

---

**Forward and Reverse Genetics  
in  
Industrially Important Clostridia**

---

**Alexander Grosse-Honebrink, MSc**

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

**March 2017**



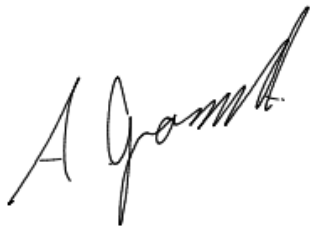
For my parents





# 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.

A handwritten signature in black ink, appearing to read 'A Grosse-Honebrink', written in a cursive style.

Alexander Grosse-Honebrink  
March 2017



# Abstract

The bacterial genus *Clostridium* is composed of Gram-positive, spore-forming rods with widespread biotechnological applications. This study focused mainly on *Clostridium pasteurianum* DSM 525, a saccharolytic species which is able to convert glycerol, the by-product of the biodiesel industry, into the valuable chemical commodities *n*-butanol, ethanol and 1,3-propanediol. The aim was to formulate reproducible methods for the creation of mutants, both directed and random, and use the tools developed to investigate genes, and their products, important in solvent production.

A prerequisite for the deployment of the envisaged genetic tools was a reproducible means for their introduction into the cell. Following the observation of low frequencies of plasmid transfer by electroporation it was hypothesised that the low level of transformants observed were a consequence of the presence of rare hypertransformable variants within the population. Accordingly, successfully transformed clonal populations were cured of their acquired plasmid and retransformed. In a number of instances the cured cell lines proved hypertransformable, with plasmid transformation frequencies obtained that were 5 orders of magnitude higher than those obtained with the progenitor strain. All of the hypertransformable strains isolated were shown by whole genome sequence to contain single nucleotide polymorphisms (SNPs) in one or more genes. In one instance, the single SNP present was shown to be directly responsible for the increased transformation frequency by its deliberate restoration to wild type using the allelic exchange procedures subsequently developed.

Having established reproducible, high frequencies of plasmid transformation reverse genetics was employed to establish gene function. Accordingly, allelic exchange gene knock-out procedures were used to target genes coding for enzymes of the central energy metabolism in *C. pasteurianum* and the phenotypes of the mutants obtained were analysed in laboratory scale fermentations. Strains in which the genes encoding the redox response regulator (*rex*) and a hydrogenase (*hyd*) were deleted showed increased *n*-butanol titres, representing first steps towards utilisation of *C. pasteurianum* as a chassis for this important chemical. With the inactivation of the *dhaBCE* gene, encoding

---

glycerol dehydratase, production of 1,3-propanediol was entirely eliminated, demonstrating the importance of the reductive pathway for growth and redox homeostasis of this organism when grown on glycerol.

In order to allow forward genetic approaches, a *mariner*-transposon system previously exemplified in *Clostridium difficile* was adapted for use in alternative clostridial hosts. In the absence of an efficient transformation system for *C. pasteurianum*, the initial exemplification of the system was undertaken in *Clostridium acetobutylicum* and *Clostridium sporogenes*. Successful transposon delivery was demonstrated through the use of a plasmid conditional for replication and through the insertion of a gene encoding an alternate  $\sigma$ -factor, TcdR, into their genomes. Transposition was shown to be entirely random and the libraries obtained of sufficient size to allow the isolation of both auxotrophic and sporulation/germination deficient mutants. Steps were taken to develop the same system in *C. pasteurianum* which was successful by using a suicide delivery plasmid, which was only possible with the high transformation efficiency achieved as part of this study.

This study presents an essential forward genetics procedure for industrially important *Clostridium* species and a comprehensive genetic engineering approach for the important biofuel producer *C. pasteurianum*.

# Publications

Copies of publications can be found in Appendix Section 6.8.

KM. Schwarz\*, [A. Grosse-Honebrink](#)\* (\*equal contribution), K. Derecka, C. Rotta, Y. Zhang, NP Minton. Towards improved butanol production through targeted genetic modification of *Clostridium pasteurianum*. *Metabolic Engineering*, 2017, doi: 10.1016/j.ymben.2017.01.009

Y. Zhang, [A. Grosse-Honebrink](#) and NP. Minton. A universal *mariner* transposon system for forward genetic studies in the genus *Clostridium*. *PLOS ONE*, 10(4), 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0122411.

A. Poehlein, [A. Grosse-Honebrink](#), Y. Zhang, NP. Minton and R. Daniel. Complete genome sequence of the nitrogen-fixing and solvent-producing *Clostridium pasteurianum* DSM 525. *Genome announcements*, 2015. doi: 10.1128/genomeA.01591-14.



# Acknowledgements

A lot of effort was involved in conceiving this work and it could not have been done without the help of many people, be it professionally or in private. Firstly, I want to thank my supervisors Ying Zhang and Nigel P. Minton for the opportunity of this project and the guidance from my first steps in the lab to the final piece of work. In the same way I want to thank for the funding I received from the University of Nottingham's Vice Chancellor's Scholarship of Research Excellence and the Foreman Hardy Foundation.

I'm deeply grateful to Katrin Schwarz and Anne Henstra for their guidance at the bench but foremost for their friendship and the myriads of tips which made me understand our work and the system we are working in better and which played a huge part in forming my scientific mind. I'm also grateful to Wouter Kuit who taught me how to set up a meticulous experiment and introduced me to my first steps in bioinformatics. A big thank you to Catherine Tuckey for her support during the final stages of this work and for her conscientious corrections. I further want to thank all my colleagues and friends in the Clostridia Research Group, the Beeston Musical Theatre Group, my trusty friends in Switzerland and everyone else that accompanied me on my way and stood by my side with knowledgeable inputs, encouragements, distractions and one or the other pint. I could not have done this without you and I'm fortunate to have so many people that care about and support me.

I'm grateful to my family in Switzerland and Germany especially my grandparents Elisabeth and Johann Schmid and my grandmother Traute Grosse-Honebrink for their support and understanding. Finally, I want to thank my parents, to whom this work is dedicated, Elisabeth and Siegmund Grosse-Honebrink and my brother Thomas Grosse-Honebrink for bearing with me over all these years I spent in education and for supporting me in every way possible to reach my goal. This is as much your work as it is mine!





# Table of Contents

Declaration . . . . .	v
Abstract . . . . .	vii
Publications Associated With This Study . . . . .	ix
Acknowledgements . . . . .	xi
List of Figures . . . . .	xxi
List of Tables . . . . .	xxiv
Abbreviations . . . . .	xxv
<b>1 Introduction</b>	<b>1</b>
1.1 The Genus <i>Clostridium</i> . . . . .	1
1.1.1 <i>Clostridium pasteurianum</i> . . . . .	3
1.1.1.1 PBE Fermentation in <i>C. pasteurianum</i> . . . . .	4
1.1.1.2 Medium Optimisation . . . . .	6
1.1.1.3 Multi-Substrate Feeding . . . . .	8
1.1.1.4 Crude Glycerol Feeding . . . . .	9
1.1.1.5 Chemical Mutagenesis . . . . .	9
1.1.1.6 Genetic Engineering in <i>C. pasteurianum</i> . . . . .	10
1.1.2 Other <i>Clostridium</i> spp. . . . .	11
1.1.2.1 <i>Clostridium acetobutylicum</i> . . . . .	11
1.1.2.2 <i>Clostridium sporogenes</i> . . . . .	12
1.2 Bioenergy . . . . .	13
1.2.1 The Modern Energy Market . . . . .	13
1.2.2 Glycerol - A Waste Product? . . . . .	15
1.2.3 Microbial Glycerol Fermentation . . . . .	16
1.2.3.1 Butanol . . . . .	17
1.2.3.2 1,3-Propanediol . . . . .	17
1.2.3.3 Ethanol and Butyric Acid . . . . .	18
1.3 Tools for Genetic Engineering of <i>Clostridium</i> spp. . . . .	18
1.3.1 Clostridial Roadmap . . . . .	18
1.3.2 Transformation . . . . .	20
1.3.2.1 Restriction-Modification Systems . . . . .	20

## TABLE OF CONTENTS

---

1.3.3	The ClosTron . . . . .	21
1.3.4	Allelic Exchange and Allelic Coupled Exchange . . . . .	23
1.3.4.1	Homologous Recombination . . . . .	23
1.3.4.2	Allelic Exchange . . . . .	24
1.3.4.3	ACE - Allele Coupled Exchange . . . . .	25
1.3.5	CRISPR Mediated Genome Editing . . . . .	27
1.3.6	Transposable Elements . . . . .	29
1.4	Aims of This Study . . . . .	29
<b>2</b>	<b>Materials and Methods</b>	<b>31</b>
2.1	General Microbiology . . . . .	31
2.1.1	Aerobic Strains and Culture Conditions . . . . .	31
2.1.1.1	Preparation of Chemically Competent <i>E. coli</i>	31
2.1.1.2	Transformation of Chemical-Competent <i>E. coli</i>	33
2.1.2	Anaerobic Strains and Culture Conditions . . . . .	33
2.1.2.1	Media Used . . . . .	34
2.1.2.2	Storage of Anaerobic Cultures . . . . .	36
2.1.2.3	Serum Flask Cultures . . . . .	36
2.1.2.4	Batch Fermentation in Bioreactor . . . . .	37
2.1.2.5	Preparation of Electro-Competent <i>C. acetobutylicum</i> . . . . .	38
2.1.2.6	Transformation of Electro-Competent <i>C. acetobutylicum</i> . . . . .	38
2.1.2.7	Preparation of Electro-Competent <i>C. sporogenes</i> . . . . .	38
2.1.2.8	Transformation of Electro-Competent <i>C. sporogenes</i> . . . . .	38
2.1.2.9	Preparation of Electro-Competent <i>C. pasteurianum</i> . . . . .	39
2.1.2.10	Transformation of Electro-Competent <i>C. pasteurianum</i> . . . . .	39
2.2	Molecular Biology Methods . . . . .	39
2.2.1	Plasmid Extraction and Purification . . . . .	39
2.2.2	Genomic DNA Extraction and Purification . . . . .	40
2.2.2.1	Routine Extractions . . . . .	40
2.2.2.2	Phenol-Chloroform Extraction . . . . .	40
2.2.2.3	Ethanol Precipitation . . . . .	40
2.2.3	Analysis of Extracted DNA . . . . .	40
2.2.4	Polymerase Chain Reaction . . . . .	41

	2.2.4.1	Polymerases . . . . .	41
	2.2.4.2	Splicing by Overlap Extension PCR . . . . .	41
	2.2.4.3	Inverse PCR . . . . .	43
	2.2.4.4	Colony PCR . . . . .	43
	2.2.4.5	Touch-Down PCR . . . . .	43
2.2.5		Cloning Techniques . . . . .	43
	2.2.5.1	Restriction Digestion of DNA . . . . .	43
	2.2.5.2	DNA Blunting Using Klenow Fragment . . . . .	43
	2.2.5.3	DNA Dephosphorylation Using Antarctic Phosphatase . . . . .	43
	2.2.5.4	T4 Ligation . . . . .	44
	2.2.5.5	Agarose Gel Electrophoresis . . . . .	44
	2.2.5.6	Purification of DNA Fragments from Agarose Gel . . . . .	44
	2.2.5.7	Purification of DNA Fragments from PCR Products . . . . .	44
2.3		Plasmid Construction . . . . .	44
	2.3.1	pMTL-KS12-tcdR . . . . .	44
	2.3.2	pMTL-AGH12 . . . . .	46
	2.3.3	pMTL-AGH15 . . . . .	46
	2.3.4	pMTL-AGH15-c30550 . . . . .	46
	2.3.5	pMTL-KS12-rexcomp . . . . .	47
2.4		Mutant Construction in <i>C. pasteurianum</i> . . . . .	47
	2.4.1	Allele Coupled Exchange . . . . .	47
		2.4.1.1 Truncation of <i>pyrE</i> . . . . .	47
		2.4.1.2 <i>pyrE</i> Correction in Gene-Deletion Mutants . . . . .	48
		2.4.1.3 Rex Mutant Complementation . . . . .	48
		2.4.1.4 Genome Insertion of <i>tcdR</i> . . . . .	48
	2.4.2	Allelic Exchange Employing <i>pyrE</i> . . . . .	49
	2.4.3	Allelic Exchange Employing <i>codA</i> . . . . .	49
	2.4.4	Minimal Inhibitory Concentration Testing . . . . .	49
	2.4.5	Procedure of Allelic Exchange with <i>codA</i> . . . . .	50
		2.4.5.1 SNP Correction of CLPA_c30550 . . . . .	50
	2.4.6	Transposon Mutagenesis . . . . .	50
		2.4.6.1 Transposon Mutant Library Generation . . . . .	50
		2.4.6.2 Transposon Library Analysis . . . . .	51
		2.4.6.3 Auxotroph Mutant Screen . . . . .	51
		2.4.6.4 Sporulation/ Germination Deficient Mutant Screen . . . . .	52

## TABLE OF CONTENTS

---

2.4.7	Calculations of Growth and Fermentation Products . . .	52
2.4.8	High Performance Liquid Chromatography . . . . .	52
2.4.9	Chloramphenicol Acetyltransferase Assay . . . . .	53
2.4.9.1	Sample Preparation by Cell Lysis . . . . .	53
2.4.9.2	Analysis of CAT Assay . . . . .	53
2.4.9.3	Calculations . . . . .	54
2.5	<i>In Silico</i> Methods . . . . .	55
2.5.1	Genome Sequencing and Resequencing . . . . .	55
2.5.1.1	Illumina Sequencing of Strains . . . . .	55
2.5.1.2	Acquisition of Other Sequencing Data . . . . .	55
2.5.1.3	Read Mapping . . . . .	55
2.5.1.4	Small Nucleotide Polymorphism Calling . . . . .	56
2.5.2	Other Bioinformatic Tools . . . . .	56
2.5.3	Data Analysis and Visualisation . . . . .	56
<b>3</b>	<b>Development of a Hypertransformable Strain of <i>Clostridium pasteurianum</i></b>	<b>59</b>
3.1	Introduction . . . . .	59
3.1.1	Bacterial Competence . . . . .	59
3.1.1.1	<i>C. pasteurianum</i> Restriction Systems . . . . .	60
3.1.2	Electroporation . . . . .	61
3.1.2.1	Electroporation of Bacteria . . . . .	63
3.1.3	Aim of This Study . . . . .	64
3.2	Results . . . . .	65
3.2.1	Development of Electro-Transformation . . . . .	65
3.2.1.1	Testing Different Wild Type Strains . . . . .	65
3.2.1.2	Screening for a Hypertransformable Clone . . . . .	65
3.2.1.3	Optimisation of the Transformation Protocol . . . . .	66
3.2.2	Next Generation Sequencing of Relevant Strains . . . . .	68
3.2.2.1	Genomic Variant in <i>C. pasteurianum</i> -H1 . . . . .	72
3.2.2.2	Genomic Variants in <i>C. pasteurianum</i> -H3 and -H4 . . . . .	73
3.2.2.3	Genomic Variants in the Pyne Genome . . . . .	73
3.2.3	Growth Characteristics of Hypertransformable Clones . . . . .	74
3.2.4	SNP in CLPA_c30550 is the Direct Cause of Hypertransformability in <i>C. pasteurianum</i> . . . . .	76
3.2.4.1	<i>codA</i> for Counterselection . . . . .	76
3.2.4.2	Correction of SNP in <i>C. pasteurianum</i> -H3 . . . . .	77
3.2.5	<i>In Silico</i> Analysis of Gene CLPA_c30550 . . . . .	79

3.3	Discussion . . . . .	81
3.3.1	Initial Electroporation Experiments . . . . .	81
3.3.2	Hypertransformable Sub-population . . . . .	82
3.4	Conclusion . . . . .	85
<b>4</b>	<b>Metabolic Engineering of <i>C. pasteurianum</i> for Improved Solvent Production</b>	<b>87</b>
4.1	Introduction . . . . .	87
4.1.1	<i>spo0A</i> - Sporulation Master Regulator . . . . .	88
4.1.2	<i>rex</i> - Redox Responsive Repressor . . . . .	89
4.1.3	<i>hyd</i> - Iron Only Hydrogenase . . . . .	90
4.1.4	<i>dhaBCE</i> - Glycerol Dehydratase . . . . .	92
4.1.5	Aim of This Study . . . . .	94
4.2	Results . . . . .	94
4.2.1	Mutant Development . . . . .	94
4.2.1.1	<i>pyrE</i> Repair . . . . .	95
4.2.1.2	<i>rex</i> -Complementation . . . . .	95
4.2.2	Spo0A Deletion Phenotype . . . . .	96
4.2.3	Rex Regulator . . . . .	98
4.2.3.1	<i>In silico</i> Analysis . . . . .	98
4.2.3.2	Rex Deletion Phenotype . . . . .	104
4.2.4	Hydrogenase . . . . .	116
4.2.4.1	Hydrogenase Deletion Phenotype . . . . .	116
4.2.5	Glycerol Dehydratase . . . . .	122
4.2.5.1	Glycerol Dehydratase Deletion Phenotype . . . . .	122
4.2.6	1,2-Propanediol Analysis . . . . .	128
4.3	Discussion . . . . .	131
4.3.1	Spo0A Sporulation Master Regulator . . . . .	131
4.3.2	Rex Regulator . . . . .	134
4.3.3	Hydrogenase Deletion . . . . .	136
4.3.4	Glycerol Dehydratase . . . . .	139
4.3.5	1,2-Propanediol . . . . .	140
4.3.6	Butyrate Re-assimilation . . . . .	141
4.4	Conclusion . . . . .	142
<b>5</b>	<b>Development of a <i>Mariner</i>-Transposon System in <i>Clostridium</i> spp.</b>	<b>143</b>
5.1	Introduction . . . . .	143
5.1.1	The <i>Clostridium</i> Transposon System . . . . .	144

## TABLE OF CONTENTS

---

5.1.2	Aim of This Study . . . . .	147
5.2	Results . . . . .	147
5.2.1	Transposon Mutagenesis in <i>C. acetobutylicum</i> . . . . .	147
5.2.1.1	Assessment of Random Insertions . . . . .	147
5.2.1.2	Isolation of Phenotypic Mutants . . . . .	149
5.2.2	Transposon Mutagenesis in <i>C. sporogenes</i> . . . . .	152
5.2.2.1	Assessment of Random Insertions . . . . .	152
5.2.2.2	Isolation of Phenotypic Mutants . . . . .	153
5.2.3	Assessment of the Transposon System in <i>C. pasteurianum</i> . . . . .	154
5.2.3.1	<i>tcdR</i> -Dependent Expression . . . . .	154
5.2.3.2	The Conditional Replicon . . . . .	159
5.2.3.3	Transposon Mutagenesis in <i>C. pasteurianum</i> . . . . .	160
5.3	Discussion . . . . .	162
5.3.1	Transposon Mutagenesis in <i>C. acetobutylicum</i> . . . . .	162
5.3.2	Transposon Mutagenesis in <i>C. sporogenes</i> . . . . .	164
5.3.3	Transposon Mutagenesis in <i>C. pasteurianum</i> . . . . .	165
5.3.3.1	<i>tcdR</i> Dependent Expression . . . . .	165
5.3.3.2	Conditional Replicon . . . . .	166
5.3.3.3	The Conditional System . . . . .	166
5.3.3.4	The Suicide System . . . . .	167
5.3.4	Conclusion . . . . .	167
<b>6</b>	<b>General Discussion</b> . . . . .	<b>169</b>
6.1	Background . . . . .	169
6.2	Key Findings of This Work . . . . .	171
6.2.1	Initial Steps of the <i>C. pasteurianum</i> Roadmap . . . . .	171
6.2.2	Genotypic Analysis of Hypertransformable Clones . . . . .	172
6.2.3	Development of Reverse Genetics . . . . .	173
6.2.3.1	Creation of Mutants With Improved Solvent Titres . . . . .	174
6.2.3.2	Forward Genetic Studies in Clostridia . . . . .	177
6.3	Limitations of This Study and Future Work . . . . .	179
6.4	Concluding Remarks . . . . .	181
	<b>Appendix</b> . . . . .	<b>A</b>
6.5	Supplementary Material Chapter 3 . . . . .	A
6.6	Supplementary Material Chapter 4 . . . . .	D
6.7	Supplementary Material Chapter 5 . . . . .	G
6.8	Publications Originating From This Study . . . . .	N

# List of Figures

1.1	Unrooted phylogenetic tree with a subset of solventogenic and pathogenic clostridia. . . . .	4
1.2	<i>C. pasteurianum</i> central energy metabolism. . . . .	6
1.3	Current energy production and use world-wide, with focus on the United Kingdom. . . . .	14
1.4	Reaction mechanism of acid- and base-catalysed transesterification reaction. . . . .	16
1.5	Main fermentation products of <i>C. pasteurianum</i> using glycerol. . . . .	17
1.6	The clostridial roadmap implemented in <i>C. pasteurianum</i> . . . . .	19
1.7	The mechanism of retrohoming a mobile group II intron. . . . .	22
1.8	Allele-coupled exchange (ACE) procedure. . . . .	26
1.9	CRISPR mediated adaptive bacterial immunity. . . . .	28
3.1	Theory of electroporation with an example of exponential decay and square wave and a scheme of an electroporator . . . . .	62
3.2	Results of screening for hypertransformable strains. . . . .	68
3.3	Histogram of transformation efficiencies of exponential decay versus square wave electroporation. . . . .	69
3.4	Conserved domains of CLPA_c13710 and CLPA_c33080. . . . .	72
3.5	Bottle fermentation analysis of different hypertransformable strains and WT <i>C. pasteurianum</i> DSM 525 . . . . .	74
3.6	Minimal inhibitory concentration determination of 5-FC and 5-FU. . . . .	76
3.7	Minimal inhibitory concentration testing of 5-FC when <i>codA</i> is present in the strain. . . . .	77
3.8	Histogram showing transformation efficiencies of <i>C. pasteurianum</i> -H3 SNP corrected mutants. . . . .	78
3.9	Conserved domains of CLPA_c30550. . . . .	79
3.10	Phyre2 predicted 3D model of CLPA_c30550 with amino acid residue of interest (F1102). . . . .	80

4.1	Proposition of nomenclature of the reductive part of the <i>dha</i> regulon in <i>C. pasteurianum</i> DSM 525 genome. . . . .	93
4.2	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta spo0A$ , <i>C. pasteurianum</i> $\Omega spo0A$ and the wild type in serum bottles in CGM with glycerol . . . . .	97
4.3	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta spo0A$ , <i>C. pasteurianum</i> $\Omega spo0A$ and wild type in serum bottles in Biebl medium glycerol . . . . .	99
4.4	Clustal $\omega$ alignment of reported Rex proteins. . . . .	102
4.5	The sequence logo of Rex DNA recognition motif. . . . .	104
4.6	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta rex$ and <i>C. pasteurianum</i> DSM 525 in bioreactor on glycerol . . . . .	105
4.7	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta rex$ and <i>C. pasteurianum</i> DSM 525 in bioreactor on glucose . . . . .	107
4.8	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta rex$ , <i>C. pasteurianum</i> $\Omega rex$ and <i>C. pasteurianum</i> DSM 525 in serum bottles on glycerol . . . . .	110
4.9	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta rex$ , <i>C. pasteurianum</i> $\Omega rex$ and <i>C. pasteurianum</i> DSM 525 in serum bottles on glucose . . . . .	111
4.10	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta hyd$ and <i>C. pasteurianum</i> DSM 525 in bioreactor on glycerol . . . . .	117
4.11	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta hyd$ and <i>C. pasteurianum</i> DSM 525 in bioreactor on glucose . . . . .	118
4.12	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta hyd$ , <i>C. pasteurianum</i> $\Omega hyd$ and <i>C. pasteurianum</i> DSM 525 in serum bottles on glycerol . . . . .	120
4.13	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta hyd$ , <i>C. pasteurianum</i> $\Omega hyd$ and <i>C. pasteurianum</i> DSM 525 in serum bottles on glucose . . . . .	121
4.14	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta dhaBCE$ , <i>C. pasteurianum</i> $\Omega dhaBCE$ and <i>C. pasteurianum</i> DSM 525 in serum bottles on glycerol . . . . .	123
4.15	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta dhaBCE$ and <i>C. pasteurianum</i> DSM 525 in bioreactor on glucose . . . . .	126
4.16	Growth and solvent profiles of <i>C. pasteurianum</i> $\Delta dhaBCE$ , <i>C. pasteurianum</i> $\Omega dhaBCE$ and <i>C. pasteurianum</i> DSM 525 in serum bottles on glucose . . . . .	127
4.17	1,2-propanediol analysis in mutants and wild type. . . . .	130



5.1	The pMTL-YZ14 transposon system used in clostridia. . . . .	145
5.2	Mode of action of the <i>Clostridium</i> transposon system. . . . .	146
5.3	Mini-transposon insertion loci on <i>C. acetobutylicum</i> chromosome from 60 individual mutants. . . . .	148
5.4	Putative tricarballic acid cycle in anaerobic bacteria. . . . .	150
5.5	Mini-transposon insertion loci on <i>C. sporogenes</i> chromosome from 60 individual mutants. . . . .	152
5.6	CAT assay in <i>E. coli</i> , showing CatP activity when its gene is expressed from different promoters in different growth phases. . . . .	155
5.7	CAT assay of <i>C. pasteurianum</i> -H1 with and without <i>tcdR</i> on the chromosome to test gene expression from $P_{tcdB}$ . . . . .	158
5.8	Test of conditional replicon under permissive and non-permissive conditions. . . . .	159
5.9	Inverse PCR of mini-transposon insertions in <i>C. pasteurianum</i> . . . . .	161
A.1	Multiblast analysis of <i>C. pasteurianum</i> cluster <i>agrB-agrD-resE</i> . . . . .	B
A.2	Growth and product formation of <i>C. pasteurianum</i> $\Delta dhaBCE$ and <i>C. pasteurianum</i> DSM 525 in bioreactor on glycerol . . . . .	D
A.3	Butanol toxicity assay of <i>C. pasteurianum</i> -H1 in 2xYTG. . . . .	F
A.4	Growth curve comparison between <i>C. pasteurianum</i> -H1 $\Omega tcdR$ 16 and <i>C. pasteurianum</i> -H1 on glucose and glycerol . . . . .	G
A.5	Growth curve for sampling of chloramphenicol acetyltransferase (CAT) assay of <i>E. coli</i> . . . . .	H
A.6	Growth curve and sampling points of CAT assay testing $P_{tcdB}$ in <i>C. pasteurianum</i> . . . . .	I



# List of Tables

1.1	A selection of biotechnologically valuable <i>Clostridium</i> species.	2
1.2	A selection of pathogenic clostridia and clostridial strains with medical interest. . . . .	3
2.1	Bacterial strains used in this study. . . . .	32
2.2	Oligonucleotides used in this study. . . . .	42
2.3	Plasmids used in this study . . . . .	45
3.1	Electroporation parameters tested for transformation optimisation . . . . .	67
3.2	Summary reports of Illumina read-mapping of different <i>C. pasteurianum</i> strains against reference genome. . . . .	70
3.3	SNPs called in different strains against <i>C. pasteurianum</i> reference. . . . .	71
3.4	Growth characteristics and product yields of different hypertransformable strains and WT <i>C. pasteurianum</i> DSM 525 . . . . .	75
4.1	Growth characteristics and product yields of <i>spo0A</i> deletion and complementation strain . . . . .	100
4.2	Table of putatively Rex regulated genes with target sequence and distance from closest start codon. . . . .	103
4.3	Growth characteristics and product yields of bioreactor fermentations . . . . .	113
4.4	Growth characteristics and product yields of fermentations in serum bottles with glycerol. . . . .	114
4.5	Growth characteristics and product yields of fermentations in serum bottles with glucose . . . . .	115
4.6	Product formation comparison between different <i>C. pasteurianum</i> hydrogenase studies . . . . .	137
5.1	Transposon insertion sites in <i>spo/ger</i> <sup>-</sup> and auxotrophic mutants of <i>C. acetobutylicum</i> . . . . .	149

## LIST OF TABLES

---

5.2	Growth of <i>C. acetobutylicum</i> -Tn- <i>citB</i> on medium supplemented with different tricarboic and amino acids. . . . .	151
5.3	Transposon insertion sites in <i>spo/ger</i> <sup>-</sup> and auxotrophic mutants of <i>C. sporogenes</i> . . . . .	153
5.4	SNPs called in different mutant strains filtered against SNPs in progenitor strain. . . . .	157
A.1	Electroporation buffers and other agents used for electroporation in different bacterial species. . . . .	C
A.2	Data from Sandoval et al. (2015) fermentations . . . . .	E
A.3	Transposon insertion sites and concomitantly disrupted genes in <i>C. acetobutylicum</i> . . . . .	J
A.4	Transposon insertion sites and concomitantly disrupted genes in <i>C. sporogenes</i> . . . . .	K
A.5	Summary reports of Illumina read-mapping of different <i>C. pasteurianum</i> strains against reference genome. . . . .	L

# Abbreviations

<b>GHG</b>	greenhouse gas
<b>1,3-PDO</b>	1,3-propanediol
<b>1,2-PDO</b>	1,2-propanediol
<b>3-HPA</b>	3-hydroxypropionaldehyde
<b>3-HP</b>	3-hydroxypropionic acid
<b>ABE</b>	acetone, butanol, ethanol
<b>NHOC</b>	net heat of combustion
<b>CDEPT</b>	<i>Clostridium</i> -directed enzyme pro-drug therapy
<b>PTT</b>	polytrimethylene terephthalate
<b>EMS</b>	ethane methyl sulfonate
<b>NTG</b>	<i>N</i> -methyl- <i>N'</i> nitro- <i>N</i> -nitrosoguanidine
<b>R-M systems</b>	restriction-modification systems
<b>RT</b>	reverse-transcriptase
<b>RAM</b>	Retrotransposition-Activated Marker
<b>ORF</b>	open reading frame
<b>FOA</b>	5-fluoroorotic acid
<b>DOG</b>	2-deoxygalactose
<b>5-FC</b>	5-fluorocytosine
<b>5-FU</b>	5-fluorouracil
<b>ACE</b>	Allele Coupled Exchange
<b>RHA</b>	right homology arm
<b>LHA</b>	left homology arm
<b>CRISPR</b>	Clustered Regularly Interspaced Short Palindromic Repeat
<b>Cas</b>	CRISPR associated
<b>TE</b>	transposable element
<b>DSMZ</b>	Deutsche Sammlung von Mikroorganismen und Zellkulturen
<b>ATCC</b>	American Type Culture Collection

## LIST OF TABLES

---

<b>PWM</b>	Position Weight Matrix
<b>Em</b>	erythromycin
<b>Tm</b>	thiamphenicol
<b>Kan</b>	kanamycin
<b>NMR</b>	Nuclear magnetic resonance spectroscopy
<b>SNP</b>	Single Nucleotide Polymorphism
<b>InDel</b>	insertion/deletion
<b>SMC</b>	structural maintenance of chromosome
<b>IPTG</b>	$\beta$ -D-1-thiogalactopyranoside
<b>CAT</b>	chloramphenicol acetyltransferase
<b>TCA</b>	tricarboxylic acid
<b>PBS</b>	Phosphate Buffer Saline
<b>RBS</b>	ribosomal binding site
<b>DPA</b>	dipicolinic acid
<b>TraDIS</b>	transposon directed insertion-site sequencing
<b>ITR</b>	inverted terminal repeat
<b>FDR</b>	flanking direct repeat

# Chapter 1

## Introduction

“Ob mir durch Geistes Kraft und Mund  
Nicht manch Geheimnis würde kund;  
Da ich nicht mehr mit saurem Schweiß  
Zu sagen brauch, was ich nicht weiß;  
Da ich erkenne, was die Welt  
Im Innersten zusammenhält,  
Schau alle Wirkenskraft und Samen,  
Und tu nicht mehr in Worten kramen.”

---

J.W. von Goethe, *Faust*

### 1.1 The Genus *Clostridium*

The genus *Clostridium*, which is classified into the phylum Firmicutes, class Clostridia, order Clostridiales, family Clostridiaceae, includes a phylogenetically wide variety of organisms, which were grouped together on the basis of a small set of features. These are: possession of a genome low in GC content (the majority around 28 mole%), being obligate anaerobes, rod-shaped and Gram-positive, able to form an endospore and being unable to undertake dissimilatory sulphate reduction (Jones and Keis, 2005). However, with data emerging from whole genome sequencing, this simplistic classification is increasingly being challenged. Recently, for example, *C. difficile* was reclassified into the family of *Peptostreptococcaceae*. However, the species' genus is unchanged (*Clostridium*), which can lead to confusion (Yutin and Galperin, 2013).

The genus *Clostridium* has been separated into 19 clusters with the three main clusters I, XI and XIVa (Keis et al., 1995; Stackebrandt et al., 1999; Collins et al., 1994). In current research special attention is paid to pathogenic strains for medical purposes and on the biotechnologically important solvent-producers.

**Table 1.1** – A selection of biotechnologically valuable *Clostridium* species.

Name	Type strain	Cluster	Notes	Key-references
<i>C. pasteurianum</i>	DSM 525 ATCC 6013	I	Glycerol and sugar fermentation to butanol, ethanol and 1,3-propanediol (1,3-PDO); Atmospheric nitrogen fixation	Nakas et al. (1983); Dabrock et al. (1992); Biebl (2001)
<i>C. butyricum</i>	DSM 10702 ATCC 19398	I	Type-strain of the genus; Glycerol fermentation to acids and 1,3-PDO; Butyric acid producer	Biebl et al. (1992); Keis et al. (1995)
<i>C. kluyveri</i>	DSM 555 ATCC 12489	I	Ethanol and acetate fermentation to butyrate, caproate and H <sub>2</sub> amongst others; Grows on <i>n</i> -propanol	Barker and Beck (1942); Seedorf et al. (2008); Weimer and Stevenson (2012)
<i>C. beijerinckii</i>	DSM 791 ATCC 25752	I	ABE-fermentation	Ezeji et al. (2007); Wang et al. (2012)
<i>C. acetobutylicum</i>	DSM 792 ATCC 824	I	ABE-fermentation; Best studied solventogenic <i>Clostridium</i>	Obrien and Morris (1971); Nolling et al. (2001); Lutke-Eversloh and Bahl (2011)
<i>C. ljungdahlii</i>	DSM 13528 ATCC 55383	I	C1- (CO, CO <sub>2</sub> ) and H <sub>2</sub> -fermenter (syngas)	Tanner et al. (1993); Köpke et al. (2010); Liew et al. (2013)
<i>C. autoethanogenum</i>	DSM 10061 JA1-1	I	C1- (CO, CO <sub>2</sub> ) and H <sub>2</sub> -fermenter (syngas) producing ethanol and 2,3-butanediol	Abrini et al. (1994); Bruno-Barcena et al. (2013); Liew et al. (2013)
<i>C. saccharobutylicum</i>	DSM 13864 ATCC BAA117	I	ABE-fermentation; Formerly classified as <i>C. acetobutylicum</i>	Keis et al. (2001); Poehlein et al. (2013)
<i>C. saccharoperbutylacetonicum</i>	DSM 14923 ATCC 27021	I	ABE-fermentation	Ishizaki et al. (1999); Tashiro et al. (2004)
<i>C. aurantibutyricum</i>	DSM 793 ATCC 17777	I	ABE-fermentation	Somrutai et al. (1996); George et al. (1983)
<i>C. thermocellum</i>	DSM 1237 ATCC 27405	III	Possesses cellulosome; Ferments cellulose to ethanol	Bayer et al. (1985); Feinberg et al. (2011)



**Table 1.2** – A selection of pathogenic clostridia and clostridial strains with medical interest.

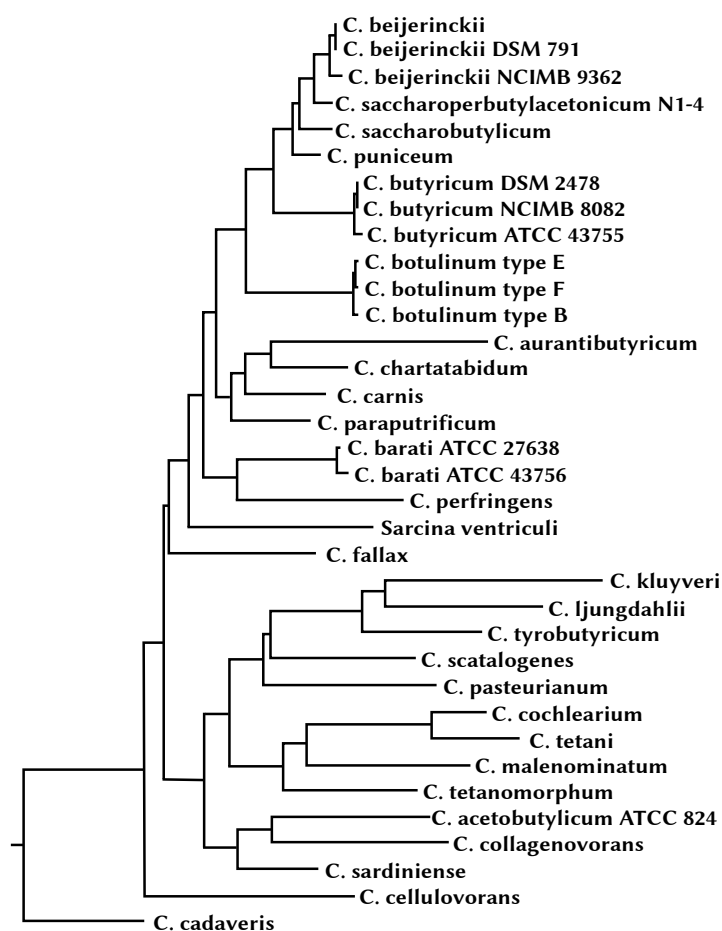
Name	Type strain	Cluster	Notes	Key-references
<i>C. difficile</i>	DSM 1296 ATCC 9689	XI	Most important cause of nosocomial infection; Symptoms: from watery diarrhoea to pseudomembranous colitis	Poxton et al. (2001); Poutanen (2004)
<i>C. botulinum</i>	ATCC 3502	I	Food-borne pathogen; Causes botulism; Symptoms: blurred vision, bilateral flaccid paralysis, flaccid paralysis of respiratory cardiac muscle	Peck (2010); Xiao et al. (2011)
<i>C. sporogenes</i>	DSM 795 ATCC 3584	I	Non-pathogenic; Used as pro-drug delivery vector in cancer treatment	Minton (2003); Theys et al. (2006)
<i>C. tetani</i>	ATCC 19406	I	Causes tetanus; Symptoms: from rigidity and muscle spasms to autonomic dysfunction	Cook et al. (2001)
<i>C. perfringens</i>	DSM 726 ATCC 13124	I	Causes gas gangrene; Symptoms: fever, pain, edema, myonecrosis, and gas production	Rood and Cole (1991)

Tables 1.1 and 1.2 give an overview over the most important strains with their application and selected key references.

The phylogenetic tree proposed by Keis et al. (1995) in Figure 1.1 shows a similar structure to cluster I proposed by Collins et al. (1994) and it can be seen that *Clostridium pasteurianum*, the main focus of this study, is closely related to *Clostridium acetobutylicum* a well-studied solventogenic organism for which molecular genetic tools have been available for several decades. Thus, insights gained in *C. acetobutylicum* might prove of relevance to *C. pasteurianum*.

### 1.1.1 *Clostridium pasteurianum*

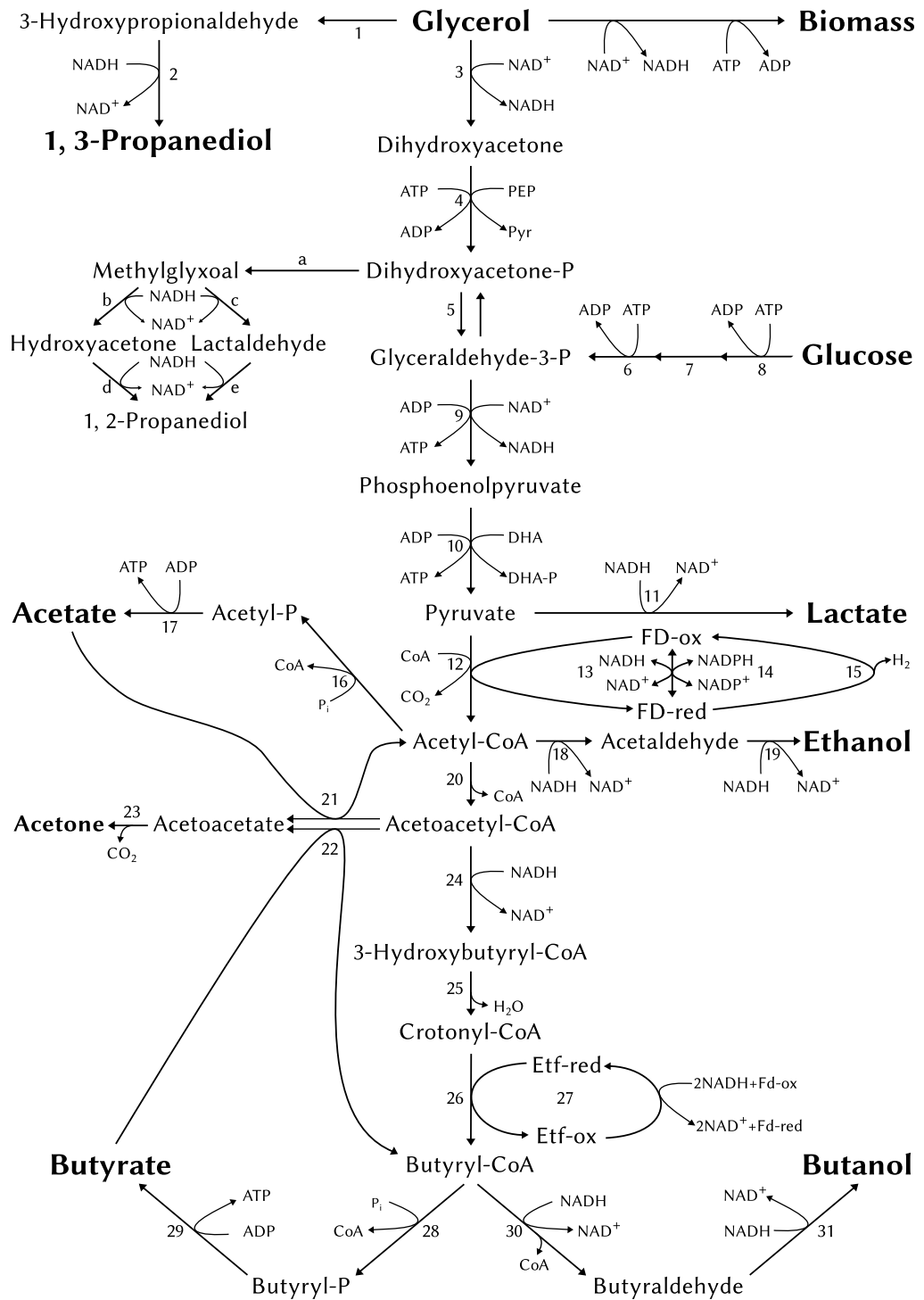
*C. pasteurianum* was first described in 1895 by Sergey Nikolayevich Winogradsky and named in honour of the French microbiologist Louis Pasteur (Winogradsky, 1895). Initial interest resided in the ability of *C. pasteurianum* to fix atmospheric nitrogen (Zelitch, 1951). However, in recent years interest has shifted to the ability of the strain to produce solvents from glycerol, which was first mentioned by Nakas et al. (1983). The pathway for glycerol utilisation in *C. pasteurianum* was proposed by Dabrock et al. (1992) (Figure 1.2).



**Figure 1.1** – Unrooted phylogenetic tree with a subset of solventogenic and pathogenic clostridia. For the tree a part of the 16s rRNA gene was examined (positions 830 to 1383; *E. coli* numbering). The tree resembles cluster I proposed by Collins et al. (1994). Figure adapted from Keis et al. (1995).

### 1.1.1.1 PBE Fermentation in *C. pasteurianum*

The 1,3-propanediol- butanol- ethanol (PBE) fermentation pathway in *C. pasteurianum* from glycerol is divided in two pathways, the reductive and the oxidative pathway. In the oxidative branch glycerol is oxidised by glycerol dehydrogenase to dihydroxyacetone which is further phosphorylated by dihydroxyacetone kinase. Dihydroxyacetone-phosphate enters glycolysis to produce pyruvate in six steps. The oxidation of pyruvate by the pyruvate-ferredoxin oxidoreductase yields acetyl-Coenzyme A (acetyl-CoA). From acetyl-CoA three possible pathways lead to acetate, ethanol or butyrate and butanol in a similar manner to the acetone, butanol, ethanol (ABE)-fermentation pathway in *C. acetobutylicum*. In the reductive branch glycerol is dehydrated to the toxic intermediate 3-hydroxypropionaldehyde (3-HPA) and further reduced by 1,3-propanediol dehydrogenase to 1,3-propanediol (1,3-PDO). It was shown that



**Figure 1.2** – *C. pasteurianum* central energy metabolism. Full description see opposite page.

**Figure 1.2** – Possible acid and solvent metabolism in *C. pasteurianum* from glucose and glycerol based on reported pathway for glycerol utilization in *C. pasteurianum* Biebl (2001) and *Klebsiella pneumonia* (Vasconcelos et al., 1994). Enzymes involved are numbered as follows: 1, glycerol dehydratase; 2, 1,3-propanediol oxidoreductase. 3, glycerol-3-phosphate dehydrogenase; 4, dihydroxyacetone kinase; 5, triose-phosphate isomerase; 6, phosphofructokinase; 7, phosphoglucose isomerase; 8, hexokinase; 9, glyceraldehyde-3-phosphate dehydrogenase; 10, pyruvate kinase; 11, lactate dehydrogenase; 12, pyruvate-ferredoxin oxidoreductase; 13, ferredoxin-NADP reductase; 14, NADPH-ferredoxin oxidoreductase; 15, ferredoxin hydrogenase; 16, phosphate acetyltransferase; 17, acetate kinase; 18, acetaldehyde dehydrogenase; 19, ethanol dehydrogenase; 20, thiolase; 21, acetoacetyl-CoA:acetate:CoA transferase; 22, acetoacetyl-CoA:butyrate:CoA transferase; 23, acetoacetate decarboxylase; 24,  $\beta$ -hydroxybutyryl-CoA dehydrogenase; 25, crotonase; 26, butyryl-CoA dehydrogenase; 27, electron transfer flavoprotein; 28, phosphotransbutyrylase; 29, butyrate kinase; 30, butyraldehyde dehydrogenase; 31, butanol dehydrogenase. The possible pathways to 1,2-propanediol as proposed by Pyne et al. (2016c) consist of: a, methylglyoxal synthase; b, aldo/keto reductase; c, methylglyoxal reductase; d, glycerol dehydrogenase; e, 1,2-propanediol dehydrogenase.

---

at least 9 % of the initial glycerol has to be converted to 1,3-PDO for reducing equivalents of both the fermentative pathway and biomass accumulation to be balanced (Biebl, 2001).

### 1.1.1.2 Medium Optimisation

Until recently no genetic tools were available for *C. pasteurianum*, therefore, early research efforts were limited to culture optimisation and the deployment of chemical mutagens as a means to isolate the strains with favourable phenotypes. Biebl (2001) tested basic growth conditions in batch and continuous culture. Growth and fermentation time was very similar over the tested pH range from pH 5.0 to pH 7.5. When growing the cultures at pH 4.5 the fermentation time increased considerably from around 20 h to over 50 h in batch culture. This confirmed an earlier finding by Mallette et al. (1974) who found an initial pH of 5.8 - 6.5 in a pH-uncontrolled batch culture to be favourable for growth. A pH of 6.0 was found to favour butanol production whereas neutral pH (pH 7.0 - 7.5) showed highly increased ethanol yields. Biebl (2001) showed that the initial glycerol concentration should not exceed 60 g/l. Higher levels led to product inhibition and consequently reduced product yields. A

glycerol concentration of 54.2 g/l favoured butanol production over 1,3-PDO whereas all the other tested concentrations were slightly in favour of 1,3-PDO. In continuous culture in chemostats, a low dilution rate favoured 1,3-PDO production and higher dilution rates butanol and ethanol production. Dilution rates higher than 0.225 h<sup>-1</sup> considerably decreased solvent yields. Glycerol feed showed maximal butanol production at 400- 800 mM, whereas all over glycerol recovery was highest at 800 mM glycerol feed. Glycerol fermentation in pH auxostats, where medium was supplied in periods of pH correction by alkali addition, led mainly to 1,3-PDO whereas butanol and ethanol were by-products. (Biebl, 2001)

A very similar study was published recently by Johnson and Rehm (2016), however, in this case different results were obtained to those of Biebl (2001). In the pH range tested, from 4.7 to 5.9, yields of butanol were highest at a lower pH (5.0) with the highest levels of 1,3-PDO being obtained at a pH of 5.9. Ethanol yields were found not to be significantly influenced by pH, but low pH (4.7) favoured ethanol production. From their findings it can be concluded that a low pH should be employed to maximise butanol and minimise by-product formation. Discrepancies between the two studies were explained by the high glycerol concentrations used by Biebl (2001) (50 g/l, compared to 30 g/l by Johnson and Rehm (2016)) which led to product inhibition and a masking of the pH effect.

Dabrock et al. (1992) tested the influence of iron limitation on growth and solvent production of *C. pasteurianum* DSM 525 with glucose or glycerol as sole carbon sources. With both substrates, increased production of lactate was found when iron was limiting. Solvent and butyrate production was decreased under this condition. However, with glycerol as sole carbon source 1,3-PDO production changed dramatically under iron limitation with an increase of around 225 % compared to excess iron. The cause of this behaviour is not known but it is hypothesised that the alcohol dehydrogenases could be iron dependent. Iron limitation would then make it crucial for the 1,3-PDO pathway to be more active to regulate the internal redox balance. (Dabrock et al., 1992) This finding was supported by Moon et al. (2011) in their multi-factorial analysis of medium composition. Dabrock et al. (1992) further found that increased carbon monoxide (CO) levels increased solvent production levels, which was attributed to an inhibition of hydrogenase activity (further discussed in Chapter 4).

Moon et al. (2011) employed a fractional factorial experiment to find the optimal medium compositions for production of butanol and 1,3-PDO. They found that concentrations of the following three factors influence fermentation

patterns significantly:  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ ,  $(\text{NH}_4)_2\text{SO}_4$  and yeast extract. Employing the widely used M2 medium (Baer et al., 1987) as control and starting point they found that butanol production was favoured by 0.06 g/l of  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 7.35 g/l of  $(\text{NH}_4)_2\text{SO}_4$ , and 5 g/l of yeast extract resulting in an average productivity of  $1.01 \text{ g l}^{-1}\text{h}^{-1}$  compared to  $0.7 \text{ g l}^{-1}\text{h}^{-1}$  in the control medium. 1,3-PDO production, on the other hand, was favoured if  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  and  $(\text{NH}_4)_2\text{SO}_4$  were omitted and 8.0 g/l of yeast extract was added, resulting in an average maximal productivity of  $1.19 \text{ g l}^{-1}\text{h}^{-1}$  compared to  $0.15 \text{ g l}^{-1}\text{h}^{-1}$  in the control. (Moon et al., 2011)

### 1.1.1.3 Multi-Substrate Feeding

Two groups, Kao et al. (2013) with *C. pasteurianum* CH4 and Sabra et al. (2014) with *C. pasteurianum* DSM 525, investigated the possibility of multi-substrate feeding of *C. pasteurianum*. It is known that butyrate concentrations have an impact on butanol fermentation in *C. acetobutylicum* where butyrate addition to the growth medium can increase butanol yields by assimilation of butyrate (Bahl et al., 1982; Hartmanis et al., 1984). Both groups showed that this holds true in *C. pasteurianum*-fermentation with an increase in productivity from 0.38 g/l to 0.79 g/l of butanol when adding 6 g/l butyrate (Kao et al., 2013) and an increase from 1.85 g/l to 7 g/l when adding 5 g/l butyrate (Sabra et al., 2014). However, butyrate is relatively expensive and thus the ability of the organism to convert glucose to butyrate can be used to make the process more economical. Both sequential- and dual-substrate fermentation were tried. In sequential fermentation the culture is initially grown on glucose and produces mainly acids, after which glycerol is added and solvents and 1,3-PDO are produced. This gave final concentrations of 5.5 g/l and 5.9 g/l of butanol and 1,3-PDO, respectively (Kao et al., 2013). Dual-substrate fermentation with addition of glucose and glycerol in the initial medium, on the other hand gave 13.2 g/l butanol (initial concentrations of 60 g/l glycerol and 20 g/l glucose) for Kao et al. (2013) and the highest reported final butanol concentration without *in-situ* removal of solvents of 21.1 g/l (initial concentrations of 50 g/l glycerol and 50 g/l glucose). Both groups also used crude substrates to verify their method. Kao et al. (2013) reached final butanol concentrations of 11.8 g/l by growing the bacteria on 25 g/l enzymatic hydrolysate of acid-pretreated bagasse (lignocellulose) and 60 g/l crude glycerol. Experiments undertaken by Sabra et al. (2014) achieved final butanol concentrations of 17.4 g/l when grown on 50 g/l spruce hydrolysate in combination with 50 g/l technical grade glycerol. (Kao et al., 2013; Sabra et al., 2014)

#### 1.1.1.4 Crude Glycerol Feeding

Crude glycerol, derived directly from biodiesel production, is a mixture of glycerol, alcohol (usually methanol), water, inorganic salts (as residues from catalysts), free fatty acids, non-reacted mono-, di- and triacylglycerols, methyl esters, and a variety of other organic materials (Chatzifragkou and Papanikolaou, 2012; Pagliaro and Rossi, 2010). Some of these contaminants were reported to inhibit growth or solvent formation of *C. pasteurianum*. It was found that methanol and inorganic salts (KCl and K<sub>2</sub>SO<sub>4</sub>) had no effect on *C. pasteurianum* growth and only a very limited decrease of 1,3-PDO yield was reported (Venkataramanan et al., 2012). Butanol yield was not affected by these compounds. The fatty acids stearic-, oleic- and linoleic acid on the other hand had strong effects on growth and fermentation products with the most profound effects being caused by unsaturated oleic- (one double-bond) and linoleic acids (two double-bonds) (Venkataramanan et al., 2012). Unsaturated fatty acids are present in crude biodiesel derived glycerol as a consequence of a failure to undergo transesterification. Crude glycerol should, therefore, be partially purified by acid precipitation. Fermentation using glycerol prepared in this way showed the same growth and butanol yield as does growth on technical glycerol, although, 1,3-PDO yield was reduced (Venkataramanan et al., 2012). Taconi et al. (2009) reported similar behaviour using crude glycerol that only contained methanol, salt and water as contaminants which is comparable to the acid precipitated crude glycerol.

Jensen et al. (2012a) investigated other possibilities for crude glycerol purification. These included: carbonation similar to the method used in sugar juice purification; electrodialysis for salt removal; addition of activated stone carbon for adsorption of toxic compounds and as protective attachment surface for growing microorganisms; and prolonged storage of crude glycerol. Of these techniques storage and activated carbon were advantageous to the growth and fermentation of *C. pasteurianum* and combination of both conditions lead more effective glycerol fermentation than could be achieved with technical grade glycerol. (Jensen et al., 2012a)

#### 1.1.1.5 Chemical Mutagenesis

Chemical mutagenesis of *C. pasteurianum* have only been reported recently. Jensen et al. (2012b) used ethane methyl sulfonate (EMS) to isolate a mutant showing greater resistance to crude glycerol and in turn showing 33 % and 46 % increased 1,3-PDO and butanol production, respectively, compared to the wild type grown on pretreated glycerol (Jensen et al., 2012b). Malaviya et al.



(2012) screened for butanol hyper-producing strains using *N*-methyl-*N'* nitro-*N*-nitrosoguanidine (NTG) and found a 37 % increased butanol titre when compared to the wild type under non-optimised conditions. When optimising the continuous culture and running it for 720 h, a maximal butanol productivity of 7.8  $\text{gl}^{-1}\text{h}^{-1}$  at a dilution rate of 0.9  $\text{h}^{-1}$  was achieved (Malaviya et al., 2012).

#### 1.1.1.6 Genetic Engineering in *C. pasteurianum*

In recent years several studies focused on genetic engineering of *C. pasteurianum*. Pyne et al. (2013) developed a transformation protocol for the organism based on electroporation with a maximal transformation efficiency of  $7.5 \times 10^4$  cfu/ $\mu\text{g}$  plasmid DNA. They used *in vivo* (*E. coli*) M.FnuDII methylation and addition of glycine as a cell-wall weakening agent in the growth phase. Chapter 3 will describe this transformation protocol in detail. Based on successful transformation the same group published several tools in *C. pasteurianum* and applied these tools to inactivate crucial genes. A group II intron system (based on pSY6 developed by Shao et al. (2007)) was used to inactivate the type II restriction system CpaAI (Pyne et al., 2014a) and the resulting mutant was able to accept unmethylated DNA. The group II intron system was further used to disrupt *dhaT*, the gene encoding the 1,3-propanediol oxidoreductase catalysing the second step in the reductive branch from 3-hydroxypropionaldehyde to 1,3-PDO. This inactivation was only possible with concurrent intron insertion in a separate gene (*speA*) thought to be essential under the growth conditions employed. In theory, the reductive pathway should be inactivated in this mutant, while in reality 1,3-PDO production was only reduced by 81 % compared to the wild type rather than entirely abolished. This *dhaT* mutant did, however, produce more butanol (approximately 30 % more compared to wild type). The study, furthermore, found traces of 1,2-propanediol (1,2-PDO) in culture supernatants of all tested strains. 1,2-PDO concentrations were higher in the *dhaT* mutant, in which 1,2-PDO was proposed to act as NADH sink in replacement of the knocked down 1,3-PDO pathway. The proposed pathway to 1,2-PDO can be seen in Figure 1.2.

The same group developed an anti-sense RNA-mediated knock down technique which they used to down-regulate the *hydA* gene (encoding hydrogenase) leading to increased acid production and a slight increase in the yields of solvents (ethanol and butanol) as well as a decrease in 1,3-PDO concentrations (Pyne et al., 2015). Furthermore, the group identified a type I-B Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) system in the genome of *C. pasteurianum* (Pyne et al., 2016b). This system, as well as the widely used



heterologous CRISPR-Cas9 system from *Streptococcus pyogenes* (Wang et al., 2015; Nagaraju et al., 2016; Xu et al., 2015), was developed for genome editing in *C. pasteurianum* and used to disrupt gene *cpaAI* encoding a restriction endonuclease (Pyne et al., 2016a). Finally, Sandoval et al. (2015) disrupted the sporulation master regulator *spo0A* with an allelic exchange method based on the use of *mazF* as a counter-selection marker (see Section 1.3.4.2). The mutant obtained exhibited increased butanol production and depressed yields of 1,3-PDO.

### 1.1.2 Other *Clostridium* spp.

Besides *Clostridium pasteurianum*, the use of a further two *Clostridium* species was explored in the present study. They were employed to demonstrate the general applicability of a *mariner*-transposon system to the genus *Clostridium* in Chapter 5 and shall be introduced briefly.

#### 1.1.2.1 *Clostridium acetobutylicum*

According to Jones and Woods (1986) historical review of ABE fermentation, *Clostridium acetobutylicum* was first isolated by Chaim Weizmann between 1912 and 1914 for the purpose of butanol synthesis from starch for rubber production. The production strain was then used during the World War I to produce acetone as colloidal solvent for nitrocellulose which was used in smokeless gunpowder. History tells us that Weizmann's involvement with the UK war effort played a role in the establishment of the state of Israel of which he became the first president (term of office 1949-1952). Interest in acetone declined with the end of the war and ABE-fermentation was again used to produce the solvent butanol until in the 1960s when the petrochemical route out-competed the biological production of this chemical. (Jones and Woods, 1986)

Interest was renewed in the 1980s due to the oil crisis (Lutke-Eversloh and Bahl, 2011) and still today this organism is an important cornerstone of industrial biotechnology. It is a model-organism for anaerobic bacterial research due to the extensive academic literature available (for extensive reviews see Lutke-Eversloh and Bahl (2011) and Papoutsakis (2008)). The type strain *C. acetobutylicum* ATCC 824 is reported to be able to ferment sugars and starch but also whey, pectin, inulin and xylan (Nolling et al., 2001). The organisms genome consists of a 4 Mbp chromosome and a 200 kbp megaplasmid called pSOL1 carrying the solventogenic genes.

Unlike *C. pasteurianum*, where acids and solvents are produced at the same

time, *C. acetobutylicum* has a biphasic metabolism producing acids (lactate, acetate, butyrate) during the exponential growth phase and switching to solvent production (acetone, butanol) in stationary phase. Ethanol is produced in both phases in similar quantities. Other than this the central metabolism from glucose is similar to the one presented in *C. pasteurianum* (Figure 1.2). In batch culture acetate and butyrate are produced during exponential growth and are re-assimilated by CtfA/B (acetoacetyl-CoA:acyl-CoA transferase) to acetyl-CoA and butyryl-CoA, respectively, which can be further metabolised to butanol or ethanol. The acetoacetate produced by the re-assimilation process is further converted to acetone. During acidogenesis the pH drops as a consequence of acid production which leads to a hostile environment to the organism. In poorly buffered medium the pH can drop below 4.5 which leads to ‘acid crash’, the inability of the organism to de-acidify its environment and cell death. If the medium is buffered a metabolic shift takes place, during which acids are re-assimilated and solvents produced, leading to an increase of environmental pH. Simultaneous with solventogenesis, endospore formation is initiated. (Dürre, 2005; Ezeji et al., 2005)

Nowadays, the biotechnological focus for the use of *C. acetobutylicum* is its ability to produce butanol. *C. acetobutylicum* is able to produce between 1- 2 % butanol in batch fermentation from glucose with 2.1 % being the theoretical maximum because of alcohol toxicity (Ezeji et al., 2010). Even though the alcohol toxicity is a major problem limiting solvent generation in this organism most studies to date have focused on the catabolic pathway directly involved with butanol production. The highest butanol concentration reported to date and reached by rational metabolic engineering was 238 mM, compared to 158 mM in the wild type. This was achieved by deleting *solR* and overexpressing *adhE*. (Harris et al., 2001) For an overview of *C. acetobutylicum* strains improved by metabolic engineering see Kuit (2013).

### 1.1.2.2 *Clostridium sporogenes*

Proteolytic *Clostridium sporogenes* is a harmless (biosafety level 1) close relative of the toxin producer *C. botulinum*. The two organisms share 99- 100 % nucleotide similarity of 16S rRNA (Kubiak et al., 2015). Due to its similar sporulation properties, compared to *C. botulinum*, *C. sporogenes* is widely used as a model organism for food safety studies (La Torre et al., 2016). There are also research efforts into using the organism for biofuel production and 5.52 g/l butanol was produced from rice straw (Gottumukkala et al., 2013) and 11.9 g/l butanol together with 12.1 g/l ethanol was produced in a dual substrate fer-

mentation of glucose and glycerol (Kaushal et al., 2017). These initial studies indicate a potential for the organism to be used as a biofuel producer and further research could have valuable impact.

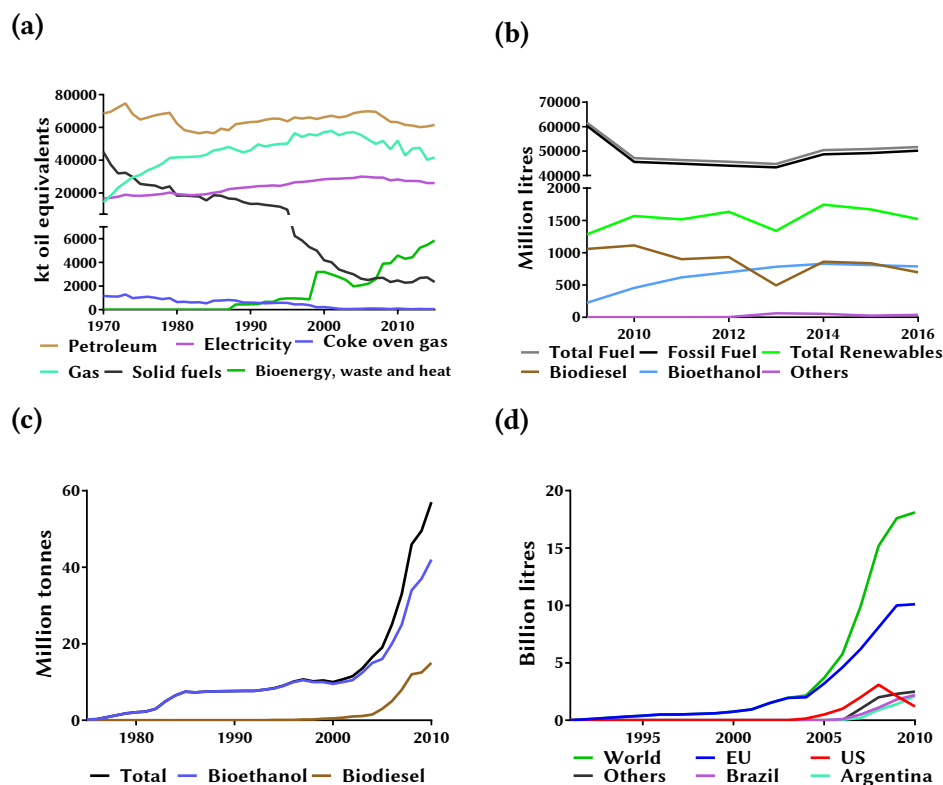
Recently another application was found for this spore producer: a cancer treatment strategy termed *Clostridium*-directed enzyme pro-drug therapy (CDEPT) in which *C. sporogenes* spores are used to deliver therapeutics to the hypoxic/necrotic regions of tumours. *Clostridium* spores will only germinate in hypoxic environment (Kubiak, 2013) which can be used to deliver toxic agents directly to the tumour and exclude toxic effects on healthy tissue. The mechanism envisaged is to administer an intravenous spore suspension. The spores are dispersed throughout the body, but only those that infiltrate tumours will be able to germinate due to the hypoxic environment encountered. After colonization a pro-drug can be given to the patient which, by action of an active pro-drug converting enzyme produced by the vegetative cells in the tumour, is converted to the active drug. (Minton, 2003) This approach was shown to work with pro-drug CB1954 (Theys et al., 2006) and was improved by Heap et al. (2014).

## 1.2 Bioenergy

### 1.2.1 The Modern Energy Market

The scientific evidence for an anthropogenic climate change is accumulating. Unprecedented increase in average global temperatures (Marcott et al., 2013), rising sea-levels (Meysignac et al., 2012), local heat-waves and more frequent droughts (Hansen et al., 2012), are clear signs of a transformation with a possibly devastating outcome. Carbon dioxide (CO<sub>2</sub>) is thought to be the main driver of global warming (Stocker et al., 2013) and regulations to reach the 2 °C Scenario (limiting the global temperature increase to 2 °C) are in operation or envisaged to reduce CO<sub>2</sub> emissions (International Energy Agency, 2016). CO<sub>2</sub> adds to over 90 % of greenhouse gas (GHG) emissions and a significant reduction is thought to stop the increase of the global average temperature (International Energy Agency, 2013).

The energy market contributes over two-thirds of GHG (International Energy Agency, 2013). As a consequence investments into renewable and sustainable energy would have the greatest impact on reducing GHG emissions. The transportation sector, with its dependency on fossil oil, represents an obvious target for improving sustainability. Even though improving vehicle efficiency is the most important element to reduce CO<sub>2</sub> emission, biofuels will have to



**Figure 1.3** – Current energy production and use world-wide, with focus on the United Kingdom. a) Energy Consumption in the UK from 1970 to 2012. Solid fuels include coal, coke and breeze and from 1994 manufactured liquid fuels. Gas includes town gas and natural gas. (data from [Department for Business Energy & Industrial Strategy \(2016\)](#)) b) Contribution of different fuel types to the total transport fuel mix in the UK (Data from ([Renewable Fuel Agency, 2009, 2010, 2011, 2013a,b, 2015, 2016a,b](#))). c) World biodiesel and bioethanol production from 1975 to 2010. Mtoe; million tonnes of oil equivalent. Figure after [Guyomard et al. \(2011\)](#). d) World biodiesel production from 1991 to 2010 with world leading countries. Figure after [Guyomard et al. \(2011\)](#)

play an important role in planes, marine vessels and heavy transport modes ([International Energy Agency, 2011](#)). Transport fuels from renewable sources have been used since the late 19<sup>th</sup> century with the emergence of the modern road vehicles. For example, one of the first engines designed by Rudolph Diesel was modified by the company Otto to run on Peanut-oil ([Knothe et al., 2010](#)). Interest in industrially produced fuels declined over the course of the 20<sup>th</sup> century due to cheaper fossil alternatives. However, with the 1970's oil crisis and approaching of 'peak oil' (maximal oil extraction capacity, with a decline thereafter), predicted to be between 2020 and 2030 ([Miller and Sorrell, 2014](#)), interest in biofuel research and development gathered new momentum.

The UK's energy mix has changed over the last forty years with solid fu-

els such as coal nearly disappearing and increasing amounts of energy being derived from sustainable or recycled sources (Figure 1.3a). However, despite a considerable increase in the production of liquid fuels from renewable sources in the UK, biofuels still only contribute 3 % to the total transportation fuel (Figure 1.3b). Nonetheless, worldwide the production of biofuel has exhibited a steady increase, with bioethanol making the largest contribution (Figure 1.3c). In this respect, the US and Brazil are the main producers. The EU on the other hand is holding the market share of biodiesel production which has increased since the turn of the millennium (Figure 1.3d). Production of biodiesel in the EU has stagnated in recent years, explained by high vegetable oil prices, but is expected to accelerate again (Guyomard et al., 2011).

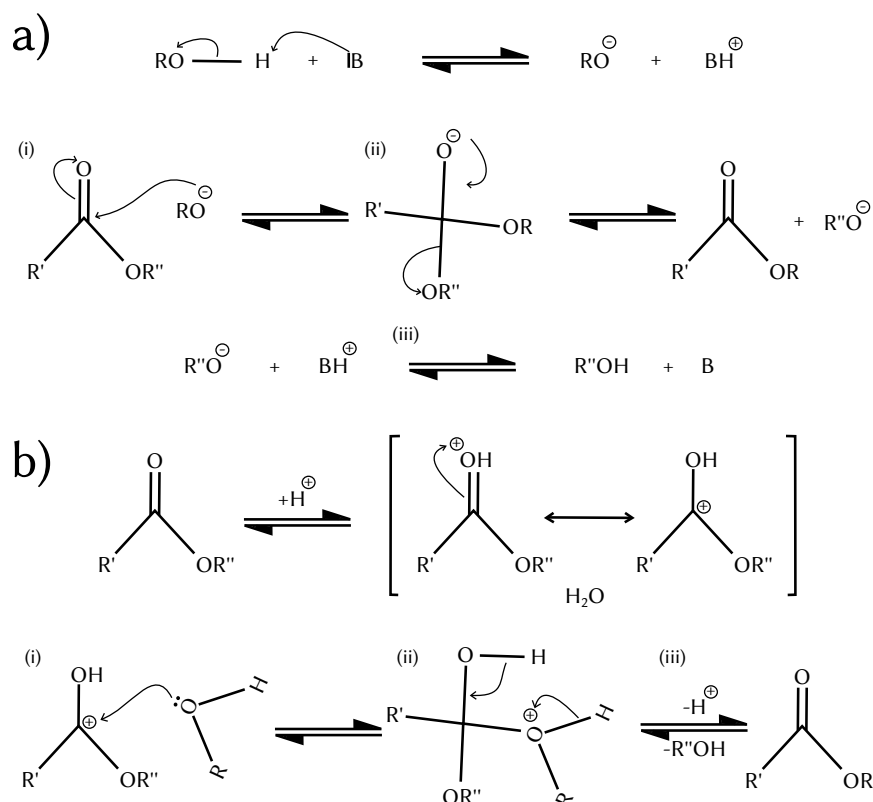
Investing in the development of advanced technologies for biodiesel production and ancillary technologies is therefore an important consideration for European countries if they are to maintain their leading position in biodiesel production. Allied to such an intensification of efforts is the requirement to find further routes to add value to glycerol waste streams generated by biodiesel manufacture. One such route is through the use of *C. pasteurianum* and its ability to convert waste glycerol into the valuable chemical commodities butanol and PDO.

### 1.2.2 Glycerol - A Waste Product?

Generally, biodiesel is chemically synthesised from lipids and alcohols together with an alkaline or acidic catalyst. The lipids originate from vegetable or animal sources and more recently from fats of other plant-like organisms such as algae and cyanobacteria (Issariyakul and Dalai, 2014). These fats or oil, together with alcohol, are converted to esters and glycerol through a transesterification reaction (Figure 1.4).

Depending on the source of the lipid the catalyst should be carefully chosen. Alkali catalysts are most commonly used, due to their low cost and wide availability. However, if the feedstock is rich in free fatty acids, acid catalysts are more suitable (Aransiola et al., 2014). Transesterification theoretically converts one mole of lipids and three moles of alcohol into three moles of esters and one mole of glycerol. However, due to the reversible nature of the reaction, the use of an excess of alcohol is advisable (Freedman et al., 1984).

The process produces 10 kg of glycerol per 100 kg of engine combustible methyl esters. Purified glycerol serves as a component of a multitude of products and as a feed stock for the synthesis of several chemicals. Only 17 years ago, glycerol was regarded as a higher value chemical and microbial produc-

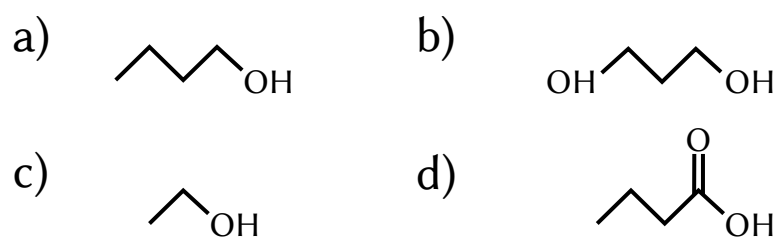


**Figure 1.4** – Reaction mechanism of a) base-catalysed transesterification and b) acid-catalysed transesterification reaction. The breakdown of triglyceride requires three steps: (i) the first step is to produce a tetrahedral intermediate; (ii) the second step is the breakdown of the unstable intermediate tetrahedral into diglyceride ion and fatty acid ester; (iii) the final step is the recovery of the catalyst by proton transfer. Figure adapted from [Aransiola et al. \(2014\)](#).

tion was considered ([Wang et al., 2001](#)). However, with the advent of industrial biodiesel production the price of glycerol fell considerably and it is nowadays considered a waste product with an associated disposal cost ([Yazdani and Gonzalez, 2007](#)). This makes crude glycerol a cheap precursor for microbial fermentation to value-added chemicals which can potentially increase biodiesel process economy while at the same time reducing waste streams of biodiesel production ([Almeida et al., 2012](#)).

### 1.2.3 Microbial Glycerol Fermentation

Several microorganisms are able to anaerobically ferment glycerol to higher value products, e.g. *Klebsiella pneumonia*, *Klebsiella oxytoca*, *Clostridium butyricum*, *C. acetobutylicum*, *Escherichia coli*, *Anaerobiospirillum succiniciproducens*, *Citrobacter freundii* and *Enterobacter aerogenes* ([Clomburg and Gonzalez, 2013](#)). The main products that can be obtained from microbial conversion in-



**Figure 1.5** – Main fermentation products of *C. pasteurianum* using glycerol. a) *n*-butanol, b) 1,3-propanediol, c) ethanol, d) butyric acid

clude 1,3-propanediol, 1,2-propanediol, succinic acid, propionic acid, 1,3-butanediol, ethanol and butanol (Clomburg and Gonzalez, 2013). Further chemicals or precursors which can be produced from glycerol by biotechnological means are dihydroxyacetone, glyceric acid, lactic acid, citric acid, oxalic acid, mannitol, erythritol, arabitol and poly-3-hydroxybutyrate (PHB). For an exhaustive description of each of these compounds see Almeida et al. (2012).

### 1.2.3.1 Butanol

Butanol is a four carbon alcohol ( $C_4H_9OH$ , Figure 1.5a). The alcohol group can be attached to any of the carbons. However, biological conversion mostly produces *n*-butanol with the hydroxyl-group on the terminal carbon. The combustion characteristics are closer to those of petrol than the specificities of shorter chain alcohols with a net heat of combustion (NHOC) of 26.8 MJ/l compared to 32.3 MJ/l for petrol and opposed to 21.1 MJ/l for ethanol and 15.6 MJ/l for methanol. The density of *n*-butanol is 0.88 g/ml and its boiling point and flash point are relatively high with 117 °C and 29 °C, respectively, which makes butanol safer to work than shorter chain alcohols (Harvey and Meylemans, 2011).

Not only is butanol a superior transportation fuel, it can also easily be converted to acrylates, ethers, and butyl acetate which are utilized in paints, resin formulations, and lacquers (Harvey and Meylemans, 2011). This makes butanol an economically valuable product.

### 1.2.3.2 1,3-Propanediol

1,3-propanediol is a three carbon compound with two alcohol groups attached to the first and last carbon ( $CH_2(CH_2OH)_2$ , Figure 1.5b). It is a valuable substrate for the production of polyesters, polyurethans and polyethers (Biebl et al., 1999). Noteworthy is its conversion to polytrimethylene terephthalate (PTT) which is commercially available as Sorona<sup>TM</sup> from Dupont (Biebl et al., 1999; Kraus, 2008). This textile-like material has excellent properties, such as increased resilience, inherent strain resistance and low static generation



(Biebl et al., 1999; Saxena et al., 2009). It could, thus, be used in textiles like carpets, upholstery, or even wearable textiles. It should be noted that even though these materials are sustainably sourced they are not biodegradable (Witt et al., 1994). However, small chemical changes in the structure can lead to increased biodegradability making them more environmentally friendly (Witt et al., 1995).

### 1.2.3.3 Ethanol and Butyric Acid

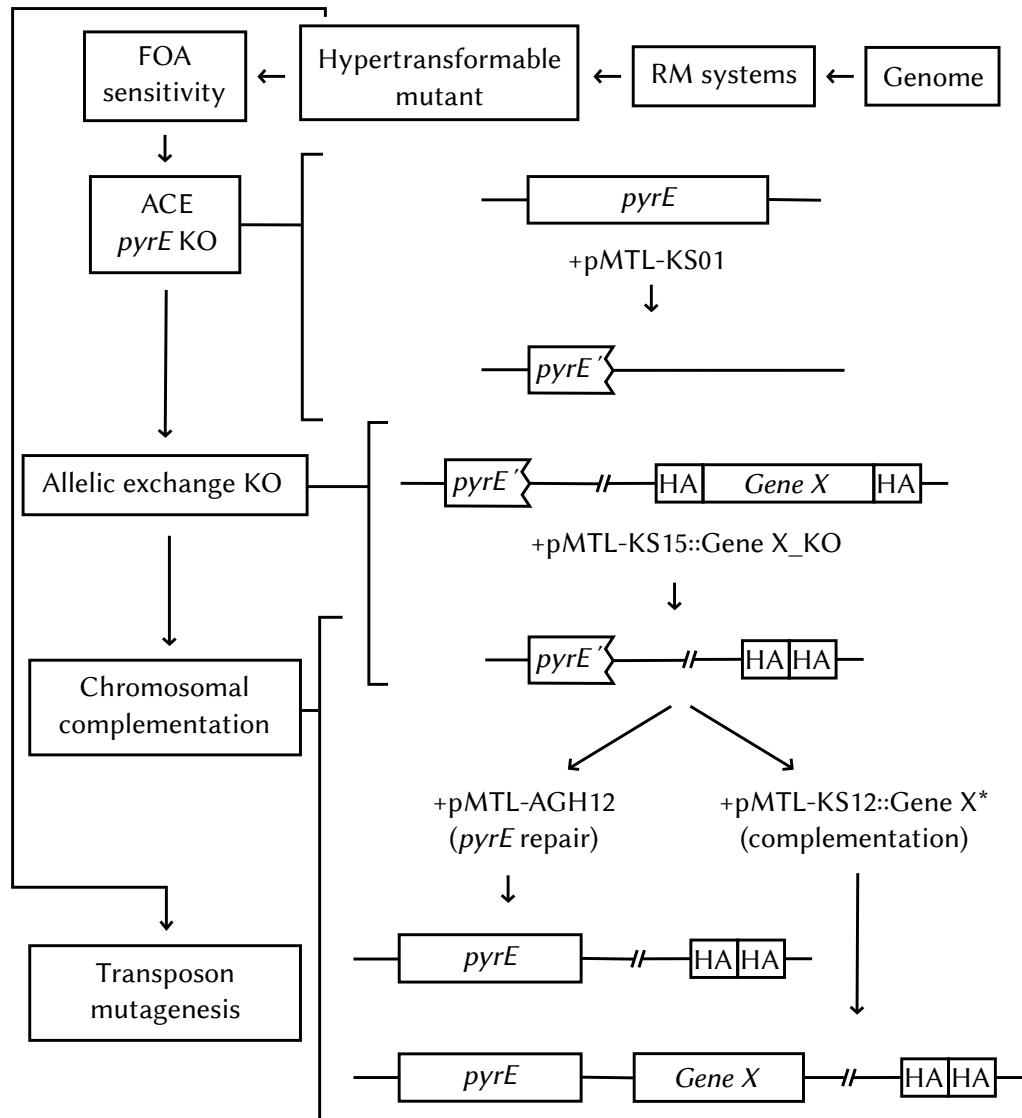
Further compounds produced by *C. pasteurianum* are the two-carbon alcohol ethanol ( $C_2H_5OH$ , Figure 1.5c) and butyric acid ( $C_4H_7OOH$ , Figure 1.5d). Butyric acid is used in the dairy industry or for food additives to enhance fruit flavours. Furthermore, it is used in the textile and plastic industry (Zigová and Šturdík, 2000). Ethanol serves, apart from drinking alcohol, as a solvent in many pharmaceuticals, cosmetics, household products, and flavourings (Equistar Chemicals LP, 2003). Bioethanol is currently the major biofuel produced from sugars, starch, or cellulose (Wang et al., 2013; Kocar and Civas, 2013).

