kinase. Taking all data together, this is strong evidence that at the analysed time point, the QsrB overproduction strain remained in an acidogenic state.

Swimming and swarming assays were inconclusive within the organism, and studies assaying these physical attributes, rather than gene expression indicators, are not widely reported (Gutierrez and Maddox, 1987; Christopher A. Tomas et al., 2003). Obtaining an indication for increased motility through omics data can be useful, but it is difficult to discern a rationale for organism motility through this method.

Attempts to purify QsrB through strep-tagging was also unsuccessful, a route of experimentation which could have yielded a number of interesting avenues of research. This would be an integrally important section of further work, and it is highly recommended that should the work proceed. That this is done from a purified protein the molecular and structural basis of regulator activity can unfold, by analysing both the ligand bound and unbound structures (Cabrera et al., 2014).

6.4.2 Further research and summary

The proteomics data set gave a much needed insight into the function of QssB, and led to several observations that were previously only speculation. Proteomics was utilised at the 24 hour time point of organism growth, and it is possible that further insight into the function of the QsrB regulator could be obtained with a greater resolution of data. It would be useful to obtain a proteomics data set as before, but for time points traditionally seen as acidogenic phase (before 13 hours of fermentation under our conditions) and solventogenic (post 13 hours of fermentation). This would

allow for a greater picture of any possible alterations in the actions of the QsrB regulator over time. These could then be correlated to transcriptomic data, obtained for instance by RNAseq (Venkataramanan et al., 2015). Given that QsrB possesses an helix-turn-helix, Xre-type DNA binding domain (Kotte, 2013), and therefore most likely is a transcriptional regulator, transcriptional changes occurring in the *qsrB* overexpressing strain shortly after addition of synthetic peptide may be more indicative of the genes under its direct control.

The specific motility of the organism and its *qsrB* derivatives could be analysed to confirm the changes implied by the proteomic analysis. Methods into analysing chemotaxis and the swimming and swarming motility of *C. acetobutylicum* ATCC 824 are poorly documented, and could indicate practical difficulty of the protocol or lack of meaningful data obtained from them (Gutierrez and Maddox, 1987; Steiner et al., 2012; Christopher A Tomas et al., 2003). Motility of many oxygen-tolerant bacteria has been quantified with the help of a Hobson BacTracker (Karim et al., 1998; Murray et al., 2002) but its utilisation for strictly anaerobic bacteria would require the instrument to be placed in a suitable anaerobic enclosure. Alternatively, a microcapillary-based method has recently been established in the laboratory to assess chemotactic behavior of various *C. acetobutylicum* ATCC 824 mutants (Mastrangelo, 2014), which may be adopted for the quantification of motility in *qsrB* overexpressing strains (Shonnard et al., 1992).

Enzyme assays analysing the activity of only four enzymes were conducted due to the time constraints of the project. The portfolio of assays and time points at which they are conducted could be extended. For example, acetate and butyrate kinase activity as well as solventogenic key enzymes such as CoA-transferase, acetoacetate decarboxylase, and alcohol and aldehyde dehydrogenases could be tested at regular intervals post inoculation, and a comparison made between the wild type strain and a strain that is overexpressing *qsrB*. Furthermore, it should be analysed how supplementation of the overexpression strain with QspB alters the activity at these time points. Together, these data would help to establish how and when the switch from acidogenic to solventogenic state is altered by QsrB.

The GusA assays gave insight into the possible self-repression of the QsrB-QspB complex. However, testing this hypothesis is not straightforward. Ultimate proof would require the identification of the *qsrB* promoter region and the potential QsrB binding site. This could be achieved by a combination of primer extension analysis (Carey et al. 2013), or differential RNAseq approach (Thomason et al. 2015), and successive truncation of the promoter containing sequence of the *gusA* reporter construct. Once narrowed down, this region could be subjected to band shift assays (Yuan & Wong 1995; Narberhaus 1999) and DNA footprinting analysis (Nguyen et al. 2016), requiring purified, active QsrB. Direct binding of the QsrB protein could then be demonstrated through a shift of the promoter/operator-containing DNA fragment during gel electrophoresis or its protection against DNase degradation, respectively, in either absence or presence of synthetic QspB peptide. However, these experiments would rely on the provision of purified, active QsrB protein.

If such protein was available, another major task could be tackled; the identification of target genes directly controlled by QsrB. Knowledge of these genes would mark a major advance in unravelling the biological role of this protein and go a long way towards answering many of the questions raised here and in the individual results chapters. Genes directly controlled by QsrB would be expected to contain binding sites in their promoter region, which could be identified through the application of ChIP-seq, a combination of chromatin immunoprecipitation (ChIP) with massively parallel DNA sequencing (Bonocora & Wade 2015). This would also require the availability of an antibody against QsrB or the use of specific tags for which antibodies exist. A complementary approach to this would be In Vitro DNA Affinity Purification and Deep Sequencing (IDAP-Seq), an in vitro analog of ChIP-Seq. In IDAP-Seq, purified tagged protein is incubated with genomic DNA and the DNA bound to protein is isolated by affinity purification rather than immunoprecipitation (Smith & Grossman 2015)

As the above experiments all rely on the availability of active, pure QsrB, further work is required to achieve purification. Work was initiated exchanging the Strep-tag attempted with a conventional His-tag, and exchanging the strain used for an expression strain. Future work may explore C- or N-terminal tagging, codon-

optimisation of the gene for expression in *E. coli* or alternative hosts and the use of different temperatures and induction regimes (Kimple et al., 2015; Young et al., 2012).

In addition to the identification of binding sites, a tagged and purified QsrB would give great insight into the function of the regulator. Through crystallisation the oligomerization state of the regulator could be determined. Many other RNPP-type regulators have been shown to not only operate as oligomers, but also the state of these oligomers may change upon interaction with the conjugate peptide (Gallego del Sol and Marina, 2013; Grenha et al., 2012; Perchat et al., 2016b). Furthermore, purified QsrB may be used to study its interaction with different peptide variants through isothermal titration calorimetry or surface plasmon resonance (Cabrera et al., 2014; Mohanraj et al., 2014). The tagging of QsrB for purification and further study is likely the most integral recommendation for further study from this body of work.

6.5 General Conclusion

To conclude this section, the body of work presented here led to a number of novel findings that fit well within the somewhat limited body of literature. Further work was performed showing that the QsrB overproduction strain lead to a drastically reduced solvent profile and sporulation phenotype. Alleviation of this phenotype could be performed via the addition of synthetically produced conjugate QspB oligopeptides of various lengths, and in doing so the minimum oligopeptide length required for activity was discovered to be the heptamer, AEPTWGW. This discovery was in keeping with wider literature concerning conjugate peptides of RNPP-type cell-cell communication systems in other organisms (Bouillaut et al., 2008).

The proteomics data set reinforced the idea that the organism remained in the acidogenic phase of growth during QsrB overproduction, and also suggested the organism is engaged in motility. Certain sugar uptake systems were shown to be in lower abundance during QsrB overproduction, despite glycolysis being generally upregulated. It was theorised that an increase in glycolysis and acidogenesis streamlines the production of ATP required for chemotaxis, despite glucose consumption being low, ATP generated from these molecules could be used for

motility (Jones and Woods, 1986). Attempts to verify swimming and swarming characteristics of the organism were unsuccessful, but would be helpful for defining the action of QsrB in further study (Gutierrez and Maddox, 1987). Enzyme assays are a standard in *C. acetobutylicum* ATCC 824 for analysing the phase of its growth, and during QsrB overproduction both an acetate and butyrate kinase assay reinforced the concept that organism was maintain acidogenic phase (Winzer et al., 1997). Attempts to purify the QsrB regulator were unsuccessful, but its purification should be considered a priority for future work on the QssB system. Protein purification has formed a standard for RNPP-type regulators, and has led to many influential works describing the regulators in detail (Cabrera et al., 2014; Grenha et al., 2012). Overall, the QssB system has been shown to be responsible for the control of chemotaxis and maintaining acidogenesis within *C. acetobutylicum* ATCC 824. The regulator is pleiotropic, like other RNPP-type regulators (Agaisse et al., 1999). While its action is still unknown, evidence here further shows that it is more complex than our previous conception (Kotte, 2013).

7. Bibliography

7 Bibliography

- Abd-Alla, M.H., Elsadek El-Enany, A.-W., 2012. Production of acetone-butanol-ethanol from spoilage date palm (*Phoenix dactylifera* L.) fruits by mixed culture of *Clostridium acetobutylicum* and *Bacillus subtilis*. Biomass Bioenergy 42, 172–178. <https://doi.org/10.1016/j.biombioe.2012.03.006>
- Abd-Alla, M.H., Zohri, A.-N.A., El-Enany, A.-W.E., Ali, S.M., 2015. Acetone–butanol–ethanol production from substandard and surplus dates by Egyptian native *Clostridium* strains. Anaerobe 32, 77–86. <https://doi.org/10.1016/j.anaerobe.2014.12.008>
- Agaisse, H., Gominet, M., Okstad, O. a, Kolstø, a B., Lereclus, D., 1999. PlcR is a pleiotropic regulator of extracellular virulence factor gene expression in *Bacillus thuringiensis*. Mol. Microbiol. 32, 1043–1053. <https://doi.org/10.1046/j.1365-2958.1999.01419.x>
- Al-Hinai, M.A., Jones, S.W., Papoutsakis, E.T., 2015. The *Clostridium* Sporulation Programs: Diversity and Preservation of Endospore Differentiation. Microbiol. Mol. Biol. Rev. 79, 19–37. <https://doi.org/10.1128/MMBR.00025-14>
- Al-Hinai, M. a, Fast, A.G., Papoutsakis, E.T., 2012. Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. Appl. Environ. Microbiol. 78, 8112–21. <https://doi.org/10.1128/AEM.02214-12>
- Alonso, D.M., Bond, J.Q., Dumesic, J. a, 2010. Catalytic conversion of biomass to biofuels. Green Chem. 12, 1493–1513. <https://doi.org/10.1039/c004654j>
- Alsaker, K., Paredes, C., Papoutsakis, E., 2005. Design, optimization and validation of genomic DNA microarrays for examining the *Clostridium acetobutylicum* transcriptome. Biotechnol. Bioprocess Eng. 10, 432–443. <https://doi.org/10.1007/bf02989826>
- Alsaker, K. V, Papoutsakis, E.T., 2005. Transcriptional program of early sporulation and stationary-Phase Events in *Clostridium acetobutylicum* J. Bacteriol. 187, 7103–7118. <https://doi.org/10.1128/JB.187.20.7103>
- Andersch, W., Bahl, H., Gottschalk, G., 1983. Level of enzymes involved in acetate, butyrate, acetone and butanol formation by *Clostridium acetobutylicum*. Eur. J. Appl. Microbiol. Biotechnol. 18, 327–332. <https://doi.org/10.1007/BF00504740>
- Anetzberger, C., Reiger, M., Fekete, A., Schell, U., Stambrau, N., Plener, L., Kopka, J., Schmitt-Kopplin, P., Hilbi, H., Jung, K., 2012. Autoinducers act as biological timers in *Vibrio harveyi*. PLoS One 7. <https://doi.org/10.1371/journal.pone.0048310>
- Aristilde, L., Lewis, I.A., Park, J.O., Rabinowitz, J.D., 2015. Hierarchy in pentose sugar metabolism in *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 81, 1452–1462. <https://doi.org/10.1128/AEM.03199-14>
- Atkinson, S., Sockett, R.E., Cámara, M., Williams, P., 2006. Quorum sensing and the lifestyle of *Yersinia*. Curr. Issues Mol. Biol. 8, 1–10.
- Baban, S.T., Kuehne, S.A., Barketi-Klai, A., Cartman, S.T., Kelly, M.L., Hardie, K.R., Kansau, I., Collignon, A., Minton, N.P., 2013. The Role of Flagella in *Clostridium difficile* Pathogenesis: Comparison between a Non-Epidemic and an Epidemic Strain. PLoS One 8. <https://doi.org/10.1371/journal.pone.0073026>
- Bahl, H., Andersch, W., Braun, K., Gottschalk, G., 1982. Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium*

- acetobutylicum* grown in continuous culture. Eur. J. Appl. Microbiol. Biotechnol. 14, 17–20. <https://doi.org/10.1007/BF00507998>
- Bahl, H., Andersch, W., Gotschalk, G., 1983. Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a two-stage phosphate limited chemostat. Eur. J. Appl. Microbiol. Biotechnol. 17, 73. <https://doi.org/10.1007/BF00510576>
 - Bahl, H., Müller, H., Behrens, S., Joseph, H., Narberhaus, F., 1995. Expression of heat shock genes in *Clostridium acetobutylicum*. FEMS Microbiol. Rev. 17, 341–348. <https://doi.org/10.1111/j.1574-6976.1995.tb00217.x>
 - Bao, G., Dong, H., Zhu, Y., Mao, S., Zhang, T., Zhang, Y., Chen, Z., Li, Y., 2014. Comparative genomic and proteomic analyses of *Clostridium acetobutylicum* Rh8 and its parent strain DSM 1731 revealed new understandings on butanol tolerance. Biochem. Biophys. Res. Commun. 450, 1612–8. <https://doi.org/10.1016/j.bbrc.2014.07.052>
 - Barnard, A.M., Bowden, S.D., Burr, T., Coulthurst, S.J., Monson, R.E., Salmond, G.P., 2007. Quorum sensing, virulence and secondary metabolite production in plant soft-rotting bacteria. Philos. Trans. R. Soc. B Biol. Sci. 362, 1165–1183. <https://doi.org/10.1098/rstb.2007.2042>
 - Bean, Z., 2016. Role of RNPP and Agr quorum sensing in the regulation of stationary phase phenotypes in *Clostridium acetobutylicum*.
 - Beek, J. ter, Guskov, A., Slotboom, D.J., 2014. Structural diversity of ABC transporters. J. Gen. Physiol. 143, 419–35. <https://doi.org/10.1085/jgp.201411164>
 - Borisova, M., Gaupp, R., Duckworth, A., Schneider, A., Dal??gge, D., M??hleck, M., Deubel, D., Unsleber, S., Yu, W., Muth, G., Bischoff, M., G??tz, F., Mayer, C., 2016. Peptidoglycan recycling in gram-positive bacteria is crucial for survival in stationary phase. MBio 7, 1–10. <https://doi.org/10.1128/mBio.00923-16>
 - Bouillaut, L., Perchat, S., Arold, S., Zorrilla, S., Slamti, L., Henry, C., Gohar, M., Declerck, N., Lereclus, D., 2008. Molecular basis for group-specific activation of the virulence regulator PlcR by PapR heptapeptides. Nucleic Acids Res. 36, 3791–3801. <https://doi.org/10.1093/nar/gkn149>
 - Bruder, M.R., Pyne, M.E., Moo-Young, M., Chung, D.A., Chou, C.P., 2016. Extending CRISPR-Cas9 technology from genome editing to transcriptional engineering in the genus *Clostridium*. Appl. Environ. Microbiol. 82, 6109–6119. <https://doi.org/10.1128/AEM.02128-16>
 - Cabrera, R., Rocha, J., Flores, V., Vázquez-Moreno, L., Guarneros, G., Olmedo, G., Rodríguez-Romero, A., de la Torre, M., 2014. Regulation of sporulation initiation by NprR and its signaling peptide NprRB: molecular recognition and conformational changes. Appl. Microbiol. Biotechnol. 98, 9399–9412. <https://doi.org/10.1007/s00253-014-6094-8>
 - Cai, Y., Chandrangsu, P., Gaballa, A., Helmann, J.D., 2017. Lack of formylated methionyl-tRNA has pleiotropic effects on *Bacillus subtilis*. Microbiology 163, 185–196. <https://doi.org/10.1099/MIC.0.000413>
 - Cartman, S.T., Kelly, M.L., Heeg, D., Heap, J.T., Minton, N.P., 2012. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the tcdC genotype and toxin production. Appl. Environ. Microbiol. 78, 4683–4690. <https://doi.org/10.1128/AEM.00249-12>
 - Cartman, S.T., Minton, N.P., 2010. A mariner-Based transposon system for in vivo random mutagenesis of *Clostridium difficile*. Appl. Environ. Microbiol. 76, 1103–1109. <https://doi.org/10.1128/AEM.02525-09>