## 1.3 Tools for Genetic Engineering of *Clostridium* spp.

### 1.3.1 Clostridial Roadmap

New biotechnological innovations demand the use of different clostridial species with properties that are appropriate to task at hand. To streamline the process of development of new species and the known synthetic biology tools needed for innovation Minton et al. (2016) proposed a roadmap which is set out to help the scientist making decisions when researching a new organism for biotechnological applications. The *C. pasteurianum* specific implementation is outlined in Figure 1.6. Firstly, the genome of the new organism is sequenced and a high quality genome assembled. This can be used to identify restriction-modification systems against which the plasmids can be protected by methylation. Only then can a transformation procedure be developed by either conjugation from an *E. coli* host or by electroporation. In *C. pasteurianum* electroporation was used but was not effective which led to the development of a screening protocol for hypertransformable mutants not originally in the *Clostridium* roadmap and explained in detail in Chapter 3. Once transformation is established reverse genetic tools can be developed to investigate gene functions. One such tool is the ClosTron, which can be used to inactivate the *pyrE*





**Figure 1.6** – The clostridial roadmap implemented in *C. pasteurianum* (Minton et al., 2016; Schwarz et al., 2017). The genome was published by Poehlein et al. (2015) and restriction modification systems were identified. Plasmid DNA was protected against restriction by *in vitro* methylation with *E. coli* carrying plasmid pMTL-CR1 expressing the M.BepI methylase. Due to low transformation efficiency in the wild type a screen was done for hypertransformable mutants (not in the original roadmap). 5-Fluorouracil (FOA) sensitivity was assayed based on which a *pyrE* truncation strain (*pyrE'*) was produced with the ACE plasmid pMTL-KS01. This strain allowed the use of allelic exchange technology to make knock-outs of genes of interest (Gene X) guided by homology arms (HAs) up- and downstream of the gene. Knock-out mutants were complemented at the *pyrE* locus by repairing the *pyrE* allele and supplying the gene with its native promoter downstream of *pyrE* using plasmids pMTL-KS12::Gene X\*. The *pyrE* allele was repaired to wild type without additional cargo with plasmid pMTL-AGH12.

gene, encoding orotate phosphoribosyl transferase responsible for the conversion of orotic acid into orotidine 5'-monophosphate, and rendering the mutant an uracil auxotroph. 5-fluoroorotic acid (FOA) being an analogue of orotic acid and can be converted by PyrE to 5-fluoroorotidine monophosphate and subsequently to 5-fluorouridine monophosphate which is toxic to the cell. The *pyrE* inactivated mutant is not able to use FOA and is thus resistant.

The concentration of FOA and uracil to be used with such a mutant is assessed at this stage of the roadmap. This is needed for further use with Allele Coupled Exchange (ACE) and allelic exchange methods for gene knock-in and knock-out based on a *pyrE* truncation mutant. Furthermore, forward genetics for the genus is established with a *mariner*-transposon. These techniques allow for comprehensive genetic engineering of an organism by allowing in frame deletion of genes with subsequent complementation of the mutant on the chromosome.

The steps of the clostridial roadmap are explained below and details for *C. pasteurianum* are given.

### 1.3.2 Transformation

Attempts to increase solvent yields are limited in *Clostridium* species where genetic tools are not available. Gene transfer in *C. acetobutylicum* was first reported by Oultram and co-workers in the mid-1980s using conjugative plasmid transfer from *Streptococcus lactis* or *Bacillus subtilis* donors (Oultram and Young, 1985; Oultram et al., 1987). Soon after, Reysset et al. (1988) reported plasmid DNA transfer by protoplast transformation. However, both of these rather laborious methods were soon made redundant with the development of an electro-transformation protocol by Oultram et al. (1988). Due to its simplicity and speed this approach is still in use in most *Clostridium* spp.

In *C. pasteurianum* the first reported example of DNA transfer was based on electroporation (Pyne et al. (2013), see Chapter 3). Before that, Richards et al. (1988) reported high efficiency conjugational plasmid transfer from *Streptococcus lactis* to *C. pasteurianum*, however neither protocol nor formal results were supplied.

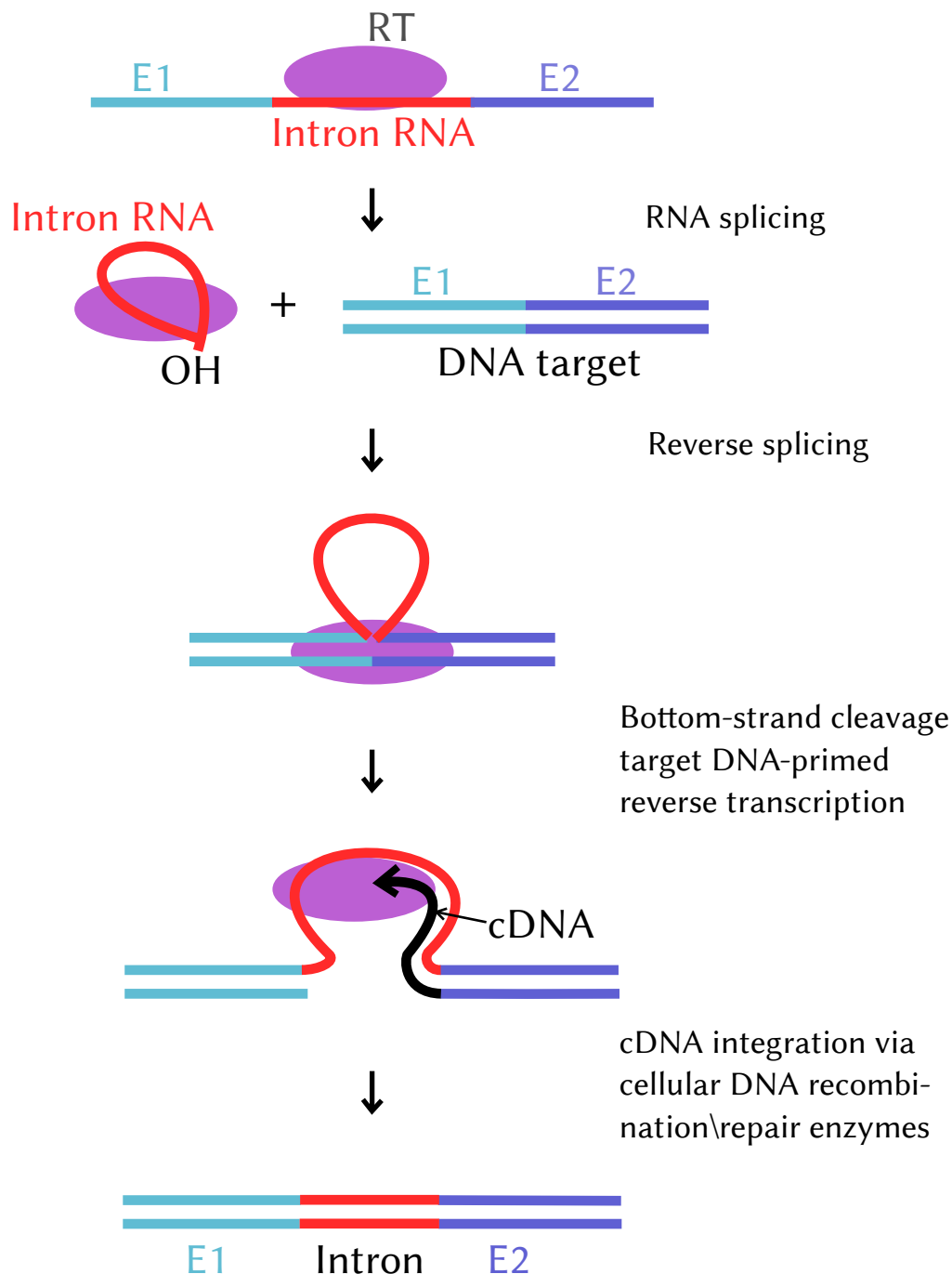
#### 1.3.2.1 Restriction-Modification Systems

For a transformation to be successful the natural barrier of the bacteria against introduction of foreign DNA has to be overcome. Bacteria express restriction-modification systems (R-M systems) which are comprised of pairs of enzymes: an endodeoxyribonuclease (restriction) and a DNA-methyltransferase (modifi-

cation). The methyltransferase adds a methyl group to a nucleotide at its recognition sequence which inhibits restriction by the endodeoxyribonuclease. As newly introduced DNA lacks the methylation pattern of the host this DNA is recognised by the endodeoxyribonuclease and subsequently cleaved and degraded. This very simple system is known as type II-restriction and enzymes found to act accordingly are widely used in laboratories for basic molecular procedures. Besides type II-systems, bacteria may also carry more complex I-, III- and IV-systems (Wilson and Murray, 1991). Type I-systems consist of multifunctional enzymes built from three subunits. The complex acts simultaneously as restriction and modification enzyme. Two enzymes translocate bound DNA in an ATP-dependent manner until they meet, which initiates DNA cleavage. Thus, cleavage sites are independent from recognition sites (Murray, 2000). Type III systems rely on the modification component for sequence recognition (Bickle and Kruger, 1993). In the case of type III-systems, however, the DNA does not need to be methylated if the recognition sequences occur in inversely repeated head-to-head orientation (Rao et al., 2014). Finally, type IV enzymes only cut methylated or otherwise modified DNA (Wilson and Murray, 1991). Restriction modification systems for *C. pasteurianum* are discussed in detail in Section 3.1.1.1.

#### 1.3.3 The ClosTron

The previously mentioned group II intron system is a quick way of disrupting genes in bacteria and was developed independently several times for the use in different *Clostridium* species (Shao et al., 2007; Gupta and Chen, 2008; Heap et al., 2007; Mohr et al., 2013). Most of these systems are based on the TargeTron<sup>®</sup>-system (<http://www.sigmaaldrich.com>, last accessed 19/01/2017) with the mobile group II intron Ll.ltrB of *Lactococcus lactis*. Mohr's system is to be used in *Clostridium thermocellum* a thermophile with a growth optimum around 60 °C, and they used TeI4h\* derived from the thermophile *Thermosynechococcus elongates*. Mobile group II introns consist of a catalytically active intron RNA and a reverse-transcriptase (RT). The intron with the adjacent exons is transcribed and the RT translated. The RT folds the intron RNA into an active catalytic structure which splices itself yielding the ligated exons and the lariat intron RNA tightly bound to the RT. An intricate mechanism, involving several independent recognition sequences, initiates retro-splicing at the target sequence in the host genome starting in one strand before cleaving the other. In the latter, cDNA is synthesised using the 3'-end of the genomic DNA as primer for retro transcription before integration by cellular DNA recombina-



**Figure 1.7** – The mechanism of retrohoming a mobile group II intron. The reverse-transcriptase binds to the intron RNA and assists its self-splicing. The ribonucleoprotein complex then recognises a target sequence in the hosts genome and retro-splices the RNA into one strand. The opposite strand is cleaved and cDNA synthesised by reverse transcription which is then integrated by the hosts recombination and repair enzymes. Figure adapted from [Enyeart et al. \(2014\)](#).

nation and repair mechanisms (see Figure 1.7, Enyeart et al. (2014)). The target mechanism of Ll.ltrB is empirically established and algorithms exist to retarget the mobile intron to disrupt genes of interest by simple SOE-PCR or by ordering synthetic DNA (<http://www.clostron.com>, <http://www.sigmaaldrich.com>, last accessed 19/01/2017).

In *Clostridium* the ClosTron system is widely used. It combines the mobile group II intron with a nested antibiotic resistance gene (erythromycin) interrupted by a group I intron. The latter allows insertion and activation of the resistance gene only if spliced out of the group II intron. This marker, called a Retrotransposition-Activated Marker (RAM), thus overcomes the potentially low intron insertion frequencies. (Heap et al., 2007, 2010) ClosTron usability was exemplified in all major *Clostridium* species (Heap et al., 2010; Minton et al., 2016) and is exemplified in *C. pasteurianum* by Rotta (2014).

### 1.3.4 Allelic Exchange and Allelic Coupled Exchange

The ClosTron is well suited to making rapid gene disruptions. However, it has disadvantages: off-target and multiple insertions were observed and this makes Southern-blot analysis a laborious necessity. Also, the ClosTron mediated disruption of the gene does not guarantee the gene to be inactive depending on the insertion-site within the gene (Enyeart et al., 2014). Furthermore, disruption of a gene might have unforeseen polar effects (Kuehne and Minton, 2012). Finally, it has been reported in *L. lactis* that disruption can be reversed if the *ltrA*-gene stays active in the host, which is due to excision of the intron from the mRNA of the target gene (Frazier et al., 2003).

#### 1.3.4.1 Homologous Recombination

If one wants to obtain gene knock-outs without the aforementioned disadvantages it is advisable to delete the gene cleanly by deleting large stretches of the gene of interest in whole codon steps. Such a deletion (called in-frame deletion) should not interfere with downstream open reading frames (ORFs) and therefore not cause polar effects. It can be achieved by allelic exchange, a method based on the naturally observed homologous recombination between two strands of DNA that show a certain degree of sequence similarity. For an in-depth review of homologous recombination in *E. coli* see Kowalczykowski et al. (1994). Briefly: the recombination event passes through four steps: (i) Initiation: for the recombination to happen single stranded DNA (ssDNA) is required, several enzymes can process the double stranded DNA (dsDNA) to ssDNA. (ii) Homologous pairing: ssDNA of one molecule invades, by aid of

more enzymes, the dsDNA of the other reaction partner and binds to a homologous region. A so called Holliday-junction is built (Holliday, 1964). (iii) Heteroduplex extension: Presumably protein-promoted branch migration extends the DNA in the heteroduplex region in both directions. (iv) Resolution: Nucleolytic resolution of the Holliday-junction is again achieved by enzymes and leads to patched-, with only the initial ssDNA exchanged, or spliced-molecules in which the flanking markers have been exchanged and thus yielding heteroduplex DNA.

#### 1.3.4.2 Allelic Exchange

The combination of two independent homologous recombination events can be used to knock-out genes by in-frame deletion or to knock-in genes by adding a cargo between two homology arms. In readily transformable yeast species gene knock-out can be achieved by transforming linear DNA with two homology sites flanking the gene to be knocked out and carrying an antibiotic resistant marker between the homology arms (Baudin et al., 1993). The marker can then be cut out via the use of flipases (Datsenko and Wanner, 2000). However, most bacterial species including *Clostridium* are not easily transformed with DNA and to achieve the same goal homologous regions have to be transformed on plasmids. This demands a more elaborate mechanism of insertion and plasmid excision because plasmid-loss has to be selected for. The first crossover event will insert the whole plasmid into the host's chromosome, this event can be selected for by plasmid conferred antibiotic resistance and faster growth of single crossover colonies. The second recombination will, depending on which of the homology arms recombination happens, either cut the whole plasmid back out of the genome (if it happens in the same homology arm as the first crossover-event) or will leave the cargo behind (if it happens via the other homology arm). To screen for the second crossover event single-crossover mutants are exposed to negative selection. Mutants still harbouring the plasmid on the genome or in the cytoplasm will be selected against. This allows the selection of excised derivatives in which a double crossover event has resulted in excision of the plasmid and is subsequently lost from the cell. Colonies having lost the plasmid can then be screened for successful double crossover-events against wild type revertants by PCR and sequencing (Heap et al., 2012).

Several independent systems have been developed in different *Clostridium* species depending on different counter-selectable markers. In *C. thermocellum* and *C. difficile* similar systems have been exemplified depending on orotidine-5-phosphatedecarboxylase (PyrF) (*C. thermocellum*) or orotate phosphoribosyl-

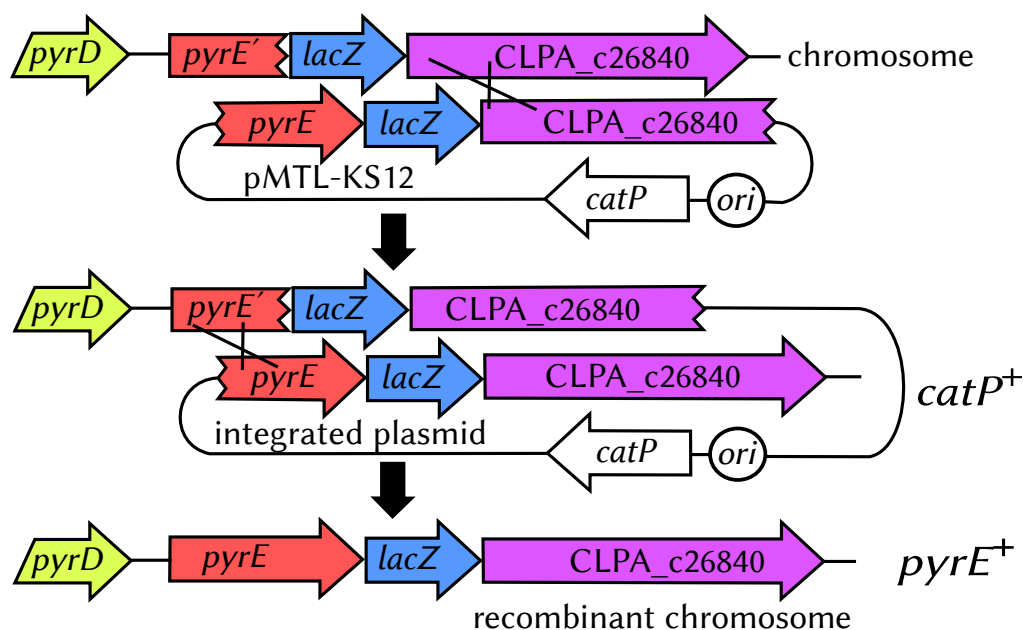
transferase (PyrE) (*C. difficile*) both involved in *de novo* pyrimidine synthesis (Ng et al., 2013; Tripathi et al., 2010). Cells lacking *pyrF* or *pyrE* are uracil auxotrophs. Both enzymes are involved in pyrimidine biosynthesis and thus catalyse the incorporation of fluorinated nucleotides (from 5-fluorouracil FOA) into DNA or RNA. This is lethal and can thus be used for counter-selection. To use the genes for allelic exchange either of the wild type genes have to be inactivated. In *C. perfringens* GalK was applied in a similar fashion. For this the *galKT*-operon has to be deleted in the organism and *galK*-gene from *C. acetobutylicum* can be used as counterselection marker on the plasmid (Nariya et al., 2011). GalK catalyses the phosphorylation of galactose to galactose-1-phosphate in normal metabolism but can also phosphorylate 2-deoxygalactose (DOG) which cannot be metabolised further. The build-up of DOG-1-phosphate is toxic to the cell (Warming et al., 2005). Another gene that can be used is *codA* an enzyme from *E. coli* encoding cytosine deaminase which converts cytosine to uracil. It can similarly convert 5-fluorocytosine (5-FC) into the highly toxic compound 5-fluorouracil (5-FU). However, toxicity to cells relies on the presence of *upp* (uracil phosphoribosyltransferase) and absence of *codA* in the host which might not always be the case (Cartman et al., 2012). A final counter-selectable marker is the *E. coli* derived *mazF*-gene which encodes an mRNA interferase targeting mRNA 5'-ACA sequences, leading to growth arrest and subsequent cell-death (Al-Hinai et al., 2012).

In all introduced systems the counter-selectable marker acts when a compound is added to the growth medium once selection for the double-crossover is desired. Each of the systems has advantages and disadvantages. The dependence on a genotypic background for all but the MazF-system might be called a disadvantage and it was found that FOA introduces chromosome alterations in *Candida albicans* (Wellington et al., 2006) which was not confirmed in *Clostridium* (NP Minton, unpublished data). The MazF system on the other hand relies on an inducible promoter which might not be applicable to all *Clostridium* spp.

#### 1.3.4.3 ACE - Allele Coupled Exchange

With the PyrE-based allelic exchange method described above Heap et al. (2012) developed a simple system for gene knock-in illustrated in Figure 1.8. As in allelic exchange, if the native *pyrE* gene is inactive the bacteria are resistant to FOA. Thus, a background *pyrE* strain was constructed with a truncated *pyrE* gene only containing 300 bp at the 5' end. Now an ACE plasmid was constructed with a longer right homology arm (RHA) (1200 bp) downstream of the *pyrE*-gene and a shorter 300 bp long left homology arm (LHA). A cargo





**Figure 1.8** – Allele-coupled exchange (ACE) procedure. From top to bottom: (i) first crossover event at the 1200 bp right homology arm (RHA) and integration of the whole plasmid; (ii) second crossover event at the shorter (300 bp) left homology arm (LHA) and excision of plasmid backbone. This procedure repairs the formerly truncated *pyrE* gene and cargo can be inserted into the genome instead of the *lacZ* gene.

gene for knock-in is transported between the homology arms as well as the complementation of the *pyrE* gene. The RHA should direct the first crossover event into the downstream region and *catP* mediated thiamphenicol resistance and quicker growing colonies indicates insertion of the plasmid. A second crossover event should then happen in the LHA and if successful restore the *pyrE* gene. The cells are again able to grow on medium lacking uracil. The plasmid is subsequently lost by segregation due to an unstable replicon (Ng et al., 2013; Heap et al., 2012, 2009).

If one wishes to knock-in multiple genes this can be done by creating a new RHA corresponding to the previously inserted cargo downstream of *pyrE* and using the same LHA which after recombination will truncate the *pyrE* gene again rendering the host resistant to FOA. This flip-flop like, subsequent gene delivery was used by Heap et al. (2012) to deliver over 40 kbp of lambda phage DNA in three fragments by in turn restoration and disruption of PyrE activity.

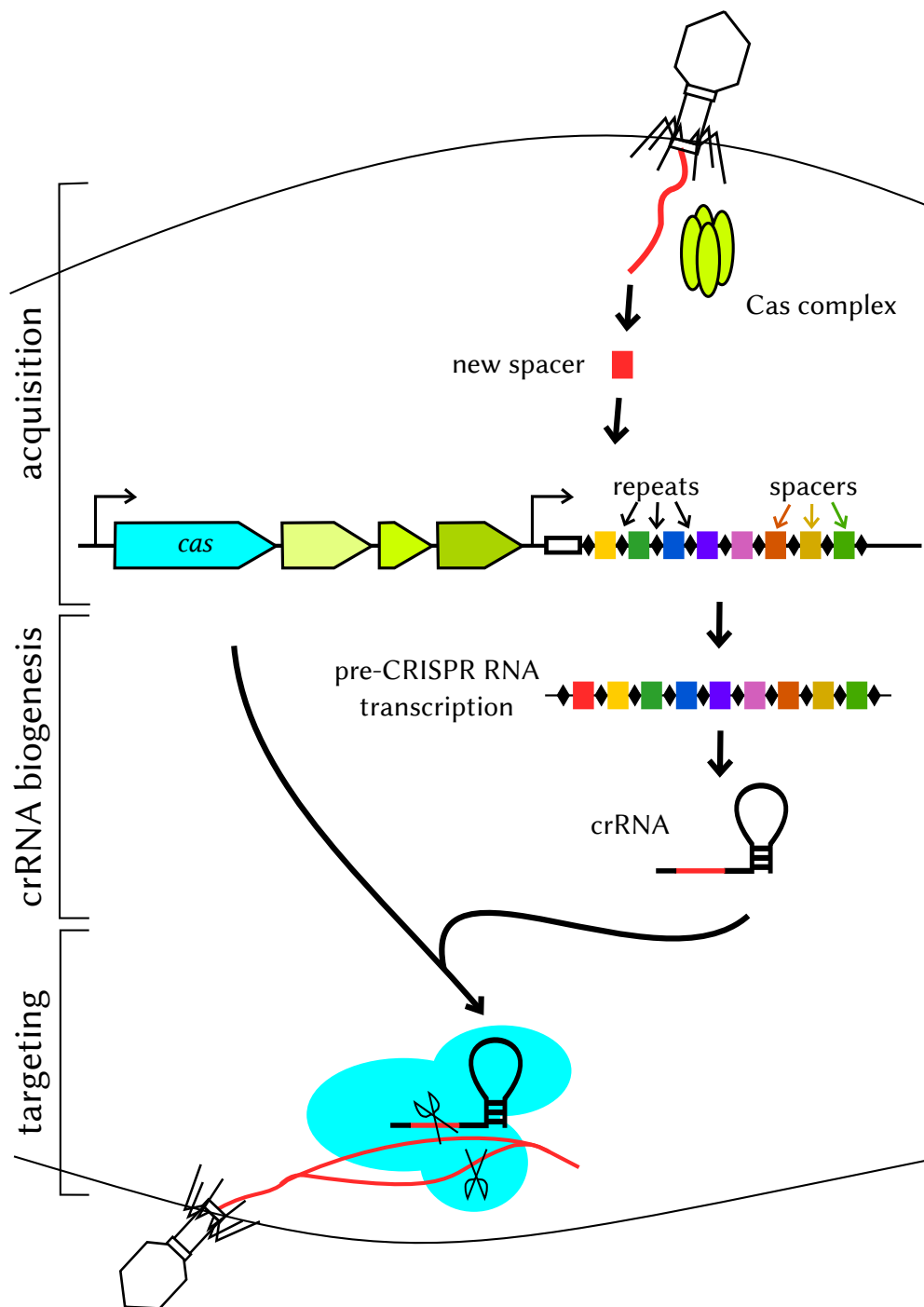


### 1.3.5 CRISPR Mediated Genome Editing

The newest addition in the bacteriologist molecular tool box is CRISPR mediated genome editing. CRISPRs work with CRISPR associated (Cas) proteins to build an adaptive immune system in bacteria (Figure 1.9, Barrangou and Marraffini (2014)). Short nucleic acid sequences from invading pathogens can be captured and integrated in the CRISPR array. The CRISPR array is transcribed and endonucleases process the pre-CRISPR RNA transcript into individual crRNAs. Cas proteins bind the crRNAs which guide the CRISPR complex to the targets where the crRNA binds to the complementary sequence. Cas proteins nucleolytically cleave the crRNA target complex and the invading DNA is degraded. (Makarova et al., 2011)

This system has since been adapted into a powerful tool in genome engineering where editing is achieved by introducing a cell with the Cas protein along with a specially designed guide RNA (gRNA) that directs the cut through hybridization with its matching genomic sequence. When the cell repairs that break, errors can occur to generate a knock-out of that gene or additional modifications can be introduced. For example, CRISPR can be used to improve the above described allelic exchange procedure (Section 1.3.4.2). By expressing a synthetic pre-CRISPR RNA transcript, targeting genes of interest in the organism, a heterologous or endogenous Cas machinery can be guided against a self-gene and cleave it. By concomitant transformation of homology arms up and downstream of the target gene the double-strand break can be repaired by the host's homologue recombination machinery. Such a mutant lacks the target sequence of the CRISPR-Cas machinery and is not cleaved again. Thus, only deletion mutants should be able to survive. (Cong et al., 2013; Jinek et al., 2013)

The CRISPR methodology was developed in several *Clostridium* species. In all species the CRISPR-Cas9 system of *Streptococcus pyogenes* (Jinek et al., 2012) was used with different levels of success. In *C. beijerinckii* high efficient gene deletion (100 %) was achieved targeting the sporulation master regulator *spo0A* (Wang et al., 2015). In *C. autoethaogenum* efficiency was 50 % based on two target genes (*adh* and *2,3-bdh*) both involved in alcohol metabolism (Nagaraju et al., 2016). In the close relative *C. ljungdahlii* editing efficiencies were 100 %, 75 %, 100 % and 50 % with targets *pta*, *adhE1*, *ctf* and *pyrE*, respectively. Xu et al. (2015) found over 95 % efficiency in *C. cellulolyticum* with very short homology arms (0.2 kbp) when targeting *pyrF*. Finally, Pyne et al. (2016a) found the *S. pyogenes* system to be less efficient (25 %) than the endogenous CRISPR-Cas system (100 %) in *C. pasteurianum*. These studies show that genome editing can



**Figure 1.9** – CRISPR mediated adaptive bacterial immunity. In the acquisition phase the Cas complex incorporates foreign DNA ( e.g. from an invading virus) as new spacer into the CRISPR loci. Every spacer is flanked by two repeats. Thus, during incorporation of a new spacer a new repeat is synthesised as well. CRISPR RNA (crRNA) or guide RNA is synthesised from the pre-CRISPR RNA transcript. The transcript is processed by Cas endoribonucleases for crRNA generation. The crRNA matches with the target sequence (so called proto-spacer) and the Cas-machinery nucleolytically cleaves the target DNA (scissors). Figure adapted from Barrangou and Marraffini (2014) and Jinek et al. (2013).

be efficiently achieved with the heterologous *S. pyogenes* CRISPR-Cas9 system accompanying some modifications. Such modifications include: codon optimisation of the Cas enzyme (Nagaraju et al., 2016; Xu et al., 2015) and controlled Cas9 expression with an inducible promoter (Nagaraju et al., 2016), which is also recommended by Pyne et al. (2016a) due to the detrimental effects of high Cas9 expression on *C. pasteurianum*.

### 1.3.6 Transposable Elements

Forward genetics is an approach used to identify genes (or sets of genes) responsible for a particular phenotype of an organism. Tools to do so are transposable elements (TEs), also known as ‘jumping genes’ or transposons, which are sequences of DNA that move (or jump) from one location in the genome to another. The ability to insert randomly into host genomes, makes TEs perfect tools to study genes important for *Clostridium* specific phenotypes, namely sporulation and solvents production. Chapter 5 is dedicated to transposon mutagenesis with a *mariner*-transposon (Cartman and Minton, 2010; Zhang et al., 2015) and will introduce TEs in more detail.

## 1.4 Aims of This Study

This study aimed to develop forward and reverse genetics in *Clostridium pasteurianum*. To apply a genetic tool set, an effective transformation procedure had to be developed first. Initial efforts of this study focused on developing an electro-transformation procedure and subsequently increasing the transformation efficiency in *C. pasteurianum*.

With a successful DNA transfer protocol, reverse genetics could be employed. The aim was to produce deletion mutants of crucial genes in the central metabolic pathways. For a proof-of-principle, the *spo0A* gene was targeted as a clear sporulation deficient phenotype is expected from similar studies in closely related organisms. Once successfully deleted *spo0A*, genes *rex*, *hyd* and *dhaBCE* were chosen to be deleted individually with focus on improved solvent production.

Finally, this study adopted a forward genetics approach using a random mutagenesis *mariner*-transposon system in *C. pasteurianum*. Due to the initial lack of transformation procedure, the *mariner*-transposon system was first exemplified in the established species *C. acetobutylicum* and *C. sporogenes* before it was adapted in *C. pasteurianum*.

The outcome of this study should provide a good basis for further studies in

*C. pasteurianum* and a deeper understanding of its central energy metabolism. Furthermore, the development of the *mariner*-transposon system in three crucial species will accelerate forward genetic studies in *Clostridium*.

# Chapter 2

## Materials and Methods

“THAT’S MORTALS FOR YOU”, Death continued.  
“THEY’VE ONLY GOT A FEW YEARS IN THIS WORLD  
AND THEY SPEND THEM ALL IN MAKING THINGS  
COMPLICATED FOR THEMSELVES. FASCINATING.”

---

Terry Pratchett, *Mort*

### 2.1 General Microbiology

#### 2.1.1 Aerobic Strains and Culture Conditions

*Escherichia coli* was cultured in LB-medium consisting of, per litre: tryptone, 10 g; yeast extract, 5 g; NaCl, 5 g; agar (for plates only), 10 g. Cultures were typically grown at 37 °C and liquid cultures were shaken at 200 rpm. Difficult or toxic constructs were grown at 30 °C or room temperature. Antibiotics were used in following concentrations: erythromycin (500 µg/ml), kanamycin (50 µg/ml) and chloramphenicol (25 µg/ml in agar plates or 12.5 µg/ml in liquid media). *E. coli* cultures were stored at -80 °C on microbeads (Microbank™ Bacterial and Fungal Preservation System, Pro-Lab Diagnostics, UK) according to supplier’s instructions.

##### 2.1.1.1 Preparation of Chemically Competent *E. coli*

Competent *E. coli* were prepared using the CaCl<sub>2</sub> method (Mandel and Higa, 1970). In short: (i) suspend one colony of desired *E. coli* in 5 ml liquid LB medium (with appropriate antibiotic) and incubate over-night shaking at 37 °C; (ii) use 0.5 ml of the over-night to inoculate a 50 ml main culture (LB with appropriate antibiotic); (iii) grow to OD<sub>600</sub> 0.4- 0.6 shaking at 37 °C; (iv) cool on

**Table 2.1** – Bacterial strains used in this study. *C. pasteurianum* is abbreviated to *C. p.*, *C. acetobutylicum* to *C. a.*, *C. sporogenes* to *C. s.* and *E. coli* *E. c.*. Table continued on next page.

Strain	CRG strain designation	Properties	Source
<i>C. p.</i> DSM 525	CRG4091	wild type strain	DSMZ
<i>C. p.</i> DSM 526	CRG4092	wild type strain	DSMZ
<i>C. p.</i> DSM 12136	CRG4093	wild type strain	DSMZ
<i>C. p.</i> DSM 9989	CRG4094	wild type strain	DSMZ
<i>C. p.</i> ATCC 6013	CRG4080	wild type strain	ATCC
<i>C. p.</i> ATCC 6013U	CRG2520	wild type strain	Unilever Ltd.
<i>C. p.</i> H1	CRG4111	hypertransformable strain based on DSM 525, working title H0D0R1	this study
<i>C. p.</i> H1 $\Delta$ <i>pyrE</i> 9	CRG4273	<i>pyrE</i> truncation strain for Allele Coupled Exchange (ACE) and allelic exchange	Schwarz et al. (2017)
<i>C. p.</i> H1 $\Omega$ <i>tcdR</i> 16		<i>tcdR</i> insertion strain for orthogonal expression system based on CRG4273	this study
<i>C. p.</i> H3	CRG5512	hypertransformable strain based on DSM 525	this study
<i>C. p.</i> H3 $\Delta$ <i>pyrE</i> 3	CRG4817	<i>pyrE</i> truncation strain for ACE and allelic exchange	this study
<i>C. p.</i> H3 $\Delta$ <i>pyrE</i> 4	CRG4818	<i>pyrE</i> truncation strain for ACE and allelic exchange	this study
<i>C. p.</i> H3 $\Delta$ <i>pyrE</i> 5	CRG4819	<i>pyrE</i> truncation strain for ACE and allelic exchange	this study
<i>C. p.</i> H3 $\Omega$ <i>tcdR</i> 5		<i>tcdR</i> insertion strain for orthogonal expression system based on CRG4817	this study
<i>C. p.</i> H3 SNP repair 3	CRG5538	SNP in CLPA_c30550 repaired	this study
<i>C. p.</i> H3 SNP repair 4	CRG5539	SNP in CLPA_c30550 repaired	this study
<i>C. p.</i> H3 SNP repair 8	CRG5540	SNP in CLPA_c30550 repaired	this study
<i>C. p.</i> H3 SNP revert 5	CRG5541	SNP in CLPA_c30550 repair assayed but reverted to H3 allele	this study
<i>C. p.</i> H4	CRG5513	hypertransformable strain based on DSM 525	this study
<i>C. p.</i> H4 $\Delta$ <i>pyrE</i> 3	CRG4820	<i>pyrE</i> truncation strain for ACE and allelic exchange	this study
<i>C. p.</i> H4 $\Delta$ <i>pyrE</i> 5	CRG4821	<i>pyrE</i> truncation strain for ACE and allelic exchange	this study
<i>C. p.</i> H4 $\Delta$ <i>pyrE</i> 23	CRG4822	<i>pyrE</i> truncation strain for ACE and allelic exchange	this study
<i>C. p.</i> H4 $\Omega$ <i>tcdR</i> 1	CRG5230	<i>tcdR</i> insertion strain for orthogonal expression system based on CRG4821	this study
<i>C. p.</i> $\Delta$ <i>spo0A</i> $\Delta$ <i>pyrE</i> 4-44	CRG5514	<i>spo0A</i> deletion and <i>pyrE</i> truncation strain	Schwarz et al. (2017)
<i>C. p.</i> $\Delta$ <i>spo0A</i> $\Delta$ <i>pyrE</i> 4-15	CRG5515	<i>spo0A</i> deletion and <i>pyrE</i> truncation strain	Schwarz et al. (2017)
<i>C. p.</i> $\Delta$ <i>spo0A</i> 1	CRG5516	<i>spo0A</i> deletion strain	this study
<i>C. p.</i> $\Delta$ <i>spo0A</i> 2	CRG5517	<i>spo0A</i> deletion strain	this study
<i>C. p.</i> $\Omega$ <i>spo0A</i> 9	CRG5518	<i>spo0A</i> complementation at <i>pyrE</i> locus	Schwarz et al. (2017)
<i>C. p.</i> $\Omega$ <i>spo0A</i> 10	CRG5519	<i>spo0A</i> complementation at <i>pyrE</i> locus	Schwarz et al. (2017)
<i>C. p.</i> $\Delta$ <i>rex</i> $\Delta$ <i>pyrE</i> 6-49	CRG5520	<i>rex</i> deletion and <i>pyrE</i> truncation strain	Schwarz et al. (2017)
<i>C. p.</i> $\Delta$ <i>rex</i> $\Delta$ <i>pyrE</i> 6-5-1	CRG5521	<i>rex</i> deletion and <i>pyrE</i> truncation strain	Schwarz et al. (2017)
<i>C. p.</i> $\Delta$ <i>rex</i> 4	CRG5522	<i>rex</i> deletion strain	this study
<i>C. p.</i> $\Delta$ <i>rex</i> 7	CRG5523	<i>rex</i> deletion strain	this study
<i>C. p.</i> $\Omega$ <i>rex</i> 1	CRG5524	<i>rex</i> complementation at <i>pyrE</i> locus	this study
<i>C. p.</i> $\Omega$ <i>rex</i> 2	CRG5525	<i>rex</i> complementation at <i>pyrE</i> locus	this study
<i>C. p.</i> $\Delta$ <i>hyd</i> $\Delta$ <i>pyrE</i> 47	CRG5526	<i>hyd</i> deletion and <i>pyrE</i> truncation strain	Schwarz et al. (2017)
<i>C. p.</i> $\Delta$ <i>hyd</i> $\Delta$ <i>pyrE</i> 20	CRG5527	<i>hyd</i> deletion and <i>pyrE</i> truncation strain	Schwarz et al. (2017)

**Continued Table 2.1** – Bacterial strains used in this study. *C. pasteurianum* is abbreviated to *C. p.*, *C. acetobutylicum* to *C. a.*, *C. sporogenes* to *C. s.* and *E. coli* *E. c.*.

Strain	CRG strain designation	Properties	Source
<i>C. p.</i> Δ <i>hyd</i> 1	CRG5528	<i>hyd</i> deletion strain	this study
<i>C. p.</i> Δ <i>hyd</i> 2	CRG5529	<i>hyd</i> deletion strain	this study
<i>C. p.</i> Ω <i>hyd</i> 3	CRG5530	<i>hyd</i> complementation at <i>pyrE</i> locus	Schwarz et al. (2017)
<i>C. p.</i> Ω <i>hyd</i> 12	CRG5531	<i>hyd</i> complementation at <i>pyrE</i> locus	Schwarz et al. (2017)
<i>C. p.</i> Δ <i>dhaBCE</i> Δ <i>pyrE</i> 6	CRG5532	<i>dhaBCE</i> deletion and <i>pyrE</i> truncation strain	Schwarz et al. (2017)
<i>C. p.</i> Δ <i>dhaBCE</i> Δ <i>pyrE</i> 10	CRG5533	<i>dhaBCE</i> deletion and <i>pyrE</i> truncation strain	Schwarz et al. (2017)
<i>C. p.</i> Δ <i>dhaBCE</i> 2	CRG5534	<i>dhaBCE</i> deletion strain	this study
<i>C. p.</i> Δ <i>dhaBCE</i> 9	CRG5535	<i>dhaBCE</i> deletion strain	this study
<i>C. p.</i> Ω <i>dhaBCE</i> 9	CRG5536	<i>dhaBCE</i> complementation at <i>pyrE</i> locus	Schwarz et al. (2017)
<i>C. p.</i> Ω <i>dhaBCE</i> 5	CRG5537	<i>dhaBCE</i> complementation at <i>pyrE</i> locus	Schwarz et al. (2017)
<i>C. a.</i> ATCC 824	CRG20	<i>C. acetobutylicum</i> type strain	DSMZ
<i>C. a.</i> ATCC 824Ω <i>tcdR</i>	CRG3011	<i>tcdR</i> insertional mutant	Zhang et al. (2015)
<i>C. s.</i> NCIMB 10696	CRG21	<i>C. sporogenes</i> type strain	DSMZ
<i>C. s.</i> NCIMB 10696Ω <i>tcdR</i>	CRG3821	<i>tcdR</i> insertional mutant	Zhang et al. (2015)
<i>E. c.</i> HB101		Hypercompetent <i>E. coli</i> strain	Promega, UK
<i>E. c.</i> Top10	CRG27	Cloning <i>E. coli</i> strain	CRG
<i>E. c.</i> XL1-Blue		Hypercompetent <i>E. coli</i> strain	Agilent, UK
<i>E. c.</i> M.BepI	CRG3131	Strain harbouring the methylation plasmid to methylate against CpaI restriction Kan <sup>R</sup>	Carlo Rotta

ice for 10 min; (v) centrifuge at 2700 g for 10 min at 4 °C; (vi) decant supernatant and resuspend in 10 ml ice-cold 100 mM CaCl<sub>2</sub> with 20 % glycerol, leave on ice for 15 min; (vii) repeat step vi but resuspend in 2 ml and leave on ice for 1- 2h; (viii) prepare 100 μl aliquots and store at -80 °C for future use or use immediately as described in Section 2.1.1.2.

### 2.1.1.2 Transformation of Chemical-Competent *E. coli*

To transform cells prepared according to Section 2.1.1.1 it was proceeded as follows: (i) add 50 μl plasmid DNA to 1 aliquot of competent cells; (ii) incubate on ice for 30 min; (iii) heat-shock at 42 °C for 45 sec; (iv) incubate on ice for 2 min; (v) add 400 μl warm (37 °C) LB or SOC medium (20 g/L tryptone, 5 g/L yeast extract, 4.8 g/L MgSO<sub>4</sub>, 3.603 g/L dextrose, 0.5 g/L NaCl, 0.186 g/L KCl [Sigma Aldrich, UK]) and incubate shaking at 37 °C for 45 min; (vi) plate on selective LB-plates and incubate at 37 °C or 30 °C as appropriate.

## 2.1.2 Anaerobic Strains and Culture Conditions

*Clostridium* spp. were grown anaerobically at 37 °C in artificial atmosphere of N<sub>2</sub>:H<sub>2</sub>:CO<sub>2</sub> in an 80:10:10 volume% ratio in an anaerobic workbench (Don

Whitley, Yorkshire, UK) using medium pre-reduced over-night under the same conditions. Antibiotics were used in the following concentrations: erythromycin (50  $\mu\text{g/ml}$ ) and thiamphenicol (25  $\mu\text{g/ml}$ ).

### 2.1.2.1 Media Used

**CGM Agar** Clostridial Growth Medium was used as solid medium for *C. acetobutylicum* cultures. For one litre medium the following was added:  $(\text{NH}_4)_2\text{SO}_4$ , 2 g;  $\text{KH}_2\text{PO}_4$ , 0.5 g;  $\text{K}_2\text{HPO}_4$ , 1 g;  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.1 g;  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.015 g;  $\text{CaCl}_2$ , 0.01 g;  $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ , 0.01 g;  $\text{CoCl}_2$ , 0.002 g;  $\text{ZnSO}_4$ , 0.002 g; tryptone, 2 g; yeast extract, 1 g; glucose, 50 g; agar, 12 g. Medium was adjusted to pH 7, and made up with distilled water to 1 l. The following compounds were used as 20 g/l stock solutions:  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ ,  $\text{CaCl}_2$ ,  $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ ,  $\text{CoCl}_2$  and  $\text{ZnSO}_4$ .

**RCM Agar** Reinforced Clostridial Growth Medium was used as solid medium for *C. pasteurianum* cultures. A prepared mix (CM0151) was bought from Oxoid (Thermo Scientific, UK) and used by supplier's instructions. The composition of RCM is per litre: yeast extract, 3 g; 'Lab Lemco' powder (beef extract), 10 g; peptone, 10 g; glucose, 5 g; soluble starch, 1 g; NaCl, 5 g; Na-acetate, 3 g; cystein hydrochloride, 0.5 g; agar, 15 g; pH 6.8.

**CBM** Clostridial Basal Medium was used as solid and liquid defined medium. For one litre the following was added:  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 200 mg;  $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ , 7.58 ml of a 1 mg/ml stock;  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 10 ml of a 1 mg/ml stock; casein enzymatic hydrolysate, 4 g; agar (for plates only), 10 g, make up with distilled water to 880 ml (plates) or 860 ml (liquid). The medium was autoclaved and after autoclaving the following were added: *p*-aminobenzoic acid, 1 ml of a 1 mg/ml stock; biotin, 20  $\mu\text{l}$  of a 0.1 mg/ml stock; thiamine-HCl, 1 ml of a 1 mg/ml stock; glucose, 100 ml of a 50 % w/v stock;  $\text{K}_2\text{HPO}_4$ , 10 ml of a 50 mg/ml stock;  $\text{KH}_2\text{PO}_4$ , 10 ml of a 50 mg/ml stock;  $\text{CaCO}_3$  (for liquid medium only), 20 ml of a 250 g/l stock.

**P2 Minimal Medium** P2 was used as solid medium. For liquid use agar addition was neglected. The following solutions were prepared: Solution 1: 890 ml of desired sugar concentration (for plates add 15 g agar to this solution); the solution was autoclaved. Solution 2:  $\text{H}_2\text{O}$ , 100 ml;  $\text{K}_2\text{HPO}_4$ , 0.5 g;  $\text{KH}_2\text{PO}_4$ , 0.5 g; ammonium acetate, 2.2 g; autoclave solution. Solution 3:  $\text{H}_2\text{O}$ , 100 ml;  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 2 g;  $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ , 0.1 g; NaCl, 0.1 g;  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.1 g; . Solution 4:  $\text{H}_2\text{O}$ , 100 ml; *p*-aminobenzoic acid, 0.1 g; thiamine, 0.1 g; biotin, 0.01 g; filter sterilise solution. Solutions 1 and 2 were autoclaved and solutions 3 and 4 were



filter sterilised. For one litre of P2 medium the following were added to make up one litre of medium: Solution 1, 890 ml; Solution 2, 100 ml; Solution 3, 10 ml; Solution 4, 1 ml.

**MACC - Amino Acid Complete Medium** was used as solid medium for minimal *C. sporogenes* medium. It is based on low phosphate basal medium (LPBM) as described by Lovitt et al. (1987) and adjusted with an amino acid solution described in Dixon et al. (1987). LPBM was prepared by adding to one litre of H<sub>2</sub>O: KH<sub>2</sub>PO<sub>4</sub>, 2 g; K<sub>2</sub>HPO<sub>4</sub>, 2 g; MgCl<sub>2</sub>, 93 mg; (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 5 g. This solution was autoclaved and after autoclaving the following was added: 50 % sterile glucose, 10 ml; 10 % NaHCO<sub>3</sub>, 25 ml; simple vitamin solution, 1 ml; trace element solution, 10 ml; amino acids as described below. The simple vitamin solution was filter sterilised and consists of per litre: thiamine HCl, 50 mg; biotin, 5 mg; *p*-aminobenzoic acid, 5 mg; nicotinic acid, 500 mg. The trace element solution was filter sterilised and consists of per litre: nitrilotriacetic acid, 12.8 g; FeSO<sub>4</sub>·7H<sub>2</sub>O, 0.1 g; MnCl<sub>2</sub>, 63 mg; CoCl<sub>2</sub>·6H<sub>2</sub>O, 0.24 g; CaCl<sub>2</sub>, 75 mg; ZnCl<sub>2</sub>, 0.1 g; CuCl<sub>2</sub>·2H<sub>2</sub>O, 13 mg; H<sub>3</sub>BO<sub>4</sub>, 0.01 g; NaMoO<sub>4</sub>·2H<sub>2</sub>O, 0.01 g; NaCl, 1 g; NaSeO<sub>3</sub>, 0.017 g; NiSO<sub>4</sub>·6H<sub>2</sub>O, 0.026 g; NaWO<sub>4</sub>·2H<sub>2</sub>O, 0.1 g. Amino acids stocks were all autoclaved apart from tryptophane which was filter sterilised. Per litre of final medium the following concentrations of amino acids were added: glycine, 3 mmol; valine, 2 mmol; isoleucine, 2 mmol; arginine, 2 mmol; leucine, 1 mmol; histidine, 1 mmol, methionine, 1 mmol; phenylalanine, 1 mmol; tryptophane, 1 mmol; tyrosine, 0.125 mmol.

**TYG** TYG medium was used as solid and liquid medium for *C. sporogenes* cultures. For one litre the following was added to H<sub>2</sub>O: tryptone, 30 g; yeast extract, 20 g; Na-thioglycolate, 1 g; agar (plates only), 10 g.

**2xYT** 2xYT was used as liquid medium. For one litre the following was added to H<sub>2</sub>O: tryptone, 16 g; yeast extract, 10 g; NaCl, 5 g; adjust pH to pH 6.2; autoclave. After autoclaving add 100 ml sterile 50 % glucose solution (or glycerol).

**2xYP** For *C. pasteurianum* electroporation procedure 2xYP was used instead of 2xYT. Preparation was similar to 2xYT but tryptone was substituted by peptone. Veggie™ yeast extract and Veggie™ peptone for this medium were purchased from Novagen (Merck Milipore, UK).

**Biebl Medium** was produced according to Biebl (2001) with SL7 micronutrient solution according to Biebl and Pfennig (1981). For one litre of medium

the following was added to H<sub>2</sub>O: KH<sub>2</sub>PO<sub>4</sub>, 0.5 g; K<sub>2</sub>HPO<sub>4</sub>, 0.5 g; (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 5 g; MgSO<sub>4</sub>·7H<sub>2</sub>O; CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.02 g; FeSO<sub>4</sub>·7H<sub>2</sub>O, 9.14 mg and, if glycerol was fermented, 1 g yeast extract. This mixture was adjusted to pH 6.2 with HCl and sterilised by autoclaving. After autoclaving 2 ml trace element solution SL7 and 120 ml of a 500 g/l sterile stock of carbon source (glucose or glycerol) was added and depending on culture condition 250 µl of a 0.1 mg/ml biotin stock (filter sterilised, only needed if no yeast extract was used), 20 ml of an autoclaved 250 g/l CaCO<sub>3</sub> stock (not needed in pH controlled bioreactors) and 1 ml of an autoclaved 2 mg/ml stock of resazurin (if redox indication was required). Trace element SL7 solution consists of, per litre: MnCl<sub>2</sub>·H<sub>2</sub>O, 100 mg; ZnCl<sub>2</sub>, 70 mg; H<sub>3</sub>BO<sub>3</sub>, 60 mg; NaMoO<sub>4</sub>·2H<sub>2</sub>O, 40 mg; CoCl<sub>2</sub>·6H<sub>2</sub>O, 28.7 mg; CuCl<sub>2</sub>·2H<sub>2</sub>O, 20 mg; NiCl<sub>2</sub>·6H<sub>2</sub>O, 20 mg; HCl (25 %), 1 ml. The mixture was sterilised by filtration through a 0.2 µm filter.

### 2.1.2.2 Storage of Anaerobic Cultures

*C. acetobutylicum* was stored with as a liquid culture with 15 % glycerol at -80 °C or as spores at room temperature. Spores were produced by growing a 2xYT culture with glucose for five days in the cabinet after which the culture was spun down and washed once with PBS (Phosphate Buffered Saline: phosphate buffer, 0.01 M; KCl, 0.0027 M; NaCl, 0.137 M; pH 7.4, Sigma-Aldrich) before it was stored in PBS. *C. pasteurianum* was stored from an over-night liquid culture with 15 % DMSO at -80 °C. Spores were produced by growing *C. pasteurianum* for 5 days on RCM plates and scraping colonies into PBS; or by growing *C. pasteurianum* in Biebl medium for 5 days and spinning the culture down and washing with PBS before it was stored at room temperature in PBS. *C. sporogenes* was stored as a liquid culture with 15 % glycerol at -80 °C. Spore stocks were produced by growing *C. sporogenes* for 5 days in TYG after which the culture was spun down and washed with PBS before it was stored in PBS at room temperature.

### 2.1.2.3 Serum Flask Cultures

For growth experiments 250 ml serum flasks were used with 50 ml growth medium. The serum flasks were sterilised by autoclaving with a cotton stopper and aluminium foil in the top. The growth medium and the butyl-rubber stoppers were sterilised separately and was then filled into the serum bottles. The bottles with medium were transferred into the anaerobic cabinet for at least 24 h to reduce before inoculation. Precultures were grown in falcon tubes with 20 ml growth medium inside the cabinet. The first tube was inoculated with a

loop of culture from plate and left to grow for a few hours before it was serially diluted up to  $10^{-3}$ . The next day the optical density ( $OD_{600}$ ) was measured and the over-night culture grown to mid-exponential phase was selected to inoculate the serum bottle main culture inside the cabinet. For Biebl medium with glycerol a pre-pre culture was grown in Biebl medium with glucose for 16 h and the initial glycerol culture was inoculated from there to ascertain growth. Depending on growth medium and carbon source different final inoculum densities were chosen: for 2xYT with glycerol and glucose, CGM with glucose, Biebl medium with glucose an inoculum of  $OD_{600} = 0.05$  was aimed for. For Biebl medium with glycerol and inoculum of 0.075 was aimed for. Once inoculated the bottles were closed with sterile butyl-rubber stoppers and transferred out of the cabinet for incubation at 37 °C. Samples for OD, pH and solvent analysis were taken by sterilising the top of the butyl-rubber stopper with an alcohol swab and then taking the sample with a syringe through the rubber.

#### **2.1.2.4 Batch Fermentation in Bioreactor**

Bioreactor cultures were grown in a final medium volume of 350 ml in an Infors HT Multifors 2 bioreactor (Infors UK Ltd., UK). The reactors were prepared with 308 ml Biebl medium without carbon source, micro nutrients and vitamins. The reactors were then autoclaved after which the already sterile carbon source, micronutrient solution and vitamins (for glucose only as yeast extract is added to the medium in glycerol fermentation) were added with a syringe through a spare port to make up to 350 ml. The reactors were then bubbled with nitrogen to reduce at 5 volumes/minute for at least 20 h with stirring at 400 rpm and the medium was held at 37 °C. Redox and pH was monitored online and was adjusted before inoculation. Redox was adjusted if needed to -350 mV with L-cystein and pH was adjusted with HCl and KOH to pH 6. The pH was held constant with KOH during the whole fermentation but and was not regulated with acid. Precultures were prepared similar to the serum flask precultures but 35 ml were grown for the bioreactor. After OD measurement of the over-night culture the appropriate volume of inoculum was pipetted into a syringe with a stopper and the syringe closed with the piston. The syringe was transferred out of the cabinet and used to inoculate the bioreactor through a spare port. During fermentation the stirrer speed was held at 200 rpm, nitrogen was bubbled at 2 v/min and temperature was held at 37 °C. The pH and redox values were measured online and pH was additionally measured offline as a control and  $OD_{600}$  was measured offline. Samples for OD, pH and solvent analysis were taken with a syringe through a valve port.

### 2.1.2.5 Preparation of Electro-Competent *C. acetobutylicum*

*C. acetobutylicum* electro-competent cells were derived by the method of [Mermelstein et al. \(1992\)](#) with slight adjustments. In short: (i) a 5 ml over-night (2xYT with 5 % glucose) culture was inoculated with a loop of wild type *C. acetobutylicum* ATCC 824; (ii) the over-night was serially diluted to  $10^{-4}$  in tubes with 5 ml medium and the cultures were inoculated in the cabinet at 37 °C over-night; (iii) The most dilute over-night that still showed growth was used to inoculate a main-culture of 60 ml (2xYT with 5 % glucose); (iv) the main-culture was grown to OD<sub>600</sub> of 0.6- 0.8 and harvested in falcon tubes; (v) the cells were centrifuged for 10 min at 7000 rpm in a cooled centrifuge (4 °C) (vi) supernatant was discarded and cells were washed with 10 ml per tube of cool electroporation buffer (EPB: 272 mM sucrose, 5 mM NaH<sub>2</sub>PO<sub>4</sub>) and centrifuged again; (vii) supernatant was discarded and cells taken up in 2.1 ml EPB. Competent cells were used immediately or stored at -80 °C by 15 % glycerol addition.

### 2.1.2.6 Transformation of Electro-Competent *C. acetobutylicum*

*C. acetobutylicum* electroporation was performed according to [Mermelstein et al. \(1992\)](#). In short: (i) 0.7 ml of competent cells were added to a 4 mm cuvette with 0.1- 10 µg pAN2 methylated plasmid DNA (pAN1 ([Mermelstein and Papoutsakis, 1993](#)) with Tet resistance instead of Cm<sup>R</sup> ([Heap et al., 2007](#))); (ii) the cuvette was incubated for 5 min on ice inside the cabinet; (iii) the suspension was electroporated with settings 2 kV, infinite resistance, 25 µF capacitance; (iv) the electroporated culture was immediately recovered in 10 ml prewarmed 2xYT plus glucose and incubated for 4 h inside the cabinet; (v) dilutions of the transformed cells were plated on antibiotic containing CGM plates and incubated until colonies appeared.

### 2.1.2.7 Preparation of Electro-Competent *C. sporogenes*

*C. sporogenes* electro-competent cells were prepared in the same way as *C. acetobutylicum* competent cells but TYG was used as medium.

### 2.1.2.8 Transformation of Electro-Competent *C. sporogenes*

*C. sporogenes* was electroporated by the same procedure used in *C. acetobutylicum*, but TYG and TYG agar was used for growth.

### 2.1.2.9 Preparation of Electro-Competent *C. pasteurianum*

*C. pasteurianum* was transformed with an optimised protocol based on the protocol published by Pyne et al. (2013). In short: (i) diluted over-night cultures were grown; (ii) 100 ml main-culture was inoculated to  $OD_{600} = 0.05$ ; (iii) culture was grown to  $OD_{600} = 0.3-0.4$  when sucrose was added to 0.4 M end concentration and glycine was added to give an end concentration of 1.25 w/v%; (iv) culture was further grown to  $OD_{600} = 0.6-0.8$ ; (v) culture was washed once in ice-cold sucrose-magnesium-phosphate (SMP) buffer (270 mM sucrose, 1 mM  $MgCl_2$ , 5 mM  $NaH_2PO_4$ , pH 6.5) after which the pellet was taken up in 3000  $\mu l$  SMP buffer, enough for five transformations. For daily routine a short protocol was used which neglected the addition of sucrose and glycine in step (iii) and instead cells were harvested at  $OD_{600} = 0.4-0.6$  and further treated according to (v).

### 2.1.2.10 Transformation of Electro-Competent *C. pasteurianum*

For the electro-transformation 580  $\mu l$  of competent cells, 30  $\mu l$  ethanol (100 %) and 5  $\mu g$  methylated plasmid DNA of plasmid of interest were added to a 4 mm cuvette. Inside the anaerobic cabinet the culture was electroporated with a square wave protocol at 2500 V and no parallel resistance. Transformations were recovered in 10 ml 2x YPG with 0.2 M sucrose overnight before plating dilutions on plates with antibiotics. The original Pyne et al. (2013) protocol suggests an exponential decay pulse of 1800 V without parallel resistance.

## 2.2 Molecular Biology Methods

### 2.2.1 Plasmid Extraction and Purification

Plasmid extraction from *E. coli* for small concentrations was performed with several different kits: (i) QIAprep<sup>®</sup> Miniprep Kit (Qiagen, UK) with either centrifuge or vacuum method according to supplier's instructions using 50  $\mu l$  water for elution. (ii) GenElute<sup>™</sup> Plasmid Miniprep Kit (Sigma Aldrich, UK) with the centrifuge or vacuum method according to supplier's instructions using 50  $\mu l$  water for elution. (iii) PureYield<sup>™</sup> Plasmid Miniprep System (Promega, UK) with the alternative protocol for larger culture volumes and the centrifuge or vacuum method according to supplier's instructions and elution with 30  $\mu l$  water. Plasmids of Gram-positive organisms were extracted for plasmid rescue with the DNeasy<sup>®</sup> Blood & Tissue Kit (Qiagen, UK) as described in Section 2.2.2.1 and 5  $\mu l$  of the extract was transformed into competent *E. coli* ac-

ording to Section 2.1.1.2 for amplification and further work.

Higher concentrated plasmid DNA was extracted using ZymoPURE™ Plasmid Midiprep Kit (Zymo Research, USA) or CompactPrep Plasmid Midi Kit (Qiagen, UK) according to supplier's instructions.

## 2.2.2 Genomic DNA Extraction and Purification

### 2.2.2.1 Routine Extractions

Genomic DNA of Gram-positive organisms was routinely extracted using the DNeasy® Blood & Tissue Kit (Qiagen, UK) with pretreatment of Gram-positive bacteria and extraction method with centrifuge or vacuum method as per supplier's instructions and elution with 50  $\mu$ l water.

### 2.2.2.2 Phenol-Chloroform Extraction

High quality, high quantity Genomic DNA was extracted with Manual Phase Lock Gel™ Heavy (5Prime, UK). Thus, a pellet of bacteria from a 10 ml culture was lysed according to the instructions of 'pretreatment of Gram-positive bacteria' of the DNeasy® Blood & Tissue Kit (Qiagen, UK) (compare Section 2.2.2.1). The lysate was then used in a phenol-chloroform extraction according to the instructions of Manual Phase Lock Gel™ Heavy but procedure was repeated three times as opposed to only once in the protocol. DNA was then precipitated according to Section 2.2.2.3.

### 2.2.2.3 Ethanol Precipitation

(i) 2 volumes of ice-cold ethanol were added to 1 volume of aqueous phase with DNA (ii) mixture was frozen at -80 °C for 1 hour (iii) mixture was centrifuged at 20000 rpm for 20 min in a cooled centrifuge (iv) the liquid was decanted (v) the pellet was washed with 70 % ethanol and step iii-iv repeated (vi) the rest of the ethanol was evaporated in a vacuum desiccator (vii) the DNA was taken up in 100  $\mu$ l water.

## 2.2.3 Analysis of Extracted DNA

Genomic and plasmid DNA purity and concentration was routinely analysed using a NanoDrop Lite (Thermo Scientific™, UK) measuring nucleic acid concentration at 260 nm and purity using the 260/280 nm ratio. For analysis of genomic DNA extracted for next generation sequencing DNA was additionally analysed using Qubit™ Fluorometric Quantitation (ThermoFisher Scien-

tific, UK) according to supplier's instructions and DNA was further run on 0.8 % agarose gel to detect possible RNA contaminations.

## 2.2.4 Polymerase Chain Reaction

### 2.2.4.1 Polymerases

DreamTaq Green PCR Master Mix (Thermo Scientific™, UK) was used for routine and analytical PCR with 20  $\mu$ l reaction volume according to supplier's instructions. Phusion® High-Fidelity DNA Polymerase (NEB, UK) was used for amplification of fragments subsequently used in cloning procedures according to supplier's instructions. KOD Hot Start Master Mix (Merck-Millipore, UK) was used for transposon insertion inverse PCR in 20  $\mu$ l reactions according to supplier's instructions.

### 2.2.4.2 Splicing by Overlap Extension PCR

Splicing by overlap extension (SOE) PCR ([Higuchi et al., 1988](#)) is the method of choice to amplify homology arms left and right of a target gene for deletion and splicing them together. Thus, four primers have to be prepared as follows: left primer (LHA\_F), the primer upstream the left homology arm (LHA) annealing to the minus strand of the target; right primer (RHA\_R), the primer downstream of the right homology arm (RHA) annealing to the plus strand of the target; two complementary SOE primers (SOE\_F, annealing to the minus strand and SOE\_R, annealing to the plus strand), these primers need to have about 20 bp upstream of the target gene including the start-codon and 20 bp downstream of the target gene including the stop-codon and are complementary to each other. In a first round of PCR (using a high fidelity polymerase [Phusion]) LHA\_F and SOE\_R amplify the LHA and in a separate reaction SOE\_F and RHA\_R amplify the RHA. The fragments will then be isolated either from gel or by PCR clean up and 0.2  $\mu$ l of each clean up is used together in one reaction to splice the two fragments by only adding primers LHA\_F and RHA\_R. The complementary ends attached by SOE\_F and SOE\_R will anneal to each other and allow the splicing.

The protocol can be shortened into one step by adding all primers (LHA\_F, SOE\_F, SOE\_R, RHA\_R) together with the DNA template in the ratio 4:1:1:4. The resulting fragments can then be separated on gel and the spliced fragment will be recovered.

**Table 2.2** – Oligonucleotides used in this study. Underlined sequences indicate added restrictions sites with 5' nonsense sequences 5'-CCCGGG. <sup>1</sup> Messing and Vieira (1982), <sup>2</sup> designed by Edward Farries, <sup>3</sup> compare (Schwarz et al., 2017), <sup>4</sup> compare Zhang et al. (2015)

Oligonucleotide	Sequence 5'-3'
SHA_F	ATCGTTGGACCTGCTATGGG
LHA_R	GCACTTTCACCTCCTCTTCCA
M13R <sup>1</sup>	CAGGAAACAGCTATGACC
M13F <sup>1</sup>	TGTAAAACGACGGCCAGT
MCS_Rev	CAGGGTGCTATCTTCGTCAT
traj_F	GCTTGGCAAGGTCATGATG
MCS_85121_Rev	CGCTTATTCGCTTCGCTCAT
Trsf_F	TTGATGGAGGAAGTGCAAAGA
Trsf_R	AGAAGCACTTGAAATTCTCCTGA
Kin_F	CTTTCTATCAGCCCAAGGAACC
Kin_R	GTTTAGCTGCACTTTCATTTACG
85XXX-RF <sup>2</sup>	TACTAATGAGAGGCGACGAC
8X1XX-LR <sup>2</sup>	GCTGTAGGTACAAGGTACAC
KS001_SHA12_Fw <sup>3</sup>	CCCGGGCCTGCAGGTAAAAGAGTCAGAGGCTCTTTTAGAAGGACA
KS002_LHA_Rev <sup>3</sup>	CCCGGGGCGCGCCAAGCTACTTTCTGCTACAGAGTAAGTTTTAT
Rex_genome_Fw	GGCTGTCCAAATTTTGCTTAGTAGAAGCTATT
Rex_genome_Rev	CCCTTATAAGCTATAAGATACTGTTTCGAGCCATTC
catP_Fw	GGGCCGGCCAGTGGGCAAGTTGAAAAATTCAC
catP_Rev	GTTTAAACTTAGGGTAACAAAAACACCGTATTTCTAC
Rex_CT_Fw <sup>3</sup>	GTTAATATAGTATATCTTGTTATTGTAGGATACGTTAACATTG
Rex_CT_Rev <sup>3</sup>	CTTAAAACTATCATAATTTAATCACCAGCCTAATATTAAGTG
Rex_compl_Fw <sup>3</sup>	CCCGGGGCGGCCGCAATTTAATCACCAGCCTAATATTAAGTG
Rex_compl_Rev <sup>3</sup>	CCCGGGGCTAGCTTATATTGTCTTATTATTAGGCAAGTC
catP_INV_F1 <sup>4</sup>	TAAATCATTTTTAGCAGATTATGAAAGTGATACGCAACGGTATGG
catP_INV_R1 <sup>4</sup>	TATTGTATAGCTTGGTATCATCTCATATATCCCCAATTCACC
catP_INV_R2 <sup>4</sup>	TATTTGTGTGATATCCACTTTAACGGTCATGCTGTAGGTACAAGG
Segr_LHA_SbfI_F	GAATTCcctgcaggTGCTTTGCTCTCTTATTCTCTCA
Segr_RHA_NheI_R	AATTCgctagcAAAGGAATGTAGCAGTGCGA
Segr_outside_F	TCCAATTCCAACAACCTCCATCA
Segr_outside_R	GGAAGGCCTAAGTGCAATACA



### 2.2.4.3 Inverse PCR

Genomic DNA was digested with HindIII at a concentration of 200 ng/ $\mu$ l after which the enzyme was inactivated by heating it to 65 °C for 30 minutes. The DNA was then self-ligated with T4 DNA ligase at a concentration of 5 ng/ $\mu$ l. Ligation was allowed to proceed for one hour after which the ligase was inactivated by heating the mixture to 65 °C for 30 min. 100 ng of the ligated DNA was used in a PCR reaction with KOD Hot Start with primers catP\_INV\_F1 and catP\_INV\_R1. The product was run on agarose gel and fragments purified and sent for Sanger sequencing with primer catP\_INV\_R2.

### 2.2.4.4 Colony PCR

One colony was picked from plate and resuspended in a 20  $\mu$ l mix of DreamTaq Green PCR Master Mix with appropriate primers. The initial denaturing step was prolonged to 10 min.

### 2.2.4.5 Touch-Down PCR

When unspecific bands were observed with a standard PCR a touch-down PCR was performed. After the initial denaturing step of a PCR an eight step touch-down protocol was inserted. The annealing temperature was decreased in every step to reach the calculated annealing temperature. After reaching the annealing temperature a standard PCR was performed with 25 steps.

## 2.2.5 Cloning Techniques

### 2.2.5.1 Restriction Digestion of DNA

FastDigest<sup>™</sup> (Thermo Scientific<sup>™</sup>, UK) restriction enzymes were used throughout this work according to supplier's instructions but incubation time was increased to 1 hour.

### 2.2.5.2 DNA Blunting Using Klenow Fragment

DNA Polymerase I, Large (Klenow) Fragment (NEB, UK) was used to produce blunt ended DNA after restriction digest with restriction enzymes according to supplier's instructions.

### 2.2.5.3 DNA Dephosphorylation Using Antarctic Phosphatase

DNA fragments were dephosphorylated using Antarctic Phosphatase (NEB, UK) according to supplier's instructions.

#### 2.2.5.4 T4 Ligation

Ligations were carried out with Quick Ligation™ Kit (NEB, UK) according to supplier's instructions and incubated for at least 15 minutes at room temperature.

#### 2.2.5.5 Agarose Gel Electrophoresis

A Biometra Compact (analytikjena, DE) system with a PowerPac™ HC High-Current Power Supply (BioRad, UK) was used for gel electrophoresis. Agarose gel was routinely cast at 0.8 % agarose with SYBR® Safe DNA Stain (ThermoFisher Scientific, UK) at half the recommended concentration. Electrophoresis was run for 30- 60 minutes at 100 V. Gel was visualised in a Gel Doc™ XR+ System (BioRad, UK) and if needed fragments were visualised and cut from gel on a Syngene UltraBright-LED blue light transilluminator (Syngene, UK).

#### 2.2.5.6 Purification of DNA Fragments from Agarose Gel

After DNA fragments were cut from agarose gel they were extracted using a Zymoclean™ Gel DNA Recovery Kit (Zymo Research, USA) according to supplier's instructions.

#### 2.2.5.7 Purification of DNA Fragments from PCR Products

PCR amplified fragments were purified using QIAquick PCR Purification Kit (Qiagen, UK) according to supplier's instructions.

## 2.3 Plasmid Construction

Plasmids used in this study are listed in Table 2.3.

### 2.3.1 pMTL-KS12-tcdR

To construct the ACE plasmid for *tcdR* insertion at downstream the *pyrE* locus in *C. pasteurianum* plasmid pMTL-ME6C-tcdR (Zhang et al., 2015) and plasmid pMTL-KS12 (Schwarz et al., 2017) were digested with SbfI and NheI and fragments were separated on agarose gel. The 4.4 kb fragment of pMTL-KS12 and the 1450 bp fragment of pMTL-ME6C-tcdR were purified from gel and ligated together prior to transformation into *E. coli* XL1-Blue. Several clones were sequenced and appropriate stored as CRG5989.

**Table 2.3** – Plasmids used in this study. If there is no mention of a strain they are to be used in *C. pasteurianum*. Table continued on next page.

Plasmid	<i>E. coli</i> strain designation	Properties	Source
pMTL-CR1	CRG3131	plasmid encoding M.BepI methyltransferase of <i>Brevibacterium epidermidis</i> . Kan <sup>R</sup>	Rotta (2014)
pMTL-AGH12	CRG4978	ACE <i>pyrE</i> repair plasmid. Cm <sup>R</sup>	this study
pMTL-KS12-rexcomp	CRG4979	complement <i>rex</i> with native promoter downstream of <i>pyrE</i> . Cm <sup>R</sup>	this study
pMTL-ME6C-tcdR	CRG3262	ACE plasmid for for <i>tcdR</i> integration at <i>pyrE</i> locus in <i>C. acetobutylicum</i> 824. Cm <sup>R</sup>	Zhang et al. (2015)
pMTL-KS12	CRG4153	ACE plasmid with <i>lacZ</i> cargo. Cm <sup>R</sup>	Schwarz et al. (2017)
pMTL-KS15	CRG4156	Allelic exchange plasmid with <i>C. acetobutylicum pyrE</i> and <i>lacZ</i> as cargo. Cm <sup>R</sup>	Schwarz et al. (2017)
pMTL-KS10	CRG4151	<i>pyrE</i> truncation plasmid. Cm <sup>R</sup>	Schwarz et al. (2017)
pMTL-KS12-tcdR	CRG5989	ACE plasmid for for <i>tcdR</i> integration at <i>pyrE</i> locus. Cm <sup>R</sup>	this study
pMTL-SC7515	CRG1351	pMTL85151 with <i>codA</i> on backbone	Stephen T. Cartman
pMTL-AGH15	CRG4848	Plasmid based on pMTL-KS15 with <i>codA</i> on the backbone for allelic exchange in wild type (non- <i>pyrE</i> truncation) background Cm <sup>R</sup>	this study
pMTL-AGH15-c30550	CRG4851	Allelic exchange plasmid to repair the SNP in Cpas_c30550 in <i>C. pasteurianum</i> -H4. Cm <sup>R</sup>	this study
pMTL007cc-E2	CRG4673	Experimental ClosTron plasmid with <i>codA-catP</i> fragment on backbone for counterselection with 5-fluorocytosine (5-FC)	Hengzheng Wang
pMTL83251	CRG1234	<i>Clostridium</i> modular plasmid used for construction of IPTG inducible promoter system	Heap et al. (2007)
pMTL83251-lacI	CRG3975	<i>Clostridium</i> modular plasmid with conditional replicon	Zhang et al. (2015)
pMTL83251-lacI-T	CRG3976	<i>Clostridium</i> modular plasmid with conditional replicon and a <i>fdx</i> terminator	Zhang et al. (2015)
pMTL82254	CRG1971	Plasmid for CAT assay without promoter in front of <i>catP</i> , Em <sup>R</sup>	Zhang et al. (2015)
pMTL82254-P <sub>fdx</sub>	CRG2108	Plasmid for CAT assay with strong P <sub>fdx</sub> promoter in front of <i>catP</i> , Em <sup>R</sup>	Zhang et al. (2015)
pMTL82254-P <sub>tcdB</sub>	CRG2964	Plasmid for CAT assay with P <sub>tcdB</sub> promoter in front of <i>catP</i> , Em <sup>R</sup>	Zhang et al. (2015)

**Continued Table 2.3** – Plasmids used in this study. If there is no mention of a strain they are to be used in *C. pasteurianum*. Table continues on next page.

Plasmid	<i>E. coli</i> strain designation	Properties	Source
pMTL-YZ14	CRG3261	Pseudo-suicide transposon mutagenesis plasmid with <i>mariner</i> transposon carrying <i>catP</i> , $P_{tcdB}$ promoted <i>Himar1 C9</i> transposase and inducible <i>lacI</i> promoter in front of <i>repH</i> , Em <sup>R</sup>	Zhang et al. (2015)
pMTL-GL15	CRG4205	Suicide transposon mutagenesis plasmid similar to pMTL-YZ14, but excised Gram-positive conditional replicon, Em <sup>R</sup>	Gareth Little

### 2.3.2 pMTL-AGH12

To repair the truncated *pyrE* locus the wild type allele was amplified with primers KS001\_SHA12\_Fw and KS002\_LHA\_Rev which add a 5'-SbfI and a 3'-AscI site for cloning. The PCR fragment and vector pMTL-KS12 were digested with SbfI and AscI and fragments were run on gel. The 3406 bp fragment of pMTL-KS12 and the 1752 bp fragment of the PCR product were purified from gel and ligated together before transforming into *E. coli* XL1-Blue. Several clones were sequenced and appropriate stored as CRG4978.

### 2.3.3 pMTL-AGH15

The allelic exchange plasmid based on *codA* was constructed digesting a PCR fragment incorporating *catP* and *codA* (Ying Zhang, personal communication) and pMTL-KS15 with FseI and PmeI and run on agarose gel. The 3253 bp fragment of pMTL-KS15 and the 2260 bp PCR fragment were purified from gel and ligated together before transformed into *E. coli* XL1-Blue. Several clones were sequenced and appropriate stored as CRG4848.

### 2.3.4 pMTL-AGH15-c30550

To change the SNP in *C. pasteurianum*-H3 gene Cpas\_c30550 back to wild type, the allelic exchange plasmid pMTL-AGH15-Cpas\_c30550 was constructed by amplifying Cpas\_c30550 with primers Segr\_LHA\_SbfI\_F and Segr\_RHA\_NheI\_R to give a fragment of 1124 bp and add a 5'-SbfI and a 3'-NheI site. Plasmid pMTL-AGH12 and the PCR fragment were digested with SbfI and NheI and subsequently run on agarose gel. The 4930 bp pMTL-AGH15 fragment and the PCR product were purified from gel, ligated together and transformed into *E. coli*

XL1-Blue. Several clones were sequenced and appropriate stored as CRG4851.

### 2.3.5 pMTL-KS12-rexcomp

To complement *rex* downstream of *pyrE* in the *rex* deletion mutant plasmid pMTL-KS12-rexcomp was constructed by amplifying a 816 bp fragment from *C. pasteurianum* wild type and attaching NotI upstream and NheI downstream with primers Rex\_compl\_Fw and Rex\_compl\_Rev. Plasmid pMTL-KS12 and the fragment were then digested with NotI and NheI and subsequently run on agarose gel. The 4956 bp plasmid fragment and the PCR product were purified from gel, ligated together and transformed into *E. coli* XL1-Blue. Clones were Sanger sequenced and appropriate clone was stored as CRG4979.

## 2.4 Mutant Construction in *C. pasteurianum*

### 2.4.1 Allele Coupled Exchange

In general allelic coupled exchange to knock-in genes of interest is done according to [Ng et al. \(2013\)](#) and [Ehsaan et al. \(2016\)](#). In short:

1. Transform plasmid
2. Restreak 24 faster growing colonies on RCM+Tm
3. Restreak on CBM plates and RCM+Tm (if growing, *pyrE* is repaired)
4. Single colonies were re-streaked onto CBM and RCM+Tm agar plates until no growth on RCM+Tm agar plates was visible (plasmid loss) and then 3 times more on CBM
5. Grow over-night culture in 2xYTG use 1 ml for -80 °C and 5-9 ml for gDNA extraction
6. Screen with SHA\_F and LHA\_R for insertion and plasmid primers for absence of plasmid

#### 2.4.1.1 Truncation of *pyrE*

*pyrE*-truncation mutants were made by transforming *C. pasteurianum*-H3 and *C. pasteurianum*-H4 with plasmid pMTL-KS10. Faster growing transformants were restreaked twice on RCM+Tm before they were restreaked on CBM with

600  $\mu\text{g/ml}$  5-fluoroorotic acid (FOA) and 5  $\mu\text{g/ml}$  uracil ( $\text{CBM}_{\text{FOA/U}}$ ). Single colonies were replica plated onto CBM, CBM+Tm+U and CBM+U. Colonies showing growth only on CBM+U were restreaked three times on CBM+U plates before over-nights were made for storage at  $-80^\circ\text{C}$  and for gDNA extraction and verification of the genotype. gDNA was analysed via PCR with primers SHA\_F and LHA\_R for *pyrE* truncation and with primers 85XXX\_RF and 85XXX\_LR for absence of the plasmid. Positive clones were stored as CRG4817-4822.

#### 2.4.1.2 *pyrE* Correction in Gene-Deletion Mutants

The *pyrE* gene had to be repaired in the deletion mutants *C. pasteurianum* -  $\Delta\text{spo0A}\Delta\text{pyrE}$  4-44, *C. pasteurianum*  $\Delta\text{rex}\Delta\text{pyrE}$  6-49, *C. pasteurianum*  $\Delta\text{hyd}\Delta\text{pyrE}$  47 and *C. pasteurianum*  $\Delta\text{dhaBCE}\Delta\text{pyrE}$  6. Thus in all these strains pMTL-AGH12 was transformed and *pyrE* repair mutants were selected as described above (Section 2.4.1). Repaired mutants were screened with primers SHA\_F and LHA\_R for the insertion and checked with primers 85XXX\_RF and 85XXX\_LR for absence of the plasmid. Correct mutants have a fragment size of insertion of 684 bp and were sent for Sanger sequencing of the insert. Correct mutants were stored with CRG strain designations indicated in Table 2.1.

#### 2.4.1.3 Rex Mutant Complementation

For a complementation study, the gene *rex* with its own promoter was inserted down stream of *pyrE* in *C. pasteurianum*  $\Delta\text{rex}$ . Thus, plasmid pMTL-KS12-*rexcomp* was transformed into *C. pasteurianum*  $\Delta\text{rex}\Delta\text{pyrE}$  6-49 and *rex* insertion mutants were selected as described above (see Section 2.4.1). Insertional mutants were screened for with primers SHA\_F and LHA\_R for the insertion and checked with primers 85XXX\_RF and 85XXX\_LR for absence of the plasmid. Correct mutants have an insertion size of 1564 bp and were sent for Sanger sequencing of the insert. Two correct mutants were stored as CRG5522 and CRG 5523.

#### 2.4.1.4 Genome Insertion of *tcdR*

The gene *tcdR* was inserted downstream *pyrE* via allele coupled exchange in the *pyrE* truncation mutants of *C. pasteurianum*-H4. Plasmid pMTL-KS12-*tcdR* was transformed and insertion mutants were selected as described above (Section 2.4.1). Potential double crossover insertion mutants were screened with primers SHA\_F and LHA\_R for the insertion and checked with primers 85XXX\_RF and 85XXX\_LR for absence of the plasmid. Clones with an inserted

fragment size of 1391 bp and absence of the plasmid were sent for Sanger sequencing of the insert. One clone with a correct insert was stored at -80 °C as CRG5230.

### 2.4.2 Allelic Exchange Employing *pyrE*

Generally, allelic exchange employing *pyrE* was done as described by [Heap et al. \(2012\)](#). In short:

1. Transform plasmid
2. Restreak faster growing colonies twice on RCM+Tm
3. Optional: screen for single crossover with PCR
4. Restreak twice on CBM+FOA 600 µg/ml and Uracil 40 µg/ml
5. Restreak on RCM +Tm, CBM and RCM (should grow only on RCM)
6. Over-night of successful colonies in 2xYTG, store 1 ml at -80 °C and use 5-9 ml for gDNA extraction
7. Screen with flanking primers (flanking KO site and outside of homology arms and primers for plasmid)

This technique was used by [Schwarz et al. \(2017\)](#) to delete the genes in *spo0A*, *rex*, *hyd* and *dhaBCE*.

### 2.4.3 Allelic Exchange Employing *codA*

### 2.4.4 Minimal Inhibitory Concentration Testing

To test if allelic exchange employing *codA* is feasible sensitivity to 5-fluorouracil (5-FU) and 5-fluorocytosine (5-FC) was tested. Thus, over-night cultures of *C. pasteurianum* DSM 525, *C. pasteurianum*-H1, *C. pasteurianum*-H3 and *C. pasteurianum*-H4 were grown and diluted in PBS to 10<sup>-6</sup>. 5 µl of dilutions 10<sup>-4</sup>, 10<sup>-5</sup> and 10<sup>-6</sup> were dropped onto plates with increasing concentrations of 5-FC and 5-FU.

Furthermore, the minimal inhibitory concentration of 5-FC was tested when a plasmid with *codA* on the backbone was present in the host (pMTL-SC7515). Thus, *C. pasteurianum*-H1, *C. pasteurianum*-H3 and *C. pasteurianum*-H4 were transformed with pMTL-7515 and as a control *C. pasteurianum*-H3 was transformed with pMTL85151 not carrying the *codA* gene. The resulting colonies

were restreaked on RCM+Tm to confirm plasmid conferred antibiotic resistance and were further restreaked on plates with thiamphenicol and increasing concentrations of 5-FC. Controls were made by restreaking the same colonies on plates with 5-FU and plates with thiamphenicol only.

### 2.4.5 Procedure of Allelic Exchange with *codA*

In general allelic exchange with *codA* is done as described in [Cartman et al. \(2012\)](#) and similar to allelic exchange employing *pyrE*. In short:

1. Transform plasmid
2. Restreak faster growing colonies twice on RCM+Tm
3. Optional: screen for single crossover with PCR
4. Restreak twice on CBM+FC 500  $\mu\text{g/ml}$
5. Restreak on RCM +Tm and CBM (should grow only on CBM)
6. Over-night of successful colonies in 2xYTG, store 1 ml at  $-80\text{ }^{\circ}\text{C}$  and use 5-9 ml for gDNA extraction
7. Screen with flanking primers (flanking KO site and outside of homology arms and primers for plasmid)

#### 2.4.5.1 SNP Correction of CLPA\_c30550

The SNP in CLPA\_c30550 in *C. pasteurianum*-H3 was corrected using allelic exchange employing *codA*. Plasmid pMTL-AGH12-CLPA\_c30550 was transformed into *C. pasteurianum*-H3 and double crossover mutants were selected as described above. The corrections were screened with primers Segr\_outside\_F and Segr\_outside\_R and absence of the plasmid was screened with 85XXX\_RF and 85XXX\_LR. All screened mutants were sent for Sanger sequencing for confirmation of the correct genotype. Double crossover mutants with corrected allele were stored as CRG5538 - CRG5540 and one mutant showing a double crossover but being a wild type (*C. pasteurianum*-H3) revertant was stored as CRG5541 (compare Table 2.1).

### 2.4.6 Transposon Mutagenesis

#### 2.4.6.1 Transposon Mutant Library Generation

pMTL-YZ14 was transformed into the organism and transformants were selected on erythromycin (Em) supplemented rich medium (RCM for *C. pasteurianum*-



*anum*, CGM for *C. acetobutylicum* and TYG for *C. sporogenes*). After 48- 72 h all growth was harvested by scraping the plates and collecting all growth in PBS. Serial dilutions to single colony were plated on rich medium (to calculate transposition frequency), rich medium with thiamphenicol (**Tm**) and  $\beta$ -D-1-thiogalactopyranoside (**IPTG**) (1 mM) to select for transposition and induce plasmid loss simultaneously. After 16-24 h colonies appearing on **Tm** plus **IPTG** plates were picked and replica plated on plates supplemented with **Tm** and plates with **Em** to check for transposon insertion ( $Tm^R$ ) and plasmid loss ( $Em^S$ ). Colonies only growing on plates with **Tm** are putative transposon insertional mutants without plasmid.

**pMTL-GL15** was transformed into *C. pasteurianum* and transformants directly selected on RCM+Tm. These were patch plated on RCM+Tm and RCM+Em to check for plasmid loss in the same way as above.

#### 2.4.6.2 Transposon Library Analysis

A transposon mutant library was produced and 60 clones were chosen at random. Genomic DNA of each clone was extracted and transposon insertion assessed by inverse PCR as described in Section 2.2.4.3 and subsequent Sanger sequencing. Location of insertion was then assessed by searching the transposon adjacent sequence in the genome of the appropriate organism. Genomes are freely available online: *C. pasteurianum* under GenBank accession number CP009268 (Poehlein et al., 2015), *C. acetobutylicum* under GenBank accession number AE001437 (chromosome) and AE001438 (megaplasmid) (Nolling et al., 2001) and for *C. sporogenes* under GenBank accession number CP009225 (Kubiak et al., 2015).

#### 2.4.6.3 Auxotroph Mutant Screen

Auxotroph mutants were screened for by patch plating transposon mutant libraries onto minimal medium and rich medium plates. Colonies that only grew on rich medium were considered auxotroph mutants. These were checked again and analysed by inverse PCR and Sanger sequencing as explained above (Section 2.4.6.2). For *C. acetobutylicum* rich medium used was CGM and minimal medium used was P2 Minimal Medium and for *C. sporogenes* TYG rich- and MACC minimal medium was used. All medium was supplemented with **Tm** to select for transposon mutants.

#### 2.4.6.4 Sporulation/ Germination Deficient Mutant Screen

For sporulation/ germination deficient mutants clones from a transposon mutant library were inoculated in liquid rich medium (2xYTG for *C. acetobutylicum* and TYG for *C. sporogenes*) in 2ml 96 well plates with Tm and grown for 5-7 days in the cabinet to allow spores to form. This was done in duplicate. One replicate was then heat shocked for 10 mins at 80 °C and the other replicate left in the anaerobic cabinet. Of both replicates 5  $\mu$ l were plated onto rich medium. Also, 5  $\mu$ l of non-heat-shocked spore solution was plated as a control for growth. Clones showing growth when not heat-shocked but no growth when heat shocked were considered sporulation or germination deficient. These were checked again and analysed by inverse PCR and Sanger sequencing as explained above (Section 2.4.6.2).

#### 2.4.7 Calculations of Growth and Fermentation Products

Growth characteristics were calculated by calculating the best approximation of slope and intercept of the exponential growth phase in a logarithmic scale. From this two theoretical points N and  $N_0$  with a time interval t were calculated and thus the specific growth rate  $\mu$  can be calculated as follows:

$$\mu = \frac{\ln(N/N_0)}{t} \quad (2.1)$$

The doubling time ( $t_d$ ) in hours is then calculated as:

$$t_d = \frac{\ln(2)}{\mu} \quad (2.2)$$

Carbon recovery was calculated by assuming 3.5 g/l dry-weight per 10 OD values (Sarchami et al., 2016), carbon dioxide desorption as described by Percheron et al. (1995) and the assumption that 46.2 % of dry-weight is carbon (Papoutsakis and Meyer, 1985). The carbon fraction of yeast extract was neglected as it was with under 5 % w/w a very small fraction of the available carbon (calculated with 43 % yeast carbon fraction (Lange and Heijnen, 2001)).

#### 2.4.8 High Performance Liquid Chromatography

2 ml samples of fermentation cultures were collected, centrifuged (10 min, 16,000 g, 4 °C) and cell-free supernatant stored at -80 °C until HPLC analysis was performed. For the HPLC analysis, cell-free supernatant was thawed on ice, mixed with an equal volume (200  $\mu$ l) of internal standard solution (80 mM

valeric acid [Sigma-Aldrich, UK] in 0.005 M H<sub>2</sub>SO<sub>4</sub>), filtered through a 0.22  $\mu$ m HPLC certified syringe filter (Whatman<sup>®</sup> Spartan<sup>®</sup> 13/0.2 RC [GE Healthcare Life Sciences, UK]) and transferred into a HPLC vial with a 100  $\mu$ l insert. Subsequently, substrate (glucose, glycerol) and fermentation products (acetate, acetone, butanol, butyrate, lactate, ethanol, 1,3-propanediol (1,3-PDO), 1,2-propanediol (1,2-PDO)) were analysed by the use of a Dionex UltiMate 3000 HPLC system (Thermo Fisher Scientific, UK) equipped with a Bio-Rad Aminex HPX-87H (BioRad, UK) column, a refractive index (RI) and diode array detector (DAD) at UV 210 nm at an isocratic flow rate of 0.5 ml/min of 0.005 M H<sub>2</sub>SO<sub>4</sub> as mobile phase and a column temperature of 35 °C for 55 min. The injection volume was 20  $\mu$ l. If required samples were diluted using reverse osmosis (RO) water. Standard concentrations ranged from 0.98 mM to 250 mM. For glycerol two additional concentrations of 500 mM and 750 mM were employed. Signal analysis was performed using the Chromeleon 7.2 Chromatography Data System (Thermo Fisher Scientific, UK).

### 2.4.9 Chloramphenicol Acetyltransferase Assay

An adapted protocol similar to [Shaw \(1975\)](#) was used for the chloramphenicol acetyltransferase (CAT) assay. Samples were taken from a growth curve at different time points to account for eventual non-constitutive expression. Samples were taken to with a relative OD<sub>600</sub> of 5. Samples were centrifuged down and pellets were stored at -20 °C until use.

#### 2.4.9.1 Sample Preparation by Cell Lysis

Samples were prepared working on ice as follows: (i) take up pellet in 1 ml 10mM Tris pH 7.8 plus 1:50 Calbiochem protease inhibitor cocktail VII (Merck Millipore, UK) make one blank control in triplicates (water); (ii) use 300  $\mu$ l per sample and sonicate in Bioruptor<sup>®</sup> (Diagenode, UK) at 4 °C with 5 cycles (30 sec on/ 30 sec off) at high intensity, sonicated samples can be stored on ice until further use; (iii) centrifuge at 21000 rpm at 4 °C for 30 min and store 250  $\mu$ l supernatant in a clean Eppendorf tube.

#### 2.4.9.2 Analysis of CAT Assay

For the assay a Tecan Infinite<sup>®</sup> M1000 PRO (TECAN, CH) 96-well plate reader with automatic injection was used. (i) A mastermix was prepared: 500  $\mu$ l chloramphenicol of a 0.3 % stock in H<sub>2</sub>O; 1 ml acetyl-CoA sodium salt of a 5 mM stock in H<sub>2</sub>O; 1 ml 5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent) of a 2.5 mM stock in 100 mM Tris pH 7.8; 20 ml Tris buffer pH 7.8 (100 mM). (ii)

Prewarm the mastermix to room temperature and prepare for injection with plate-reader. (iii) In a 96-well plate prepare 10  $\mu\text{l}$  of the lysate and 40  $\mu\text{l}$  Tris pH 7.8 in each well. (iv) run TECAN program as follows:

```
Temperature:
  On: 20 °C
Wait for temperature:
  Minimum: 24.5 °C
  Maximum: 25.5 °C
Part of plate: (Use 48 wells maximal)
Well:
  Injection:
    Injector A, Volume: 150  $\mu\text{l}$ 
    Speed: 100  $\mu\text{l/s}$ 
    Refill speed: 100  $\mu\text{l/s}$ 
    Refill mode: Standard
  Wait (Timer):
    Wait for injection
  Kinetic cycle:
    Number of cycles: 50
    Use kinetic interval:
      Time: 1 s
      Absorbance (Well-wise):
        Wavelength: Measurement: 412 nm
        Number of Flashes: 25
        Settle time: 20 ms

Move plate: out
```

To analyse the protein concentration in the cell lysate, a Bradford assay with Coomassie (Bradford) Protein Assay (ThermoFisher Scientific, UK) was used according to supplier's instructions.

### 2.4.9.3 Calculations

According to [Shaw \(1975\)](#) the **CAT** activity was then be calculated as follows: In the measured curves the linear response is taken and an absorption difference between two points calculated ( $\Delta A/\text{min}$ ) and the same for the blank control ( $\Delta A_{\text{blank}}/\text{min}$ ) this together with the total volume of the assay ( $V_{\text{total}}$ , 0.2 ml), the dilution factor (df, 20), the micromolar extinction coefficient for DTNB at 412 nm ( $E_c$ , 0.0136 ([Silverstein, 1975](#))) and the volume of enzyme used ( $V_{\text{enzyme}}$ , 0.01 ml) can be used to calculate the units of **CAT** per ml lysate as follows:

$$\text{Units/ml lysate} = \frac{(\Delta A/\text{min} - \Delta A_{\text{blank}}/\text{min}) \times V_{\text{total}} \times \text{df}}{E_c \times V_{\text{enzyme}}} \quad (2.3)$$

To get the units per mass protein this can then be divided by the result of the Bradford assay ( $m_{\text{protein}}$ ) as follows:

$$\text{Units/mg protein} = \frac{\text{Units/ml lysate}}{m_{\text{protein}}/\text{ml}} \quad (2.4)$$

## 2.5 *In Silico* Methods

### 2.5.1 Genome Sequencing and Resequencing

#### 2.5.1.1 Illumina Sequencing of Strains

Illumina sequencing of genomic DNA was carried out by Deep-Seq (Next Generation Sequencing Facility, University of Nottingham, UK) and raw data was used for further processing. *C. pasteurianum* DSM 525 data was sent to the Göttingen Genomics Laboratory (Göttingen, DE) for gap closing and *de novo* assembly as well as genome annotation (Poehlein et al., 2015).

#### 2.5.1.2 Acquisition of Other Sequencing Data

Raw data of Pyne et al. (2014b) was acquired using the SRA tool-kit from the Sequence Read Archive (Leinonen et al., 2010). Three datasets were downloaded and converted into fastq format; the Illumina data (SRX207911), PacBio RS-I sequencing data (SRX543191) and the PacBio RS-II sequencing data (SRX543193). The reads were extracted under MS Windows as follows:

```
.\ fastq-dump -split-files -I SRR# % for Illumina paired-  
end library  
.\ fastq-dump -fasta 0 SRR# % for PacBio libraries, does  
not extract quality scores
```

#### 2.5.1.3 Read Mapping

Paired-end reads were mapped against the published *C. pasteurianum* DSM 525 genome (Poehlein et al., 2015) in CLC Genomics Workbench 8.0.2 (Qiagen, DK) with the following parameters:

```
masking mode = no masking  
mismatch cost = 2  
cost of insertions and deletions = linear gap cost  
insertion cost = 3  
deletion cost = 3  
insertion open cost = 6  
insertion extend cost = 1  
deletion open cost = 6  
deletion extend cost = 1  
length fraction = 0.5  
similarity fraction = 0.8  
global alignment = No  
auto-detect paired distances = yes  
non-specific match handling = map randomly
```

### 2.5.1.4 Small Nucleotide Polymorphism Calling

CLC Genomics Workbench 8.0.2 (Qiagen, DK) was used for Single Nucleotide Polymorphism (SNP) calling on the basis of mapped reads (Section 2.5.1.3). Parameters for the SNP call were:

```
ploidy = 1
ignore positions with coverage above = 100,000
restrict calling to target regions = not set
ignore broken pairs = yes
ignore non-specific matches = reads
minimum coverage = 10
minimum count = 2
minimum frequency (%) = 40.0
base quality filter = yes
neighbourhood radius = 5
minimum central quality = 20
minimum neighbourhood quality = 15
read direction filter = no
relative read direction filter = yes
significance (%) = 1.0
remove pyro-error variants = yes
in homopolymer regions with minimum length = 5
with frequency below = 0.7
```

### 2.5.2 Other Bioinformatic Tools

Plasmids were designed and visualised using GENtle (Manske, 2006). Genomes were visualised with Artemis (Rutherford et al., 2000) and genome comparisons were performed using DoubleAct ([http://www.hpa-bioinfotools.org.uk/pise/-double\\_actv2.html](http://www.hpa-bioinfotools.org.uk/pise/-double_actv2.html), last access 1/3/2017). Genomes were visualised using Artemis ACT (Carver et al., 2005). Primers were designed with Primer3 (Untergasser et al., 2012). BLAST was used for standard nucleotide or protein similarity searches (Altschul et al., 1990) and the Conserved Domain Database (CDD) was used to find conserved domains (Marchler-Bauer et al., 2011) as well as Conserved Domain Architecture Retrieval Tool (CDART) to find homologue genes by domain architecture rather than amino acid similarity (Geer et al., 2002). Proteins were aligned with Clustal  $\omega$  (Sievers et al., 2011). The Phyre2 web portal was used for 3D model prediction of CLPA\_c30550 (Kelley et al., 2015) and UGene for protein visualisation (Okonechnikov et al., 2012).

### 2.5.3 Data Analysis and Visualisation

Data was analysed and visualised using Microsoft Excel 2013 (Microsoft, USA) and Graphpad Prism (GraphPad Software, USA). Text was processed with the open source visual interface  $\text{\TeX}$ nicCenter (<http://www.texniccenter.org>, last access 23/08/2016) for  $\text{\LaTeX}$  (Lamport, 1986a). The reference manager used was

JabRef (<http://www.jabref.org>, last access 23/08/2016) running BibTex (Lamport, 1986b). For picture composition PowerPoint 2013 (Microsoft, USA) was used and for vector graphics Inkscape was employed (<http://inkscape.org>, last access 23/08/2016).





# Chapter 3

## Development of a Hyper-transformable Strain of *Clostridium pasteurianum*

“Hodor.”

---

G.R.R. Martin, *A Game of Thrones*

### 3.1 Introduction

Modern molecular microbiology relies on the ability of microorganisms to be transformed with foreign DNA. For a successful transformation of a microorganism several factors have to be taken into consideration: (i) competence, the ability (natural or induced) to take up DNA; (ii) DNA passing the cell-wall and membrane(s); (iii) overcoming internal (enzymatic) defence mechanisms.

#### 3.1.1 Bacterial Competence

Bacterial competence for transformation was first observed by [Griffith \(1928\)](#) when mixing heat-killed virulent *Streptococcus pneumoniae* (Pneumococcus) with living cultures of a non-virulent strain resulting in adopted virulence of the former non-virulent strain. [Avery et al. \(1944\)](#) later confirmed these findings and were able to show that the deoxyribonucleic acid fraction was responsible for the transformation of the virulence behaviour. Not only was this the first proof of bacterial competence and DNA transformation capability, it also “marked the opening of the contemporary era of genetics, its molecular phase” ([Lederberg, 1994](#)).

Bacterial competence can be separated into natural (inherent) competence of which only certain species are capable (Chen et al., 2005) and induced (or artificial) competence. The latter is imposed in laboratory settings which might be possible by most cultivatable species (Thomas and Nielsen, 2005).

Natural competence is the ability of bacteria to absorb free DNA from the environment. It is a wide spread phenomenon in both Gram-positive (e.g. *Bacillus subtilis* and *Streptococcus pneumoniae*) and Gram-negative bacteria (e.g. *Neisseria gonorrhoeae* and *Haemophilus influenzae*) (Dubnau, 1999). Competence is dependent on environmental conditions such as altered growth characteristics, nutrient access, cell density (by quorum sensing) or starvation. The ratio of competent cells inside a population ranges from very little to almost complete competence. (Thomas and Nielsen, 2005)

During the advances in research into natural competence it was found that most bacterial species are not readily transformed with free DNA. Cosloy and Oishi (1973) showed that *E. coli* K12 can be transformed with R-factor DNA (antibiotic-resistance bestowing plasmid DNA) by CaCl<sub>2</sub> treatment and heat-shock at 42 °C, a technique developed by Mandel and Higa (1970). Based on this finding, protocols for other non-competent bacterial species were developed based around permeabilisation of the membrane by molecular interaction and physical breaching of the membrane by e.g. heat-shock or electroporation. The latter is the technique of choice in most organisms as it yields high efficiency transformations. An examination of electroporation protocols for different species reveals similarities (compare Table A.1). They resemble each other in that the bacterial culture is harvested in a specific growth phase (mostly exponential) and the bacterial membrane is weakened by different cell wall interacting agents. Salts of the growth medium are then washed off and the resulting concentrated cell suspension is electroporated with species specific settings (Section 3.1.2).

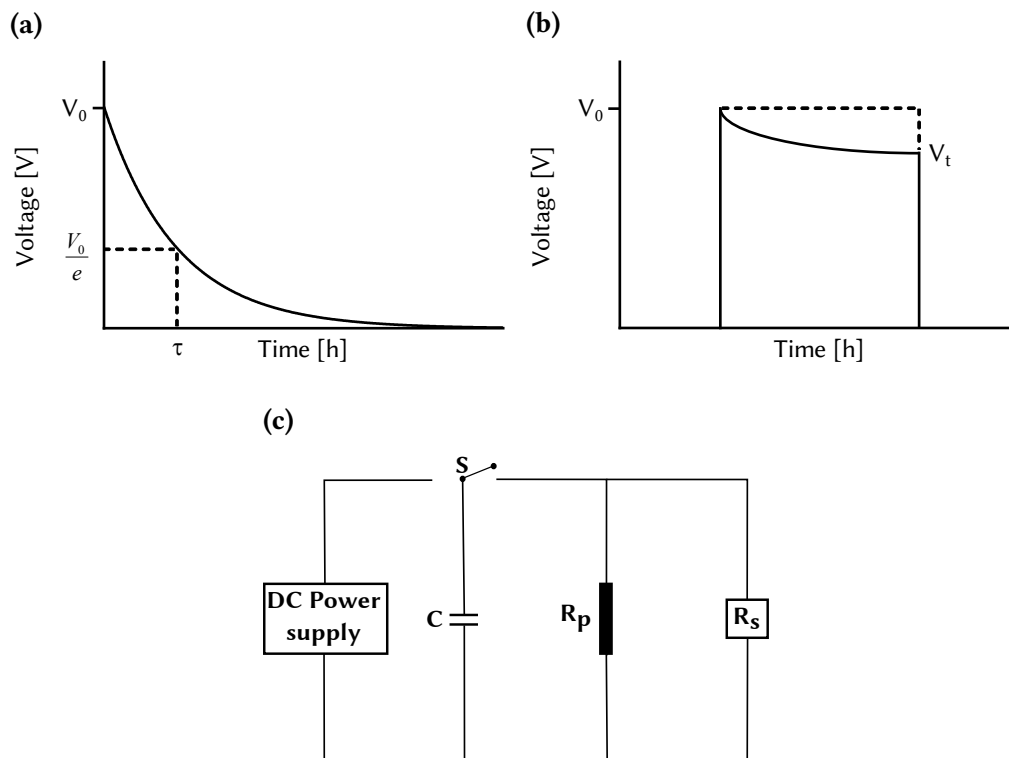
### 3.1.1.1 *C. pasteurianum* Restriction Systems

Restriction systems in bacteria in general were introduced in Section 1.3.2.1. In *C. pasteurianum* DSM 525 two type II-restriction systems have been identified in the genome (Poehlein et al., 2015). The first is CpaAI (Richards et al., 1988), an isoschizomer of *ThaI* and *FnuDII* with the recognition sequence 5'-CGCG. Secondly, CpaPI (Venetianer 1992, unpublished observation according to REBASE), an isoschizomer of *MboI* cutting the recognition sequence 5'-GATC. Furthermore, a type I system was found by similarity and was confirmed with a predicted recognition sequence of 5'-AAGNNNNNCTCC by Pyne et al. (2014a).

These enzymes could all present hurdles when trying to transform *C. pasteurianum*. However, [Pyne et al. \(2013\)](#) managed the first reported electrotransformation. They protected plasmid DNA by *in-vivo* methylation in *dam*<sup>+</sup>-*E. coli* (protecting the CpaPI 5'-GATC recognition sequence) and by co-harboring plasmid pFnuDIIMKn expressing the M.FnuDII methyltransferase with the plasmid of interest. M.FnuDII methylates the CpaAI recognition sequence on the first cytosine (5'-<sup>m</sup>CGCG) in both strands, rendering it inaccessible to CpaAI ([Pyne et al., 2013](#)). By methylation of the plasmid DNA a first successful transformation was observed with low efficiencies (approximately 10 transformants per  $\mu\text{g}$  plasmid DNA) compared to other *Clostridium* species. Pyne and co-workers were, however, successful in increasing the number of transformants to  $7.5 \times 10^4 \mu\text{g}^{-1}$  plasmid DNA by optimising the protocol ([Pyne et al., 2013](#)). These findings could not be reproduced in our laboratory by [Rotta \(2014\)](#) who used M.BepI for protection against CpaAI. This study follows on from his initial efforts.

### 3.1.2 Electroporation

The method of choice for simple (plasmid) DNA transformation is electroporation using a strong electric field. This is due to the ease of the technique when compared with more labour intensive and, at the same time, less efficient techniques like conjugation, where a donor organism is mated with the organisms of choice (see *e.g.* [Bouillaut et al. \(2011\)](#)), or transduction, which needs a suitable bacteriophage donor (see *e.g.* [Thomason et al. \(2007\)](#)). Furthermore, heat shock is not a viable technique in many organisms due to low efficiencies or total failure of DNA uptake and the need for protoplast development in some species (see *e.g.* [Lin and Blaschek \(1984\)](#)). For a more comprehensive review of methods of transformation with assets and drawbacks the reader is referred to [Aune and Aachmann \(2010\)](#). Electroporation offers a physical way of delivering DNA inside the cell which works in many species ([Aune and Aachmann, 2010](#); [Kotnik et al., 2015](#); [Tracye, 2013](#)). The principle of electroporation is to apply an electric field to a membrane (be it a cell/bacterial suspension, individual cells or planar bilayer lipid membranes). If the electric field is strong enough this leads to an increase in electric conductivity or permeability (hence the synonymously used name 'electropermeabilisation') and in some cases mechanical rupture of the membranes ([Weaver and Chizmadzhev, 1996](#)). If the voltage is chosen carefully the lipids of the membrane bilayer invert and create a hydrophilic pore allowing molecules to pass which is hindered under normal conditions ([Rosazza et al., 2016](#)). The permeabilisa-



**Figure 3.1** – Theory of electroporation. (a) Exponential decay curve as it occurs during electroporation.  $V_0$  is the initial voltage.  $\tau = V_0/e$  is the time constant of an electroporation and a measure for purity of the sample. (b) Square wave form as it occurs during electroporation. Due to the set-up of modern electroporators a square wave is only approximated with a truncated exponential decay wave.  $V_0$  is the initial,  $V_t$  the final voltage on the sample, with the difference between the two being called droop. (c) Simplified scheme of the electrical set-up of a modern electroporator. A DC power source charges a capacitor (C), the switch (S) closes the circuit and C discharges through the sample ( $R_s$ ) and the parallel resistance ( $R_p$ ).

---

tion is short lived ( $\sim$ sec) (Aune and Achmann, 2010) and is still the matter of scientific discussion, as is the nature and development of the ‘pores’ (Rosazza et al., 2016).

Once pores are formed the molecules have to enter the cell. For this phenomenon several models are proposed but no final evidence is given for any of the following: electropores, stable macropores allowing for diffusion of molecules into the cells; electrophoresis, a high polarisation between intracellular and extracellular membrane which pushes the DNA through electropores; endocytosis, entrapment of molecules in unilamellar vesicles and subsequent internalisation (Rosazza et al., 2016). Once the target molecule has entered the cell, the membrane has to reseal which happens in minutes after the electric discharge

and the whole cell has to repair physiological changes due to the electrical shock which can take hours (Rosazza et al., 2016).

### 3.1.2.1 Electroporation of Bacteria

Two methods of electrically pulsing a bacterial cell mixture are used: exponential decay (Figure 3.1a), where a voltage is applied from a capacitor, which is allowed to discharge freely over time, and square wave (Figure 3.1b), where a steady voltage is applied for a certain amount of time. Based on these techniques there are subcategories like the application of several pulses or cutting short of an exponential decay curve. Commercial electroporators largely work by the same principle: a capacitor is charged to a chosen potential, the capacitor is discharged through an electrical circuit with the biological sample and a resistance connected in parallel (see Figure 3.1c). The following theory is based on Tracye (2013).

**Exponential Decay Electroporation** In exponential decay electroporation the capacitor discharges quickly through the sample with a peak voltage  $V_0$  which declines exponentially over time as indicated in Figure 3.1a.

The time-point when the voltage reaches a value of  $V_0/e$  is called  $\tau$  or simply time-constant which is a convenient measure for the pulse length and gives the researcher insight into the quality of the sample (resistance of the sample due to left over salts in the medium). The electric field (E) over the sample is dependent on the voltage (V) and the distance (d) of the electrodes according to Equation 3.1. The distance is a fixed value dependent on the cuvettes used to transform the samples.

$$E = V/d \quad (3.1)$$

Equation 3.1 shows that the electric field over by the sample changes over time with the voltage. In a conventional set up, the sample resistance ( $R_s$ ) will be much higher than the parallel resistance ( $R_p$ ) and thus the circuits resistance ( $R_c$ ) is mainly defined by the parallel resistance according to Equation 3.2 (compare Figure 3.1c). This will influence the time constant of the pulse and reduce its value.

$$R_c = \frac{R_s * R_p}{R_s + R_p} \quad (3.2)$$

**Square Wave Electroporation** In square wave electroporation the voltage should be the same over the whole electric discharge through the sample. How-

ever, due to using a capacitor to produce this wave form in modern electroporators, this can be approximated only with a truncated exponential decay wave (see Figure 3.1b). The square wave is produced by discharging the capacitor (C, capacitance) for a chosen time (t) through a sample with the resistance  $R_s$  (see Figure 3.1c). According to Equation 3.3 this leads to a voltage decay over time with a final voltage  $V_t$  (compare Figure 3.1b).

$$\ln\left(\frac{V_0}{V_t}\right) = \frac{t}{R_s C} \quad (3.3)$$

The droop is then given as a fractional decrease of the initial voltage according to Equation 3.4 which is usually displayed by the electroporator and can give insights into the sample quality similar to the time constant  $\tau$  in exponential decay electroporation.

$$Droop = \frac{V_0 - V_t}{V_0} \quad (3.4)$$

To most closely approximate true square wave the droop must be minimised which can be accomplished by maximising sample resistance. This can be achieved by cooling the sample and by washing it carefully to reduce salts in the medium.

Irrespective of choosing exponential decay or square wave electroporation for the transformation of microorganisms the researcher has to carefully adjust voltage and parallel resistance as well as using an appropriate electroporation buffer to be successful.

### 3.1.3 Aim of This Study

This study aimed to enhance electro-transformation efficiency for transformation of *C. pasteurianum*. Shortly after commencing the study a transformation protocol for the organism was published (Pyne et al., 2013) but it soon became clear that the method could not be reproduced satisfactorily in our laboratory. This led to the development of a screening method to find hypertransformable clones (transformable with high efficiency). The resulting clones were used to optimise the existing transformation protocol and for subsequent studies (Chapter 4). Furthermore, this study aims to elucidating the underlying reason for the hypertransformable phenotype by next generation sequencing.

## 3.2 Results

### 3.2.1 Development of Electro-Transformation

When this work began no transformation technique for *Clostridium pasteurianum* was available. Rotta (2014) tried different techniques with minor success rate ( $<10$  cfu/ $\mu$ g). He tried electroporation with methylation against the *C. pasteurianum* CpaAI restriction system. Based on these early results measures were taken to optimise the electroporation protocol to increase efficiency to a useful level ( $\sim 10^4$  cfu/ $\mu$ g).

#### 3.2.1.1 Testing Different Wild Type Strains

Initial attempts of altering buffers and electroporation parameters (data not shown) were unsuccessful. Pyne et al. (2013) published a protocol which was successful in *C. pasteurianum* ATCC 6013 and gave transformation efficiencies of up to  $7.5 \times 10^4$  transformants per  $\mu$ g plasmid. However, duplication of the described method failed in our wild type strain *C. pasteurianum* DSM 525. *C. pasteurianum* ATCC 6013 and DSM 525 were at the time reported to be the same strain but Poehlein et al. (2015) and Rotta et al. (2015) showed that there were distinctive differences between the strains. For confirmation of the electroporation protocol reported by Pyne et al. (2013), several different *C. pasteurianum* strains were ordered and subjected to the published electroporation protocol (Pyne et al., 2013).

Strains tested were *C. pasteurianum* DSM 525, DSM 526, DSM 9989, DSM 12136 and ATCC 6013. Strain *C. pasteurianum* DSM 12136 failed to grow after repeated assays of culturing. Type-strain DSM 525 had a success rate of 8 cfu/ $\mu$ g, DSM 9989 of 20 cfu/ $\mu$ g and DSM 526 was not able to recover after glycine addition and we were thus unable to transform this strain using this protocol. Strain ATCC 6013, the strain used by Pyne et al. (2013) could not be transformed in three independent attempts by different experimentators.

#### 3.2.1.2 Screening for a Hypertransformable Clone

Based on the fact that *C. pasteurianum* DSM 525 could repeatedly be transformed with low efficiencies it was hypothesised that the few transformants were most likely a result of the presence of rare mutant variants within the culture that are highly competent for DNA transfer. This could lead to the observed low transformation efficiencies if only this sub-population takes up the DNA. Furthermore, having unwittingly cultured such a sub-population, this would explain the ability of Pyne et al. (2013) to transform *C. pasteurianum*.



*anum*. Attempts were made to obtain the strain used in the study of Pyne et al. (2013) but the inquiry was refused. To test this hypothesis a transformed colony that exhibited antibiotic resistance was picked and restreaked onto an antibiotic bearing plate to confirm resistance and successful plasmid transformation. The colony was then grown over-night in non-selective liquid medium and culture plated on selective and non-selective plates to select for plasmid loss. 36 colonies were tested on selective and non-selective medium, all of which have lost the plasmid judged by sensitivity to thiamphenicol. The high rate of plasmid loss shows the instability of the Gram-positive replicon pIM13 in *C. pasteurianum*. Of the 36 clones one (denoted *C. pasteurianum*-H1) was kept and subjected to electro-transformation. Transformation of *C. pasteurianum*-H1 was considerably improved to  $2.6 \times 10^5$  cfu/ $\mu$ g compared to the wild type DSM 525 control with  $1.6 \times 10^2$  cfu/ $\mu$ g. The higher transformation efficiency of the wild type strain compared to initial assays (Section 3.2.1.1) are presumably due to the practice of the experimenter but were variable between experiments. Reproduction of wild type transformation is demonstrated below (Figure 3.8).

To demonstrate reproducibility of the screening process for a hypertransformable clone the assay described above was repeated six times independently. After plasmid loss, isolates 1, 2, 5 and 6 showed the same efficiency as the wild type control and isolates 3 and 4 showed comparable efficiencies to *C. pasteurianum*-H1 (Figure 3.2). These isolates 3 and 4 were denoted *C. pasteurianum*-H3 and *C. pasteurianum*-H4, respectively. The results corroborate our hypothesis and validated the screening methods for a hypertransformable sub-population in *C. pasteurianum*. All of the hypertransformable clones were transformed routinely in our laboratory with transformation efficiencies in the order of magnitude of  $10^4$  cfu/ $\mu$ g plasmid.

### 3.2.1.3 Optimisation of the Transformation Protocol

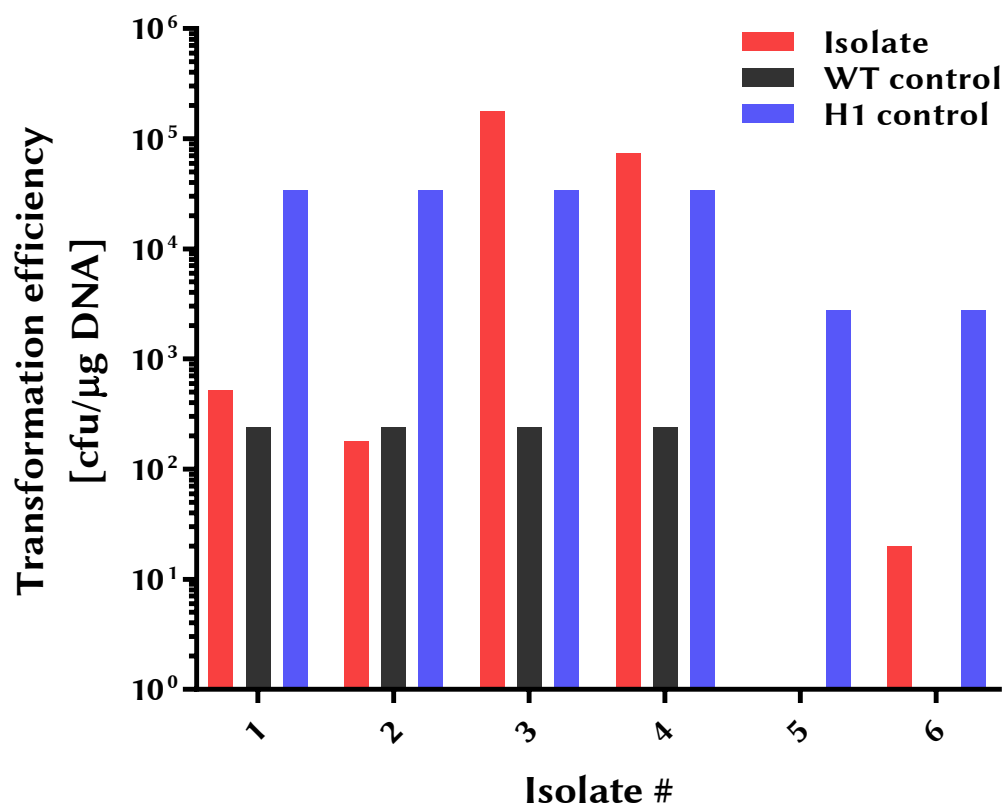
With successful isolation of *C. pasteurianum*-H1, optimisation of the transformation protocol was planned to further increase transformation efficiencies. Pyne et al. (2013) has already optimised buffers used in the protocol. Hence, altering of the wave form to square wave was tested, along with the alteration of the field strength.

The electric field was altered by changing the voltage and leaving the cuvette distance the same ( $d = 0.4$  cm). Furthermore, time of the pulse and, if multiple pulses were applied, the interval between them, was varied. Table 3.1 shows the varied parameters and the achieved transformation efficiencies. When ap-



**Table 3.1** – Parameters assayed for electroporation optimisation and the results as cfu/ $\mu\text{g}$  plasmid and the ratio of sample transformation over control transformation. N/A = unsuccessful transformations due to contaminations. Each parameter was initially tested once and promising parameters were further inspected (Figure 3.3).

Wave form	Voltage [kV]	Time [ms]	Interval [ms]	Sample size	Efficiency [cfu/ $\mu\text{g}$ ]	Control efficiency [cfu/ $\mu\text{g}$ ]	Ratio [sample/control]
Exponential decay	1800				-	control	-
	2.5			1	$3.0 \times 10^4$	$1.0 \times 10^5$	0.3
	3.0			1	$2.0 \times 10^3$	$1.0 \times 10^5$	0.0
Square wave	1.8	5		1	$3.0 \times 10^4$	$1.0 \times 10^5$	0.3
	2.5	5		1	$4.0 \times 10^4$	$1.0 \times 10^5$	0.4
	1.5	1		1	$7.0 \times 10^3$	$1.6 \times 10^4$	0.4
	1.5	1		1	$1.2 \times 10^2$	$2.6 \times 10^5$	0.0
	2.0	1		1	$3.8 \times 10^2$	$2.6 \times 10^5$	0.0
	2.0	2		1	N/A	$2.6 \times 10^5$	N/A
	2.0	3		1	N/A	$2.6 \times 10^5$	N/A
	2.0	3		1	$2.4 \times 10^3$	$1.6 \times 10^4$	1.5
	2.0	4		1	N/A	$1.6 \times 10^4$	N/A
	2.0	5		1	$2.3 \times 10^4$	$1.6 \times 10^4$	1.4
	2.5	1		1	$2.0 \times 10^2$	$1.6 \times 10^4$	0.0
	2.5	1		1	$3.6 \times 10^2$	$2.6 \times 10^5$	0.0
	2.5	2		1	$3.4 \times 10^4$	$1.6 \times 10^4$	2.1
	2.5	3		1	$2.2 \times 10^4$	$1.6 \times 10^4$	1.4
	2.5	4		1	$5.0 \times 10^5$	$1.6 \times 10^4$	31.3
	2.5	5		1	$6.6 \times 10^5$	$1.6 \times 10^4$	41.3
	2.5	5		2	$2.1 \times 10^3$	$5.5 \times 10^2$	3.9
	2.5	5		2	$3.0 \times 10^3$	$1.0 \times 10^5$	0.0
	3.0	5		1	$4.0 \times 10^3$	$1.0 \times 10^4$	0.4
	2.0	1	5	1	$2.5 \times 10^0$	$5.5 \times 10^2$	0.0
2.0	5	5	1	$1.4 \times 10^2$	$5.5 \times 10^2$	0.2	
2.5	1	5	1	$1.7 \times 10^1$	$5.5 \times 10^2$	0.0	
2.5	5	5	1	$6.5 \times 10^2$	$5.5 \times 10^2$	1.2	
2.5	5	15	1	$2.6 \times 10^1$	$5.5 \times 10^2$	0.0	

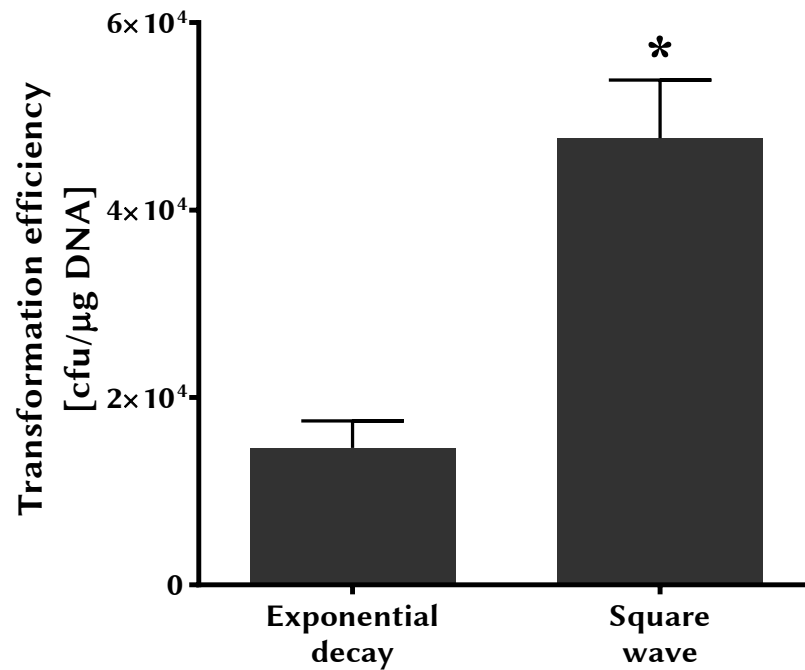


**Figure 3.2** – Results of screening for hypertransformable strains. The screening was done in six individual assays (isolates). Wild type and the hypertransformable strain *C. pasteurianum*-H1 were used as controls in the re-transformation experiment.

plying an exponential decay pulse the voltage chosen as control was the most efficient (1800 V, compare [Pyne et al. \(2013\)](#)). Shortening the exponential decay pulse to 5 ms negatively influenced the transformation efficiency. However, with increasing voltage and increasing pulse times, when applying a square wave pulse, the efficiency increased. 2500 V for 5 ms showed a greater than 40 fold increase in transformation efficiency compared to the control (Table 3.1). This setting was further tested in three biological repeats with technical triplicates each against the published setting and the square wave approach was found to be more than three times more efficient than exponential decay (see Figure 3.3).

### 3.2.2 Next Generation Sequencing of Relevant Strains

The genomes of strains *C. pasteurianum* DSM 525, H1, H3 and H4 were subjected to Illumina sequencing. Sequencing data of type strain *C. pasteurianum* DSM 525 is published as [Poehlein et al. \(2015\)](#). The genome serves as reference for resequencing analysis. Illumina data of the published genome [Pyne et al.](#)



**Figure 3.3** – Histogram showing the transformation efficiencies of 5 individual exponential decay transformations and 6 individual square wave transformations. Square wave electroporation ( $m= 47667$  cfu/ $\mu$ g, SD= 17550 cfu/ $\mu$ g) is 3.6 times more efficient than exponential decay electroporation ( $m= 14625$  cfu/ $\mu$ g, SD= 7619 cfu/ $\mu$ g),  $t(15)= 4.629$ ,  $p= 0.0003$ .

(2014b) was collected from the Sequence Read Archive (SRA).

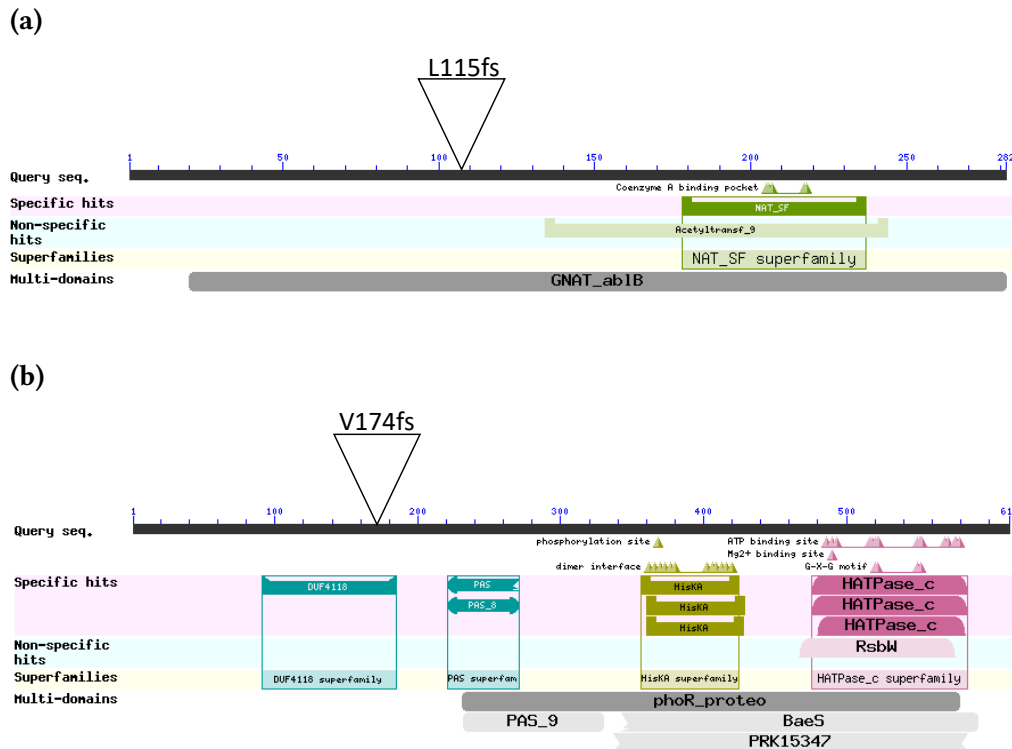
All sequenced strains were mapped against the reference genome (Poehlein et al., 2015) and statistics are compiled in Table 3.2. All mappings were done with high rate of reads mapped (>97 %) with most reads mapped in pairs (>94 %). The quality of the reads mapped was good with an average read length >249 bp for genomes sequenced in this study and >149 bp for the genome sequenced by Pyne et al. (2014b). The discrepancy can be explained by different technology used with a maximal read length of 250 bp for this study and 150 bp for Pyne et al. (2014b). The read-mapping was used to call Single Nucleotide Polymorphisms (SNPs) and small structural variants. Newly emerged variants are listed in Table 3.3.

**Table 3.2** – Summary reports of Illumina read-mapping against reference genome *C. pasteurianum* DSM 525 (Poehlein et al., 2015) of reported strains *C. pasteurianum* H1, H3 and H4 as well as published reads of Pyne et al. (2014b).

	Read count	Reads [%]	Average length	Bases count	Bases [%]
<i>C. pasteurianum</i> -H1					
Mapped reads	1353861	99.06	249.75	338129803	99.06
Not mapped reads	12851	0.94	249.02	3200170	0.94
Reads in pairs	1314700	96.19	550.42	329102299	96.42
Broken paired reads	39161	2.87	230.52	9027504	2.64
Total reads	1366712	100.00	249.75	341329973	100.00
<i>C. pasteurianum</i> -H3					
Mapped reads	2053015	98.86	249.74	512722204	98.86
Not mapped reads	23703	1.14	248.94	5900584	1.14
Reads in pairs	1974576	95.08	528.56	494557068	95.36
Broken paired reads	78439	3.78	231.58	18165136	0.35
Total reads	2076718	100.00	249.73	518622788	100.00
<i>C. pasteurianum</i> -H4					
Mapped reads	2060145	97.19	249.54	514096309	97.19
Not mapped reads	59505	2.81	249.68	14857121	2.81
Reads in pairs	2010532	94.85	519.92	501741186	94.86
Broken paired reads	49613	2.34	249.03	12355123	2.34
Total reads	2119650	100	249.55	528953430	100.00
Pyne et al. (2014b)					
Mapped reads	10095012	98.76	149.93	1513530521	98.76
Not mapped reads	126450	1.24	149.68	18926728	1.24
Reads in pairs	9920260	97.05	454.69	1487464442	97.06
Broken paired reads	174752	1.71	149.16	26066079	0.17
Total reads	10221462	100.00	149.93	1532457249	100.00

**Table 3.3** – SNPs and small structural variants called from read-mapping of the hypertransformable strains (H1, H3, H4) and the published genome of [Pyne et al. \(2013\)](#) against *C. pasteurianum* DSM 525 reference genome ([Poehlein et al., 2015](#)).

Reference position	Reference	Allele	Count	Coverage	Frequency	Average quality	Amino acid change
<i>C. pasteurianum</i> -H1							
1458053	-	T	70	74	94.59	37.69	CLPA_c13710 ( <i>ablB</i> ): Val174fs
3564785	-	T	83	86	96.51	37.96	CLPA_c33080 ( <i>resE9</i> ): Leu115fs
<i>C. pasteurianum</i> -H3							
3284805	A	C	104	104	100.00	37.38	CLPA_c30550: Phe1102Val
<i>C. pasteurianum</i> -H4							
3284805	A	C	109	109	100.00	36.72	CLPA_c30550: Phe1102Val
<a href="#">Pyne et al. (2014b)</a>							
828795	G	-	136	306	44.44	35.77	CLPA_c07500: Arg91fs
828798	-	A	130	309	42.07	37.38	CLPA_c07500: Gln93fs
2418487	ACTCAATAG	-	90	131	68.70	36.60	



**Figure 3.4** – Conserved domains (Marchler-Bauer et al., 2011) of CLPA\_c13710 (a) and CLPA\_c33080 (b). a) Gene CLPA\_c13710 (*ablB*) annotated as  $\beta$ -lysine N<sub>6</sub>-acetyltransferase with the location of the variants indicated as large triangle. Conserved domains as follows: NAT\_SF (N-acyltransferase superfamily), various enzymes that characteristically catalyse the transfer of an acyl group to a substrate (e-value=  $3.5 \times 10^{-4}$ ); GNAT\_ab1B, putative  $\beta$ -lysine N-acetyltransferase (e-value=  $5.5 \times 10^{-88}$ ). b) Gene CLPA\_c33080 (*resE9*) annotated as sensory box histidine kinase with location of variant. Putatively relevant conserved domains: DUF4118 (domain of unknown function), likely to be a transmembrane domain involved in ligand sensing (e-value=  $10^{-8}$ ); PAS domain, have been found to bind ligands, and to act as sensors for light and oxygen in signal transduction (e-value=  $5 \times 10^{-4}$ ); HisKA, histidine kinase A (e-value=  $1.5 \times 10^{-12}$ ); HATPase\_c, Histidine kinase-like ATPases (e-value=  $3 \times 10^{-38}$ ); phoR\_proteo, phosphate regulon sensor kinase PhoR (e-value=  $1 \times 10^{-29}$ ).

### 3.2.2.1 Genomic Variant in *C. pasteurianum*-H1

The hypertransformable clone from the first screening assay *C. pasteurianum*-H1 has two SNP differences to the wild type. Both SNPs were confirmed by Sanger sequencing. The SNPs are found in genes CLPA\_c13710 and CLPA\_c33080. Gene CLPA\_c13710 annotated as  $\beta$ -lysine N<sub>6</sub>-acetyltransferase (*ablB*) has an insertion of a thymine after position 520 leading to a frame-shift (Figure 3.4a). The frame-shift leads to a premature stop codon at position 529 in the

849 bp long gene. A search against the conserved domain database confirmed the annotation, with the gene showing domains from N-acetyltransferases and specifically  $\beta$ -lysine N-acetyltransferases (see Figure 3.4a). The putative active centre and several putative coenzyme-A binding pockets are located downstream of the mutation and it is thus likely that the frame-shift inactivates the protein completely.

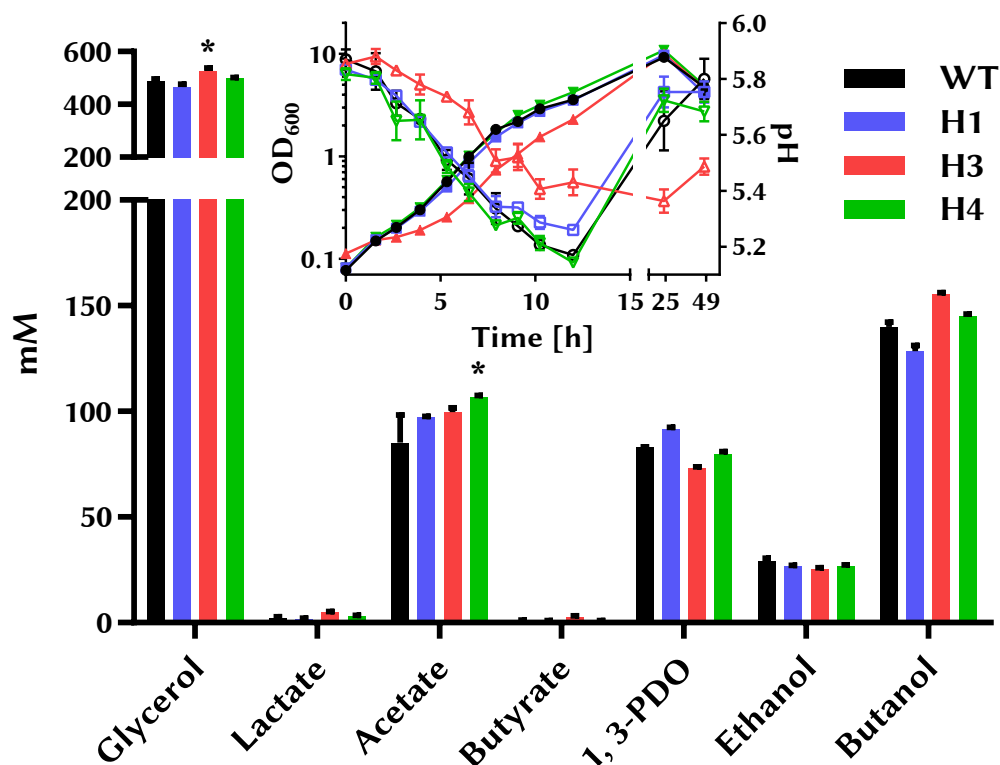
The insertion of a thymine after position 344 bp introduces a frame-shift in gene CLPA\_c33080, as well. The frame-shift leads to a premature stop-codon at position 481 bp in the 1845 bp long sensory box histidine kinase gene (*resE9*) (Figure 3.4b). The conserved domain database finds domains from histidine kinases with conserved phosphorylation sites and ATP-binding sites (Figure 3.4b). It, furthermore, finds DUF4118- (hypothetical ligand sensor) and PAS-domains (ligand binding and signal transduction) suggesting membrane location of the protein. The N-terminal thymine insertion in this multi-domain protein and subsequent premature stop-codon likely deactivate this gene.

### 3.2.2.2 Genomic Variants in *C. pasteurianum*-H3 and -H4

Two independent clones with high transformation efficiency were found in the subsequent experiments aiming at the validation of the hypertransformable screening method. Illumina sequencing and SNP calling revealed the same single nucleotide variation in both strains. The SNP is located in gene CLPA\_c30550 (encoding a hypothetical protein) where T3304 is replaced by a guanine leading to an amino acid change Phe1102Val (Figure 3.9). The SNP was confirmed in both strains by Sanger sequencing and the absence of variants found in *C. pasteurianum*-H1 was confirmed. Protein analysis of CLPA\_c30550 is detailed in Section 3.2.5.

### 3.2.2.3 Genomic Variants in the Pyne Genome

A SNP analysis between *C. pasteurianum* DSM 525 (Poehlein et al., 2015) used in this study and ATCC 6013 used by Pyne et al. (2013) might reveal the reason for the ability of Pyne et al. (2013) to transform *C. pasteurianum* with high efficiencies, which could not be repeated in this study. There were two SNP differences and one insertion found distinguishing the two genomes. Variants are located at reference positions: 828795, 828798 and 2418487. All of these variants are called with low frequency which usually indicates that they are only present in a sub-population. Both variants 828795 and 828798 are located in gene CLPA\_c07500 which encodes a small (103 amino acids) hypothetical protein for which a BLAST and conserved domain database search does not



**Figure 3.5** – Bottle fermentation analysis of different hypertransformable strains *C. pasteurianum*-H1 (H1), *C. pasteurianum*-H3 (H3) and *C. pasteurianum*-H4 (H4) and *C. pasteurianum* DSM 525 (WT). Final fermentation product concentrations measured after 24 h. Glycerol is presented as glycerol consumed in 24 h. Symbols for OD<sub>600</sub> are filled and pH empty. Asterisks indicate significant difference of compound concentration of hypertransformable strains compared to wild type determined by independent t-test with Holm-Sidak multiple comparison correction ( $\alpha=0.05$ ). Data presented as mean with error-bars indicating standard error (SEM).

find any additional information. The deletion at 2418487 is located in the intergenic region downstream of CLPA\_c22570 (transposase) and upstream of CLPA\_c22560 (hypothetical protein), and it is likely to be benign.

### 3.2.3 Growth Characteristics of Hypertransformable Clones

To test if the mutations in the hypertransformable strains had any influence on solvent production in *C. pasteurianum* DSM 525, all hypertransformable strains were tested in bottle fermentations in Biebl medium with 1 g/l yeast extract and 60 g/l glycerol. Figure 3.5 shows the growth pattern and solvent profile of the hypertransformable strains *C. pasteurianum*-H1, *C. pasteurianum*-H3 and *C. pasteurianum*-H4 using the wild type *C. pasteurianum* DSM 525 as con-

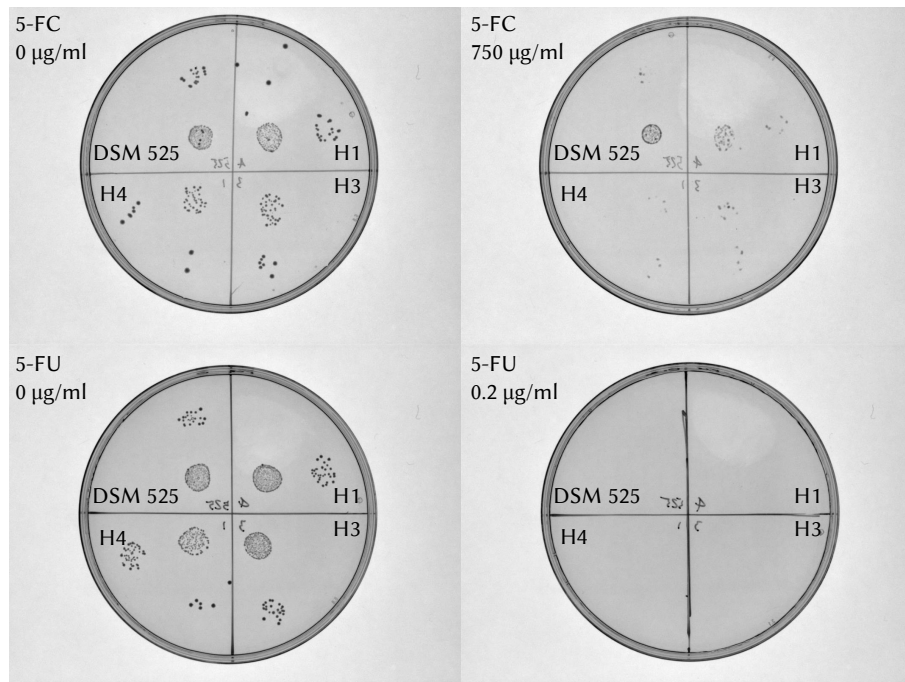


**Table 3.4** – Growth characteristics and product yields of different hypertransformable strains and WT *C. pasteurianum* DSM 525. Data from serum bottle fermentations in biological triplicates. Data given as mean $\pm$ SEM.

Carbon source Medium Strain	60 g/l glycerol Biebl plus 1 g/l yeast extract			
	WT	H1	H3	H4
Growth characteristics				
Specific Growth Rate [h <sup>-1</sup> ]	0.40 $\pm$ 0.00	0.37 $\pm$ 0.00	0.37 $\pm$ 0.01	0.38 $\pm$ 0.00
Doubling Time [min]	103 $\pm$ 0.1	113 $\pm$ 0.3	113 $\pm$ 1.8	109 $\pm$ 0.5
Max. OD <sub>600</sub> []	9.24 $\pm$ 0.16	9.71 $\pm$ 0.12	9.97 $\pm$ 0.16	10.82 $\pm$ 0.04
Carbon recovery [%]	117 $\pm$ 3	122 $\pm$ 1	116 $\pm$ 1	123 $\pm$ 1
Selectivity [M/M]				
BuOH/Solvents	0.55 $\pm$ 0.01	0.52 $\pm$ 0.00	0.61 $\pm$ 0.00	0.58 $\pm$ 0.00
Yield [M/M]				
BuOH/C-Source	0.31 $\pm$ 0.00	0.30 $\pm$ 0.00	0.32 $\pm$ 0.00	0.32 $\pm$ 0.00
EtOH/C-Source	0.07 $\pm$ 0.00	0.06 $\pm$ 0.00	0.05 $\pm$ 0.00	0.06 $\pm$ 0.00
PDO/C-Source	0.19 $\pm$ 0.00	0.22 $\pm$ 0.00	0.15 $\pm$ 0.00	0.18 $\pm$ 0.00
Solvents/C-Source	0.57 $\pm$ 0.00	0.58 $\pm$ 0.01	0.52 $\pm$ 0.00	0.56 $\pm$ 0.01
Acids/C-Source	0.20 $\pm$ 0.02	0.24 $\pm$ 0.01	0.22 $\pm$ 0.01	0.25 $\pm$ 0.00

trol. Their growth characteristics and yield calculations were summarised in Table 3.4. In general all strains show comparable characteristics. *C. pasteurianum*-H3 has a longer lag-phase and its pH decreases slower and less than in the other mutants (Figure 3.5). The difference in growth of *C. pasteurianum*-H3 is likely to be explained by differences in the inoculum e.g. inoculum age, since clones H3 and H4 exhibit the same genotype and H4 does not exhibit the same growth. It was reported that small differences in culture condition can lead to large differences in growth and solvent formation (Moon et al., 2011; Johnson and Rehmann, 2016).

Glycerol consumption is significantly increased in H3 (m= 527.9 mM, SD= 14.3 mM) compared to the wild type (m= 487.1 mM, SD= 12.0 mM), t(28)= 5.67, p= 0.00. For acid production H4 (m= 106.6 mM, SD= 1.2 mM) exhibits a slight but significant difference to the wild type (m= 85.3 mM, SD= 18.6 mM), t(28)=3.46, p= 0.012. There are no significant differences in solvent production between any strain and the wild type. As expected from the genotype all hypertransformable strains behave similarly to the wild type in respect to solvent production what makes them useful chassis for metabolic engineering with the aim to increase solvent production.



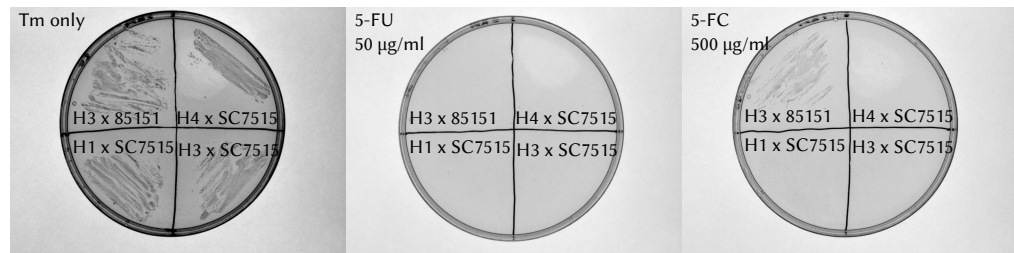
**Figure 3.6** – Minimal inhibitory concentration determination of 5-fluorocytosine (5-FC) and 5-fluorouracil (5-FU). Solid CBM medium with different concentrations of 5-FC and 5-FU to test sensitivity of *C. pasteurianum* wild type and hypertransformable strains *C. pasteurianum* DSM 525, *C. pasteurianum*-H1, *C. pasteurianum*-H3 and *C. pasteurianum*-H4. Strains can grow on plates with 750 µg/ml 5-FC and are sensitive to as little as 0.2 µg/ml 5-FU.

### 3.2.4 SNP in CLPA\_c30550 is the Direct Cause of Hypertransformability in *C. pasteurianum*

Based on the SNP analysis of the hypertransformable strains it was decided to analyse *C. pasteurianum*-H3. This strain only carries one SNP as opposed to *C. pasteurianum*-H1 with two SNPs. Furthermore, the genotype of *C. pasteurianum*-H3 is identical to the one of *C. pasteurianum*-H4. To analyse the involvement of the SNP in the hypertransformable phenotype it was corrected back to wild type by allelic exchange and the resulting strain evaluated for transformation efficiency. To overcome the need of producing a *pyrE* truncated strain for the application of allelic exchange for 5-fluoroorotic acid (FOA) based counterselection, a different approach was chosen based on allelic exchange with *codA* as a counterselection marker (Cartman et al. (2012), Section 1.3.4.2).

#### 3.2.4.1 *codA* for Counterselection

To use the *codA* gene as a counterselection marker in *C. pasteurianum*, presence of gene *upp*, encoding the uracil phosphoribosyltransferase, had to be ascer-

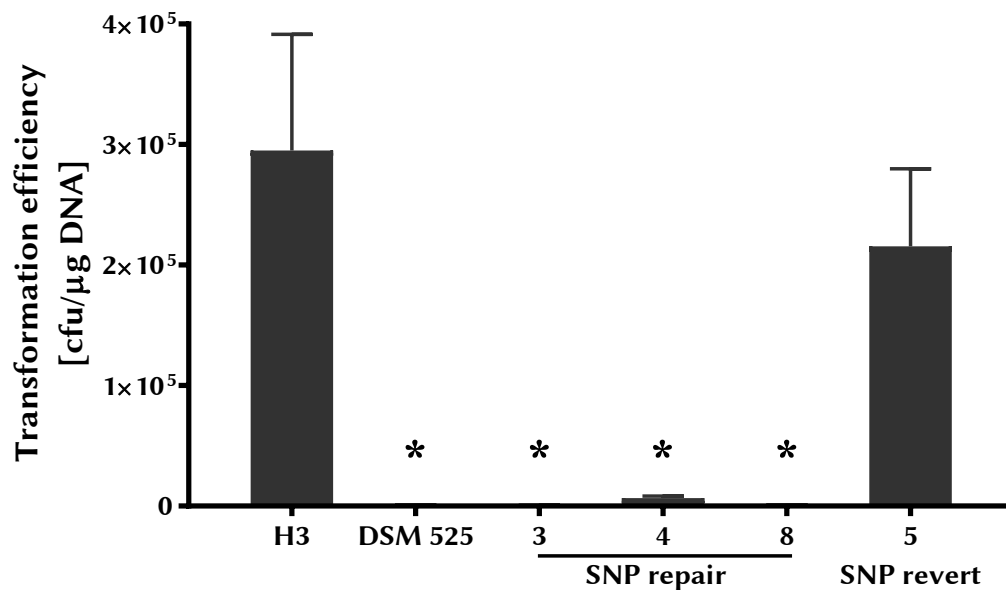


**Figure 3.7** – Minimal inhibitory concentration testing of 5-fluorocytosine (5-FC) when *codA* (on plasmid pMTL-SC7515) is present in the strain. All plates are supplemented with thiamphenicol. 5-fluorouracil (5-FU) serves as positive control inhibiting growth of all strains and 5-FC inhibits all strains carrying plasmid pMTL-SC7515. Strains with plasmids tested: *C. pasteurianum*-H3 x pMTL85151 (control plasmid without *codA*), *C. pasteurianum*-H4 x pMTL-SC7515, *C. pasteurianum*-H3 x pMTL-SC7515, and *C. pasteurianum*-H1 x pMTL-SC7515.

tained. It was found by BLAST sequence alignment that gene CLPA\_c02720 is 67 % identical on protein level to the *upp* gene of *C. difficile* 630. Furthermore, absence of a homologous *codA*-gene in *C. pasteurianum* was confirmed by BLAST alignment. Next, the sensitivity of *C. pasteurianum* to 5-fluorocytosine (5-FC) and 5-fluorouracil (5-FU) was tested (Figure 3.6). *C. pasteurianum* is resistant to high levels of 5-FC (750  $\mu\text{g/ml}$ ) and sensitive to the lowest tested level of 5-FU (0.2  $\mu\text{g/ml}$ ) suggesting that *codA* based allelic exchange is feasible if 5-FC can be converted to 5-FU. This was further tested in *C. pasteurianum* by transforming all hypertransformable strains with plasmid pMTL-SC7515 carrying *codA* and *catP* (bestowing thiamphenicol resistance). Viability was tested on 5-FC plus thiamphenicol (Tm) plates. Figure 3.7 shows that all strains grow on CBM plates supplemented with Tm only (negative control) and that 50  $\mu\text{g/ml}$  5-FU kills all strains (positive control). Strains carrying pMTL-SC7515 (Table 2.3) are rendered sensitive to 5-FC as can be seen on plates with 500  $\mu\text{g/ml}$  5-FC plus Tm where only *C. pasteurianum*-H3 with control plasmid pMTL85151 (not carrying *codA*) is able to grow.

#### 3.2.4.2 Correction of SNP in *C. pasteurianum*-H3

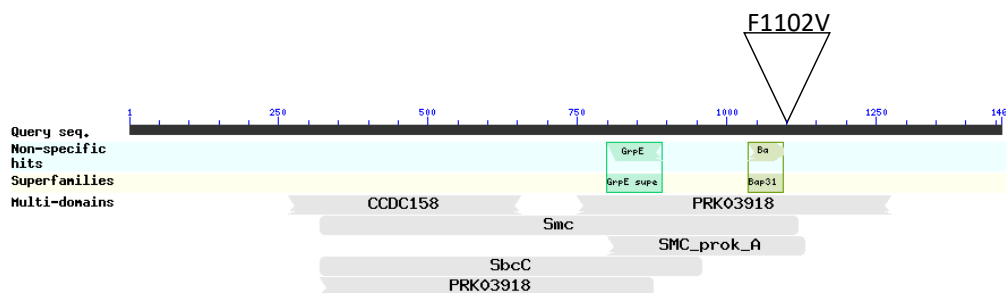
To correct the SNP in *C. pasteurianum*-H3 back to wild type, a *codA* (Cartman et al., 2012) allelic exchange plasmid was constructed. For this, the fragment *catP-pyrE* in pMTL-KS15 (Schwarz et al., 2017) was substituted with a module containing *catP-codA* (Ying Zhang, personal communication) between sites FseI and PmeI and the new vector was named pMTL-AGH15. Further-



**Figure 3.8** – Histogram showing transformation efficiencies of *C. pasteurianum*-H3 SNP corrected mutants. Efficiencies are given of positive control *C. pasteurianum*-H3 (H3); negative control *C. pasteurianum* DSM 525 (DSM 525); the mutants in which the SNP in CLPA\_c30550 is corrected back to wild type (DSM 525) genotype (SNP repair 3, 4, 8); and the method control strain going through double crossover but reverting to *C. pasteurianum*-H3 genotype (SNP revert 5). Asterisks indicate statistical significance difference to H3 calculated by one-way ANOVA ( $F(5,18) = 7.716$ ,  $p = .0005$ ) with *a posteriori* Dunnett’s multiple comparisons test. Error bars represent standard errors from the mean (SEM).

more, a 1127 bp fragment of the wild type CLPA\_c30550 allele was amplified flanking the SNP in *C. pasteurianum*-H3 by 599 bp down- and 527 bp upstream. The fragment was then cloned into pMTL-AGH15 via SbfI and NheI, yielding the SNP repair plasmid pMTL-AGH15-c30550. This correction plasmid was transformed into *C. pasteurianum*-H3 to correct the SNP back to wild type. Of colonies growing after transformation 24 were restreaked and 14 were streaked to double crossover. Sanger sequencing revealed that exactly half of the double crossovers corrected the SNP back to wild type *C. pasteurianum* and the other half remained as the background *C. pasteurianum*-H3.

Three of the SNP changed clones were chosen (*C. pasteurianum*-H3 SNP repair 3, 4, 8) for transformation with *C. pasteurianum*-H3 as positive and *C. pasteurianum* DSM 525 as negative control as well as one wild type revertant (*C. pasteurianum*-H3 SNP revert 5) as a control of the allelic exchange procedure. Results can be seen in Figure 3.8 and show that correction of the SNP from *C. pasteurianum*-H3 genotype to *C. pasteurianum* DSM 525 genotype has the



**Figure 3.9** – Conserved domains (Marchler-Bauer et al., 2011) of CLPA\_c30550. Triangle shows SNP location where amino acid change F1101V took place. Putatively relevant superfamilies: GrpE, is the adenine nucleotide exchange factor of DnaK (Hsp70)-type ATPases (e-value=  $8 \times 10^{-3}$ ); Bap31, B-cell receptor-associated protein 31-like (e-value=  $9.5 \times 10^{-3}$ ); PRK03918, chromosome segregation protein (provisional) (e-value=  $4 \times 10^{-10}$ ); SMC, chromosome segregation ATPase (Cell cycle control, cell division, chromosome partitioning) (e-value=  $3 \times 10^{-9}$ ); SMC\_prok\_A, chromosome segregation protein SMC, primarily archaeal type (e-value=  $3 \times 10^{-8}$ ); SbcC, DNA repair exonuclease SbcCD ATPase subunit (Replication, recombination and repair) (e-value=  $2 \times 10^{-6}$ ); PRK03918, chromosome segregation protein (provisional) (e-value=  $3.5 \times 10^{-3}$ ).

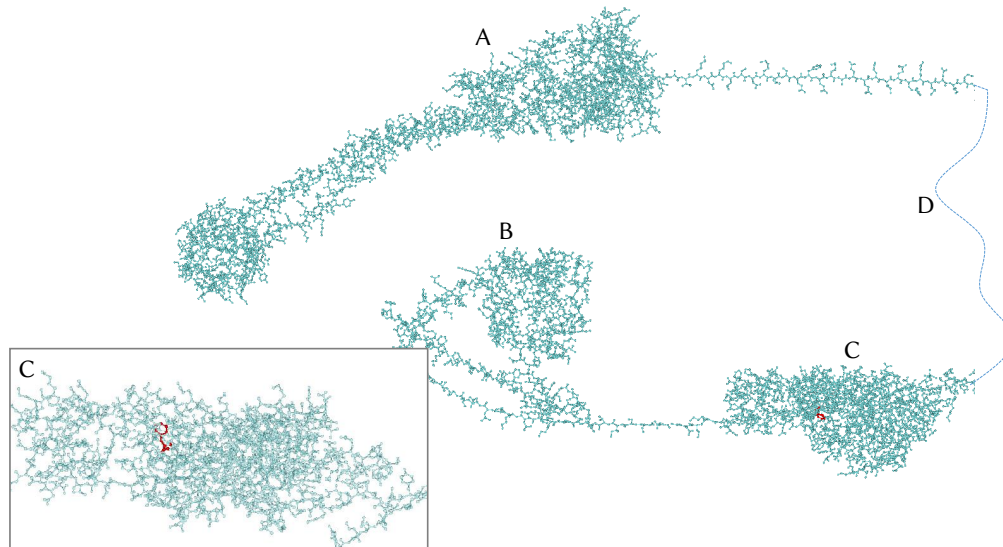
effect to reverse hypertransformability in all SNP repair clones. As expected SNP revertant clone 5 shows the same phenotype as *C. pasteurianum*-H3, concluding the SNP in CLPA\_c30550 in *C. pasteurianum*-H3, and by homology in *C. pasteurianum*-H4, to be responsible for the hypertransformable phenotype. This makes gene CLPA\_c30550 an intriguing target for further investigations into bacterial competence and attempts were made below to illuminate the putative function of this gene.

### 3.2.5 *In Silico* Analysis of Gene CLPA\_c30550

Having ascertained the SNP in gene CLPA\_c30550 to be the reason for hypertransformability of *C. pasteurianum*-H3 a more in-depth *in silico* analysis of the gene was undertaken. Gene CLPA\_c30550 encodes a 1462 amino acid long predicted protein and lies in a gene cluster with CLPA\_c30550 to CLPA\_c30580 with a predicted promoter sequence upstream of CLPA\_c30580. A protein BLAST of CLPA\_c30550 against the BLAST database revealed the protein to be present in a wide range of Firmicutes but not well described and mostly annotated as ‘hypothetical protein’ or ‘putative protein’.

Specific homology searches of CLPA\_c30550 against *C. acetobutylicum*, *C. beijerinckii*, *C. difficile* and *C. sporogenes*, all species with reported electropo-





**Figure 3.10** – Phyre2 (Kelley et al., 2015) predicted 3D model of CLPA\_c30550. Four distinct regions have been modelled based on similarity to the following: A) region M1 to S559: DNA-binding-, membrane-, hydrolase-, transport- and replication-proteins; B) region L1280 to L1433: motor-, transport- and hydrolase-type-proteins; C) region F911 to K1230: protein-transport protein and gene regulation (chaperone) protein; D) region S560 to N910 was not modelled sufficiently due to lack of structural information in Phyre database. The inlaid picture shows the region of interest (C) with amino acid residue F1102, which is mutated in *C. pasteurianum*-H3 and-H4 to V1102, in red.

---

ration protocol, revealed no hits or hits with very low probability (e-value > 0.03). A search of the whole gene cluster CLPA\_c30550-CLPA\_c30580 revealed 94 % coverage with 72 % identity (e-value= 0.0) in the genome of *C. perfringens* SM101. Regions up- and downstream of the cluster are not preserved, suggesting preservation of the whole 7730 bp cluster as one entity. Closer analysis of the genes in *C. pasteurianum* and *C. perfringens* show that residue CLPA\_c30550 F1102 (mutated in *C. pasteurianum*-H3) is preserved in *C. perfringens*.

The conserved domain database search (Figure 3.9) reveals high similarity to structural maintenance of chromosome (SMC) proteins and the Conserved Domain Architecture Retrieval Tool (CDART) found hits with: GrpE, a nucleotide exchange factor of DnaK and Bap31; a B-cell receptor-associated protein as well as hits with membrane associated proteins. Furthermore, the 3D model in Figure 3.10 was produced based on similarities to proteins involved in cell cycle maintenance and chromosome repair or DNA replication. Even though this model is not a reliable representation of the actual protein struc-

ture it can be seen that residue of interest F1102 lies within a modelled area (region B in Figure 3.10) based on protein database (PTB) protein c2oevA involved in protein transport. The Phyre2 model was made with 72 % of residues modelled at >90 % confidence; low confidence residues are only found in regions without tertiary structure (region D in Figure 3.10).

### 3.3 Discussion

Initially this study was intended to develop an efficient transformation protocol for DNA transfer in *C. pasteurianum* to lay the foundation for genetic tool development and metabolic engineering work in this organism. Previous work (Rotta, 2014) investigated the restriction systems of the organism. A combination of genomic data, published studies and original experimental data showed that *C. pasteurianum* type strain ATCC 6013, which only differs in three variants (SNPs and insertion/deletions (InDels)) from type strain DSM 525, possesses two type II restriction methylation systems (CpaI and CpaAI). Rotta (2014) further showed that methylation in a *dam*<sup>+</sup> *E. coli* strain protects against restriction by CpaI and that co-harboring the plasmid of interest with a plasmid expressing the methylase M.BepI (Kupper et al., Kupper1989) protected against restriction by CpaAI by methylating the first cytosine in the target site (5'-<sup>m</sup>CGCG). These findings were later confirmed by Pyne et al. (2013) who used methylase M.FnuDII for protection against CpaAI restriction.

#### 3.3.1 Initial Electroporation Experiments

The method of choice when developing DNA transformation for bacteria is electroporation due to ease of use and reported high success-rates in organisms from all kingdoms of life (Tracye, 2013). This work was associated with previous studies by Rotta (2014) where very low transformation efficiencies were achieved by electroporation (~10 cfu/ $\mu$ g). The efficiency was too low for the development of molecular tools in *C. pasteurianum* and initial effects of altering electroporation buffers and electroporator parameters were unsuccessful. At that time Pyne et al. (2013) published an improved transformation protocol, using cell-weakening agents, reaching a frequency of transformants of 10<sup>4</sup> cfu/ $\mu$ g, the efficiency desired for reliable and repeatable metabolic engineering. However, repeated efforts by several researchers in our laboratory were not able to reproduce the findings, which led to the suspicion that there might be a difference between the strains used by Pyne et al. (2013) and us. A request to obtain the wild type used by Pyne et al. (2013) was denied, and

I was being assured that the strain used in their study was *C. pasteurianum* ATCC 6013 which can be obtained from the American Type Culture Collection. Transformation of several *C. pasteurianum* wild type strains, including ATCC 6013, purchased after failing to obtain Pyne's strain, was tried without success (Section 3.2.1.1). All findings here strengthened our hypothesis that variants between the strains are to be held accountable for the differences of transformation efficiency observed between our experiments and [Pyne et al. \(2013\)](#).

### 3.3.2 Hypertransformable Sub-population

These findings led to formulate a hypothesis based on the small set of transformants found: Transformable clones might exist in small numbers within the wild type population. This notion was tested by curing one of the rare transformants of its acquired plasmid, and then testing the frequency of DNA transfer in the isolated plasmid-free strain obtained. The first plasmid-cured strain, designated *C. pasteurianum*-H1, was found to transform at frequencies of up to  $10^5$  transformants per  $\mu\text{g}$  DNA. Apart from the increased transformation efficiency *C. pasteurianum*-H1 did not exhibit any difference in phenotype and growth and solvent profile were shown to be unchanged. This strain was used for subsequent metabolic engineering exemplified in Chapter 4 and in [Schwarz et al. \(2017\)](#). High throughput genome sequencing of its genome revealed that frame-shift mutations had arisen in two independent genes causing possible detrimental effects on both gene functions. The one (CLPA\_c13710, *ablB*) encoded a predicted  $\beta$ -lysine N<sub>6</sub>-acetyltransferase while the other (CLPA\_c33080, *resE9*) encoded a histidine kinase likely to be involved in quorum sensing. Neither gene appeared to encode a protein which played any direct role in restriction/ modification systems.

N-acetyltransferases are known to be involved in antibiotic resistance and biosynthesis of cross-links in branched cell walls ([Vetting et al., 2005](#)). *C. pasteurianum*  $\beta$ -lysine N-acetyltransferase CLPA\_c13710 *ablB* shows high similarity ( $e= 3 \times 10^{-47}$ ) to the  $\beta$ -lysine N<sub>6</sub>-acetyltransferase (*ablB*) of the anaerobic Archaea *Methanosarcina mazei*, found to be regulated by extracellular salt levels and involved in salt stress ([Pfluger et al., 2003](#)). A mutation of such an enzyme might lead to a weakened cell-wall due to missing cross-links similar to cell-wall weakening agents used when inducing competence (Table A.1) and thus increasing transformation efficiency. The histidine kinase CLPA\_c33080, *resE9* possesses a N-terminal transmembrane stretching domain which links it to the cell-wall and presumptively allows it to sense external stimuli. Histi-



dine kinases in general are known to be involved in sensing the cellular environment and in turn regulating intracellular processes by phosphorylating a histidine of a response regulator protein (West and Stock, 2001). Such regulations can involve *e.g.* basic housekeeping and expression of toxins and other proteins. ResE seems, however, not to be part of a two component system as no gene encoding a cognate response regulator was located in the immediate vicinity. Interestingly, the gene adjacent to the histidine kinase annotated as CLPA\_c33090 encodes a protein with homology to an AgrA quorum sensing protein. AgrA processes and secretes the AgrD signal molecule which can be found as an unannotated ORF immediately downstream of *agrA*. The cluster *agrB-agrD-resE* is conserved in many clostridial species (Figure A.1) and was first noted by Cooksley et al. (2010) in *C. botulinum*. It is possible that this quorum sensing system regulates a process involved with DNA uptake similar to the *com* operon in *Streptococcus pneumoniae*. In this organism the histidine kinase (*comD*) binds competence-stimulating peptides which are released by the population as part of a quorum sensing system (Spoering and Gilmore, 2006). When a cell density of  $10^7$  cells per ml is reached the histidine kinase phosphorylates its cognate response regulator (*comE*), which in turn activates transcription of several genes. It is, furthermore, assumed that this activates the sigma factor ComX, which activates transcription of a competence regulon with proteins involved in binding, processing, uptake, and recombination of DNA (Steinmoen et al., 2002). It remains to be determined if the *C. pasteurianum agrBD-resE* serves a similar function.