- Chen, L., Wu, L., Wang, J., Zhang, W., 2014. Butanol tolerance regulated by a two-component response regulator Slr1037 in photosynthetic *Synechocystis* sp. PCC 6803. *Biotechnol. Biofuels* 7, 89. <https://doi.org/10.1186/1754-6834-7-89>
- Chen, X., Schauder, S., Potier, N., Van Dorsselaer, A., Pelczar, I., Bassler, B.L., Hughson, F.M., 2002. Structural identification of a bacterial quorum-sensing signal containing boron. *Nature* 415, 545–549. <https://doi.org/10.1038/415545a>
- Cook, L.C., Federle, M.J., 2014. Peptide pheromone signaling in *Streptococcus* and *Enterococcus*. *FEMS Microbiol. Rev.* 38, 473–492. <https://doi.org/10.1111/1574-6976.12046>
- Cooksley, C.M., Davis, I.J., Winzer, K., Chan, W.C., Peck, M.W., Minton, N.P., 2010. Regulation of neurotoxin production and sporulation by a putative agrBD signaling system in proteolytic *Clostridium botulinum*. *Appl. Environ. Microbiol.* 76, 4448–4460. <https://doi.org/10.1128/AEM.03038-09>
- Cooksley, C.M., Zhang, Y., Wang, H., Redl, S., Winzer, K., Minton, N.P., 2012. Targeted mutagenesis of the *Clostridium acetobutylicum* acetone-butanol-ethanol fermentation pathway. *Metab. Eng.* 14, 630–41. <https://doi.org/10.1016/j.ymben.2012.09.001>
- Cornillot, E., Nair, R., Papoutsakis, E.T., Soucaille, P., 1997. The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. *J. Bacteriol.* 179.
- Croux, C., Nguyen, N.-P.-T., Lee, J., Raynaud, C., Saint-Prix, F., Gonzalez-Pajuelo, M., Meynial-Salles, I., Soucaille, P., 2016. Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. *Biotechnol. Biofuels* 9, 23. <https://doi.org/10.1186/s13068-016-0432-2>
- Dam, P., Olman, V., Harris, K., Su, Z., Xu, Y., 2007. Operon prediction using both genome-specific and general genomic information. *Nucleic Acids Res.* 35, 288–298. <https://doi.org/10.1093/nar/gkl1018>
- Declerck, N., Bouillaut, L., Chaix, D., Rugani, N., Slamti, L., Hoh, F., Lereclus, D., Arold, S.T., 2007. Structure of PlcR: Insights into virulence regulation and evolution of quorum sensing in Gram-positive bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 104, 18490–5. <https://doi.org/10.1073/pnas.0704501104>
- Desjardins, P.R., Conklin, D.S., 2011. Microvolume quantitation of nucleic acids. *Curr. Protoc. Mol. Biol.* 1–4. <https://doi.org/10.1002/0471142727.mba03js93>
- Diggle, S.P., Gardner, A., West, S.A., Griffin, A.S., 2007. Evolutionary theory of bacterial quorum sensing: when is a signal not a signal? *Philos. Trans. R. Soc. B Biol. Sci.* 362, 1241–1249. <https://doi.org/10.1098/rstb.2007.2049>
- Do, H., Kumaraswami, M., 2016. Structural Mechanisms of Peptide Recognition and Allosteric Modulation of Gene Regulation by the RRNPP Family of Quorum-Sensing Regulators. *J. Mol. Biol.* 428, 2793–2804. <https://doi.org/10.1016/j.jmb.2016.05.026>
- Doeven, M.K., Kok, J., Poolman, B., 2005. Specificity and selectivity determinants of peptide transport in *Lactococcus lactis* and other microorganisms. *Mol. Microbiol.* 57, 640–649. <https://doi.org/10.1111/j.1365-2958.2005.04698.x>
- Drepper, T., Eggert, T., Circolone, F., Heck, A., Krauß, U., Guterl, J.-K., Wendorff, M., Losi, A., Gärtner, W., Jaeger, K.-E., 2007. Reporter proteins for in vivo fluorescence without oxygen. *Nat. Biotechnol.* 25, 443–445. <https://doi.org/10.1038/nbt1293>
- Dubois, T., Perchat, S., Verplaetse, E., Gominet, M., Lemy, C., Aumont-Nicaise, M., Grenha, R., Nessler, S., Lereclus, D., 2013. Activity of the *Bacillus*

- thuringiensis* NprR-NprX cell-cell communication system is co-ordinated to the physiological stage through a complex transcriptional regulation. Mol. Microbiol. <https://doi.org/10.1111/mmi.12168>
- Dürre, P., 2014. Physiology and Sporulation in *Clostridium*. Bact. Spore from Mol. to Syst. 315–329. <https://doi.org/10.1128/microbiolspec.TBS-0010-2012>
 - Dürre, P., 2005. Formation of solvents in *clostridia*, in: Handbook on Clostridia. CRC Press, Boca Raton, Florida, pp. 671–693.
 - Dürre, P., Hollergschwandner, C., 2004. Initiation of endospore formation in *Clostridium acetobutylicum*. Anaerobe 10, 69–74. <https://doi.org/10.1016/j.anaerobe.2003.11.001>
 - Dürre, P., Kuhn, A., Gottwald, M., Gottschalk, G., 1987. Enzymatic investigations on butanol dehydrogenase and butyraldehyde dehydrogenase in extracts of *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. 26, 268–272. <https://doi.org/10.1007/BF00286322>
 - Ehsaan, M., Kuit, W., Zhang, Y., Cartman, S.T., Heap, J.T., Winzer, K., Minton, N.P., 2016. Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. Biotechnol. Biofuels 9, 4. <https://doi.org/10.1186/s13068-015-0410-0>
 - Engebrecht, J., Nealson, K., Silverman, M., 1983. Bacterial bioluminescence: Isolation and genetic analysis of functions from *Vibrio fischeri*. Cell 32, 773–781. [https://doi.org/10.1016/0092-8674\(83\)90063-6](https://doi.org/10.1016/0092-8674(83)90063-6)
 - Errington, J., 2001. Septation and chromosome segregation during sporulation in *Bacillus subtilis*. Curr. Opin. Microbiol. 4, 660–666. [https://doi.org/10.1016/S1369-5274\(01\)00266-1](https://doi.org/10.1016/S1369-5274(01)00266-1)
 - Feustel, L., Nakotte, S., Dürre, P., Du, P., 2004. Characterization and Development of Two Reporter Gene Systems for *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 70, 798–803. <https://doi.org/10.1128/AEM.70.2.798>
 - Gallego del Sol, F., Marina, A., 2013. Structural Basis of Rap Phosphatase Inhibition by Phr Peptides. PLoS Biol. 11. <https://doi.org/10.1371/journal.pbio.1001511>
 - Gamba, P., Hamoen, L.W., Daniel, R.A., 2016. Cooperative recruitment of FtsW to the division Site of *Bacillus subtilis*. Front. Microbiol. 7, 1–9. <https://doi.org/10.3389/fmicb.2016.01808>
 - George, H.A., Johnson, J.L., Moore, W.E.C., 1983. Acetone, isopropanol, and butanol production by *Clostridium beijerinckii* (syn. *Clostridium butylicum*) and *Clostridium aurantibutyricum*. Appl. Environ. Microbiol. 45, 1160–1163.
 - Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C. a, Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345. <https://doi.org/10.1038/nmeth.1318>
 - Girbal, L., Mortier-barrière, I., Rouanet, C., Croux, C., Mortier-barrie, I., Soucaille, P., 2003. Development of a Sensitive Gene Expression Reporter System and an Inducible Promoter-Repressor System for *Clostridium acetobutylicum* Appl. Environ. Microbiol. 69, 4985–4988. <https://doi.org/10.1128/AEM.69.8.4985>
 - Girbal, L., Soucaille, P., 1998. Regulation of solvent production in *Clostridium acetobutylicum*. Trends Biotechnol. 16, 11–16. [https://doi.org/10.1016/S0167-7799\(97\)01141-4](https://doi.org/10.1016/S0167-7799(97)01141-4)
 - Gogliettino, M., Balestrieri, M., Pocsfalvi, G., Fiume, I., Natale, L., Rossi, M., Palmieri, G., 2010. A highly selective oligopeptide binding protein from the archaeon *Sulfolobus solfataricus*. J. Bacteriol. 192, 3123–3131. <https://doi.org/10.1128/JB.01414-09>

- Gominet, M., Slamti, L., Gilois, N., Rose, M., Lereclus, D., 2001. Oligopeptide permease is required for expression of the *Bacillus thuringiensis* plcR regulon and for virulence. *Mol. Microbiol.* 40, 963–975. <https://doi.org/10.1046/j.1365-2958.2001.02440.x>
- González-Peñas, H., Lu-Chau, T. a, Moreira, M.T., Lema, J.M., 2014. Solvent screening methodology for in situ ABE extractive fermentation. *Appl. Microbiol. Biotechnol.* 98, 5915–24. <https://doi.org/10.1007/s00253-014-5634-6>
- Gottwald, M., Hippe, H., Gottschalk, G., 1984. Formation of n-Butanol from D-Glucose by Strains of the “*Clostridium tetanomorphum*” Group. *Appl. Environ. Microbiol.* 48, 573–576.
- Grenha, R., Slamti, L., Nicaise, M., Refes, Y., Lereclus, D., Nessler, S., 2012. Structural basis for the activation mechanism of the PlcR virulence regulator by the quorum-sensing signal peptide PapR. *Proc. Natl. Acad. Sci.* 110, 1047–1052. <https://doi.org/10.1073/pnas.1213770110>
- Grimmmler, C., Janssen, H., Krausse, D., Fischer, R.-J., Bahl, H., Dürre, P., Liebl, W., Ehrenreich, A., 2011. Genome-wide gene expression analysis of the switch between acidogenesis and solventogenesis in continuous cultures of *Clostridium acetobutylicum*. *J. Mol. Microbiol. Biotechnol.* 20, 1–15. <https://doi.org/10.1159/000320973>
- Gutierrez, N. a, Maddox, I.S., 1987. Role of Chemotaxis in Solvent Production by *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* 53, 1924–7.
- Harris, L.M., Welker, N.E., Papoutsakis, E.T., 2002. Northern, morphological, and fermentation analysis of spo0A inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *J. Bacteriol.* 184, 3586–3597. <https://doi.org/10.1128/JB.184.13.3586-3597.2002>
- Heap, J.T., Kuehne, S.A., Ehsaan, M., Cartman, S.T., Cooksley, C.M., Scott, J.C., Minton, N.P., 2010. The ClosTron : Mutagenesis in *Clostridium* refined and streamlined. *J. Microbiol. Methods* 80, 49–55. <https://doi.org/10.1016/j.mimet.2009.10.018>
- Heap, J.T., Pennington, O.J., Cartman, S.T., Carter, G.P., Minton, N.P., 2007. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J. Microbiol. Methods* 70, 452–64. <https://doi.org/10.1016/j.mimet.2007.05.021>
- Heap, J.T., Pennington, O.J., Cartman, S.T., Minton, N.P., 2009. A modular system for *Clostridium* shuttle plasmids. *J. Microbiol. Methods* 78, 79–85. <https://doi.org/10.1016/j.mimet.2009.05.004>
- Henke, J.M., Bassler, B.L., 2004. Bacterial social engagements. *Trends Cell Biol.* 14. <https://doi.org/10.1016/j.tcb.2004.09.012>
- Holt, R.A., Cairns, A.J., Morris, J.G., 1988. Production of butanol by *Clostridium puniceum* in batch and continuous culture. *Appl. Microbiol. Biotechnol.* 27, 319–324. <https://doi.org/10.1007/BF00251761>
- Horneck, G., Bücker, H., Reitz, G., 1994. Long-term survival of bacterial spores in space. *Adv. Sp. Res.* 14, 41–45. [https://doi.org/10.1016/0273-1177\(94\)90448-0](https://doi.org/10.1016/0273-1177(94)90448-0)
- Huang, H., Chai, C., Li, N., Rowe, P., Minton, N.P., Yang, S., Jiang, W., Gu, Y., 2016. CRISPR/Cas9-Based Efficient Genome Editing in *Clostridium ljungdahlii*, an Autotrophic Gas-Fermenting Bacterium. *ACS Synth. Biol.* 5, 1355–1361. <https://doi.org/10.1021/acssynbio.6b00044>
- Huillet, E., Tempelaars, M.H., André-Leroux, G., Wanapaisan, P., Bridoux, L., Makhzami, S., Panbangred, W., Martin-Verstraete, I., Abee, T., Lereclus, D., 2012. PlcRa, a new quorum-sensing regulator from *Bacillus cereus*, plays a role in

- oxidative stress responses and cysteine metabolism in stationary phase. PLoS One 7, e51047. <https://doi.org/10.1371/journal.pone.0051047>
- Jabbari, S., Steiner, E., Heap, J.T., Winzer, K., Minton, N.P., King, J.R., 2013. The putative influence of the agr operon upon survival mechanisms used by *Clostridium acetobutylicum*. Math. Biosci. 243, 223–39. <https://doi.org/10.1016/j.mbs.2013.03.005>
 - Jang, Y.-S., Han, M.-J., Lee, J., Im, J.A., Lee, Y.H., Papoutsakis, E.T., Bennett, G., Lee, S.Y., 2014. Proteomic analyses of the phase transition from acidogenesis to solventogenesis using solventogenic and non-solventogenic *Clostridium acetobutylicum* strains. Appl. Microbiol. Biotechnol. 98, 5105–15. <https://doi.org/10.1007/s00253-014-5738-z>
 - Jang, Y.-S., Lee, J.Y., Lee, J., Park, J.H., Im, J.A., Eom, M.-H., Lee, J., Lee, S.-H., Song, H., Cho, J.-H., Seung, D.Y., Lee, S.Y., 2012. Enhanced butanol production obtained by reinforcing the direct butanol-forming route in *Clostridium acetobutylicum*. MBio 3. <https://doi.org/10.1128/mBio.00314-12>
 - Janssen, H., Döring, C., Ehrenreich, A., Voigt, B., Hecker, M., Bahl, H., Fischer, R.-J., 2010. A proteomic and transcriptional view of acidogenic and solventogenic steady-state cells of *Clostridium acetobutylicum* in a chemostat culture. Appl. Microbiol. Biotechnol. 87, 2209–26. <https://doi.org/10.1007/s00253-010-2741-x>
 - Jiang, M., Tzeng, Y.L., Feher, V.A., Perego, M., Hoch, J.A., 1999. Alanine mutants of the Spo0F response regulator modifying specificity for sensor kinases in sporulation initiation. Mol. Microbiol. 33, 389–395. <https://doi.org/10.1046/j.1365-2958.1999.01481.x>
 - Jin, L., Zhang, H., Chen, L., Yang, C., Yang, S., Jiang, W., Gu, Y., 2014. Combined overexpression of genes involved in pentose phosphate pathway enables enhanced D-xylose utilization by *Clostridium acetobutylicum*. J. Biotechnol. 173, 7–9. <https://doi.org/10.1016/j.jbiotec.2014.01.002>
 - Jones, D., 2001. Applied acetone-butanol fermentation, in: Dürre, P., Bahl, H. (Eds.), Biotechnology and Medical Applications. Wiley-VCH, Verlag GmbH, Weinheim, Germany, pp. 125–168.
 - Jones, D.T., 2014. The Strategic Importance of Butanol for Japan during the WWII: A Case study of the Butanol Fermentation Process in Taiwan and Japan, in: Durre, P. (Ed.), Systems Biology of Clostridium. Imperial College Press, London, pp. 220–272. https://doi.org/10.1142/9781783264414_0009
 - Jones, D.T., Woods, D.R., 1986. Acetone-butanol fermentation revisited. Microbiol. Rev. 50, 484–524. <https://doi.org/3540574>
 - Jones, S.W., Paredes, C.J., Tracy, B., Cheng, N., Sillers, R., Senger, R.S., Papoutsakis, E.T., 2008. The transcriptional program underlying the physiology of *clostridial* sporulation. Genome Biol. 9, R114. <https://doi.org/10.1186/gb-2008-9-7-r114>
 - Karberg, M., Guo, H., Zhong, J., Coon, R., Perutka, J., Lambowitz, a M., 2001. Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. Nat. Biotechnol. 19, 1162–1167. <https://doi.org/10.1038/nbt1201-1162>
 - Karim, Q.N., Logan, R.P., Puels, J., Karnholz, a, Worku, M.L., 1998. Measurement of motility of *Helicobacter pylori*, *Campylobacter jejuni*, and *Escherichia coli* by real time computer tracking using the Hobson BacTracker. J. Clin. Pathol. 51, 623–628. <https://doi.org/10.1136/jcp.51.8.623>
 - Kariminik, A., Baseri-Salehi, M., Kheirkhah, B., 2017. *Pseudomonas aeruginosa* quorum sensing modulates immune responses: An updated review article. Immunol. Lett. 190, 1–6. <https://doi.org/10.1016/j.imlet.2017.07.002>

- Kimple, M.E., Brill, A.L., Pasker, R.L., 2015. Overview of Affinity Tags for Protein Purification 608–616. <https://doi.org/10.1002/0471140864.ps0909s73>
- Klepsch, M.M., Kovermann, M., Löw, C., Balbach, J., Permentier, H.P., Fusetti, F., De Gier, J.W., Slotboom, D.J., Berntsson, R.P.A., 2011. *Escherichia coli* peptide binding protein OppA has a preference for positively charged peptides. J. Mol. Biol. 414, 75–85. <https://doi.org/10.1016/j.jmb.2011.09.043>
- Koide, A., Perego, M., Hoch, J. a., 1999. ScoC regulates peptide transport and sporulation initiation in *Bacillus subtilis*. J. Bacteriol. 181, 4114–4117.
- Kotte, A., 2013. RNPP-type Quorum Sensing in *Clostridium acetobutylicum*. The University of Nottingham.
- Kotte, A., Severn, O.J., Bean, Z., Schwarz, K., Minton, N.P., Winzer, K., 2017. RNPP-type quorum sensing regulates solvent formation and sporulation in *Clostridium acetobutylicum*. bioRxiv Prepr.
- Kovács, K., Willson, B.J., Schwarz, K., Heap, J.T., Jackson, A., Bolam, D.N., Winzer, K., Minton, N.P., 2013. Secretion and assembly of functional mini-cellulosomes from synthetic chromosomal operons in *Clostridium acetobutylicum* ATCC 824. Biotechnol. Biofuels 6, 117. <https://doi.org/10.1186/1754-6834-6-117>
- Kozłowicz, B.K., Bae, T., Dunny, G.M., 2004. *Enterococcus faecalis* pheromone-responsive protein PrgX: Genetic separation of positive autoregulatory functions from those involved in negative regulation of conjugative plasmid transfer. Mol. Microbiol. 54, 520–532. <https://doi.org/10.1111/j.1365-2958.2004.04286.x>
- Kumar, M., Gayen, K., Saini, S., 2013. Role of extracellular cues to trigger the metabolic phase shifting from acidogenesis to solventogenesis in *Clostridium acetobutylicum*. Bioresour. Technol. 138, 55–62. <https://doi.org/10.1016/j.biortech.2013.03.159>
- Kumar, M., Saini, S., Gayen, K., 2014. Elementary mode analysis reveals that *Clostridium acetobutylicum* modulates its metabolic strategy under external stress. Mol. Biosyst. 10, 2090–105. <https://doi.org/10.1039/c4mb00126e>
- LaRock, C.N., Yu, J., Horswill, A.R., Parsek, M.R., Minion, F.C., 2013. Transcriptome Analysis of Acetyl-Homoserine Lactone-Based Quorum Sensing Regulation in *Yersinia pestis*. PLoS One 8, 1–8. <https://doi.org/10.1371/journal.pone.0062337>
- Lazazzera, B. a, Solomon, J.M., Grossman, A.D., 1997. An exported peptide functions intracellularly to contribute to cell density signaling in *B. subtilis*. Cell 89, 917–25. [https://doi.org/10.1016/S0092-8674\(00\)80277-9](https://doi.org/10.1016/S0092-8674(00)80277-9)
- Le, K.Y., Otto, M., 2015. Quorum-sensing regulation in *staphylococci*, an overview. Front. Microbiol. 6, 1–8. <https://doi.org/10.3389/fmicb.2015.01174>
- LeDeaux, J.R., Solomon, J.M., Grossman, A.D., 1997. Analysis of non-polar deletion mutations in the genes of the spoOK (opp) operon of *Bacillus subtilis*. FEMS Microbiol. Lett. 153, 63–69. [https://doi.org/10.1016/S0378-1097\(97\)00232-2](https://doi.org/10.1016/S0378-1097(97)00232-2)
- Lee, S., Lee, J.H., Mitchell, R.J., 2015. Analysis of *Clostridium beijerinckii* NCIMB 8052's transcriptional response to ferulic acid and its application to enhance the strain tolerance. Biotech. Biofuels 8, 68. <https://doi.org/10.1186/s13068-015-0252-9>
- Leggett, M.J., McDonnell, G., Denyer, S.P., Setlow, P., Maillard, J.Y., 2012. Bacterial spore structures and their protective role in biocide resistance. J. Appl. Microbiol. 113, 485–498. <https://doi.org/10.1111/j.1365-2672.2012.05336.x>
- Lemy, C., Buisson, C., Dubois, T., Faegri, K., Nielsen-Ieroux, C., Gohar, M., Jacques, P., Ramarao, N., Kolstø, A., 2012. Necrotrophism Is a Quorum-Sensing-

- Regulated Lifestyle in *Bacillus thuringiensis*. PLoS Pathog. 8.
<https://doi.org/10.1371/journal.ppat.1002629>
- Lereclus, D., Agaisse, H., Grandvalet, C., Salameitou, S., Gominet, M., 2000. Regulation of toxin and virulence gene transcription in *Bacillus thuringiensis*. Int. J. Med. Microbiol. 290, 295–299. [https://doi.org/10.1016/S1438-4221\(00\)80024-7](https://doi.org/10.1016/S1438-4221(00)80024-7)
 - Levdikov, V.M., Blagova, E. V., Brannigan, J.A., Wright, L., Vagin, A.A., Wilkinson, A.J., 2005. The structure of the oligopeptide-binding protein, AppA, from *Bacillus subtilis* in complex with a nonapeptide. J. Mol. Biol. 345, 879–892. <https://doi.org/10.1016/j.jmb.2004.10.089>
 - Li, H., Luo, W., Wang, Q., Yu, X., 2014. Direct fermentation of gelatinized cassava starch to acetone, butanol, and ethanol using *Clostridium acetobutylicum* mutant obtained by atmospheric and room temperature plasma. Appl. Biochem. Biotechnol. 172, 3330–41. <https://doi.org/10.1007/s12010-014-0765-x>
 - Li, J., Chen, J., Vidal, J.E., McClane, B.A., 2011. The Agr-like quorum-sensing system regulates sporulation and production of enterotoxin and beta2 toxin by *Clostridium perfringens* type a non-food-borne human gastrointestinal disease strain F5603. Infect. Immun. 79, 2451–2459. <https://doi.org/10.1128/IAI.00169-11>
 - Li, J., Chen, X., Qi, B., Luo, J., Zhang, Y., Su, Y., Wan, Y., 2014. Efficient production of acetone-butanol-ethanol (ABE) from cassava by a fermentation-pervaporation coupled process. Bioresour. Technol. 169C, 251–257. <https://doi.org/10.1016/j.biortech.2014.06.102>
 - Liew, F., Henstra, A.M., Köpke, M., Winzer, K., Simpson, S.D., Minton, N.P., 2017. Metabolic engineering of *Clostridium autoethanogenum* for selective alcohol production. Metab. Eng. 40, 104–114. <https://doi.org/10.1016/j.ymben.2017.01.007>
 - Liew, F.M., Martin, M.E., Tappel, R.C., Heijstra, B.D., Mihalcea, C., Köpke, M., 2016. Gas Fermentation-A flexible platform for commercial scale production of low-carbon-fuels and chemicals from waste and renewable feedstocks. Front. Microbiol. 7, 694 <https://doi.org/10.3389/fmicb.2016.00694>
 - Liu, J., Guo, T., Shen, X., Xu, J., Wang, J., Wang, Y., Liu, D., Niu, H., Liang, L., Ying, H., 2016. Engineering *Clostridium beijerinckii* with the Cbei_4693 gene knockout for enhanced ferulic acid tolerance. J. Biotechnol. 229, 53–57. <https://doi.org/10.1016/j.jbiotec.2016.04.052>
 - Liu, W., Caffrey, M., 2006. Interactions of tryptophan, tryptophan peptides, and tryptophan alkyl esters at curved membrane interfaces. Biochem. 45, 11713–11726. <https://doi.org/10.1021/bi0608414>
 - Locher, K.P., 2004. Structure and mechanism of ABC transporters. Curr. Opin. Struct. Biol. 14, 426–431. <https://doi.org/10.1016/j.sbi.2004.06.005>
 - Lowy, F., 1998. *Staphylococcus aureus* Infections. N. Engl. J. Med. 339, 520–532. <https://doi.org/10.1056/NEJM199808203390806>
 - Lund, B.M., Brocklehurst, T.F., Wyatt, G.M., 1981. Characterization of Strains of *Clostridium puniceum* sp.nov., a Pink-pigmented, Pectolytic Bacterium. Trans. R. Soc. Trop. Med. Hyg. 17–26.
 - Lütke-Eversloh, T., 2014. Application of new metabolic engineering tools for *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. 98, 5823–37. <https://doi.org/10.1007/s00253-014-5785-5>
 - Lütke-Eversloh, T., Bahl, H., 2011. Metabolic engineering of *Clostridium acetobutylicum*: recent advances to improve butanol production. Curr. Opin. Biotechnol. 22, 634–47. <https://doi.org/10.1016/j.copbio.2011.01.011>

- Lyon, G.J., Novick, R.P., 2004. Peptide signaling in *Staphylococcus aureus* and other Gram-positive bacteria. *Peptides* 25, 1389–1403.
<https://doi.org/10.1016/j.peptides.2003.11.026>
- Maddox, I.S., Steiner, E., Hirsch, S., Wessner, S., Gutierrez, N. a, Gapes, J.R., Schuster, K.C., 2000. The cause of “acid-crash” and “acidogenic fermentations” during the batch acetone-butanol-ethanol (ABE-) fermentation process. *J. Mol. Microbiol. Biotechnol.* 2, 95–100.
- Madihah, M.S., Ariff, A.B., Sahaid, K.M., Suraini, A.A., Karim, M.I.A., 2001. Direct fermentation of gelatinized sago starch to acetone-butanol-ethanol by *Clostridium acetobutylicum*. *World J. Microbiol. Biotechnol.* 17, 567–576.
<https://doi.org/10.1023/A:1012351112351>
- Mairhofer, J., Scharl, T., Marisch, K., Cserjan-Puschmann, M., Striedner, G., 2013. Comparative transcription profiling and in-depth characterization of plasmid-based and plasmid-free *Escherichia coli* expression systems under production conditions. *Appl. Environ. Microbiol.* 79, 3802–3812.
<https://doi.org/10.1128/AEM.00365-13>
- Mann, M., 2006. Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* 7, 952–958. <https://doi.org/10.1038/nrm2067>
- Mao, S., Luo, Y., Bao, G., Zhang, Y., Li, Y., Ma, Y., 2011. Comparative analysis on the membrane proteome of *Clostridium acetobutylicum* wild type strain and its butanol-tolerant mutant. *Mol. Biosyst.* 7, 1660–1677.
<https://doi.org/10.1039/c0mb00330a>
- Maqbool, A., Horler, R.S.P., Muller, A., Wilkinson, A.J., Wilson, K.S., Thomas, G.H., 2015. The substrate-binding protein in bacterial ABC transporters: dissecting roles in the evolution of substrate specificity. *Biochem. Soc. Trans.* 43, 1011–7. <https://doi.org/10.1042/BST20150135>
- Marraffini, L., Sontheimer, E., 2011. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11, 181–190.
<https://doi.org/10.1038/nrg2749.CRISPR>
- Marshall, K.M., Nowaczyk, L., Raphael, B.H., Skinner, G.E., Rukma Reddy, N., 2014. Identification and genetic characterization of *Clostridium botulinum* serotype A strains from commercially pasteurized carrot juice. *Food Microbiol.* 44, 149–55. <https://doi.org/10.1016/j.fm.2014.05.009>
- Martin-Verstraete, I., Peltier, J., Dupuy, B., 2016. The regulatory networks that control *Clostridium difficile* toxin synthesis. *Toxins*. 8, 1–24.
<https://doi.org/10.3390/toxins8050153>
- May, A., Hillmann, F., Riebe, O., Fischer, R.J., Bahl, H., 2004. A rubrerythrin-like oxidative stress protein of *Clostridium acetobutylicum* is encoded by a duplicated gene and identical to the heat shock protein Hsp21. *FEMS Microbiol. Lett.* 238, 249–254. <https://doi.org/10.1016/j.femsle.2004.07.042>
- Meltzer, J.L., 2010. Chaim Weizmann. *Encycl. Br.*
- Menon, V., Rao, M., 2012. Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept. *Prog. Energy Combust. Sci.* 38, 522–550. <https://doi.org/10.1016/j.pecs.2012.02.002>
- Millat, T., Winzer, K., 2017. Mathematical modelling of clostridial acetone-butanol-ethanol fermentation. *Appl. Microbiol. Biotechnol.* 101, 2251–2271.
<https://doi.org/10.1007/s00253-017-8137-4>
- Minton, N.P., Ehsaan, M., Humphreys, C.M., Little, G.T., Baker, J., Henstra, A.M., Liew, F., Kelly, M.L., Sheng, L., Schwarz, K., Zhang, Y., 2016. A roadmap

- for gene system development in *Clostridium*. *Anaerobe* 41, 104–112.
<https://doi.org/10.1016/j.anaerobe.2016.05.011>
- Miyashiro, T., Ruby, E.G., 2013. Shedding light on bioluminescence regulation in *Vibrio fischeri* 84, 795–806. <https://doi.org/10.1111/j.1365-2958.2012.08065.x>.Shedding
 - Mohanraj, S., Kodhaiyolii, S., Rengasamy, M., Pugalenth, V., 2014. Green synthesized iron oxide nanoparticles effect on fermentative hydrogen production by *Clostridium acetobutylicum*. *Appl. Biochem. Biotechnol.* 173, 318–31. <https://doi.org/10.1007/s12010-014-0843-0>
 - Murray, E., Khamri, W., Walker, M.M.M., Eggleton, P., Moran, A.P., Ferris, J.A., Knapp, S., Karim, Q.N., Worku, M., Strong, P., Reid, K.B.M., Thursz, M.R., 2002. Expression of surfactant protein D in the human gastric mucosa and during *Helicobacter pylori* infection. *Innate Immun.* 70, 1481–1487. <https://doi.org/10.1128/IAI.70.3.1481>
 - Naik, S.N., Goud, V. V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: A comprehensive review. *Renew. Sustain. Energy Rev.* 14, 578–597. <https://doi.org/10.1016/j.rser.2009.10.003>
 - Nakayama, J., Cao, Y., Horii, T., Sakuda, S., Akkermans, A.D.L., De Vos, W.M., Nagasawa, H., 2001. Gelatinase biosynthesis-activating pheromone: A peptide lactone that mediates a quorum sensing in *Enterococcus faecalis*. *Mol. Microbiol.* 41, 145–154. <https://doi.org/10.1046/j.1365-2958.2001.02486.x>
 - NCBI Resource Coordinators, 2017. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 45, D12–D17. <https://doi.org/10.1093/nar/gkw1071>
 - Neiditch, M.B., Capodagli, G.C., Prehna, G., Federle, M.J., 2017. Genetics and structural analysis of RRNPP Intercellular Peptide Signaling of Gram-positive Bacteria. *Annu. Rev. Genet.* 51, 311–333. <https://doi.org/10.1146/annurev-genet-120116>
 - Ng, Y.K., Ehsaan, M., Philip, S., Collery, M.M., Janoir, C., Collignon, A., Cartman, S.T., Minton, N.P., 2013. Expanding the Repertoire of Gene Tools for Precise Manipulation of the *Clostridium difficile* Genome: Allelic Exchange Using pyrE Alleles. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0056051>
 - Nguyen, N.-P.-T., Sonja, L., Flitsch, S.K., Schiel-Bengelsdorf, B., Dürre, P., Soucaille, P., 2016. Cap0037, a Novel Global Regulator of *Clostridium acetobutylicum* Metabolism 7, 1–10. <https://doi.org/10.1128/mBio.01218-16>.Editor
 - Ni, Y., Sun, Z., 2009. Recent progress on industrial fermentative production of acetone-butanol-ethanol by *Clostridium acetobutylicum* in China. *Appl. Microbiol. Biotechnol.* 83, 415–23. <https://doi.org/10.1007/s00253-009-2003-y>
 - Nicholson, W.L., 2003. Using Thermal Inactivation Kinetics to Calculate the Probability of Extreme Spore Longevity: Implications for Paleomicrobiology and Lithopanspermia. *Orig. Life Evol. Biosph.* 33, 621–631. <https://doi.org/10.1023/A:1025789032195>
 - Nigam, P.S., Singh, A., 2011. Production of liquid biofuels from renewable resources. *Prog. Energy Combust. Sci.* 37, 52–68. <https://doi.org/10.1016/j.pecs.2010.01.003>
 - Nölling, J., Breton, G., Omelchenko, M. V., Makarova, K.S., Zeng, Q., Gibson, R., Lee, H.M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y.I., Tatusov, R.L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M.J., Bennett, G.N., Koonin, E. V., Smith, D.R., Smith, D.R., 2001. Genome sequence and comparative analysis of