Even though at this point it cannot conclusively be determined if one or both of the defective genes are responsible for the susceptibility of *C. pasteurianum*-H1 to electroporation, both enzymes seem to be involved in environmental sensing and response regulation and could hypothetically play a role. Further research could help elucidate this question and advance understanding of DNA uptake in general or at least specifically in *C. pasteurianum*.

To investigate if the screening for hypertransformable clones can be reproduced, a more methodical approach was undertaken with six individual repetitions. Out of the six screens two strains arose with increased transformation efficiency: *C. pasteurianum*-H3 and *C. pasteurianum*-H4. Next generation sequencing of the two mutants revealed both to carry the same SNP variant, introducing a change on amino acid level (F1102V) in gene CLPA\_c30550 encoding a hypothetical protein. Both strains *C. pasteurianum*-H3 and -H4 were shown to be virtually unchanged in their ability to produce solvents from glycerol compared to wild type *C. pasteurianum*. The SNP was further shown to be directly responsible for the observed hypertransformable phenotype. Correc-

tion of the SNP to wild type restored DNA transfer efficiency from high to low in *C. pasteurianum*-H3.

Gene CLPA\_c30550 lies in a cluster of four genes (CLPA\_c30550- CLPA\_c30580) which is conserved in *C. perfringens*. All of the genes are annotated as hypothetical proteins and are separately conserved in other bacteria but not as a cluster. Gene CLPA\_c30570 has a FtsY domain involved in intracellular transport and CLPA\_c30580 possesses a conserved domain of unknown function. According to the conserved domain database and Phyre2 modelling, the protein encoded by CLPA\_c30550 has nucleolytic activity and resembles structural maintenance of chromosome (SMC) proteins indicating it to be involved in the segregation of chromosomes during cell division. SMC proteins possess conserved domains at the C- and N-terminus called Walker A (consensus sequence: [AG]XXXXGK[ST]) and B (consensus sequence [h= hydrophobe]: [RK]XXX-XGXXXXLhhhhD), respectively (Walker et al., 1982). Sequences resembling these motifs can be found in CLPA\_c30550 at amino acid position 37 (Walker A) and amino acid position 968 (Walker B) and the predicted tertiary protein structure resembles the hinged architecture of SMCs (Hirano, 2002).

In *B. subtilis* SMC proteins are involved in chromosome condensation in the later stage of chromosome segregation (Graumann, 2000; Lindow et al., 2002). It is unclear how a point mutation of such an enzyme could lead to increased transformation efficiency in *C. pasteurianum*. *In silico* analysis of the target protein is not conclusive and gives only indications of putative function, but the protein is likely interacting with DNA. A mutation of such an enzyme can hypothetically increase or decrease the protein's natural ability of one or more functions. In case of involvement of the protein in chromosome segregation, higher efficiency could enhance the ability to segregate multi-copy plasmids. A similar mechanism was observed by Moriya et al. (1998) reporting a strain of *B. subtilis* with a disrupted SMC protein being able to grow at 23 °C but not at 37 °C, indicating that chromosome segregation was SMC independent in slow cell-growth while with faster growth chromosome segregation relies on SMC activity due to faster segregation of chromosomes. If the SMC in *C. pasteurianum* is naturally inefficient this could lead to cell-death by the inability of plasmid segregation in a transformant at 37 °C in the presence of antibiotic selection. However, CLPA\_c30550 might have a different, yet unknown function similar to restriction enzymes. The SNP could in this case inactivate or reduce such activity and allow foreign DNA to be taken up.

### 3.4 Conclusion

This study laid the basis for genetic work with *C. pasteurianum* by developing DNA transformation. A strategy to find hypertransformable clones in a wild type population was presented and square wave electroporation was shown to be a viable way for bacterial researchers to increase transformation efficiency. The hypertransformable screening procedure yielded three hypertransformable clones of *C. pasteurianum*. Contrary to expectations the mutations in these strains were not found in restriction/ modification or CRISPR systems but in seemingly unrelated enzymes of the organism. Further studies of these enzymes might shed more light on the mechanisms of bacterial competence in general and electroporation in particular, and could yield solutions to increase transformation efficiency in other organisms.



# Chapter 4

## Metabolic Engineering of *C. pasteurianum* for Improved Solvent Production

Some humans would do anything to see if it was possible to do it. If you put a large switch in some cave somewhere, with a sign on it saying 'End-of-the-World Switch. PLEASE DO NOT TOUCH', the paint wouldn't even have time to dry.

---

Terry Pratchett, *Thief of Time*

### 4.1 Introduction

Until recently there was a lack of available genetic tools to alter molecular pathways in *Clostridium pasteurianum* mainly due to difficulties in genetic accessibility of the organism. The ability to transform the organism demonstrated in Chapter 3 allowed the development of molecular tools through exemplification of the clostridial roadmap proposed by [Minton et al. \(2016\)](#). For this purpose allelic exchange based on the *pyrE* locus ([Ng et al., 2013](#)) and Allele Coupled Exchange (ACE) ([Heap et al., 2012](#)) as well as the ClosTron and in a later phase transposon mutagenesis (Chapter 5) have to be developed to build a complete tool set for genetic modification of this organism. Recently, [Rotta \(2014\)](#) employed the ClosTron to disrupt the *spo0A* gene (encoding the sporulation mas-

---

This chapter is published as [Schwarz et al. \(2017\)](#)

ter regulator) and *pyrE*, the orotate phosphoribosyltransferase. Schwarz et al. (2017) then went on to truncate *pyrE* by ACE and used the deletion background to employ allelic exchange to delete *spo0A* in frame and complement the gene in the *pyrE* locus. The *spo0A* gene was chosen to exemplify the techniques as a deletion mutant exhibits a clear sporulation deficient phenotype similar to *C. acetobutylicum* (Harris et al., 2002) and presented in *C. pasteurianum* (Rotta, 2014). The sporulation phenotype of these strains was analysed by Schwarz et al. (2017) and in this study the phenotype of these mutants in regard to solvent production in different media is assessed.

With the Spo0A proof-of-principle for a *C. pasteurianum* tool set, the next step is to delete central energy pathways in *C. pasteurianum* to increase solvent production. As shown in Section 1.1.1.6 there are only a few recent studies that rationally alter *C. pasteurianum* on a genetic level with the aim of improving product levels. Genetic engineering was done by downregulating the expression of *hydA* by anti-sense RNA (Pyne et al., 2015) and disruption of *dhaT* with a targetron system (Pyne et al., 2016c). For this study it was decided to target genes involved in intracellular redox balance. The chosen targets were the *rex* redox regulator, involved in sensing NADH/NAD<sup>+</sup> ratios and accordingly repressing or de-repressing genes (Brekasis and Paget, 2003); *hydA*, which acts in regenerating reduction equivalents by generating hydrogen; *dhaBCE*, the first step in 1,3-propanediol (1,3-PDO) biosynthesis which is a mean for *C. pasteurianum* to regenerate reduction equivalents for cell growth (Ruch et al., 1974). By deleting these genes it should be possible to free oxidative power for the production of butanol and the genes shall be introduced in detail below.

#### 4.1.1 *spo0A* - Sporulation Master Regulator

Spo0A is known as the central transcriptional regulatory protein that initiates the sporulation cascade upon integration of environmental and physiological signals like nutrient depletion, cell density, tricarboxylic acid cycle, DNA synthesis and DNA damage in *Bacillus subtilis* (Stragier and Losick, 1996). In *B. subtilis* the relay protein Spo0F is phosphorylated by at least three different histidine kinases and Spo0F transfers the phosphate to Spo0B which finally phosphorylates Spo0A. Spo0A then regulates self-expression and expression of other phosphatase genes but most importantly Spo0A~P governs transcription of genes that initiate sporulation (Stragier and Losick, 1996). Even though *spo0A* exists in clostridia, no genes resembling *spo0F* or *spo0B* from *B. subtilis* could be found (Dürre and Hollergschwandner, 2004). In these organisms Spo0A is activated directly by histidine kinases (Steiner et al., 2011). Inactiva-

tion of Spo0A in *C. beijerinckii* (Heap et al., 2010), *C. perfringens* (Huang et al., 2004) and *C. acetobutylicum* showed a sporulation deficient phenotype and *spo0A* over-expression led to accelerated sporulation (Harris et al., 2002). Rotta (2014) showed that inactivation of *spo0A* in *C. pasteurianum* led to a sporulation deficient phenotype. All of this data suggests that Spo0A is the master regulator of sporulation in clostridia too. Unlike *Bacillus*, however, clostridia do not sporulate in response to the above mentioned stimuli but rather to escape acidic conditions generated by acidogenesis (Dürre et al., 2002). Furthermore, it was reported that Spo0A controls expression of solvent pathways in *C. acetobutylicum* and *C. beijerinckii* (Ravagnani et al., 2000). Solvent production was nearly eliminated in a *C. acetobutylicum spo0A* disruption strain and acid generation was increased (Harris et al., 2002). On the other hand, when *spo0A* was over-expressed the strain showed a 20 % increase in solvent levels when compared to the control (Harris et al., 2002). Similarly, in a *C. beijerinckii* strain with *spo0A* disruption butanol and acetone formation was abolished whereas ethanol and acid production was unchanged (Ravagnani et al., 2000). Recently, Sandoval et al. (2015) found that solvent production is not dependent on Spo0A in *C. pasteurianum*, and contrary to *C. acetobutylicum*, the solvent levels were significantly increased when *spo0A* was deleted. However, sporulation was abolished in this mutant as expected.

#### 4.1.2 *rex* - Redox Responsive Repressor

The Rex regulator was found to respond to intracellular redox potential by allosterically binding NADH or NAD<sup>+</sup> and thus de-repressing or repressing gene transcription, respectively (Brekasis and Paget, 2003). The dimeric Rex binds to DNA via the winged helix domains on both subunits. NAD<sup>+</sup> binding to Rex does not change the conformation of the enzyme complex and the protein stays bound to DNA. However, increasing NADH levels lead to dissociation of NAD<sup>+</sup> and association of NADH. This in turn leads to a rotation of the subunits and release of the enzyme complex from the Rex operator site allowing transcription of the downstream gene. (Sickmier et al., 2005; McLaughlin et al., 2010) Rex homologues were found in the sequenced genomes of most Gram-positive bacteria but are absent in Gram-negatives (Brekasis and Paget, 2003).

The binding site of Rex in different phyla was investigated by Ravcheev et al. (2012) and a co-evolution of the Rex protein and its DNA recognition motif was found. In Clostridiaceae the modified Rex-binding motif has a consensus sequence 5'-TTGTAAANNNTTAACAA. This consensus sequences was used to identify Rex regulons in eleven diverse clostridial species. It was found that

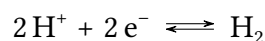
Rex mainly targets enzymes involved in fermentation that use NADH or other reducing equivalents. This was confirmed in *C. kluyveri* where it was shown via electrophoretic mobility shift assay that Rex binds to promoter regions of *rnfG*, *hydA*, *bdh2*, *thlA* and its own promoter, suggesting self-regulation (Hu et al., 2016). In *C. acetobutylicum* the 17 genes in 7 operons regulated are *adhE2*, *ldh*, *crt-bcd-etfBA-hbd*, *thlA*, *asrTABC*, *ptb-buk* and *nadABC*. The genes are involved in fermentation, NAD<sup>+</sup> biosynthesis and sulfite reduction and all of them were upregulated in a *C. acetobutylicum*- $\Delta$ *rex* mutant with *adhE2* being expressed  $\geq 160$  fold confirming negative regulation of Rex on these genes. (Zhang et al., 2014) Furthermore, inactivation of Rex in *C. acetobutylicum* led to decreased acid and increased alcohol titres with a substantial decrease of the by-product acetone (Wietzke and Bahl, 2012).

Rex controls the expression of cytochrome *bd* terminal oxidase in *Streptomyces coelicolor* (Brekasis and Paget, 2003) which in *E. coli* is switched on under anaerobic conditions and has a high affinity to oxygen (Anraku and Gennis, 1987). In *B. subtilis* the NADH dehydrogenase *Ndh* is controlled by Rex allowing balancing of NADH/NAD<sup>+</sup> ratios in the cell (Gyan et al., 2006). In *Staphylococcus aureus*, Rex regulates genes in the fermentative pathway. A knock-out mutant shows increased production of metabolites produced by the regulated genes. (Pagels et al., 2010)

These findings make Rex an interesting target for metabolic pathway engineering in *C. pasteurianum* as it can be anticipated that Rex has a similar effect in this strain as it has in *C. acetobutylicum* with higher solvent titres being a major aim of this study.

### 4.1.3 *hyd* - Iron Only Hydrogenase

Hydrogenases catalyse the reversible reaction of two hydrogen atoms to molecular hydrogen according to the reaction:



There are three phylogenetically distinct classes of hydrogenases. The best studied class is the [NiFe]-hydrogenases in Bacteria and Archaea. As the name suggests they have a nickel and an iron atom at their active site. The second group, [FeFe]-hydrogenases with a two iron centre, is found in eukaryotes and prokaryotes. And lastly the H<sub>2</sub> forming methylenetetrahydromethanopterin dehydrogenases, which rely on an iron containing cofactor. (Vignais and Colbeau, 2004) The latter are also called [Fe]-hydrogenases or [Fe]-only-hydrogenases and are only found in methanogenic Archaea (Calusinska et al., 2010)



and thus here the focus will be laid on the other two classes.

[NiFe]- and [FeFe]-hydrogenases have an iron-sulfur cluster that is bound by cysteinyl sulfur atoms to the protein. The iron in both types of enzymes is also bound to carbon monoxide and cyanide. In [NiFe]-hydrogenases the most likely site of hydrogen activation is the nickel atom whereas in [FeFe]-hydrogenases it is at one of the iron atoms which can bind a ligand and, depending on conditions, is blocked by carbon monoxide. (Heinekey, 2009) This observation fits with reports of inhibition of H<sub>2</sub> generation by carbon monoxide (Hoberman and Rittenberg, 1943) and strengthens the hypothesis of this location being the active site of the enzyme.

Among the best studied [FeFe]-hydrogenases are the ones of *Desulfovibrio desulfuricans* (Nicolet et al., 1999) and *C. pasteurianum* itself (Cornish et al., 2011). The *C. pasteurianum* hydrogenase has been crystallographically solved to a resolution of 1.39 Å (Pandey et al., 2008) and the proton transfer in the enzyme is largely understood (Cornish et al., 2016, 2011).

In *Bacteria* hydrogenases regulate the H<sup>+</sup> levels in the cells and thus are actively involved in redox regulation and in maintaining the proton motive force. Clostridia rely foremost on [FeFe]-hydrogenases but a few species possess [NiFe]-hydrogenases (Calusinska et al., 2010). In *C. pasteurianum* 4 hydrogenases were found in the genome (CLPA\_c00280, CLPA\_c07060-70, CLPA\_c33960 and CLPA\_c37830) (Poehlein et al., 2015). With CLPA\_c07060 and CLPA\_c07070 *C. pasteurianum* seems to possess a rare [NiFe]-hydrogenase with the genes encoding the small and large subunit, respectively. The other genes encode iron-only hydrogenases with CLPA\_c37830 designated as the fermentative hydrogenase *hydA*. BLAST searches and the Pyne et al. (2014b) annotation on the other hand identify the respective gene of CLPA\_c00280 as *hydA* indicating that the hydrogenases found are very similar and it remains to be seen which one is involved in dark fermentation in *C. pasteurianum* or if there is an interplay between the different hydrogenases.

Previous work for improvement of solvent generation in clostridia by manipulation of hydrogenases involves carbon monoxide inhibition in *C. pasteurianum* (Dabrock et al., 1992) and *C. acetobutylicum* (Datta and Zeikus, 1985; Kim et al., 1984), *hydA* knock-down via anti-sense RNA in *C. pasteurianum* (Pyne et al., 2015), attempted disruption of *C. acetobutylicum hydA* (Cooksley et al., 2012), deletion of the hydrogenase maturase gene *hydG* in *C. thermocellum* (Biswas et al., 2015) and increasing hydrogen formation was assayed via over-expression of *hydA* (Klein et al., 2010). Carbon monoxide inhibition led to increased solvent titres and decreased total acid concentrations when *C. pasteurianum* was grown on glucose (Table 4.6, Dabrock et al. (1992)). In

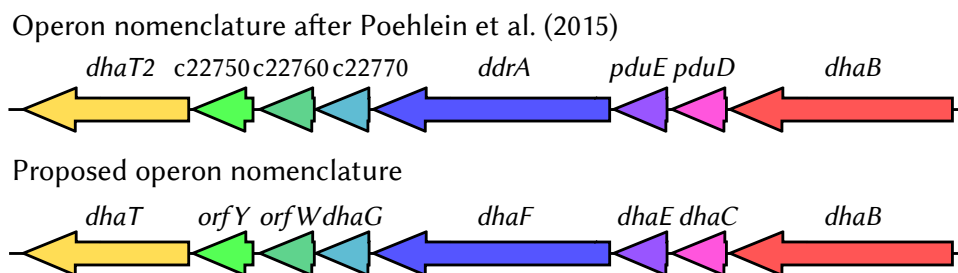
*C. acetobutylicum* solventogenesis started much earlier in carbon monoxide inhibited fermentations as in normal two phasic growth and acetone production was abolished whereas ethanol stayed largely the same and butanol production increased. Acids, on the other hand, were produced much less. (Datta and Zeikus, 1985) Another carbon monoxide inhibition experiment showed similar results: acetone production was reduced but not abolished, ethanol and butanol production were increased (Kim et al., 1984). The knock-down experiment of *C. pasteurianum* hydrogenase with anti-sense RNA showed a slightly different outcome with increased solvent-production of ethanol and butanol but also an increase in organic acid production. The levels of the by-product 1,3-PDO was reduced in this mutant. (Pyne et al., 2015) Over-expression of hydrogenase in *C. acetobutylicum* DSM 729 did not show any influence on fermentation pattern and it was concluded that intracellular hydrogenase concentration does not limit hydrogen formation in this strain (Klein et al., 2010).

The most promising way of shuttling NADH away from hydrogen formation to production of solvents is the complete inactivation of intracellular hydrogenase by deletion of the gene or genes. Cooksley et al. (2012) tried to disrupt the *C. acetobutylicum* hydrogenase via ClosTron integration but no mutants could be obtained leading them to conclude that under the conditions employed inactivation of *hydA* was not possible. Biswas et al. (2015) took a slightly different approach in *C. thermocellum* to inactivate all three [FeFe]-hydrogenases at once. They targeted the maturase gene *hydG* and, to abolish all hydrogenase activity, deleted the [NiFe]-hydrogenase as well. The resulting strain did not produce any detectable hydrogen but 63 % more ethanol and showed a dramatic decrease in acid production.

To my knowledge there is no report of a single [Fe]-hydrogenase deletion in *Clostridium* to date but data of inhibition or inactivation of all hydrogenases in an organism or knock-down of a single hydrogenase are promising in terms of increasing solvent yield and decreasing by-product formation and acid formation in particular. In this study the fermentation pattern of a *C. pasteurianum* *hyd* deletion mutant is presented.

#### 4.1.4 *dhaBCE* - Glycerol Dehydratase

Glycerol can be used by micro-organisms in two pathways. The oxidative pathway leads from glycerol to pyruvate to enter the fermentative pathway. The second pathway is reductive and goes from glycerol via 3-hydroxypropionaldehyde (3-HPA) to 1,3-PDO. (Biebl et al., 1999; Malaviya et al., 2012) In the reduction of 3-HPA to 1,3-PDO one NADH is used and a NAD<sup>+</sup> regenerated



**Figure 4.1** – Proposition of a nomenclature of the reductive part of the *dha* regulon in *C. pasteurianum* DSM 525 genome (Poehlein et al., 2015) based on publications by Macis et al. (1998), Seifert et al. (2001) and Jiang et al. (2016) and manual BLAST searches.

which is crucial for microbial growth as glycerol is more reduced than biomass (Ruch et al., 1974). The conversion of glycerol is done by enzymes encoded by the *dha* regulon (Forage and Foster, 1982; Seyfried et al., 1996).

In the oxidative branch glycerol dehydrogenase (*dhaD*) converts glycerol to dihydroxyacetone using  $\text{NAD}^+$  and dihydroxyacetone is phosphorylated by dihydroxyacetone kinase (*dhaK*) in a coupled reaction with pyruvate kinase to dihydroxyacetone-phosphate which enters glycolysis and fermentation. The reductive branch of the *dha* regulon starts with the, for most organisms, vitamin  $\text{B}_{12}$  dependent multi-subunit enzyme glycerol dehydratase, encoded by *dhaBCE*. DhaBCE is built as a complex  $\alpha_2\beta_2\gamma_2$  (Knietzsch et al., 2003) which converts glycerol to 3-HPA. This conversion inactivates the enzyme by cleavage of the coenzyme resulting in a tight bond between the coenzyme and glycerol dehydratase (Toraya, 2000). It has been shown that this effect can be reversed in several organisms by reactivation with a two subunit protein complex encoded by *dhaFG* (Mori et al., 1997; Seifert et al., 2001). Other organisms express coenzyme  $\text{B}_{12}$  independent glycerol dehydratases (Hartmanis and Stadtman, 1986; Raynaud et al., 2003). In the last step of glycerol conversion 1,3-PDO is formed by reduction of 3-HPA by the 1,3-propanediol dehydrogenase (*dhaT*) and thus regeneration of reducing equivalents (González-Pajuelo et al., 2006; Biebl et al., 1999; Luers et al., 1997).

In *C. pasteurianum* DSM 525 glycerol dehydratase is encoded by the first three genes in the operon *dhaB-pduD-pduE-ddrA-CLPA\_c22770-CLPA\_c22760-CLPA\_c22750-dhaT2* (Poehlein et al., 2015). In Figure 4.1 a new nomenclature for the 1,3-PDO forming operon is proposed based on literature (Seifert et al., 2001; Macis et al., 1998; Jiang et al., 2016) and BLAST searches due to the gene names being firmly established in other species. For the same reason and confirmation by further BLAST searches with the protein sequences of *dhaT1* and *dhaT2* (Poehlein et al., 2015) it was revealed *dhaT2* to form the 1,3-propanediol

dehydrogenase and *dhaT1* having a different if similar function that is not yet elucidated. For the purpose of this study this naming convention shall be used to fit in with established literature.

According to [Johnson and Rehm \(2016\)](#), 1,3-PDO production in *C. pasteurianum* is important for biomass production as outlined above. In their study this was demonstrated with high 1,3-PDO production rates in exponential growth when most biomass was produced and thus most electrons and NADH freed. Nonetheless, the pathway to 1,3-PDO is an obvious target for genetic engineering of *C. pasteurianum* as 1,3-PDO production directly competes with butanol production for the substrate carbon. There are several strategies to eliminate 1,3-PDO production. Either the production can be stopped in the first step from glycerol with a deletion of *dhaBCE* or the second step can be inactivated with by a knock-out of *dhaT*. The latter strategy could lead to accumulation of the intermediate 3-HPA which is reported to be toxic ([Barbirato et al., 1996](#); [Maervoet et al., 2014](#)). Surprisingly, however, a recent report of a *dhaT* ClosTron disruption in *C. pasteurianum* showed very reduced levels of 1,3-PDO but no buildup of 3-HPA or derivatives thereof (acrolein, 3-hydroxypropionate) ([Pyne et al., 2016c](#)). To prevent 3-HPA inhibition of microbial growth for this study, it was decided to inactivate the first step in the pathway by deleting all three subunits of the glycerol dehydratase (*dhaBCE*) and analyse fermentation patterns of the strain.

#### 4.1.5 Aim of This Study

The aim of this study was to analyse fermentative phenotypes of in-frame deletion mutants of *C. pasteurianum* developed by [Schwarz et al. \(2017\)](#) and completed here. Spo0A was chosen to be deleted due to a clear sporulation deficient phenotype and its reported influence on fermentative pathways. Furthermore, three genes actively involved in balancing internal levels of *C. pasteurianum* redox equivalents were chosen as targets for in frame deletions. Their ability to produce solvents and butanol in particular was analysed in batch fermentation.

## 4.2 Results

### 4.2.1 Mutant Development

In this work several clean deletion mutants were used which were created by [Schwarz et al. \(2017\)](#). As these mutants were constructed in a *pyrE* trunca-

tion background as described in Minton et al. (2016) the *pyrE* locus had to be restored in this study. Furthermore, complementations of the gene with the native promoter were used. Schwarz et al. (2017) created the complementations of  $\Delta spo0A$ ,  $\Delta hyd$  and  $\Delta dhaBCE$  and the final  $\Delta rex$  complementation strain was produced in this study. The cloning procedures of plasmids used to integrate DNA into the host is outlined in Chapter 2 and the integration procedure is outlined below.

#### 4.2.1.1 *pyrE* Repair

The truncated *pyrE* gene in all allelic exchange knock-out mutants produced by Katrin Schwarz and Kamila Derecka were repaired using pMTL-AGH12 which essentially carries a wild type allele of the *pyrE* locus between the ACE homology arms (construction see Section 2.3.2). For this purpose, strains *C. pasteurianum*  $\Delta rex \Delta pyrE$  6-49, *C. pasteurianum*  $\Delta hyd \Delta pyrE$  47, *C. pasteurianum*  $\Delta dhaBCE \Delta pyrE$  6 and *C. pasteurianum*  $\Delta spo0A \Delta pyrE$  4-44 were transformed with the methylated plasmid and 24 faster growing colonies having putatively undergone single crossover were restreaked on RCM+Tm twice to single colonies. Colonies were then restreaked on both CBM and RCM+TM plates until growth could only be observed on CBM plates and not on RCM+TM plates, indicating insertion of wild type allele and loss of the plasmid from the organism. To ascertain loss of the plasmid, successfully screened colonies were restreaked thrice more on CBM before the strain was grown in 2xYTG overnight for genomic DNA extraction and storage with 15 % DMSO at -80 °C. The genomic DNA of putatively *pyrE* repaired strains was extracted and a PCR was performed with DreamTaq Green Mastermix and primers SHA\_F/LHA\_R to check for successful insertion and primers 85XXX-RF/8X1XX-LR to check for absence of the plasmid. Successful clones exhibit a 684 bp band and no band from the plasmid whereas the *pyrE* truncation strain shows a 718 bp band at the *pyrE* locus and strains still possessing the plasmid show a 322 bp band. The *pyrE* truncated strain shows a longer fragment than the repair strain as it features the *lacZ* $\alpha$  downstream of the truncation. Bands of successful candidates not showing a plasmid band were purified and sent for sequencing. Two successful clones of each mutant were stored and used for further genotypic analysis (see Table 2.1 for details of the clones).

#### 4.2.1.2 *rex*-Complementation

To check if the observed phenotype of *C. pasteurianum*  $\Delta rex$  was due to the deletion of the gene or if it had other underlying reasons a *rex* complementa-

tion strain was produced featuring *rex* under its own promoter downstream of the repaired *pyrE* locus. The plasmid pMTL-KS12-*rexcomp* carrying *rex* with its native promoter between the ACE homology arms (for construction see Section 2.3.5) was transformed into *C. pasteurianum*  $\Delta_{rex} \Delta_{pyrE}$  6-49 and 24 faster growing colonies with a putative single crossover were restreaked twice to single colony on RCM+Tm. Clones were then restreaked on each CBM and RCM+Tm until only growth on CBM could be observed indicating that a double crossover was successful and suggesting plasmid loss. The clones were restreaked another three times on CBM to ascertain plasmid loss before they were grown in 2xYTG over-night for DNA extraction and storage with 15 % DMSO at -80 °C. Genomic DNA was extracted and putative complementation strains were tested by PCR with DreamTaq Green mastermix with primers SHA\_F/LHA\_R for the insertion of the complementation and primers 85XXX\_RF/8X1XX\_LR for absence of the plasmid. Expected band sizes for successful mutants were 1564 bp for the genome locus and absence of a band with the plasmid based primers. Double crossovers having reverted back to wild type show a band of 718 bp and if the plasmid was present a band of 322 bp. Bands showing the expected size were extracted and sent for Sanger sequencing of the insert. Two successful complementation clones were stored under the designations CRG5524 and CRG5525 for further phenotypic analysis (compare Table 2.1).

#### 4.2.2 Spo0A Deletion Phenotype

To analyse the fermentation phenotype of a *spo0A* knock-out the deletion strain *C. pasteurianum*  $\Delta_{spo0A}$ , the complementation *C. pasteurianum*  $\Omega_{spo0A}$  and the wild type were grown in two different media. A defined medium (Biebl medium) with yeast extract to aid growth (1 g/l) and a complex medium, CGM. In both media the fermentations reached an endpoint after 24 h after which only acetate concentration changed when grown on Biebl medium due to re-assimilation.

When grown on complex medium (CGM), all three strains show a similar growth rate (Table 4.1) but the pH in *C. pasteurianum*  $\Delta_{spo0A}$  is lower after 24 h and even though it does increase after 48 h, as does the pH in the other strains, it stays more acidic all through the experiment. This coincides with a significantly higher glycerol consumption for the deletion strain ( $284.7 \pm 4.0$  mM) as opposed to the wild type ( $236.4 \pm 3.7$  mM) as determined by t-test with conditions;  $t(4) = 7.4$ ,  $p = 0.002$ . Lactate was produced with the same efficiency for 12 h and increased over the next 12 h in *C. pasteurianum*  $\Delta_{spo0A}$  to reach a signifi-





cantly higher level after 24 h ( $14.9 \pm 0.2$  mM) than in the wild type ( $7.6 \pm 0.5$  mM) (t-test conditions,  $t(4) = 10.5$ ,  $p = 0.001$ ). Interestingly the mutant produces much more acetate in 12 h but re-assimilates the acid after that whereas the wild type and the complementation strain both produce acetate until 12 h after which the concentration stays the same over the rest of the fermentation. In contrast, butyrate production is similar in all strains until 12 h after which the wild type and the complementation re-assimilate butyrate significantly. The deletion strain is only able to re-assimilate a small portion of the produced acid.

The *spo0A* mutation does not seem to influence 1,3-PDO production and all three tested strains reach a similar final concentration after 12 h of fermentation. Ethanol production on the other hand is significantly decreased in the deletion mutant with a concentration of  $17.4 \pm 0.4$  mM after 24 h and in the wild type of  $22.4 \pm 0.4$  mM (t-test conditions,  $t(4) = 8.0$ ,  $p = 0.001$ ). Butanol concentration after 24 h is significantly increased in *C. pasteurianum*  $\Delta spo0A$  ( $84.7 \pm 2.1$  mM) as opposed to the wild type ( $71.8 \pm 1.1$  mM) (t-test conditions,  $t(4) = 4.5$ ,  $p = 0.011$ ).

Complementation of the *spo0A* mutation restores product levels back to those of the wild type, with only minor differences in acid formation. Glycerol consumption and solvent production is nearly perfectly compensated. Butyrate re-assimilation is reinstated. Only lactate and acetate levels differ slightly with less lactate and more acetate produced in the complementation strain.

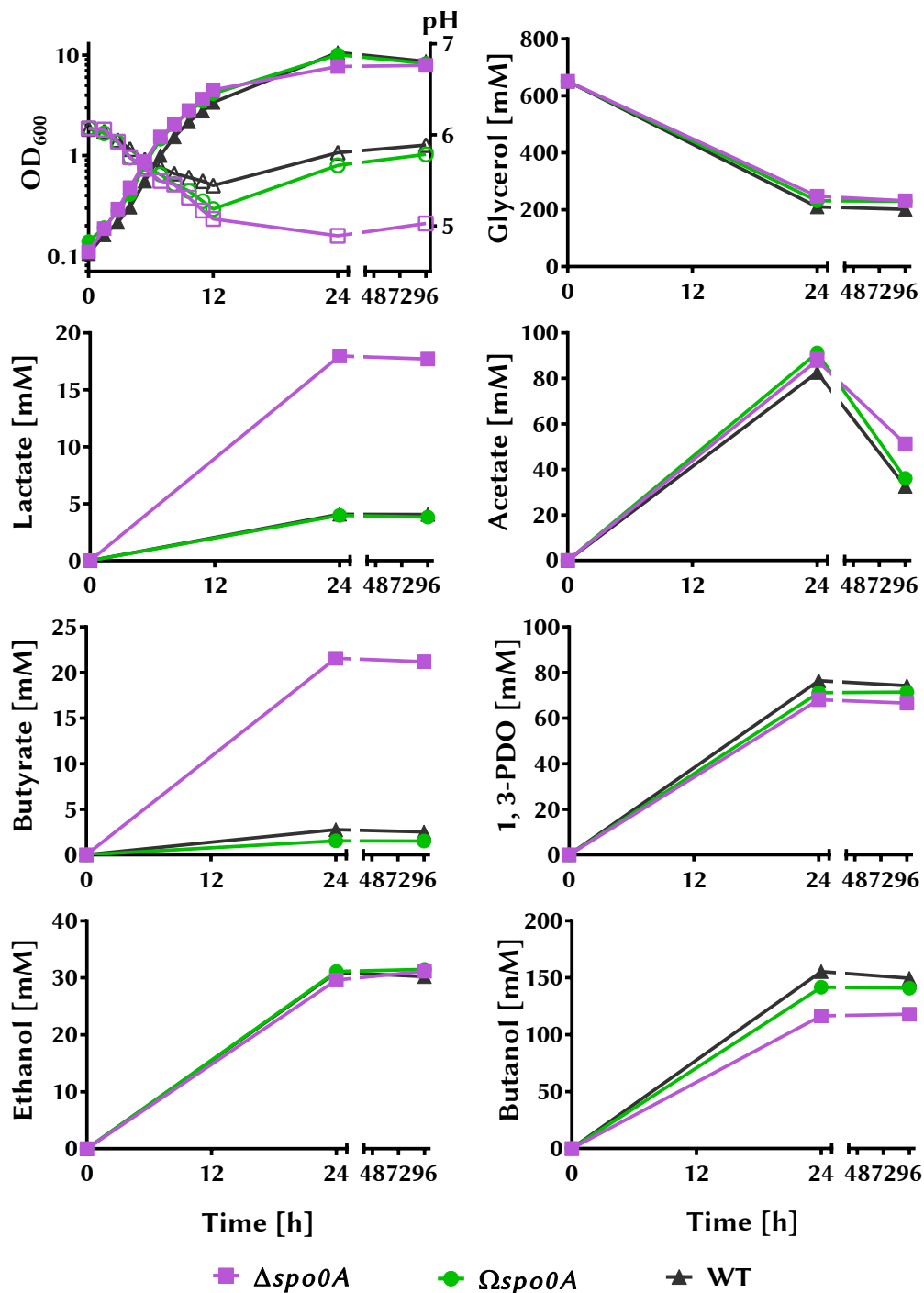
When the *spo0A* deletion strain is grown in Biebl medium similar differences in growth as in CGM can be observed. In this medium the growth rate in all strains is similar (Table 4.1). The deletion strain grows to a lower maximal  $OD_{600}$  and pH decreases further than in the wild type. In Biebl medium the pH of the mutant recovers marginally. Glycerol uptake, acetate production and re-assimilation and ethanol production are similar in all strains, whereas lactate and butyrate production are increased in the knock-out mutant. Butyrate is not visibly re-assimilated in this fermentation by any of the strains tested. Finally, butanol and 1,3-PDO levels are slightly lower in the deletion mutant in this medium. The complementation strain seems to over-compensate these levels suggesting a higher influence of Spo0A over synthesis of these products.

## 4.2.3 Rex Regulator

### 4.2.3.1 *In silico* Analysis

**Rex Protein** Rex in *C. pasteurianum* is encoded by gene CLPA\_c28640 expressing a 213 amino acid protein showing high similarity to known and reported Rex proteins in other species. The Rex proteins of *C. pasteurianum* and





**Figure 4.3** – Batch cultures were grown in serum bottles of the *spo0A* deletion strain *C. pasteurianum*  $\Delta spo0A$ , its complementation *C. pasteurianum*  $\Omega spo0A$  and the wild type (WT) strain. Open symbols are used for pH. Strains were grown in three biological replicates in Biebl medium supplemented with 1 g/l yeast extract and 6 % glycerol. Data is given as mean with error-bars indicating standard error.

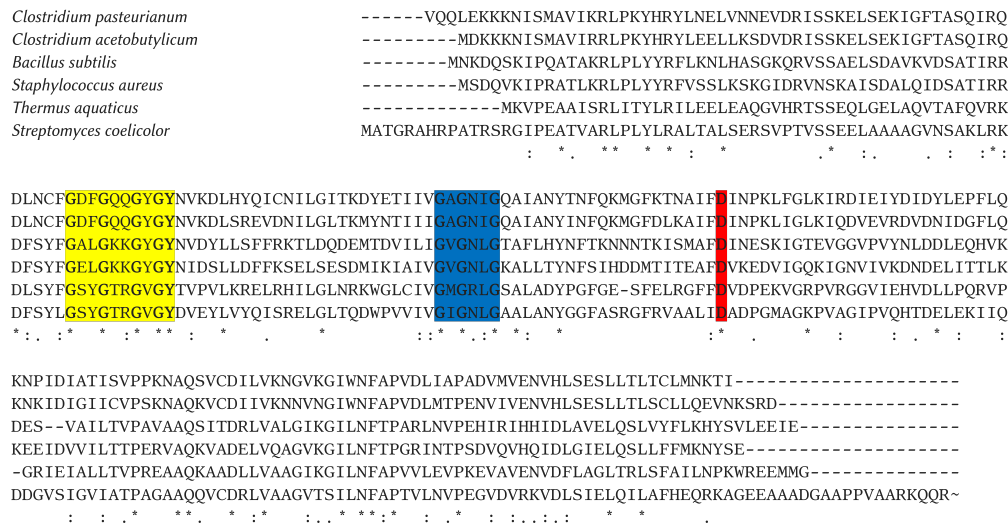
**Table 4.1** – Growth characteristics and product yields of *spo0A* deletion and complementation strain. Data of *spo0A* serum bottle fermentations in biological triplicates at 24 h time point. Data given as mean±SEM.

Carbon source Medium Strain	60 g/l glycerol					
	Biebl plus 1 g/l yeast extract			CGM		
	$\Delta spo0A$	$\Omega spo0A$	WT	$\Delta spo0A$	$\Omega spo0A$	WT
Growth characteristics						
Specific Growth Rate [ $h^{-1}$ ]	0.37±0.00	0.37±0.00	0.35±0.01	0.50±0.01	0.47±0.01	0.45±0.01
Doubling Time [min]	113±0.7	113±0.3	120±2.2	83±1.8	88±1.0	92±1.4
Max. OD <sub>600</sub> []	7.9±0.0	8.3±0.2	8.7±0.1	4.3±0.2	5.1±0.0	5.3±0.1
Carbon recovery [%]	129±1.4	125±1.5	125±1.9	n.a.	n.a.	n.a.
Product concentrations [mM]						
Carbon source used	403.7±1.7	419.6±4.1	441.1±1.4	284.7±4.0	240.9±3.1	236.4±3.7
Lactate	18.0±0.2	4.0±0.1	4.1±0.3	14.9±0.2	8.8±0.3	7.6±0.5
Acetate	87.9±2.7	91.2±2.4	82.6±1.2	46.3±0.8	54.4±1.2	60.1±1.3
Butyrate	21.6±0.4	1.5±0.3	2.8±0.0	16.8±0.6	1.6±0.0	1.5±0.0
1,3-Propanediol	68.1±1.1	71.2±0.8	76.5±1.7	42.5±1.1	43.1±0.1	45.1±0.5
Ethanol	29.6±0.3	31.1±0.7	31.0±0.1	17.4±0.4	22.4±0.2	22.4±0.4
Butanol	116.6±1.3	141.7±2.7	155.4±1.8	84.7±2.1	75.2±1.9	71.8±1.1
Selectivity [M/M]						
BuOH/Solvents	0.544±0.001	0.581±0.003	0.591±0.002	0.586±0.009	0.534±0.006	0.515±0.003
Yield [M/M]						
BuOH/C-Source	0.289±0.003	0.338±0.004	0.352±0.005	0.297±0.005	0.312±0.004	0.304±0.001
EtOH/C-Source	0.073±0.001	0.074±0.002	0.070±0.000	0.061±0.001	0.093±0.002	0.095±0.002
PDO/C-Source	0.169±0.002	0.170±0.000	0.173±0.004	0.150±0.006	0.179±0.002	0.191±0.002
Solvents/C-Source	0.531±0.005	0.582±0.004	0.596±0.010	0.508±0.006	0.584±0.004	0.589±0.003
Acids/C-Source	0.316±0.006	0.231±0.006	0.203±0.003	0.274±0.008	0.269±0.008	0.293±0.006

*C. acetobutylicum* show the highest identity of the aligned species with 76 % and the other Rex proteins range from 33 % to 40 % identity. The gene can be found upstream of the *crt-bcd-etfAB-hbd* operon responsible for butyryl-CoA biosynthesis (*bcs*) in the same way as reported in *C. acetobutylicum* (Wietzke and Bahl, 2012) and in *C. kluyveri* (Hu et al., 2016). In the protein sequence the wing segment (GXXGXXGXY), responsible for docking to DNA, is conserved starting from position 60. The characteristic ‘Rossmann fold’ GXGXXG is conserved following V93. Also the Asp residue, presumptively distinguishing NADPH from NADH, is conserved (D120) (Figure 4.4). This data gives a good indication that the identified protein could have similar function as other well-described Rex proteins.

**DNA Recognition Motif** Ravcheev et al. (2012) reported 5'-TTGTAAANNNTTAACAA as the consensus sequence of the Rex target region. Using this sequence the *C. pasteurianum* genome was analysed for Rex binding boxes with the ‘Virtual Footprint’ algorithm (Münch et al., 2005) allowing 2 mismatches in a similar fashion to Wietzke and Bahl (2012). As the *C. pasteurianum* genome was not available in the PRODORIC database (Münch et al., 2003), to which ‘Virtual Footprint’ is linked, the resulting table had to be analysed manually. Thus, every hit was searched against the genome of *C. pasteurianum* DSM 525 (Poehlein et al., 2015) using the Artemis genome browser (Rutherford et al., 2000). Information extracted from this search was locus tag, gene name (if applicable), protein function and location of the target sequence in respect to the genome context. As in other organisms confirmed regulatory sequences were found exclusively in intergenic regions. In the next step the list was filtered and only targets in intergenic regions with a putatively regulated gene up or downstream of the site were considered. For every entry in the list the distance to the closest start codon was calculated and a frequency plot with a bin size of 10 bp was plotted. The result is a positively skewed distribution with the maximal target site occurrence between 80 bp and 90 bp from the start codon. After a distance of 250 bp only single occurrences are noted. This observation led to further filtering of the list with a threshold of 250 bp target distance to the start codon. The resulting list comprised 40 sequences which were used in an iterative approach to run ‘Virtual Footprint’ with a new Position Weight Matrix (PWM). The resulting list of 113 targets was filtered again excluding hits in coding regions of genes and in intergenic regions with antidromic genes. This final list comprises 47 targets which putatively regulate downstream genes (Table 4.2). Figure 4.5 shows the DNA motif extracted from all the target sequences in this final list. From this a new consensus sequence

## CHAPTER 4. METABOLIC ENGINEERING OF *C. PASTEURIANUM* FOR IMPROVED SOLVENT PRODUCTION



**Figure 4.4** – Clustal  $\omega$  (Sievers et al., 2011) alignment of reported Rex proteins. For this representation 21 amino acids at the C-terminus of *S. coelicolor* were withheld as they did not carry additional information. The conserved glycine residues are shown in bold and the conserved motifs are shown in colour: the winged helix motif in yellow, the 'Rossmann fold' in blue and the aspartate residue presumptively distinguishing NADPH from NADH in red. Rex species are reported: *C. pasteurianum* in this study, *Clostridium acetobutylicum* in Zhang et al. (2014). *B. subtilis* in Gyan et al. (2006), *S. aureus* in Pagels et al. (2010), *T. aquaticus* in Sickmier et al. (2005) and *S. coelicolor* in Brekasis and Paget (2003).

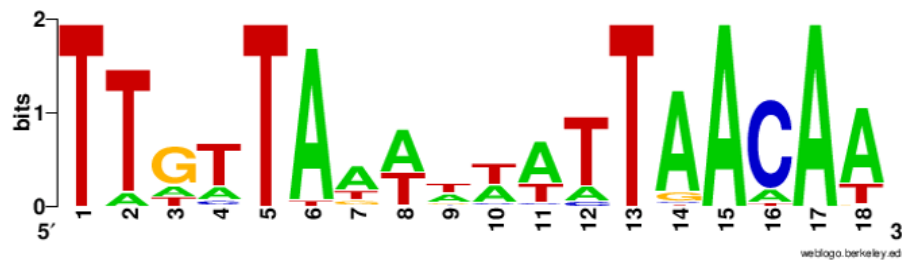
can be derived for *C. pasteurianum* 5'-TTGTAAWNNNTTAACAA.

The targets found by this approach might give an indication of which genes are regulated by Rex in *C. pasteurianum*. In the fermentative pathway the following targets can be found: *adhE1*, *adhE2* and *adhE4* (two target sites upstream), *crt-bcd-etfBA-hbd*, two *thlA*-genes (*thlA1* and *thlA2*), *pta* and *dhaB*. All of these target sites apart from *pta* were previously described based on either *in silico* data in other Clostridiaceae (*dhaB* as the first gene in the *dha* regulon *dhaB-pduD-pduE-ddrA-CLPA\_c22770-CLPA\_c22760-CLPA\_c22750-dhaT2* in Ravcheev et al. (2012)) or were experimentally confirmed in *C. acetobutylicum* (all but *dhaB* and *pta* in Zhang et al. (2014)). To date there is no mention in literature of regulation of *pta* by Rex in any of the analysed organisms. Based on the similarity of the *C. pasteurianum* and *C. acetobutylicum* Rex regulator and predicted target site consensus sequence it was expected to find target sites upstream of *ldh* and *hydA* but even with relaxed algorithm parameters none were found.

Of the non-fermentative genes reported none were found with this approach. However, other notable putative Rex targets found include CLPA\_

**Table 4.2** – Table of putatively Rex regulated genes with target sequence and distance from closest start codon.

Genome position	Sequence	Locus tag	Gene	Protein function	ATG-Gene on distance
185124	TTGATAAATTTTAAATAA	CLPA_c01790	<i>mpg1</i>	putative mannose-1-phosphate guanyltransferase	84 complement
193348	TTATTATAAATATAACAA	CLPA_c01840		transcriptional regulator	81 complement
442350	TTGTTATGTACTTAACAA	CLPA_c04110		transcriptional regulator, AraC family	119 same strand
470707	TTGTTATATTTATAACAT	CLPA_c04410		hypothetical protein	146 same strand
542964	TTGTTAAATGTTAACAA	CLPA_c05110		iron (metal) dependent repressor, DtxR family	377 same strand
580610	TTGTTATATAAATAACAA	CLPA_c05450		nitroreductase	413 complement
617573	TTGTTAAATAAATAACAA	CLPA_c05730	<i>adhE1</i>	aldehyde-alcohol dehydrogenase	37 same strand
946770	TAGTTAAAAATTTAACAT	CLPA_c08620	<i>thlA1</i>	acetyl-CoA acetyltransferase	114 complement
955099	TTGTTAATACACTAACAG	CLPA_c08710		MarR family transcriptional regulator	15 same strand
1058613	TTTTTATTTTATTGACAA	CLPA_c09900		nicotinamidase family protein	74 same strand
1324031	TTGATAAAATACTAACAT	CLPA_c12290		ABC-type Fe3+-siderophore transport system, permease component	99 same strand
1351981	TTATTATAAAATTAACAA	CLPA_c12470	<i>cbiM</i>	cobalt transport protein CbiM	270 complement
1390593	TTTTTATTTATTAACAT	CLPA_c12900		hypothetical protein	498 complement
1393572	TTATTACAATATTAACAA	CLPA_c12920		peptidase family M50	71 same strand
1425190	TTGATAGTGTTAAAAAA	CLPA_c13250	<i>cysE</i>	serine acetyltransferase	568 same strand
1429806	TTGTTAAAAATTTGACAA	CLPA_c13290		hypothetical protein	86 same strand
1429918	TTTTTATATCAATAACAA	CLPA_c13300		hypothetical protein UPF0178	221 same strand
1615045	TTATTAGAAAAATAACAA	CLPA_c15130		hypothetical protein	35 same strand
1617865	TTGCTAATTTTAAACAT	CLPA_c15160	<i>adhE2</i>	aldehyde-alcohol dehydrogenase	97 complement
1881969	TTGATAATTTTAAACAA	CLPA_c17480		flavodoxin	83 same strand
1952030	TTATTAAATAATTGACAA	CLPA_c18120	<i>flgB</i>	flagellar basal body rod protein FlgB	428 complement
2225549	TTGATAACAAATAACAT	CLPA_c20830	<i>pta</i>	phosphate acetyltransferase	256 same strand
2450806	TTGATAATTTTAAACAA	CLPA_c22810	<i>dhaB</i>	glycerol dehydratase large subunit	64 complement
2545028	TTGATAATTTTAAACAT	CLPA_c23760	<i>aldA</i>	putative aldehyde dehydrogenase AldA	81 complement
2563784	TAGATAATTTTAAACAA	CLPA_c23901		hypothetical protein	307 same strand
2757266	TTTATACACTCTTAACAA	CLPA_c25790		hypothetical protein	173 complement
2776821	TTGTTAAAAACTAACAA	CLPA_c26000		NifU-like domain containing protein	121 same strand
2843829	TTATTATACTATTAACAT	CLPA_c26620		hypothetical protein	82 complement
2937664	TTGTTAATATATTAACAT	CLPA_c27550		LysR family transcriptional regulator	75 same strand
2984508	TTATTAATTAATTAACAA	CLPA_c27880		putative transketolase N-terminal section	156 same strand
3015230	TAGGTAAAGTATTAACAA	CLPA_c28170	<i>gpmI</i>	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	162 same strand
3067448	TTGTTAGAATATTAACAA	CLPA_c28630	<i>crt2</i>	3-hydroxybutyryl-CoA dehydratase	27 same strand
3092582	TTGATAATATATTAACAA	CLPA_c28840	<i>mtlA</i>	PTS system mannitol-specific E1-ICB component	307 same strand
3106226	TTGTTTGACAATTAACAA	CLPA_c28930	<i>adhE4</i>	aldehyde-alcohol dehydrogenase	121 same strand
3106321	TAGTTAAATTTTAAACAA	CLPA_c28930	<i>adhE4</i>	aldehyde-alcohol dehydrogenase	216 same strand
3259481	TTATTAATGGATTAACAA	CLPA_c30260	<i>ssuC1</i>	putative aliphatic sulfonates transport permease protein SsuC	216 same strand
3612217	TTTTTAATTCATTGACAA	CLPA_c33470		hypothetical protein	365 same strand
3634909	TTTTTAAAGGTTTAAACAA	CLPA_c33660	<i>mdeA3</i>	methionine gamma-lyase	44 same strand
3707143	TTGCTAGTTTATTAACAT	CLPA_c34300		hypothetical protein	17 complement
3835308	TTGTTATTTAATTCACAA	CLPA_c35710	<i>thlA2</i>	acetyl-CoA acetyltransferase	47 complement
3846267	TATTTAAAAATTAACAA	CLPA_c35810	<i>spoVD2</i>	stage V sporulation protein D	207 same strand
3971344	TTTTTAGAAATATAACAA	CLPA_c37090		hypothetical protein	34 same strand
4077063	TTGCTAATTAATTAATAA	CLPA_c38200	<i>nifH6</i>	nitrogenase iron protein	813 same strand
4082067	TTGTTATTGTATTAACAA	CLPA_c38230	<i>opuCA</i>	carnitine transport ATP-binding protein OpuCA	111 same strand
4160172	TTGCTAAAAATTTAAAAA	CLPA_c39050	<i>asd</i>	aspartate-semialdehyde dehydrogenase	589 same strand
4166771	TTATTAATTTATTAACAA	CLPA_c39090		NAD(P)H dehydrogenase	136 same strand
4264175	TTGTTAATATATTTACAA	CLPA_c40020		hypothetical protein	14 same strand



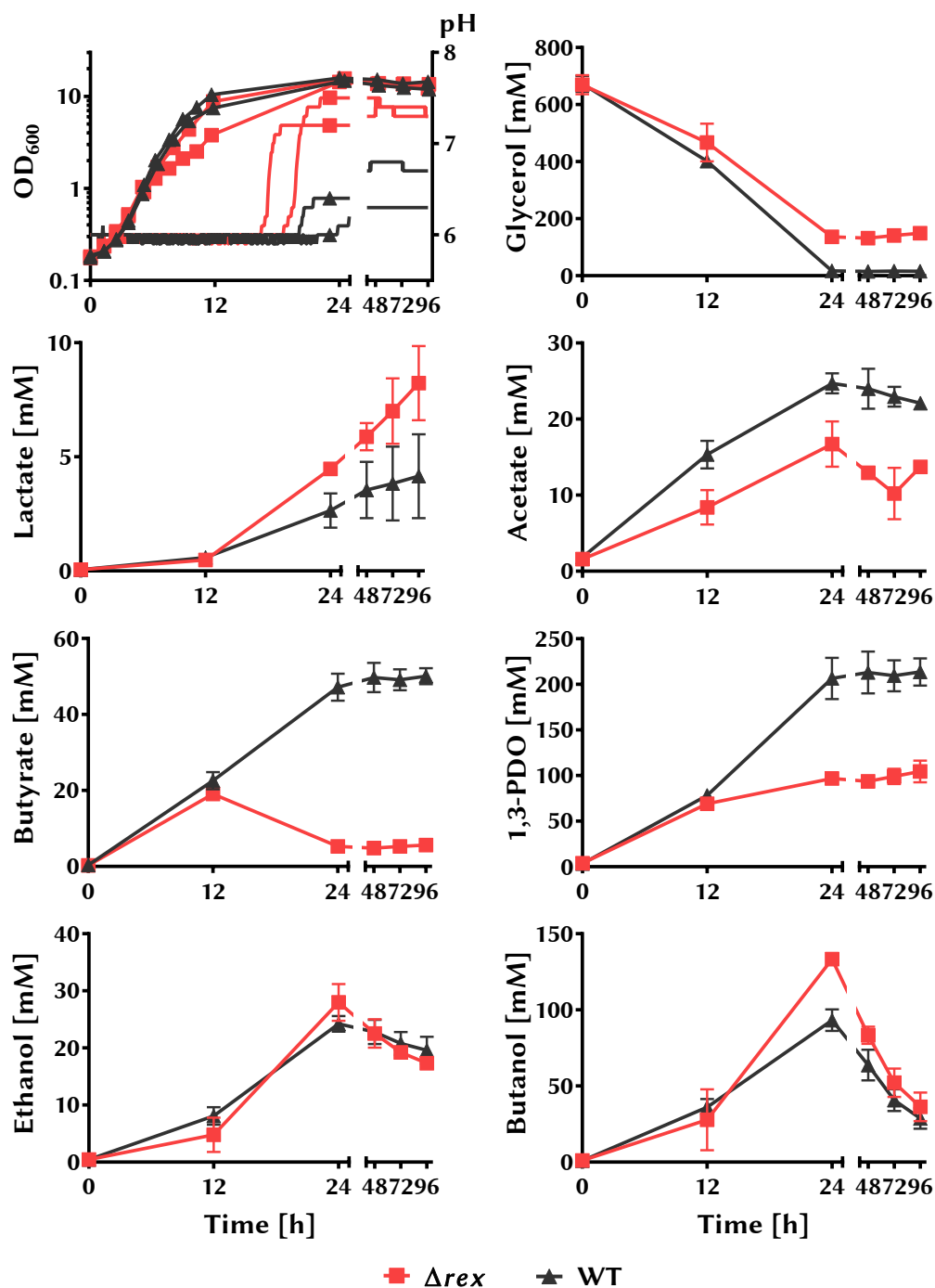
**Figure 4.5** – The sequence logo of Rex DNA recognition motif derived from 47 putative target sequences of *C. pasteurianum* Rex in Table 4.2. The size of individual nucleotides indicates the frequency of occurrence (y-axis) at individual nucleotide positions of the Rex recognition site. DNA logo was generated with WebLogo (Crooks et al., 2004).

c05450, a nitroreductase not related to *narAB* or *narK* reported in *C. carboxidivorans* (Zhang et al., 2014), CLPA\_c09900, a nicotinamidase family protein, CLPA\_c17480 encoding flavodoxin, *spoVD2*, the stage V sporulation protein and CLPA\_c39090, a NADPH dehydrogenase.

The *C. pasteurianum* Rex regulator shows high similarity with well-described Rex proteins and target sequences were found in genes of the fermentative pathway and the alcohol dehydrogenases in particular. Thus, the *in silico* approach gives an idea of which genes might be negatively regulated by the *C. pasteurianum* Rex redox regulator and it can be concluded that a deletion of *rex* might have beneficial effect on solvent generation as it was reported in *C. acetobutylicum* (Zhang et al., 2014; Wietzke and Bahl, 2012).

#### 4.2.3.2 Rex Deletion Phenotype

**Glycerol Fermentation in a Bioreactor** The *rex* deletion mutant was grown in a bioreactor alongside the *C. pasteurianum* DSM 525 wild type strain. It was grown in defined medium with a supplementation of 1 g/l yeast extract to aid growth and 60 g/l glycerol. It seems that the deletion strain grows slower than the wild type though not statistically significant (Table 4.3). Both strains reach a similar optical density of about 15 after one day of fermentation. During acid production in the exponential phase of growth the pH was held at pH 6 but an increase of pH can be observed in early stationary phase. The increase in pH happens drastically over the course of 1- 2 h. The increase is much more distinct in the Rex knock-out mutant with the pH increasing by about 1.2 in 2 h and then reaching a plateau with the pH dropping again in late phases of the fermentation (+48 h). In the wild type the pH only increases by maximal 0.3 in an initial phase to then increase gradually until the fermentation is finished. The onset of the pH shift is earlier in the mutant with an onset of on average



**Figure 4.6** – Bioreactor batch growth and product profile of the *rex* deletion strain *C. pasteurianum*  $\Delta rex$  and the wild type (WT) strain in two individual fermentations. Strains were grown in Biebl medium supplemented with 1 g/l yeast extract and 6 % glycerol. OD<sub>600</sub> was measured manually and pH was measured with an automatic internal probe (lines with single symbols). For clarity, OD and pH is given separately for each replicate. The product formation is given as mean with bars indicating the range of the two samples.

17.5 h after inoculation as opposed to about 22 h in the wild type.

Glycerol is consumed faster in the wild type than in the mutant which correlates with faster growth observed in the wild type. The mutant does not use all the glycerol available. The mutant uses  $533.6 \pm 17.2$  mM glycerol and cannot use 20 % of the initial concentration which suggests limitation by other factors. The wild type consumes with  $656.0 \pm 19.7$  mM nearly all of the glycerol not using only 2 % of the initial glycerol. The difference in final glycerol concentration tested at 24 h is significant for *C. pasteurianum*  $\Delta rex$  ( $m = 136.4$  mM,  $SD = 12.7$  mM) and wild type ( $m = 16.7$  mM,  $SD = 0.06$  mM) conditions;  $t(2) = 13.4$ ,  $p = 0.006$ .

Lactate is produced in notable amounts in both strains only after 12 h and lactate production seems to be highest between 12 and 24 h. However, both strains produce lactate over the whole course of the fermentation with the mutant producing greater quantities and more rapidly.

Acetate production starts in both strains early and a linear production rate can be observed until 24 h, after which production stalls and acetate was re-assimilated in the deletion mutant. The wild type strain produces acetate faster and more than the deletion mutant however not to a significant extent. The production rate could again be due to the faster growth rate in the wild type.

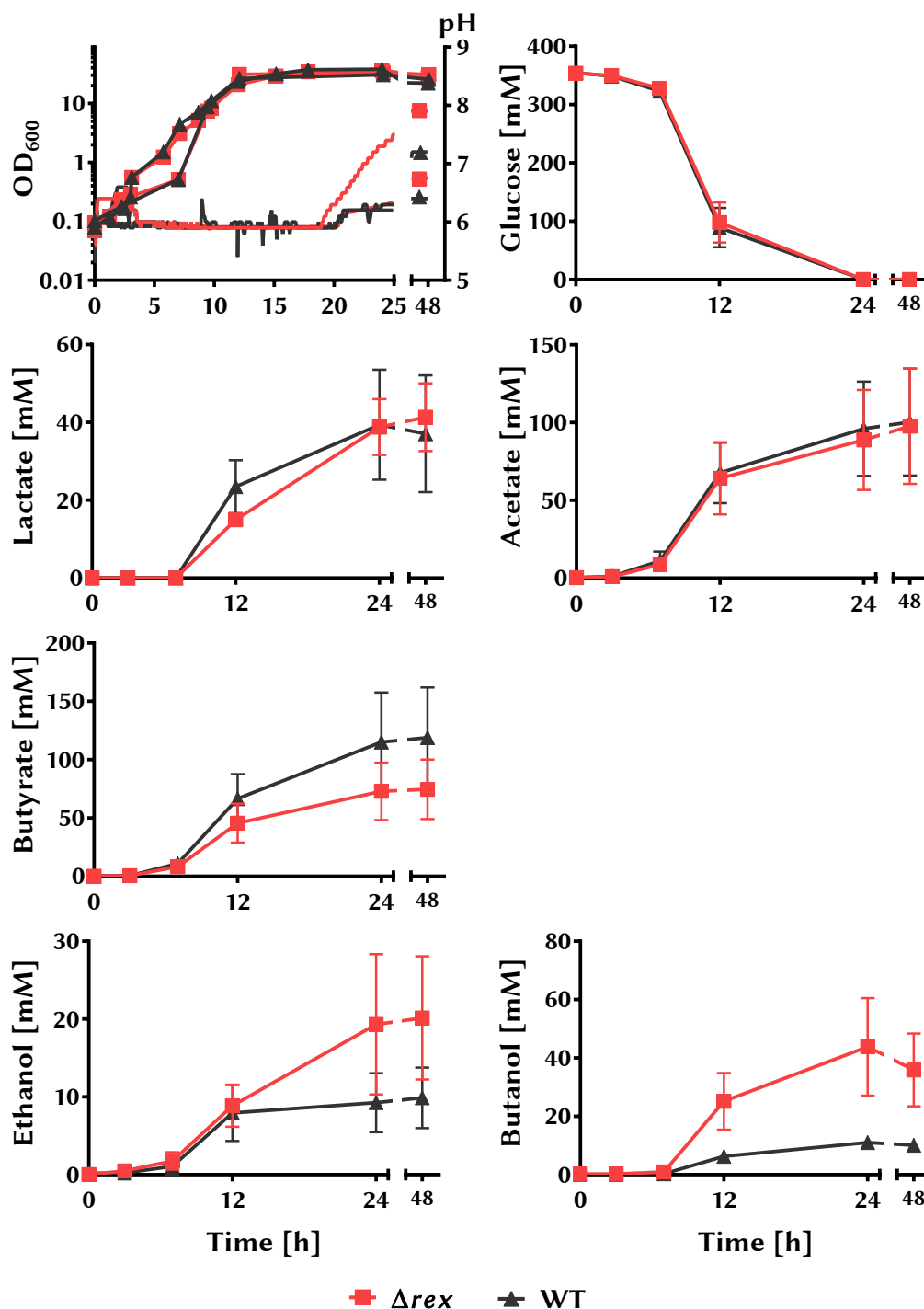
Butyrate production is similar in both strains up to 12 h but diverges largely in later phases of the fermentation. The wild type produces more butyrate until 24 h and stops production after that. The mutant on the other hand re-assimilates the butyrate produced. This leads to a significant difference at 24 h in butyrate produced for the mutant ( $m = 5.3$  mM,  $SD = 0.5$  mM) and wild type ( $m = 47.2$  mM,  $SD = 5.0$  mM) conditions;  $t(2) = 11.7$ ,  $p = 0.007$ .

Ethanol production is unchanged in the mutant compared to the wild type and increases in both strains up to 24 h and decreases slightly after that probably due to evaporation of the solvent. 1,3-PDO production is impaired in the deletion strain. In both strains 1,3-PDO is produced until 24 h with production ceasing after that. At 24 h there is a significant difference between the strains with the mutant ( $m = 96.9$  mM,  $SD = 6.1$  mM) and wild type ( $m = 206.5$  mM,  $SD = 32.0$  mM) conditions;  $t(2) = 4.8$ ,  $p = 0.04$ .

Finally, up to 12 h butanol production is similar in the tested strains after which the deletion strain produces significantly more butanol than the wild type. Tested at 24 h *C. pasteurianum*  $\Delta rex$  ( $m = 133.3$  mM,  $SD = 3.7$  mM) and wild type ( $m = 93.2$  mM,  $SD = 10.0$  mM) conditions;  $t(2) = 5.3$ ,  $p = 0.03$ . After that the butanol concentration in both strains decreases due to evaporation.

Total solvent and acid yield at 24 h is similar in both strains. Solvent yield is  $0.48 \pm 0.02$  mM/mM in the mutant and  $0.5 \pm 0.03$  mM/mM in the wild type.





**Figure 4.7** – Bioreactor batch growth and product profile of the *rex* deletion strain *C. pasteurianum*  $\Delta rex$  and the wild type (WT) strain in two individual fermentations. Strains were grown in Biebl medium with 6 % glucose. OD<sub>600</sub> was measured manually and pH was measured with an automatic internal probe (lines with single symbols). For clarity, OD and pH is given separately for each replicate. The product formation is given as mean with error-bars indicating the range of the two samples.

Acid yields differ widely (mutant;  $0.05 \pm 0.01$  mM/mM, wild type;  $0.11 \pm 0.02$  mM/mM) if butyrate is included but are similar if excluded (mutant;  $0.04 \pm 0.01$  mM/mM, wild type;  $0.11 \pm 0.00$  mM/mM). This suggests that butyrate is solely responsible for the difference in pH. It seems that the mutant strain differs to the wild type in that it re-assimilates butyrate at a higher rate. It cannot be concluded from the data if the wild type takes up the acid at all or if production is higher than re-assimilation.

**Glucose Fermentation in Bioreactor** *C. pasteurianum*  $\Delta rex$  and *C. pasteurianum* wild type were grown in multifors bioreactors in two independent replicates in Biebl minimal medium with 6 % glucose (Figure 4.7). Due to small sample size during the growth phase in the first fermentation growth characteristics in Table 4.3 were calculated from the second fermentation only. However, in the graph it can be seen that both strains grow at a similar rate during exponential phase and reach a similar maximal OD<sub>600</sub> of  $34.9 \pm 1.0$  for the mutant and  $34.3 \pm 2.7$  for the wild type. The pH was held at pH 6 with KOH and was not controlled with acid. It can also be observed, that the pH in the wild type of the second fermentation shows random spikes which is due to a faulty electrode. This does not have an influence on the fermentation itself but is a technical artefact. After 19- 20 h of fermentation the pH starts to increase rapidly with a more distinct increase in the deletion mutant. The increase in pH probably marks the end of sugar fermentation and thus the end of acid production and switch into acid re-assimilation and increased solvent production.

When looking at product formation and glucose uptake no significant difference can be observed. Glucose was used entirely in all fermentations over the course of 24 h. Lactate and acetate production is comparable in both strains and production seems to stall after 24 h. Ethanol production rate is similar in both strains up to 12 h when it ceases in the wild type which leads to a higher final ethanol concentration in the deletion mutant.

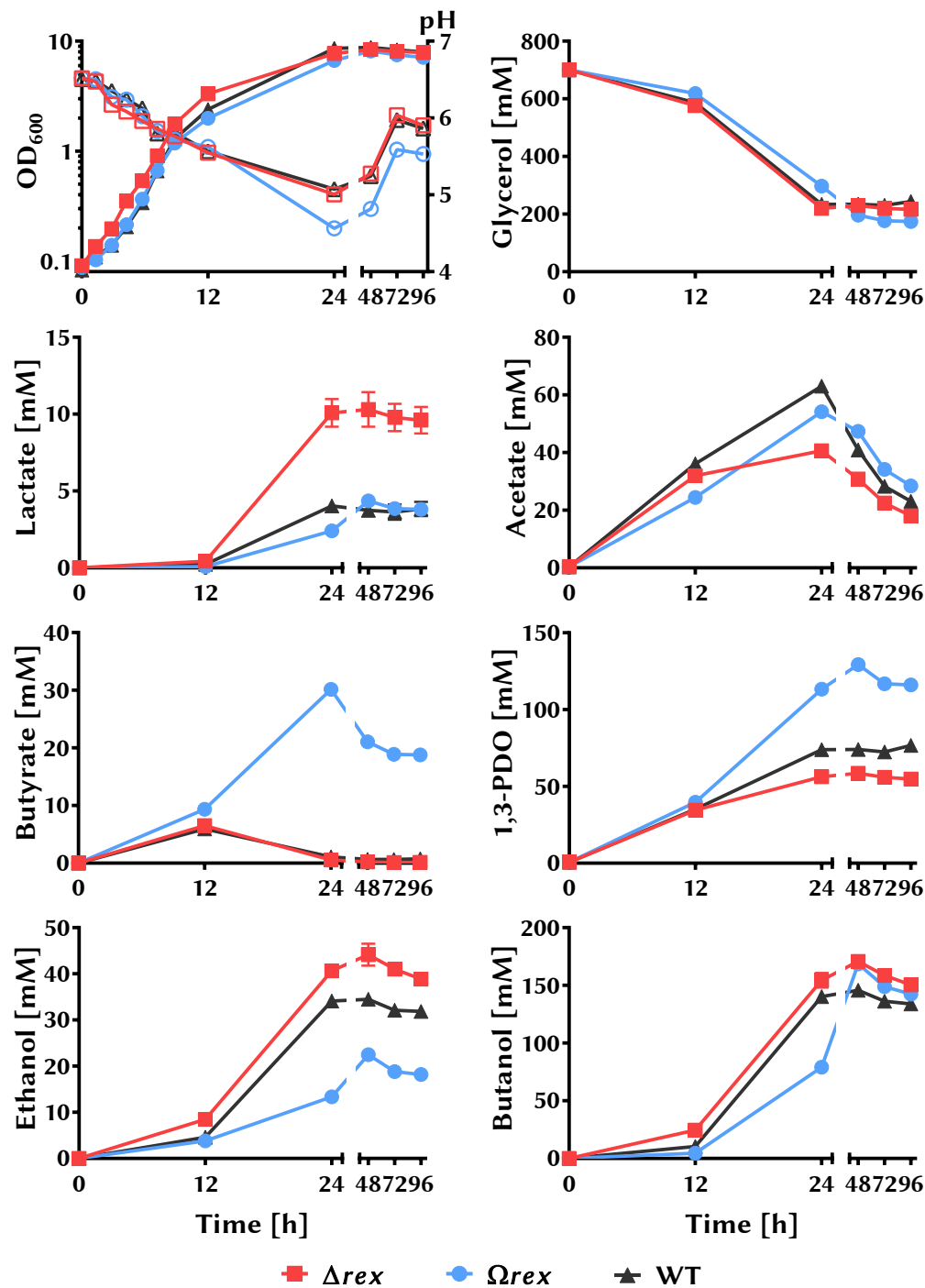
There is a trend to more butyrate production in the wild type when compared to the  $\Delta rex$  strain and inversely the mutant produces more butanol than the wild type. This suggests that the increase in butanol in the mutant is due to lower butyrate production.

The glucose fermentation in the bioreactor gives some indication on differences between the wild type strain and the *rex* deletion mutant. It seems that the mutant produces more solvents; probably due to it being able to re-assimilate butyrate in the case of butanol production and by favourable regulation of ethanol production.

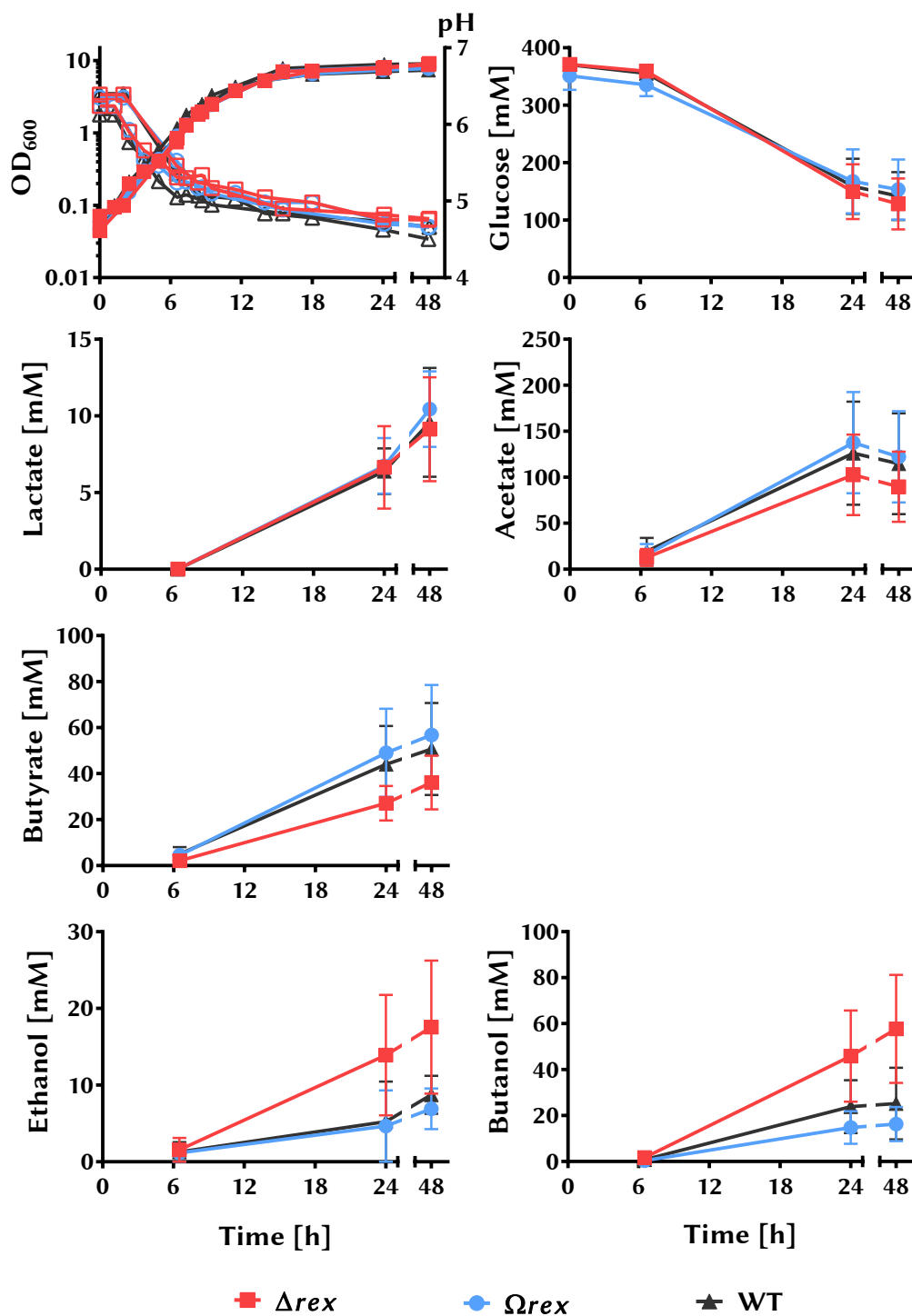
**rex Complementation Analysis** To test if the effects seen in the *C. pasteurianum*  $\Delta rex$  mutant did have their basis in the knock-out of the gene and are not based on other mutations in the genome the strain was complemented downstream of the *pyrE* locus as described in Section 4.2.1.2. The complementation strain *C. pasteurianum*  $\Omega rex$  was tested alongside the deletion mutant and the wild type in fermentations in serum bottles on glycerol and glucose. The results of fermentations of these strains can be seen in Figures 4.8 and 4.9.

When grown on glycerol the wild type and the complementation strain show a slight lag phase but the growth rate after that is comparable to the deletion mutant. The pH drops in the same way in all strains but the pH of the complemented mutant strain drops lower than the other two. This is most likely due to an increase in butyrate production which is not found in either wild type or *rex* deletion mutant. Glycerol uptake is similar in all strains with the complementation strain using it slower but being able to use more. Lactate production is increased in the deletion mutant compared to the wild type but complemented well by *C. pasteurianum*  $\Omega rex$ . Acetate levels are reduced when *rex* is deleted and its complementation brings the levels back close to wild type. Butyrate is re-assimilated in all three strains but the complementation strain behaves differently to the wild type and the deletion. Butyrate levels are much higher reaching a maximum of  $30.1 \pm 0.5$  mM after 24 h as compared to the wild type with  $5.9 \pm 0.1$  mM and the deletion strain with  $6.5 \pm 0.1$  mM reached after 12 h already. 1,3-PDO production is higher in the wild type than in the deletion strain and the complementation seems to overcompensate and produces 1.5 times as much as the wild type. For ethanol production, on the other hand, the compensation is inverse with the wild type producing less ethanol than the deletion strain and the complementation producing only two thirds of the wild type levels and produces it in a later stage of fermentation reaching the maximum level only after 48 h when the other strains reach the maximal ethanol levels already after 24 h. Differences in butanol production are small. The complementation strain produces butanol slower than the other strains and goes up to similar levels as the deletion strain.

To summarise the findings of the *rex* complementation strain it can be said that the wild type phenotype was largely restored for the acids lactate and acetate while butyrate production was highly increased in this strain compared to both the *rex* deletion and the wild type strain. Solvent production in the oxidative pathway is highly decreased compared to the wild type while the reductive pathway to 1,3-PDO is more active. This inverted product profile compared to *C. pasteurianum*  $\Delta rex$  could indicate over-expression of *rex* in the complementation strain. Such variation can be due to genome location



**Figure 4.8** – Batch growth in serum bottles of the *rex* deletion strain *C. pasteurianum*  $\Delta rex$ , its complementation *C. pasteurianum*  $\Omega rex$  and the wild type (WT) strain. Strains were grown in three biological replicates in Biebl medium supplemented with 1 g/l yeast extract and 6 % glycerol. Open symbols are used for pH. Data is given as mean of the three replicates with error-bars indicating the standard error.



**Figure 4.9** – Batch growth in serum bottles of the *rex* deletion strain *C. pasteurianum*  $\Delta rex$ , its complementation *C. pasteurianum*  $\Omega rex$  and the wild type (WT) strain. Strains were grown in two biological replicates in Biebl medium with 6 % glucose. Open symbols are used for pH. Data is given as mean of the two replicates with error-bars showing the range of the samples.

downstream of *pyrE* of the complementation as opposed to its native location upstream of the *bcs* operon.

In glucose fermentation all strains show similar growth characteristics but pH levels of the wild type and complementation strain are a bit lower than the deletion mutant's. Glucose consumption and lactate levels are similar between all three strains. Acetate and butyrate levels are reduced in the deletion strain compared to the wild type and the complemented strain. Solvent production is increased in *C. pasteurianum*  $\Delta rex$  and complementation to wild type levels is achieved in *C. pasteurianum*  $\Omega rex$ .

**Table 4.3** – Growth characteristics and product yields of *rex*, *hyd*, *dhaBCE* mutant bioreactor fermentations with wild type control. Data given as mean±SEM from biological duplicates.

Medium	Biebl plus 1 g/l yeast extract 60 g/l glycerol			Biebl 60 g/l glucose			
Carbon source	$\Delta rex$	$\Delta hyd$	WT	$\Delta rex$	$\Delta hyd$	$\Delta dhaBCE$	WT
Strain							
Growth characteristics							
Specific Growth Rate [ $h^{-1}$ ]	0.32±0.04	0.35±0.01	0.42±0.01	0.15±0.02	0.11±0.01	0.13±0.03	0.16±0.03
Doubling Time [min]	135±15	119±3	100±1	90±9	103±10	85±4	81±12
Max. OD <sub>600</sub> []	15.0±0.5	17.5±0.6	15.3±0.3	34.9±1.0	36.2±0.1	34.8±2.5	34.3±2.7
Carbon recovery [%]	88.9±2.7	91.4±1.3	90.5±4.4	70.7±14.3	83.5±15.0	72.3±13.9	73.3±14.9
Product concentrations [mM]							
Carbon source used	533.6±17.2	632.5±14.3	656.0±19.7	353.9±1.0	356.8±2.9	355.2±2.3	354.4±1.7
Lactate	4.5±0.1	14.4±1.4	2.6±0.5	38.8±5.1	27.5±10.7	34.2±6.7	39.4±10.0
Acetate	16.7±2.1	44.1±0.0	24.7±1.0	88.8±22.7	116.9±28.0	101.0±20.6	96.0±21.4
Butyrate	5.3±0.3	8.7±0.0	47.2±2.5	72.9±17.4	57.1±14.3	107.8±26.3	115.1±30.0
1,3-Propanediol	96.9±3.1	167.6±0.9	206.5±16.0				
Ethanol	28.0±2.3	64.3±2.2	24.2±1.0	19.3±6.4	21.7±6.0	8.8±2.6	9.3±2.7
Butanol	133.3±1.8	105.1±0.0	93.2±5.0	43.8±11.8	55.7±10.0	14.5±2.8	11.1±1.1
Selectivity [M/M]							
BuOH/Solvents	0.516±0.002	0.312±0.003	0.289±0.024	0.703±0.015	0.729±0.021	0.635±0.027	0.563±0.051
Yield [M/M]							
BuOH/C-Source	0.250±0.005	0.166±0.004	0.142±0.003	0.124±0.034	0.156±0.029	0.041±0.008	0.031±0.003
EtOH/C-Source	0.053±0.006	0.102±0.001	0.037±0.000	0.055±0.018	0.061±0.017	0.025±0.007	0.026±0.008
PDO/C-Source	0.182±0.000	0.265±0.005	0.317±0.034				
Solvents/C-Source	0.484±0.011	0.533±0.007	0.496±0.030	0.179±0.052	0.217±0.046	0.066±0.016	0.057±0.001
Acids/C-Source	0.050±0.006	0.106±0.000	0.114±0.008	0.567±0.129	0.567±0.153	0.686±0.156	0.708±0.177

**Table 4.4** – Growth characteristics and product yields of *rex*, *hyd*, *dhaBCE* mutant fermentations with wild type control in serum bottles. Data given as mean±SEM from biological triplicates.

Carbon source	60 g/l glycerol							
	Biebl plus 1 g/l yeast extract					2xYT		
Strain	$\Delta rex$	$\Omega rex$	$\Delta hyd$	$\Omega hyd$	WT	$\Delta cde$	$\Omega cde$	WT
Growth characteristics								
Specific Growth Rate [ $h^{-1}$ ]	0.34±0.01	0.34±0.00	0.27±0.00	0.35±0.00	0.34±0.00	0.39±0.00	0.38±0.00	0.40±0.00
Doubling Time [min]	121±2	124±1	154±3	119±1	123±1	106±1	110±1	105±0
Max. OD <sub>600</sub> []	7.72±0.08	6.673±0.05	4.99±0.05	8.50±0.21	8.59±0.09	4.38±0.12	4.44±0.05	4.48±0.01
Carbon recovery [%]	102.7±2.1	109.6±0.6	128.5±1.0	115.1±2.5	107.5±0.9	114.4±6.0	119.4±5.8	127.8±5.4
Product concentrations [mM]								
Carbon source used	480.4±5.4	403.2±3.2	350.1±7.5	444.0±14.6	466.3±7.9	244.5±11.7	248.2±12.8	241.9±6.9
Lactate	10.1±0.7	2.4±0.0	12.0±0.1	3.9±0.1	4.0±0.3	6.7±0.9	4.7±0.5	5.0±0.3
Acetate	40.6±1.0	54.2±0.8	92.1±2.1	71.1±2.6	63.1±0.5	9.7±0.2	11.6±0.3	14.0±0.2
Butyrate	0.5±0.1	30.1±0.7	0.0±0.0	1.2±0.3	1.1±0.3	13.1±0.6	13.6±0.2	13.8±0.4
1,3-Propanediol	56.3±2.8	113.3±2.4	83.1±0.7	78.9±2.4	73.9±1.2	0.0±0.0	2.2±0.0	11.0±0.2
Ethanol	40.6±0.6	13.3±0.5	57.1±1.6	33.6±1.0	34.1±0.3	15.5±0.2	15.9±0.5	15.8±0.6
Butanol	154.4±5.2	79.0±4.1	76.3±3.1	136.9±3.6	140.2±3.2	96.9±0.5	103.2±2.3	103.8±2.6
Selectivity								
BuOH/Solvents	0.614±0.005	0.384±0.015	0.352±0.007	0.549±0.003	0.565±0.007	0.862±0.002	0.851±0.001	0.795±0.001
Yield [M/M]								
BuOH/C-Source	0.321±0.009	0.196±0.009	0.218±0.005	0.309±0.002	0.301±0.002	0.399±0.019	0.418±0.020	0.431±0.021
EtOH/C-Source	0.085±0.002	0.033±0.001	0.163±0.001	0.076±0.005	0.073±0.001	0.064±0.002	0.065±0.003	0.066±0.003
PDO/C-Source	0.117±0.004	0.281±0.008	0.238±0.004	0.178±0.000	0.159±0.005	0.000±0.000	0.009±0.000	0.046±0.002
Solvents/C-Source	0.523±0.012	0.510±0.004	0.618±0.003	0.563±0.007	0.533±0.005	0.463±0.021	0.492±0.023	0.542±0.026
Acids/C-Source	0.107±0.001	0.215±0.003	0.297±0.003	0.173±0.011	0.146±0.003	0.122±0.009	0.122±0.007	0.136±0.004



**Table 4.5** – Growth characteristics and product yields of *rex*, *hyd*, *dhaBCE* mutant fermentations with wild type control in serum bottles. Data given as mean±SEM from biological triplicates.

Medium	Biebl plus 60 g/l glucose						
Strain	$\Delta rex$	$\Omega rex$	$\Delta hyd$	$\Omega hyd$	$\Delta cde$	$\Omega cde$	WT
Growth characteristics							
Specific Growth Rate [ $h^{-1}$ ]	0.44±0.01	0.46±0.00	0.35±0.02	0.47±0.00	0.49±0.01	0.45±0.00	0.47±0.01
Doubling Time [min]	96±3	91±0	119±7	89±0	85±2	93±1	89±2
Max. OD <sub>600</sub> []	7.95±0.07	7.74±0.27	8.41±0.75	8.69±0.72	8.45±0.60	8.39±0.53	7.92±0.65
Carbon recovery [%]	85.1±32.6	138.0±68.1	88.5±33.3	95.2±39.4	90.0±35.0	96.6±39.8	95.2±39.4
Product concentrations [mM]							
Carbon source used	243.1±25.6	198.0±44.4	214.0±40.8	237.7±20.2	225.9±26.9	233.4±27.5	228.5±25.9
Lactate	9.1±2.0	10.4±1.4	9.6±2.4	9.5±1.2	9.4±1.0	9.3±1.4	9.6±2.1
Acetate	89.6±22.1	122.2±28.6	126.8±30.7	124.5±34.2	119.6±30.3	122.9±31.5	114.6±31.6
Butyrate	36.1±6.7	56.8±12.6	33.3±5.2	50.7±10.8	51.3±9.8	50.8±11.1	50.7±11.5
1,3-Propanediol							
Ethanol	17.6±5.0	6.9±1.5	9.4±2.1	7.6±2.0	7.4±1.8	7.3±1.8	8.7±1.4
Butanol	57.7±13.6	16.3±4.3	49.5±13.9	33.6±10.2	27.8±7.9	31.7±8.2	25.2±9.0
Selectivity							
BuOH/Solvents	0.755±0.011	0.695±0.010	0.833±0.009	0.809±0.008	0.782±0.008	0.811±0.002	0.695±0.052
Yield [M/M]							
BuOH/C-Source	0.264±0.084	0.114±0.047	0.301±0.122	0.156±0.056	0.141±0.052	0.155±0.054	0.129±0.054
EtOH/C-Source	0.081±0.029	0.047±0.018	0.056±0.021	0.035±0.011	0.037±0.013	0.036±0.012	0.042±0.011
PDO/C-Source							
Solvents/C-Source	0.345±0.113	0.161±0.066	0.357±0.143	0.191±0.068	0.179±0.064	0.191±0.065	0.171±0.065
Acids/C-Source	0.615±0.191	1.296±0.505	1.005±0.371	0.845±0.266	0.902±0.289	0.888±0.293	0.866±0.296

## 4.2.4 Hydrogenase

### 4.2.4.1 Hydrogenase Deletion Phenotype

**Glycerol Fermentation in Bioreactor** To analyse fermentation patterns of a hydrogenase knock-out, *C. pasteurianum*  $\Delta hyd$  and *C. pasteurianum* DSM 525 wild type were grown in 350 ml fermentations in defined medium with 60 g/l glycerol. The pH was held static at 6.0 with KOH and allowed to increase into the alkaline. The mutant strain grew significantly slower with a specific growth rate of  $0.36 \pm 0.0 \text{ h}^{-1}$  compared to  $0.42 \pm 0.01 \text{ h}^{-1}$  ( $m \pm SD$ ) in the wild type with conditions;  $t(2) = 6.5$ ,  $p = 0.02$ . The knock-out mutant seems to grow to a slightly elevated OD as compared to the wild type, although the difference was not significant. As observed in the *C. pasteurianum*  $\Delta rex$  strain the onset of increase of pH in the *hyd* deletion mutant was earlier than in the wild type with 18 h and 19.3 h as opposed to 20.2 h and 23.8 h. The pH also increases much more to a final pH of  $8.5 \pm 0.2$  versus  $6.6 \pm 0.3$  in the wild type.

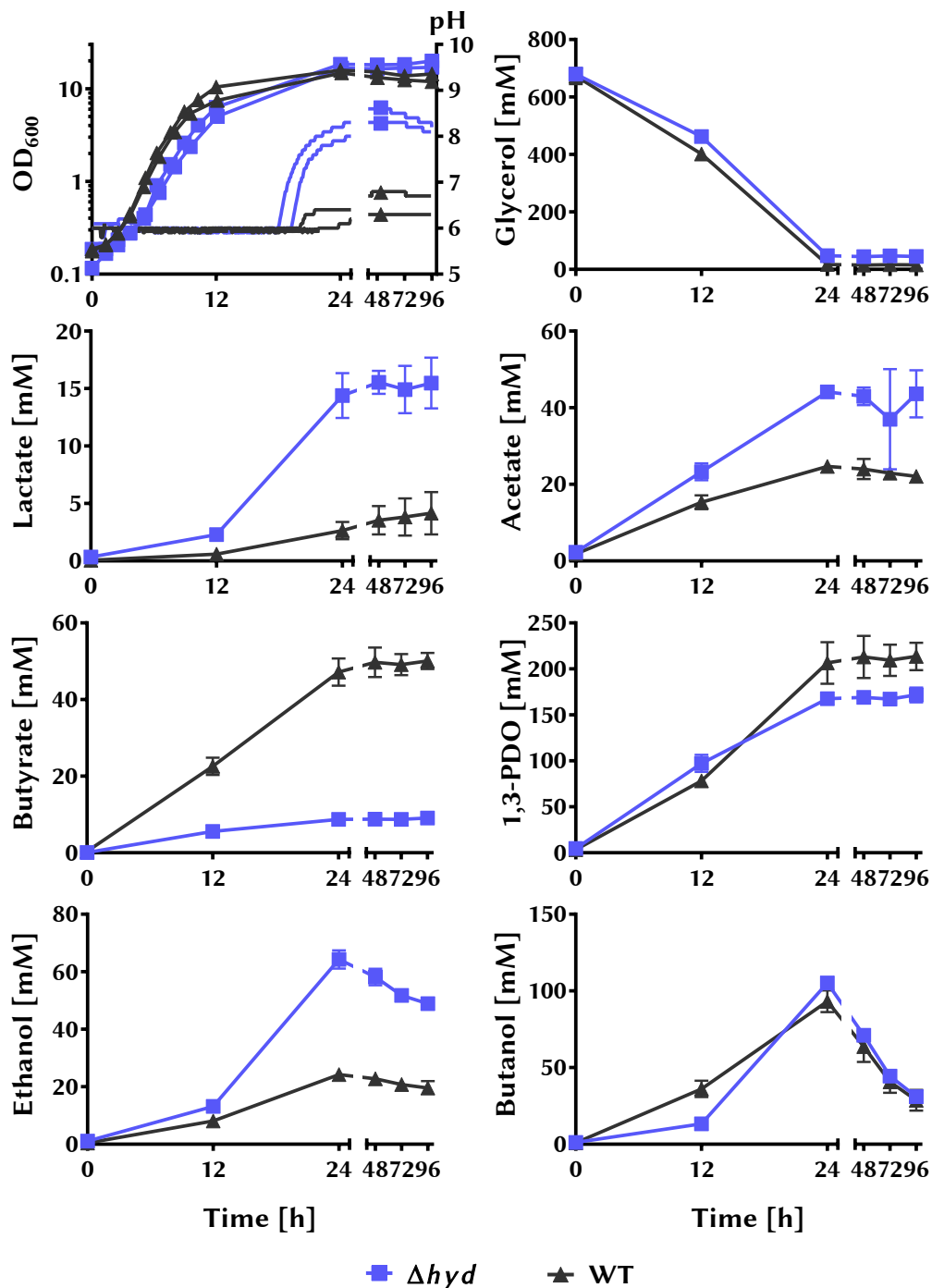
Glycerol is not entirely exhausted in the cultures of either strain. Consumption in the *hyd* mutant takes place at a slower rate than that of the wild type. Final glycerol concentration is 7 % of the initial glycerol in the *hyd* mutant, as opposed to 2 % in the wild type. The acids acetate and lactate are produced much faster and to higher final concentration in the mutant than in the wild type. 5.5 times more lactate and 1.8 times more acetate is produced after 24 h. Lactate production is  $14.4 \pm 2.5 \text{ mM}$  in the mutant and  $2.6 \pm 1.1 \text{ mM}$  in the wild type with conditions;  $t(2) = 5.6$ ,  $p = 0.03$ . Acetate production is  $44.1 \pm 0.0 \text{ mM}$  and  $24.7 \pm 1.9 \text{ mM}$  in the mutants and wild type, respectively with conditions;  $t(2) = 14.8$ ,  $p = 0.005$ .

Butyrate production and end concentration is significantly reduced in the knock-out mutant. At 24 h *C. pasteurianum*  $\Delta hyd$  produced  $8.7 \pm 0.0 \text{ mM}$  and the wild type  $47.2 \pm 3.6 \text{ mM}$  butyrate ( $m \pm SD$ ) conditions;  $t(2) = 10.8$ ,  $p = 0.008$ . Butyrate re-assimilation is not observed in any of the strains.

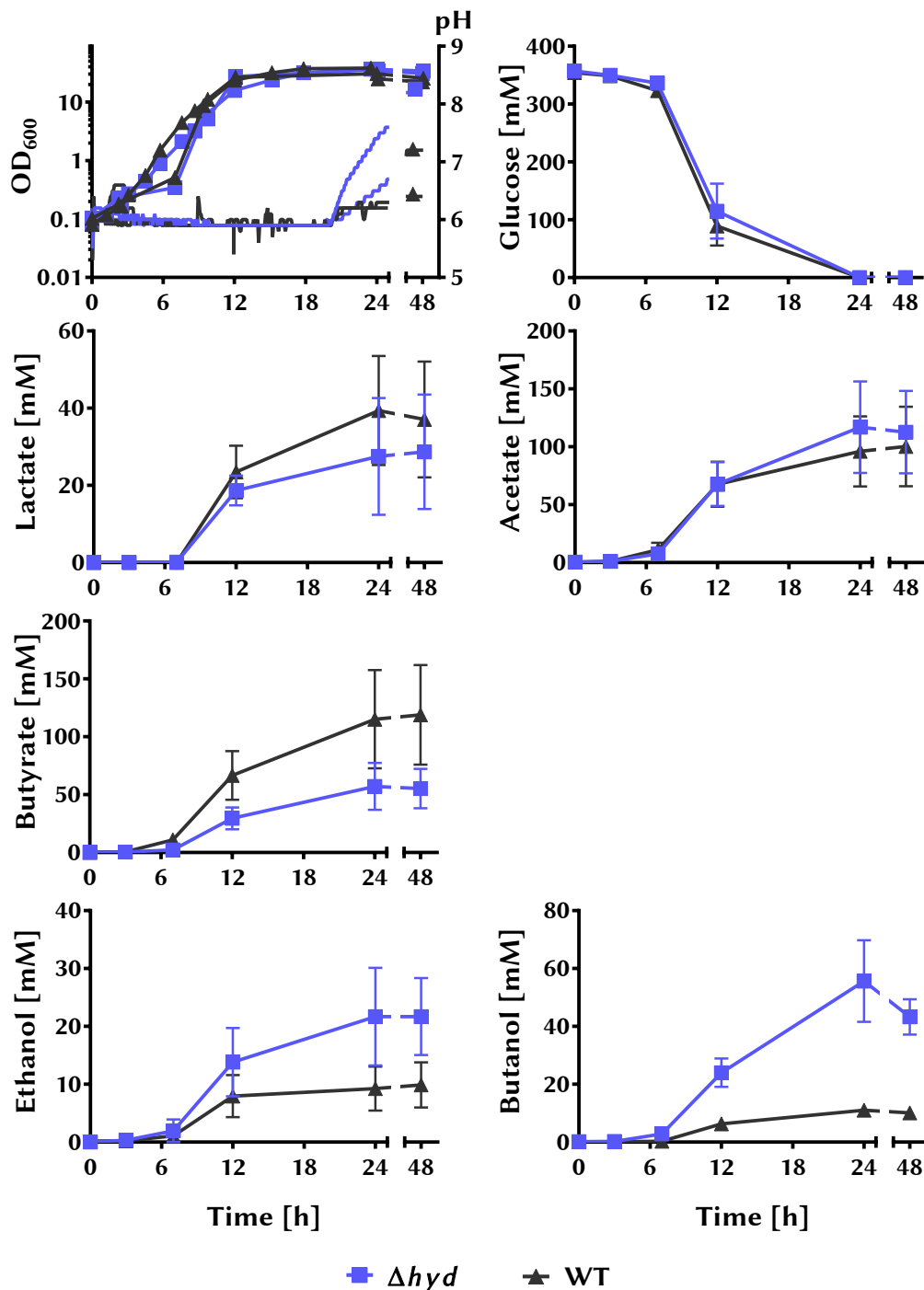
1,3-PDO is production rate similarly in both strains but the mutant's final concentration is about 19 % lower than the wild type's. Ethanol on the other hand is produced significantly (2.7 times) more in the mutant. The deletion strain produces  $64.3 \pm 3.2 \text{ mM}$  and the wild type  $24.2 \pm 1.4 \text{ mM}$  ethanol in 24 h, conditions;  $t(2) = 11.6$ ,  $p = 0.007$ . After 24 h the ethanol concentration decreases due to solvent extraction by nitrogen bubbling through the reactor.

Butanol production starts later in the deletion strain but production is faster after that and reaches a maximum concentration at 24 h which is slightly higher in the mutant ( $105.1 \pm 0.0 \text{ mM}$ ) than in the wild type ( $93.2 \pm 7.1 \text{ mM}$ ).

Total levels of solvents and acids are similar in both strains. The total acid



**Figure 4.10** – Bioreactor batch growth and product profile of the *hyd* deletion strain *C. pasteurianum*  $\Delta hyd$  and the wild type (WT) strain in two individual fermentations. Strains were grown in Biebl medium supplemented with 1 g/l yeast extract and 6 % glycerol. OD<sub>600</sub> was measured manually and pH was measured with an automatic internal probe (lines with single symbols). For clarity, OD and pH is given separately for each replicate. The product formation is given as mean with error-bars indicating the range of the two samples.



**Figure 4.11** – Bioreactor batch growth and product profile of the *hyd* deletion strain *C. pasteurianum*  $\Delta hyd$  and the wild type (WT) strain in two individual fermentations. Strains were grown in Biebl medium with 6 % glucose. OD<sub>600</sub> was measured manually and pH was measured with an automatic internal probe (lines with single symbols). For clarity, OD and pH is given separately for each replicate. The product formation is given as mean with error-bars indicating the range of the two samples.

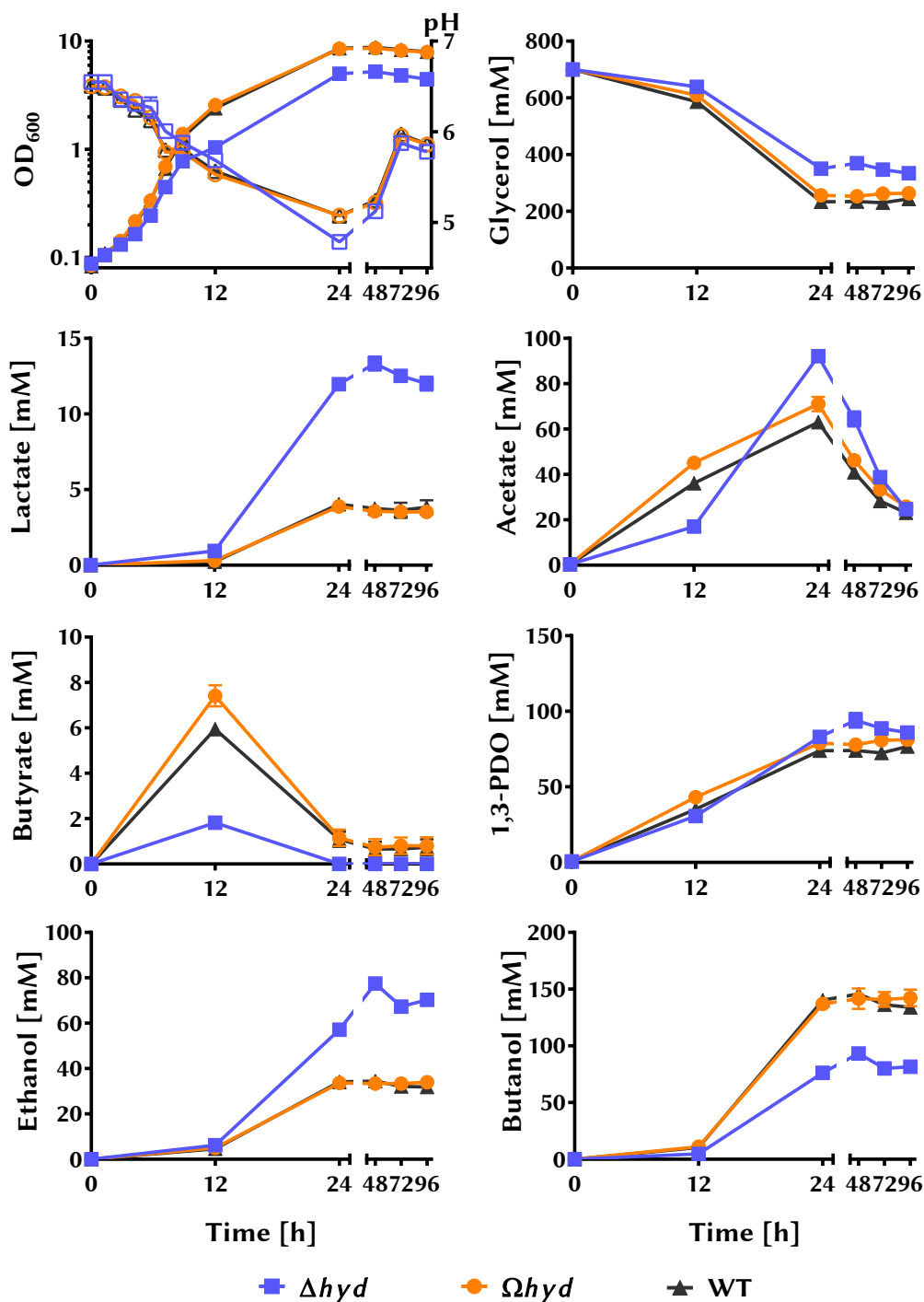
concentration at 24 h in the mutants is  $67.2 \pm 2.0$  mM and in the wild type  $74.5 \pm 4.1$  mM and total solvents reach  $337.0 \pm 4.5$  mM and  $323.9 \pm 14.2$  mM in mutant and wild type, respectively. Thus, there is less acid in the mutant strain which correlates with the increased pH as compared to the wild type and the mutant overall produces more solvents, however this is largely due to the increase in ethanol production.

**Glucose Fermentation in Bioreactor** When grown on glucose the deletion strain grows slightly slower than the wild type. Both strains reach a similar maximal OD<sub>600</sub> of  $36 \pm 0$  ( $\Delta hyd$ ) and  $34 \pm 4$  (wild type) after 24 h. The pH is held stable at pH 6 over the initial phase and with entering of solventogenesis and acid re-assimilation pH increases in both strains. This can be observed in both strains simultaneously. However, the extent of the increase differs largely with the mutant reaching a much higher pH than the wild type. The pH of the wild type in the second experiment shows extensive jumps which is due to a faulty connection of the pH probe but does not have an influence on the outcome of the fermentation.

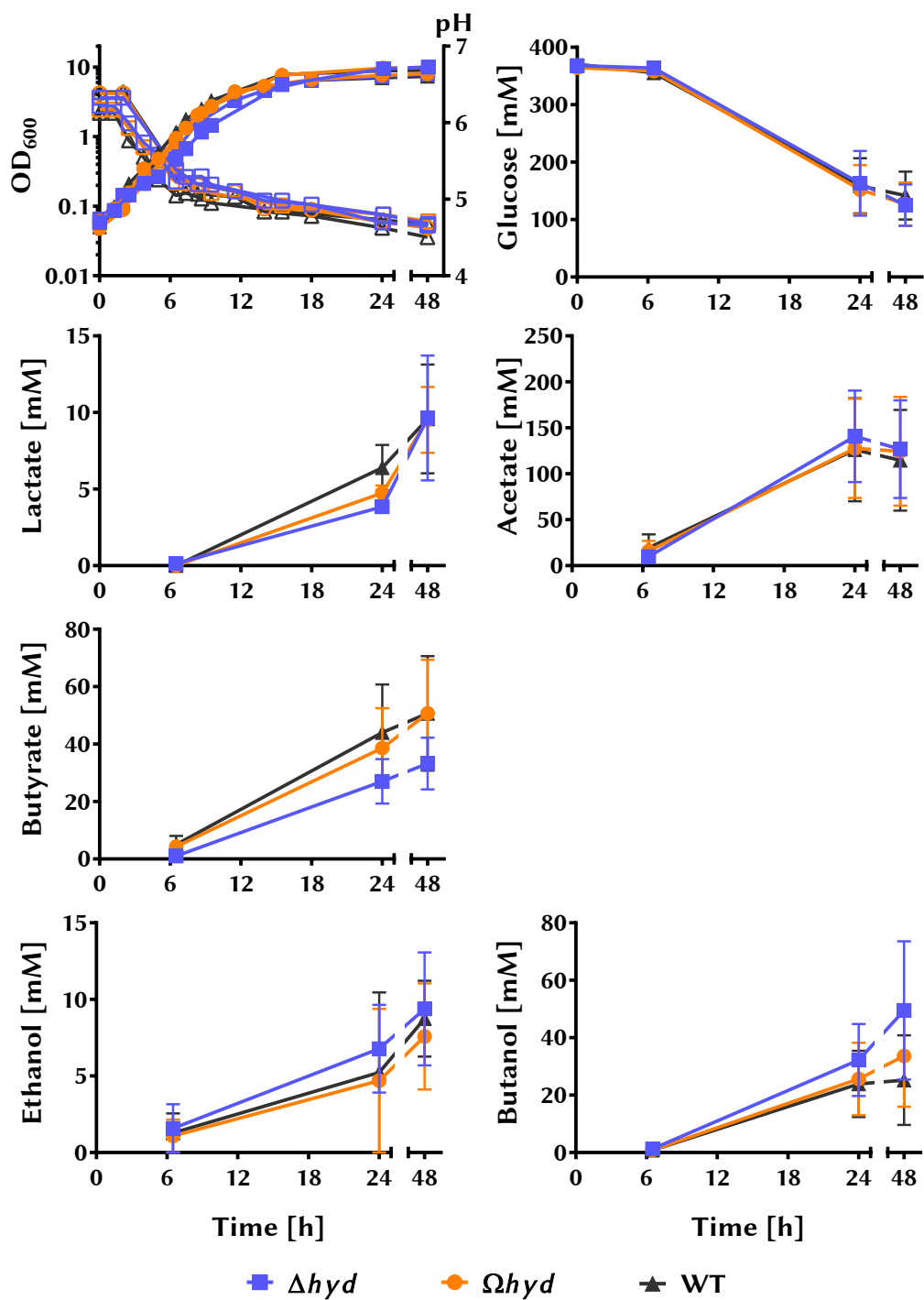
Glucose uptake is similar in the tested strains and both use the available substrate entirely over the course of 24 h. The acids lactate and acetate were produced at the same time and to a similar extend in both strains with the wild type producing slightly more lactate and less acetate. However, due to sample variance this is not conclusive. Butyrate production and final levels are decreased in the mutant strain compared to the wild type.

The *hyd* deletion mutant seems to produce much more solvents than the wild type when grown on glucose. Ethanol production reaches a maximum concentration after 24 h of 2.3 times more in the mutant ( $21.7 \pm 8.4$  mM) than the wild type ( $9.3 \pm 3.8$  mM) and butanol maximal concentration is 5 times higher in the mutant  $55.7 \pm 14$  mM than the wild type ( $11.1 \pm 1.5$  mM). Thus, there is a trend that *C. pasteurianum*  $\Delta hyd$  produces more solvents than the wild type on glucose and especially butanol production is increased.

***hyd* Complementation Analysis** The complementation study was done in 50 ml medium in serum bottles. When grown on glycerol the deletion mutant grows slower and to lower final OD than the wild type which is corrected in the complementation strain *C. pasteurianum*  $\Omega hyd$ . pH reaches a minimum in all strains after 24 h and pH in the deletion strain reaches a lower minimum and stays slightly lower than the wild type. The complementation strain shows the same pH as the wild type over the whole fermentation. Glycerol is consumed slower in the deletion mutant than in the other two strains in accordance with



**Figure 4.12** – Batch growth in serum bottles of the *hyd* deletion strain *C. pasteurianum*  $\Delta hyd$ , its complementation *C. pasteurianum*  $\Omega hyd$  and the wild type (WT) strain. Strains were grown in three biological replicates in Biebl medium supplemented with 1 g/l yeast extract and 6 % glycerol. Open symbols are used for pH. Data is given as mean of the three replicates with error-bars indicating the standard error.



**Figure 4.13** – Batch growth in serum bottles of the *hyd* deletion strain *C. pasteurianum*  $\Delta hyd$ , its complementation *C. pasteurianum*  $\Omega hyd$  and the wild type (WT) strain. Strains were grown in two biological replicates in Biebl medium with 6 % glucose. Open symbols are used for pH. Data is given as mean of the two replicates and error-bars indicating the range of the samples.

the slower growth. The knock-out mutant can use much less substrate than the others using only 50 % of the initial glycerol when the complementation strain uses 63 % and the wild type 67 %.

Lactate and acetate concentration at 24 h are increased in the knock-out mutant. Lactate levels are 3 times higher in the deletion strain and complementation brings it back to wild type levels. Acetate production on the other hand starts slower in *C. pasteurianum*  $\Delta hyd$  and production speeds up after 12 h. After reaching the maximum concentration at 24 h in all strains acetate is re-assimilated.

Much less butyrate is produced in the deletion mutant and the complementation brings the levels back up to wild type levels. After 12 h of fermentation the butyrate levels differ largely between deletion and wild type strain. In all strains butyrate is re-assimilated and no butyrate is detected in the deletion strain after 24 h and only trace amounts are detected in the other strains after this time point.

Not a lot of difference can be observed between all three strains in 1,3-PDO production and concentrations. The solvents ethanol and butanol show a big difference between deletion mutant and wild type which is complemented in *C. pasteurianum*  $\Omega hyd$ . 1.7 times more ethanol is produced by the knock-out strain than the wild type in 24 h. On the other hand only slightly more than half the level (55.7 %) of butanol is produced in the deletion mutant when compared to the wild type. This decrease in butanol production was not observed in the bioreactor whereas increased ethanol production in the strain was, which suggests that the two phenomenon are not linked.

Growth of the deletion, wild type and complementation strain on glucose does not reveal much difference between the three strains and product patterns are also very similar between the three strains. Trends show that glycerol consumption and lactate, acetate and ethanol production are very similar in all three strains. It seems that the deletion strain produces less butyrate and more butanol than the wild type and the complementation strain both of which show concentrations of the two compounds of very similar levels.

## 4.2.5 Glycerol Dehydratase

### 4.2.5.1 Glycerol Dehydratase Deletion Phenotype

**Glycerol Fermentation** To analyse the fermentation phenotype of a *dhaBCE* knock-out mutant, it was tried to grow the strain in a bioreactor in Biebl medium with 1 g/l yeast extract and 6 % glycerol. The wild type grew well under these conditions but the deletion strain was not able to grow (Figure A.2). The dele-





tion strain was able to grow in the cabinet in precultures of 35 ml buffered with  $\text{CaCO}_3$ . But also this growth was diminished when compared to the wild type in that only the initial inoculum would grow properly and dilutions would not reach noteworthy OD values as opposed to the wild type where the  $10^{-2}$  dilution would show OD values around 1 after 16 h of growth and thus could be used to inoculate the bioreactors. It is interesting to note that in all strains grown under the same conditions and reported here *C. pasteurianum*  $\Delta rex$  (Section 4.2.3.2), *C. pasteurianum*  $\Delta hyd$  (Section 4.2.4.1) and the *C. pasteurianum* wild type all showed a sharp drop in redox-potential of the culture (data not shown) directly after inoculation and then a slow decrease of the potential to final fermentation potentials of about -500 mV. The sharp drop which is due to the reduction state of the inoculum can be observed in the *dhaBCE* deletion strain as well but after that it is not able to reduce the medium further.

It was then decided to grow the strain in complex medium in serum bottles instead. In 2xYT with glycerol all three tested strains (*C. pasteurianum*  $\Delta dhaBCE$ , *C. pasteurianum*  $\Omega dhaBCE$  and the wild type) show a similar growth-profile. Growth rates are similar (Table 4.4) and pH drops faster in the wild type than in the other two strains. After 24 h the pH reaches a minimum of  $4.4 \pm 0.0$  in the deletion and  $4.5 \pm 0.0$  in the other strains and increases slightly after that. Growth reaches stationary phase after 24 h and increases over the whole fermentation and is maximal at the last time point with  $4.9 \pm 0.2$  in  $\Delta dhaBCE$  and  $5.0 \pm 0.0$  in the other strains. Glycerol is taken up very fast over the first 24 h of the fermentation and uptake carries on until 72 h after which it ceases. The tested strains cannot use all glycerol provided and the deletion strain can use the least with only  $279 \pm 1$  mM, when the complementation strain uses  $300 \pm 27$  mM and the wild type  $293 \pm 12$  mM.

The deletion strain produces more lactate than the wild type and the complementation strain brings levels back to wild type levels. All three strains start producing lactate only after 9 h of fermentation and reach maximal concentration after 24 h after which the concentration slightly drops. Acetate levels on the other hand are decreased in the deletion mutant and acetate production starts for all strains directly from the start of the fermentation when already a small concentration of acetate is found in the medium. Concentration reaches a maximum at 24 h after which acetate is re-assimilated. The complementation does not perfectly correct the phenotype but levels are decreased when compared to the wild type.

Butyrate is produced from the start and reaches a maximum already after 9 h of fermentation after which small amounts of butyrate are re-assimilated. Maximal concentrations of butyrate reached are slightly higher in the wild type

strain than in the deletion or complementation strain but even equalise over the course of the fermentation.

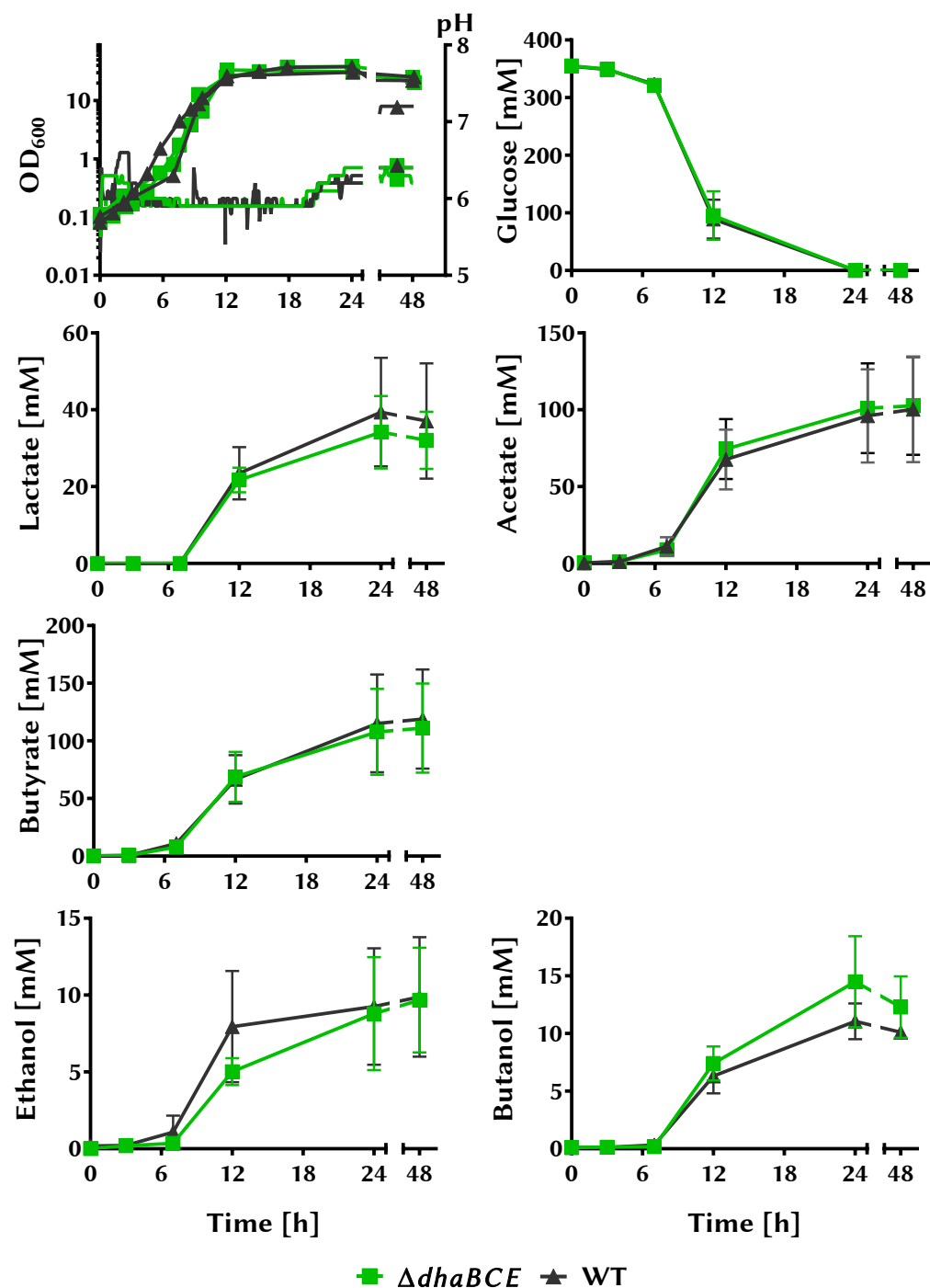
Solvent production is unchanged in the deletion strain when compared to the wild type and complementation. Ethanol reaches its maximal concentration after 24 h with  $15.5\pm 0.3$  mM,  $15.9\pm 0.8$  mM and  $15.8\pm 1.0$  mM in  $\Delta dhaBCE$ ,  $\Omega dhaBCE$  and wild type, respectively. Butanol concentrations after the same time reach  $96.9\pm 0.8$  mM,  $103.2\pm 4.0$  mM and  $103.8\pm 4.6$  mM in the respective strains.

Finally, as expected 1,3-PDO production is abolished in the *dhaBCE* mutant as the pathways to 1,3-PDO are no longer present. The strain is, thus, not able to produce 1,3-PDO via a different route. 1,3-PDO production in this medium is reduced also in the wild type (0.04 mM/mM glycerol) compared to levels in minimal medium (0.16 mM/mM glycerol, Section 4.2.4.1) and the complementation reintroduces 1,3-PDO production although at lower levels than the wild type.

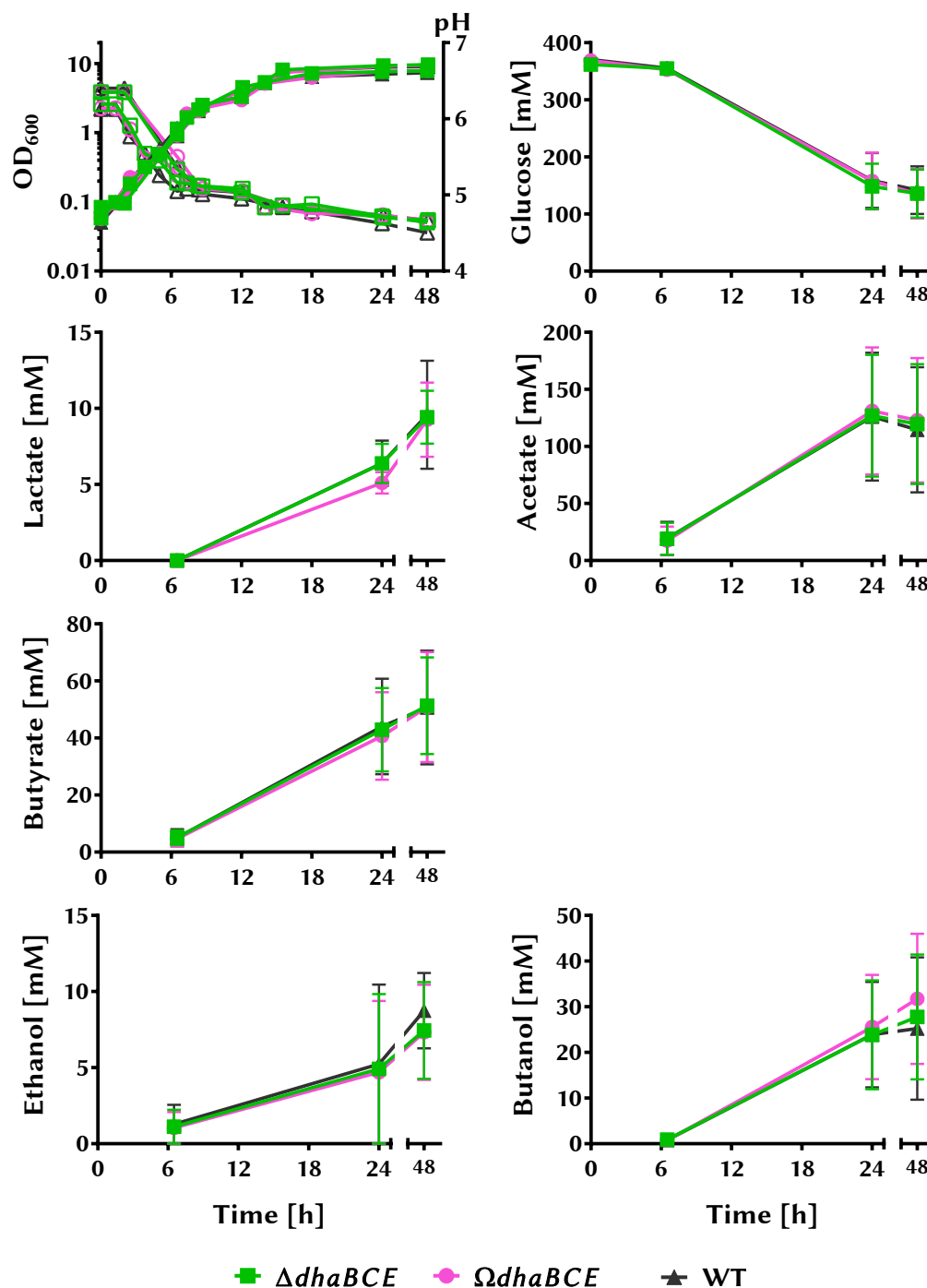
**Glucose Fermentation in Bioreactor** Growth on glucose of *C. pasteurianum*  $\Delta dhaBCE$  is expected to be very similar to wild type growth as 1,3-PDO is only produced directly from glycerol and not from glucose. Batch growth in the bioreactor with minimal Biebl medium and 6 % glucose shows just that. The strains  $\Delta dhaBCE$  and wild type grow very similar with similar growth rate and reach comparable maximal OD. Also the pH follows the same pattern with only slight increase after about 20 h of growth. Fluctuations in pH can be observed in the first *C. pasteurianum*  $\Delta dhaBCE$  fermentation and in the second wild type fermentation which are due to connection faults of the electrodes. pH was measured externally (data not shown) and the fluctuations measured online were not found to have an influence on the fermentation.

In both strains all of the glucose was utilised after 24 h and it seems that the onset of pH increase coincides with depletion of the carbon source. Lactate production did not start until after 7 h and reached a maximum after 24 h which was very similar in both strains. Both acetate and butyrate started to be produced three hours into the fermentation and maximum concentrations of both compounds were reached after 24 h which in both strains again were very similar. No re-assimilation after this time of any of the two acids can be observed in neither wild type nor mutant.

Solventogenesis seems to start a bit later than acetate and butyrate production with ethanol being produced before butanol production starts. Butanol then reaches a maximum concentration after 24 h after which concentration decreases due to evaporation. The same would be expected for ethanol, how-



**Figure 4.15** – Bioreactor batch growth and product profile of the *dhaBCE* deletion strain *C. pasteurianum*  $\Delta dhaBCE$  and the wild type (WT) strain in two individual fermentations. Strains were grown in Biebl medium with 6 % glucose. OD<sub>600</sub> was measured manually and pH was measured with an automatic internal probe (lines with single symbols). For clarity, OD and pH is given separately for each replicate. The product formation is given as mean with error-bars indicating the range of the two samples.



**Figure 4.16** – Batch growth in serum bottles of the *dhaBCE* deletion strain *C. pasteurianum*  $\Delta dhaBCE$ , its complementation *C. pasteurianum*  $\Omega dhaBCE$  and the wild type (WT) strain. Strains were grown in two biological replicates in Biebl medium with 6 % glucose. Open symbols are used for pH. Data is given as mean of the two replicates with error-bars indicating the range of the samples.

ever, ethanol concentration increases slightly over the whole 48 h. Together with the ethanol that is expected to be lost to evaporation this would make for a considerable ethanol increase long after the rest of the fermentation seemingly has stopped. Ethanol production is faster in the wild type and the wild type produces slightly less butanol but this is down to batch to batch differences which is also expected to be seen in different wild type batches (Biebl, 2001; Taconi et al., 2009).

***dhaBCE* Complementation Analysis** As there was no difference expected between the wild type and the *dhaBCE* deletion mutant when grown on glucose and indeed no difference could be seen in bioreactors, a complementation analysis was done only to confirm the findings. The analysis was done in serum bottles in Biebl medium with 6 % glucose. No difference in growth or pH could be observed between all three strains tested (*C. pasteurianum*  $\Delta$ *dhaBCE*, *C. pasteurianum*  $\Omega$ *dhaBCE* and wild type). Glucose uptake was the same in all strains and not all glucose could be used. Only just over 60 % of the initial carbon source was used on average.

Acid formation of lactate, acetate and butyrate is very similar for all strains and happens over the whole time scale of the fermentation for lactate and butyrate. Lactate production starts later than acetate and butyrate production and acetate production ceases after 24 h and it even seems that it is re-assimilated, however within the margin of variance.

Ethanol is produced earlier than butanol and is produced over the whole time scale of the fermentation with the same rate, whereas butanol production starts only after 6.5 h and the production rate decreases after 24 h.

#### 4.2.6 1,2-Propanediol Analysis

During this work, Pyne et al. (2016c) reported a novel pathway in *C. pasteurianum* which allows the organism to produce 1,2-propanediol (1,2-PDO) from dihydroxyacetone-phosphate via methylglyoxal and either hydroxyacetone or lactaldehyde to 1,2-PDO. To see if 1,2-PDO is detected as well in the fermentations reported here, an HPLC was run with increasing concentrations of 1,2-PDO and 50 mM of valerate as internal standard (Figure 4.17-A). From this a standard curve was drawn which was used to estimate 1,2-PDO concentrations based on detected areas in previously run HPLC analysis. The HPLC spectrum of all fermentations was analysed for a 1,2-PDO peak and the compound was found to be produced in very small quantities in all strains and under most conditions. The highest concentration detected was of 3.3 mM (0.25 g/l) in

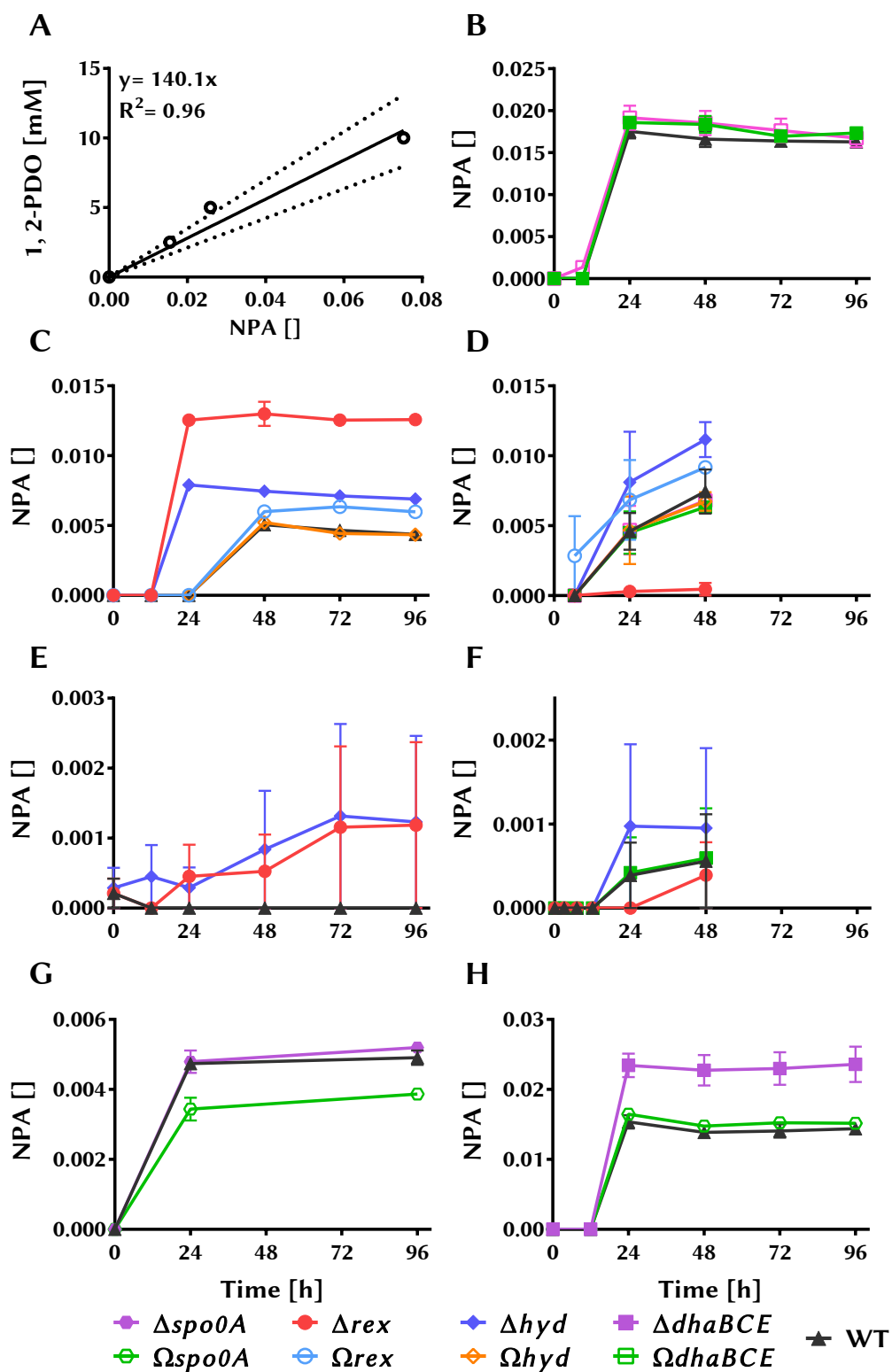


Figure 4.17 – 1,2-propanediol analysis in mutants and wild type. Full description see opposite page.

**Figure 4.17** – 1,2-propanediol production of mutants and wild type. Data is given as normalised HPLC peak area (NPA); the area of the 1,2-PDO peak divided by the area of the peak of the internal standard (valerate) and presented as mean±SEM.

A) Externally run standard curve of 1,2-propanediol (black line) used to estimate amounts in the text with the 95 % confidence bands (dashed lines) and measured samples (empty circles) (n=1).

B) *dhaBCE* mutant and complement strain grown on 2xYT with glycerol in serum bottles (n=3).

C) *rex* and *hyd* mutants and complements grown in Biebl medium with 1 g/l yeast extract and glycerol in serum bottles.

D) *rex*, *hyd* and *dhaBCE* mutants and complements grown in Biebl medium with glucose in serum bottles.

E) *rex*, *hyd* mutants grown in bioreactor in Biebl medium with 1 g/l yeast extract and glycerol.

F) *rex*, *hyd* and *dhaBCE* mutants grown in bioreactor in Biebl medium with glucose.

G) *spo0A* mutant and complement grown in Biebl medium with 1 g/l yeast extract and glycerol in serum bottles.

H) *spo0A* mutant and complement grown in CGM with glycerol in serum bottles.

---

*C. pasteurianum*  $\Delta spo0A$  grown in CGM (Figure 4.17-H) which is in the order of magnitude of what Pyne et al. (2016c) found in the *dhaT* disruption strain. In *C. pasteurianum*  $\Delta dhaBCE$  fermentations in serum bottles with 2xYT on glycerol (Figure 4.17-B) no difference between the strains was found and between 2.5- 2.7 mM 1,2-PDO was produced. The *rex* and *hyd* deletion mutants started producing 1,2-PDO earlier than the wild type and the complements and to higher levels (Figure 4.17-C). The wild type produced 0.7 mM and the *rex* and *hyd* complementations 0.7 mM and 0.8 mM, respectively. *C. pasteurianum*  $\Delta rex$  produced 1.8 mM and *C. pasteurianum*  $\Delta hyd$  1.0 mM. On glucose (Figure 4.17-D) *C. pasteurianum*  $\Delta rex$  showed barely any production, whereas *C. pasteurianum*  $\Omega rex$  is over-compensating showing higher 1,2-PDO levels than the wild type. *C. pasteurianum*  $\Delta dhaBCE$  and its complementation produced similar levels of around 1.0 mM. *C. pasteurianum*  $\Delta hyd$  showed the highest 1,2-PDO production in this fermentation with 1.6 mM. Fermentations in bioreactors on both glycerol and glucose (Figure 4.17-E, F) gave very low levels of 1,2-PDO and variable between duplicates with one of the fermentations on each carbon source not yielding any 1,2-PDO. This shows that production of this compound is not sufficiently reproducible. Both *C. pasteurianum*  $\Delta spo0A$  fermentations in Biebl medium and CGM (Figure 4.17-G, H) gave good yields of 1,2-PDO with the CGM fermentation giving about 5 times more end product. In Biebl



medium the wild type and the *spo0A* deletion gave the same amount of product (0.7 mM) and the complementation strain produced less (0.5 mM). In CGM the deletion mutant produced 3.3 mM and the complementation and wild type strain produced 2.3 mM and 2.2 mM, respectively.

## 4.3 Discussion

This study is part of work presented in Schwarz et al. (2017). Three in frame deletion mutants of *Clostridium pasteurianum*, *rex*, *hyd* and *dhaBCE*, crucial genes in the fermentative pathway, as well as the deletion of *spo0A* were analysed for their solvent profiles on glycerol and glucose. As each deletion was created in a *pyrE* truncation background (Minton et al., 2016), all four mutants were complemented at chromosomal level concomitant with correction of the *pyrE* allele and restoration of uracil prototrophy. The fermentative capabilities of all strains is discussed and compared to existing studies.

### 4.3.1 Spo0A Sporulation Master Regulator

The sporulation master regulator Spo0A initiates the sporulation cascade in *B. subtilis* (Stragier and Losick, 1996) and is found playing important roles besides sporulation in most clostridial species (Heap et al., 2010; Huang et al., 2004; Harris et al., 2002). In this work the *spo0A* deletion mutant of *C. pasteurianum* (Schwarz et al., 2017) was analysed for its influence on the production of acids and solvents. A defined medium (Biebl medium) was used for fermentations which, due to its nature, increases reproducibility of fermentations between batches and is more cost effective on an industrial scale compared to complex medium. In this medium, acid production of the mutant was increased compared to the wild type, whereas solvent production was lower and especially butanol showing a  $25 \pm 0.84$  % decreased concentration in the mutant compared to wild type. Similar to this study, Sandoval et al. (2015) reported a random *C. pasteurianum* mutant with a deletion in *spo0A* (M150B) and an engineered mutant with a part deletion of *spo0A*. Both of these strains were reported to show increased butanol production in complex medium (CGM) fermentations, which contradicted data of the deletion strain produced in this study when fermented in defined medium. For this reason it was decided to analyse the mutant and its complementation strain in complex medium (CGM) as well to verify the findings of Sandoval et al. (2015).

*C. pasteurianum*  $\Delta spo0A$  showed increased lactate production and reduced ability to re-assimilate butyrate in complex and defined medium fermentations.

The increase in acid generation led to lower pH in  $\Delta spo0A$  as compared to the wild type. The strain does not acid crash as it is known in *C. acetobutylicum* in unbuffered medium (Maddox et al., 2000) or in degenerate strains (Cornillot et al., 1997). The mutant still produces solvents and is also able to re-assimilate acetate and thus increases pH of the medium in later stages of the fermentation. As expected, all strains, wild type and mutants, grow faster in complex (CGM) than in defined medium (Biebl medium) due to CGM being richer in carbon sources and vitamins. No changes of metabolites can be observed after 24 h in CGM whereas in Biebl medium there are still signs of acetate re-assimilation after the 24 h time point. It is interesting to note that more glycerol is consumed in Biebl medium than in CGM suggesting that higher concentrations of yeast extract inhibit glycerol uptake. This is likely due to the fact that yeast extract is easily available and offers less reduced carbon sources. However, *C. pasteurianum*  $\Delta spo0A$  does use more glycerol than the wild type in yeast extract rich CGM but not in Biebl medium with low concentration of yeast extract. The ability to use more glycerol in CGM might not be due to the strains superior capacity of glycerol use, but rather its inability to use certain compounds in the yeast extract and thus having to use more glycerol to grow. Another explanation could be that compounds in yeast extract trigger glycerol uptake mechanisms in the *spo0A* deletion strain which are not active or not as active in the wild type. In both cases these compounds are not, or only very limited, available in Biebl medium and therefore no difference in growth between the mutant and the wild type can be observed. Finally, the lower butanol yield of the *spo0A* mutant in Biebl medium (0.29 M/M) compared to the wild type (0.35 M/M) is most likely due to the inability of the mutant to re-assimilate butyrate. The difference in butyrate re-assimilation capability of the mutant between CGM and Biebl medium might explain the differences in solvent generation observed.

Sandoval et al. (2015) reported similar product patterns in a mutant obtained by chemical mutagenesis (M150B) using both crude and pure glycerol as substrate. For the case of pure glycerol, 75 % increased glycerol uptake, 80 % more butanol and 40 % less 1,3-PDO was found to be produced by the M150B. Similar results were presented for the engineered strain ( $\Delta Spo0A$ .Sandoval), obtained by deletion of the promoter region and parts of *spo0A*, for which data is only available on 100 g/l crude glycerol and no glycerol uptake data is given. In this strain the butanol titre was increased and 1,3-PDO titre was decreased compared to the wild type. Compared to that, the fermentation of strain *C. pasteurianum*  $\Delta spo0A$  in this study gave a 20 % increase in glycerol uptake and an 18 % increase in final butanol titre but no difference was found in 1,3-PDO pro-

duction compared to the wild type in rich medium (CGM). However, a 22 % decrease in ethanol concentration was noticed. Ethanol was only marginally lower in the reports of [Sandoval et al. \(2015\)](#). Interestingly, when calculating yields from glycerol in [Sandoval et al. \(2015\)](#) chemically induced mutant, M150B, fermentation (compare Appendix Table A.2), the yield of butanol is the same in the mutant strain and in the wild type. The same is true for butanol yields in this study's *spo0A* mutant strain compared to the wild type. This indicates that the productivity of a *spo0A* inactivational strain and wild type *C. pasteurianum* is the same and the increased butanol concentration on CGM is only due to the mutant strain's ability to use more glycerol under these conditions. This was not considered by [Sandoval et al. \(2015\)](#).

Differences between fermentations reported by [Sandoval et al. \(2015\)](#) and in this study are likely to have arisen from culture conditions. There is no mention of fermentation vessel in [Sandoval et al. \(2015\)](#). I have noticed that fermentations in serum bottles with pressure build up differ largely to those in open vessels inside an anaerobic cabinet. However, as the main difference observed between the studies is the reduction of 1,3-PDO titres, it is more likely to have arisen from the medium composition with the only undefined part being yeast extract. Yeast extract is a non-defined substance and can differ largely from batch to batch and between producers. Furthermore, yeast extract supplies redox mediators riboflavin and nicotinamide ([Corrêa et al., 2009](#)). Differences in redox balance can have an influence on product pattern as is shown with other mutants in this chapter. Another explanation for the different fermentation patterns is the nature of their mutants with their random mutant having 82 differences on nucleotide level compared to the wild type genome used in this study ([Poehlein et al., 2015](#)). Furthermore, their *spo0A* deletion mutant is not a clean deletion, as it was in this work, but an insertional mutant with deletion of only the promoter region and 61 codons. [Sandoval et al. \(2015\)](#) did not present a complementation of their rational mutant and therefore no final proof was given that the phenotype arose from the mutation and not polar effects in the genome as described by [Ehsaan et al. \(2016\)](#). Here, a clean in frame deletion of *spo0A* was made and its solvent production was analysed. The mutant's phenotype was then reverted back to wild-type by complementing *spo0A* at the *pyrE* locus at chromosomal level. All of these make this study a preferred case concerning the function of *spo0A* than the one previously reported by [Sandoval et al. \(2015\)](#).

I showed that the *C. pasteurianum* Spo0A master regulator acts similarly on acidogenesis as it does in *C. acetobutylicum*, with a knock-out increasing acid production in both organisms ([Harris et al., 2002](#)). However, opposite

to Spo0A's regulatory action in *C. acetobutylicum* and *C. beijerinckii* where a knock-out abolishes solvent production (Harris et al., 2002; Ravagnani et al., 2000), it does not seem to control butanol levels in the same way in *C. pasteurianum*. In *C. pasteurianum* only a minor decrease in butanol levels was detected in Biebl medium and butanol production was increased in CGM in the *spo0A* deletion strain. With the comparison between a rich medium (CGM) and a medium low in yeast extract concentration, this study showed that changes in glycerol uptake in *C. pasteurianum*  $\Delta spo0A$  are only observed with high concentrations of yeast extract implying that, to some degree Spo0A regulation is dependent on compounds therein. This effect is further shown in reduced ethanol and 1,3-PDO yields found in rich medium as opposed to Biebl medium where yields of these compounds were unchanged. It would be interesting to see which compounds in yeast extract are responsible for this effect. Finally, it was shown that the clean deletion mutant has a distinct phenotype in fermentation which can be reversed to wild type by complementing the gene at the *pyrE* locus.

### 4.3.2 Rex Regulator

The Rex redox regulator has become a target of interest in metabolic engineering of industrially relevant organisms. Most work to date concentrated on Rex recognition sites and the mechanism by which Rex acts on gene expression (Brekasis and Paget, 2003; Anraku and Gennis, 1987; Gyan et al., 2006; Ravcheev et al., 2012; Hu et al., 2016; Zhang et al., 2014). Two studies reported physiological effects of *rex* inactivations in *C. acetobutylicum* and *Staphylococcus aureus* (Wietzke and Bahl, 2012; Pagels et al., 2010). In this study the presence of a *rex* gene in the *C. pasteurianum* genome was confirmed by homology to the *rex* gene in *C. acetobutylicum* and it was shown that it possesses the reported conserved DNA binding regions (Sickmier et al., 2005). *In silico* work further revealed a more stringent recognition sequence for *C. pasteurianum*, 5'-TTGTAAWNNNTTAACAA, than the one reported by Ravcheev et al. (2012) for bacteria in general. Based on this recognition sequence 47 genes or operons were predicted to interact with Rex, many of which are found in the fermentative pathway of *C. pasteurianum*. Since a high ratio of NADH/NAD<sup>+</sup> inhibits the binding of Rex to DNA, thus de-repressing expression of the target gene, Rex-binding sites were expected to be found upstream of enzymes utilising NADH. This was true for the alcohol dehydrogenases putatively involved in butanol and ethanol production, indirectly the thiolases that channel carbon into the reductive part of the fermentative pathway, the *bcs* operon which uses

2 NADH to reduce acetyl-CoA to butyryl-CoA, and the reductive part of the *dha* operon. Based on the same argument, being NADH reducing steps, regulatory sequences were expected to be found upstream of *ldh* and *hyd*, which was true in the related species, *C. acetobutylicum* and *C. kluyveri* (Wietzke and Bahl, 2012; Hu et al., 2016), but none was found in *C. pasteurianum*. A putative Rex binding sequence upstream of *pta* was detected which was not reported before in other bacterial species. As Rex is also involved in pathways of ATP synthesis (Pagels et al., 2010) this *pta* regulatory sequence is a target of interest for future research.

Based on these observations, these Rex regulated pathways should be up-regulated in a *rex* deletion mutant even when NADH/NAD<sup>+</sup> ratios are low and thus more of the respective products are expected. In bioreactor growth on glycerol this was found to be true for butanol and ethanol, both showed enhanced production when compared to the wild type. 1,3-PDO production, on the other hand, is reduced which might be a result of carbon being diverted by a more active oxidative part of the glycerol metabolism. The main ATP producing pathways to acids acetate and butyrate are down-regulated in the mutant; even though a Rex recognition sequence upstream of *pta* was found suggesting up-regulation of acetate production in absence of Rex. On the other hand, more lactate was found to be produced without a Rex recognition sequence upstream of *ldh*. A similar picture to the glycerol fermentation is found when growing the strain on glucose but much more acid and less solvent is produced with this substrate. When grown on glucose the mutant is able to use all of the substrate whereas it cannot use glycerol entirely. Compared to all the mutant strains of *C. pasteurianum* grown in this study (*spo0A*, *hyd*, *dhaBCE*) the *rex* mutant produced the highest concentration of butanol, which could indicate product inhibition stopping the strain from using glycerol completely. This fits well with a butanol toxicity assay done in rich medium (2xYTG) showing a concentration between 0.87 and 1.17 % to be toxic to a well growing culture of *C. pasteurianum*-H1 (Figure A.3). However, when grown in bottles, a higher maximal butanol concentration than in the bioreactor was found. This suggests that *C. pasteurianum* can grow with higher butanol concentrations and Sabra et al. (2014) found *C. pasteurianum* DSM 525 to produce 21 g/l butanol in dual substrate batch culture which is nearly double the amount reached in this study. This does not rule product inhibition out as the reason for the *rex* deletion mutants inability to use all of the substrate, as already a slight difference in culture condition can have a big influence on butanol resistance.

Since Rex is balancing redox levels in the cell, one would expect that its absence will have consequences on redox and energy equivalence which can

lead to impaired growth. However, this study shows this not to be the case. Growth rate of *C. pasteurianum*  $\Delta rex$  is the same as it is in the wild type and the concentration of butanol is increased despite the reduced ability to take up glycerol. Furthermore, *C. pasteurianum*  $\Delta rex$  has 50.4 % decreased by-product and 43.0 % increased butanol production levels which is in accordance with the *C. acetobutylicum* *rex* mutant analysed by [Wietzke and Bahl \(2012\)](#) showing increased alcohol levels. With 1,3-PDO production reduced by 53.1 % the butanol selectivity in the *C. pasteurianum* mutant is vastly superior to the wild type and butanol production levels make up for 51.6 % of total solvents as opposed to 28.9 % in the wild type. With these superior production characteristics, *C. pasteurianum*  $\Delta rex$  should be considered the obvious target for metabolic engineering in *C. pasteurianum* for industrial bioprocess optimisation.

### 4.3.3 Hydrogenase Deletion

In *C. pasteurianum* it was possible to delete the putative main hydrogenase CLPA\_c00280 which shows 71% identity to the well described *C. acetobutylicum* hydrogenase *hydA* ([Santangelo et al., 1995](#)). An earlier attempt of disruption of *hydA* in *C. acetobutylicum* was unsuccessful ([Cooksley et al., 2012](#)) and down-regulation of the *C. pasteurianum* *hydA* gene was published recently ([Pyne et al., 2015](#)). A deletion of the hydrogenase gene was thought impossible due to it being a sink for excess reducing power in the fermentative pathway. However, [Schwarz et al. \(2017\)](#) were able to delete the gene in *C. pasteurianum* but H<sub>2</sub> measurements did not show impaired hydrogen production when compared to the wild type (private communication) which suggested other hydrogenases were able to maintain hydrogen generation of the cell. A homology search on protein level confirmed two more [Fe]-hydrogenases CLPA\_c33960 and CLPA\_c37830 (*hydA*, according to [Poehlein et al. \(2015\)](#)) and a [NiFe]-hydrogenase (CLPA\_c07060/70) in the genome as outlined above (Section 4.1.3). Results obtained in this study clearly show that the hydrogenase deleted for this report (CLPA\_c00280) is tightly involved with the fermentative pathway in *C. pasteurianum*.

The solventogenic pathway in clostridia accumulates reduced ferredoxin (fd-red) when producing acetyl-CoA from pyruvate via the pyruvate-ferredoxin oxidoreductase. Fd-red is then oxidised again to fd-ox via the ferredoxin-NAD reductase thus producing NADH which in turn is oxidised back to NAD<sup>+</sup> via production of solvents, butyrate, lactate or production of H<sub>2</sub> by the [Fe]-hydrogenase (*hydA*) (Figure 1.2). Thus, hypothetically the deletion of *hyd* should lead to increased intracellular NADH which has to be consumed in other NADH ox-



**Table 4.6** – Product formation comparison between different *C. pasteurianum* hydrogenase studies. Fold change indicates the product ratio mutant/wild type. In the study by Pyne et al. (2015) the hydrogenase was knocked down with antisense RNA. In the study by Dabrock et al. (1992) all [FeFe]-hydrogenases were inhibited by 15 % carbon monoxide (CO).

Reference Conditions	Pyne et al. (2015)			Dabrock et al. (1992)		
	Shake flask fermentation Biebl medium 40 g/l glycerol, initial pH 6			Serum bottle fermentation in defined minimal medium, 20 g/l glucose, initial pH 7		
	mutant	wild type	Fold change	15 % CO	0 % CO	Fold change
Fermentation time [h]	43	52.5	0.82	27	27	1
OD <sub>600</sub> []	10.3	7.64	1.35	2.4	4.2	0.57
Substrate used [g/l]	37.7	38.6	0.98	5.95	7.93	0.75
Lactate [g/l]				1.02	0.18	5.65
Acetate [g/l]	0.9	0.38	2.37	0.77	1.30	0.59
Butyrate [g/l]	0.21	0.15	1.40	0.52	2.61	0.20
1,3-propanediol [g/l]	2.3	3.3	0.70			
Ethanol [g/l]	2	1.6	1.25	0.37	0.09	4.00
Butanol [g/l]	10	8.9	1.12	0.93	0.15	6.25
Total acids [g/l]	1.1	0.53	2.08	2.31	4.09	0.56

idising pathways and lead to an increase in solvents, butyrate and lactate.

In bioreactor growth of *C. pasteurianum*  $\Delta hyd$  on glycerol dramatic changes in product formation were found. While under the chosen conditions 1,3-PDO was produced as major solvent in the wild type, this shifted in favour of butanol in *C. pasteurianum*  $\Delta hyd$  but still only accounting for 31 % of total solvents. The largest increase is ethanol, of which nearly 2.7 times more is produced by the *hyd* mutant.

The product profile of *C. pasteurianum*  $\Delta hyd$  in glycerol fermentation fits well with the expectation of NADH consuming pathways being upregulated. More ethanol (2.7 times) and butanol (1.1 times) are produced with both pathways being major NADH sinks. The same was observed by Pyne et al. (2015) in a knock down of *hydA* and Dabrock et al. (1992) by carbon monoxide inhibition of hydrogenases in *C. pasteurianum* (Table 4.6). Lactate production was increased considerably (5.5 times) which was also found in CO inhibited *C. acetobutylicum* (Datta and Zeikus, 1985) and *C. pasteurianum* wild type strains (Dabrock et al., 1992), using glucose as carbon source. Unfortunately, no lactate data can be found in Pyne et al. (2015), which would have been interesting to compare. In wild type *C. pasteurianum*, lactate is produced only in small amounts and the pathway is thought to be non-functional in *C. acetobutylicum* under normal conditions unless hydrogenases are inhibited or the cells are depleted of iron (Jones and Woods, 1986). Datta and Zeikus (1985) suggest that in

*C. acetobutylicum* lactate is produced under conditions with increased NADH pool and is then converted to acetate and butyrate. Here, however, only increased levels of lactate can be observed but no re-assimilation thereof.

Butyrate production of *C. pasteurianum*  $\Delta hyd$  was reduced 5 times compared to wild type, confirmed in a similar study by [Dabrock et al. \(1992\)](#). [Datta and Zeikus \(1985\)](#) find *C. acetobutylicum*, inhibited in hydrogenase activity (by CO), to re-assimilate added butyrate early in the fermentation which leads to overall less butyrate being produced. Butyrate production uses NADH only in the initial steps from acetyl-CoA to butyryl-CoA. Butyrate re-assimilation and consecutive butanol production allows 2 equivalents more NADH to be used. It would, therefore, make sense that butyrate is produced at a lower rate in the *hyd* mutant than in the wild type. Concomitant re-assimilation of butyrate would then lead to increased butanol levels and decreased butyrate levels.

Surprisingly, a 1.8 times increase in acetate production was observed in the hydrogenase deletion strain. Acetate production being an ATP producing step of the fermentative pathway and not related to NADH oxidation. As noted by [Pyne et al. \(2015\)](#) this correlates with the increased biomass which is 114 % higher in the mutant than the wild type judged by OD<sub>600</sub> and the shorter fermentation time judged by the earlier on-set in pH increase (after 18.7 h as opposed to 22.0 h). This was, on the other hand, not manifested in faster growth rate; contrarily, growth was slower in the mutant. Inhibition of all hydrogenases did not show the same pattern and acetate production was nearly halved for [Dabrock et al. \(1992\)](#) which suggests that only the inactivation of the fermentative hydrogenase deleted here channels carbon to increased ATP production and inhibition of all hydrogenases counteracts that effect.

Another unexpected observation was the decrease in 1,3-PDO production. With 1,3-PDO being a major redox pathway its production was expected to be increased. However, this study and reports by [Pyne et al. \(2015\)](#) found a 20 % and 30 % decrease in final concentrations, respectively. [Pyne et al. \(2015\)](#) explains this with 1,3-PDO levels varying widely between fermentations which was observed as well by others in the field ([Taconi et al., 2009](#); [Biebl, 2001](#)). This is further underpinned by the serum bottle fermentation in this study in which the  $\Delta hyd$  mutant's 1,3-PDO levels were 12 % increased and butanol levels halved as opposed to the wild type. It is also interesting to note that in opposition to [Biebl \(2001\)](#) statement that the production of reducing equivalents by biomass build up has to be equalled by production of 1,3-PDO to regenerate the pool, both this reports and [Pyne et al. \(2015\)](#) find that a higher biomass coincides with lower 1,3-PDO (*C. pasteurianum*  $\Delta hyd$  bioreactor fermentation, [Figure 4.10](#)) and vice versa (*C. pasteurianum*  $\Delta hyd$  serum bottle fermentation,



Figure 4.12). However, in the case of this study the postulated condition (Biebl, 2001) is met that more than 9 % glycerol is converted to 1,3-PDO which is not true for the fermentations done by Pyne et al. (2015). The discrepancy could be explained by the new insight that *C. pasteurianum* does also produce trace amounts of 1,2-PDO which regenerates the same amount of reducing equivalents as the reduction to 1,3-PDO. Again, this fits with observations made here that 1,2-PDO concentration is increased in the *hyd* deletion strain (Figure 4.17 and Section 4.3.5) suggesting a new way for the organisms to regenerate the intracellular NAD<sup>+</sup> pool in redox imbalanced strains.

In summary, the deletion of the *C. pasteurianum* hydrogenase *hyd* elevated solvent production with the major increase observed in ethanol but also reasonable improvement of butanol, and at the same time diminishing by-product formation of 1,3-PDO. Optimisation of fermentation conditions towards butanol production could further increase the efficacy of this strain for industrial applications. Furthermore, it would be interesting to delete other hydrogenases and see the effect of single deletions as well as combinations of deletions of the hydrogenases in this organism to demonstrate feasibility and the effect on solvent titres of such genetic alterations.

#### 4.3.4 Glycerol Dehydratase

Thought to be crucial for growth on glycerol as sole carbon source, 1,3-PDO is more reduced than biomass. Production of 1,3-PDO absorbs the excess reducing equivalents from biomass accumulation. 1,3-PDO production was, until recently, thought to be the only pathway in *C. pasteurianum* to give a net-regeneration of reducing equivalents. In a recent study it was found that *C. pasteurianum* also produces 1,2-propanediol (1,2-PDO) which in the proposed pathway leads to the same net regeneration of NAD<sup>+</sup> (Pyne et al., 2016c) (Section 4.3.5). This theory is strengthened by my findings of *C. pasteurianum* lacking *dhaBCE*, the genes encoding the glycerol dehydratase, to be only able to grow in rich medium with glycerol. The mutant did not grow in defined medium with glycerol as only carbon source, at least not under the conditions chosen. This suggests that the 1,2-PDO route is not enough to compensate for the loss of function of *dhaBCE*, reflected in the small amounts of 1,2-PDO detected (<0.23 g/l), which are similar to the wild type. The fact that *C. pasteurianum*  $\Delta$ *dhaBCE* only grows in rich medium seems to verify its inability to regenerate reducing equivalents without 1,3-PDO production and using reducing equivalents and alternative carbon from yeast extract and peptone instead.

This is the first report of a clean in frame *dhaBCE* deletion in any mi-

croorganism to date. Pyne et al. (2016c) attempted ClosTron disruption of the individual genes (*dhaA*, *dhaB*, *dhaC*) but failed to isolate integration mutants. They succeeded in isolating a *dhaT* insertional mutant but only by co-harboring an insertion in *speA*. A similar approach was taken by Ashok et al. (2011) who deleted *dhaT* in *Klebsiella pneumoniae* to channel carbon towards 3-hydroxypropionic acid (3-HP). In both studies 1,3-PDO production was not entirely abolished; in fact Ashok et al. (2011) found highly increased 1,3-PDO concentration as compared to the wild type under aerobic conditions with the strain failing to grow anaerobically. These studies suggest that multiple 1,3-propanediol dehydrogenases are active in *K. pneumoniae* and *C. pasteurianum*. Furthermore, Pyne et al. (2016c) report a novel 1,2-PDO pathway which is more active in the *dhaT* disrupted mutant than in the wild type (Section 4.3.5).

The *dhaBCE* clean deletion strain presented here is superior to the above mentioned mutants as it abolishes 1,3-PDO production completely. However, the other major fermentation products and solvent levels in particular were unchanged, which did not allow to identify an alternative recipient for the carbon saved from 1,3-PDO production. This is not surprising seeing that even wild type 1,3-PDO levels were very low when grown in rich medium (maximal 11.2 mM) as compared to defined medium (maximal 74 mM). This suggests that the 1,3-PDO pathway is not very active when alternative carbon sources are available and a distribution of such a small amount of carbon is difficult to track with the methods used in this study.

#### 4.3.5 1,2-Propanediol

Pyne et al. (2016c) found 1,2-propanediol (1,2-PDO) production in *C. pasteurianum* as mentioned above. With this novel insight into the *C. pasteurianum* pathway the HPLC chromatograms in this study were re-analysed. Amounts of 1,2-PDO were estimated with an external standard curve and comparisons between fermentations were done via normalised peak area. Maximal production levels of this metabolite were small with 3.3 mM (0.25 g/l) found in complex medium fermentation which could explain why the metabolite was not detected before. By comparing wild type levels of 1,2-PDO it becomes clear that higher concentrations of yeast extract increase product end concentrations as the fermentations in 2xYT and CGM yielded the highest 1,2-PDO concentrations. Carbon sources, glycerol or glucose, showed no influence on 1,2-PDO production in wild type *C. pasteurianum*, suggesting that the entry point of glucose into the central energy metabolism is not at glyceraldehyde-3-phosphate as suggested by Malaviya et al. (2012). Since carbon flux by the proposed

1,2-PDO pathway (Pyne et al., 2016c) has to go through dihydroxyacetone-phosphate, it is probable that carbon from glucose enters through fructose-bisphosphate aldolase (*fba*, CLPA\_c13070). Another explanation would be that glyceraldehyde-3-phosphate is converted via a bidirectional triose-phosphate isomerase to dihydroxyacetone-phosphate. This would allow the production of 1,2-PDO via the proposed pathway by Pyne et al. (2016c) from dihydroxyacetone-phosphate when grown on glucose. Deletions of *rex*, *hyd* and *spo0A* seem to increase 1,2-PDO production whereas deletion of *dhaBCE* does not have an influence. This stands opposed to findings by Pyne et al. (2016c) who found increased 1,2-PDO in the *dhaT* disruption mutant compared to the wild type. However, Pyne et al. (2016c) grew the *dhaT* disruption in Biebl medium whereas *C. pasteurianum*  $\Delta$ *dhaBCE* was grown in 2xYT and the higher yeast extract concentration might have changed product patterns. The deletion of the Rex regulator increases 1,2-PDO production suggesting regulation of this NADH consuming pathway. This is further emphasised by the finding of a regulator sequence upstream of gene *aldA* (CLPA\_c23760) suggested to catalyse the NADH consuming step methylglyoxal to hydroxyacetone as one of five possible aldo/keto reductases (Pyne et al., 2016c) and thus strongly suggesting this gene to encode the appropriate enzyme which could not be definitely determined by Pyne et al. (2016c). It has to be emphasised that levels of 1,2-PDO were very low in both this and the Pyne et al. (2016c) studies. Even though, Pyne et al. (2016c) reports a significant change in 1,2-PDO product titres between their *dhaT* deletion mutant and wild type, this was not confirmed here with *C. pasteurianum*  $\Delta$ *dhaBCE*. Therefore, further investigation of the 1,2-PDO pathway in *C. pasteurianum* is advised.

#### 4.3.6 Butyrate Re-assimilation

During this work it was noted that *C. pasteurianum* is able to re-assimilate acids without the production of acetone, as opposed to the widely accepted dogmatic notion in *C. acetobutylicum* that butyrate and acetate re-assimilation have to be coupled to acetone formation via *cftAB* (acetoacetyl-CoA:acetate/butyrate:CoA-transferase) (Jones and Woods, 1986). However, uptake could be accomplished through either the reverse reaction of the butyrate kinase (*buk*) and the phosphotransbutyrylase (*ptb*) as proposed by Hüsemann and Papoutsakis (1989) and Millat et al. (2014) or via a third mechanism that is not yet described as proposed by Millat et al. (2014). These findings support similar reports in a *C. acetobutylicum* ClosTron mutant lacking CoA-transferase activity (*ctfA*) (Millat et al., 2014), with insertions in *ptb* and *adc* (acetoacetate decarboxylase)

(Lehmann et al., 2012) and an insertional deletion mutant of the butyrate kinase (*buk*) (Desai et al., 1999). All of these mutants lack parts of the pathway responsible for acetone production with concomitant butyrate re-assimilation. Acetone was not produced in these mutants but butyrate was re-assimilated.

The fact that *C. pasteurianum* does not produce acetone under any condition tested here or elsewhere (Biebl, 2001; Sandoval et al., 2015; Pyne et al., 2016c) makes this organism an ideal target for investigation in *ctfAB* independent butyrate and acetate uptake mechanisms.

## 4.4 Conclusion

In the present study, a robust tool kit for the metabolic engineering of *C. pasteurianum* was developed. The *spo0A* gene was deleted and complemented as a proof of principle for the system. For the first time, in frame deletion mutants of pivotal genes involved in solvent production, namely *hyd* (hydrogenase), *rex* (Rex response regulator) and *dhaBCE* (glycerol dehydratase), were constructed. The phenotype analysis of these mutants has provided valuable data on the mechanisms involved in solvent production and will guide future engineering approaches in this organism. For the first time in *C. pasteurianum*, elimination of 1,3-PDO synthesis was presented and production of this metabolite was demonstrated to be essential for growth on glycerol as sole carbon source. Inactivation of both *rex* and *hyd* resulted in increase in *n*-butanol titres, representing the first steps towards improving the utilisation of *C. pasteurianum* as a chassis for the industrial production of this important chemical.

# Chapter 5

## Development of a *Mariner*-Transposon System in *Clostridium* spp.

Un tas de pierres cesse d'être un tas de pierres, dès qu'un seul homme le contemple avec, en lui, l'image d'une cathédrale.