- the solvent-producing bacterium *Clostridium acetobutylicum*. J. Bacteriol. 183, 4823–38. <https://doi.org/10.1128/JB.183.16.4823-4838.2001>
- Novick, R.P., Geisinger, E., 2008. Quorum Sensing in *Staphylococci*. Annu. Rev. Genet. 42, 541–564. <https://doi.org/10.1146/annurev.genet.42.110807.091640>
 - O'Brien, R.W., Morris, J.G., 1971. Oxygen and the Growth and Metabolism of *Clostridium acetobutylicum*. J. Gen. Microbiol. 68, 307–318.
 - Okonechnikov, K., Golosova, O., Fursov, M., Varlamov, A., Vaskin, Y., Efremov, I., German Grehov, O.G., Kandrov, D., Rasputin, K., Syabro, M., Tleukenov, T., 2012. Unipro UGENE: A unified bioinformatics toolkit. Bioinformatics 28, 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>
 - Olson, M.E., Todd, D.A., Schaeffer, C.R., Paharik, A.E., Van Dyke, M.J., Buttner, H., Dunman, P.M., Rohde, H., Cech, N.B., Fey, P.D., Horswill, A.R., 2014. *Staphylococcus epidermidis* agr quorum-sensing system: Signal identification, cross talk, and importance in colonization. J. Bacteriol. 196, 3482–3493. <https://doi.org/10.1128/JB.01882-14>
 - Otto, M., 2004. Quorum-sensing control in *Staphylococci* - A target for antimicrobial drug therapy? FEMS Microbiol. Lett. 241, 135–141. <https://doi.org/10.1016/j.femsle.2004.11.016>
 - Otto, M., Echner, H., Voelter, W., Gotz, F., 2001. Pheromone Cross-Inhibition between *Staphylococcus aureus* and *Staphylococcus epidermidis*. Infect. Immun. 69, 1957–1960. <https://doi.org/10.1128/IAI.69.3.1957>
 - Ow, S.Y., Salim, M., Noirel, J., Evans, C., Rehman, I., Wright, P.C., 2009. iTRAQ Underestimation in Simple and Complex Mixtures: “The Good, the Bad and the Ugly.” J. Proteome Res. 8, 5347–5355. <https://doi.org/10.1021/pr900634c>
 - Papenfort, K., Bassler, B., 2016. Quorum-Sensing Signal-Response Systems in Gram-Negative Bacteria. Nat. Rev. Microbiol. 1343, 113–119. <https://doi.org/10.1111/nyas.12666>
 - Papoutsakis, E.T., 2008. Engineering solventogenic *clostridia*. Curr. Opin. Biotechnol. 19, 420–429. <https://doi.org/10.1016/j.copbio.2008.08.003>
 - Papoutsakis, E.T., Mermelstein, L.D., 1993. In Vivo Methylation in *Escherichia coli* by the *Bacillus subtilis* Phage +3T I Methyltransferase To Protect Plasmids from Restriction upon Transformation of *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. 59, 1077–1081.
 - Paredes, C.J., Rigoutsos, I., Papoutsakis, T., 2004. Transcriptional organization of the *Clostridium acetobutylicum* genome. Nucleic Acids Res. 32, 1973–1981. <https://doi.org/10.1093/nar/gkh509>
 - Parsiegla, G., Reverbel, C., Tardif, C., Driguez, H., Haser, R., 2008. Structures of Mutants of Cellulase Cel48F of *Clostridium cellulolyticum* in Complex with Long Hemithiocellooligosaccharides Give Rise to a New View of the Substrate Pathway during Processive Action. J. Mol. Biol. 375, 499–510. <https://doi.org/10.1016/j.jmb.2007.10.039>
 - Peng, X., Nelson, E.S., Maier, J.L., Demali, K. a, 2015. A Novel Regulator Controls *Clostridium difficile* Sporulation, Motility and Toxin Production 100, 954–971. <https://doi.org/10.1016/B978-0-12-386043-9.00005-0>
 - Perchat, S., Dubois, T., Zouhir, S., Gominet, M., Poncet, S., Lemy, C., Aumont-Nicaise, M., Deutscher, J., Gohar, M., Nessler, S., Lereclus, D., 2011. A cell-cell communication system regulates protease production during sporulation in bacteria of the *Bacillus cereus* group. Mol. Microbiol. 82, 619–633. <https://doi.org/10.1111/j.1365-2958.2011.07839.x>

- Perchat, S., Talagas, A., Poncet, S., Lazar, N., Li de la Sierra-Gallay, I., Gohar, M., Lereclus, D., Nessler, S., 2016a. How Quorum Sensing Connects Sporulation to Necrotrophism in *Bacillus thuringiensis*. PLoS Pathog. 12, 1–21. <https://doi.org/10.1371/journal.ppat.1005779>
- Perchat, S., Talagas, A., Zouhir, S., Poncet, S., Bouillaut, L., Nessler, S., Lereclus, D., 2016b. NprR, a moonlighting quorum sensor shifting from a phosphatase activity to a transcriptional activator. Microb. Cell 3, 573–575. <https://doi.org/10.15698/mic2016.11.542>
- Perego, M., Higgins, C.F., Pearce, S.R., Gallagher, M.P., Hoch, J. a., 1991. The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. Mol. Microbiol. 5, 173–185. <https://doi.org/10.1111/j.1365-2958.1991.tb01838.x>
- Perutka, J., Wang, W., Goerlitz, D., Lambowitz, A.M., 2004. Use of Computer-designed Group II Introns to Disrupt *Escherichia coli* DExH/D-box Protein and DNA Helicase Genes. J. Mol. Biol. 336, 421–439. <https://doi.org/10.1016/j.jmb.2003.12.009>
- Pfromm, P.H., Amanor-Boadu, V., Nelson, R., Vadlani, P., Madl, R., 2010. Bio-butanol vs. bio-ethanol: A technical and economic assessment for corn and switchgrass fermented by yeast or *Clostridium acetobutylicum*. Biomass Bioenergy 34, 515–524. <https://doi.org/10.1016/j.biombioe.2009.12.017>
- Pich, A., Narberhaus, F., Bahl, H., 1990. Induction of heat shock proteins during initiation of solvent formation in *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. 33, 697–704. <https://doi.org/10.1007/BF00604941>
- Poehlein, A., Solano, J.D.M., Flitsch, S.K., Krabben, P., Winzer, K., Reid, S.J., Jones, D.T., Green, E., Minton, N.P., Daniel, R., Dürre, P., 2017. Microbial solvent formation revisited by comparative genome analysis. Biotechnol. Biofuels 10, 58. <https://doi.org/10.1186/s13068-017-0742-z>
- Pogliano, J., Osborne, N., Sharp, M.D., Mello, A.A. De, Perez, A., Sun, Y.L., Pogliano, K., 1999. A vital stain for studying membrane dynamics in bacteria: A novel mechanism controlling septation during *Bacillus subtilis* sporulation. Mol. Microbiol. 31, 1149–1159. <https://doi.org/10.1046/j.1365-2958.1999.01255.x>
- Pottathil, M., Lazazzera, B.A., 2003. The extracellular Phr peptide-Rap phosphatase signaling circuit of *Bacillus subtilis*. Front. Biosci. 8, d32–45.
- Pyne, M.E., Bruder, M., Moo-Young, M., Chung, D. a, Chou, C.P., 2014. Technical guide for genetic advancement of underdeveloped and intractable *Clostridium*. Biotechnol. Adv. 32, 623–41. <https://doi.org/10.1016/j.biotechadv.2014.04.003>
- Queck, S.Y., Jameson-Lee, M., Villaruz, A.E., Bach, T.H.L., Khan, B.A., Sturdevant, D.E., Ricklefs, S.M., Li, M., Otto, M., 2008. RNAIII-Independent Target Gene Control by the agr Quorum-Sensing System: Insight into the Evolution of Virulence Regulation in *Staphylococcus aureus*. Mol. Cell 32, 150–158. <https://doi.org/10.1016/j.molcel.2008.08.005>
- Qureshi, N., Liu, S., Ezeji, T.C., 2013. Cellulosic Butanol Production from Agricultural Biomass and Residues: Recent Advances in Technology, in: Lee, J.W. (Ed.), Advanced Biofuels and Bioproducts. Springer New York, New York, NY, pp. 247–265. https://doi.org/10.1007/978-1-4614-3348-4_15
- Raju, D., Waters, M., Setlow, P., Sarker, M.R., 2006. Investigating the role of small, acid-soluble spore proteins (SASPs) in the resistance of *Clostridium perfringens* spores to heat. BMC Microbiol. 6, 50. <https://doi.org/10.1186/1471-2180-6-50>

- Rauniyar, N., Yates, J.R., 2014. Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics. *J. Proteome Res.* 13, 5293–5309.
<https://doi.org/10.1021/pr500880b>
- Raut, M.P., Couto, N., Pham, T.K., Evans, C., Noirel, J., Wright, P.C., 2016. Quantitative proteomic analysis of the influence of lignin on biofuel production by *Clostridium acetobutylicum* ATCC 824. *Biotechnol. Biofuels* 9, 113.
<https://doi.org/10.1186/s13068-016-0523-0>
- Ravagnani, A., Jennert, K.C.B., Steiner, E., Grünberg, R., Jefferies, J.R., Wilkinson, S.R., Young, D.I., Tidswell, E.C., Brown, D.P., Youngman, P., Gareth Morris, J., Young, M., 2000. Spo0A directly controls the switch from acid to solvent production in solvent-forming *clostridia*. *Mol. Microbiol.* 37, 1172–1185.
<https://doi.org/10.1046/j.1365-2958.2000.02071.x>
- Redfield, R.J., 2002. Is quorum sensing a side effect of diffusion sensing? *Trends Microbiol.* 10, 365–370.
- Ren, C., Gu, Y., Wu, Y., Zhang, W., Yang, C., Yang, S., Jiang, W., 2012. Pleiotropic functions of catabolite control protein CcpA in Butanol-producing *Clostridium acetobutylicum*. *BMC Genomics* 13, 349.
<https://doi.org/10.1186/1471-2164-13-349>
- Rocha-Estrada, J., Aceves-Diez, A.E., Guarneros, G., de la Torre, M., 2010. The RNPP family of quorum-sensing proteins in Gram-positive bacteria. *Appl. Microbiol. Biotechnol.* 87, 913–23. <https://doi.org/10.1007/s00253-010-2651-y>
- Rossetto, O., Pirazzini, M., Montecucco, C., 2014. Botulinum neurotoxins: genetic, structural and mechanistic insights. *Nat. Rev. Microbiol.* 12, 535–549.
<https://doi.org/10.1038/nrmicro3295>
- Rudner, D.Z., LeDeaux, J.R., Ireton, K., Grossman, A.D., 1991. The spo0K locus of *Bacillus subtilis* is homologous to the oligopeptide permease locus and is required for sporulation and competence. *J. Bacteriol.* 173, 1388–1398.
- Ryan, R.P., An, S. qi, Allan, J.H., McCarthy, Y., Dow, J.M., 2015. The DSF Family of Cell–Cell Signals: An Expanding Class of Bacterial Virulence Regulators. *PLoS Pathog.* 11, 1–14. <https://doi.org/10.1371/journal.ppat.1004986>
- Sanchez-Contreras, M., Bauer, W.D., Gao, M., Robinson, J.B., Allan Downie, J., 2007. Quorum-sensing regulation in *rhizobia* and its role in symbiotic interactions with legumes. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 1149–1163.
<https://doi.org/10.1098/rstb.2007.2041>
- Santangelo, J., Kuhn, A., Treuner-Lange, A., Dürre, P., 1998. Sporulation and time course expression of σ -factor homologous genes in *Clostridium acetobutylicum*. *FEMS Microbiol. Lett.* 161, 157–164.
- Sauer, U., Durre, P., 1995. Differential induction of genes related to solvent formation during the shift from acidogenesis to solventogenesis in continuous culture of *Clostridium acetobutylicum*. *FEMS Microbiol. Lett.* 125, 115–120.
[https://doi.org/10.1016/0378-1097\(94\)00483-8](https://doi.org/10.1016/0378-1097(94)00483-8)
- Schuchmann, K., Müller, V., 2014. Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* 12, 809–821. <https://doi.org/10.1038/nrmicro3365>
- Schwarz, K.M., Grosse-Honebrink, A., Derecka, K., Rotta, C., Zhang, Y., Minton, N.P., 2017. Towards improved butanol production through targeted genetic modification of *Clostridium pasteurianum*. *Metab. Eng.* 40, 124–137.
<https://doi.org/10.1016/j.ymben.2017.01.009>
- Shao, L., Hu, S., Yang, Y., Gu, Y., Chen, J., Yang, Y., Jiang, W., Yang, S., 2007. Targeted gene disruption by use of a group II intron (targetron) vector in

- Clostridium acetobutylicum*. Cell Res. 17, 963–965.
<https://doi.org/10.1038/cr.2007.91>
- Shi, K., Brown, C.K., Gu, Z., Kozłowicz, B.K., Dunny, G.M., Ohlendorf, D.H., Earhart, C.A., 2005. Structure of peptide sex pheromone receptor PrgX and PrgX⁺ pheromone complexes and regulation of conjugation in *Enterococcus faecalis*. Proc. Natl. Acad. Sci. 102, 18596–18601.
 - Shindo, Y., Dobashi, Y., Sakai, T., Monma, C., Miyatani, H., Yoshida, Y., 2015. Epidemiological and pathobiological profiles of *Clostridium perfringens* infections: Review of consecutive series of 33 cases over a 13-year period. Int. J. Clin. Exp. Pathol. 8, 569–577.
 - Shonnard, D.R., Taylor, R.T., Tompson, A., Knapp, R.B., 1992. Hydrodynamic effects on microcapillary motility and chemotaxis assays of *Methylosinus trichosporium* OB3b. Appl. Environ. Microbiol. 58, 2737–2743.
 - Sivagnanam, K., Raghavan, V., Shah, M., Hettich, R., Verberkmoes, N., Lefsrund, M., 2013. Characterization of *Clostridium acetobutylicum* Protein Interaction Network from Butanol Fermentation. J. Anal. Bioanal. Tech. 4.
<https://doi.org/10.4172/2155-9872.S3-002>
 - Sivagnanam, K., Raghavan, V.G., Shah, M., Hettich, R.L., Verberkmoes, N.C., Lefsrud, M.G., 2011. Comparative shotgun proteomic analysis of *Clostridium acetobutylicum* from butanol fermentation using glucose and xylose. Proteome Sci. 9, 66. <https://doi.org/10.1186/1477-5956-9-66>
 - Slamti, L., Perchat, S., Huillet, E., Lereclus, D., 2014. Quorum Sensing in *Bacillus thuringiensis* Is Required for Completion of a Full Infectious Cycle in the Insect. Toxins (Basel). 6, 2239–55. <https://doi.org/10.3390/toxins6082239>
 - Sleight, S.H., Seavers, P.R., Wilkinson, A.J., Ladbury, J.E., Tame, J.R., 1999. Crystallographic and calorimetric analysis of peptide binding to OppA protein. J. Mol. Biol. 291, 393–415. <https://doi.org/10.1006/jmbi.1999.2929>
 - Solomon, J., Su, L., Shyn, S., Grossman, A.D., 2003. Isolation and Characterization of Mutants of the *Bacillus subtilis* Oligopeptide Permease with Altered Specificity of Oligopeptide Transport. J. Bacteriol. 185, 6425–6433. <https://doi.org/10.1128/JB.185.21.6425-6433.2003>
 - Steiner, E., Dago, A.E., Young, D.I., Heap, J.T., Minton, N.P., Hoch, J. a, Young, M., 2011. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. Mol. Microbiol. 80, 641–54. <https://doi.org/10.1111/j.1365-2958.2011.07608.x>
 - Steiner, E., Scott, J., Minton, N.P., Winzer, K., 2012. An agr quorum sensing system that regulates granule formation and sporulation in *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 78, 1113–22. <https://doi.org/10.1128/AEM.06376-11>
 - Strickmann, D.B., Küng, A., Keppler, B.K., 2002. Changes in protein synthesis and identification of proteins specifically induced during solventogenesis in *Clostridium acetobutylicum*. Electrophoresis 23, 110–121. [https://doi.org/10.1002/1522-2683\(200201\)23:1<110::AID-ELPS110>3.0.CO;2-G](https://doi.org/10.1002/1522-2683(200201)23:1<110::AID-ELPS110>3.0.CO;2-G)
 - Sturme, M.H.J., Kleerebezem, M., Nakayama, J., Akkermans, a D.L., Vaughan, E.E., de Vos, W.M., 2002. Cell to cell communication by autoinducing peptides in gram- positive bacteria. Int. J. Gen. Mol. Microbiol. 81, 233–243. <https://doi.org/10.1023/A:1020522919555>
 - Sullivan, L., Bennett, G.N., 2006. Proteome analysis and comparison of *Clostridium acetobutylicum* ATCC 824 and Spo0A strain variants. J. Ind. Microbiol. Biotechnol. 33, 298–308. <https://doi.org/10.1007/s10295-005-0050-7>