---

Antoine de Saint-Exupéry, *Le Petit Prince*

### 5.1 Introduction

Transposable elements (TEs), or transposons, are DNA sequences which are able to move from one genome location to another in a host factor independent manner. They are abundant in the genomes of all organisms, prokaryotic and eukaryotic, and can in some instances comprise up to 80 % of the genome, as it is the case for maize (SanMiguel et al., 1996). Transposons are involved in disease development (in vertebrates) or deleterious phenotypes, when transposition into essential genes occurs (Pray, 2008). They are also involved in driving genome evolution by reshuffling of exons, double-strand break repair, translocation of genomic sequences and alteration of gene regulations (Pray, 2008). Although most of the transposon fragments investigated to date are inactive some are still active and some were reactivated from inactive genes (Ivics et al., 1997). There are two classes of TEs: class I, which function through a RNA-intermediate and subsequent reverse transcription (see also Section 1.3.3) and

---

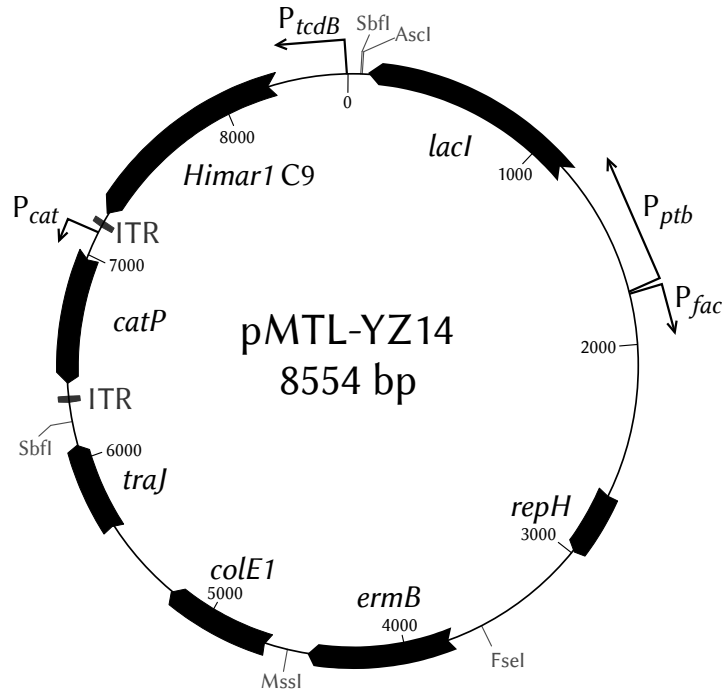
Parts of this chapter are published as Zhang et al. (2015)

class II, transposons which are DNA based and consist of a transposase gene between two inverted terminal repeats (ITRs). The latter predominantly move by a cut-and-paste mechanism (Munoz-Lopez and Garcia-Perez, 2010). Here the focus will be on class II TEs and the TC1/*mariner* superfamily in particular since a *mariner*-transposon is used in this work.

Class II TEs are flanked by ITRs of 9- 40 bp length which are recognised by the transposase in its cut and paste action. These TEs border flanking direct repeats (FDRs) which are not cut and pasted by the transposase but remain in the original location leaving a scar. (Pray, 2008) In the case of the *mariner*-transposon used in this study the ITRs are 5'-TAACAGGTTGGCTGATAAGTCCCCGGTATGACA and its inversion. The *mariner* FDR is a TA dinucleotide which marks the point of insertion in the genome and which is duplicated during the transposon insertion. (Figure 5.2, Cartman and Minton (2010); Zhang et al. (2015)) This last characteristic is especially interesting when working with genomes with low GC content as is the case with the genus *Clostridium*.

### 5.1.1 The *Clostridium* Transposon System

The ability of transposons to randomly insert into host genomes is often used in molecular genetics to isolate mutants and has been employed in several *Clostridium* species. Two approaches have been pursued based on the use of either conjugative or non-conjugative transposons. Conjugative transposons exhibit a circular, plasmid-like (without being able to replicate) intermediary structure and do not duplicate the target-site upon integration (Salyers et al., 1995). Their use, however, suffers from a number of drawbacks including their large size, low transposition efficiency, restricted host range, chromosomal 'hot-spots' (preferred integration sites as opposed to random insertions) and insertion of multiple copies which complicate correlation of phenotype to genotype (Schwarz et al., 2015). Such a conjugative transposon, Tn916, was employed in *C. botulinum* (Lin and Johnson, 1991), *C. perfringens* (Awad and Rood, 1997), *C. acetobutylicum* (Babb et al., 1993) and *C. difficile* (Hussain et al., 2010). However, in every case the mutants studied contained multiple insertions. More recently, the utilisation of the non-conjugative transposon system based on the *Himar1* C9 transposase, a hyperactive derivative of the *mariner*-transposon *Himar1* (Lampe et al., 1999), has been explored. When used in *C. difficile*, only a small percentage of mutants showing double insertions of the transposon were obtained. (Cartman and Minton, 2010). The same transposon was also found to be effective in *C. perfringens* (Liu et al., 2013).

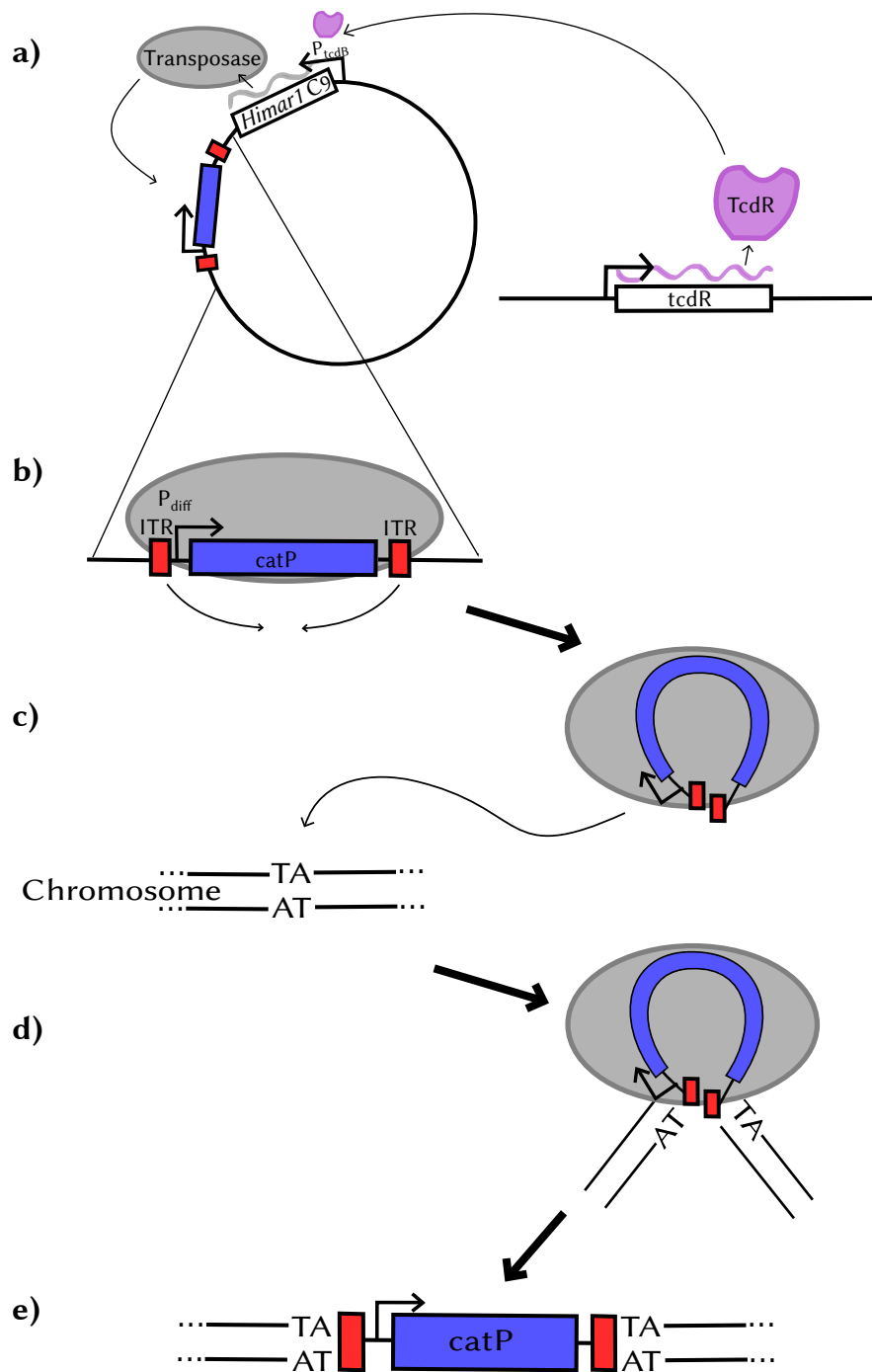


**Figure 5.1** – Plasmid pMTL-YZ14 with the transposon system used in clostridia developed by [Cartman and Minton \(2010\)](#) and refined by [Zhang et al. \(2015\)](#) with, clockwise: the inducible *lacI* system for putative over-expression of Gram-positive replicon *repH* which leads to plasmid loss (conditional replicon), the erythromycin resistance marker *ermB*, Gram-negative replicon *colE1*, conjugational transfer function *traJ*, the mini-transposon with the chloramphenicol resistant marker *catP* flanked by two inverted terminal repeats (ITR), the *Himar1 C9* transposase with the  $P_{tcdB}$  promoter which is only active when the  $\sigma$  factor TcdR is present in the host (see also [Figure 5.2](#)).

The *Clostridium* transposon plasmid, developed by [Cartman and Minton \(2010\)](#), was designated pMTL-SC1 carrying the *Himar1 C9* transposase gene under the control of the  $P_{tcdB}$  promoter which is separate from the mini-transposon element. The latter comprises a *catP* (encoding chloramphenicol acetyltransferase, bestowing resistance to thiamphenicol) gene flanked by ITRs ([Figure 5.1](#)). Expression of the transposase is under the transcriptional control of the promoter ( $P_{tcdB}$ ) of the *C. difficile* toxin B gene, *tcdB*. Transcription from the  $P_{tcdB}$  promoter is mediated by a specialized sigma factor, TcdR ([Figure 5.1](#)), which is absent in *E. coli*. The  $P_{tcdB}$  promoter is, therefore, relatively inactive in *E. coli* ([Moncrief et al., 1997](#)), a factor which prevents appreciable transposition events taking place in the donor strain before the plasmid is transferred to the clostridial recipient.

pMTL-SC1 bears a major drawback: the delivering vehicle for the mariner transposon, a segregationally unstable pseudo-suicide vector which requires





**Figure 5.2** – Mode of action of the *Clostridium* transposon system. a) The transposase *Himar1 C9* is expressed by the  $P_{tcdB}$  promoter when the *TcdR*  $\sigma$  factor is present and expressed. b) The transposase recognises the *ITRs* and binds to the mini-transposon (in this case *catP*) and initiates excision. c) The transposase-transposon-complex finds a target-site in the host chromosome. In transposons of the *mariner* family this is a TA-dinucleotide and a target capture complex is built. d) The transposase initiates a double strand break of the host chromosome which builds a strand transfer complex for insertion of the mini-transposon. e) The mini-transposon is inserted and the transposase dissociates. Figure after [Dornan et al. \(2015\)](#); [Lampe et al. \(1998\)](#)



a minimum of two passages of the recipient bacteria to eradicate the *Himar1* C9 transposase encoding plasmid. When considering generating large random mutant libraries and applying high-throughput mutagenesis strategies, the two additional passages required for screening using this system are very labour intensive and time consuming, prohibiting its implementation. This limitation however has been addressed in a follow up study by Zhang et al. (2015) in *C. acetobutylicum* and *C. sporogenes*. Zhang et al. (2015) presented a conditional replicon by which plasmid loss can be induced using  $\beta$ -D-1-thiogalactopyranoside (IPTG). This system, together with the transposase and mini-transposon from Cartman and Minton (2010), was combined in plasmid pMTL-YZ14.

### 5.1.2 Aim of This Study

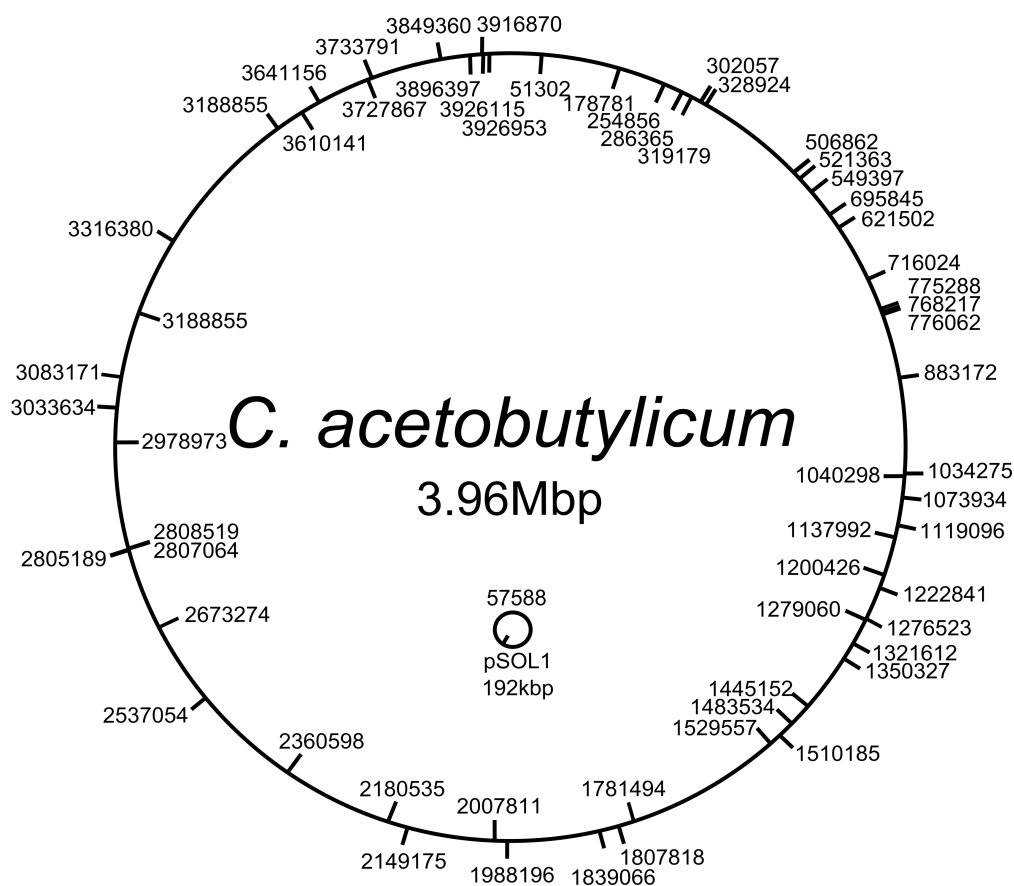
To apply forward genetics, this study aimed to adapt the *mariner*-transposon system (pMTL-YZ14) in *C. pasteurianum*. Due to initially low transformation efficiencies in this organism the system was exemplified in two established *Clostridium* species: the biofuel producer *C. acetobutylicum* and *C. sporogenes*, a strain of medical importance. The transposon system had to be tested for randomness of insertion as well as occurrence of multiple insertions. Furthermore, appropriate screening of both libraries resulted in the isolation of auxotrophic mutants as well as cells deficient in spore formation/germination. With a successful proof of principle in these organisms, this strategy is capable of being implemented in *C. pasteurianum*.

## 5.2 Results

### 5.2.1 Transposon Mutagenesis in *C. acetobutylicum*

#### 5.2.1.1 Assessment of Random Insertions

Prior to the initiation of the work described here, Ying Zhang had constructed the transposon delivery plasmid pMTL-YZ14 (Zhang et al., 2015) and established that its principle component parts were functional in *C. acetobutylicum*. Thus, the *tcdR* gene within the *C. acetobutylicum* recipient had been shown to be functional, and able to bring about expression of the transposase gene, while loss of the plasmid could be induced through addition of IPTG. In the following series of experiments the suitability of the system for generating a random transposon library was tested. Plasmid pMTL-YZ14 was transformed into *C. acetobutylicum* ATCC 824 and erythromycin resistant ( $\text{Em}^{\text{R}}$ ) colonies were



**Figure 5.3** – Mini-transposon insertion loci on *C. acetobutylicum* chromosome from 60 individual mutants. The insertions show random distribution around the chromosome without loci of preference. 61 insertions were noted due to the fact that one mutant carried a double-insertion.

pooled in PBS. Serial dilutions were plated to single colony on plates with IPTG and thiamphenicol (**Tm**) to promote plasmid loss. After 12 h to 16 h **Tm** resistant colonies were visible with a frequency of  $2.6 \pm 0.6 \times 10^{-4}$  cfu(**Tm**<sup>R</sup>)/cfu(total). Of these **Tm**<sup>R</sup> colonies 100 were patch plated onto agar media supplemented with either **Tm** or erythromycin (**Em**, plasmid backbone) to ascertain plasmid loss. In *C. acetobutylicum* 80 % of the colonies tested had lost the plasmid.

To establish the authenticity of the putative transposon insertion mutants, 60 **Tm**<sup>R</sup> and **Em**<sup>S</sup> colonies were randomly selected for analysis by inverse PCR. Since the mini-transposon was located on a circular DNA fragment, the outward facing primers amplified the (initially) circular region flanking the mini-transposon. The DNA fragments amplified were then gel purified and their nucleotide sequence determined by Sanger sequencing. The data obtained revealed the site of insertion of the mini-transposon element in each individual clone. Their location was mapped to the genome as illustrated in Figure 5.3.

**Table 5.1** – Transposon insertion sites in *spo/ger*<sup>-</sup> and auxotrophic mutants of *C. acetobutylicum*. <sup>a</sup> *C. acetobutylicum* is abbreviated to *C. a.*  
<sup>b</sup>Showing the chromosomal context of the transposon insertion with the *mariner*-transposon target dinucleotide (TA) in upper-case letters.

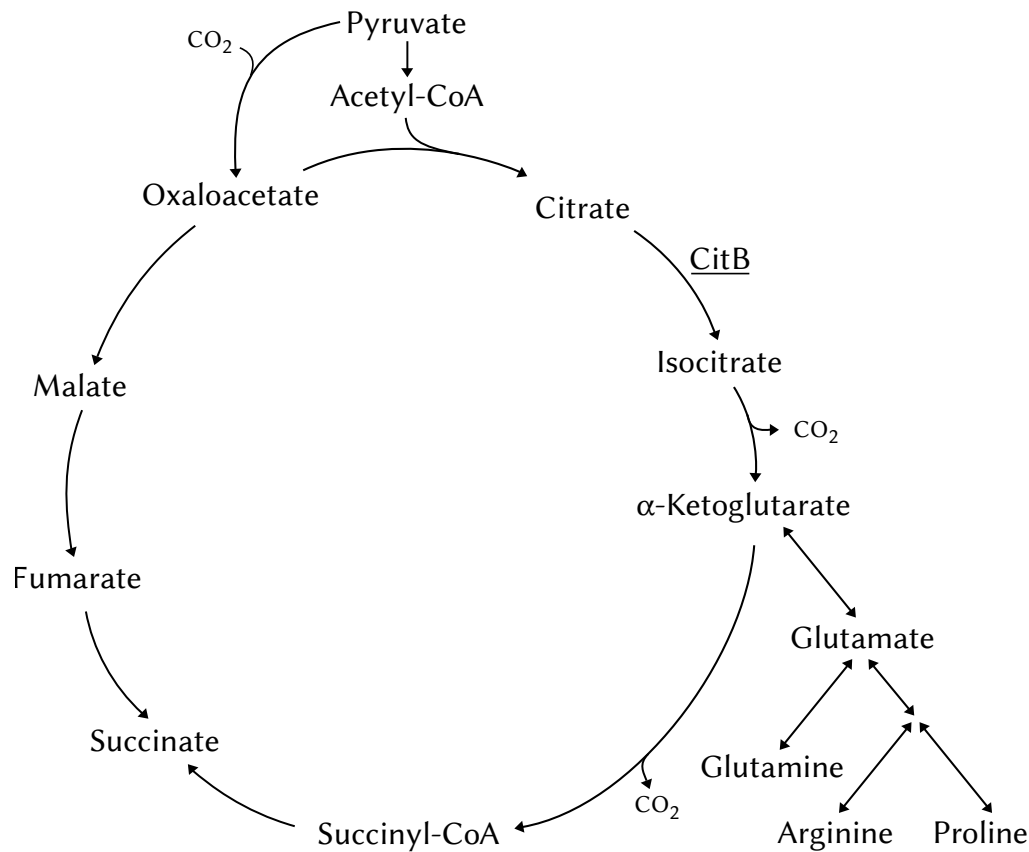
Strain <sup>a</sup>	Phenotype	Transposon insertion site <sup>b</sup>	Locus	ORF	Description
<i>C. a.</i>	<i>spo/ger</i> <sup>-</sup>	atactaaac- ttgatattaTA- Tn- TAAatataact- ttcttcttt	2740454	CA_c2631	hypothetical protein (L,D-transpeptidase catalytic domain protein; YkuD (spore protein of <i>Bacillus subtilis</i> ))
<i>C. a.</i>	auxotroph	ctaaatcat- ttgcaagaaTA- Tn- TAcacaaggct- aatctaadc	1119096	<i>citB</i> (CA_c0971)	aconitate hydratase / dehydrogenase, Catalyses the conversion of citrate to isocitrate

In the majority of cases (53 of 60) insertion took place within a coding region. Those open reading frames affected are listed in Table A.3. Of the 60 chosen clones characterized, only one carried a double-insertion of the mini-transposon. The majority of insertion were in the chromosome, with only one in the 192 kbp pSOL1 megaplasmid. In total, 34 insertions were located on the forward strand, with the other 27 having inserted into the reverse strand. In every case, insertions were flanked by a TA dinucleotide.

### 5.2.1.2 Isolation of Phenotypic Mutants

Having demonstrated that system was capable of generating random mutants, it was important to show that the transposon could be employed to isolate mutants affected in particular phenotypes. That is to say, use the transposon in a forward genetics approach. One particularly attractive target is the process of spore formation and germination, given the large number of genes and their products that participate in these processes. Mutants affected in these processes can be easily selected by their inability to form colony forming units (cfu) after heat shock. A second target is to screen for auxotrophy. Mutants that are no longer able to grow on simple minimal media, but are able to grow on media supplemented with, for instance amino acids, nucleotides or vitamins.

**Sporulation/Germination Deficient Mutants** were screened for by growing a representative library of transposon mutants of about 3000 individual clones in duplicate in CBM broth for one week to allow sporulation to occur. One of the duplicate cultures was then heat-shocked at 80 °C for 10 mins, whereas the other was left in the cabinet. A 5  $\mu$ l aliquot of each culture was



**Figure 5.4** – Putative tricarbalic acid cycle in anaerobic bacteria according to Amador-Noguez et al. (2010); Crown et al. (2011). The auxotroph *C. acetobutylicum* mutant showed a transposon insertion in the gene encoding aconitate hydratase responsible for conversion of citrate to isocitrate. Growth of the mutant was restored by addition of glutamate and partly by addition of glutamine to the growth medium. Growth was not restored with addition of citrate, isocitrate,  $\alpha$ -ketoglutarate, arginine or proline (see Table 5.2).

plated on CGM+Tm and clones that did not grow after heat treatment but did grow in the untreated culture were considered possible sporulation or germination deficient strains. Each of those mutants was tested with the same procedure in triplicate to confirm the spore deficient phenotype and one *spo/ger*<sup>-</sup> mutant was isolated. In this mutant the transposon disrupts the open reading frame (ORF) CA\_c2631 which is annotated as a hypothetical protein exhibiting 70 % homology to YkuD, a spore protein of *B. subtilis* (Table 5.1).

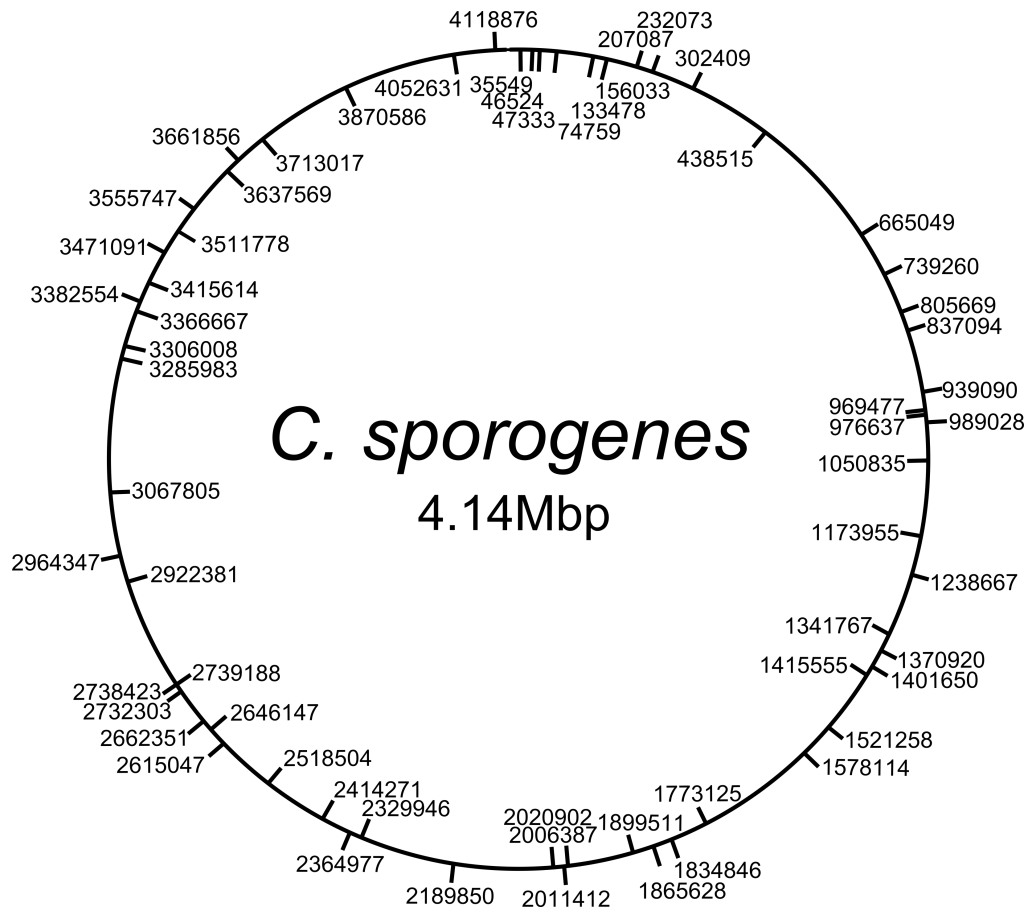
**Auxotroph Mutants** were screened for by replica plating the 3000 clone *C. acetobutylicum* transposon library onto rich medium (CGM) and minimal medium (P2) both supplemented with Tm. Plates were compared and colonies that grew on rich medium but failed to grow on minimal medium were consid-

**Table 5.2** – Growth of *C. acetobutylicum*-Tn-*citB* on medium supplemented with different tricarboic and amino acids. Growth on solid medium was visually observed and qualitatively noted as: - - indicating no growth, + indicating partially restored growth, ++ indicating good growth on wild type levels.

Supplementation	<i>C. a.ΩtcdR</i> -Tn- <i>citB</i>	<i>C. a.ΩtcdR</i>
Citrate	- -	++
Iso-citrate	- -	++
$\alpha$ -Ketoglutarate	- -	++
Glutamate	++	++
Glutamine	+	++
Proline	- -	++
Arginine	- -	++
Without supplementation	- -	++

ered to be auxotrophic. Such clones were tested in triplicate on both plates to confirm the auxotrophic phenotype. One *C. acetobutylicum* mutant was found with these characteristics and the transposon insertion was located in gene CA\_c0971 (*citB*) encoding aconitate hydratase (Table 5.1). In aerobic bacteria this enzyme is responsible for the conversion of citrate to isocitrate as part of the tricarboxylic acid (TCA) cycle (Figure 5.4). Isocitrate is further converted into the TCA cycle intermediate  $\alpha$ -ketoglutarate, and later into the non-TCA cycle metabolite glutamate, which can subsequently be converted into arginine, proline and glutamine.

To further characterise the mutant phenotype, the auxotroph mutant, together with *C. acetobutylicum*  $\Omega tcdR$  as a control, was plated onto P2 minimal agar, individually supplemented (1 mM) with the tricarboxylic and amino acids: citrate, iso-citrate,  $\alpha$ -ketoglutarate, glutamate, glutamine, proline and arginine (Table 5.2). Qualitative growth of the mutant was only fully restored to wild type by exogenous glutamate, whereas growth was only partially restored by glutamine supplementation. Intriguingly, the mutant remained unable to grow on the minimal media when supplemented with  $\alpha$ -ketoglutarate, which is suggested to be the product of the conversion of iso-citrate by isocitrate dehydrogenase (CA.c0972) and the substrate for the synthesis of glutamate (Amador-Noguez et al., 2010).



**Figure 5.5** – Mini-transposon insertion loci on *C. sporogenes* chromosome from 60 individual mutants. The insertions show random distribution around the chromosome without loci of preference. 62 insertions were noted due to the fact that two mutants carried double-insertions.

## 5.2.2 Transposon Mutagenesis in *C. sporogenes*

### 5.2.2.1 Assessment of Random Insertions

Having tested the individual components of plasmid pMTL-YZ14 (Zhang et al., 2015) in *C. sporogenes* the randomness of insertion was assessed by constructing a transposon mutant library. Thus, pMTL-YZ14 was transformed via electroporation into *C. sporogenes*  $\Omega$ *tcdR* and plated on TYG+Em plates after a six hour recovery period. Colonies were pooled in PBS and plated onto TYG+IPTG+Tm plate to induce plasmid loss. After 12 to 16 h Tm<sup>R</sup> colonies appeared with a frequency of  $3.2 (\pm 0.5) \times 10^{-4}$  cfu Tm<sup>R</sup>/total cfu. A total of 100 colonies were picked and plated onto TYG+Tm and TYG+Em plates to check for plasmid loss efficiency. All of the *C. sporogenes* mutants tested were Tm<sup>R</sup> and Em<sup>S</sup> indicating 100 % IPTG-induced plasmid loss. Of these mutants, 60 were selected and

**Table 5.3** – Transposon insertion sites in *spo/ger*<sup>-</sup> and auxotrophic mutants of *C. sporogenes*. <sup>a</sup> *C. sporogenes* is abbreviated to *C. s.*. <sup>b</sup>Showing the chromosomal context of the transposon insertion with the *mariner*-transposon target dinucleotide (TA) in upper-case letters.

Strain <sup>a</sup>	Phenotype	Transposon insertion site <sup>b</sup>	Locus	ORF	Description
<i>C. s.</i>	<i>spo/ger</i> <sup>-</sup>	attgcactc- taatggaaaTA- Tn- TAgataaatat- ctataatag	2565770	CS_2380	putative flavodoxin oxidoreductase
<i>C. s.</i>	<i>spo/ger</i> <sup>-</sup>	cagcgcata- gcatatctaTA- Tn- TAtctgtatcc- tttagatta	3418727	<i>spoVAC</i> (CS_3252)	stage V sporulation protein AC
<i>C. s.</i>	auxotroph	actgtaagt- tttactatgTA- Tn- TAagtaccagg- atctttata	1791546	<i>colA</i> (CS_1694)	microbial collagenase
<i>C. s.</i>	auxotroph	ctactacta- aattatttcTA- Tn- TAtatctacta- gtaactgga	2319668	CS_2157	KWG leptospira repeat protein
<i>C. s.</i>	auxotroph	attgtactg- ccatagaaaTA- Tn- TAAatataata- ctcaaacta	3233851	CS_3053	putative membrane protein

site of insertion determined by Sanger sequencing of amplified PCR fragments generated by the PCR procedure outlined in Chapter 2. The sequences were searched against the genome of *C. sporogenes* (Kubiak et al., 2015). The position of these insertions in the genome are shown in in Figure 5.5 and the features of the insertion sites detailed in Table A.4. In all cases, the mini-transposon was flanked by the expected TA-dinucleotide. In this case, 2 out of the 60 mutants contained double-insertions. Insertion of the mini-transposon had occurred within coding regions in 45 of the 60 mutants, with the remainder of the insertions being in intergenic regions.

### 5.2.2.2 Isolation of Phenotypic Mutants

Having demonstrated the random nature of the transposon in *C. sporogenes*, a similar screen to that undertaken with *C. acetobutylicum* was implemented, concentrating on screening for mutants affected in either sporulation/ germination or prototrophy.



**Sporulation/Germination Deficient Mutants** were screened for by growing approximately 6000 single colonies in liquid TYG+Tm for one week in duplicate after which one replicate was heat shocked at 80 °C for 10 minutes and the other kept in the anaerobic cabinet. Mutants that were unable to produce colonies after heat shock but able to grow if not heat shocked were putative sporulation/germination deficient mutants. These mutants were tested again in triplicate culture for defects in sporulation/germination with the same procedure. Two *spo/ger*<sup>-</sup> mutants were found with this strategy (Table 5.3). The insertions in these mutants were in CLSPO\_c23320, encoding a flavodoxin oxidoreductase, and CLSPO\_c31790, encoding for the stage V sporulation protein SpoVAD, respectively.

**Auxotroph Mutants** in *C. sporogenes* were screened for by patch plating the transposon mutant library on rich medium (TYG) and minimal medium (MACC) both supplemented with Tm. Mutants showing growth only on rich medium were restreaked again on rich and minimal medium for verification and finally three mutants were isolated showing an auxotroph phenotype. Inverse PCR and sequencing of the insertion flanking site revealed insertions in CLSPO\_c16750, encoding a collagenase (annotated as ColA), CLSPO\_c21170, coding for a hypothetical protein and CLSPO\_c29860, annotated as an ABC-2 family transporter protein (Table 5.3). All of the hits were subjected to BLAST searches to confirm the annotations. CLSPO\_c21170 did not return any known conserved domains and implication in growth support of this protein cannot be verified. Gene CLSPO\_c29860 encodes an ABC2-transporter involved in drug and carbohydrate export (Reizer et al., 1992). It is not clear why a defect in this gene should result in auxotrophy. Finally, ColA (CLSPO\_c16750) is not known to be essential for growth on minimal medium.

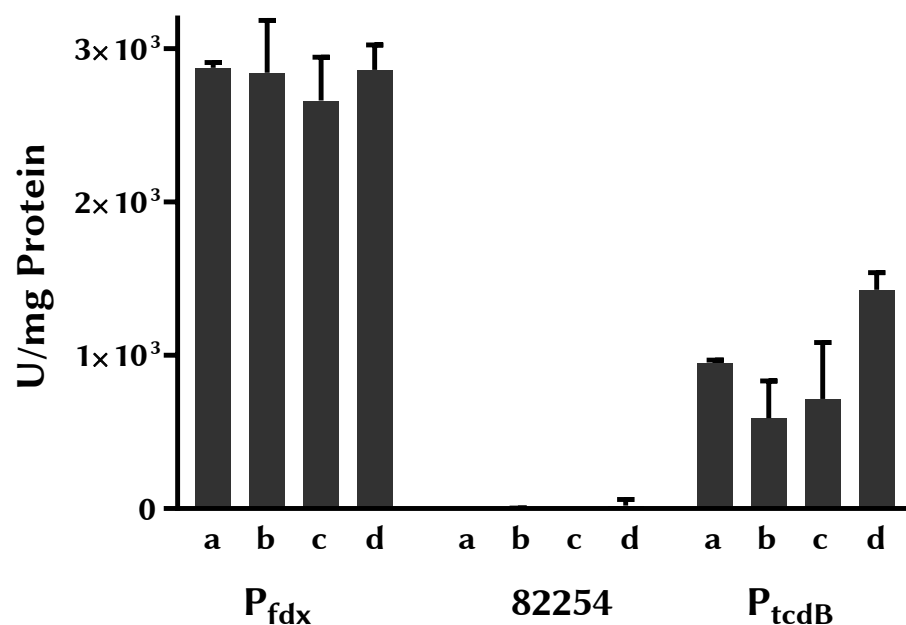
## 5.2.3 Assessment of the Transposon System in *C. pasteurianum*

### 5.2.3.1 *tcdR*-Dependent Expression

The usefulness of the  $\sigma$ -factor *tcdR* to drive expression of genes downstream of the  $P_{tcdB}$  promoter was tested in two clostridial species (*C. acetobutylicum* and *C. sporogenes*) (Zhang et al., 2015). Here, the effectiveness *tcdR*-dependent expression system was tested in *E. coli* and *C. pasteurianum*.

**$P_{tcdB}$  in *E. coli*** Transcription from the  $P_{tcdB}$  promoter should only occur if TcdR is present in the cell. Accordingly, as  $P_{tcdB}$  is reported to function inef-





**Figure 5.6** – **CAT** assay in *E. coli*, showing CatP activity when its gene is expressed from different promoters in different growth phases. Corresponding growth phases: a) mid-exponential-, b) late exponential-, c) early stationary-, d) late stationary stage. Growth stages are indicated in growth curve in Figure A.5.  $P_{fdx}$  is a known strong promoter in clostridia and serves as positive control. 82254 is the plasmid without a promoter in front of *catP* serving as negative control.  $P_{tcdB}$  has the *tcdB* promoter in front of *catP* which should be only active when the TcdR  $\sigma$ -factor is expressed in the host which is not the case in *E. coli*. Data is given as mean of six replicates with error bars presenting standard error.

fectively in *E. coli* (Moncrief et al., 1997), transposition should not occur in the intermediate host. However, as the clostridial recipient, in this case *C. pasteurianum*, carries a functional copy of the *tcdR* gene, transposition will take place when the plasmid is introduced into the clostridial recipient. To test this assumption, plasmids with *catP* downstream of different promoters ( $P_{fdx}$  [strong promoter, positive control], no promoter [pMTL82254, negative control],  $P_{tcdB}$ ) were transformed into *E. coli* and samples from a growth curve (Figure A.5) were taken at different growth-phases to test **CAT** activity with a colourimetric assay. Figure 5.6 shows that **CAT** is expressed constitutively from  $P_{tcdB}$ . The expression is about one third of the level of  $P_{fdx}$  known to be a strong promoter in clostridia. This experiment was done twice in biological triplicates as it contradicted previous results in our laboratory not showing expression of *catP* driven from  $P_{tcdB}$  (Yang, Gu, unpublished data) as well as the report of Moncrief et al. (1997) not showing any expression of their reporter gene (toxin A repeating units) from  $P_{tcdB}$  without a *tcdR* gene present in *E. coli*.

**P<sub>tcdB</sub> in *C. pasteurianum* Integration of *tcdR* Into the *C. pasteurianum* Genome by ACE.** To use Allele Coupled Exchange (ACE) in all hypertransformable *C. pasteurianum* strains (*C. pasteurianum*-H1, -H3 and -H4) *pyrE* truncation mutants were constructed, as described in Section 2.4.1.1. Strain *C. pasteurianum*-H1 $\Delta$ *pyrE* was constructed by Katrin Schwarz and Kamila Derecka. Here, equivalent *pyrE* mutant strains were constructed in *C. pasteurianum*-H3 and -H4, as described in Chapter 2. Each of the mutant strains generated were subjected to whole genome sequencing to ascertain whether any additional, unintended mutations (Single Nucleotide Polymorphisms (SNPs) or insertions/deletions (InDels)) had arisen. Table A.5 shows the outcome of mapping the paired-end Illumina reads obtained to the reference *C. pasteurianum* DSM 525 genome. In all *pyrE* truncation strains more than 98 % of reads were mapped and less than 6 % of the reads could not be paired. The average mapped read length was over 249 bp in most experiments. *C. pasteurianum*-H1  $\Delta$ *pyrE* 9 had a shorter average mapped read length with 244.72 bp.

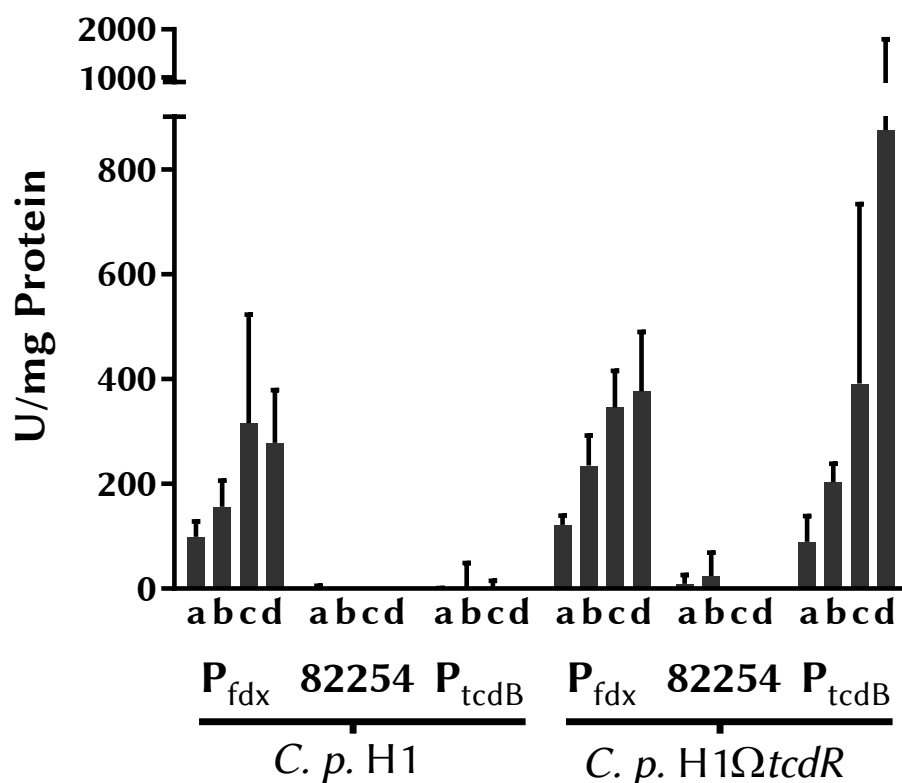
Variants are shown in Table 5.4 and for every strain newly emerged variants are listed and variants already present in the progenitor strain (Table 2.1) were neglected. A newly emerged SNP is noted in *C. pasteurianum*-H1  $\Delta$ *pyrE* 9 in the DNA topoisomerase II (*topB2*) changing an arginine into a histidine. Despite the presence of this SNP, phenotypic assays suggest no change in growth behaviour or fermentation pattern of this strains (Katrin Schwarz, personal communication and Chapter 4). Interestingly, the SNP at position 2883343 is absent in this strain as well as in strain *C. pasteurianum*-H4  $\Delta$ *pyrE* 5 even though it is present in the progenitor cells. The SNP is located in gene CLPA\_c26970 at the end of the gene and results in the introduction of a frameshift in the progenitor strains and correcting it in this clone. However, both the genome annotation and a confirmatory BLAST analysis indicate the gene encodes a as transposase. Transposases are non-essential to the host and as such prone to mutations. The same is true for the highly mutated region in strain *C. pasteurianum*-H3  $\Delta$ *pyrE* 4 (reference number 850742- 850795). The gene (CLPA\_c07760) encodes a hypothetical protein. The rate of mutation in this gene suggest it to be non-essential. However, there are no plans to use this strain for any future studies.

Based on these findings it was decided to use strains *C. pasteurianum*-H1  $\Delta$ *pyrE* 9 (as it was the first strain developed), *C. pasteurianum*-H3  $\Delta$ *pyrE* 3 and *C. pasteurianum*-H4  $\Delta$ *pyrE* 5 to be endowed with gene *tcdR*. Plasmid pMTL-KS12-*tcdR* was constructed as described in Section 2.3.1 and methylated before it was transformed into the *pyrE* truncation mutants. pMTL-KS12-*tcdR* is the ACE plasmid carrying *tcdR* with its native ribosomal binding site (RBS) be-

**Table 5.4** – SNPs and small structural variants called from read-mapping against *C. pasteurianum* DSM 525 reference genome (Poeblein et al., 2015). Only newly emerged SNPs are shown. \*SNP is called in progenitor but is not called in new strain.

Reference position	Reference	Allele	Count	Coverage	Frequency	Average quality	Amino acid change
<i>C. pa.</i> -H1 $\Delta$ <i>pyrE</i> 9							
2883343							present in H1, absent in this strain*
4219058	C	T	60	60	100	36.73	CLPA_c39530 ( <i>topB2</i> ): Arg193His
<i>C. pa.</i> -H1 $\Omega$ <i>tcdR</i> 16							
2522043	TTTA	-	16	22	72.73	34.43	CLPA_c23510: Ile138fs
2883343	T	-	22	22	100	29.82	CLPA_c26970 : Asn416fs
<i>C. pa.</i> -H3 $\Delta$ <i>pyrE</i> 3							
<i>C. pa.</i> -H3 $\Delta$ <i>pyrE</i> 4							
850742	AA	GT	58	58	100	36.35	CLPA_c07760: Phe122Thr
850749	CG	GA	57	57	100	37.08	CLPA_c07760: Ala120Pro
850762	-	C	43	51	84.31	37.77	CLPA_c07760:Pro116fs
850763	C	T	40	44	90.91	37.5	CLPA_c07760: Arg115Lys
850767	AA	G	48	50	96	33.08	CLPA_c07760: Ser114fs
850770	C	T	33	33	100	33.79	CLPA_c07760: Ala113Thr
850774	A	G	40	40	100	36.25	CLPA_c07760: silent
850776	C	A	40	40	100	35.33	CLPA_c07760: Ala111Ser
850792	A	G	35	68	51.47	36.51	CLPA_c07760: silent
850795	A	G	37	75	49.33	32.76	CLPA_c07760: silent
<i>C. pa.</i> -H3 $\Delta$ <i>pyrE</i> 5							
no SNP							
<i>C. pa.</i> -H4 $\Delta$ <i>pyrE</i> 3							
no SNP							
<i>C. pa.</i> -H4 $\Delta$ <i>pyrE</i> 5							
2883343							present in H4, absent in this strain*
<i>C. pa.</i> -H4 $\Delta$ <i>pyrE</i> 23							
no SNP							

tween the homology arms for insertion downstream of *pyrE* (Schwarz et al., 2017). After going through the ACE procedure one putative *pyrE* insertional clone of each strain was stored for future work. *C. pasteurianum*-H1  $\Omega$ *tcdR* 16 being the first strain developed was sent for next generation sequencing and read mapping was successful with nearly 98 % reads mapped of which nearly 95 % were mapped in pairs. The mapped reads had an average length of 249.69 bp (Table A.5). The read-mapping was used to call variants. One small deletion in gene CLPA\_c23510 and the recurrence of SNP at reference position 2883343 discussed above. Gene CLPA\_c23510 encodes a hypothetical protein with resemblance to a DNA-directed RNA polymerase  $\sigma$ -70 factor and the deletion introduces a frameshift 23 amino acids before the C-terminus. The frameshift

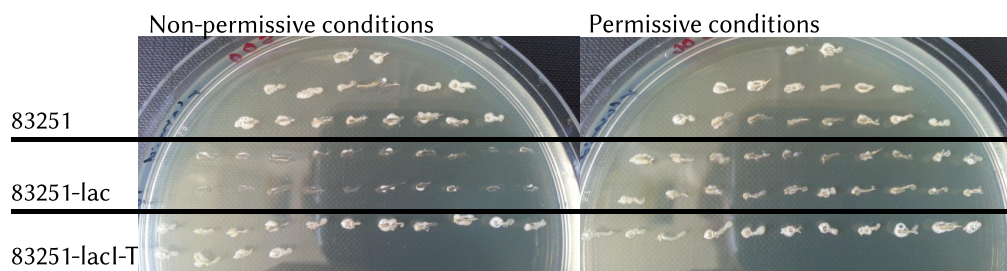


**Figure 5.7** – **CAT** assay of *C. pasteurianum*-H1 with and without *tcdR* (*C. pasteurianum*-H1 $\Omega$ *tcdR*) on the chromosome to test gene expression from  $P_{tcdB}$ . 82254 is the negative control plasmid with promoter-less *catP*,  $P_{fdx}$  is the plasmid with a strong promoter in front of *catP* and  $P_{tcdB}$  is the plasmid with the *tcdB* promoter in front of *catP*.  $P_{tcdB}$  should only be active if the TcdR  $\sigma$ -factor is present in the host. Expression at different growth phases is shown, as follows: a) mid-exponential-, b) late exponential-, c) early stationary-, d) late stationary stage; according to growth curve Figure A.6). Data is given as mean from six replicates with error bars indicating the standard error (SEM).

leads to alteration of the coding region down-stream and introduces a premature stop codon shortening the enzyme by 18 amino acids.

Strains *C. pasteurianum*-H3  $\Omega$ *tcdR* 5 and H4 $\Omega$ *tcdR* 1 were not sent for next generation sequencing as work has begun with *C. pasteurianum*-H1  $\Omega$ *tcdR* 16 and it was only later decided to switch strain since *C. pasteurianum*-H3 has less SNPs than -H1.

**Expression levels of the *tcdB*-Promoter.** To show expression levels of  $P_{tcdB}$  with and without *tcdR* on the chromosome of the host plasmids pMTL-82254- $P_{fdx}$ , pMTL82254 and plasmid pMTL82254- $P_{tcdB}$  were transformed into *C. pasteurianum*-H1 and *C. pasteurianum*-H1  $\Omega$ *tcdR* 16. All strains were grown in 2xYT with 6 % glucose in two biological replicates with three technical replicates each (Figure A.6) and samples were taken for a **CAT** assay at different



**Figure 5.8** – Test of conditional replicon under permissive and non-permissive conditions. *C. pasteurianum*-H1 endowed with indicated plasmids was replica plated from non-permissive- (IPTG) and permissive- (RCM only) conditions onto solid medium with thiamphenicol to test if the plasmid is retained. Plasmid 83251 serves as positive control, carrying the native *repH* replicon. Plasmid 83251-lac-T carries the conditional replicon with a terminator in front of *repH* which hinders putative over-expression and plasmid loss. Plasmid 83251-lac carries the functional conditional replicon which should induce plasmid loss during non-permissive conditions.

growth stages (mid-exponential, late exponential, early stationary and late stationary) to see if promoters are expressed constitutively. Finally, the **CAT** assay was performed by lysing the cells by sonication and analysing enzyme activity in the lysate as well as protein mass by colourimetric assays in a plate reader. The relative activities can be seen in Figure 5.7. It was evident from the low **CAT** activity observed in the wild type cell lines that the  $P_{tcdB}$  promoter was essentially inactive in the absence of TcdR. In contrast the level of **CAT** expression seen in *C. pasteurianum*-H1  $\Omega tcdR$  cells carrying pMTL82254-PtcdB was broadly equivalent or even higher to that seen in cells carrying plasmid pMTL82254-Pfdx, where expression was due to the strong  $P_{fdx}$  promoter.

The results prove that expression of a gene downstream of  $P_{tcdB}$  is reliant on the presence of TcdR in the *C. pasteurianum* host as it was shown in *C. acetobutylicum* and *C. sporogenes* (Zhang et al., 2015).

### 5.2.3.2 The Conditional Replicon

The conditional replicon consist of an inducible (*lacI*) promoter system in front of replicon *repH*. Induction with IPTG leads to plasmid loss through a currently unknown mechanism (Zhang et al., 2015). To test the functionality of the conditional replicon in *C. pasteurianum* plasmids pMTL83251 with the natural replicon *repH*, pMTL83251-lacI with the IPTG inducible *lacI* system in front of *repH* and pMTL83521-lacI-T with the *lacI* system followed by a *fdx* terminator before *repH* were transformed into *C. pasteurianum*-H1. Single colonies of each

plasmid transformant were replica plated onto RCM plates with 1 mM IPTG to induce plasmid loss and without IPTG as a control. The next day colonies were replica plated onto RCM plates with erythromycin on which only mutants still bearing the plasmid should grow. Figure 5.8 shows that all colonies can grow when subjected to permissive conditions (RCM). On the other hand, when plated on IPTG bearing plates, only mutants with plasmids pMTL83251 or pMTL83251-lacI-T were able to grow. Mutants formerly harbouring plasmid pMTL83251-lacI only show much reduced growth which putatively is an artefact of too thick a streak when replica plating. Thus, plasmid loss can be induced with IPTG on solid media in *C. pasteurianum*.

### 5.2.3.3 Transposon Mutagenesis in *C. pasteurianum*

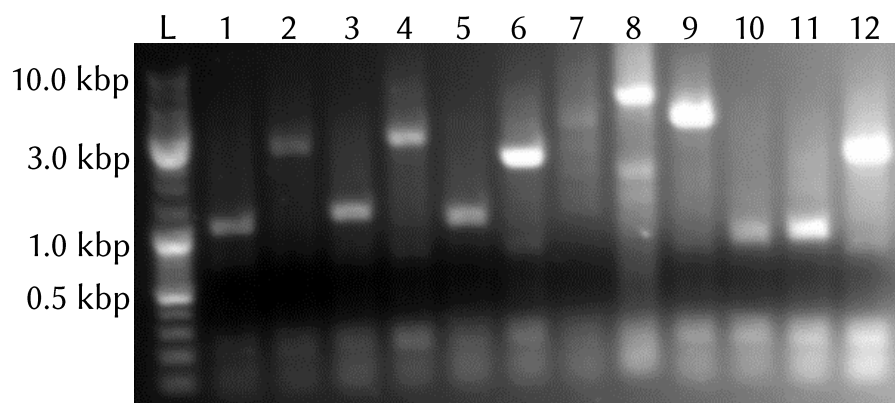
Having shown all the parts (conditional replicon, *tcdR*-P<sub>*tcdB*</sub>) of the pMTL-YZ14 transposon system to be working in *C. pasteurianum* the complete system could be tested.

**pMTL-YZ14** The transposon plasmid (pMTL-YZ14) with the P<sub>*tcdB*</sub> in front of the transposase and the IPTG inducible promoter in front of *repH* was transformed into *C. pasteurianum*-H1  $\Omega$ *tcdR* and Em<sup>R</sup> colonies selected for. Grown colonies were pooled in Phosphate Buffer Saline (PBS), serially diluted and plated on RCM agar media supplemented with 1 mM IPTG. Plates with single colonies were chosen and colonies replica plated onto RCM+Em and RCM+Tm agar media to check for transposon insertion and plasmid loss. All colonies grew on Em and Tm indicating the plasmid had not been lost.

This experiment was repeated another two times with the same result indicating that that plasmid pMTL-YZ14 was non-functional presumably due to a problem with the IPTG inducible replicon in pMTL-YZ14, which may have degenerated in the *E. coli* host and will need to be cloned again. However, instead of testing every component of the system again it was decided to attempt suicide delivery of the plasmid using a delivery vehicle lacking a Gram-positive replicon, plasmid pMTL-GL15. This alternate strategy was made possible by the improved transformation efficiency described in Chapter 3.

**pMTL-GL15** Plasmid pMTL-GL15 (constructed by Gareth Little) is similar to pMTL-YZ14 with the only difference being the absence of the IPTG inducible Gram-positive replicon. Thus, this plasmid acts as a suicide plasmid in that it is transformed into the organism but cannot be replicated. The exposure time of the plasmid in the organism, however, should be sufficient for the transposase to be expressed and the transposon to be inserted into the genome at





**Figure 5.9** – Inverse PCR of mini-transposon insertions in *C. pasteurianum*. Each band has a different size which indicates different insertion loci on the *C. pasteurianum* chromosome. Clones 6 and 8 show possible double insertions. L: gene size ladder (Generuler mix).

random. This was tested in *C. pasteurianum*-H3  $\Omega$ tcdR, which was used because it only carries one SNP as opposed to two in *C. pasteurianum*-H1. The strain was transformed with methylated pMTL-GL15. After recovery and plating on RCM+Tm, the number of potential transposon insertion events corresponded to 10 cfu/ $\mu$ g of plasmid DNA. Individual colonies were grown overnight in liquid medium and genomic DNA extracted for inverse PCR and subsequent Sanger sequencing of the amplified DNA product. Figure 5.9 shows the fragments after inverse PCR. All of the 12 clones show different fragment sizes, indicating that insertion of the mini-transposon has most likely occurred at different positions in the genome. Two mutants show two bands (lanes 6 and 8 in Figure 5.9) suggesting double insertion (two insertions of the mini-transposon in one chromosome). Bands were excised and sent for Sanger sequencing for insertion point analysis. Only two of the 12+2 bands could be sequenced due to low elution concentration of gel extraction. Both extracted fragments showed insertion in different loci. Clone 3 had the insertion on the forward strand at nucleotide number 3968816 which causes disruption of the 3'-end of gene CLPA\_c37060, annotated as a putative ATP:guanido phosphotransferase CKL. Clone 12 had its insertion on the forward strand at nucleotide number 3851193 which disrupts gene CLPA\_c35840 annotated as an ABC-type cobalamin/Fe<sup>3+</sup>-siderophore transport system, ATPase component. Both insertions were flanked by a TA-dinucleotide.

This work demonstrated individual random mutagenesis in *C. pasteurianum* with the *mariner*-transposon system. Unfortunately, due to time constraints this work could not be finished and further investigations are advised.

## 5.3 Discussion

The *mariner*-transposon system, used in this study, was originally developed by [Cartman and Minton \(2010\)](#) for the use in *C. difficile* and was further improved with a conditional replicon by [Zhang et al. \(2015\)](#), resulting in plasmid pMTL-YZ14. Here the applicability of the complete system in *C. acetobutylicum* and *C. sporogenes* is assessed. Furthermore, the system is used to screen for mutants of phenotypic interest. Having shown the system to work in these industrially (*C. acetobutylicum*) and medically (*C. sporogenes*) important strains it was assumed that the strategy could be used in all clostridial species. This study was further set out to test the system and its parts in *C. pasteurianum* which was possible with alteration of the Gram-positive replicon from conditional to suicide.

### 5.3.1 Transposon Mutagenesis in *C. acetobutylicum*

The *mariner*-transposon system with the *tcdR*-P<sub>*tcdB*</sub> expression system and the conditional replicon on plasmid pMTL-YZ14 was exemplified in the butanol producer *C. acetobutylicum* by assessing randomness of insertion of the mini-transposon on the chromosome and by directed screening for mutants of interest. Proof of principles for the separate parts (*tcdR*-P<sub>*tcdB*</sub> expression system and the conditional replicon) of the system were tested by [Zhang et al. \(2015\)](#). Here, it was shown that the insertion of the mini-transposon is only reliant on a TA-dinucleotide which is duplicated in the process of insertion to flank the mini-transposon as reported by [Cartman and Minton \(2010\)](#). In over 98 % of the tested mutants a single insertion was detected in either the chromosome or the mega-plasmid which largely simplifies determination of the underlying genetic cause of an observed phenotype. Furthermore, 85.3 % of the random insertions were located in protein coding sequences, fitting well with the fact that 80 % of the clostridial genome is protein coding. This proof of principle showed that the *mariner*-transposon can be used to construct large transposon mutant libraries in *C. acetobutylicum* without redundant clones or many mutants with double insertions that will have to be assessed separately for a conclusive genotype.

To show that the transposon system can be used to screen for phenotypically interesting mutants, proof of principle screenings for auxotroph mutants and sporulation/germination deficient mutants were conducted. In this work the focus was on feasibility of such a screen rather than assessing fundamental basics of the phenotypes. A library of approximately 3000 mutants was used to screen for mutants of interest. With *C. acetobutylicum* being a spore forming



organism there is an interest in its sporulation and germination ability. However, not all genes involved in this system are known and the transposon is the perfect tool to investigate this mechanism. The screening led to one mutant with the transposon inserted in CA\_c2631, a hypothetical protein with homology to YkuD involved in sporulation in *B. subtilis*. According to Kodama et al. (2000) YkuD is expressed in sporulation state T<sub>4</sub> in *B. subtilis* and bound to the cell membrane. The specific function of CA\_c2631 has to be investigated further but it can be assumed that it is involved in the sporulation state and possibly transcribed similarly to *ykuD* and other sporulation proteins by mother-cell specific transcription factors like *sigE* or *sigK* (Kodama et al., 2000).

Auxotroph mutants give valuable insight into the essentiality of genes for an organism and at the same time the ability of an organism to synthesise certain metabolites. In *C. acetobutylicum* the screen found gene CA\_c0971 (*citB*) by the inability of the transposon insertion mutant to grow on minimal medium. The gene encodes aconitate hydratase which in aerobic bacteria is responsible for the conversion of citrate to iso-citrate in the tricarboxylic acid (TCA) cycle. Iso-citrate is further converted into the TCA cycle intermediate  $\alpha$ -ketoglutarate which is further converted into the non-TCA cycle metabolite glutamate, which can subsequently be converted into arginine, proline and glutamine. A TCA cycle was not thought to be part of strict anaerobes; however, two independent research groups have recently suggested that *C. acetobutylicum* possesses a complete, albeit bifurcated, TCA cycle (Figure 5.4). In the TCA cycle, oxaloacetate is converted to succinate both through citrate/ $\alpha$ -ketoglutarate and via malate/ fumarate (Amador-Noguez et al., 2010; Crown et al., 2011). This study showed that growth of the CA\_c0971 (*citB*) auxotroph mutant can be restored to wild type by exogenous glutamate and partially restored by glutamine addition. Intriguingly, the mutant remained unable to grow on minimal media when supplemented with  $\alpha$ -ketoglutarate, which is suggested to be the product of the conversion of iso-citrate by isocitrate dehydrogenase (CA\_c0971) and the substrate for the synthesis of glutamate (Amador-Noguez et al., 2010). This study clearly shows the gene CA\_c0971 to be involved in glutamate biosynthesis and suggests that the proposed TCA cycle (Figure 5.4) has to be investigated further in anaerobic organisms. There might exist another mechanism for glutamate synthesis other than from  $\alpha$ -ketoglutarate seeing that this addition did not restore growth, assuming that exogenous  $\alpha$ -ketoglutarate can be assimilated by *C. acetobutylicum*. The fact that glutamine addition restored growth partially suggests that glutamine is derived from glutamate by the ATP dependent glutamine synthetase (Ussdin et al., 1986), but also accentuates that the reverse reaction glutamine to glutamate is possible but not thermodynamically

favourable under laboratory conditions.

### 5.3.2 Transposon Mutagenesis in *C. sporogenes*

The transposon mutagenesis pMTL-YZ14 system was exemplified in *C. sporogenes* in similar fashion to the work reported in *C. acetobutylicum* above. The delivery in *C. sporogenes* is more convenient due to the fact that plasmid loss is induced and successful in 100 % of the screened mutants and the transposon is inserted. A library of *C. sporogenes* transposon mutants is easily generated without having a subset of plasmid bearing strains. Randomness of insertion, only guided by a TA-dinucleotide, was demonstrated with 60 independent transposon mutants. Multiple insertions were found in only 3.3 % of the mutants which simplifies genotyping of putatively interesting phenotypes. Of the 62 insertions 49 were present in gene encoding regions signifying a slightly higher ratio than found in *C. acetobutylicum*.

From a library of 6000 transposon mutants, two were found to be sporulation/germination deficient. One was gene CLSPO\_c23320, encoding a flavodoxin oxidoreductase. Oxidoreductases are related to aerotolerance and provide protection against oxidative stress (Hillmann et al., 2009). Spores by their very nature are important to disseminate the clostridial species and survive otherwise deleterious conditions. It is therefore likely that proteins scavenging oxidative radicals are present in the layers of the spores and play roles in structural build-up of spore coat layers by mediating di-tyrosine cross-linking among proteins and/or in resistance against oxidative stress (Lehmann et al., 1996; Riebe et al., 2009). However, the flavodoxin oxidoreductase found in this screening could also have a different and hitherto unknown function. The other gene disrupted to give a sporulation/germination negative phenotype was found to be CLSPO\_31790 encoding the stage V sporulation protein SpoVAD. In *B. subtilis* the *spoVA* operon encodes six proteins which have been suggested together to be involved in dipicolinic acid (DPA) transport into the developing forespore and the release of Ca<sup>2+</sup>-DPA and other small molecules during spore germination (Tovar-Rojo et al., 2002; Vepachedu and Setlow, 2007). This fits well with the observed phenotype of the CSPO\_31790 disruption mutant.

The same transposon library screened for sporulation/germination deficient mutants was also screened for auxotroph mutants. Out of the 6000 mutants three were found unable to grow on minimal medium. The disrupted genes were CLSPO\_c16750, encoding a collagenase (*colA*), CLSPO\_c21170, coding for a hypothetical protein and CLSPO\_c29860, annotated as an ABC-2 fam-

ily transporter protein. BLAST searches of the proteins did not reveal any more information or different conserved regions. Thus, for the hypothetical protein (CLPA\_c21170) a conclusive determination of the cause of the auxotrophy is not possible. Further experimental work would be required to determinate its involvement which is out of the scope of this study. The same is true for the ABC-2 transporter protein CLSPO\_c29860. However, it is likely involved in the translocation of, as yet unknown, nutrient(s)/substrate(s) into the cell. Finally, even though no direct evidence suggests that collagenases are essential for growth under minimal conditions, it was noted that some organisms, such as e.g. *Peptostreptococcus asaccharolyticus*, *Flavobacterium meningosepticum* and *Porphyromonas* spp., expressing collagenolytic enzymes are asaccharolytic (Harrington, 1996). As the metabolism of such organisms is dependent on the uptake of small peptides and amino acids, the production of collagenolytic enzymes may be essential for growth and survival. Thus, *colA* could be an enzyme involved in such uptake in *C. sporogenes*.

### 5.3.3 Transposon Mutagenesis in *C. pasteurianum*

To this date there are no reports of transposon mutagenesis in *C. pasteurianum*. Having shown successful transposon mutagenesis in *C. acetobutylicum* and *C. sporogenes* the pMTL-YZ14 transposon plasmid was chosen to be used in *C. pasteurianum* and its individual components were assessed.

#### 5.3.3.1 *tcdR* Dependent Expression

The coupled expression system with the *tcdR*  $\sigma$ -factor and the  $P_{tcdB}$  promoter was previously developed in *E. coli* (Moncrief et al., 1997) and adapted for use in *C. difficile* (Cartman and Minton, 2010). The characteristic of  $P_{tcdB}$  to only be expressed when induced by TcdR (Moncrief et al., 1997) is used in the transposon system to limit expression of the transposase to the clostridial recipient. Specifically, transposase expression has to be prevented in the transient host *E. coli*. Activity of  $P_{tcdB}$  in *E. coli* was determined in a chloramphenicol acetyltransferase (CAT) assay, measuring CatP expression levels colourimetrically, to compare promoter strength against the known strong *C. sporogenes*  $P_{fdx}$  promoter. It was found that  $P_{tcdB}$  is constitutively expressed in *E. coli* Top10. However, expression levels are low (a third of the levels of  $P_{fdx}$ ). Thus, if the gene-product expressed by  $P_{tcdB}$  is not lethal to the intermediate host, the *tcdR*- $P_{tcdB}$  expression system can be used to reduce expression compared to the use of a strong promoter, with the advantage that expression will still be high in the final clostridial host endowed with *tcdR*. This was shown in *C. acetobutylicum*

where products toxic to *E. coli* (parts of cellulosomes) were cloned successfully in this organism and expressed in *Clostridium* (Katrin Schwarz, unpublished data).

Having shown the low expression profile in the intermediate host *E. coli* by  $P_{tcdB}$ ; expression of the promoter in *C. pasteurianum*, with and without the  $\sigma$ -factor *tcdR* on the chromosome, was tested. Host strains were constructed in all hypertransformable clones (*C. pasteurianum*-H1,-H3 and -H4, Chapter 3) by Allele Coupled Exchange (ACE) knock-in downstream of the *pyrE* locus. For this purpose *pyrE*-truncation mutants had to be constructed first and the *tcdR* fragment added with simultaneous correction of the *pyrE* locus. Most of these strains apart from *C. pasteurianum*-H3/4 *tcdR* insertions were subjected to whole genome sequencing and checked for the presence of ancillary mutations (Single Nucleotide Polymorphisms (SNPs) and/or insertion/deletions (InDels)). This located a SNP in *C. pasteurianum*-H1 $\Delta$ *pyrE* and two additional SNPs in *C. pasteurianum*-H1 $\Omega$ *tcdR*. The latter strain was initially used for transposon mutagenesis and showed lower transformation efficiency than its progenitor. However, at the time *C. pasteurianum*-H3/4 were not available and so initial work was undertaken with this strain before changing to *C. pasteurianum*-H3 $\Omega$ *tcdR* for the pMTL-GL15 suicide transposon mutagenesis. The gene *catP* was barely expressed in *C. pasteurianum*-H1 without *tcdR* and was highly expressed in *C. pasteurianum*-H1 $\Omega$ *tcdR*, indicating tight regulation of  $P_{tcdB}$  by the expression couple *tcdR*- $P_{tcdB}$ .

### 5.3.3.2 Conditional Replicon

Functionality of the conditional replicon was shown in *C. pasteurianum* with similar efficiency to *C. acetobutylicum* and *C. sporogenes* (Zhang et al., 2015). Plasmid loss was successfully induced with 1 mM  $\beta$ -D-1-thiogalactopyranoside (IPTG) which de-represses the  $P_{fac}$  promoter putatively leading to increased transcription of the pCB102 replicon region and potentially a metabolic burden on cells carrying the plasmid. The exact mechanism of this system is not yet fully understood (Zhang et al., 2015).

### 5.3.3.3 The Conditional System

With all parts of the conditional transposon plasmid pMTL-YZ14 assessed, and shown to work in *C. pasteurianum*, transposon mutagenesis could be assayed. Transformation was successful with good transformation efficiencies and  $Em^R$  colonies harbouring the plasmid were harvested. However, trying to induce plasmid loss with 1 mM IPTG was unsuccessful in three separate experiments.

It was assumed that the conditional replicon could have degenerated in the intermediate *E. coli* host and it will have to be cloned again. Before this could be further investigated another route of transposon mutagenesis was successfully tested involving the use of a non-replicative suicide plasmid.

#### 5.3.3.4 The Suicide System

A suicide delivery system is preferred to a conditional delivery system because every colony growing after transformation is most likely an independently generated transposon insertion mutant. The plasmid is lost from the population immediately after transformation as it cannot be replicated in the host due to absence of the Gram-positive replicon. One hurdle to the use of a suicide delivery method is transformation efficiency because the possibility of obtaining direct transposon mutants is small. To obtain a mutant, the plasmid has to be taken up during transformation, the transposase has to be expressed and a transposon event has to happen, before the plasmid is segregated out due to its inability to replicate. With the high efficiency transformation protocol developed in Chapter 3 suicide transformation was assessed. 12 putative transposon mutants were generated in one transformation with a standard *C. pasteurianum* transformation. These were confirmed by inverse PCR but only two were confirmed to be inserted in the chromosome of *C. pasteurianum* by DNA sequencing of the inverse PCR amplified DNA product. This shows that transposon mutagenesis is possible with the pMTL-GL15 suicide plasmid. By optimising the transformation procedure for suicide transformation and using a newly developed high-transformation protocol from *C. beijerinckii* (Gareth Little, personal communication) it should be possible to increase the initial library size to a capacity useful for screening for mutants of interest or transposon directed insertion-site sequencing (TraDIS) (Langridge et al., 2009) which needs an insertion at every possible chromosomal locus to be applicable.

This study represents the first report of transposon mutagenesis in *C. pasteurianum* with a slight variation to plasmid pMTL-YZ14 used in *C. acetobutylicum* and *C. sporogenes*. This system has to be optimised in future work for it to be used in transposon based screenings and further developments.

#### 5.3.4 Conclusion

In this work the application of a *mariner*-transposon system was presented. The system was first introduced by Cartman and Minton (2010) and refined by Zhang et al. (2015). The system was used to build exhaustive transposon mutant libraries in *C. acetobutylicum* and *C. sporogenes*. In both strains randomness of

transposon insertion was proven and it was shown that transposons would insert only once per mutant in a majority of the cases. The value of the system was further shown by screening for mutants of interest and sporulation/ germination defective as well as auxotroph mutants were found in both species. When the system was assessed in *C. pasteurianum* all parts seemed to work well but use of the whole conditional system in this organism was unsuccessful as plasmid loss could not be induced with IPTG. Instead, a suicide plasmid (pMTL-GL15) was used and transposition was shown to be effective. Due to time constraints this work could not be finished in its entirety and further improvements have to be done to apply the transposon system on a regular basis.

This work is a first step towards the application of a universal *mariner*-transposon system in a wide range of clostridial hosts.

# Chapter 6

## General Discussion

Sometimes the truth is arrived at by adding all the little lies together and deducting them from the totality of what is known.

---

Terry Pratchett, *Going Postal*

### 6.1 Background

The genus *Clostridium* includes a wide range of biotechnologically and medically important species. These anaerobic, Gram-positive spore formers are used to produce biofuels and other valuable chemical commodities from a wider range of first, second and fourth generation feedstocks and clostridia are in line for promising cancer treatments (Minton, 2003). Furthermore, extensive research is ongoing to prevent diseases caused by the pathogenic members of the genus with its noteworthy representative *C. difficile* causing infectious diarrhoea (Poxton et al., 2001). This work focused on three *Clostridium* species: *C. acetobutylicum*, *C. sporogenes* and special attention was paid to the glycerol fermenter *C. pasteurianum*. *C. sporogenes* is a close relative of *C. botulinum* without producing the potent neurotoxin which makes it a suitable and safe surrogate strain for studies of this medically important strain (Kubiak et al., 2015). Even more importantly, *C. sporogenes* is used in a novel approach to cancer treatment called *Clostridium*-directed enzyme pro-drug therapy (CDEPT). CDEPT uses *C. sporogenes* spores to shuttle pro-drug converting enzymes to the hypoxic regions of tumours where the anaerobe organism germinates and expresses the enzymes. The administered pro-drug will then be converted to the active drug only in the vicinity of the tumour and kill it. (Minton, 2003) Modern research in both *C. acetobutylicum* and *C. pasteurianum* is primarily



targeting their ability to produce biofuels from renewable feedstocks. Even though both produce ethanol their ability to produce *n*-butanol is of paramount importance as this commodity shows higher energy density and an increased boiling point, which makes it safer to manipulate (Harvey and Meylemans, 2011). *C. acetobutylicum* can use different feedstocks such as sugars and starch but also whey, pectin, inulin and xylan (Nolling et al., 2001). With the debate of food vs fuel the use of sugar and starch is controversial and more attention is paid to the adoption of non-food crops for butanol production by *C. acetobutylicum*. The feedstock of *C. pasteurianum*, on the other hand, is glycerol which is the major waste product of biodiesel production. Biodiesel production generates 10 w/w% of unrefined (crude) glycerol which cannot be used by other traditional glycerol utilising industries. *C. pasteurianum* was shown to grow well in the crude effluent (Venkataramanan et al., 2012; Taconi et al., 2009) and in one instance it was even shown to be advantageous to growth compared to refined glycerol (Jensen et al., 2012a).

The abilities of *Clostridium* make these microorganisms attractive targets for modern molecular biological research. For this purpose, Minton et al. (2016) proposed a roadmap for the development of clostridial hosts. It outlines a procedure to develop tools needed for comprehensive genetic engineering of this genus. This procedure starts with genome sequencing and analysis of restriction/ modification systems, followed by development of a transformation procedure with subsequent evaluation of applicable plasmid replication origins and antibiotic resistance genes. Once transformation is established, reverse genetic tools can be developed starting with the ClosTron insertional inactivation tool, followed by Allele Coupled Exchange (ACE) to generate a *pyrE* truncation mutant used in allelic exchange with a heterologous *pyrE* as the counter-selection marker. ACE can further be used for gene insertions for either complementation of knocked-out genes at chromosomal level or synthetic genes at the *pyrE* locus with concomitant repair of *pyrE*. By inserting the *tcdR* gene into the host's chromosome, a *mariner* based plasmid system, in which the *Himar1* C9 transposase gene was placed under the control of the  $P_{tcdB}$  promoter, can be utilised for forward genetics. This promoter ( $P_{tcdB}$ ) is exclusively recognised by a specialised class of sigma factor, TcdR. The roadmap is comprehensive insofar as it offers forward and reverse genetic tools. In a next iteration it should, however, be completed with a CRISPR system for simplification of ACE and allelic exchange. The reverse genetic tools of the roadmap were successfully established in *C. sporogenes* and *C. acetobutylicum* when this work began (Minton et al., 2016) and here forward genetics is exemplified in these species with the application of the *mariner*-transposon system. In *C. pasteurii-*



*anum*, on the other hand, this work together with developments reported in Schwarz et al. (2017) established the complete roadmap.

## 6.2 Key Findings of This Work

### 6.2.1 Initial Steps of the *C. pasteurianum* Roadmap

When this study was conceived no reliable transformation protocol was available for *C. pasteurianum*. The genome was sequenced and one type I and two type II restriction modification systems were identified (Poehlein et al., 2015). Rotta (2014) was able to protect DNA against the endogenous type II restriction system CpaAI by *in vivo* methylation with M.BepI in a *dam*<sup>+</sup> (protecting against CpaPI) *E. coli* host. However, he was not able to establish transformation efficiencies needed for genetic engineering. By using a transformation protocol published by Pyne et al. (2013) efficiencies of transformation could be achieved to the order of magnitude of 10<sup>2</sup> but the reported 10<sup>4</sup> could not be reached with any *C. pasteurianum* strain obtained from different culture collections. Judging from the small transformation efficiencies reached it was hypothesised that there might exist a sub-population of the wild type that is susceptible to transformation as opposed to the majority of the population. This could explain the lack of success in reproduction of the published transformation method of Pyne et al. (2013), who may have unwittingly transformed a mutant *C. pasteurianum* ATCC 6013 strain by culturing such a sub-population. *C. pasteurianum* ATCC 6013 was at the time considered to be the same as DSM 525. However, later the two the strains were shown to differ by several variants (SNPs and InDels) as shown by the ATCC 6013 genome published by Rotta et al. (2015) in comparison to the DSM 525 genome published by Poehlein et al. (2015). To test the hypothesis of a hypertransformable sub-population, type strain *C. pasteurianum* DSM 525 was tested by curing one of the rare transformants of its acquired plasmid, and then testing the frequency of DNA transfer in the isolated plasmid-free strain obtained. The screening experiments were carried out several times independently, and 3 out of 7 screens yielded clones that were hypertransformable (transformable with increased transformation efficiency) compared to the wild type. Such a screening procedure yielding hypertransformable mutants is reported for the first time and it could be a valuable method to overcome DNA transfer barriers in other organisms presenting low transformation efficiency.

## 6.2.2 Genotypic Analysis of Hypertransformable Clones

The hypertransformable clones found by the screening procedure were analysed by whole genome re-sequencing for putative genetic reasons for hypertransformability. Next generation sequencing revealed two SNPs in one clone (*C. pasteurianum*-H1) and one identical SNP in two individual clones (*C. pasteurianum*-H3 and -H4) to be potentially responsible for the observed phenotype. Both SNPs in the first clone, *C. pasteurianum*-H1, introduced frame-shifts in the respective genes which potentially led to inactivation of the genes. One SNP was located in gene CLPA\_c13710 encoding a  $\beta$ -lysine N<sub>6</sub>-acetyltransferase showing high similarity to *ablB* of *Methanosarcina mazei*, in which it is regulating salt stress (Pfluger et al., 2003). The second SNP is located in CLPA\_c33080 (*resE9*) encoding an ‘orphan’ histidine kinase without a cognate response regulator in the genomic vicinity and thus is not part of a two component system (West and Stock, 2001). Interestingly, the gene adjacent to and in the same operon as CLPA\_c33080 is *agrB*, part of a quorum sensing system and responsible for pre-processing of the AgrD signal molecule to generate the cyclic *agr* signal peptide (Zhang et al., 2002). The structural arrangement *agrB* → *agrD* → *resE* is conserved in many clostridial species and was first noted by Cooksley et al. (2010). Presently, *resE* has not been shown to be directly involved in *agr* quorum sensing system; it was suggested that the two homologous orphan kinases encoded by CBO\_0336 and CBO\_0340 in *C. botulinum* may play a role in sensing the two AgrD peptides (Cooksley et al., 2010). It is further possible that this quorum sensing system regulates processes involved in DNA uptake similar to the *com* operon in *Streptococcus pneumoniae*. In this organism its histidine kinase *comD* interacts with competence-stimulating peptides, which, when a cell density of about 10<sup>7</sup> cells per ml is reached, triggers a signal cascade leading to expression of a competence regulon (Steinmoen et al., 2002). This induction leads to cell lysis of a subpopulation of *S. pneumoniae* and the released DNA is taken up by the rest of the population (Spoering and Gilmore, 2006). It remains to be determined if the *C. pasteurianum agrBD-resE* serves a similar function.

The SNP (T3304G), introducing the amino acid change Phe1102Val in gene CLPA\_c30550 in *C. pasteurianum*-H3 and by genetic homology in *C. pasteurianum*-H4, was shown to be the cause of hypertransformability. Correction of the variant back to wild type returned the hypertransformable phenotype to wild type low transformation efficiency. The affected gene CLPA\_c30550 shows homology to structural maintenance of chromosome (SMC) proteins with nucleolytic activity (Graumann, 2000; Lindow et al., 2002). CLPA\_c30550 pos-

sesses the characteristic motives (Walker et al., 1982) as well as the structural hinged arrangement (Hirano, 2002) of typical SMC proteins. *In silico* modelling of the protein further revealed similarities to nucleolytic enzymes. This substantiates the hypothesis that the gene is involved in protein-DNA interaction with possible implication in DNA uptake by transformation. Phyre2 modelling suggested that the SNP is located in a region putatively responsible for protein transport. However, without further experimental evidence, it cannot be unravelled how one minor change in protein CLPA\_c30550 could cause the drastic increased DNA transfer efficiency in *C. pasteurianum* and more work is needed to investigate the function of CLPA\_c30550. Nevertheless, it does show that more factors can influence the transfer of foreign DNA into a host organism than restriction/ modification and CRISPR systems. Furthermore, research into bacterial competence and transformation is still missing key insights (Thomas and Nielsen, 2005) and this work might help for further comprehension of this topic.

### 6.2.3 Development of Reverse Genetics

With firmly established transformation, reverse genetic tools were developed in *C. pasteurianum*. Rotta (2014) showed Clostron inactivation of genes was feasible with the small adaptation of changing the  $P_{fdx}$  promoter against the  $P_{fac}$  promoter controlling group II intron expression. Schwarz et al. (2017) constructed all basic vectors needed for ACE and allelic exchange, then exemplified both by making in-frame deletions of *spo0A* encoding the sporulation master regulator using allelic exchange and subsequently complemented it at chromosomal level through ACE. The *spo0A* deletion strain showed the expected sporulation deficient phenotype (Schwarz et al., 2017) and complementation restored sporulation fully. The deletion of the Spo0A sporulation master regulator was found to lead to increased acid levels and decreased butanol levels in defined medium, while both metabolites increased when grown in complex medium containing yeast extract. This stands opposed to findings in *C. acetobutylicum* and *C. beijerinckii* where solvent production was entirely abolished when *spo0A* was deleted or insertionally deactivated, respectively (Harris et al., 2002; Ravagnani et al., 2000). These findings suggest the regulatory functions of Spo0A in *C. pasteurianum* to share similarity in terms of sporulation but also to be different compared to its close relatives, namely in its influence over the solvent metabolism. This study confirmed findings by Sandoval et al. (2015) that deletion of *C. pasteurianum spo0A* leads to increased butanol production in complex medium. However, closer analysis of the data-sets of both studies

showed that the increased levels were due to the strains superior capability of glycerol uptake demonstrated by unchanged butanol yields (c/c) between wild type and mutant. Fermentation in defined medium in this study, on the other hand, led to decreased butanol titres of the mutant compared to the wild type. With yeast extract concentration being five times lower in this medium than in the complex medium, it is likely that the ample yeast extract concentration led to the differences in metabolite pattern. A similar influence of yeast extract over *C. pasteurianum* fermentation was noted by Moon et al. (2011), with increased concentration favouring 1,3-propanediol (1,3-PDO) and lower concentrations favouring butanol production.