- Taga, M.E., Bassler, B.L., 2003. Chemical communication among bacteria. Proc. Natl. Acad. Sci. U. S. A. 100 Suppl, 14549–14554. <https://doi.org/10.1073/pnas.1934514100>
- Tal-Gan, Y., Stacy, D.M., Foegen, M.K., Koenig, D.W., Blackwell, H.E., 2013. Highly potent inhibitors of quorum sensing in *Staphylococcus aureus* revealed through a systematic synthetic study of the group-III autoinducing peptide. J. Am. Chem. Soc. 135, 7869–82. <https://doi.org/10.1021/ja3112115>
- Tenenbaum, D.J., 2008. Food vs. fuel diversion of crops could cause more hunger. Environ. Health Perspect. 116, 254–257. <https://doi.org/10.1289/ehp.116-a254>
- Thomas, G.H., 2010. Homes for the orphans: Utilization of multiple substrate-binding proteins by ABC transporters: MicroCommentary. Mol. Microbiol. 75, 6–9. <https://doi.org/10.1111/j.1365-2958.2009.06961.x>
- Thormann, K., Feustel, L., Lorenz, K., Dürre, P., Nakotte, S., 2002. Control of butanol formation in *Clostridium acetobutylicum* by transcriptional activation. J. Bacteriol. 184, 1966–1973. <https://doi.org/10.1128/JB.184.7.1966>
- Tomas, C.A., Welker, N.E., Eleftherios, T., Tomas, C.A., Welker, N.E., Papoutsakis, E.T., 2003. Overexpression of groESL in *Clostridium acetobutylicum* Results in Increased Solvent Production and Tolerance , Prolonged Metabolism , and Changes in the Cells Transcriptional Program. Inc. Appl. Environ. Microbiol. 69, 4951–4965. <https://doi.org/10.1128/AEM.69.8.4951>
- Tomas, C.A., Welker, N.E., Papoutsakis, E.T., 2003. Overexpression of groESL in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. Appl. Environ. Microbiol. 69, 4951–4965. <https://doi.org/10.1128/AEM.69.8.4951-4965.2003>
- Tomii, K., Kanehisa, M., 1998. A comparative analysis of ABC transporters in complete microbial genomes. Genome Res. 8, 1048–1059. <https://doi.org/10.1101/gr.8.10.1048>
- Tracy, B.P., Jones, S.W., Fast, A.G., Indurthi, D.C., Papoutsakis, E.T., 2012. *Clostridia*: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. Curr. Opin. Biotechnol. 23, 364–81. <https://doi.org/10.1016/j.copbio.2011.10.008>
- Tummala, S.B., Welker, N.E., Papoutsakis, E.T., 1999. Development and characterization of a gene expression reporter system for *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. 65, 3793–3799.
- Turner, K.H., Wessel, A.K., Palmer, G.C., Murray, J.L., Whiteley, M., 2015. Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. Proc. Natl. Acad. Sci. U. S. A. 112, 4110–5. <https://doi.org/10.1073/pnas.1419677112>
- Varahan, S., Harms, N., Gilmore, M.S., Tomich, J.M., Hancock, L.E., 2014. An ABC transporter is required for secretion of peptide sex pheromones in *Enterococcus faecalis*. MBio 5, 1–9. <https://doi.org/10.1128/mBio.01726-14>
- Venkataramanan, K.P., Min, L., Hou, S., Jones, S.W., Ralston, M.T., Lee, K.H., Papoutsakis, E.T., 2015. Complex and extensive post-transcriptional regulation revealed by integrative proteomic and transcriptomic analysis of metabolite stress response in *Clostridium acetobutylicum*. Biotechnol. Biofuels 8, 81. <https://doi.org/10.1186/s13068-015-0260-9>
- Wallace, B.D., Roberts, A.B., Pollet, R.M., Ingle, J.D., Biernat, K.A., Pellock, S.J., Venkatesh, M.K., Guthrie, L., O'Neal, S.K., Robinson, S.J., Dollinger, M., Figueroa, E., McShane, S.R., Cohen, R.D., Jin, J., Frye, S. V., Zamboni, W.C., Pepe-Ranney, C., Mani, S., Kelly, L., Redinbo, M.R., 2015. Structure and

- Inhibition of Microbiome β -Glucuronidases Essential to the Alleviation of Cancer Drug Toxicity. *Chem. Biol.* 22, 1238–1249.
<https://doi.org/10.1016/j.chembiol.2015.08.005>
- Wang, S., Zhang, Y., Dong, H., Mao, S., Zhu, Y., Wang, R., Luan, G., Li, Y., 2011. Formic acid triggers the “Acid Crash” of acetone-butanol-ethanol fermentation by *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* 77, 1674–80. <https://doi.org/10.1128/AEM.01835-10>
 - Wang, X.-G., Kidder, J.M., Scagliotti, J.P., Klempner, M.S., Noring, R., Hu, L.T., 2004. Analysis of differences in the functional properties of the substrate binding proteins of the *Borrelia burgdorferi* oligopeptide permease (Opp) operon. *J. Bacteriol.* 186, 51–60. <https://doi.org/10.1128/JB.186.1.51-60.2004>
 - Wang, Y., Zhang, Z.T., Seo, S.O., Lynn, P., Lu, T., Jin, Y.S., Blaschek, H.P., 2016. Bacterial Genome Editing with CRISPR-Cas9: Deletion, Integration, Single Nucleotide Modification, and Desirable “clean” Mutant Selection in *Clostridium beijerinckii* as an Example. *ACS Synth. Biol.* 5, 721–732.
<https://doi.org/10.1021/acssynbio.6b00060>
 - Weizmann, C., 1915. Improvements in the Bacterial Fermentation of Carbohydrates and in Bacterial Cultures for the same. 4845.
 - West, S.A., Winzer, K., Gardner, A., Diggle, S.P., 2012. Quorum sensing and the confusion about diffusion. *Trends Microbiol.* 20, 586–594.
<https://doi.org/10.1016/j.tim.2012.09.004>
 - Wetzel, D., Fischer, R.J., 2015. Small acid-soluble spore proteins of *Clostridium acetobutylicum* are able to protect DNA in vitro and are specifically cleaved by germination protease GPR and spore protease YyaC. *Microbiol.* 161, 2098–2109.
<https://doi.org/10.1099/mic.0.000162>
 - Willson, B.J., Kovács, K., Steele, T.W., Markus, R., Winzer, K., Minton, N.P., 2016. Production of a functional cell wall - anchored minicellulosome by recombinant *Clostridium acetobutylicum* ATCC 824. *Biotechnol. Biofuels* 9, 1–22. <https://doi.org/10.1186/s13068-016-0526-x>
 - Winterberg, W.S.R., M., K., 2008. Transposon-Based Strategies for the Identification of Essential Bacterial Genes William, in: *Microbial Gene Essentiality: Protocols and Bioinformatics*. pp. 13–26.
<https://doi.org/10.1038/224747b0>
 - Winzer, K., Hardie, K.R., Burgess, N., Doherty, N., Kirke, D., Holden, M.T., Linforth, R., Cornell, K.A., Taylor, A.J., Hill, P.J., Williams, P., 2002. LuxS: its role in central metabolism and the in vitro synthesis of 4-hydroxy-5-methyl-3(2H)-furanone. *Microbiology* 148, 909–922. <https://doi.org/10.1099/00221287-148-4-909>
 - Winzer, K., Lorenz, K., Dürre, P., 1997. Acetate kinase from *Clostridium acetobutylicum*: A highly specific enzyme that is actively transcribed during acidogenesis and solventogenesis. *Microbiology* 143, 3279–3286.
<https://doi.org/10.1099/00221287-143-10-3279>
 - Xue, C., Zhao, J., Chen, L., Yang, S.T., Bai, F., 2017. Recent advances and state-of-the-art strategies in strain and process engineering for biobutanol production by *Clostridium acetobutylicum*. *Biotechnol. Adv.* 35, 310–322.
<https://doi.org/10.1016/j.biotechadv.2017.01.007>
 - Yoo, M., Bestel-Corre, G., Croux, C., Riviere, A., Meynial-Salles, I., Soucaille, P., 2015. A quantitative system-scale characterization of the metabolism of

Clostridium acetobutylicum. MBio 6, 1–12. <https://doi.org/10.1128/mBio.01808-15>

- Young, C.L., Britton, Z.T., Robinson, A.S., 2012. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. Biotechnol. J. 7, 620–634. <https://doi.org/10.1002/biot.201100155>
- Yung, P.T., Ponce, A., 2008. Fast sterility assessment by germinable-endospore biodosimetry. Appl. Environ. Microbiol. 74, 7669–7674. <https://doi.org/10.1128/AEM.01437-08>
- Zhang, Y., Grosse-Honebrink, A., Minton, N.P., 2015. A universal mariner transposon system for forward genetic studies in the genus *clostridium*. PLoS One 10, 1–21. <https://doi.org/10.1371/journal.pone.0122411>
- Zhong, J., Karberg, M., 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 Res. 31, 1656–1664. <https://doi.org/10.1093/nar/gkg248>
- Zouhir, S., Perchat, S., Nicaise, M., Perez, J., Guimaraes, B., Lereclus, D., Nessler, S., 2013. Peptide-binding dependent conformational changes regulate the transcriptional activity of the quorum-sensor NprR. Nucleic Acids Res. 41. <https://doi.org/10.1093/nar/gkt546>

8. Appendices

Appendix

Appendix Table 1: iTRAQ Proteomic analysis of *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* vs *C. acetobutylicum* ATCC 824 pMTL85143. Identified RNPP type regulators within the proteomics data set are highlighted in turquoise, proteins of interest investigated during enzyme or GusA assays are highlighted in green, where they don't fall into the latter category.

UniProt ID	Protein names	Gene name and loci	Fold Change	p-value
Q97N21	Serine--tRNA ligase 2	<i>CAserS2</i> CA_C0017	2.11	1.58E-07
Q97GB7	Flavodoxin	CA_C2452	1.99	2.26E-04
Q97MY3	Uncharacterized protein	CA_C0057	1.98	4.25E-06
Q97N24	Aminotransferase	CA_C0014	1.93	3.05E-12
Q97FP4	Sugar kinase, ribokinase family	CA_C2684	1.91	4.88E-05
Q97LJ0	PTS enzyme II, ABC component	CA_C0570	1.70	1.45E-05
Q97GP2	Glycosyltransferase	CA_C2324	1.62	1.63E-07
Q97KD3	Lipoprotein	CA_C0986	1.60	7.77E-07
Q97JD3	Phosphotransferase system IIC component	CA_C1353	1.56	7.68E-06
Q97CX5	Hypothetical secreted protein	CA_C3721	1.56	7.52E-12
Q97KN7	Periplasmic amino acid binding protein	CA_C0880	1.56	1.14E-12
Q97JG7	Uncharacterized protein	CA_C1318	1.55	4.79E-05
P58255	Phosphate butyryltransferase	<i>ptb</i> CA_C3076	1.55	1.75E-35
P45359	Acetyl-CoA acetyltransferase	<i>thlA</i> thl CA_C2873	1.55	2.93E-54
Q97LX2	Methyl-accepting chemotaxis protein	CA_C0432	1.55	2.38E-06
Q97GY0	Cysteine synthase	<i>cysK</i> CA_C2235	1.52	7.70E-25
Q7D462	Flagellin	<i>hag</i> CA_C2203	1.51	6.28E-17

Q97FF7	O-acetylhomoserine sulfhydrylase	<i>cysD</i> CA_C2783	1.51	1.11E-10
Q97G09	NifU-related domain containing protein	CA_C2565	1.50	4.08E-31
Q97GI6	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase	<i>dapH</i> CA_C2381	1.49	6.70E-09
Q97GK6	50S ribosomal protein L35	<i>rpmI</i> CA_C2360	1.47	4.22E-05
Q97F73	Uracil phosphoribosyltransferase	<i>upp</i> CA_C2879	1.47	5.15E-09
Q97LK4	Uncharacterised conserved protein	CA_C0556	1.47	7.39E-05
Q97MD1	L-lactate dehydrogenase 1	<i>ldh1</i> ldh CA_C0267	1.46	6.03E-04
Q9Z686	ATP synthase epsilon chain	<i>atpC</i> CA_C2864	1.45	4.05E-05
Q97G10	Uncharacterized protein	CA_C2564	1.45	3.15E-28
Q97GK7	50S ribosomal protein L20	<i>rplT</i> CA_C2359	1.45	4.51E-06
Q97IP0	Methyl-accepting chemotaxis-like protein	CA_C1600	1.44	8.77E-12
Q97K67	Membrane protease subunit, stomatin/prohibitin homolog	CA_C1052	1.44	1.20E-12
Q97DV0	2-enoate reductase	CA_C3371	1.44	1.36E-08
P52040	Electron transfer flavoprotein subunit beta	<i>etfB</i> CA_C2710	1.42	2.49E-29
Q97DC6	L-lactate dehydrogenase 2	<i>ldh2</i> CA_C3552	1.42	9.30E-06
Q97EK4	30S ribosomal protein S11	<i>rpsK</i> CA_C3106	1.40	9.59E-13
Q97EB3	Formate--tetrahydrofolate ligase	<i>fhs</i> CA_C3201	1.40	2.10E-23
Q97E27	Iron-regulated ABC-type transporter membrane component	CA_C3289	1.40	1.84E-08
Q97L52	Enolase	<i>eno</i> CA_C0713	1.39	3.97E-09
Q97ID4	L-asparaginase	<i>ansA</i> CA_C1714	1.38	2.11E-03
Q97DZ5	Periplasmic amino acid binding protein	CA_C3325	1.38	6.89E-05
Q97KC6	D-alanyl-D-alanine carboxypeptidase (Penicilin binding protein)	<i>dacF</i> CA_C0993	1.38	1.70E-04
Q97H59	Flagellar basal body protein	<i>flgE</i> CA_C2154	1.38	6.95E-07
Q97E92	Bifunctional protein GlmU	<i>glmU</i> CA_C3222	1.37	2.22E-09

Q97IR1	Methyl-accepting chemotaxis-like protein	CA_C1579	1.37	1.53E-03
Q97MX5	Uncharacterized protein	CA_C0065	1.37	2.78E-03
Q97GN1	UTP--glucose-1-phosphate uridylyltransferase	CA_C2335	1.37	1.12E-06
Q97IN6	Methylglyoxal synthase	<i>mgsA</i> CA_C1604	1.36	9.62E-04
Q97J86	ABC transporter, periplasmic binding component	CA_C1400	1.35	1.00E-03
Q97E28	Iron-regulated ABC transporter ATPase subunit	CA_C3288	1.35	8.58E-09
Q97F61	CTP synthase	<i>pyrG</i> CA_C2892	1.35	2.21E-07
Q97J94	Amidophosphoribosyltransferase	<i>purF</i> CA_C1392	1.35	2.51E-07
Q97EK3	30S ribosomal protein S13	<i>rpsM</i> CA_C3107	1.35	4.32E-05
Q97D48	Oligopeptide ABC transporter, SBP component	<i>oppA</i> CA_C3632	1.34	2.25E-16
Q97D83	Reverse rubrerythrin-2	<i>rbr3B</i> <i>hsp21</i> <i>rpr2</i> CA_C3597	1.34	1.11E-06
Q97GZ8	Chemotaxis signal transduction protein CheW	<i>cheW</i> CA_C2217	1.33	5.19E-04
Q97KB7	Nicotinate phosphoribosyltransferase	CA_C1002	1.33	1.52E-03
Q97M93	Chemotaxis motility protein B, gene <i>motB</i>	<i>motB</i> CA_C0305	1.33	2.47E-04
P30717	60 kDa chaperonin	<i>groL</i> <i>groEL</i> <i>mopA</i> CA_C2703	1.33	1.18E-34
Q97EG7	50S ribosomal protein L10	<i>rplJ</i> CA_C3146	1.33	1.09E-08
Q97EG8	50S ribosomal protein L7/L12	<i>rplL</i> CA_C3145	1.33	3.94E-22
Q97D85	Uncharacterized FAD-dependent dehydrogenase	CA_C3595	1.32	1.68E-03
Q97J07	Branched-chain-amino-acid transaminase	<i>ilvE</i> CA_C1479	1.32	1.12E-14
Q97GY6	Pyruvate-flavodoxin oxidoreductase	CA_C2229	1.31	7.23E-80
Q97I64	Uridylate kinase	<i>pyrH</i> CA_C1789	1.31	3.29E-03
O52633	Triosephosphate isomerase	<i>tpiA</i> <i>tpi</i> CA_C0711	1.31	8.51E-15