### 6.2.3.1 Creation of Mutants With Improved Solvent Titres

After the exemplification of reverse genetic tools in *C. pasteurianum*, genes *rex*, *hyd* and *dhaBCE*, crucial for redox balance, were chosen to be deleted with the expectation of enhanced solvent profiles.

***dhaBCE*** Glycerol is a highly reduced substrate and production of less reduced biomass creates excess NADH. 1,3-PDO, the final product of the reductive part of glycerol metabolism, is more reduced than its substrate and can thus serve as an acceptor for excess reducing equivalents from biomass. 1,3-PDO is thought to be crucial for growth of *C. pasteurianum* on glycerol (Biebl, 2001; Johnson and Rehmann, 2016), which was confirmed in this study with a *dhaBCE* deletion mutant. By deleting *dhaBCE* in *C. pasteurianum*, the pathway for funnelling extra reducing power during glycerol fermentation is absent, resulting in poor growth of the deletion mutant in defined medium. Growth of the deletion mutant could only be restored in complex medium with high concentrations of yeast extract, which is potentially used as alternative substrate and acts as reducing agent in *C. pasteurianum*  $\Delta$ *dhaBCE* growth. By reintroducing *dhaBCE* back into the deletion strain the phenotype could be reversed, confirming the importance of 1,3-PDO production for biomass production and redox balance of *C. pasteurianum*.

Simplistically, elimination of the reductive pathway could funnel carbon into the oxidative branch of the glycerol metabolism and increase solvent titres. However, solvents in the deletion mutant were found to be unchanged compared to the wild type. This stands opposed to a reported disruption of *dhaT*, the final step of the reductive branch, by Pyne et al. (2016c) who found 30 % increased butanol titres. However, their mutant was still able to produce small amounts of 1,3-PDO, possibly due to imperfect ClosTron mutagenesis. Pyne et al. (2016c) failed to isolate intron mutants of *dhaA*, *dhaB* and *dhaC*. An in-

tron insertion in *dhaT* was obtained but the intron delivery plasmid and its encoded LtrA protein could not be cured. They found an additional ecotopic, sense strand, intron insertion in *speA* (encoding arginine decarboxylase) which was hypothesised as being essential. One consequence of these difficulties was that Pyne et al. (2016c) were unable to eliminate 1,3-PDO production, which makes this the first report of complete elimination of 1,3-PDO production in *C. pasteurianum*.

**rex** This work presents an in-frame deletion of the *rex* gene, encoding the Rex redox regulator. Rex is sensing the internal redox balance of the cell by binding NADH and NAD<sup>+</sup> and regulates gene expression in response. High NADH/NAD<sup>+</sup> ratios release bound Rex from its recognition sequence in the promoter region of the target gene which allows expression (Brekasis and Paget, 2003). A deletion of *rex* should therefore lead to constitutive expression of its target genes.

In this work a stringent *C. pasteurianum* DNA recognition sequence was identified, 5'-TTGTTAAWNNNTTAACAA, based on a general bacterial recognition sequence (Ravcheev et al., 2012). For the *C. pasteurianum* Rex recognition sequence a potential regulon of 47 genes and operons was identified. Many of the genes were located in the NADH consuming fermentative pathway, as it was expected and reported in the related species *C. acetobutylicum* (Wietzke and Bahl, 2012) and *C. kluyveri* (Hu et al., 2016). Recognition sequences were found adjacent to the alcohol dehydrogenases putatively involved in butanol and ethanol production, the thiolases that channel carbon into the reductive part of the fermentative pathway, the *bcs* operon which reduces acetyl-CoA to butyryl-CoA, and the reductive part of the *dha* operon leading from glycerol to 1,3-PDO. Of the non-fermentative genes previously reported (Ravcheev et al., 2012; Zhang et al., 2014), none was found using our approach. However, other notable putative Rex targets found include: CLPA\_c05450, a nitroreductase related to *narAB* or *narK* in *C. carboxidivorans*; CLPA\_c09900, a nicotinamidase family protein; CLPA\_c17480 (*spoVD2*) encoding a stage V sporulation protein and CLPA\_c39090, a NADPH dehydrogenase.

The metabolite analysis of the *rex* deletion mutant fit in part to the expected de-repression by the absence of Rex. Lactate production increased even though no regulatory sequence was located upstream of *ldh* suggesting that the gene is regulated by Rex after all. The regulation of *ldh* is reported in other species (Wietzke and Bahl, 2012; Hu et al., 2016), however, even relaxed algorithm parameters could not find a regulatory sequence upstream of this gene in *C. pasteurianum*. Furthermore, butyrate production was down-regulated

even though the pathway should be de-repressed. This can be explained by increased butanol concentration which could, in part, be due to re-assimilation of butyrate. The increased production of butanol and ethanol fits well with the expected de-repression of the regulatory sequences found upstream of the alcohol dehydrogenases (*adhE1,2,4*) and the *bcs* operon encoding genes leading from acetyl-CoA to butyryl-CoA.

The *rex* deletion strain produced 1.4 times more butanol than the wild type and showed a 53 % reduction of the by-product 1,3-PDO which makes this mutant a superior industrial butanol producer.

**hyd** Hydrogenases are involved in regulation of redox balance by producing molecular hydrogen from excess reducing equivalents (Vignais and Colbeau, 2004). Inactivation of bacterial hydrogenases should thus free reducing equivalents for other oxidative pathways like solvent generation. Several reports of increased butanol and ethanol production by *in vivo* inhibition of hydrogenases support this hypothesis by e.g. CO inhibition of hydrogenases (Datta and Zeikus, 1985; Hüsemann and Papoutsakis, 1989; Dabrock et al., 1992) or by the application of an artificial electron carrier (Peguin et al., 1994). This study shows for the first time the phenotype of a hydrogenase clean deletion in *Clostridium*. Previous attempts of disruption of *hydA* in *C. acetobutylicum* were unsuccessful (Cooksley et al., 2012) and it was hypothesised that the hydrogenase is crucial as an electron sink in this species under the conditions tested. It can be assumed that the knock-out in *C. pasteurianum* was done successfully due to a homologue [Fe]-hydrogenase located in the genome (Poehlein et al., 2015) or through the action of the [NiFe]-hydrogenase. Nevertheless, the deletion of *hyd* had the effect to channel reducing equivalents towards solvent production and to increase solvent yields. In the *hyd* mutant 2.7 times more ethanol and 1.1 times more butanol was produced compared to the wild type. A similar but less defined profile was observed in a hydrogenase knock-down study in *C. pasteurianum* where butanol concentration was increased 1.1 times and ethanol 1.25 times (Pyne et al., 2015). Furthermore, lactate and acetate levels were increased. Acetate generation is not NADH-consuming but ATP generating via substrate level phosphorylation. One consequence of which was increased biomass observed in the *hyd* deletion mutant with the OD<sub>600</sub> being 114 % higher than in the wild type. This was further characterised by a shorter fermentation time as adjudged by the earlier onset of pH increase after 18.7 h in the mutant as opposed to 22.0 h in the wild type. This observation was made by Pyne et al. (2015) in their *hydA* knock-down but was not evident in the study of Dabrock et al. (1992), who inhibited hydrogenase activity by sparging the



culture with CO. Here acetate concentration was reduced by 50 %. However, CO inhibits all iron containing hydrogenases, whereas the strain used here and in the study of [Pyne et al. \(2015\)](#) was only affected in the production of *hydA* and the other [Fe]-hydrogenases (CLPA\_c33960, CLPA\_c37830) and the [NiFe]-hydrogenase (CLPA\_c07060/ CLPA\_c07070) of *C. pasteurianum* DSM 525 might compensate for the loss of *hydA*. An unexpected observation was the reduction of 1,3-PDO production by 19 % in the *hyd* deletion strain compared to the wild type. 1,3-PDO production is NADH utilising and would be expected to be upregulated in such a mutant. [Pyne et al. \(2015\)](#) reported a similar reduction (30 %) in their knock-down mutant attributing the reduction to variation of metabolite production between fermentations, an observation made in other studies before ([Biebl, 2001](#); [Taconi et al., 2009](#)). This study supports the view that 1,3-PDO production is a real phenomenon caused by ablation of *hyd* activity. It should also be noted that the reduction in 1,3-PDO and accompanying increase in biomass is contrary to the suggestion by [Biebl \(2001\)](#) that the production of reducing equivalents by biomass build up has to be equalled by production of 1,3-PDO to regenerate the pool of NAD<sup>+</sup> to NADH.

**Further Findings** During analysis of the *C. pasteurianum* knock-out mutants and especially the *rex* deletion it was noticed that the organism re-assimilates butyrate for the production of butanol as it is the case in *C. acetobutylicum*. However, the acetoacetyl-CoA: butyrate:CoA transferase in *C. acetobutylicum* and *C. pasteurianum* is thought to only function with concomitant acetoacetate production followed by conversion to acetone. In the *rex* mutant butyrate re-assimilation was observed directly for the first time and no acetone was produced which indicates that butyrate can be re-assimilated via another pathway than the proposed *ctfAB*. A similar phenomenon was observed before in *C. acetobutylicum* ([Hüsemann and Papoutsakis, 1989](#); [Lehmann et al., 2012](#); [Millat et al., 2014](#)) in deletion mutants of the acetone pathway. As *C. pasteurianum* was never reported to produce acetone it is a perfect organism to study butyrate uptake without the need of concomitant acetone production.

### 6.2.3.2 Forward Genetic Studies in Clostridia

The last step of the clostridial roadmap ([Minton et al., 2016](#)) is to develop a forward genetics approach. For this purpose a transposon random mutagenesis system was developed to generate a transposon insertional mutant library which can then be used to screen for mutants of interest and subsequently the affected gene. In the current study, two fundamental refinements were implemented to an existing *mariner* transposon system ([Cartman and Minton, 2010](#))

that greatly increase its utility for use in *Clostridium* spp. On the one hand, an improved transposon delivery vehicle was devised that is conditional for plasmid maintenance. This significantly reduces the background of antibiotic resistant transposon colonies that still retain the plasmid after plating under non-permissive condition. On the other hand, a strategy was devised whereby the benefits of using the  $P_{tcdB}$  promoter, which does not function in the *E. coli* donor but is confined to the target *Clostridium* cell, can be implemented in potentially any *Clostridium* species. This is simply achieved through the introduction of the cognate  $\sigma$  factor gene, *tcdR*, into the target host genome using the well established gene integration technology, ACE.

In the current study, the system was initially exemplified in *C. acetobutylicum* and *C. sporogenes*, two biotechnologically valuable *Clostridium* species. Transposition into clostridial genome occurred at a frequency of  $2.6 (\pm 0.6) \times 10^{-4}$  and  $3.2 (\pm 0.5) \times 10^{-4}$  in *C. acetobutylicum* and *C. sporogenes*, respectively. Transposition occurred in a random fashion, with simultaneous plasmid loss at 80 % and 100 %, generating transposon mutants with just a single insertion in an overwhelming majority of cases (98.3 % and 96.7 %, in this study). Moreover, 51 of the 61 insertions sequenced (83.6 %) in *C. acetobutylicum* and 49 of the 62 (79 %) in *C. sporogenes* were located within protein coding sequences. This is within the range that would be expected for a random mutagen in *Clostridium*, considering that around 80 % of the clostridial genome is protein coding. These observations make the *mariner*-transposon system a favourable tool for forward genetics in the genus compared to e.g. Tn916 which shows, apart from low efficiency transformation (due to large size) and low efficiency transposition (Schwarz et al., 2015), insertional ‘hot spots’ (Hussain et al., 2010) and a predisposition for multiple insertions (Lin and Johnson, 1991; Schwarz et al., 2015).

It was further demonstrated that the system could be used to isolate specific mutants with a particular phenotype in *C. acetobutylicum* and *C. sporogenes*, by screening for clones affected in sporulation/ germination as well as autotrophic strains which could no longer grow on minimal media. The intention was merely to exemplify that the method could be used for such a purpose, rather than gather fundamental information on the nature of the cellular components identified. Nevertheless, some preliminary conclusions can be made on the nature of the mutants identified. Thus, the transposon insertions, in the main sporulation/ germination mutants obtained, were in genes that encoded proteins associated with *B. subtilis* spores: spore protein YkuD in *C. acetobutylicum*, and SpoVA (stage V sporulation protein) in *C. sporogenes*. The reason for the absence of colony forming units after heat shock when the



*C. sporogenes* CS\_2380 gene (encoding a putative flavodoxin oxidoreductase) was inactivated was less clear. Auxotrophic mutant screening has clearly implicated the *C. acetobutylicum* gene CA\_c0971 (encoding CitB, aconitase) to be involved in glutamate biosynthesis, and most likely glutamine, but has called into question whether glutamate is formed from  $\alpha$ -ketoglutarate or if this is a simple annotation error in this organism. In the case of *C. sporogenes*, the reason why inactivation of the three genes (CS\_1694, a collagenase; CS\_3053, a putative membrane protein, and CS\_2157, a protein annotated as a KWG leptospira repeat protein) prevented growth on a minimal media was largely unclear, with the possible exception of the collagenase where a possible role in the uptake of small peptides and amino acids may exist.

Having proven the *Clostridium* transposon system, pMTL-YZ14, to work in the established industrial organisms *C. acetobutylicum* and *C. sporogenes* it was developed for use in *C. pasteurianum*. It was shown that the  $P_{tcdB}$ -*tcdR* expression system and the conditional replicon are functioning in this organism. However, when applying the complete pMTL-YZ14 system to *C. pasteurianum* non-permissive conditions did not lead to plasmid loss. A possibility for this could be that the plasmid degenerated in the conditional replicon region while harboured in *E. coli*. This could happen due to an active transposase since it was shown here that  $P_{tcdB}$  is active in *E. coli*. Before further investigation could be undertaken another approach was attempted using a suicide transposon plasmid pMTL-GL15, which is essentially pMTL-YZ14 without Gram-positive replicon, hence suicidal in clostridia. A suicide approach is advantageous as every mutant isolated after transformation is a transposon insertional mutant. This approach overcomes the need for plasmid curation from the transposon mutants which makes the generation of transposon libraries much faster. However, due to the low efficiency of transposition of our system ( $2.6 (\pm 0.6) \times 10^{-4}$  and  $3.2 (\pm 0.5) \times 10^{-4}$  in *C. acetobutylicum* and *C. sporogenes*, respectively), transformation efficiency in the host needs to be high. With the hypertransformable clones of *C. pasteurianum* isolated in this study, transformation efficiencies of up to  $6.6 \times 10^5$  cfu/ $\mu$ g DNA were reached. It was considered sufficient to subject the suicide transposon delivery to a trial. Indeed, transposon mutants were obtained at the rate of 10 cfu/ $\mu$ g, establishing the *mariner*-transposon system as a forward genetics tool for *C. pasteurianum*.

### 6.3 Limitations of This Study and Future Work

As part of a wider effort to develop and apply forward and reverse genetic tools to *Clostridium* species, this study proved to be invaluable to the ongo-

ing endeavour and further research. However, due to the time restraint of a PhD, several aspects of this study could be expanded. Further work is needed to investigate gene CLPA\_c30550 regarding foreign DNA transfer and DNA transformation efficiency. Restriction and CRISPR systems, which both naturally hinder foreign DNA from entering the host cell, are adapted to be valuable tools in modern molecular biology. In this light, and considering the putative DNA binding ability of protein CLPA\_c30550 suggested by *in silico* analysis, it is possible for the protein to have similar applications when studied fully. The same applies to the enzymes involved in hypertransformability of *C. pasteurianum*-H1. This study did not investigate which of the two genes, if not both, are responsible for the observed phenotype. This could be narrowed down and tested similar to CLPA\_c30550 and further enzymatic studies can be envisaged to uncover the enzymes involvement in DNA transformation.

Deletion mutants of *C. pasteurianum* were only analysed for their dissolved metabolite levels. However, in the *rex* (Wietzke and Bahl, 2012) and especially the *hyd* deletion mutant, hydrogen production can be expected to be altered drastically. Further fermentations with off-gas analysis should be undertaken which at the time of this study were not possible due to missing equipment. In the *hyd* mutant this could give indications on the function of other hydrogenases in the *C. pasteurianum* genome. Furthermore, with this study's *hyd* mutant showing a superior phenotype in solvent production, deletions of the other hydrogenases are speculated to have a similar effect or accumulating deletion of hydrogenases can further improve the solvent titres. This was shown by Dabrock et al. (1992), who found a large increase in solvent titres when inhibiting all [Fe]-hydrogenases in *C. pasteurianum* by CO sparging of the medium with glucose as the sole carbon source.

Even though the *in silico* analysis of Rex recognition sites and target genes revealed putatively regulated genes which were in part confirmed by metabolite analysis, no direct experimental evidence for interactions of the *C. pasteurianum* Rex with these promoter regions was given. Gel shift assays showing protein-DNA interaction under different NADH/NAD<sup>+</sup> ratios could be done similar to Pagels et al. (2010) to confirm gene regulations and potentially confirm a hypothesised Rex self-regulation. Furthermore qPCR or RNA-seq could be envisaged to analyse expression levels of genes, especially those involved in central metabolic pathways, in the *rex* inactivational strain. With foot-printing analysis the precise Rex recognition site can be determined and potentially confirm the predicted recognition site.

Forward genetic approaches in *C. pasteurianum* need to be further improved. Problems with the conditional replicon of the pMTL-YZ14 system

should be re-assessed and eliminated. The efficiency of the suicide system can be further improved by optimising the electroporation protocol, increasing the efficiency sufficiently for screening applications. However, modern transposon methods like transposon-insertion sequencing (Barquist et al., 2013) rely on large libraries ( $10^6$  mutants) which are unlikely to be reached with the suicide approach but could be achieved with the pMTL-YZ14 conditional replicon system with ease.

## 6.4 Concluding Remarks

This work laid the foundation for genetic engineering in *C. pasteurianum* by the discovery of a hypertransformable sub-population of the wild type, using a rational screening approach. The clones differed to the wild type only in small genotypic variants allowing high efficiency electroporation. The fermentative phenotypes of the clones were unchanged making them ideal starting strains for further genetic engineering to fit industrial requirements.

With the establishment of reproducible, high frequency plasmid transformation, the clostridial roadmap (Minton et al., 2016) was implemented in *C. pasteurianum*. A reverse genetic approach was employed to establish gene function. Allelic exchange gene knock-out procedures were used to delete genes of the central metabolic pathways and allele coupled exchange was used to complement the deletion mutants (Schwarz et al., 2017). The strains obtained were analysed in laboratory scale fermentations. With the *dhaBCE* deletion mutant it was, for the first time in *C. pasteurianum*, possible to entirely eliminate 1,3-PDO production, showing the reductive pathway to be crucial for growth of the organism on glycerol. With the *rex* and *hyd* mutants this work presents two superior solvent producers. The mutants show 140 % increased butanol production and 270 % increased ethanol production in *rex* and *hyd*, respectively, while at the same time showing reduced by-product (1,3-PDO) formation. These mutants represent the first steps towards improvement of *C. pasteurianum* as a chassis for the industrial production of solvents from glycerol.

To allow forward genetic approaches in *C. pasteurianum* a *mariner*-transposon system was adapted. Due to the initial absence of an efficient transformation procedure in *C. pasteurianum*, the system was exemplified in *C. acetobutylicum* and *C. sporogenes*. In both species a system conditional for replication and employing an alternate  $\sigma$  factor, TcdR, was shown to yield random transposon mutant libraries and sporulation/ germination as well as auxotrophic transposon mutants were isolated. The system was further developed in *C. pasteurianum* with the slight alteration of using a suicide instead of a

conditional plasmid.

This study, therefore, presents an essential forward genetics procedure for industrially important *Clostridium* species and a comprehensive genetic engineering approach for the important biofuel producer *C. pasteurianum*.

# Bibliography

- Jamal Abrini, Henry Naveau, and Edmond-Jacques Nyns. *Clostridium autoethanogenum*, sp. nov., an anaerobic bacterium that produces ethanol from carbon monoxide. *Archives of Microbiology*, 161(4):345–351, 1994.
- M. A. Al-Hinai, A. G. Fast, and E. T. Papoutsakis. Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. *Appl Environ Microbiol*, 78(22):8112–21, 2012.
- João RM Almeida, Léia CL Fávaro, and Betania F Quirino. Biodiesel biorefinery: opportunities and challenges for microbial production of fuels and chemicals from glycerol waste. *Biotechnol. Biofuels*, 5(48):48, 2012.
- S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. Basic local alignment search tool. *J Mol Biol*, 215(3):403–10, 1990.
- Daniel Amador-Noguez, Xiao-Jiang Feng, Jing Fan, Nathaniel Roquet, Herschel Rabitz, and Joshua D Rabinowitz. Systems-level metabolic flux profiling elucidates a complete, bifurcated tricarboxylic acid cycle in *Clostridium acetobutylicum*. *Journal of bacteriology*, 192(17):4452–4461, 2010.
- Yasuhiro Anraku and Robert B Gennis. The aerobic respiratory chain of *Escherichia coli*. *Trends in Biochemical Sciences*, 12:262–266, 1987.
- EF Aransiola, TV Ojumu, OO Oyekola, TF Madzimbamuto, and DIO Ikhu-Omoregbe. A review of current technology for biodiesel production: State of the art. *Biomass and bioenergy*, 61:276–297, 2014.
- Somasundar Ashok, Subramanian Mohan Raj, Chelladurai Rathnasingh, and Sunghoon Park. Development of recombinant *Klebsiella pneumoniae* delta*dhaT* strain for the co-production of 3-hydroxypropionic acid and 1, 3-propanediol from glycerol. *Applied microbiology and biotechnology*, 90(4):1253–1265, 2011.
- Trond Erik Vee Aune and Finn Lillelund Aachmann. Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. *Applied microbiology and biotechnology*, 85(5):1301–1313, 2010.
- Oswald T Avery, Colin M MacLeod, and Maclyn McCarty. Studies on the chemical nature of the substance inducing transformation of pneumococcal types induction of transformation by a desoxyribonucleic acid fraction isolated

- from pneumococcus type III. *The Journal of experimental medicine*, 79(2): 137–158, 1944.
- M. M. Awad and J. I. Rood. Isolation of alpha-toxin, theta-toxin and kappa-toxin mutants of *Clostridium perfringens* by Tn916 mutagenesis. *Microbial Pathogenesis*, 22(5):275–284, 1997.
- BL Babb, HJ Collett, SJ Reid, and DR Woods. Transposon mutagenesis of *Clostridium acetobutylicum* P262 isolation and characterization of solvent deficient and metronidazole-resistant mutants. *FEMS Microbiol Lett*, 114(3): 343–348, 1993.
- Shirley H Baer, Hans P Blaschek, and Terrance L Smith. Effect of butanol challenge and temperature on lipid composition and membrane fluidity of butanol-tolerant *Clostridium acetobutylicum*. *Applied and environmental microbiology*, 53(12):2854–2861, 1987.
- Hubert Bahl, Wolfram Andersch, Konstantin Braun, and Gerhard Gottschalk. Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium acetobutylicum* grown in continuous culture. *European journal of applied microbiology and biotechnology*, 14(1):17–20, 1982.
- Fabien Barbirato, Jean Philippe Grivet, Philippe Soucaille, and Andre Bories. 3-hydroxypropionaldehyde, an inhibitory metabolite of glycerol fermentation to 1,3-propanediol by enterobacterial species. *Applied and Environmental Microbiology*, 62(4):1448–1451, 1996.
- H. A. Barker and J. V. Beck. *Clostridium acidi-uridi* and *Clostridium cylindrosporum*, organisms fermenting uric acid and some other purines. *Journal of Bacteriology*, 43(3):291–304, 1942.
- L. Barquist, C. J. Boinett, and A. K. Cain. Approaches to querying bacterial genomes with transposon-insertion sequencing. *RNA Biol*, 10(7):1161–9, 2013.
- Rodolphe Barrangou and Luciano A Marraffini. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Molecular cell*, 54(2):234–244, 2014.
- A. Baudin, O. Ozierkalogeropoulos, A. Denouel, F. Lacroute, and C. Cullin. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Research*, 21(14):3329–3330, 1993.
- E. A. Bayer, E. Setter, and R. Lamed. Organization and distribution of the cellulosome in *Clostridium thermocellum*. *Journal of Bacteriology*, 163(2):552–559, 1985.
- T. A. Bickle and D. H. Kruger. Biology of DNA restriction. *Microbiological Reviews*, 57(2):434–450, 1993.
- H. Biebl. Fermentation of glycerol by *Clostridium pasteurianum* - batch and continuous culture studies. *Journal of Industrial Microbiology & Biotechnology*, 27(1):18–26, 2001.

- H. Biebl, S. Marten, H. Hippe, and W. D. Deckwer. Glycerol conversion to 1,3-propanediol by newly isolated Clostridia. *Applied Microbiology and Biotechnology*, 36(5):592–597, 1992.
- H. Biebl, K. Menzel, A. P. Zeng, and W. D. Deckwer. Microbial production of 1,3-propanediol. *Applied Microbiology and Biotechnology*, 52(3):289–297, 1999.
- Hanno Biebl and Norbert Pfennig. Isolation of members of the family Rhodospirillaceae. In *The prokaryotes*, pages 267–273. Springer, 1981.
- Ranjita Biswas, Tianyong Zheng, Daniel G Olson, Lee R Lynd, and Adam M Guss. Elimination of hydrogenase active site assembly blocks H<sub>2</sub> production and increases ethanol yield in *Clostridium thermocellum*. *Biotechnology for biofuels*, 8(1):1, 2015.
- Laurent Bouillaut, Shonna M McBride, and Joseph A Sorg. Genetic manipulation of *Clostridium difficile*. *Current protocols in microbiology*, pages 9A–2, 2011.
- Dimitris Brekasis and Mark SB Paget. A novel sensor of NADH/NAD<sup>+</sup> redox poise in *Streptomyces coelicolor* A3 (2). *The EMBO journal*, 22(18):4856–4865, 2003.
- J. M. Bruno-Barcena, M. S. Chinn, and A. M. Grunden. Genome sequence of the autotrophic acetogen *Clostridium autoethanogenum* JA1-1 strain DSM 10061, a producer of ethanol from carbon monoxide. *Genome Announc*, 1(4), 2013.
- Magdalena Calusinska, Thomas Happe, Bernard Joris, and Annick Wilmotte. The surprising diversity of clostridial hydrogenases: a comparative genomic perspective. *Microbiology*, 156(6):1575–1588, 2010.
- S. T. Cartman and N. P. Minton. A *mariner*-based transposon system for in vivo random mutagenesis of *Clostridium difficile*. *Appl Environ Microbiol*, 76(4):1103–9, 2010.
- S. T. Cartman, M. L. Kelly, D. Heeg, J. T. Heap, and N. P. Minton. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the *tcdC* genotype and toxin production. *Appl Environ Microbiol*, 78(13):4683–90, 2012.
- Tim J Carver, Kim M Rutherford, Matthew Berriman, Marie-Adele Rajandream, Barclay G Barrell, and Julian Parkhill. ACT: the Artemis comparison tool. *Bioinformatics*, 21(16):3422–3423, 2005.
- A. Chatzifragkou and S. Papanikolaou. Effect of impurities in biodiesel-derived waste glycerol on the performance and feasibility of biotechnological processes. *Appl Microbiol Biotechnol*, 95(1):13–27, 2012.
- Inês Chen, Peter J Christie, and David Dubnau. The ins and outs of DNA transfer in bacteria. *Science*, 310(5753):1456–1460, 2005.

- J. M. Clomburg and R. Gonzalez. Anaerobic fermentation of glycerol: a platform for renewable fuels and chemicals. *Trends in Biotechnology*, 31(1):20–28, 2013.
- M. D. Collins, P. A. Lawson, A. Willems, J. J. Cordoba, J. Fernandezgarayzabal, P. Garcia, J. Cai, H. Hippe, and J. A. E. Farrow. The phylogeny of the genus *Clostridium* - Proposal of 5 new genera and 11 new species combinations. *International Journal of Systematic Bacteriology*, 44(4):812–826, 1994.
- Le Cong, F Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D Hsu, Xuebing Wu, Wenyan Jiang, Luciano A Marraffini, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339(6121):819–823, 2013.
- T. M. Cook, R. T. Protheroe, and J. M. Handel. Tetanus: a review of the literature. *British Journal of Anaesthesia*, 87(3):477–487, 2001.
- C. M. Cooksley, I. J. Davis, K. Winzer, W. C. Chan, M. W. Peck, and N. P. Minton. Regulation of neurotoxin production and sporulation by a putative *agrBD* signaling system in proteolytic *Clostridium botulinum*. *Appl Environ Microbiol*, 76(13):4448–60, 2010.
- Clare M Cooksley, Ying Zhang, Hengzheng Wang, Stephanie Redl, Klaus Winzer, and Nigel P Minton. Targeted mutagenesis of the *Clostridium acetobutylicum* acetone-butanol-ethanol fermentation pathway. *Metabolic engineering*, 14(6):630–641, 2012.
- Emmanuel Cornillot, Ramesh V Nair, Eleftherios T Papoutsakis, and Philippe Soucaille. 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. *Journal of bacteriology*, 179(17):5442–5447, 1997.
- Adam J Cornish, Katrin Gärtner, Hui Yang, John W Peters, and Eric L Hegg. Mechanism of proton transfer in [FeFe]-hydrogenase from *Clostridium pasteurianum*. *Journal of Biological Chemistry*, 286(44):38341–38347, 2011.
- Adam James Cornish, Bojana Ginovska, Adam Thelen, Júlio Cosme Santos Da Silva, Thereza A Soares, Simone Raugai, Michel Dupuis, Wendy J Shaw, and Eric L Hegg. Single amino acid modifications reveal additional controls on the proton pathway of [FeFe]-hydrogenase. *Biochemistry*, 55(22):3165–3173, 2016.
- Cássia Aparecida Rabelo Corrêa, Sérgio Francisco de Aquino, Paula Cristina de Paula Caldas, and Silvana de Queiroz Silva. Uso de extrato de levedura como fonte de carbono e de mediadores redox, para a degradação anaeróbia de corante azo. *Engenharia Sanitaria e Ambiental*, 14:559 – 568, 12 2009. ISSN 1413-4152.
- Sharon D Cosloy and Michio Oishi. Genetic transformation in *Escherichia coli* K12. *Proceedings of the National Academy of Sciences*, 70(1):84–87, 1973.



- Gavin E Crooks, Gary Hon, John-Marc Chandonia, and Steven E Brenner. Weblogo: a sequence logo generator. *Genome research*, 14(6):1188–1190, 2004.
- Scott B Crown, Dinesh C Indurthi, Woo Suk Ahn, Jungik Choi, Eleftherios T Papoutsakis, and Maciek R Antoniewicz. Resolving the TCA cycle and pentose-phosphate pathway of *Clostridium acetobutylicum* ATCC 824: Isotopomer analysis, *in vitro* activities and expression analysis. *Biotechnology journal*, 6(3):300–305, 2011.
- B. Dabrock, H. Bahl, and G. Gottschalk. Parameters affecting solvent production by *Clostridium pasteurianum*. *Appl Environ Microbiol*, 58(4):1233–9, 1992.
- K. A. Datsenko and B. L. Wanner. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12):6640–6645, 2000.
- Rathin Datta and JG Zeikus. Modulation of acetone-butanol-ethanol fermentation by carbon monoxide and organic acids. *Applied and environmental microbiology*, 49(3):522–529, 1985.
- DBEIS Department for Business Energy & Industrial Strategy. Energy consumption in the UK (ECUK). *Online at gov.co.uk*, 2016.
- Ruchir P Desai, Latonia M Harris, Neil E Welker, and Eleftherios T Papoutsakis. Metabolic flux analysis elucidates the importance of the acid-formation pathways in regulating solvent production by *Clostridium acetobutylicum*. *Metabolic engineering*, 1(3):206–213, 1999.
- NM Dixon, RW Lovitt, DB Kell, and JG Morris. Effects of pCO<sub>2</sub> on the growth and metabolism of *Clostridium sporogenes* NCIB 8053 in defined media. *Journal of applied bacteriology*, 63(2):171–182, 1987.
- Jacqueline Dornan, Heather Grey, and Julia M Richardson. Structural role of the flanking DNA in *mariner* transposon excision. *Nucleic acids research*, 43(4):2424–2432, 2015.
- David Dubnau. DNA uptake in bacteria. *Annual Reviews in Microbiology*, 53(1):217–244, 1999.
- Peter Dürre. Formation of solvents in Clostridia. In *Handbook on Clostridia*, pages 671–693. CRC Press, 2005.
- Peter Dürre and Concha Hollergschwandner. Initiation of endospore formation in *Clostridium acetobutylicum*. *Anaerobe*, 10(2):69–74, 2004.
- Peter Dürre, Michael Böhringer, Stephan Nakotte, Steffen Schaffer, Kai Thormann, and Brigitte Zickner. Transcriptional regulation of solventogenesis in *Clostridium acetobutylicum*. *Journal of molecular microbiology and biotechnology*, 4(3):295–300, 2002.

- Muhammad Ehsaan, Wouter Kuit, Ying Zhang, Stephen T Cartman, John T Heap, Klaus Winzer, and Nigel P Minton. Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. *Biotechnology for biofuels*, 9(1):1, 2016.
- P. J. Enyeart, G. Mohr, A. D. Ellington, and A. M. Lambowitz. Biotechnological applications of mobile group II introns and their reverse transcriptases: gene targeting, RNA-seq, and non-coding RNA analysis. *Mobile DNA*, 5(1):2, 2014.
- ECLP Equistar Chemicals LP. Ethyl-alcohol handbook. *Online at industrialethanol.com*, 6th Edition, 2003.
- T. Ezeji, N. Qureshi, and H. P. Blaschek. Butanol production from agricultural residues: Impact of degradation products on *Clostridium beijerinckii* growth and butanol fermentation. *Biotechnol Bioeng*, 97(6):1460–9, 2007.
- Thaddeus Ezeji, Caroline Milne, Nathan D. Price, and Hans P. Blaschek. Achievements and perspectives to overcome the poor solvent resistance in acetone and butanol-producing microorganisms. *Applied Microbiology and Biotechnology*, 85(6):1697–1712, 2010. ISSN 1432-0614. doi: 10.1007/s00253-009-2390-0.
- Thaddeus C Ezeji, Nasib Qureshi, and Hans P Blaschek. Industrially relevant fermentations. In *Handbook on Clostridia*, pages 998–1015. CRC Press, 2005.
- L. Feinberg, J. Foden, T. Barrett, K. W. Davenport, D. Bruce, C. Detter, R. Tapia, C. Han, A. Lapidus, S. Lucas, J. F. Cheng, S. Pitluck, T. Woyke, N. Ivanova, N. Mikhailova, M. Land, L. Hauser, D. A. Argyros, L. Goodwin, D. Hogsett, and N. Caiazza. Complete genome sequence of the cellulolytic thermophile *Clostridium thermocellum* DSM1313. *J Bacteriol*, 193(11):2906–7, 2011.
- Robert G Forage and Michael A Foster. Glycerol fermentation in *Klebsiella pneumoniae*: functions of the coenzyme B12-dependent glycerol and diol dehydratases. *Journal of bacteriology*, 149(2):413–419, 1982.
- C. L. Frazier, J. San Filippo, A. M. Lambowitz, and D. A. Mills. Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: Generation of stable insertions without selection. *Applied and Environmental Microbiology*, 69(2):1121–1128, 2003.
- B. Freedman, E. H. Pryde, and T. L. Mounts. Variables affecting the yields of fatty esters from transesterified vegetable oils. *Journal of the American Oil Chemists Society*, 61(10):1638–1643, 1984.
- Lewis Y Geer, Michael Domrachev, David J Lipman, and Stephen H Bryant. CDART: protein homology by domain architecture. *Genome research*, 12(10):1619–1623, 2002.
- H. A. George, J. L. Johnson, W. E. C. Moore, L. V. Holdeman, and J. S. Chen. Acetone, isopropanol, and butanol production by *Clostridium beijerinckii* (syn *Clostridium butylicum*) and *Clostridium aurantibutyricum*. *Applied and Environmental Microbiology*, 45(3):1160–1163, 1983.

- María González-Pajuelo, Isabelle Meynial-Salles, Filipa Mendes, Philippe Soucaille, and Isabel Vasconcelos. Microbial conversion of glycerol to 1,3-propanediol: physiological comparison of a natural producer, *Clostridium butyricum* VPI 3266, and an engineered strain, *Clostridium acetobutylicum* DG1 (pSPD5). *Applied and environmental microbiology*, 72(1):96–101, 2006.
- Lalitha Devi Gottumukkala, Binod Parameswaran, Sajna Kuttavan Valappil, Kuttiraja Mathiyazhakan, Ashok Pandey, and Rajeev Kumar Sukumaran. Biobutanol production from rice straw by a non acetone producing *Clostridium sporogenes* BE01. *Bioresource technology*, 145:182–187, 2013.
- Peter L Graumann. *Bacillus subtilis* SMC is required for proper arrangement of the chromosome and for efficient segregation of replication termini but not for bipolar movement of newly duplicated origin regions. *Journal of bacteriology*, 182(22):6463–6471, 2000.
- Fred Griffith. The significance of pneumococcal types. *Journal of Hygiene*, 27(02):113–159, 1928.
- Phalguni Gupta and Yue Chen. Chromosomal engineering of *Clostridium perfringens* using group II introns. In G. D. Davis and K. J. Kayser, editors, *Methods in Molecular Biology*, volume 435 of *Methods in Molecular Biology*, pages 217–228. Humana Press Inc, 2008.
- Hervee Guyomard, Agneta Forslund, and Yves Dronne. Biofuels and world agricultural markets: Outlook for 2020 and 2050. *InTech*, 2011.
- Smita Gyan, Yoshihiko Shiohira, Ichiro Sato, Michio Takeuchi, and Tsutomu Sato. Regulatory loop between redox sensing of the NADH/NAD<sup>+</sup> ratio by Rex (YdiH) and oxidation of NADH by NADH dehydrogenase Ndh in *Bacillus subtilis*. *Journal of bacteriology*, 188(20):7062–7071, 2006.
- J. Hansen, M. Sato, and R. Ruedy. Perception of climate change. *Proc Natl Acad Sci U S A*, 109(37):E2415–23, 2012.
- Dean J Harrington. Bacterial collagenases and collagen-degrading enzymes and their potential role in human disease. *Infection and immunity*, 64(6):1885, 1996.
- Latoria M Harris, Neil E Welker, and Eleftherios T Papoutsakis. Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *Journal of bacteriology*, 184(13):3586–3597, 2002.
- LM Harris, L Blank, RP Desai, NE Welker, and ET Papoutsakis. Fermentation characterization and flux analysis of recombinant strains of *Clostridium acetobutylicum* with an inactivated *solR* gene. *Journal of Industrial Microbiology and Biotechnology*, 27(5):322–328, 2001.
- Maris GN Hartmanis and Thressa C Stadtman. Diol metabolism and diol dehydratase in *Clostridium glycolicum*. *Archives of biochemistry and biophysics*, 245(1):144–152, 1986.

- Maris GN Hartmanis, Tomas Klason, and Sten Gatenbeck. Uptake and activation of acetate and butyrate in *Clostridium acetobutylicum*. *Applied microbiology and biotechnology*, 20(1):66–71, 1984.
- B. G. Harvey and H. A. Meylemans. The role of butanol in the development of sustainable fuel technologies. *Journal of Chemical Technology and Biotechnology*, 86(1):2–9, 2011.
- J. T. Heap, O. J. Pennington, S. T. Cartman, G. P. Carter, and N. P. Minton. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J Microbiol Methods*, 70(3):452–64, 2007.
- J. T. Heap, O. J. Pennington, S. T. Cartman, and N. P. Minton. A modular system for *Clostridium* shuttle plasmids. *Journal of Microbiological Methods*, 78(1):79–85, 2009.
- J. T. Heap, S. A. Kuehne, M. Ehsaan, S. T. Cartman, C. M. Cooksley, J. C. Scott, and N. P. Minton. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *J Microbiol Methods*, 80(1):49–55, 2010.
- J. T. Heap, M. Ehsaan, C. M. Cooksley, Y. K. Ng, S. T. Cartman, K. Winzer, and N. P. Minton. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic Acids Res*, 40(8):e59, 2012.
- J. T. Heap, J. Theys, M. Ehsaan, A. M. Kubiak, L. Dubois, K. Paesmans, L. V. Mellaert, R. Knox, S. A. Kuehne, P. Lambin, and N. P. Minton. Spores of *Clostridium* engineered for clinical efficacy and safety cause regression and cure of tumors *in vivo*. *Oncotarget*, 5:1761–1769, 2014.
- D Michael Heinekey. Hydrogenase enzymes: recent structural studies and active site models. *Journal of Organometallic Chemistry*, 694(17):2671–2680, 2009.
- Russell Higuchi, Barbara Krummel, and Randall Saiki. A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic acids research*, 16(15):7351–7367, 1988.
- Falk Hillmann, Christina Döring, Oliver Riebe, Armin Ehrenreich, Ralf-Jörg Fischer, and Hubert Bahl. The role of PerR in O<sub>2</sub>-affected gene expression of *Clostridium acetobutylicum*. *Journal of bacteriology*, 191(19):6082–6093, 2009.
- Tatsuya Hirano. The ABCs of SMC proteins: two-armed ATPases for chromosome condensation, cohesion, and repair. *Genes & Development*, 16(4):399–414, 2002.
- Henry D. Hoberman and D. Rittenberg. Biological catalysis of the exchange reaction between water and hydrogen. *Journal of Biological Chemistry*, 147(1):211–227, 1943.
- R. Holliday. A mechanism for gene conversion in fungi. *Genetics Research*, 5(02):282–304, 1964.

- Liejie Hu, Haiyan Huang, Hengxin Yuan, Fei Tao, Huijun Xie, and Shuning Wang. Rex in *Clostridium kluyveri* is a global redox-sensing transcriptional regulator. *Journal of Biotechnology*, 233:17–25, 2016.
- I-Hsiu Huang, Michael Waters, Roberto R Grau, and Mahfuzur R Sarker. Disruption of the gene (*spo0A*) encoding sporulation transcription factor blocks endospore formation and enterotoxin production in enterotoxigenic *Clostridium perfringens* type A. *FEMS microbiology letters*, 233(2):233–240, 2004.
- Michael HW Hüsemann and Eleftherios T Papoutsakis. Comparison between *in vivo* and *in vitro* enzyme activities in continuous and batch fermentations of *Clostridium acetobutylicum*. *Applied microbiology and biotechnology*, 30(6):585–595, 1989.
- H. A. Hussain, A. P. Roberts, R. Whalan, and P. Mullany. Transposon mutagenesis in *Clostridium difficile*. *Clostridium Difficile: Methods and Protocols*, 646:203–211, 2010.
- IEA International Energy Agency. Technology roadmap - Biofuels for transport. *Online Publication*, 2011.
- IEA International Energy Agency. Redrawing the energy-climate map. *Online Publication*, 2013.
- IEA International Energy Agency. Energy technology perspectives 2016. *Online Publication*, 2016.
- A. Ishizaki, S. Michiwaki, E. Crabbe, G. Kobayashi, K. Sonomoto, and S. Yoshino. Extractive acetone-butanol-ethanol fermentation using methylated crude palm oil as extractant in batch culture of *Clostridium saccharoperbutylacetonicum* N1-4 (ATCC 13564). *Journal of Bioscience and Bioengineering*, 87(3):352–356, 1999.
- Titipong Issariyakul and Ajay K. Dalai. Biodiesel from vegetable oils. *Renewable and Sustainable Energy Reviews*, 31:446–471, 2014.
- Z. Ivics, P. B. Hackett, R. H. Plasterk, and Z. Izsvak. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*, 91(4):501–10, 1997.
- T. O. Jensen, T. Kvist, M. J. Mikkelsen, P. V. Christensen, and P. Westermann. Fermentation of crude glycerol from biodiesel production by *Clostridium pasteurianum*. *J Ind Microbiol Biotechnol*, 39(5):709–17, 2012a.
- T. O. Jensen, T. Kvist, M. J. Mikkelsen, and P. Westermann. Production of 1,3-PDO and butanol by a mutant strain of *Clostridium pasteurianum* with increased tolerance towards crude glycerol. *AMB Express*, 2(1):44, 2012b.
- Wei Jiang, Shizhen Wang, Yuanpeng Wang, and Baishan Fang. Key enzymes catalyzing glycerol to 1,3-propanediol. *Biotechnology for biofuels*, 9(1):1, 2016.

- Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A Doudna, and Emmanuelle Charpentier. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096):816–821, 2012.
- Martin Jinek, Alexandra East, Aaron Cheng, Steven Lin, Enbo Ma, and Jennifer Doudna. RNA-programmed genome editing in human cells. *elife*, 2:e00471, 2013.
- Erin E Johnson and Lars Rehm. The role of 1,3-propanediol production in fermentation of glycerol by *Clostridium pasteurianum*. *Bioresource technology*, 209:1–7, 2016.
- David T Jones and Stefanie Keis. Species and strain identification methods. In P Duerre, editor, *Handbook on Clostridia*, pages 3–20. CRC Press Boca Raton, FL, 2005.
- David T Jones and David R Woods. Acetone-butanol fermentation revisited. *Microbiological reviews*, 50(4):484, 1986.
- W. C. Kao, D. S. Lin, C. L. Cheng, B. Y. Chen, C. Y. Lin, and J. S. Chang. Enhancing butanol production with *Clostridium pasteurianum* CH4 using sequential glucose-glycerol addition and simultaneous dual-substrate cultivation strategies. *Bioresour Technol*, 135:324–30, 2013.
- Mehak Kaushal, Saumya Ahlawat, Mayurketan Mukherjee, Muthusivaramapandian Muthuraj, Gargi Goswami, and Debasish Das. Substrate dependent modulation of butanol to ethanol ratio in non-acetone forming *Clostridium sporogenes* NCIM 2918. *Bioresource Technology*, 225:349–358, 2017.
- S. Keis, C. F. Bennett, V. K. Ward, and D. T. Jones. Taxonomy and phylogeny of industrial solvent-producing clostridia. *International Journal of Systematic Bacteriology*, 45(4):693–705, 1995.
- S. Keis, J. T. Sullivan, and D. T. Jones. Physical and genetic map of the *Clostridium saccharobutylicum* (formerly *Clostridium acetobutylicum*) NCP 262 chromosome. *Microbiology-Sgm*, 147:1909–1922, 2001.
- Lawrence A Kelley, Stefans Mezulis, Christopher M Yates, Mark N Wass, and Michael JE Sternberg. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, 10(6):845–858, 2015.
- Byung Hong Kim, Para Bellows, Rathin Datta, and JG Zeikus. Control of carbon and electron flow in *Clostridium acetobutylicum* fermentations: utilization of carbon monoxide to inhibit hydrogen production and to enhance butanol yields. *Applied and environmental microbiology*, 48(4):764–770, 1984.
- Mathias Klein, Marion B Ansorge-Schumacher, Markus Fritsch, and Winfried Hartmeier. Influence of hydrogenase overexpression on hydrogen production of *Clostridium acetobutylicum* DSM 792. *Enzyme and Microbial Technology*, 46(5):384–390, 2010.

- Anja Knietsch, Susanne Bowien, Gregg Whited, Gerhard Gottschalk, and Rolf Daniel. Identification and characterization of coenzyme B12-dependent glycerol dehydratase-and diol dehydratase-encoding genes from metagenomic DNA libraries derived from enrichment cultures. *Applied and environmental microbiology*, 69(6):3048–3060, 2003.
- G. Knothe, J. Krahl, and J. Van Gerpen. *The Biodiesel Handbook, Second Edition*. Taylor & Francis, 2010.
- G. Kocar and N. Civas. An overview of biofuels from energy crops: Current status and future prospects. *Renewable & Sustainable Energy Reviews*, 28: 900–916, 2013.
- Takeko Kodama, Hiromu Takamatsu, Kei Asai, Naotake Ogasawara, Yoshito Sadaie, and Kazuhito Watabe. Synthesis and characterization of the spore proteins of *Bacillus subtilis* YdhD, YkuD, and YkvP, which carry a motif conserved among cell wall binding proteins. *Journal of biochemistry*, 128(4): 655–663, 2000.
- M. Köpke, C. Held, S. Hujer, H. Liesegang, A. Wiezer, A. Wollherr, A. Ehrenreich, W. Liebl, G. Gottschalk, and P. Dürre. *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29):13087–13092, 2010.
- Tadej Kotnik, Wolfgang Frey, Martin Sack, Saša Haberl Meglič, Matjaž Peterka, and Damijan Miklavčič. Electroporation-based applications in biotechnology. *Trends in biotechnology*, 33(8):480–488, 2015.
- S. C. Kowalczykowski, D. A. Dixon, A. K. Eggleston, S. D. Lauder, and W. M. Rehrauer. Biochemistry of homologous recombination in *Escherichia coli*. *Microbiological Reviews*, 58(3):401–465, 1994.
- George A. Kraus. Synthetic methods for the preparation of 1,3-propanediol. *CLEAN - Soil, Air, Water*, 36(8):648–651, 2008.
- A. M. Kubiak. *The use of Clostridium sporogenes in directed enzyme pro-drug therapy*. PhD thesis, University of Nottingham, 2013.
- Aleksandra M Kubiak, Anja Poehlein, Patrick Budd, Sarah A Kuehne, Klaus Winzer, Jan Theys, Philip Lambin, Rolf Daniel, and Nigel P Minton. Complete genome sequence of the nonpathogenic soil-dwelling bacterium *Clostridium sporogenes* strain NCIMB 10696. *Genome announcements*, 3(4):e00942–15, 2015.
- Sarah A Kuehne and Nigel P Minton. Clostron-mediated engineering of *Clostridium*. *Bioengineered*, 3(4):247–254, 2012.
- Wouter Kuit. *Metabolic engineering of acid formation in Clostridium acetobutylicum*. PhD thesis, University of Wageningen, 2013.

- D Kupper, JG Zhou Zhou, B., P Venetianer, and A Kiss. Cloning and structure of the BepI modification methylase. *Nucleic acid research*, 17(3):1077–1088, FEB 11 Kupper1989. ISSN 0305-1048. doi: 10.1093/nar/17.3.1077.
- Angela La Torre, Daniela Bassi, Teresa Zotta, Luigi Orrù, Antonella Lamontana, and Pier Sandro Cocconcelli. Draft genome sequence of *Clostridium sporogenes* strain UC9000 isolated from raw milk. *Genome announcements*, 4(2):e00244–16, 2016.
- D. J. Lampe, B. J. Akerley, E. J. Rubin, J. J. Mekalanos, and H. M. Robertson. Hyperactive transposase mutants of the Himar1 *mariner* transposon. *Proceedings of the National Academy of Sciences of the United States of America*, 96(20):11428–11433, 1999.
- David J Lampe, Theresa E Grant, and Hugh M Robertson. Factors affecting transposition of the *Himar1 mariner* transposon *in vitro*. *Genetics*, 149(1):179–187, 1998.
- Leslie Lamport. Latex, a document preparation system, 1986a.
- Leslie Lamport. Latex: User’s guide & reference manual. *Online*, 1986b.
- HC Lange and JJ Heijnen. Statistical reconciliation of the elemental and molecular biomass composition of *Saccharomyces cerevisiae*. *Biotechnology and bioengineering*, 75(3):334–344, 2001.
- Gemma C Langridge, Minh-Duy Phan, Daniel J Turner, Timothy T Perkins, Leopold Parts, Jana Haase, Ian Charles, Duncan J Maskell, Sarah E Peters, Gordon Dougan, et al. Simultaneous assay of every *Salmonella typhi* gene using one million transposon mutants. *Genome research*, 19(12):2308–2316, 2009.
- Joshua Lederberg. The transformation of genetics by DNA: an anniversary celebration of Avery, MacLeod and McCarty (1944). *Genetics*, 136(2):423, 1994.
- Dörte Lehmann, Daniel Hönicke, Armin Ehrenreich, Michael Schmidt, Dirk Weuster-Botz, Hubert Bahl, and Tina Lütke-Eversloh. Modifying the product pattern of *Clostridium acetobutylicum*. *Applied microbiology and biotechnology*, 94(3):743–754, 2012.
- Yvan Lehmann, Leo Meile, and Michael Teuber. Rubrerythrin from *Clostridium perfringens*: cloning of the gene, purification of the protein, and characterization of its superoxide dismutase function. *Journal of bacteriology*, 178(24):7152–7158, 1996.
- Rasko Leinonen, Hideaki Sugawara, and Martin Shumway. The sequence read archive. *Nucleic acids research*, page gkq1019, 2010.
- FM. Liew, M. Kopke, and S. D. Simpson. *Gas fermentation for commercial bio-fuels production*. INTECH Open Access Publisher, 2013.



- W. J. Lin and E. A. Johnson. Transposon Tn916 mutagenesis in *Clostridium botulinum*. *Applied and Environmental Microbiology*, 57(10):2946–2950, 1991.
- Yun-Long Lin and Hans P Blaschek. Transformation of heat-treated *Clostridium acetobutylicum* protoplasts with pUB110 plasmid DNA. *Applied and environmental microbiology*, 48(4):737–742, 1984.
- Janet C Lindow, Robert A Britton, and Alan D Grossman. Structural maintenance of chromosomes protein of *Bacillus subtilis* affects supercoiling *in vivo*. *Journal of bacteriology*, 184(19):5317–5322, 2002.
- H. Liu, L. Bouillaut, A. L. Sonenshein, and S. B. Melville. Use of a *mariner*-based transposon mutagenesis system to isolate *Clostridium perfringens* mutants deficient in gliding motility. *J Bacteriol*, 195(3):629–36, 2013.
- RW Lovitt, JG Morris, and DB Kell. The growth and nutrition of *Clostridium sporogenes* NCIB 8053 in defined media. *Journal of applied bacteriology*, 62(1):71–80, 1987.
- Frauke Luers, Markus Seyfried, Rolf Daniel, and Gerhard Gottschalk. Glycerol conversion to 1,3-propanediol by *Clostridium pasteurianum*: cloning and expression of the gene encoding 1,3-propanediol dehydrogenase. *FEMS microbiology letters*, 154(2):337–345, 1997.
- T. Lutke-Eversloh and H. Bahl. Metabolic engineering of *Clostridium acetobutylicum*: Recent advances to improve butanol production. *Curr Opin Biotechnol*, 22(5):634–47, 2011.
- Luciana Macis, Rolf Daniel, and Gerhard Gottschalk. Properties and sequence of the coenzyme B12-dependent glycerol dehydratase of *Clostridium pasteurianum*. *FEMS microbiology letters*, 164(1):21–28, 1998.
- IS Maddox, E Steiner, S Hirsch, S Wessner, NA Gutierrez, JR Gapes, and KC Schuster. The cause of ‘acid crash’ and ‘acidogenic fermentations’ during the batch acetone-butanol-ethanol (ABE-) fermentation process. *Journal of molecular microbiology and biotechnology*, 2(1):95–100, 2000.
- Veerle ET Maervoet, Sofie L De Maeseneire, Fatma G Avci, Joeri Beauprez, Wim K Soetaert, and Marjan De Mey. 1,3-propanediol production with *Citrobacter werkmanii* DSM17579: effect of a *dhaD* knock-out. *Microbial cell factories*, 13(1):1, 2014.
- Kira S Makarova, Daniel H Haft, Rodolphe Barrangou, Stan JJ Brouns, Emmanuelle Charpentier, Philippe Horvath, Sylvain Moineau, Francisco JM Mojica, Yuri I Wolf, Alexander F Yakunin, et al. Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, 9(6):467–477, 2011.
- A. Malaviya, Y. S. Jang, and S. Y. Lee. Continuous butanol production with reduced byproducts formation from glycerol by a hyper producing mutant of *Clostridium pasteurianum*. *Appl Microbiol Biotechnol*, 93(4):1485–94, 2012.

- M. F. Mallette, P. Reece, and E. A. Dawes. Culture of *Clostridium pasteurianum* in defined medium and growth as a function of sulfate concentration. *Applied Microbiology*, 28(6):999–1003, 1974.
- Morton Mandel and Akiko Higa. Calcium-dependent bacteriophage DNA infection. *Journal of molecular biology*, 53(1):159–162, 1970.
- Magnus Manske. *GENtle, a free multi-purpose molecular biology tool*. PhD thesis, Universität zu Köln, 2006.
- Aron Marchler-Bauer, Shennan Lu, John B Anderson, Farideh Chitsaz, Myra K Derbyshire, Carol DeWeese-Scott, Jessica H Fong, Lewis Y Geer, Renata C Geer, Noreen R Gonzales, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research*, 39(suppl 1): D225–D229, 2011.
- S. A. Marcott, J. D. Shakun, P. U. Clark, and A. C. Mix. A reconstruction of regional and global temperature for the past 11,300 years. *Science*, 339(6124): 1198–201, 2013.
- Krystle J McLaughlin, Claire M Strain-Damerell, Kefang Xie, Dimitris Brekasis, Alexei S Soares, Mark SB Paget, and Clara L Kielkopf. Structural basis for NADH/NAD<sup>+</sup> redox sensing by a Rex family repressor. *Molecular cell*, 38(4): 563–575, 2010.
- LD Mermelstein and ET Papoutsakis. *In vivo* methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. *Applied and Environmental Microbiology*, 59(4):1077–1081, 1993.
- Lee D Mermelstein, Neil E Welker, George N Bennett, and Eleftherios T Papoutsakis. Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. *Nature Biotechnology*, 10(2):190–195, 1992.
- Joachim Messing and Jeffrey Vieira. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene*, 19(3):269–276, 1982.
- B. Meyssignac, D. Salas y Melia, M. Becker, W. Llovel, and A. Cazenave. Tropical Pacific spatial trend patterns in observed sea level: internal variability and/or anthropogenic signature? *Climate of the Past Discussions*, 8(1):349–389, 2012.
- Thomas Millat, Christine Voigt, Holger Janssen, Clare M Cooksley, Klaus Winzer, Nigel P Minton, Hubert Bahl, Ralf-Jörg Fischer, and Olaf Wolkenhauer. Coenzyme A-transferase-independent butyrate re-assimilation in *Clostridium acetobutylicum* evidence from a mathematical model. *Applied microbiology and biotechnology*, 98(21):9059–9072, 2014.
- R. G. Miller and S. R. Sorrell. The future of oil supply. *Philos Trans A Math Phys Eng Sci*, 372(2006):20130179, 2014.

- N. P. Minton. Clostridia in cancer therapy. *Nature Reviews Microbiology*, 1(3): 237–242, 2003.
- Nigel P Minton, Muhammad Ehsaan, Christopher M Humphreys, Gareth T Little, Jonathan Baker, Anne M Henstra, Fungmin Liew, Michelle L Kelly, Lili Sheng, Katrin Schwarz, et al. A roadmap for gene system development in *Clostridium*. *Anaerobe*, 41:104–112, 2016.
- G. Mohr, W. Hong, J. Zhang, G. Z. Cui, Y. Yang, Q. Cui, Y. J. Liu, and A. M. Lambowitz. A targetron system for gene targeting in thermophiles and its application in *Clostridium thermocellum*. *PLoS One*, 8(7):e69032, 2013.
- J Scott Moncrief, Lisa A Barroso, and Tracy D Wilkins. Positive regulation of *Clostridium difficile* toxins. *Infection and immunity*, 65(3):1105–1108, 1997.
- C. Moon, C. H. Lee, B. I. Sang, and Y. Um. Optimization of medium compositions favoring butanol and 1,3-propanediol production from glycerol by *Clostridium pasteurianum*. *Bioresour Technol*, 102(22):10561–8, 2011.
- Koichi Mori, Takamasa Tobimatsu, Tetsuya Hara, and Tetsuo Toraya. Characterization, sequencing, and expression of the genes encoding a reactivating factor for glycerol-inactivated adenosylcobalamin-dependent diol dehydratase. *Journal of Biological Chemistry*, 272(51):32034–32041, 1997.
- Shigeki Moriya, Eitoku Tsujikawa, Anwarul KM Hassan, Kei Asai, Takeko Kodama, and Naotake Ogasawara. A *Bacillus subtilis* gene-encoding protein homologous to eukaryotic SMC motor protein is necessary for chromosome partition. *Molecular microbiology*, 29(1):179–187, 1998.
- Richard Münch, Karsten Hiller, Heiko Barg, Dana Heldt, Simone Linz, Edgar Wingender, and Dieter Jahn. PRODORIC: prokaryotic database of gene regulation. *Nucleic acids research*, 31(1):266–269, 2003.
- Richard Münch, Karsten Hiller, Andreas Grote, Maurice Scheer, Johannes Klein, Max Schobert, and Dieter Jahn. Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics*, 21(22):4187–4189, 2005.
- M. Munoz-Lopez and J. L. Garcia-Perez. DNA transposons: nature and applications in genomics. *Current Genomics*, 11(2):115–128, 2010.
- N. E. Murray. Type I restriction systems: Sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiology and Molecular Biology Reviews*, 64(2):412, 2000.
- Shilpa Nagaraju, Naomi Kathleen Davies, David Jeffrey Fraser Walker, Michael Köpke, and Sean Dennis Simpson. Genome editing of *Clostridium autoethanogenum* using CRISPR/Cas9. *Biotechnology for biofuels*, 9(1):219, 2016.
- J. P. Nakas, M. Schaedle, C. M. Parkinson, C. E. Coonley, and S. W. Tanenbaum. System development for linked-fermentation production of solvents from algal biomass. *Appl Environ Microbiol*, 46(5):1017–23, 1983.

- H. Nariya, S. Miyata, M. Suzuki, E. Tamai, and A. Okabe. Development and application of a method for counterselectable in-frame deletion in *Clostridium perfringens*. *Appl Environ Microbiol*, 77(4):1375–82, 2011.
- Yen Kuan Ng, Muhammad Ehsaan, Sheryl Philip, Mark M Collery, Clare Janoir, Anne Collignon, Stephen T Cartman, and Nigel P Minton. Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using *pyrE* alleles. *PLoS One*, 8(2):e56051, 2013.
- Yvain Nicolet, Claudine Piras, Pierre Legrand, Claude E Hatchikian, and Juan C Fontecilla-Camps. *Desulfovibrio desulfuricans* iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure*, 7(1):13–23, 1999.
- J. Nolling, G. Breton, M. V. Omelchenko, K. S. Makarova, Q. Zeng, R. Gibson, H. M. Lee, J. Dubois, D. Qiu, J. Hitti, Y. I. Wolf, R. L. Tatusov, F. Sabathe, L. Doucette-Stamm, P. Soucaille, M. J. Daly, G. N. Bennett, E. V. Koonin, and D. R. Smith. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol*, 183(16):4823–38, 2001.
- R. W. O'Brien and J. G. Morris. Oxygen and growth and metabolism of *Clostridium acetobutylicum*. *Journal of General Microbiology*, 68(Nov):307–&, 1971.
- Konstantin Okonechnikov, Olga Golosova, Mikhail Fursov, et al. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28(8):1166–1167, 2012.
- J. D. Oultram and M. Young. Conjugal transfer of plasmid Pam-Beta-1 from *Streptococcus lactis* and *Bacillus subtilis* to *Clostridium acetobutylicum*. *Fems Microbiology Letters*, 27(1):129–134, 1985.
- J. D. Oultram, A. Davies, and M. Young. Conjugal transfer of a small plasmid from *Bacillus subtilis* to *Clostridium acetobutylicum* by cointegrate formation with plasmid Pam-Beta-1. *Fems Microbiology Letters*, 42(2-3):113–119, 1987.
- J. D. Oultram, M. Loughlin, T. J. Swinfield, J. K. Brehm, D. E. Thompson, and N. P. Minton. Introduction of plasmids into whole cells of *Clostridium acetobutylicum* by electroporation. *Fems Microbiology Letters*, 56(1):83–88, 1988.
- Martin Pagels, Stephan Fuchs, Jan Pané-Farré, Christian Kohler, Leonhard Menschner, Michael Hecker, Peter J McNamarra, Mikael C Bauer, Claes Von Wachenfeldt, Manuel Liebeke, et al. Redox sensing by a Rex-family repressor is involved in the regulation of anaerobic gene expression in *Staphylococcus aureus*. *Molecular microbiology*, 76(5):1142–1161, 2010.
- M. Pagliaro and M. Rossi. *The future of glycerol*. Royal Society of Chemistry, 2010.
- Arti S Pandey, Travis V Harris, Logan J Giles, John W Peters, and Robert K Szilagyi. Dithiomethylether as a ligand in the hydrogenase H-cluster. *Journal of the American Chemical Society*, 130(13):4533–4540, 2008.

- Eleftherios T Papoutsakis. Engineering solventogenic clostridia. *Current opinion in biotechnology*, 19(5):420–429, 2008.
- Eleftherios Terry Papoutsakis and Charles L Meyer. Fermentation equations for propionic-acid bacteria and production of assorted oxychemicals from various sugars. *Biotechnology and bioengineering*, 27(1):67–80, 1985.
- M. W. Peck. *Clostridium botulinum*. *Pathogens and Toxins in Foods: Challenges and Interventions*, pages 31–52, 2010.
- S Peguin, G Goma, P Delorme, and P Soucaille. Metabolic flexibility of *Clostridium acetobutylicum* in response to methyl viologen addition. *Applied microbiology and biotechnology*, 42(4):611–616, 1994.
- G Percheron, F Brechignac, Ph Soucaille, and J-St Condoret. Carbon dioxide desorption from fermentation broth by use of oxygen vectors. *Bioprocess Engineering*, 12(1-2):11–16, 1995.
- K. Pfluger, S. Baumann, G. Gottschalk, W. Lin, H. Santos, and V. Muller. Lysine-2,3-aminomutase and beta-lysine acetyltransferase genes of methanogenic archaea are salt induced and are essential for the biosynthesis of N-epsilon-acetyl-beta-lysine and growth at high salinity. *Applied and Environmental Microbiology*, 69(10):6047–6055, 2003.
- A. Poehlein, K. Hartwich, P. Krabben, A. Ehrenreich, W. Liebl, P. Durre, G. Gottschalk, and R. Daniel. Complete genome sequence of the solvent producer *Clostridium saccharobutylicum* NCP262 (DSM 13864). *Genome Announc*, 1(6), 2013.
- Anja Poehlein, Alexander Grosse-Honebrink, Ying Zhang, Nigel P Minton, and Rolf Daniel. Complete genome sequence of the nitrogen-fixing and solvent-producing *Clostridium pasteurianum* DSM 525. *Genome announcements*, 3(1):e01591–14, 2015.
- S. M. Poutanen. *Clostridium difficile*-associated diarrhea in adults. *Canadian Medical Association Journal*, 171(1):51–58, 2004.
- I. R. Poxton, J. McCoubrey, and G. Blair. The pathogenicity of *Clostridium difficile*. *Clinical Microbiology and Infection*, 7(8):421–427, 2001.
- Leslie A Pray. Transposons: The jumping genes. *Nature Education*, 1(1):204, 2008.
- M. E. Pyne, M. Moo-Young, D. A. Chung, and C. P. Chou. Development of an electrotransformation protocol for genetic manipulation of *Clostridium pasteurianum*. *Biotechnol Biofuels*, 6(1):50, 2013.
- Michael E Pyne, Murray Moo-Young, Duane A Chung, and C Perry Chou. Expansion of the genetic toolkit for metabolic engineering of *Clostridium pasteurianum*: chromosomal gene disruption of the endogenous CpaAI restriction enzyme. *Biotechnology for biofuels*, 7(1):1, 2014a.

- Michael E Pyne, Sagar Utturkar, Steven D Brown, Murray Moo-Young, Duane A Chung, and C Perry Chou. Improved draft genome sequence of *Clostridium pasteurianum* strain ATCC 6013 (DSM 525) using a hybrid next-generation sequencing approach. *Genome announcements*, 2(4):e00790–14, 2014b.
- Michael E Pyne, Murray Moo-Young, Duane A Chung, and C Perry Chou. Antisense-RNA-mediated gene downregulation in *Clostridium pasteurianum*. *Fermentation*, 1(1):113–126, 2015.
- Michael E Pyne, Mark R Bruder, Murray Moo-Young, Duane A Chung, and C Perry Chou. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. *Scientific reports*, 6, 2016a.
- Michael E Pyne, Xuejia Liu, Murray Moo-Young, Duane A Chung, and C Perry Chou. Genome-directed analysis of prophage excision, host defence systems, and central fermentative metabolism in *Clostridium pasteurianum*. *Scientific Reports*, 6, 2016b.
- Michael E Pyne, Stanislav Sokolenko, Xuejia Liu, Kajan Srirangan, Mark R Bruder, Marc G Aucoin, Murray Moo-Young, Duane A Chung, and C Perry Chou. Disruption of the reductive 1,3-propanediol pathway triggers production of 1, 2-propanediol for sustained glycerol fermentation by *Clostridium pasteurianum*. *Applied and Environmental Microbiology*, 82(17):5375–5388, 2016c.
- D. N. Rao, D. T. F. Dryden, and S. Bheemanaik. Type III restriction-modification enzymes: a historical perspective. *Nucleic Acids Research*, 42(1):45–55, 2014.
- Adriana Ravagnani, Katrin CB Jennert, Elisabeth Steiner, Raik Gruenberg, James R Jefferies, Shane R Wilkinson, Danielle I Young, Edward C Tidswell, David P Brown, Philip Youngman, et al. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. *Molecular microbiology*, 37(5):1172–1185, 2000.
- Dmitry A Ravcheev, Xiaoqing Li, Haythem Latif, Karsten Zengler, Semen A Leyn, Yuri D Korostelev, Alexey E Kazakov, Pavel S Novichkov, Andrei L Osterman, and Dmitry A Rodionov. Transcriptional regulation of central carbon and energy metabolism in bacteria by redox-responsive repressor Rex. *Journal of bacteriology*, 194(5):1145–1157, 2012.
- Céline Raynaud, Patricia Sarçabal, Isabelle Meynial-Salles, Christian Croux, and Philippe Soucaille. Molecular characterization of the 1,3-propanediol (1,3-PD) operon of *Clostridium butyricum*. *Proceedings of the National Academy of Sciences*, 100(9):5010–5015, 2003.
- Jonathan Reizer, Aiala Reizer, and Milton H Saier. A new subfamily of bacterial ABC-type transport systems catalyzing export of drugs and carbohydrates. *Protein Science*, 1(10):1326–1332, 1992.

- RFA Renewable Fuel Agency. Report regarding the supply of biofuels under the Renewable Transport Fuel Obligation within the period 15 April 2008 to 14 April 2009. *Online at Department for Transport*, 2009.
- RFA Renewable Fuel Agency. Report regarding the supply of biofuels under the Renewable Transport Fuel Obligation from the period 15 April 2009 to 14 April 2010. *Online at Department for Transport*, 2010.
- RFA Renewable Fuel Agency. Statistics on road fuel supplied in the UK that is made up of renewable fuel and well as information on its sustainability. *Online at Department for Transport*, 2011.
- RFA Renewable Fuel Agency. Final report for Year 5, 15 April 2012 to 14 April 2013. *Online at Department for Transport*, 2013a.
- RFA Renewable Fuel Agency. Statistics on biofuels for the obligation year 15 April 2011 to 14 April 2012 (Year 4). *Online at Department for Transport*, 2013b.
- RFA Renewable Fuel Agency. Renewable Transport Fuel Obligation statistics: obligation period 6, 2013/14, report 6. *Online at Department for Transport*, 2015.
- RFA Renewable Fuel Agency. Renewable Transport Fuel Obligation statistics: period 7, 2014/15, report 6. *Online at Department for Transport*, 2016a.
- RFA Renewable Fuel Agency. Renewable Transport Fuel Obligation statistics: period 8 2015/16, report 4. *Online at Department for Transport*, 2016b.
- G. Reysset, J. Hubert, L. Podvin, and M. Sebald. Transfection and transformation of *Clostridium acetobutylicum* strain N1-4081 protoplasts. *Biotechnology Techniques*, 2(3):199–204, 1988.
- D. F. Richards, P. E. Linnett, J. D. Oultram, and M. Young. Restriction endonucleases in *Clostridium pasteurianum* ATCC 6013 and *C. thermohydrosulfuricum* DSM 568. *Journal of General Microbiology*, 134:3151–3157, 1988.
- Oliver Riebe, Ralf-Jörg Fischer, David A Wampler, Donald M Kurtz Jr, and Hubert Bahl. Pathway for H<sub>2</sub>O<sub>2</sub> and O<sub>2</sub> detoxification in *Clostridium acetobutylicum*. *Microbiology*, 155(1):16–24, 2009.
- J. I. Rood and S. T. Cole. Molecular-genetics and pathogenesis of *Clostridium perfringens*. *Microbiological Reviews*, 55(4):621–648, 1991.
- Christelle Rosazza, Sasa Haberl Meglic, Andreas Zumbusch, Marie-Pierre Rols, and Damijan Miklavcic. Gene electrotransfer: A mechanistic perspective. *Current gene therapy*, 16(2):98–129, 2016.
- Carlo Rotta. *Establishment of genetic tools in Clostridium pasteurianum and investigation of the germination mechanisms*. PhD thesis, University of Nottingham, 2014.

- Carlo Rotta, Anja Poehlein, Katrin Schwarz, Peter McClure, Rolf Daniel, and Nigel P Minton. Closed genome sequence of *Clostridium pasteurianum* ATCC 6013. *Genome announcements*, 3(1):e01596–14, 2015.
- FE Ruch, J Lengeler, and ECC Lin. Regulation of glycerol catabolism in *Klebsiella aerogenes*. *Journal of bacteriology*, 119(1):50–56, 1974.
- Kim Rutherford, Julian Parkhill, James Crook, Terry Horsnell, Peter Rice, Marie-Adèle Rajandream, and Bart Barrell. Artemis: sequence visualization and annotation. *Bioinformatics*, 16(10):944–945, 2000.
- W. Sabra, C. Groeger, P. N. Sharma, and A. P. Zeng. Improved *n*-butanol production by a non-acetone producing *Clostridium pasteurianum* DSMZ 525 in mixed substrate fermentation. *Appl Microbiol Biotechnol*, 98(9):4267–76, 2014.
- Abigail A Salyers, Nadja B Shoemaker, Ann M Stevens, and Lhing-Yew Li. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. *Microbiological reviews*, 59(4):579–590, 1995.
- Nicholas R Sandoval, Keerthi P Venkataramanan, Theodore S Groth, and Eleftherios T Papoutsakis. Whole-genome sequence of an evolved *Clostridium pasteurianum* strain reveals Spo0A deficiency responsible for increased butanol production and superior growth. *Biotechnology for biofuels*, 8(1):1, 2015.
- P. SanMiguel, A. Tikhonov, Y. K. Jin, N. Motchoulskaia, D. Zakharov, A. Melake-Berhan, P. S. Springer, K. J. Edwards, M. Lee, Z. Avramova, and J. L. Bennetzen. Nested retrotransposons in the intergenic regions of the maize genome. *Science*, 274(5288):765–768, 1996.
- Joseph D Santangelo, Peter Dürre, and David R Woods. Characterization and expression of the hydrogenase-encoding gene from *Clostridium acetobutylicum* P262. *Microbiology*, 141(1):171–180, 1995.
- Tahereh Sarchami, Erin Johnson, and Lars Rehmann. Optimization of fermentation condition favoring butanol production from glycerol by *Clostridium pasteurianum* DSM 525. *Bioresource technology*, 208:73–80, 2016.
- R. K. Saxena, P. Anand, S. Saran, and J. Isar. Microbial production of 1,3-propanediol: Recent developments and emerging opportunities. *Biotechnol Adv*, 27(6):895–913, 2009.
- K. Schwarz, Y. Zhang, W. Kuit, M Ehsaan, K. Kovács, K. Winzer, and NP Minton. New tools for the genetic modification of industrial clostridia. In ME Himmel, editor, *Direct Microbial Conversion of Biomass to Advanced Biofuels*, chapter 13, pages 241–289. Elsevier, 2015. doi: 10.1016/B978-0-444-59592-8.00013-0.
- K. Schwarz, A. Grosse-Honebrink, K. Derecka, C. Rotta, Y. Zhang, and NP Minton. Towards improved butanol production through targeted genetic modification of *Clostridium pasteurianum*. *Metabolic Engineering*, 2017. doi: 10.1016/j.ymben.2017.01.009.



- H. Seedorf, W. F. Fricke, B. Veith, H. Bruggemann, H. Liesegang, A. Strittmatter, M. Miethke, W. Buckel, J. Hinderberger, F. Li, C. Hagemeyer, R. K. Thauer, and G. Gottschalk. The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features. *Proc Natl Acad Sci U S A*, 105(6):2128–33, 2008.
- Corinna Seifert, Susanne Bowien, Gerhard Gottschalk, and Rolf Daniel. Identification and expression of the genes and purification and characterization of the gene products involved in reactivation of coenzyme B12-dependent glycerol dehydratase of *Citrobacter freundii*. *European Journal of Biochemistry*, 268(8):2369–2378, 2001.
- Markus Seyfried, Rolf Daniel, and Gerhard Gottschalk. Cloning, sequencing, and overexpression of the genes encoding coenzyme B12-dependent glycerol dehydratase of *Citrobacter freundii*. *Journal of bacteriology*, 178(19):5793–5796, 1996.
- L. Shao, S. Hu, Y. Yang, Y. Gu, J. Chen, Y. Yang, W. Jiang, and S. Yang. Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. *Cell Res*, 17(11):963–5, 2007.
- WV Shaw. Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria. *Methods in enzymology*, 43:737–755, 1975.
- E Allen Sickmier, Dimitris Brekasis, Shanthi Paranawithana, Jeffrey B Bonanno, Mark SB Paget, Stephen K Burley, and Clara L Kielkopf. X-ray structure of a Rex-family repressor/NADH complex insights into the mechanism of redox sensing. *Structure*, 13(1):43–54, 2005.
- Fabian Sievers, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, Johannes Söding, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, 7(1):539, 2011.
- Robert M Silverstein. The determination of the molar extinction coefficient of reduced DTNB. *Analytical biochemistry*, 63(1):281–282, 1975.
- W. Somrutai, M. Takagi, and T. Yoshida. Acetone-butanol fermentation by *Clostridium aurantibutyricum* ATCC 17777 from a model medium for palm oil mill effluent. *Journal of Fermentation and Bioengineering*, 81(6):543–547, 1996.
- Amy L Spoering and Michael S Gilmore. Quorum sensing and DNA release in bacterial biofilms. *Current opinion in microbiology*, 9(2):133–137, 2006.
- E. Stackebrandt, I. Kramer, J. Swiderski, and H. Hippe. Phylogenetic basis for a taxonomic dissection of the genus *Clostridium*. *Fems Immunology and Medical Microbiology*, 24(3):253–258, 1999.
- Elisabeth Steiner, Angel E Dago, Danielle I Young, John T Heap, Nigel P Minton, James A Hoch, and Michael Young. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. *Molecular microbiology*, 80(3):641–654, 2011.

- Hilde Steinmoen, Eivind Knutsen, and Leiv Sigve Håvarstein. Induction of natural competence in *Streptococcus pneumoniae* triggers lysis and DNA release from a subfraction of the cell population. *Proceedings of the National Academy of Sciences*, 99(11):7681–7686, 2002.
- T.F. Stocker, G.-K. D. Qin, M. T. Plattner, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley. IPCC, 2013: Summary for policymakers. *Climate Change 2013: The Physical Science Basis*, 2013, 2013.
- Patrick Stragier and Richard Losick. Molecular genetics of sporulation in *Bacillus subtilis*. *Annual review of genetics*, 30(1):297–341, 1996.
- Katherine A. Taconi, Keerthi P. Venkataramanan, and Duane T. Johnson. Growth and solvent production by *Clostridium pasteurianum* ATCC® 6013™ utilizing biodiesel-derived crude glycerol as the sole carbon source. *Environmental Progress & Sustainable Energy*, 28(1):100–110, 2009.
- R. S. Tanner, L. M. Miller, and D. Yang. *Clostridium ljungdahlii* sp-nov, an acetogenic species in clostridial ribosomal-RNA homology group I. *International Journal of Systematic Bacteriology*, 43(2):232–236, 1993.
- Y. Tashiro, K. Takeda, G. Kobayashi, K. Sonomoto, A. Ishizaki, and S. Yoshino. High butanol production by *Clostridium saccharoperbutylacetonicum* N1-4 in fed-batch culture with pH-stat continuous butyric acid and glucose feeding method. *J Biosci Bioeng*, 98(4):263–8, 2004.
- J. Theys, O. Pennington, L. Dubois, G. Anlezark, T. Vaughan, A. Mengesha, W. Landuyt, J. Anne, P. J. Burke, P. Durre, B. G. Wouters, N. P. Minton, and P. Lambin. Repeated cycles of *Clostridium*-directed enzyme prodrug therapy result in sustained antitumour effects *in vivo*. *Br J Cancer*, 95(9):1212–9, 2006.
- Christopher M Thomas and Kaare M Nielsen. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature reviews microbiology*, 3(9):711–721, 2005.
- Lynn C Thomason, Nina Costantino, and Donald L Court. *E. coli* genome manipulation by P1 transduction. *Current protocols in molecular biology*, pages 1–17, 2007.
- Tetsuo Toraya. Radical catalysis of B12 enzymes: structure, mechanism, inactivation, and reactivation of diol and glycerol dehydratases. *Cellular and Molecular Life Sciences CMLS*, 57(1):106–127, 2000.
- Federico Tovar-Rojo, Monica Chander, Barbara Setlow, and Peter Setlow. The products of the *spoVA* operon are involved in dipicolinic acid uptake into developing spores of *Bacillus subtilis*. *Journal of bacteriology*, 184(2):584–587, 2002.
- Bellerive Tracey. *Gene Pulser XCell™ Electroporation System Manual*. BioRad, 2013.

- S. A. Tripathi, D. G. Olson, D. A. Argyros, B. B. Miller, T. F. Barrett, D. M. Murphy, J. D. McCool, A. K. Warner, V. B. Rajgarhia, L. R. Lynd, D. A. Hogsett, and N. C. Caiazza. Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. *Appl Environ Microbiol*, 76(19):6591–9, 2010.
- Andreas Untergasser, Ioana Cutcutache, Triinu Koressaar, Jian Ye, Brant C Faircloth, Mado Remm, and Steven G Rozen. Primer3 new capabilities and interfaces. *Nucleic acids research*, 40(15):e115–e115, 2012.
- Karen P Usdin, Harold Zappe, David T Jones, and David R Woods. Cloning, expression, and purification of glutamine synthetase from *Clostridium acetobutylicum*. *Applied and environmental microbiology*, 52(3):413–419, 1986.
- I. Vasconcelos, L. Girbal, and P. Soucaille. Regulation of carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol. *Journal of Bacteriology*, 176(5):1443–1450, 1994.
- K. P. Venkataramanan, J. J. Boatman, Y. Kurniawan, K. A. Taconi, G. D. Bothun, and C. Scholz. Impact of impurities in biodiesel-derived crude glycerol on the fermentation by *Clostridium pasteurianum* ATCC 6013. *Applied Microbiology and Biotechnology*, 93(3):1325–1335, 2012.
- Venkata Ramana Vepachedu and Peter Setlow. Role of SpoVA proteins in release of dipicolinic acid during germination of *Bacillus subtilis* spores triggered by dodecylamine or lysozyme. *Journal of bacteriology*, 189(5):1565–1572, 2007.
- M. W. Vetting, L. P. S. de Carvalho, M. Yu, S. S. Hegde, S. Magnet, S. L. Roderick, and J. S. Blanchard. Structure and functions of the GNAT superfamily of acetyltransferases. *Archives of Biochemistry and Biophysics*, 433(1):212–226, 2005.
- PM Vignais and A Colbeau. Molecular biology of microbial hydrogenases. *Current issues in molecular biology*, 6(2):159–188, 2004.
- John E Walker, Matti Saraste, Michael J Runswick, and Nicholas J Gay. Distantly related sequences in the alpha-and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO journal*, 1(8):945, 1982.
- L. Wang, J. Littlewood, and R. J. Murphy. Environmental sustainability of bioethanol production from wheat straw in the UK. *Renewable & Sustainable Energy Reviews*, 28:715–725, 2013.
- Y. Wang, X. Li, Y. Mao, and H. P. Blaschek. Genome-wide dynamic transcriptional profiling in *Clostridium beijerinckii* NCIMB 8052 using single-nucleotide resolution RNA-Seq. *BMC Genomics*, 13:102, 2012.