Q97CX4	30S ribosomal protein S18	<i>rpsR</i> CA_C3722	1.31	4.12E-08
Q97EB7	Lysine--tRNA ligase	<i>lysS</i> CA_C3197	1.31	8.93E-12
Q97EL3	30S ribosomal protein S9	<i>rpsI</i> CA_C3097	1.31	1.21E-06
Q97IW7	Fructokinase	CA_C1523	1.30	1.26E-05
Q97KK8	Methyl-accepting chemotaxis protein	CA_C0909	1.30	1.54E-06
Q97CX2	30S ribosomal protein S6	<i>rpsF</i> CA_C3724	1.29	1.67E-08
Q97KP0	Cyclopropane fatty acid synthase	<i>cfa</i> CA_C0877	1.29	7.27E-11
Q97MU7	Ferredoxin-nitrite reductase	CA_C0094	1.29	1.09E-04
Q97LX1	Methyl-accepting chemotaxis protein	CA_C0433	1.29	1.52E-03
Q97EH3	30S ribosomal protein S7	<i>rpsG</i> CA_C3139	1.29	1.76E-13
Q9Z687	ATP synthase subunit beta	<i>atpD</i> CA_C2865	1.29	1.24E-16
Q97EG6	50S ribosomal protein L1	<i>rplA</i> CA_C3147	1.28	4.08E-14
Q97EI1	50S ribosomal protein L2	<i>rplB</i> CA_C3130	1.28	4.02E-18
Q97EJ7	50S ribosomal protein L15	<i>rplO</i> CA_C3114	1.28	1.09E-05
P52046	Short-chain-enoyl-CoA hydratase	<i>crt</i> CA_C2712	1.27	1.17E-17
Q97LS0	Phosphoglucosamine mutase	<i>glmM</i> CA_C0484	1.27	3.21E-06
P52039	Electron transfer flavoprotein subunit alpha	<i>etfA</i> CA_C2709	1.27	2.39E-15
Q97E26	Iron-regulated ABC-type transporter membrane component	CA_C3290	1.27	5.04E-03
Q97JL7	50S ribosomal protein L21	<i>rplU</i> CA_C1257	1.27	2.47E-03
Q97FT6	Trigger factor	<i>tig</i> CA_C2641	1.27	1.53E-19
Q97HY8	Uncharacterized secreted protein, homolog YXKC <i>Bacillus subtilis</i>	CA_C1868	1.27	1.17E-05
Q97JJ2	30S ribosomal protein S21	<i>rpsU</i> CA_C1288	1.26	2.98E-05
Q97KH5	Transketolase	<i>tkt</i> CA_C0944	1.26	3.32E-05
Q97HC7	Uncharacterized protein from alkaline shock protein family	CA_C2085	1.26	1.12E-03

Q7D472	Hydrogene dehydrogenase	<i>hydA</i> CA_C0028	1.26	5.68E-12
Q97LS8	Peptidase T	<i>pepT</i> CA_C0476	1.26	3.48E-05
Q97EH8	50S ribosomal protein L3	<i>rplC</i> CA_C3133	1.26	5.97E-12
Q97F58	D-alanine--D-alanine ligase	<i>ddl</i> CA_C2895	1.26	1.33E-04
Q97M74	TPR repeats containing protein	CA_C0324	1.25	6.31E-04
Q97J91	Bifunctional purine biosynthesis protein PurH	<i>purH</i> CA_C1395	1.25	2.73E-09
Q97F85	S-adenosylmethionine synthase	<i>metK</i> CA_C2856	1.25	5.66E-07
Q97EM3	Phosphoenolpyruvate-protein phosphotransferase	CA_C3087	1.25	3.43E-05
O08309	Pyruvate kinase	<i>pyk pykA</i>	1.25	1.80E-13
		CA_C0518		
Q97GG8	Valine--tRNA ligase	<i>valS</i> CA_C2399	1.24	1.69E-05
Q45829	Butyrate kinase 1	<i>bukI buk</i>	1.24	6.10E-09
		CA_C3075		
Q97I09	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	<i>ispH</i> CA_C1847	1.24	2.38E-07
Q97GQ1	dTDP-4-dehydrorhamnose reductase	CA_C2315	1.24	1.08E-04
Q97EI7	30S ribosomal protein S17	<i>rpsQ</i> CA_C3124	1.24	1.43E-11
Q97ED5	Proline--tRNA ligase	<i>proS</i> CA_C3178	1.23	7.57E-04
P52042	Acyl-CoA dehydrogenase, short-chain specific	<i>bcd</i> CA_C2711	1.23	5.69E-09
Q97FS3	Aspartate carbamoyltransferase	<i>pyrB</i> CA_C2654	1.23	1.98E-03
Q97K83	Pyruvate kinase	<i>pykA</i> CA_C1036	1.23	1.08E-10
Q97LQ5	Carboxyl-terminal protease	CA_C0499	1.22	2.10E-03
P45360	Putative UDP-N-acetylglucosamine 2-epimerase	CA_C2874	1.22	1.15E-05
Q97EJ4	50S ribosomal protein L18	<i>rplR</i> CA_C3117	1.22	3.17E-04
Q97LG7	Pyridoxal 5'-phosphate synthase subunit PdxS	<i>pdxS</i> CA_C0594	1.22	1.80E-06

Q97EI9	50S ribosomal protein L24	<i>rplX</i> CA_C3122	1.21	3.05E-07
Q9Z689	ATP synthase subunit alpha	<i>atpA</i> CA_C2867	1.21	1.79E-05
Q97EJ5	30S ribosomal protein S5	<i>rpsE</i> CA_C3116	1.21	1.71E-08
Q97M85	Phage shock protein A	CA_C0313	1.20	1.62E-04
Q97MD0	ABC transporter ATP-binding protein	CA_C0268	1.20	8.59E-04
Q97K92	Flavo-diiron protein FprA1	<i>fprA1</i> CA_C1027	1.20	2.02E-08
Q97EI5	50S ribosomal protein L16	<i>rplP</i> CA_C3126	1.20	3.37E-03
P30726	Protein GrpE	<i>grpE</i> CA_C1281	1.20	2.05E-04
Q97EJ3	50S ribosomal protein L6	<i>rplF</i> CA_C3118	1.20	1.05E-05
Q97ES0	Isoleucine--tRNA ligase	<i>ileS</i> CA_C3038	1.19	6.33E-05
Q97LB6	Leucine--tRNA ligase	<i>leuS</i> CA_C0646	1.19	7.24E-04
Q97I38	Ribonuclease Y	<i>rny</i> CA_C1816	1.18	1.82E-04
Q97FP8	Glucose-6-phosphate isomerase	<i>pgi</i> CA_C2680	1.18	1.60E-06
Q97IG3	Alanine--tRNA ligase	<i>alaS</i> CA_C1678	1.18	2.36E-06
Q97FJ7	Histidine--tRNA ligase	<i>hisS</i> CA_C2740	1.18	1.38E-03
Q97GK9	Phenylalanine--tRNA ligase alpha subunit	<i>pheS</i> CA_C2357	1.18	1.79E-03
Q97H95	Cell division protein DivIVA	CA_C2118	1.18	1.82E-03
Q97I33	Adenylosuccinate lyase	<i>purB</i> CA_C1821	1.18	1.06E-04
Q97GL0	Phenylalanine--tRNA ligase beta subunit	<i>pheT</i> CA_C2356	1.17	1.89E-06
Q97FM8	Inosine-5'-monophosphate dehydrogenase	<i>guaB</i> CA_C2701	1.17	4.68E-08
Q97J93	Phosphoribosylformylglycinamide cyclo-ligase	<i>purM</i> CA_C1393	1.16	6.68E-06
Q97EB2	ATP-dependent zinc metalloprotease FtsH	<i>ftsH</i> CA_C3202	1.16	8.34E-05
Q97I46	30S ribosomal protein S15	<i>rpsO</i> CA_C1807	1.16	1.01E-03
Q9Z688	ATP synthase gamma chain	<i>atpG</i> CA_C2866	1.15	1.52E-03

Q97EK5	30S ribosomal protein S4 A	<i>rspD1</i> CA_C3105	1.15	1.34E-04
Q97I93	50S ribosomal protein L19	<i>rplS</i> CA_C1759	1.15	2.96E-03
P94604	DNA gyrase subunit B	<i>gyrB</i> CA_C0006	1.15	3.02E-04
Q97EJ0	50S ribosomal protein L5	<i>rplE</i> CA_C3121	1.14	8.76E-08
Q97EL2	50S ribosomal protein L13	<i>rplM</i> CA_C3098	1.14	4.73E-04
P71103	Phosphate acetyltransferase	<i>pta</i> CA_C1742	1.13	5.75E-06
Q97IF8	Ribonuclease J	<i>rnj</i> CA_C1683	1.13	7.58E-04
Q97I65	Elongation factor Ts	<i>tsf</i> CA_C1788	1.12	4.66E-06
Q97EH9	50S ribosomal protein L4	<i>rplD</i> CA_C3132	1.11	8.40E-04
P30725	Chaperone protein DnaJ	<i>dnaJ</i> CA_C1283	-1.15	2.11E-03
Q97EG9	DNA-directed RNA polymerase subunit beta	<i>rpoB</i> CA_C3143	-1.15	4.77E-04
Q97EH0	DNA-directed RNA polymerase subunit beta	<i>rpoC</i> CA_C3142	-1.17	1.20E-04
Q97E05	Chaperone protein HtpG	<i>htpG</i> CA_C3315	-1.19	3.49E-07
Q97K29	Predicted metal-dependent phosphoesterase	CA_C1092	-1.20	2.00E-04
Q97EF6	Tryptophan synthase alpha chain	<i>trpA</i> CA_C3157	-1.20	3.79E-04
Q97JM4	Site-determining protein	<i>minD</i> CA_C1249	-1.21	3.04E-04
Q97MC0	Aspartokinase	CA_C0278	-1.22	2.67E-03
Q97F13	S-ribosylhomocysteine lyase	<i>luxS</i> CA_C2942	-1.26	8.10E-04
Q97DE1	Fragment of SECA	CA_C3537	-1.26	4.39E-05
Q97ME5	Nitrogenase iron protein	<i>nifH</i> CA_C0253	-1.26	3.73E-03
Q97GI8	4-hydroxy-tetrahydrodipicolinate reductase	<i>dapB</i> CA_C2379	-1.26	1.65E-03
Q97HJ4	Acyl carrier protein homolog	CA_C2017	-1.27	6.91E-04
P58276	ATP-dependent Clp protease proteolytic subunit	<i>clpP</i> CA_C2640	-1.29	7.32E-07
Q97M19	Methyl-accepting chemotaxis protein	CA_C0381	-1.29	7.34E-04

Q97J57	UDP-glucose 4-epimerase	<i>galE</i> CA_C1429	-1.30	1.24E-05
Q97EF3	Indole-3-glycerol phosphate synthase	<i>trpC</i> CA_C3160	-1.32	1.59E-03
Q97IG8	Large subunit of NADH-dependent glutamate synthase	<i>gltA</i> CA_C1673	-1.36	2.23E-03
Q97EC4	ATPases with chaperone activity clpC, two ATP-binding domain	<i>clpC</i> CA_C3189	-1.36	7.24E-16
Q97IU4	Pyrimidine-nucleoside phosphorylase	<i>deoA</i> CA_C1546	-1.36	2.89E-04
Q97GJ6	33 kDa chaperonin	<i>hslO</i> CA_C2370	-1.38	1.65E-03
Q97EC3	Protein-arginine kinase	<i>mcsB</i> CA_C3190	-1.38	6.06E-05
Q97IU5	Deoxyribose-phosphate aldolase	<i>deoC</i> CA_C1545	-1.38	9.15E-04
Q97I53	Uncharacterized conserved protein, YLXR B.subtilis homolog	CA_C1800	-1.40	1.55E-06
Q97H96	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	<i>pfs</i> CA_C2117	-1.40	1.79E-04
Q97EB9	UDP-N-acetylmuramoylalanine--D-glutamate ligase	<i>murD</i> CA_C3194	-1.41	3.24E-04
Q97GY7	Uncharacterized protein	CA_C2228	-1.42	8.77E-04
Q97K78	Arginine--tRNA ligase	<i>argS</i> CA_C1041	-1.43	3.82E-05
Q97HA3	GAF domain-containing protein	CA_C2110	-1.43	2.58E-03
Q97F68	Peptide chain release factor 1	<i>prfA</i> CA_C2884	-1.43	1.24E-07
Q97IL1	Putative intracellular protease/amidase, ThiJ family	CA_C1629	-1.43	1.67E-03
Q97FW3	Transcriptional regulators of NagC/XylR family	<i>glcK</i> CA_C2613	-1.44	7.65E-04
P23673	Butyrate--acetoacetate CoA-transferase subunit B	<i>ctfB</i> CA_P0164	-1.44	1.03E-03
Q04945	NADH-dependent butanol dehydrogenase B	<i>bdhB</i> CA_C3298	-1.44	8.40E-05
Q97N34	DNA polymerase III subunit beta	<i>dnaN</i> CA_C0002	-1.44	4.66E-06
Q97IG2	UPF0297 protein	CA_C1679	-1.45	1.93E-06
Q97DT0	NADH-dependent butanol dehydrogenase	CA_C3392	-1.46	4.49E-07
Q97I54	Transcription termination/antitermination protein NusA	<i>nusA</i> CA_C1799	-1.46	3.30E-05
Q97K49	Fe-S oxidoreductase	CA_C1072	-1.46	4.31E-03

Q97MB9	Peptidylprolyl isomerase	CA_C0279	-1.46	9.35E-07
Q97EH4	Elongation factor G	<i>fusA</i> CA_C3138	-1.46	2.75E-22
Q97H81	Ribosomal RNA small subunit methyltransferase H	<i>rsmH mraW</i>	-1.47	1.99E-03
		CA_C2132		
Q97N05	Uncharacterized protein	CA_C0034	-1.48	1.46E-11
Q97KL3	ATPase with chaperon activity, two ATP-binding domains, ClpC orthologs	CA_C0904	-1.48	3.88E-06
P71104	Acetate kinase	<i>ackA ack</i>	-1.48	2.85E-11
		CA_C1743		
P33744	Aldehyde-alcohol dehydrogenase	<i>adhE aad</i>	-1.48	3.65E-16
		CA_P0162		
Q97LS4	Oxygen-sensitive ribonucleoside-triphosphate reductase <i>nrdD</i>	<i>nrdD</i> CA_C0480	-1.49	5.87E-04
Q97LP4	UDP-N-acetylenolpyruvoylglucosamine reductase	<i>murB</i> CA_C0510	-1.49	2.42E-03
Q97EJ6	50S ribosomal protein L30	<i>rpmD</i> CA_C3115	-1.50	3.84E-06
Q97M99	Protein from nitrogen regulatory protein P-II	CA_C0299	-1.50	3.75E-04
P33752	Butyrate--acetoacetate CoA-transferase subunit A	<i>ctfA</i> CA_P0163	-1.51	1.97E-07
Q97F94	Protein translocase subunit SecA	<i>secA</i> CA_C2846	-1.51	1.05E-13
Q97FT3	Carbamoyl-phosphate synthase large chain	<i>carB</i> CA_C2644	-1.52	2.54E-09
Q97GY9	Enzyme of ILVE/PABC family	CA_C2226	-1.53	1.21E-05
Q97EI6	50S ribosomal protein L29	<i>rpmC</i> CA_C3125	-1.53	1.12E-05
Q97KM2	3-phosphoshikimate 1-carboxyvinyltransferase	<i>aroA</i> CA_C0895	-1.54	7.42E-04
Q97DB0	Biotin carboxylase	<i>accC</i> CA_C3570	-1.54	3.47E-04
Q97MT1	Sulfate adenylyltransferase subunit 1	<i>cysN</i> CA_C0110	-1.55	2.50E-04
Q97JC5	ATP-dependent dethiobiotin synthetase BioD	<i>bioD</i> CA_C1361	-1.55	7.39E-08
Q97KG2	Xre family DNA-binding domain and TPR-repeat-containing protein	CA_C0957	-1.55	7.85E-06

Q97DA0	Dioxygenase related to 2-nitropropane dioxygenase	CA_C3580	-1.56	5.88E-05
Q97EK2	50S ribosomal protein L36	<i>rpmJ</i> CA_C3108	-1.57	1.71E-03
P58254	RecA	<i>recA</i> CA_C1815	-1.58	2.87E-10
Q97KG0	Chaperone protein ClpB	<i>clpB</i> CA_C0959	-1.58	1.07E-07
Q97FM0	Predicted glycosyl transferase from UDP-glucuronosyltransferase family	CA_C2716	-1.58	1.40E-04
Q97KX4	D-amino acid aminotransferase	CA_C0792	-1.59	3.17E-03
Q97FC1	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase 2	<i>murE2</i> CA_C2819	-1.62	8.35E-05
Q97ED6	Cysteine--tRNA ligase	<i>cysS</i> CA_C3177	-1.62	5.45E-04
Q97LD6	Tryptophan--tRNA ligase	<i>trpS</i> CA_C0626	-1.63	8.91E-07
Q97MN6	Glutamine--fructose-6-phosphate aminotransferase	<i>glmS</i> CA_C0158	-1.63	2.41E-06
Q97LP3	Nucleotide-binding protein	CA_C0511	-1.66	1.42E-03
Q97E89	UDP-N-acetylmuramate--L-alanine ligase	<i>murC</i> CA_C3225	-1.67	6.50E-05
Q97FS5	Orotidine 5'-phosphate decarboxylase	<i>pyrF</i> CA_C2652	-1.69	4.15E-03
Q97K95	L-aspartate oxidase	<i>nadB</i> CA_C1024	-1.71	9.97E-04
Q97KC0	Threonine synthase	<i>thrC</i> CA_C0999	-1.71	8.75E-06
Q97KZ9	Fe-S oxidoreductase	CA_C0767	-1.72	3.89E-10
Q97KC9	Glutamate--tRNA ligase	<i>gltX</i> CA_C0990	-1.74	5.89E-15
Q97LC7	Zinc finger domain	CA_C0635	-1.75	2.05E-05
Q97L98	Sugar-binding periplasmic protein	CA_C0664	-1.77	1.22E-04
Q97G87	Carbonic anhydrase	CA_C2482	-1.78	1.95E-03
Q97H79	Ribosome-binding ATPase YchF	<i>ychF</i> CA_C2134	-1.78	6.55E-06
Q97FT7	ATP-dependent Clp protease ATP-binding subunit ClpX	<i>clpX</i> CA_C2639	-1.84	1.20E-07
Q97DW5	Polyketide synthase pksE	CA_C3355	-1.85	2.84E-06

Q97GU1	RelA/SpoT protein	<i>spoT</i> CA_C2274	-1.93	1.43E-05
Q97IF7	TYPB/BIPA type GTPase	CA_C1684	-2.04	5.20E-16
Q97FT9	Lon protease	<i>lon</i> CA_C2637	-2.12	1.30E-04
Q97LD2	Peptide chain release factor 3	<i>prfC</i> CA_C0630	-2.22	2.19E-03
Q97LA0	Sugar ABC transporter, periplasmic sugar-binding protein	CA_C0662	-2.23	7.80E-04
Q97IT9	Nitroreductase family protein	CA_C1551	-2.29	1.57E-06
Q97MI6	Biotin synthase	<i>bioB</i> CA_C0210	-2.53	6.97E-04
Q97HL3	ABC-type transport system, ATPase component	CA_C1998	-2.54	6.91E-05

Appendix Table 2: iTRAQ Proteomic analysis of *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* with 10 μ M QspB 7 addition vs *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB*. Identified RNPP type regulators within the proteomics data set are highlighted in turquoise, proteins of interest investigated during enzyme or GusA assays are highlighted in green, where they don't fall into the latter category.

UniProt ID	Protein names	Gene name and loci	Fold Change	p-value
Q97IT9	Nitroreductase family protein	CA_C1551	2.99	6.99922E-07
P33744	Aldehyde-alcohol dehydrogenase	<i>adhE aad</i> CA_P0162	2.70	1.15897E-30
Q97G30	FAD/FMN-containing dehydrogenase	CA_C2542	2.64	3.25993E-06
Q97LA0	Sugar ABC transporter, SBP protein	CA_C0662	2.52	1.24E-04
Q97ID5	Stage IV sporulation protein A	CA_C1713	2.48	3.39628E-05
Q97MI6	Biotin synthase	<i>bioB</i> CA_C0210	2.38	0.000328683
Q97E72	Spore cortex-lytic enzyme	CA_C3244	2.31	0.002515914
P33752	Butyrate--acetoacetate CoA-transferase subunit A	<i>ctfA</i> CA_P0163	2.29	1.53007E-06
Q97EX1	Uncharacterized protein	CA_C2985	2.26	1.11044E-05
Q97F50	LysM domain containing membrane protein	CA_C2903	2.21	6.00565E-08
P23673	Butyrate--acetoacetate CoA-transferase subunit B	<i>ctfB</i> CA_P0164	2.11	3.64946E-05
Q97HC6	Stage III sporulation protein AH, SpoIIIAH	CA_C2086	2.10	7.57895E-05
Q97I63	Ribosome-recycling factor	<i>frs</i> CA_C1790	2.06	2.44935E-10
Q97FW6	L-fucose isomerase related protein	CA_C2610	2.04	2.1706E-05
Q97EC3	Protein-arginine kinase	<i>mcsB</i> CA_C3190	2.03	2.81304E-09
Q97LI2	Serine protein kinase	<i>prkA</i> CA_C0579	1.99	8.3804E-08
Q97MB4	NAD-dependent protein deacetylase	<i>cobB</i> CA_C0284	1.96	5.79919E-05
Q04945	NADH-dependent butanol dehydrogenase B	<i>bdbB</i> CA_C3298	1.96	8.96943E-09
Q97L98	Sugar-binding periplasmic protein	CA_C0664	1.96	5.59755E-05
Q97D09	Two CBS domain containing protein	CA_C3674	1.94	0.000299478
Q97FC1	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase 2	<i>murE2</i> CA_C2819	1.93	1.58267E-05
Q97GI5	Single-strand DNA-binding protein	<i>ssb</i> CA_C2382	1.89	0.00047492

Q97LC7	Zinc finger domain	CA_C0635	1.89	0.000134938
Q97EC4	ATPases with chaperone activity clpC	<i>clpC</i> CA_C3189	1.86	4.27125E-37
Q97FE9	MoaA/NirJ family Fe-S oxidoreductase	CA_C2791	1.85	3.35777E-05
Q97EI6	50S ribosomal protein L29	<i>rpmC</i> CA_C3125	1.84	3.1234E-08
Q97K95	L-aspartate oxidase	<i>nadB</i> CA_C1024	1.84	0.001229588
Q97K49	Fe-S oxidoreductase	CA_C1072	1.83	0.000718812
Q97IL1	Putative intracellular protease/amidase	CA_C1629	1.80	0.000159192
Q97FW5	Uncharacterized protein	CA_C2611	1.80	0.001046454
Q97IF7	TYPB/BIPA type GTPase	CA_C1684	1.79	2.97106E-12
Q97FT7	ATP-dependent Clp protease ATP-binding subunit ClpX	<i>clpX</i> CA_C2639	1.79	5.64888E-06
Q97HQ1	Predicted aldo/keto reductase	CA_C1958	1.77	0.000150411
Q97N16	Aspartate-semialdehyde dehydrogenase	<i>asd</i> CA_C0022	1.76	0.001078495
Q97M74	TPR repeats containing protein	CA_C0324	1.75	4.5224E-15
Q97FT9	Lon protease	<i>lon</i> CA_C2637	1.71	0.001550144
Q97EK2	50S ribosomal protein L36	<i>rpmJ</i> CA_C3108	1.71	0.001386437
Q97N05	Uncharacterized protein	CA_C0034	1.68	8.67425E-13
Q97H96	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	<i>pfs</i> CA_C2117	1.66	0.00031513
Q97DB0	Biotin carboxylase	<i>accC</i> CA_C3570	1.66	0.000120846
Q97E89	UDP-N-acetylmuramate--L-alanine ligase	<i>murC</i> CA_C3225	1.66	0.000279693
Q97GU1	RelA/SpoT protein, (P)ppGpp synthetase/pyrophosphohydrolases	<i>spoT</i> CA_C2274	1.64	0.000321606
P30719	10 kDa chaperonin	<i>groS groES</i> CA_C2704	1.64	6.93805E-29
Q97I53	Uncharacterized conserved protein	CA_C1800	1.64	7.01402E-07
Q97KV0	Lipase-esterase related protein	CA_C0816	1.63	0.000190757
Q97GI0	Uncharacterized protein	CA_C2387	1.62	0.000116955
Q97MK0	NADH-dependent flavine oxidoreductase	CA_C0196	1.59	0.000114807
Q97JC5	ATP-dependent dethiobiotin synthetase BioD	<i>bioD</i> CA_C1361	1.58	1.30837E-05
Q97KC9	Glutamate--tRNA ligase	<i>gltX</i> CA_C0990	1.58	3.04323E-10

Q97IV8	Protein containing ChW-repeats	CA_C1532	1.58	0.000108135
Q97ED6	Cysteine--tRNA ligase	<i>cysS</i> CA_C3177	1.57	0.000132852
Q97KG0	Chaperone protein ClpB	<i>clpB</i> CA_C0959	1.57	5.2845E-07
Q97K72	Ribonucleotide reductase	CA_C1047	1.56	0.000510241
Q97MT1	Sulfate adenylyltransferase subunit 1	<i>cysN</i> CA_C0110	1.55	3.9204E-05
Q97EJ6	50S ribosomal protein L30	<i>rpmD</i> CA_C3115	1.55	1.39244E-05
Q97GY9	Enzyme of ILVE/PABC family	CA_C2226	1.49	0.00064966
Q97KC0	Threonine synthase	<i>thrC</i> CA_C0999	1.49	0.000445838
Q97F94	Protein translocase subunit SecA	<i>secA</i> CA_C2846	1.49	2.20033E-12
Q97EB9	UDP-N-acetylmuramoylalanine--D-glutamate ligase	<i>murD</i> CA_C3194	1.49	2.36978E-06
Q97DT0	NADH-dependent butanol dehydrogenase	CA_C3392	1.48	1.95624E-09
Q97TI8	Solo B3/4 domain	CA_P0113	1.44	8.78896E-06
Q97H79	Ribosome-binding ATPase YchF	<i>ychF</i> CA_C2134	1.44	0.002133756
P58254	RecA	<i>recA</i> CA_C1815	1.43	2.58601E-07
Q97KL3	ATPase with chaperon activity	CA_C0904	1.42	0.000391262
Q97MN6	Glutamine--fructose-6-phosphate aminotransferase	<i>glmS</i> CA_C0158	1.41	6.1383E-06
Q97DA5	Malonyl CoA-acyl carrier protein transacylase	<i>fabD</i> CA_C3575	1.41	0.000271002
Q97DY6	Multimeric flavodoxin WrbA family protein	CA_C3334	1.40	0.001440636
Q97EJ9	Adenylate kinase	<i>adk</i> CA_C3112	1.39	1.72001E-06
Q97I54	Transcription termination/antitermination protein NusA	<i>nusA</i> CA_C1799	1.38	0.000528304
P69000	Methionine aminopeptidase	<i>map</i> CA_C3111	1.37	0.00055822
Q97L02	NADPH-dependent glutamate synthase beta chain	CA_C0764	1.37	0.000429092
P58276	ATP-dependent Clp protease proteolytic subunit	<i>clpP</i> CA_C2640	1.36	4.65107E-05
Q97EM0	Fumarate hydratase, subunit B	CA_C3090	1.36	0.003364956
Q97DE1	Fragment of SECA	CA_C3537	1.36	1.42729E-06
Q97F13	S-ribosylhomocysteine lyase	<i>luxS</i> CA_C2942	1.36	9.76492E-05
Q97F68	Peptide chain release factor 1	<i>prfA</i> CA_C2884	1.35	1.2489E-05

Q97FT3	Carbamoyl-phosphate synthase large chain	<i>carB</i> CA_C2644	1.34	3.9761E-06
Q97D25	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	CA_C3657	1.28	0.000661643
Q97GS1	Probable transcriptional regulatory protein	CA_C2295	1.28	1.89624E-05
Q97I50	Ribosome-binding factor A	<i>rbfA</i> CA_C1803	1.25	0.001401876
Q97KG2	Xre family DNA-binding domain and TPR-repeat-containing protein	CA_C0957	1.25	0.001716712
Q97HB8	Elongation factor P	<i>efp</i> CA_C2094	1.23	0.000107847
Q97E52	Uncharacterized conserved protein	CA_C3264	1.22	0.002864718
Q97GM9	Phosphomannomutase	CA_C2337	1.20	0.000356105
O52632	Phosphoglycerate kinase	<i>pgk</i> CA_C0710	1.18	9.23482E-06
Q97KT9	Fructose-bisphosphate aldolase	CA_C0827	1.17	0.000101732
Q97EK6	DNA-directed RNA polymerase subunit alpha	<i>rpoA</i> CA_C3104	1.12	4.87452E-05
P52046	Short-chain-enoyl-CoA hydratase	<i>crt</i> CA_C2712	-1.14	1.02024E-08
P94604	DNA gyrase subunit B	<i>gyrB</i> CA_C0006	-1.16	6.16553E-05
Q97I65	Elongation factor Ts	<i>tsf</i> CA_C1788	-1.18	6.26306E-05
Q97I93	50S ribosomal protein L19	<i>rplS</i> CA_C1759	-1.18	0.001887391
Q97I09	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	<i>ispH</i> CA_C1847	-1.18	1.63927E-05
Q97KM1	Chorismate synthase	<i>aroC</i> CA_C0896	-1.19	0.004180156
Q97I33	Adenylosuccinate lyase	<i>purB</i> CA_C1821	-1.20	0.000998962
Q45829	Butyrate kinase 1	<i>buk1 buk</i> CA_C3075	-1.20	0.000440082
Q97EI7	30S ribosomal protein S17	<i>rpsQ</i> CA_C3124	-1.21	3.48385E-05
P45360	Putative UDP-N-acetylglucosamine 2-epimerase	CA_C2874	-1.21	0.000295374
Q97EJ0	50S ribosomal protein L5	<i>rplE</i> CA_C3121	-1.21	2.85178E-12
Q97EX8	Glutamyl-tRNA	<i>gatA2</i> CA_C2977	-1.21	0.000802898
Q97GK9	Phenylalanine--tRNA ligase alpha subunit	<i>pheS</i> CA_C2357	-1.21	0.000577207
Q97J93	Phosphoribosylformylglycinamide cyclo-ligase	<i>purM</i> CA_C1393	-1.22	6.10164E-05
Q97EJ4	50S ribosomal protein L18	<i>rplR</i> CA_C3117	-1.22	0.001516432
Q97I46	30S ribosomal protein S15	<i>rpsO</i> CA_C1807	-1.22	7.68801E-05

Q9Z688	ATP synthase gamma chain	<i>atpG</i> CA_C2866	-1.22	0.000608129
Q97EG6	50S ribosomal protein L1	<i>rplA</i> CA_C3147	-1.23	3.52434E-06
Q97IE9	Cell division protein FtsZ	<i>ftsZ</i> CA_C1693	-1.23	6.43714E-07
P45359	Acetyl-CoA acetyltransferase	<i>thlA</i> <i>thl</i> CA_C2873	-1.24	4.25364E-11
Q97GN4	dTDP-glucose 4,6-dehydratase	<i>spsJ</i> CA_C2332	-1.24	1.12468E-05
Q97GL0	Phenylalanine--tRNA ligase beta subunit	<i>pheT</i> CA_C2356	-1.25	4.12903E-10
Q97MT9	O-acetylhomoserine sulphydrylase	CA_C0102	-1.25	0.003546777
Q97N17	Serine--tRNA ligase 1	<i>serS1</i> <i>serS</i> CA_C0021	-1.25	6.6918E-05
Q97D48	Oligopeptide ABC transporter, SBP component	<i>oppA</i> CA_C3632	-1.25	5.02991E-10
Q97JL7	50S ribosomal protein L21	<i>rplU</i> CA_C1257	-1.25	0.000934962
Q97DC6	L-lactate dehydrogenase 2	<i>ldh2</i> CA_C3552	-1.25	0.002045845
Q97EJ5	30S ribosomal protein S5	<i>rpsE</i> CA_C3116	-1.25	2.11495E-09
Q97EL2	50S ribosomal protein L13	<i>rplM</i> CA_C3098	-1.25	2.5964E-07
Q97D85	Uncharacterized FAD-dependent dehydrogenase	CA_C3595	-1.26	0.001637343
Q97F85	S-adenosylmethionine synthase	<i>metK</i> CA_C2856	-1.26	5.46548E-05
Q97EJ3	50S ribosomal protein L6	<i>rplF</i> CA_C3118	-1.27	0.000426463
O08309	Pyruvate kinase	<i>pyk</i> <i>pykA</i> CA_C0518	-1.27	2.68561E-09
Q97F58	D-alanine--D-alanine ligase	<i>ddl</i> CA_C2895	-1.27	0.000121731
Q97GQ1	dTDP-4-dehydrorhamnose reductase, rfbD ortholog	CA_C2315	-1.28	4.32708E-05
Q97FJ7	Histidine--tRNA ligase	<i>hisS</i> CA_C2740	-1.28	0.000206638
P52042	Acyl-CoA dehydrogenase, short-chain specific	<i>bcd</i> CA_C2711	-1.28	1.07748E-15
P52039	Electron transfer flavoprotein subunit alpha	<i>etfA</i> CA_C2709	-1.29	1.91899E-08
Q97J91	Bifunctional purine biosynthesis protein	<i>purH</i> CA_C1395	-1.29	4.81846E-14
Q97I64	Uridylate kinase	<i>pyrH</i> CA_C1789	-1.29	0.003082642
Q97M93	Chemotaxis motility protein B	<i>motB</i> CA_C0305	-1.30	0.001702871
P71103	Phosphate acetyltransferase	<i>pta</i> CA_C1742	-1.30	3.60983E-13
Q97KH5	Transketolase	<i>tkt</i> CA_C0944	-1.30	3.7511E-05

Q97EG7	50S ribosomal protein L10	<i>rplJ</i> CA_C3146	-1.31	2.15081E-07
Q97EI5	50S ribosomal protein L16	<i>rplP</i> CA_C3126	-1.31	3.11596E-05
Q97EX6	Aspartyl-tRNA synthetase	<i>aspS</i> CA_C2979	-1.31	8.11964E-09
Q97MX5	Uncharacterized protein	CA_C0065	-1.31	0.001411849
Q97K83	Pyruvate kinase	<i>pykA</i> CA_C1036	-1.31	5.8995E-16
Q97LG7	Pyridoxal 5'-phosphate synthase subunit PdxS	<i>pdxS</i> CA_C0594	-1.32	5.57967E-09
Q97GY6	Pyruvate-flavodoxin oxidoreductase	CA_C2229	-1.33	3.32867E-35
Q97EI1	50S ribosomal protein L2	<i>rplB</i> CA_C3130	-1.33	8.22422E-09
Q97EB7	Lysine--tRNA ligase	<i>lysS</i> CA_C3197	-1.33	1.69177E-10
Q97GG8	Valine--tRNA ligase	<i>valS</i> CA_C2399	-1.33	9.65441E-09
Q97EB3	Formate--tetrahydrofolate ligase	<i>fhs</i> CA_C3201	-1.33	8.47056E-15
Q97JC4	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	<i>bioA</i> CA_C1362	-1.34	2.06337E-07
Q97E27	Iron-regulated ABC-type transporter membrane component	CA_C3289	-1.34	3.57885E-06
Q97KN8	ABC-type polar amino acid transport system, ATPase component	CA_C0879	-1.34	0.00199537
Q97E92	Bifunctional protein GlnU	<i>glmU</i> CA_C3222	-1.34	5.33227E-07
Q97LS0	Phosphoglucosamine mutase	<i>glmM</i> CA_C0484	-1.34	1.78907E-09
O52633	Triosephosphate isomerase	<i>tpiA tpi</i> CA_C0711	-1.34	2.67067E-08
Q97CX2	30S ribosomal protein S6	<i>rpsF</i> CA_C3724	-1.34	1.88251E-06
Q97MU7	Ferredoxin-nitrite reductase	CA_C0094	-1.35	2.99154E-05
Q97E28	Iron-regulated ABC transporter ATPase subunit	CA_C3288	-1.35	2.35647E-08
Q97F61	CTP synthase	<i>pyrG</i> CA_C2892	-1.35	1.19804E-05
Q97FT6	Trigger factor	<i>tig</i> CA_C2641	-1.36	2.303E-19
Q97TQ0	Related to methyl-accepting chemotaxis protein	CA_P0048	-1.36	0.000720079
Q97EK4	30S ribosomal protein S11	<i>rpsK</i> CA_C3106	-1.36	9.80085E-06
Q97IN6	Methylglyoxal synthase	<i>mgsA</i> CA_C1604	-1.36	0.001015582
Q97LB6	Leucine--tRNA ligase	<i>leuS</i> CA_C0646	-1.36	5.50011E-08
Q9Z689	ATP synthase subunit alpha	<i>atpA</i> CA_C2867	-1.36	1.37556E-10

Q97KU7	Phosphoribosylpyrophosphate synthetase	CA_C0819	-1.36	0.000393395
P58255	Phosphate butyryltransferase	ptb CA_C3076	-1.36	4.00533E-23
Q9Z687	ATP synthase subunit beta	atpD CA_C2865	-1.36	1.45113E-11
Q97KP0	Cyclopropane fatty acid synthase	cfa CA_C0877	-1.37	5.80475E-11
Q97G09	NifU-related domain containing protein	CA_C2565	-1.37	2.32862E-08
Q97EF5	Tryptophan synthase beta chain	trpB CA_C3158	-1.37	2.54452E-05
Q97DV0	2-enoate reductase	CA_C3371	-1.38	3.15358E-10
Q97ID4	L-asparaginase	ansA CA_C1714	-1.38	0.000238539
Q97H59	Flagellar basal body protein	flgE CA_C2154	-1.39	5.15794E-05
Q97ED5	Proline--tRNA ligase	proS CA_C3178	-1.40	4.37868E-06
Q97EH3	30S ribosomal protein S7	rpsG CA_C3139	-1.40	4.29982E-23
Q7D462	Flagellin	hag CA_C2203	-1.41	1.51965E-06
Q97IW7	Fructokinase	CA_C1523	-1.41	9.77574E-06
Q97G10	Uncharacterized protein	CA_C2564	-1.41	1.17506E-24
Q97J94	Amidophosphoribosyltransferase	purF CA_C1392	-1.41	3.91003E-07
Q97DD3	Methyl-accepting chemotaxis protein	CA_C3545	-1.42	0.001089198
Q97KB7	Nicotinate phosphoribosyltransferase	CA_C1002	-1.42	0.001150175
P52040	Electron transfer flavoprotein subunit beta	etfB CA_C2710	-1.42	3.30007E-17
Q97LG8	6,7-dimethyl-8-ribityllumazine synthase	ribH CA_C0593	-1.43	0.000281615
Q97FF7	O-acetylhomoserine sulphydrylase	cysD CA_C2783	-1.44	1.73738E-08
Q7D472	Hydrogene dehydrogenase	hydA CA_C0028	-1.47	5.56195E-22
Q97LJ0	PTS enzyme II, ABC component	CA_C0570	-1.48	0.000742272
Q97GI6	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase	dapH CA_C2381	-1.50	4.14738E-07
Q97F92	Proline/glycine betaine ABC-type transport system	CA_C2849	-1.51	8.17804E-05
Q97CX5	Hypothetical secreted protein	CA_C3721	-1.55	1.37189E-09
Q97LX2	Methyl-accepting chemotaxis protein	CA_C0432	-1.56	0.001116867
Q97KD3	Lipoprotein	CA_C0986	-1.56	6.88945E-10

Q97KN7	Periplasmic amino acid binding protein	CA_C0880	-1.57	1.52973E-10
Q97GP2	Glycosyltransferase	CA_C2324	-1.57	0.000394418
Q97F73	Uracil phosphoribosyltransferase	<i>upp</i> CA_C2879	-1.57	2.84863E-11
Q97GK6	50S ribosomal protein L35	<i>rpmI</i> CA_C2360	-1.58	0.000762185
Q97DZ5	Periplasmic amino acid binding protein	CA_C3325	-1.58	5.96046E-08
Q97GY0	Cysteine synthase	<i>cysK</i> CA_C2235	-1.58	1.74548E-26
Q97IT1	Predicted transcriptional regulator	CA_C1559	-1.62	0.000584406
Q97N24	Aminotransferase	CA_C0014	-1.65	2.20548E-09
Q97IQ1	Malic enzyme	<i>malS</i> CA_C1589	-1.71	5.78713E-16
Q97G38	Protein containing ChW-repeats	CA_C2533	-1.74	0.000458397
Q97JG7	Uncharacterized protein	CA_C1318	-1.77	1.77718E-08
Q97MY3	Uncharacterized protein	CA_C0057	-2.02	2.98019E-07
Q97N21	Serine--tRNA ligase 2	<i>serS2</i> CA_C0017	-2.02	4.58849E-09

Appendix Table 3: iTRAQ Proteomic analysis of *C. acetobutylicum* ATCC 824 pMTL85143-*qsrB* supplemented with 10 μ M QspB 7vs *C. acetobutylicum* ATCC 824 pMTL85143. Identified RNPP type regulators within the proteomics data set are highlighted in turquoise, proteins of interest investigated during enzyme or GusA assays are highlighted in green, where they don't fall into the latter category.

UniProt ID	Protein names	Gene name and loci	Fold Change	p-value
Q97EX1	Uncharacterized protein	CA_C2985	2.34	1.36101E-05
Q97ID5	Stage IV sporulation protein A	CA_C1713	2.25	4.14289E-06
Q97G30	FAD/FMN-containing dehydrogenase	CA_C2542	2.24	2.0298E-05
Q97M74	TPR repeats containing protein	CA_C0324	2.14	1.47841E-13
Q97F50	LysM domain containing membrane protein	CA_C2903	2.12	6.75769E-08
Q7DFN2	Aldehyde-alcohol dehydrogenase	<i>adhe</i> CA_P0035	2.05	0.002944878
P33744	Aldehyde-alcohol dehydrogenase	<i>adhE aad</i> CA_P0162	2.03	0.000336183
Q97FE9	MoaA/NirJ family Fe-S oxidoreductase	CA_C2791	2.01	0.000216138
Q97FW6	L-fucose isomerase related protein	CA_C2610	1.99	1.20678E-06
Q97LI2	Serine protein kinase	<i>prkA</i> CA_C0579	1.97	0.002420354
Q97F47	Spore coat protein cotS related	CA_C2906	1.89	0.000411428
Q97HC6	Stage III sporulation protein AH, SpoIIIAH	CA_C2086	1.86	7.28348E-06
Q97I63	Ribosome-recycling factor	<i>frr</i> CA_C1790	1.71	2.81824E-05
Q97DU5	Possible pectin degradation protein	CA_C3376	1.57	0.003098557
Q03928	18 kDa heat shock protein	<i>hsp18</i> CA_C3714	1.56	1.17016E-06
Q97JJ1	Uncharacterized conserved protein	CA_C1289	1.54	9.45471E-08
Q97EC3	Protein-arginine kinase	<i>mcsB</i> CA_C3190	1.49	0.000836258
Q97EJ9	Adenylate kinase	<i>adk</i> CA_C3112	1.45	3.209E-06
P30719	10 kDa chaperonin	<i>groS groES</i> CA_C2704	1.45	7.69758E-19
Q97K72	Ribonucleotide reductase	CA_C1047	1.41	0.000530474
P30717	60 kDa chaperonin	<i>groL groEL mopA</i> CA_C2703	1.39	4.57806E-37
Q97K67	Membrane protease subunit	CA_C1052	1.39	2.45784E-07

Q97EC4	ATPases with chaperone activity clpC	<i>clpC</i> CA_C3189	1.37	5.41774E-20
Q97EB6	Transcription elongation factor GreA	<i>greA</i> CA_C3198	1.35	0.000461808
Q97J07	Branched-chain-amino-acid transaminase	<i>ilvE</i> CA_C1479	1.34	6.09348E-16
Q97TI8	Solo B3/4 domain	CA_P0113	1.33	0.001862491
P69000	Methionine aminopeptidase	<i>map</i> CA_C3111	1.30	0.000207118
P45359	Acetyl-CoA acetyltransferase	<i>thlA thl</i> CA_C2873	1.26	1.78429E-17
Q97D83	Reverse rubrerythrin-2	<i>rbr3B hsp21 rpr2</i> CA_C3597	1.25	0.003143772
O52632	Phosphoglycerate kinase	<i>pgk</i> CA_C0710	1.25	6.52481E-10
Q97GS1	Probable transcriptional regulatory protein	CA_C2295	1.23	0.000411725
Q97EK6	DNA-directed RNA polymerase subunit alpha	<i>rpoA</i> CA_C3104	1.19	2.28018E-08
Q97EI9	50S ribosomal protein L24	<i>rplX</i> CA_C3122	1.13	0.001215388
P52046	Short-chain-enoyl-CoA hydratase	<i>crt</i> CA_C2712	1.11	0.000133034
Q97LB6	Leucine--tRNA ligase	<i>leuS</i> CA_C0646	-1.16	0.002156429
P71103	Phosphate acetyltransferase	<i>pta</i> CA_C1742	-1.16	1.54354E-05
Q97FT3	Carbamoyl-phosphate synthase large chain	<i>carB</i> CA_C2644	-1.16	0.000805737
Q7D472	Hydrogene dehydrogenase	<i>hydA</i> CA_C0028	-1.17	1.53036E-06
Q97GV1	Serine hydroxymethyltransferase	<i>glyA</i> CA_C2264	-1.18	8.32537E-05
Q97K33	Fructose-1,6-bisphosphatase	<i>glpX</i> CA_C1088	-1.18	0.0010816
Q97IV0	D-alanine--D-alanine ligase	CA_C1540	-1.19	0.000141464
Q97EX6	Aspartyl-tRNA synthetase	<i>aspS</i> CA_C2979	-1.20	0.001466715
Q97EX8	Glutamyl-tRNA amidotransferase subunit A 2	<i>gatA2</i> CA_C2977	-1.21	0.000423325
Q97IE9	Cell division protein FtsZ	<i>ftsZ</i> CA_C1693	-1.23	4.23366E-06
Q97EH4	Elongation factor G	<i>fusA</i> CA_C3138	-1.24	5.88319E-07
Q97KG2	Xre family DNA-binding domain and TPR-repeat-containing protein	CA_C0957	-1.25	0.000771262
P71104	Acetate kinase	<i>ackA ack</i> CA_C1743	-1.28	3.37665E-05
Q97KE7	Isocitrate dehydrogenase	<i>citC</i> CA_C0972	-1.33	0.002509545
Q97MB9	Peptidylprolyl isomerase	CA_C0279	-1.37	1.90304E-05

Q97KZ9	Fe-S oxidoreductase	CA_C0767	-1.38	9.9081E-06
Q97DW5	Polyketide synthase pksE	CA_C3355	-1.40	0.000394396
Q97LL6	Methyl-accepting chemotaxis protein	CA_C0542	-1.41	0.003117174
Q97EF5	Tryptophan synthase beta chain	<i>trpB</i> CA_C3158	-1.42	7.12256E-06
Q97EF6	Tryptophan synthase alpha chain	<i>trpA</i> CA_C3157	-1.42	1.07191E-07
Q97KE8	Aconitase A	<i>citB</i> CA_C0971	-1.43	1.3666E-05
Q97LD6	Tryptophan--tRNA ligase	<i>trpS</i> CA_C0626	-1.47	5.1067E-08
Q97KM3	3-dehydroquinate synthase	<i>aroB</i> CA_C0894	-1.50	0.000846909
Q97HL3	ABC-type transport system, ATPase component	CA_C1998	-1.66	0.001351894
Q97EF2	Anthranilate phosphoribosyltransferase	<i>trpD</i> CA_C3161	-1.81	2.51362E-05

Professional Internships for PhD Students

Reflection Form

Name of Organization

Singer Instruments

Details of Placement

Please describe your main activities during the placement (150-200 words)

As the company is an SME I spent my time performing multiple roles, primarily in-between the sales team and the research and development team.

Initially, I spent the majority of my time getting familiarising myself with the robotics and troubleshooting possible lab handling problems. This was combined with any science support that the technical and product development staff required. Time was spent in meetings with senior staff, where I was expected to use the research knowledge I had acquired to answer questions relating to usage of products and explaining protocols to none scientific audience.

A larger project was conducted showing that florescence could be detected by the PhenoBooth, an advanced colony imager. This project was eventually a success and the product brought to market during my placement.

This was supplemented with experience in a sales environment, with explaining the equipment to customers and giving advice on if the product was suitable for their needs.

Placement Achievements

Please detail all outcomes from the placement, including any publications, presentations given and reports written etc. (150-200 words)

Due to my placement a product was brought to market, and was adequately tested for function. This product has now entered the international market. I contributed to a record financial successful year of the company, through supporting the sales and product development teams. I further developed a product development plan for a software package, to allow the expedient relative identification of growth changes in yeast colonies in a high throughput manner, the report for which was my final contribution to the company. I was asked to take part in a reflationary talk to the BBSRC advisory board, detailing my PIP placement and the subsequent advantages and disadvantages, for which I received positive feedback. I feel that through this placement I have adequately supported one of the key components of my funding body, the BBSRC, by promoting the bioeconomy within the UK.

Skill development

Has this Placement helped you developed any new skills or enhanced your previous skill set? (100-150 words)

Being able to communicate with the sales team, product development team, and academic customers taught me a great deal about the communication of science. I had to operate in a cross disciplinary team of managers, computer programmers, engineers and academics, and be able to be understood by all present. I gained skills in customer facing roles in which I had previously never performed. I further developed project management skills, with a special focus on team working which enhanced my previous skill set. My understanding of product development strategies increased dramatically, giving me a deeper understanding of the bioeconomy. Overall, the placement was very successful.

Future Work

Has this Placement influenced your future career aspirations? If so, in what way?
(150-200 words)

Previously, I would not have considered working in a SME as a potential future career. However, being able to view the work ethic and exciting atmosphere of a small high-tech product development company, which places a huge emphasis on previous skills gained, has changed my mind. The variety of tasks, from sales to product development, is dynamic and rewarding. This allowed me to refocus my thought on employment satisfaction. I believe that the skills gained during a PhD do have uses to the wider bioeconomy, as outlined by my funding body the BBSRC. I was offered employment through Singer Instruments, which I have accepted.