- Yi Wang, Zhong-Tian Zhang, Seung-Oh Seo, Kijoong Choi, Ting Lu, Yong-Su Jin, and Hans P Blaschek. Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9 system. *Journal of biotechnology*, 200: 1–5, 2015.
- Z. X. Wang, J. Zhuge, H. Fang, and B. A. Prior. Glycerol production by microbial fermentation: A review. *Biotechnol Adv*, 19(3):201–23, 2001.
- Søren Warming, Nina Costantino, Nancy A Jenkins, Neal G Copeland, et al. Simple and highly efficient BAC recombineering using *galK* selection. *Nucleic acids research*, 33(4):e36–e36, 2005.
- James C Weaver and Yu A Chizmadzhev. Theory of electroporation: a review. *Bioelectrochemistry and bioenergetics*, 41(2):135–160, 1996.
- P. J. Weimer and D. M. Stevenson. Isolation, characterization, and quantification of *Clostridium kluyveri* from the bovine rumen. *Applied Microbiology and Biotechnology*, 94(2):461–466, 2012.
- M. Wellington, M. A. Kabir, and E. Rustchenko. 5-Fluoro-orotic acid induces chromosome alterations in genetically manipulated strains of *Candida albicans*. *Mycologia*, 98(3):393–398, 2006.
- A. H. West and A. M. Stock. Histidine kinases and response regulator proteins in two-component signaling systems. *Trends in Biochemical Sciences*, 26(6): 369–376, 2001.
- Mandy Wietzke and Hubert Bahl. The redox-sensing protein Rex, a transcriptional regulator of solventogenesis in *Clostridium acetobutylicum*. *Applied microbiology and biotechnology*, 96(3):749–761, 2012.
- G. G. Wilson and N. E. Murray. Restriction and modification systems. *Annual Review of Genetics*, 25:585–627, 1991.
- Serge Winogradsky. Recherches sur l'assimilation de l'azote libre de l'atmosphère par les microbes. *Unknown Binding*, 1895.
- U. Witt, R. J. Muller, J. Augusta, H. Widdecke, and W. D. Deckwer. Synthesis, properties and biodegradability of polyesters based on 1,3-propanediol. *Macromolecular Chemistry and Physics*, 195(2):793–802, 1994.
- U. Witt, R. J. Muller, and W. D. Deckwer. New biodegradable polyester-copolymers from commodity chemicals with favourable use properties. *Journal of Environmental Polymer Degradation*, 3(4):215–223, 1995.
- Y. Xiao, C. Francke, T. Abee, and M. H. Wells-Bennik. Clostridial spore germination versus bacilli: genome mining and current insights. *Food Microbiol*, 28(2):266–74, 2011.
- Tao Xu, Yongchao Li, Zhou Shi, Christopher L Hemme, Yuan Li, Yonghua Zhu, Joy D Van Nostrand, Zhili He, and Jizhong Zhou. Efficient genome editing in *Clostridium cellulolyticum* via CRISPR-Cas9 nickase. *Applied and environmental microbiology*, 81(13):4423–4431, 2015.

- S. S. Yazdani and R. Gonzalez. Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry. *Curr Opin Biotechnol*, 18(3):213–9, 2007.
- N. Yutin and M. Y. Galperin. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol*, 15(10):2631–41, 2013.
- I.h Zelitch. Simultaneous use of molecular nitrogen and ammonia by *Clostridium pasteurianum*. *Proc Natl Acad Sci U S A*, 37(9):559–65, 1951.
- Lei Zhang, Xiaoqun Nie, Dmitry A Ravcheev, Dmitry A Rodionov, Jia Sheng, Yang Gu, Sheng Yang, Weihong Jiang, and Chen Yang. Redox-responsive repressor Rex modulates alcohol production and oxidative stress tolerance in *Clostridium acetobutylicum*. *Journal of bacteriology*, 196(22):3949–3963, 2014.
- Linsheng Zhang, Lillian Gray, Richard P Novick, and Guangyong Ji. Transmembrane topology of AgrB, the protein involved in the post-translational modification of AgrD in *Staphylococcus aureus*. *Journal of Biological Chemistry*, 277(38):34736–34742, 2002.
- Ying Zhang, Alexander Grosse-Honebrink, and Nigel P Minton. A universal *mariner*-transposon system for forward genetic studies in the genus *Clostridium*. *PLoS One*, 10(4):e0122411, 2015.
- J. Zigorová and E. Šturdík. Advances in biotechnological production of butyric acid. *J. Ind. Microbiol. Biotechnol.*, 24(3):153–160, 2000.



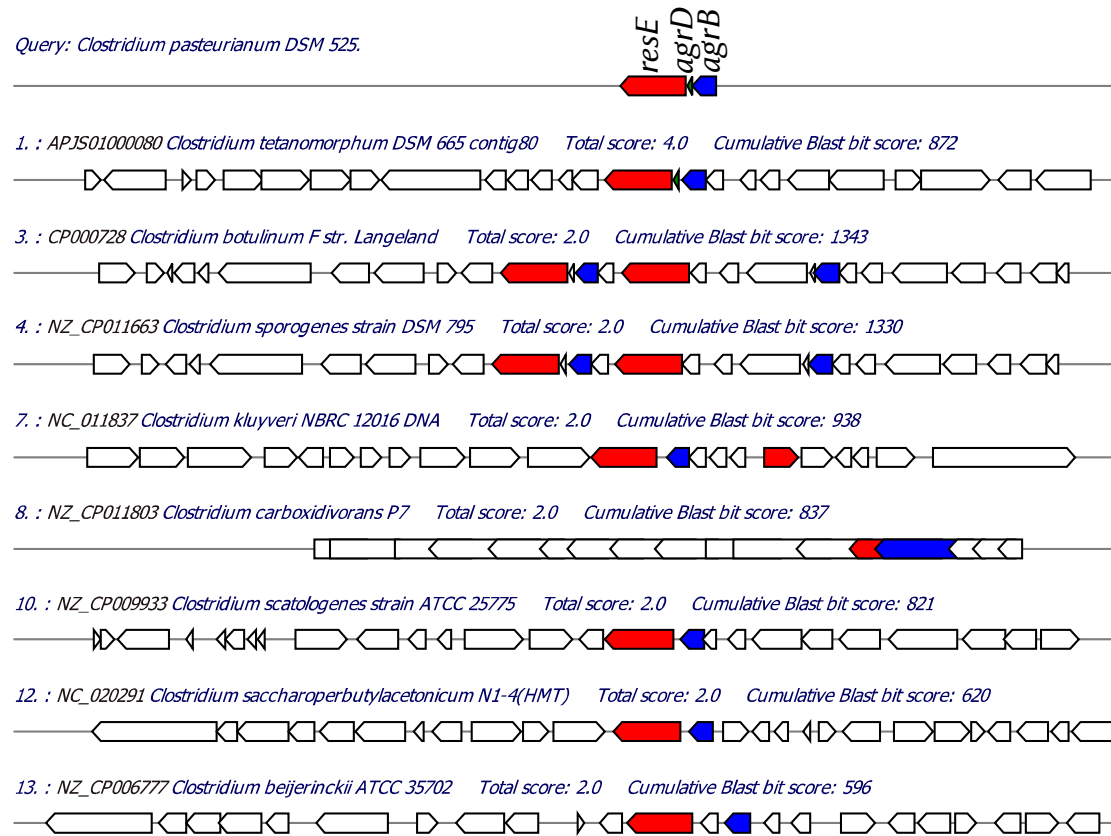
# Appendix

Things are not as they seem. They are what they are.

---

Terry Pratchett, *Thief of Time*

## 6.5 Supplementary Material Chapter 3



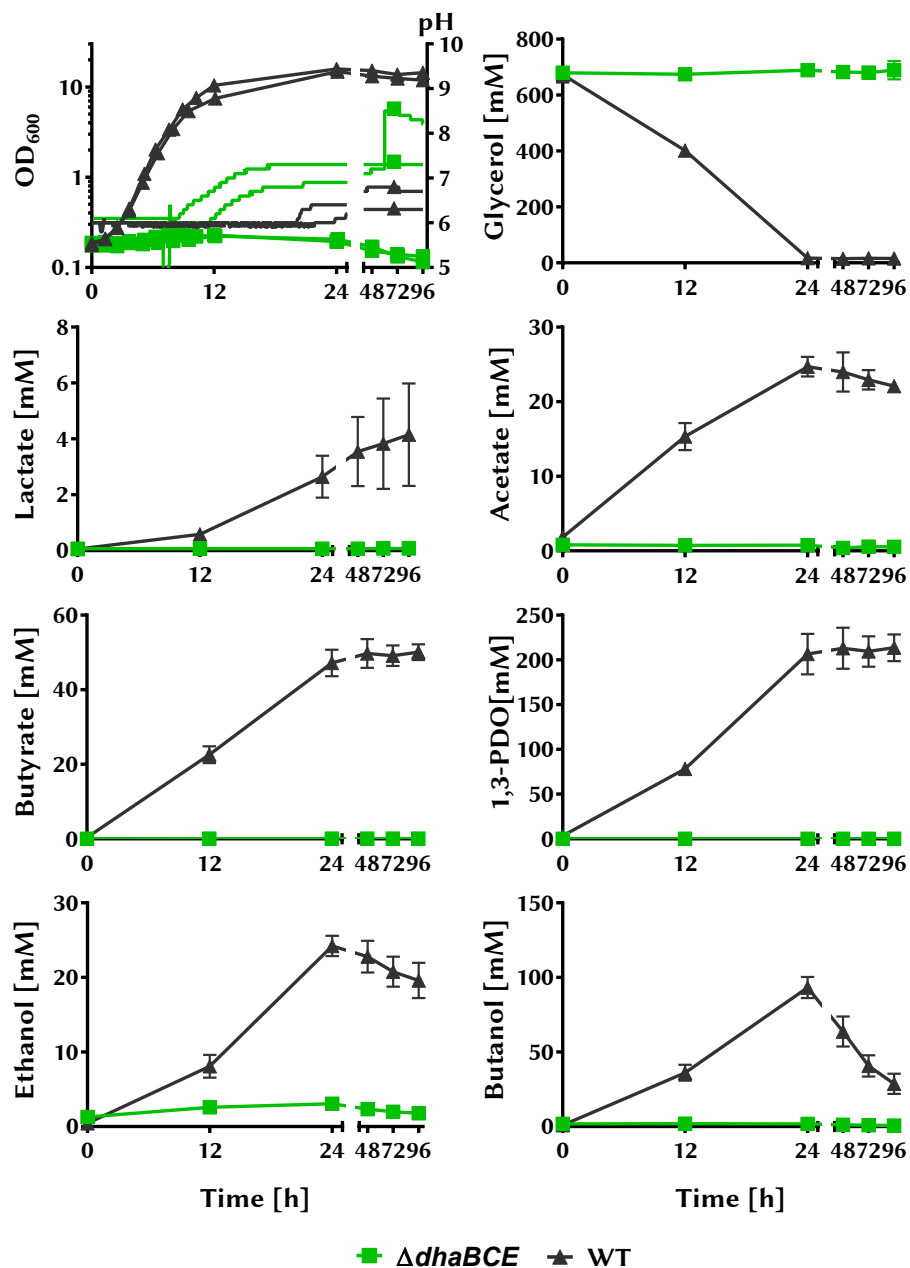
**Figure A.1** – Multiblast analysis of *C. pasteurianum* cluster *agrB-agrD-resE*. The cluster is conserved in a wide range of *Clostridium* species.



**Table A.1** – Electroporation buffers and other agents used for electroporation in different bacterial species.

Electroporation Buffer						
Use/Species	<i>Clostridium acetobutylicum</i>	<i>Campylobacter</i> spp.	<i>Clostridium botulinum</i>	<i>Pseudomonas aeruginosa</i>	<i>Pseudomonas aeruginosa</i>	<i>Clostridium beijerinckii</i>
Cell-wall stabiliser				1 mM MgCl <sub>2</sub>	1 mM MgCl <sub>2</sub>	1 mM MgCl <sub>2</sub>
Crowding agent		15 % Glycerol	10 % PEG6000			
Buffer	5 mM Na-Phosphate	2.43 mM K <sub>2</sub> HPO <sub>4</sub> 0.57 mM KH <sub>2</sub> PO <sub>4</sub>	7 mM Na-Phosphate	1 mM HEPES	1 mM HEPES	7 mM Na <sub>2</sub> HPO <sub>4</sub> 7 mM NaH <sub>2</sub> PO <sub>4</sub>
Osmotic stabiliser	270 mM Sucrose	272 mM Sucrose			300 mM Sucrose	270 mM Sucrose
pH	7.4	7.4	7.4	7.4	7.4	
Cell-wall weakening agents						
	Glycine	D-Trypsin	L-Trypsin	Threonine		
Cryopreservative						
	DMSO	Glycerol				

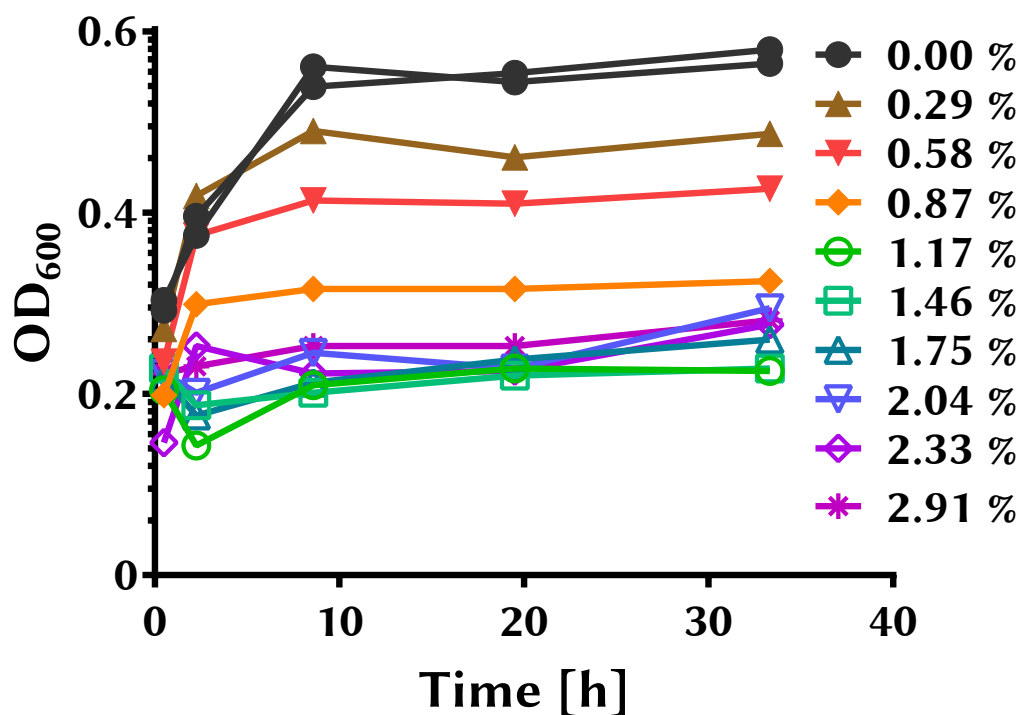
## 6.6 Supplementary Material Chapter 4



**Figure A.2** – Bioreactor batch growth and product profile of *C. pasteurianum*  $\Delta dhaBCE$  and the wild type strain in two individual fermentations. Strains were grown in Biebl medium supplemented with 1 g/l yeast extract and 6 % glycerol. Different orientations stands for individual fermentation. The measurement of  $OD_{600}$  was done manually and pH was measured with an automatic internal probe (lines with single symbols). For clarity, OD and pH is given separately for each replicate. The product formation is given as mean with bars indicating the range of the two samples.

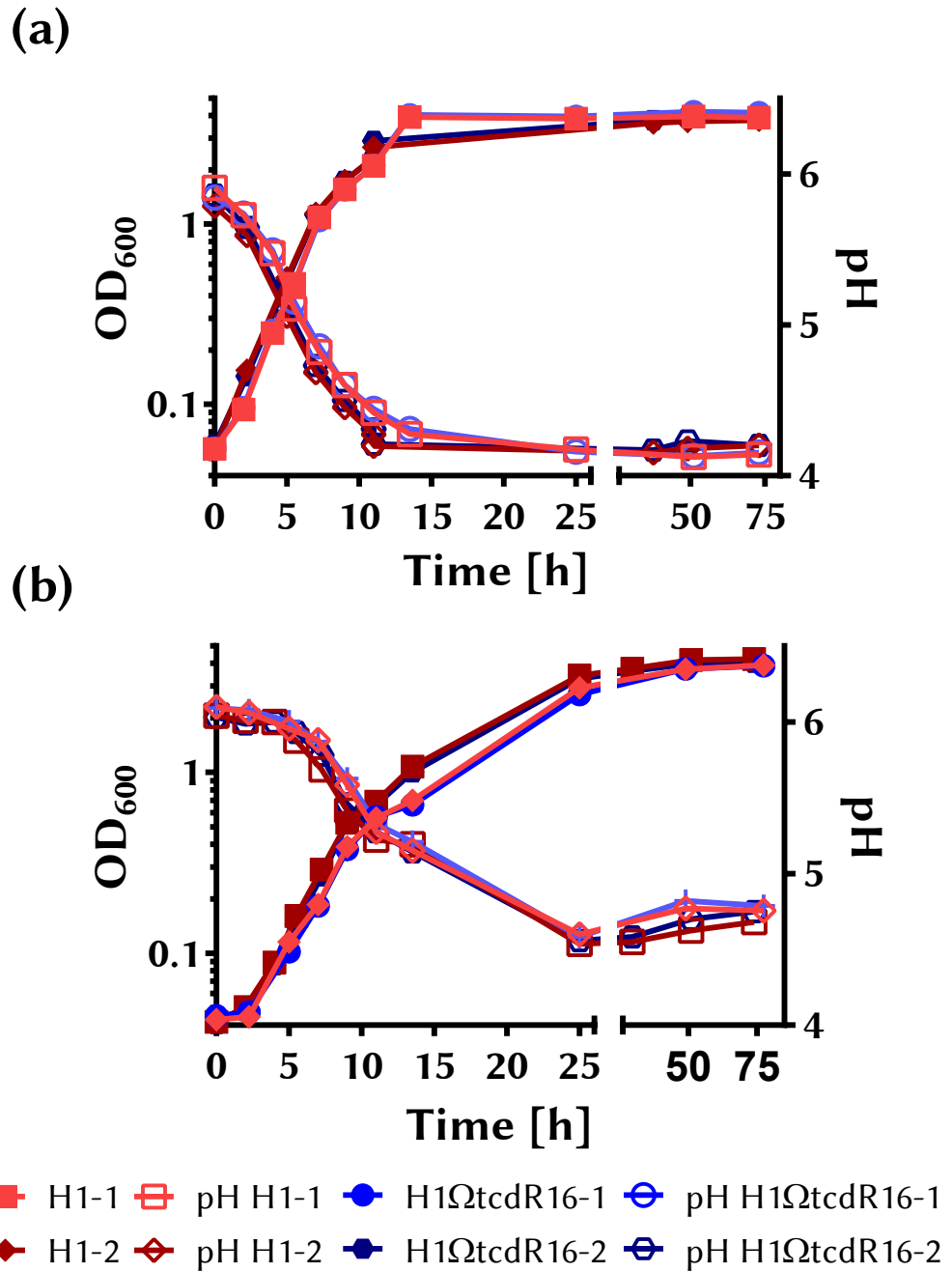
**Table A.2** – Data from fermentations in [Sandoval et al. \(2015\)](#) with random mutant M150B and *spo0A* deletion mutant ( $\Delta spo0A$ ) and their wild type. Data is based approximations from publication figures.

Medium	CGM							
	100 g/l crude glycerol		60 g/l pure glycerol		100 g/l crude glycerol			
Carbon source	M150B	ATCC 6013	M150B	ATCC 6013	M150B	$\Delta spo0A$	ATCC 6013	
Strain	(random mutant)	(WT)	(random mutant)	(WT)	(random mutant)		(WT)	
Growth characteristics								
Max. OD <sub>600</sub> []	4.7	3.9	7.5	5.6	3.9	3.9	3.9	
Product concentrations [mM]								
Carbon source used	456	261	467	250				
Lactate								
Acetate	0.0	0.0	0.0	0.0	26.6	26.6	21.6	
Butyrate	2.3	0.0	1.1	1.0	8.0	6.9	0.0	
1,3-Propanediol	10.5	17.1	10.5	18.4	26.3	26.3	39.4	
Ethanol					21.7	30.4	32.6	
Butanol	94.4	51.3	160.6	85.0	109.3	118.7	101.2	
Selectivity								
BuOH/Solvents					0.69	0.68	0.58	
Yield [M/M]								
BuOH/C-Source	0.21	0.20	0.34	0.34				
PDO/C-Source	0.02	0.07	0.02	0.07				

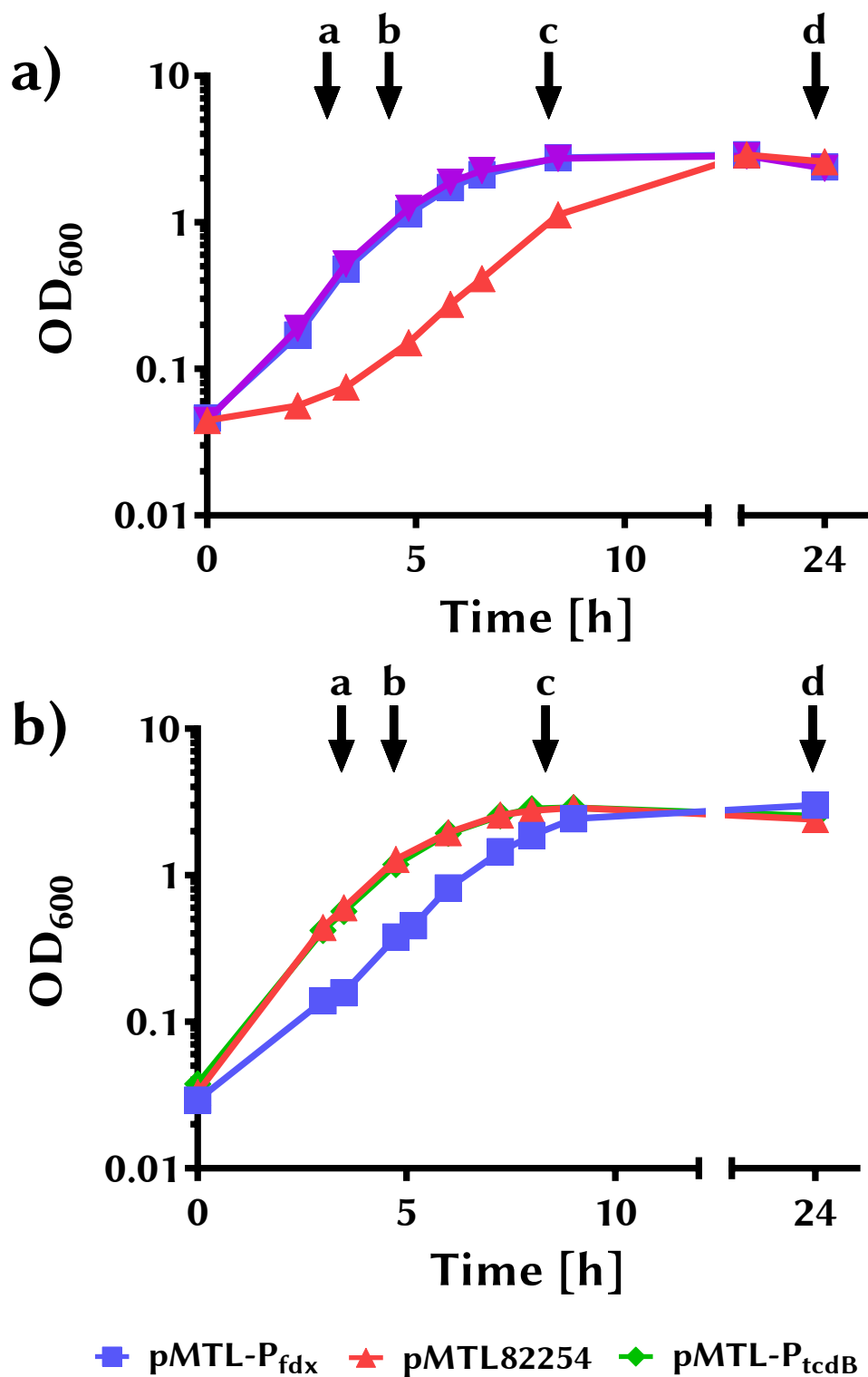


**Figure A.3** – Butanol toxicity on *C. pasteurianum*-H1 was tested by growing over-nights of the strain in 2xYTG and inoculating a culture to OD<sub>600</sub> 0.05. After main culture reached OD<sub>600</sub> 1, 25 ml aliquots were distributed into separate serum bottles and butanol added in the concentrations stated in the legend of the graph. Growth under the different condition was monitored by measuring OD<sub>600</sub> over time. External butanol addition of low concentrations (0.29 %) has an adverse effect on growth. Growth was completely inhibited with 1.17 % butanol addition.

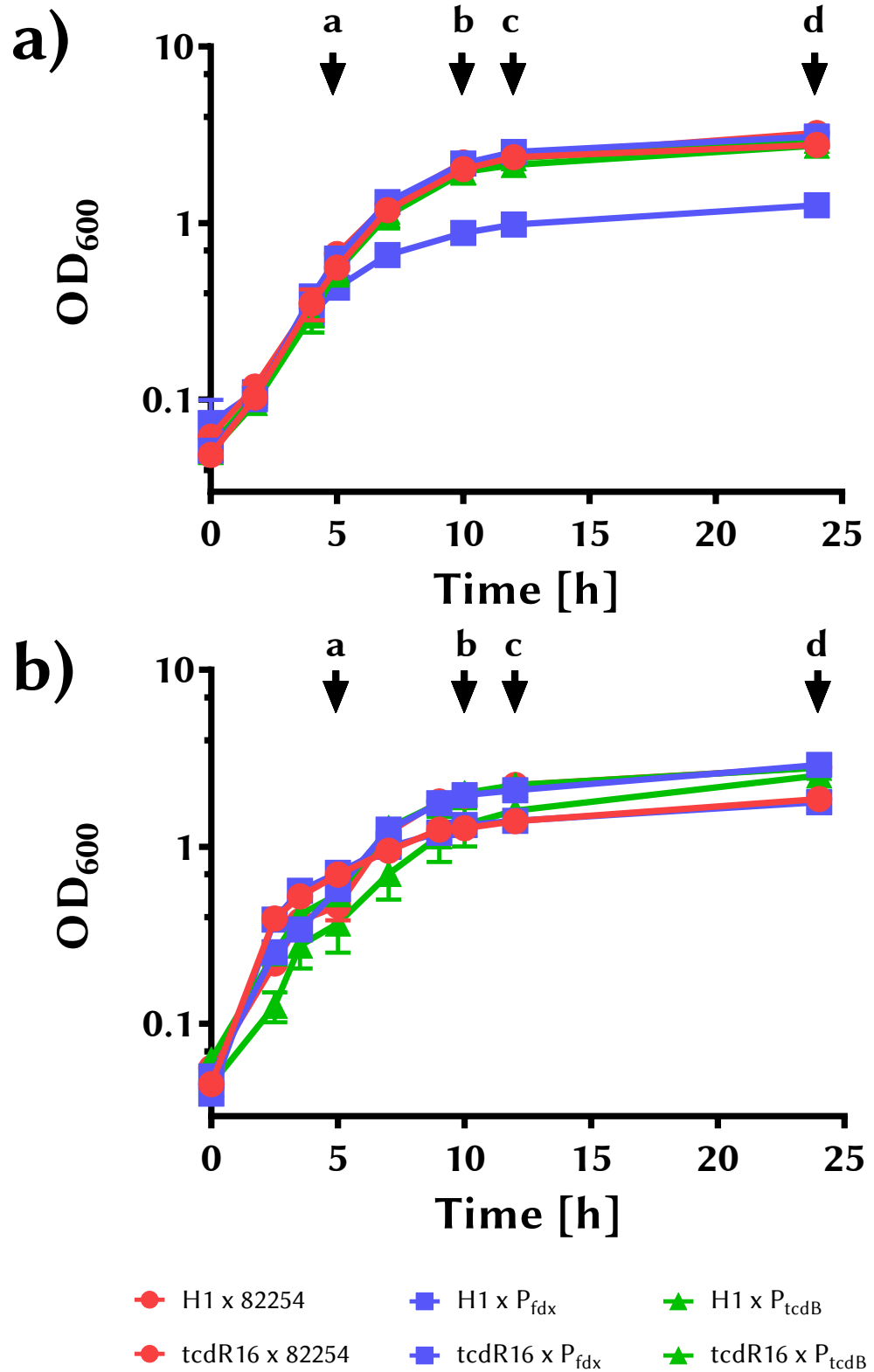
## 6.7 Supplementary Material Chapter 5



**Figure A.4** – Growth curves of *C. pasteurianum*-H1  $\Omega$ tcdR 16 and *C. pasteurianum*-H1 in 2xYT with either (a) 6 % glucose or (b) 6 % glycerol. Each figure shows two individual experiments in biological triplicates. Data is given as mean with error bars indicating the standard error (SEM).



**Figure A.5** – Growth curve for sampling of *CAT* assay of *E. coli*. (a) and (b) are the two separate sampling curves experiments. Arrows indicate sampling points with letters corresponding to *CAT* assay in Figure 5.6. Data is given as mean with error bars indicating the standard error (SEM).



**Figure A.6** – Growth curve and sampling points of CAT assay testing P<sub>tcdB</sub> in *C. pasteurianum*. (a) and (b) are the two separate sampling experiments. Arrows indicate sampling points for CAT assay corresponding to Figure 5.7. Data is given as mean with error bars indicating the standard error (SEM).

**Table A.3** – Transposon insertion sites and disrupted genes in 60 individual clones of *C. acetobutylicum*. One double-insertion is noted in clone 40.

Clone #	locus	strand	ORF	Product
1	1137992	reverse	CA_c2682	hypothetical protein
2	1781494	reverse	CA_c2052	DNA-dependent RNA polymerase sigma subunit
3	1510185	forward	upstream CA_c1365	
4	1222841	forward	CA_c1070	AcrR family transcriptional regulator
5	254856	reverse	upstream CA_c3487	
6	883172	forward	upstream CA_c0761	
7	1034275	forward	CA_c0903	sensory transduction histidine kinase
8	1040298	reverse	CA_c2768	AcrR family transcriptional regulator
9	1483534	reverse	CA_c2346	glycosyltransferase
10	57588	reverse	CA_P0124	MDR-like permease
11	775288	forward	CA_c0666	sugar permease
12	1988196	forward	CA_c1836	DNA mismatch repair protein
13	2180535	reverse	upstream CA_c1620	
14	1321612	forward	CA_c1159	hypothetical protein
15	1807818	forward	CA_c1664	glycogen phosphorylase
16	521363	forward	upstream CA_c0453	ABC-transporter, ATP-binding protein
17	178781	reverse	CA_c3563	mismatch repair protein MutS-like ATPase
18	3083171	forward	CA_c2948	ABC-transporter, ATPase
19	1276523	forward	CA_c1110	XerC/XerD family integrase/recombinase
20	595845	forward	CA_c0516	DNA polymerase III DnaE
21	3562958	forward	CA_c3385	hypothetical protein
22	302057	forward	CA_c0269	hypothetical protein
23	2537054	forward	CA_c2420	methyl-accepting chemotaxis protein
24	2673274	reverse	CA_c1098	DNA polymerase I
25	319179	reverse	CA_c3433	AcrR family transcriptional regulator
26	506862	forward	upstream CA_c0439	
27	2007811	reverse	CA_c1784	DNA uptake protein
28	1073934	forward	upstream CA_c0932	
29	286365	reverse	CA_c3457	hypothetical protein
30	2805189	forward	CA_c2685	maltose phosphorylase
31	1839066	forward	CA_c1688	penicillin-binding protein
32	51302	reverse	CA_c3684	polygalacturonase
33	3916870	forward	CA_c3710	hypothetical protein
34	3610141	reverse	CA_c0290	sensory transduction histidine kinase
35	2978973	reverse	CA_c0830	response regulator
36	3641156	forward	downstream CA_c3444	
37	1350327	forward	CA_c1200	putative phospho-adenylylsulfate sulfotransferase
38	3727867	reverse	CA_c0187	glucosamine-6-phosphate deaminase
39	3033634	forward	CA_c2899	LysM repeat-containing protein
40a	549397	forward	CA_c0475	HD-GYP domain-containing protein
40b	768217	forward	downstream CA_c0660	
41	1200426	reverse	CA_c2631	hypothetical protein
42	328924	forward	CA_c0289	response regulator
43	776062	forward	CA_c0667	sugar-binding periplasmic protein
44	3896397	reverse	CA_c0033	ABC1 family protein kinase
45	716024	forward	CA_c0611	hypothetical protein
46	3926115	reverse	CA_c3718	hypothetical protein
47	1279060	reverse	CA_c1113	hypothetical protein
48	1445152	reverse	CA_c1300	RNA polymerase sigma factor RpoD
49	3926953	reverse	CA_c3719	hypothetical protein
50	2807064	reverse	CA_c2685	maltose phosphorylase
51	3733791	forward	CA_c3538	metallo-beta-lactamase superfamily hydrolase
52	3316380	forward	CA_c3179	oligopeptide ABC transporter periplasmic binding component
53	1529557	reverse	CA_c1387	membrane associated chemotaxis sensory transducer
54	621502	forward	CA_c0535	beta-lactamase superfamily hydrolase
55	2149175	forward	CA_c2042	ABC transporter permease
56	3849360	forward	CA_c3648	flanking : acetyltransferase stage V sporulation protein T
57	2808519	reverse	CA_c2686	maltodextrin glucosidase
58	1119096	forward	CA_c0971	aconitate hydratase/ isocitrate dehydratase
59	2360598	reverse	CA_c2261	hypothetical protein
60	3188855	reverse	CA_c3039	hypothetical protein



**Table A.4** – Transposon insertion sites and disrupted genes in 60 individual clones of *C. sporogenes*. Two double-insertions are noted in clone 12 and 28.

Clone #	locus	strand	ORF	Product
1	46524	forward	CLSPO_c17110	glycosyltransferase
2	2968030	reverse	CLSPO_c10740	hypothetical protein
3	2726430	reverse	CLSPO_c13000	putative manganese efflux pump MntP
4	271399	reverse	CLSPO_c35400	FtsW-like protein
5	428968	reverse	CLSPO_c34130	AMP-binding enzyme
6	1812039	reverse	downstream CLSPO_c20910	
7	1401650	forward	upstream CLSPO_c12860	
8	665049	forward	CLSPO_c05950	spermidine/putrescine import ATP-binding protein PotA
9	232073	forward	CLSPO_c02280	hypothetical protein
10	2011412	forward	upstream CLSPO_c18080	
11	989028	forward	CLSPO_c08950	Sel1 repeat-containing protein
12a	1219604	reverse	upstream CLSPO_c26660	
12b	3172508	reverse	downstream CLSPO_c08770	
13	302409	forward	CLSPO_c02920	putative oxidoreductase, FAD-binding protein
14	805669	forward	CLSPO_c07240	NlpC/p60-like transpeptidase
15	3555747	forward	CLSPO_c32410	ABC transporter permease
16	2732303	forward	CLSPO_c24930	resolvase
17	2071115	reverse	downstream CLSPO_c18640	
18	3382554	forward	downstream CLSPO_c30780	
19	1521258	forward	CLSPO_c13910	methyl-accepting chemotaxis protein
20	1238667	forward	CLSPO_c11310	aminotransferase, class V
21	1370920	forward	CLSPO_c12560	putative transcriptional regulator
22	2121083	reverse	downstream CLSPO_c18180	
23	2964347	forward	CLSPO_c27080	putative B12-binding Fe-S oxidoreductase
24	739260	forward	CLSPO_c06630	AraC family regulatory protein
25	1402797	reverse	CLSPO_c25010	decaprenyl-phosphate phosphoribosyltransferase
26	2242474	reverse	CLSPO_c17220	NADPH-dependent butanol dehydrogenase Adh
27	1495838	reverse	CLSPO_c23710	molybdenum cofactor cytidylyltransferase MocA
28a	939090	forward	CLSPO_c08480	Crp/Fnr family transcriptional regulator
28b	4106436	reverse	downstream CLSPO_c00370	
29	630207	reverse	upstream CLSPO_c32090	
30	89354	reverse	CLSPO_c37390	putative phage head-tail adaptor
31	3661856	reverse	CLSPO_c33420	dihydroorotase PyrC
32	775318	reverse	CLSPO_c30640	putative poly(beta-D-mannuronate) O-acetylase AlgI
33	2662351	reverse	CLSPO_c23940	hypothetical protein
34	2189850	forward	CLSPO_c19720	hypothetical protein
35	4118876	forward	CLSPO_c38150	Lon protease
36	835977	reverse	CLSPO_c30040	Non-specific serine/threonine protein kinase
37	3091150	reverse	downstream CLSPO_c09620	
38	504416	reverse	CLSPO_c33180	formate acetyltransferase Pfl
39	726371	reverse	CLSPO_c31130	hypothetical protein
40	1578114	forward	CLSPO_c14440	putative amino-acid ABC transporter permease protein PatM
41	2368860	reverse	CLSPO_c16230	coenzyme A disulfide reductase Cdr
42	3703470	reverse	CLSPO_c04100	thiamine biosynthesis lipoprotein ApbE
43	3394395	reverse	CLSPO_c06710	hypothetical protein
44	837094	forward	CLSPO_c07520	downstream complement
45	2581655	reverse	upstream CLSPO_c14280	
46	4008507	reverse	CLSPO_c01290	hypothetical protein UPF0272
47	3165348	reverse	downstream CLSPO_c08850	
48	1074180	reverse	upstream CLSPO_c28110	
49	1727714	reverse	CLSPO_c21610	hypothetical protein
50	1865628	forward	downstream CLSPO_c17000	
51	2135598	reverse	CLSPO_c18040	ribosomal large subunit pseudouridine synthase B
52	2738423	forward	CLSPO_c25000	4-amino-4-deoxy-L-arabinose transferase
53	1623481	reverse	CLSPO_c22480	uncharacterized HTH-type transcriptional regulator FruR
54	2800218	reverse	downstream CLSPO_c12230	
55	3668655	reverse	downstream CLSPO_c04330	
56	2364977	forward	CLSPO_c21170	S54 family peptidase
57	1834846	forward	CLSPO_c16750	metallopeptidase
58	856002	reverse	CLSPO_c29860	ABC-2 family transporter protein
59	2615047	forward	CLSPO_c23320	flavodoxin oxidoreductase
60	3471091	forward	CLSPO_c31790	stage V sporulation protein AD

**Table A.5** – Summary reports of Illumina read-mapping against reference genome *C. pasteurianum* DSM 525 (Poehlein et al., 2015) of different *pyrE* truncation strains and one *tcdR* insertion strain. Table continues on next page.

	Read count	Reads [%]	Average length	Bases count	Bases [%]
<i>C. pasteurianum</i> -H1 $\Delta pyrE$ 9					
References	1	-	4352101.00	4352101	-
Mapped reads	1093394	98.77	244.72	267574719	99.09
Not mapped reads	13630	1.23	180.40	2458837	0.91
Reads in pairs	1049016	94.76	486.03	258195898	95.62
Broken paired reads	44378	4.01	211.34	9378821	3.47
Total reads	1107024	100.00	243.93	270033556	100.00
<i>C. pasteurianum</i> -H1 $\Omega tcdR$ 16					
References	1	-	4352101.00	4352101	-
Mapped reads	2643775	97.97	249.69	660113708	97.97
Not mapped reads	54659	2.03	249.93	13661008	2.03
Reads in pairs	2561914	94.94	541.37	641101924	95.15
Broken paired reads	81861	3.03	232.24	19011784	2.82
Total reads	2698434	100.00	249.69	673774716	100.00
<i>C. pasteurianum</i> -H3 $\Delta pyrE$ 3					
References	1	-	4352101.00	4352101	-
Mapped reads	2615035	98.06	249.31	651943343	98.06
Not mapped reads	51637	1.94	249.65	12890921	1.94
Reads in pairs	2539150	95.22	534.1	634971653	95.51
Broken paired reads	75885	2.85	223.65	16971690	2.55
Total reads	2666672	100.00	249.31	664834264	100.00
<i>C. pasteurianum</i> -H3 $\Delta pyrE$ 4					
References	1	-	4352101.00	4352101	-
Mapped reads	2404993	98.42	249.2	599334819	98.42
Not mapped reads	38637	1.58	249.64	9645375	1.58
Reads in pairs	2341080	95.80	524.42	585572983	96.16
Broken paired reads	63913	2.62	215.32	13761836	2.26
Total reads	2443630	100.00	249.21	608980194	100.00
<i>C. pasteurianum</i> -H3 $\Delta pyrE$ 5					
References	1	-	4352101.00	4352101	-
Mapped reads	2199639	98.06	249.31	548388783	98.06
Not mapped reads	43511	1.94	249.84	10870728	1.94
Reads in pairs	2131738	95.03	528.43	533479210	95.39
Broken paired reads	67901	3.03	219.58	14909573	2.67
Total reads	2243150	100.00	249.32	559259511	100.00
<i>C. pasteurianum</i> -H4 $\Delta pyrE$ 3					
References	1	-	4352101.00	4352101	-
Mapped reads	2413661	98.17	246.89	595896843	98.15
Not mapped reads	44903	1.83	249.7	11212503	1.85
Reads in pairs	2270760	92.36	531.31	568020296	93.56
Broken paired reads	142901	5.81	195.08	27876547	4.59
Total reads	2458564	100.00	246.94	607109346	100.00

**Continued Table A.5** – Summary reports of Illumina read-mapping against reference genome *C. pasteurianum* DSM 525 (Poehlein et al., 2015) of different *pyrE* truncation strains and one *tcdR* insertion strain.

	Read count	Reads [%]	Average length	Bases count	Bases [%]
<i>C. pasteurianum</i> -H4 $\Delta pyrE$ 5					
References	1	-	4352101.00	4352101	-
Mapped reads	2357261	98.32	249.34	587749414	98.31
Not mapped reads	40363	1.68	249.81	10082932	1.69
Reads in pairs	2309126	96.31	520.13	576510742	96.43
Broken paired reads	48135	2.01	233.48	11238672	1.88
Total reads	2397624	100.00	249.34	597832346	100.00
<i>C. pasteurianum</i> -H4 $\Delta pyrE$ 23					
References	1	-	4352101.00	4352101	-
Mapped reads	2705907	98.53	249.44	674969183	98.53
Not mapped reads	40413	1.47	249.64	10088844	1.47
Reads in pairs	2643136	96.24	533.56	661188076	96.52
Broken paired reads	62771	2.29	219.55	13781107	2.01
Total reads	2746320	100.00	249.45	685058027	100.00

## **6.8 Publications Originating From This Study**



## Towards improved butanol production through targeted genetic modification of *Clostridium pasteurianum*



Katrin M. Schwarz<sup>1</sup>, Alexander Grosse-Honebrink<sup>1</sup>, Kamila Derecka, Carlo Rotta, Ying Zhang, Nigel P. Minton\*

*Clostridia Research Group, BBSRC/EPSC Synthetic Biology Research Centre (SBRC), University of Nottingham, University Park, Nottingham NG7 2RD, UK*

### ARTICLE INFO

#### Keywords:

*Clostridium pasteurianum*  
Rex  
HydA  
DhaBCE  
Butanol  
1,3-propanediol

### ABSTRACT

Declining fossil fuel reserves, coupled with environmental concerns over their continued extraction and exploitation have led to strenuous efforts to identify renewable routes to energy and fuels. One attractive option is to convert glycerol, a by-product of the biodiesel industry, into *n*-butanol, an industrially important chemical and potential liquid transportation fuel, using *Clostridium pasteurianum*. Under certain growth conditions this *Clostridium* species has been shown to predominantly produce *n*-butanol, together with ethanol and 1,3-propanediol, when grown on glycerol. Further increases in the yields of *n*-butanol produced by *C. pasteurianum* could be accomplished through rational metabolic engineering of the strain. Accordingly, in the current report we have developed and exemplified a robust tool kit for the metabolic engineering of *C. pasteurianum* and used the system to make the first reported in-frame deletion mutants of pivotal genes involved in solvent production, namely *hydA* (hydrogenase), *rex* (Redox response regulator) and *dhaBCE* (glycerol dehydratase). We were, for the first time in *C. pasteurianum*, able to eliminate 1,3-propanediol synthesis and demonstrate its production was essential for growth on glycerol as a carbon source. Inactivation of both *rex* and *hydA* resulted in increased *n*-butanol titres, representing the first steps towards improving the utilisation of *C. pasteurianum* as a chassis for the industrial production of this important chemical.

### 1. Introduction

Declining fossil fuel reserves, coupled with environmental concerns over their continued extraction and exploitation have led to strenuous efforts to identify renewable routes to energy and fuels. The two most commonly used biofuels are bioethanol and biodiesel (REN21, 2015). Between 2004 and 2014 biodiesel production has increased more than twelvefold, from 2.4 billion litres/year to 29.7 billion litres/year (REN21, 2015). For every ton of biodiesel produced through the transesterification of vegetable oils and animals fats with short-chain alcohols (most commonly methanol) 100 kg of crude glycerol (10% [w/w]), or glycerine, are formed as a by-product. As a consequence, glycerol availability has increased significantly (Johnson and Taconi, 2007; Yang et al., 2012). The glycerol formed is almost always used in its refined or purified form to make a multitude of products, including cosmetics, pharmaceuticals and food and beverages (Quispe et al., 2013). The glut of glycerine that has resulted from the biodiesel industry has impacted crude and refined glycerol prices. Economically, crude glycerol has shifted from by-product to waste-

product (Kerr et al., 2011; Quispe et al., 2013). With the associated disposal costs, ways to convert crude glycerol into valuable products are becoming increasingly important (Yazdani and Gonzalez, 2007). Currently crude glycerol is utilised as a supplement for animal feeds, reacting agent in chemical catalytic procedures and as a feedstock for the biological production of chemicals and fuels (Yang et al., 2012). The high degree of reduction of its carbon atoms (Clomburg and Gonzalez, 2013), its wide availability and low cost have made glycerol an attractive feedstock for biorefinery (Werpy et al., 2004; Cherubini, 2009; Dobson et al., 2012), especially for anaerobic fermentation processes (Yazdani and Gonzalez, 2007).

*Clostridium pasteurianum* is, unlike other well-studied clostridia, capable of converting glycerol directly into the value-added chemicals *n*-butanol and 1,3-propanediol (PDO) (Biebl, 2001; Jensen et al., 2012; Malaviya et al., 2012). Butanol is an attractive biofuel as it offers a high energy density, low water solubility, low vapour pressure and good blending abilities and can be used in regular combustion engines without the need of modification (Dürre, 2007). The organic compound PDO is used as a precursor for the production of useful polymers such

\* Corresponding author.

E-mail address: [nigel.minton@nottingham.ac.uk](mailto:nigel.minton@nottingham.ac.uk) (N.P. Minton).

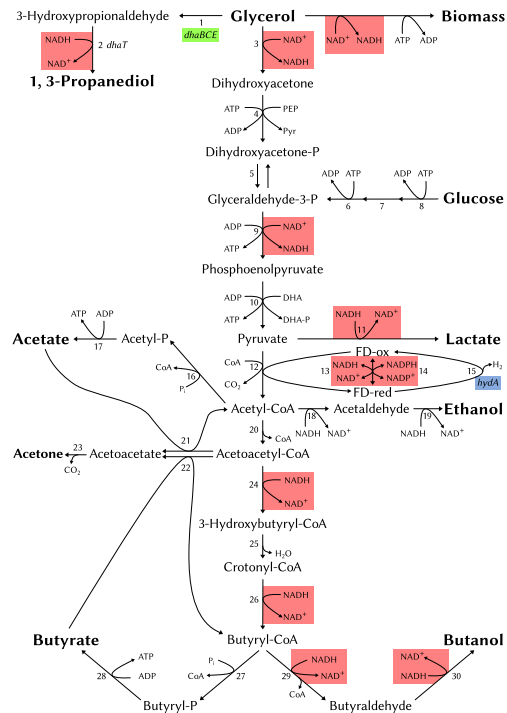
<sup>1</sup> These authors contributed equally to the work.

<http://dx.doi.org/10.1016/j.ymben.2017.01.009>

Received 2 December 2016; Received in revised form 18 January 2017; Accepted 20 January 2017

Available online 22 January 2017

1096-7176/© 2017 The Authors. Published by Elsevier Inc. on behalf of International Metabolic Engineering Society. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



**Fig. 1.** Central metabolic energy pathway in *C. pasteurianum* from glucose and glycerol. Figure based on Malaviya et al. (2012) and Biebl (2001). The knocked out genes (*hydA*, *dhaBCE*) and the NADH/NAD<sup>+</sup> utilising pathways putatively influenced by Rex are highlighted. Other enzymes involved in the central energy pathway are numbered as follows: 1, glycerol dehydratase; 2, 1,3-propanediol oxidoreductase; 3, glycerol-3-phosphate dehydrogenase; 4, dihydroxyacetone kinase; 5, triose-phosphate isomerase; 6, phosphofructokinase; 7, phosphoglucose isomerase; 8, hexokinase; 9, glyceraldehyde-3-phosphate dehydrogenase; 10, pyruvate kinase; 11, lactate dehydrogenase; 12, pyruvate-ferredoxin oxidoreductase; 13, ferredoxin-NADP reductase; 14, NADPH-ferredoxin oxidoreductase; 15, ferredoxin hydrogenase; 16, phosphate acetyltransferase; 17, acetate kinase; 18, acetaldehyde dehydrogenase; 19, ethanol dehydrogenase; 20, thiolase; 21, acetoacetyl-CoA: acetate:CoA transferase; 22, acetoacetyl-CoA: butyrate:CoA transferase; 23, acetoacetate decarboxylase; 24, β -hydroxybutyryl-CoA dehydrogenase; 25, crotonase; 26, butyryl-CoA dehydrogenase; 27, phosphotransbutyrylase; 28, butyrate kinase; 29, butanaldehyde dehydrogenase; 30, butanol dehydrogenase.

as polyesters, polyurethans and polyethers. The polyester polytrimethylene terephthalate (PTT) accounts for 90.0% of the total PDO market which is expected to be worth \$621.2 million by 2021 (Market and Market, 2015). It shares many of the features of its polyester counterpart polybutylene terephthalate (PBT) and polyethylene terephthalate (PET) but offers higher tensile and flexural strengths and stiffness (Zhang, 2004). In addition, PDO is used in composites, adhesives, laminates, coatings, moldings, aliphatic polyesters and antifreeze (Liu et al., 2010).

Under certain growth conditions *C. pasteurianum* has been shown (Dabrock et al., 1992; Jensen et al., 2012; Moon et al., 2011; Taconi et al., 2009) to predominantly produce *n*-butanol, ethanol and PDO with trace amounts of organic acids (see Fig. 1). In optimized batch cultures up to 17 g/l *n*-butanol have been produced from glycerol, in pH auxostat cultures around 7 g/l PDO (Biebl, 2001). Recently, a *C. pasteurianum* mutant strain was generated producing 17.8 g/l *n*-butanol and 8.7 g/l PDO from refined glycerol (Malaviya et al.,

2012). In comparison, engineered *C. acetobutylicum* and *C. beijerinckii* mutant strains have been described to produce 18.9 g/l and 20.9 g/l of butanol (Chen and Blaschek, 1999; Jang et al., 2012). Further increases in yields of *n*-butanol produced by *C. pasteurianum* could be accomplished through rational metabolic engineering of the strain. Accordingly, in the present study we have developed the requisite gene modification systems through the implementation of our previously described roadmap for gene tool development (Minton et al., 2016), and then used the developed allelic exchange vectors to make in-frame deletions of *spo0A* (for exemplification purposes) and a selection of mutants likely to affect solvent yields, namely, *dhaBCE* (encoding glycerol dehydratase), *rex* (coding for the redox responsive repressor Rex) and *hydA* (encoding an iron-coupled hydrogenase).

## 2. Methods

### 2.1. Bacterial strains, growth and maintenance conditions

Bacterial strains used are listed in Table 1. *E. coli* was grown aerobically at 30 °C or 37 °C in Luria-Bertani (LB) broth or agar supplemented with 25 µg/ml chloramphenicol (LB<sub>CM25</sub>) or 100 µg/ml kanamycin (LB<sub>Km100</sub>), respectively. *C. pasteurianum* was routinely grown at 37 °C in an anaerobic workstation (Don Whitley, Yorkshire, UK) in 2x YTG broth (16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl, 5 g/l glucose, pH 6.2) or on RCM agar (Oxoid Ltd) supplemented with 15 µg/ml thiamphenicol (2x YTG<sub>Tm15</sub>, RCM<sub>Tm15</sub>), 20 µg/ml erythromycin (2x YTG<sub>Em20</sub>, RCM<sub>Em20</sub>) or 40 µg/ml uracil (2x YTG<sub>Ura40</sub>, RCM<sub>Ura40</sub>), if required. Selections using 5-fluoroorotic acid (FOA, 600 µg/ml; Sigma-Aldrich, Dorset, UK) and uracil (5 µg/ml) were carried out in modified clostridial basal medium containing 0.5% (w/v) CaCO<sub>3</sub> and 5% (w/v) glucose (CBM-S) (Steiner et al., 2012) or on standard clostridial basal medium (CBM) agar (O'Brien and Morris, 1971). All solidified media contained 1.5% [w/v] agar. Solvent/ acid profiling was undertaken in CGM (Sandoval et al., 2015), Biebl medium (Biebl, 2001) or 2x YT media supplemented with 60 g/l glycerol or glucose. Biebl medium was additionally supplemented with 1 g/l yeast extract for glycerol fermentations Table 2.

### 2.2. Standard molecular biology techniques

Standard plasmids used in this study are detailed in Table 1. Electroporation of *E. coli* was carried out using a Gene-Pulsar (Bio-Rad, Hemel Hempstead, UK) in agreement with the manufacturer's instructions. Plasmid DNA and genomic DNA was isolated using a QIAprep spin miniprep or a DNeasy blood and tissue kit (Qiagen, Manchester, UK), respectively. Restriction digest by endonucleases, ligation reactions and agarose gel electrophoresis were performed according to the manufacturer's instructions. Restriction endonucleases were purchased from ThermoFisher Scientific (Loughborough, UK) and a LigaFast™ Rapid DNA Ligation kit from Promega (Southampton, UK). PCR was carried out using the BIO-X-ACT™ Short Mix (Bioline Reagents, London, UK) in accordance with the manufacturer's instructions. Site-directed mutagenesis (SDM) was carried out using the Quik Change II Site-Directed Mutagenesis Kit (Agilent Technologies, Stockport, UK). Oligonucleotides were ordered from Eurofins Genomics (Ebersberg, Germany) and are given in Supplementary Table S1. Sanger sequencing was carried out by Source BioSciences (Nottingham, UK). Synthetic DNA fragments were ordered either from Biomatik (Cambridge, Ontario/Canada) or GeneArt® Gene Synthesis (Life Technologies, Paisley, UK).

### 2.3. Plasmid methylation and plasmid transfer into *C. pasteurianum*

Plasmid DNA for the transformation into *C. pasteurianum* strain was *in vivo* methylated by propagation in the *dam*<sup>+</sup>, *dcm*<sup>+</sup> *E. coli* host CR1, an *E. coli* Top10 strain harbouring the plasmid pCR1, comprising

**Table 1**  
Strains and plasmids used in this study. *C. pa.* = *C. pasteurianum*.

Name	Designation	Properties	Source
<i>C. pa.</i> ATCC 6013	CRG4080	wild type, type strain	ATCC
<i>C. pa.</i> DSM 525	CRG4091	wild type, type strain	DSMZ
<i>C. pa.</i> DSM 525-H1	CRG4111	hypertransformable strain based on DSM 525	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>pyrE</i>	CRG4273	ACE <i>pyrE</i> truncation mutant	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>pyrE</i> $\Delta$ <i>spo0A</i>	CRG5514	ACE <i>spo0A</i> deletion mutant with $\Delta$ <i>pyrE</i> background	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>spo0A</i>	CRG5516	ACE <i>spo0A</i> deletion mutant with <i>pyrE</i> repaired	This study
<i>C. pa.</i> DSM 525-H1:: <i>spo0A</i> *	CRG5518	ACE <i>spo0A</i> complementation in <i>pyrE</i> locus	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>pyrE</i> $\Delta$ <i>rex</i>	CRG5520	ACE <i>rex</i> deletion mutant with $\Delta$ <i>pyrE</i> background	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>rex</i>	CRG5522	ACE <i>rex</i> deletion mutant with <i>pyrE</i> repaired	This study
<i>C. pa.</i> DSM 525-H1:: <i>rex</i> *	CRG5524	ACE <i>rex</i> complementation in <i>pyrE</i> locus	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>pyrE</i> $\Delta$ <i>hydA</i>	CRG5526	ACE <i>hyd</i> deletion mutant with $\Delta$ <i>pyrE</i> background	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>hydA</i>	CRG5528	ACE <i>hyd</i> deletion mutant with <i>pyrE</i> repaired	This study
<i>C. pa.</i> DSM 525-H1:: <i>hydA</i> *	CRG5530	ACE <i>hyd</i> complementation in <i>pyrE</i> locus	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>pyrE</i> $\Delta$ <i>dhaBCE</i>	CRG5532	ACE <i>dhaBCE</i> deletion mutant with $\Delta$ <i>pyrE</i> background	This study
<i>C. pa.</i> DSM 525-H1 $\Delta$ <i>dhaBCE</i>	CRG5534	ACE <i>dhaBCE</i> deletion mutant with <i>pyrE</i> repaired	This study
<i>C. pa.</i> DSM 525-H1:: <i>dhaBCE</i> *	CRG5536	ACE <i>dhaBCE</i> complementation in <i>pyrE</i> locus	This study
<i>E. coli</i> Top10 x CR1	CRG3131	Strain harbouring plasmid CR1 With M.BepI methylase	This study
pMTL85151		<i>E. coli</i> -Clostridium shuttle vector (pIM13, <i>catP</i> , ColE1 <i>traJ</i> , <i>lacZa</i> ORF/MCS, $T_{Cpa}$ <i>fdx</i> )	Heap et al. (2009)
pMTL-AMH101		<i>catP</i> - <i>pyrE</i> module used for pMTL-KS15	Liew et al. (2017)
pMTL-KS01		pMTL85151, 300-bp internal <i>pyrE</i> fragment, 1200-bp fragment immediately downstream of <i>pyrE</i>	This study
pMTL-KS03		pMTL85151, 300-bp internal <i>pyrE</i> fragment, 1200-bp fragment immediately downstream of <i>pyrE</i> , with $T_{Ctet}$ <i>glaRS</i>	This study
pMTL-KS04		pMTL85151, 937-bp fragment immediately downstream of <i>pyrE</i> , 300-bp internal <i>pyrE</i> fragment, $T_{Ctet}$ <i>glaRS</i> , $T_{Cpa}$ <i>fdx</i>	This study
pMTL-KS05		pMTL85151, 548-bp internal <i>pyrE</i> fragment, 1200-bp fragment immediately downstream of <i>pyrE</i>	This study
pMTL-KS08		pMTL85151, 937-bp fragment immediately downstream of <i>pyrE</i> , 548-bp internal <i>pyrE</i> fragment, $T_{Ctet}$ <i>glaRS</i> , $T_{Cpa}$ <i>fdx</i>	This study
pMTL-KS10		pMTL-KS01 features, $T_{Ctet}$ <i>glaRS</i> , $T_{Cpa}$ <i>fdx</i> , <i>AsiSI</i>	This study
pMTL-KS12		pMTL85151, 548-bp internal <i>pyrE</i> fragment, 937-bp fragment immediately downstream of <i>pyrE</i> , $T_{Ctet}$ <i>glaRS</i> , $T_{Cpa}$ <i>fdx</i> , <i>AsiSI</i>	This study
pMTL-KS15		pMTL85151, <i>pyrE</i> ( <i>C. acetobutylicum</i> ), $T_{Ctet}$ <i>glaRS</i> , $T_{Cpa}$ <i>fdx</i> , 300-bp SHA <i>pyrE</i> , <i>lacZa</i> ORF/MCS	This study
pMTL-KS16		pMTL85151, <i>pyrE</i> ( <i>C. acetobutylicum</i> ), $T_{Ctet}$ <i>glaRS</i> , $T_{Cpa}$ <i>fdx</i> , <i>AsiSI</i> , 300-bp SHA <i>pyrE</i> , <i>lacZa</i> ORF/MCS	This study
pMTL-AGH12		1748-bp fragment comprising <i>pyrE</i> (35–582 nt) and 1200-bp immediately downstream of <i>pyrE</i> cloned into <i>SbfI</i> / <i>AscI</i> recognition sites of pMTL-KS12	This study
pMTL-KS12:: <i>spo0A</i>		1095-bp fragment comprising the 267-bp sequence upstream of <i>spo0A</i> and the <i>spo0A</i> gene cloned into the <i>NotI</i> / <i>NheI</i> recognition sites of pMTL-KS12	This study
pMTL-KS12:: <i>rex</i>		816-bp fragment comprising the 183-bp sequence upstream of <i>rex</i> and the <i>rex</i> gene cloned into the <i>NotI</i> / <i>NheI</i> recognition sites of pMTL-KS12	This study
pMTL-KS12:: <i>hyd</i>		2118-bp fragment comprising the 393-bp sequence upstream of <i>hyd</i> and the <i>hyd</i> gene cloned into the <i>NotI</i> / <i>NheI</i> recognition sites of pMTL-KS12	This study
pMTL-KS12:: <i>dhaBCE</i>		2984-bp fragment comprising the 294-bp sequence upstream of <i>dhaB</i> and the <i>dhaBCE</i> genes cloned into the <i>NotI</i> / <i>NheI</i> recognition sites of pMTL-KS12	This study
pMTL-KS15::KO_ <i>spo0A</i>		1400-bp KO out cassette for <i>spo0A</i>	This study
pMTL-KS15::KO_ <i>rex</i>		1364-bp KO out cassette for <i>rex</i>	This study
pMTL-KS16::KO_ <i>hyd</i>		2002-bp KO out cassette for <i>hyd</i>	This study
pMTL-KS15::KO_ <i>dhaBCE</i>		1602-bp KO out cassette for <i>dhaBCE</i>	This study

a RSP1030-derived RSF origin of replication (Som and Tomizawa, 1982, Novagen, Merck KGaA, Darmstadt, Germany), a kanamycin resistance marker and a gene encoding the M.BepI methyltransferase of *Brevibacterium epidermidis* under the transcriptional control of the *C. sporogenes fdx* promoter ( $P_{Csp}$  *fdx*). Alternatively, *in vitro* methylation was performed by incubation of plasmid DNA with CpG methyltransferase (M.SssI; NEB, Hitchin, UK), according to the manufacturer's instructions. Methylated plasmids 0.5–5  $\mu$ g DNA were electroporated into *C. pasteurianum* as detailed in Supplementary Information.

#### 2.4. Construction of Allele-Coupled Exchange vectors

All Allele-Coupled Exchange (ACE) vectors generated in this study were based on the modular plasmid pMTL85151 (Heap et al., 2009) and conform to previous design principles (Heap et al., 2012). The *pyrE* ACE KO vector (pMTL-KS01) used to generate a  $\Delta$ *pyrE* mutant (strain CRG4273) of *C. pasteurianum* DSM 525 was based on pMTL-JH12 (Heap et al., 2012). The requisite *pyrE* KO cassette comprised a

long (1200 bp) right homology arm (RHA) homologous to a region downstream of *pyrE* and a short (300 bp) left homology arm (LHA) composed of an internal portion (nt 35–334) of the *pyrE* gene (full details in Supplementary Information). The *pyrE* ACE correction (pMTL-AGH12, Fig. 2) and complementation (pMTL-KS12, Fig. 2) vectors were constructed equivalent to pMTL-YN1/YN2 and pMTL-YN1C/YN2C (Ng et al., 2013). Plasmid pMTL-AGH12 was generated through the PCR amplification of a 1748-bp fragment comprising the *pyrE* nucleotides 35–582 (RHA) and a contiguous region 1200-bp downstream of its stop codon (LHA) using primers AGH0001\_ *pyrE*\_Fw and AGH0001\_ *pyrE*\_Rev (Supplementary Table S1) and CRG4111 genomic DNA. Following amplification, the DNA product generated was cleaved with *SbfI* and *AscI*, and inserted between the *SbfI* and *AscI* sites of pMTL-KS10 (More information in Supplementary Information). Plasmid pMTL-KS12 (Fig. 2) was essentially the same as pMTL-AGH12, except in this case the RHA and LHA were separated by a segment of DNA encompassing a *lacZ*-encoding region containing a MCS region. In contrast to pMTL-YN1C/YN2C, however, the RHA was followed by a 77-bp fragment encoding the *C.*

**Table 2**

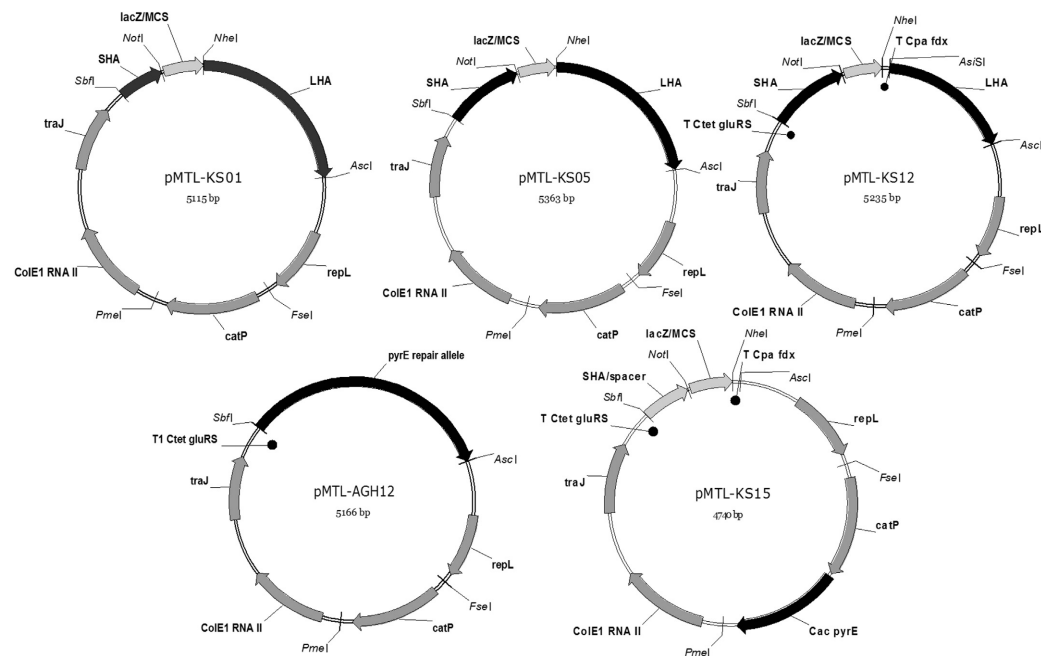
Solvent and acid yields of *C. pasteurianum* DSM 525 (WT) and its various mutant derivatives when grown in bioreactors on either glucose or glycerol as the carbon source. Mutants were: *C. pasteurianum* DSM 525-H1 $\Delta$ rex ( $\Delta$ rex, CRG5522); *C. pasteurianum* DSM 525-H1 $\Delta$ hyd ( $\Delta$ hyd, CRG5528), and; *C. pasteurianum* DSM 525-H1 $\Delta$ dhaBCE ( $\Delta$ dhaBCE, CRG5534). Abbreviations used: butanol (BuOH), ethanol (EtOH), 1,3-propanediol (PDO), solvents (EtOH, BuOH, PDO), acids (acetate, butyrate, lactate). Carbon recovery was calculated by assuming 3.5 g/l dry-weight per 10 OD values (Sarchami et al., 2016), carbon dioxide desorption as described by Percheron et al. (1995) and the assumption that 46.2% of dry-weight is carbon (Papoutsakis and Meyer, 1985). The carbon fraction of yeast extract was neglected. \*Fermentation of  $\Delta$ dhaBCE was undertaken in serum bottles along with a wild type control.

Carbon source	60 g/l glycerol					60 g/l glucose			
	Biebl plus 1 g/l yeast extract			2x YT*		Biebl			
Strain	$\Delta$ rex	$\Delta$ hyd	WT	$\Delta$ dhaBCE	WT	$\Delta$ rex	$\Delta$ hyd	$\Delta$ dhaBCE	WT
<b>Growth characteristics</b>									
Specific growth rate [h <sup>-1</sup> ]	0.32 ± 0.04	0.35 ± 0.01	0.42 ± 0.01	0.39 ± 0.00	0.40 ± 0.00	0.15 ± 0.02	0.11 ± 0.01	0.13 ± 0.03	0.16 ± 0.03
Doubling time [min]	135 ± 15	119 ± 3	100 ± 1	106 ± 1	105 ± 0	90 ± 9	103 ± 10	85 ± 4	81 ± 12
Max. OD	15.0 ± 0.5	17.5 ± 0.6	15.3 ± 0.3	4.38 ± 0.12	4.48 ± 0.01	34.9 ± 1.0	36.2 ± 0.1	34.8 ± 2.5	34.3 ± 2.7
Carbon recovery [%]	88.9 ± 2.7	91.4 ± 1.3	90.5 ± 4.4	114.4 ± 6.0	127.8 ± 5.4	70.7 ± 14.3	83.5 ± 15.0	72.3 ± 13.9	73.3 ± 14.9
<b>Selectivity [M/M]</b>									
BuOH/Solvents	0.516 ± 0.002	0.312 ± 0.003	0.289 ± 0.024	0.862 ± 0.002	0.795 ± 0.001	0.703 ± 0.015	0.729 ± 0.021	0.635 ± 0.027	0.563 ± 0.051
<b>Yield [M/M]</b>									
BuOH/C-Source	0.250 ± 0.005	0.166 ± 0.004	0.142 ± 0.003	0.399 ± 0.019	0.431 ± 0.021	0.124 ± 0.034	0.156 ± 0.029	0.041 ± 0.008	0.031 ± 0.003
EtOH/C-Source	0.053 ± 0.006	0.102 ± 0.001	0.037 ± 0.000	0.064 ± 0.002	0.066 ± 0.003	0.055 ± 0.018	0.061 ± 0.017	0.025 ± 0.007	0.026 ± 0.008
PDO/C-Source	0.182 ± 0.000	0.265 ± 0.005	0.317 ± 0.034	0.000 ± 0.000	0.046 ± 0.002	n.a.	n.a.	n.a.	n.a.
Solvents/C-Source	0.484 ± 0.011	0.533 ± 0.007	0.496 ± 0.030	0.463 ± 0.021	0.542 ± 0.026	0.179 ± 0.052	0.217 ± 0.046	0.066 ± 0.016	0.057 ± 0.001
Acids/C-Source	0.050 ± 0.006	0.106 ± 0.000	0.114 ± 0.008	0.122 ± 0.009	0.136 ± 0.004	0.567 ± 0.129	0.567 ± 0.153	0.686 ± 0.156	0.708 ± 0.177

*tetani* E88 glutamyl-tRNA synthetase terminator ( $T_{Ctet\ gluRS}$ ), while the LHA was preceded by a 42-bp DNA region encompassing the *C. pasteurianum* transcriptional terminator of the ferredoxin gene ( $T_{Cpa\ fdx}$ ). Full details of the construction of pMTL-KS12 is given in [Supplementary Information](#).

For the generation of the complementation vectors target genes including their native promoter were PCR amplified using CRG4111

genomic DNA as template and a gene specific oligonucleotide pair (Fw and Rev) encompassing 5'-*NotI* and a 3'-*NheI* recognition sites (see [Supplementary Table S1](#)). Following amplification, the DNA fragment was digested with *NotI* and *NheI*, purified and ligated into appropriately cleaved pMTL-KS12. Full details of the construction of the complementation vectors is given in [Supplementary Information](#).



**Fig. 2.** Plasmid maps of major plasmids used in this study.



### 2.5. Construction of allelic exchange KO vectors

All generated allelic exchange vectors are based on the modular plasmid pMTL85151 (Heap et al., 2009) designed in accordance with Ng et al. (2013). For the marker-less, in-frame deletion of genes in strain CRG4273 (Table 1), the *pyrE* gene of *C. acetobutylicum* was employed as a heterologous counter (negative) selection marker. Therefore, the *catP-pyrE* module of the plasmid pMTL-AMH101 (Liew et al., 2017) was cloned as a *FseI-PmeI*-fragment into the equivalent sites of pMTL-KS08. The resulting plasmid, pMTL-KS15 (Table 1, Fig. 2), carried strong transcriptional terminators (5' *T<sub>CTet</sub>* *gluRS*, 3' *T<sub>Cpa fids</sub>*) that flanked the site of insertion (between *SbfI* and *NheI*) of the KO cassette. Cassettes were generated by SOE (splicing by overlap extension) PCR as described previously (Ng et al., 2013) and cloned into pMTL-KS15 via *SbfI/NheI* sites. Full details of the construction of AE KO plasmids can be found in Supplementary Information.

### 2.6. Allele-Coupled Exchange procedure

Allele-Coupled Exchange (ACE) vectors were transformed into *C. pasteurianum* DSM 525-H1Δ*pyrE* (CRG4273) and plated onto RCM<sub>Tm15</sub> agar. Putative single crossover integrants were identified as faster growing, larger colonies (Cartman et al., 2012). These were restreaked onto RCM<sub>Tm15</sub> agar and PCR screened for the presence of single crossovers, using primers KS004\_Cpa\_pyrE\_gen\_Fw and KS007\_85151\_LHA\_Rev2 (Supplementary Table S1). In the case of *pyrE* KO using pMTL-KS01, clones confirmed as single crossovers were re-streaked onto CBM<sub>FOA600, Ura5</sub> agar plates and incubated for 24–28 h. Colonies that developed represented putative double crossover mutants. These were patch plated onto CBM, CBM<sub>Tm15U5</sub> and CBM<sub>U5</sub> agar plates to identify double crossover; uracil auxotrophs that had lost the excised plasmid grow only on the latter media. Deletion mutants were verified by Sanger sequencing of the PCR amplified DNA fragment generated using the above primer pair.

The procedure was the same for restoration of the *pyrE* locus to wild type using the *pyrE* ACE correction vector (pMTL-AGH12) and the *pyrE* ACE complementation vectors (pMTL-KS12) carrying functional copies of the gene to be complemented, but in this case the faster growing, single crossover integrants were re-streaked twice onto RCM<sub>Tm15</sub> agar plates to purify and incubated for 16–24 h. Single colonies were then restreaked onto CBM agar plates, and replica plated onto RCM<sub>Tm15</sub> agar plates to check for plasmid loss. Following, successive re-streaking (3–5 times) onto fresh CBM and RCM<sub>Tm15</sub>, clones were identified that exhibited strong growth on CBM but not on RCM<sub>Tm15</sub> agar plates. Restoration of the *pyrE* locus to WT was confirmed by Sanger sequencing of the DNA fragment PCR amplified using primers KS004\_Cpa\_pyrE\_gen\_Fw and KS004\_Cpa\_pyrE\_gen\_Rev (Supplementary Table S1).

### 2.7. Allelic exchange KO procedure

For the construction of specific gene deletions (*spo0A*, *dhaBCE*, *hydA* and *rex*) by allelic exchange, appropriate KO vectors based on pMTL-KS15 were transformed into the *pyrE* deletion strain CRG4273 (Table 1) and plated on RCM<sub>Tm15</sub> agar. Following 48 h incubation, faster growing colonies were re-streaked twice onto RCM<sub>Tm15</sub> agar plates and their identity as single crossover integrants confirmed by Sanger sequencing of the PCR amplified DNA fragment using appropriate primer pairs. Confirmed single crossover mutants were grown overnight in 5 ml CBM<sub>FOA600, Ura5</sub> broth to allow the double crossover to occur, centrifuged (10 min, 8500g, RT) and re-suspended in 250 μl PBS. A 100 μl aliquot of the suspended cells was serially diluted (up to 10<sup>-7</sup>) in PBS and 100 μl of each dilution plated onto RCM agar plates. After a 24 h incubation, 50 single, faster growing colonies were selected and re-streaked in the indicated order onto RCM<sub>Tm15</sub>, CBM and RCM agar plates. Colonies, which lost the plasmid and either reverted back

to the WT or carried the desired deletion exhibited no growth on RCM<sub>Tm15</sub> and CBM agar plates but grew on RCM agar plates. These restreaks were subjected to colony PCR, using gene specific primers that flanked the intended deletion, and the amplified DNA fragment subjected to Sanger sequencing to confirm the expected genotype. The *pyrE* deletion of verified mutants was restored to WT using the ACE plasmid pMTL-AGH12 (Fig. 2).

### 2.8. Next-generation sequencing and resequencing analysis

Illumina sequencing of genomic DNA was done by the Deep Seq: Next Generation Sequencing Facility (University of Nottingham, UK) using a MiSeq Illumina system (Illumina, USA). Paired-end reads were mapped against the published *C. pasteurianum* DSM 525 (strain CRG4091) genome (Poeblein et al., 2015) CLC Genomics Workbench 8.0.2 (Qiagen, DK). Single nucleotide polymorphisms (SNPs) were analysed using the basic variant caller in CLC Genomics Workbench 8.0.2 (Qiagen, DK).

### 2.9. Fermentation in serum flasks

To analyse butanol production, strains were grown either in 50 ml CGM (Sandoval et al., 2015), Biébl medium (Biébl, 2001) or 2x YTG broth in serum bottles at 37 °C with a starting pH of 6.2. Media was supplemented with 60 g/l glucose or glycerol as carbon source. Biébl medium was additionally supplemented with 1 g/l yeast extract for glycerol fermentations. Cultures were initiated in an anaerobic workstation (Don Whitley, Yorkshire, UK) by inoculating and re-suspending several colonies from CBM agar plates into 10 ml CGM, Biébl or 2x YTG broth containing 60 g/l glucose. After overnight incubation, 1 ml of culture was used to inoculate 20 ml CGM or 2x YTG or 35 ml Biébl broth supplemented with 60 g/l glucose or glycerol and incubated for 6 h prior to diluting up to 10<sup>-3</sup> and incubating in fresh media and growing overnight. Those cultures that were in mid-exponential phase were used to inoculate 50 ml CGM, Biébl or 2x YTG broth supplemented with 60 g/l glucose or glycerol to a starting OD<sub>600</sub> of 0.05 (CGM, 2x YTG, Biébl<sub>glucose</sub>) or 0.075 (Biébl<sub>glycerol</sub>). At this point, serum bottles were removed from the anaerobic workstation and incubated 37 °C for up to 48 h. The OD<sub>600</sub> and pH were monitored and samples for product analysis taken every 3 h initially then at 24 h and close to the end of fermentation. All culturing was carried out in triplicate.

### 2.10. pH-controlled batch fermentation

To enable optimum butanol production, pH-controlled batch fermentations were carried out in a Multifors 2 parallel reactor system (Infors UK, Reigate, UK) containing 350 ml Biébl broth supplemented with 60 g/l glucose or glycerol and 1 g/l yeast extract were incubated at 37 °C, 250 rpm and continuous sparging with sterile N<sub>2</sub> (1 l/min). The initial pH of the medium was 6.0. Following inoculation the pH was held above 6.0 by the automatic addition of 4 M KOH. Pre-cultures were grown as described above for the serum flask fermentation, with the exception, that the volume of the final 16 h dilution series was 35 ml. The main fermentation broth was inoculated to a starting OD<sub>600</sub> of 0.75. Batch fermentations were run for 48 h. Growth was monitored through online optical density (OD) as well as through the external measurement of triplicate samples employing a BioMate 3 spectrophotometer (Thermo Fisher Scientific, Loughborough, UK). Samples for solvent analysis were taken at the same time. Additionally the pH, redox potential and temperature of the culture broth were monitored online.

To enable optimum sporulation, pH-controlled batch fermentations were carried out at 37 °C, 200 rpm, with continuous sparging of sterile N<sub>2</sub> (0.02 ml/ml) in the same Multifors 2 parallel reactor system (Infors UK, Reigate, UK) containing 400 ml of CBM broth comprising no CaCO<sub>3</sub> but 6% (w/v) glucose. To set up the pH-controlled batch

fermentation, the strain of interest was streaked from  $-80\text{ }^{\circ}\text{C}$  stocks onto RCM agar plates, incubated overnight at  $37\text{ }^{\circ}\text{C}$  and subsequently used to inoculate 10 ml CBM-S broth containing 6% (w/v) glucose. The 10 ml culture was grown overnight and used to inoculate 50 ml CBM-S broth at a 2% (v/v) inoculum. A 50 ml aliquot of this pre-culture was used to start the pH-controlled batch culture at an inoculum of 6.5% (v/v). The pH was allowed to drop from 6.5 to 5.5, before being maintained at 5.5 by the addition of 4 M KOH. Batch fermentations were run for 120 h.

### 2.11. Spore assay

Spore assays were carried out either in flasks for a quick screen or in pH controlled batch cultures (see Section 2.10) to enable efficient sporulation. In flasks, 200 ml CBM-S broth supplemented with 0.25 mM phosphate buffer (pH 7.3, Steiner et al., 2012) were inoculated from CBM-S overnight culture to a final  $\text{OD}_{600}$  of 0.1 and grown anaerobically for 120 h at  $37\text{ }^{\circ}\text{C}$ . After incubation cultures were shaken thoroughly and 100  $\mu\text{l}$  samples taken in duplicate to be treated at  $80\text{ }^{\circ}\text{C}$  for 10 min and plated on RCM agar plates to quantify the number of heat resistant colony forming units (CFUs) per ml. pH-controlled (5.5) batch cultures were grown for 120 h. Cells were normalized to the lowest  $\text{OD}_{600}$  in a final volume of 5 ml PBS buffer, centrifuged (10 min, 10,000g,  $4\text{ }^{\circ}\text{C}$ ), re-suspended in 200  $\mu\text{l}$  PBS buffer and heated to  $80\text{ }^{\circ}\text{C}$  for 30 min. Serial dilutions in a total volume of 1 ml PBS were carried out and 10  $\mu\text{l}$  aliquots were spotted onto RCM agar plates. The plates were incubated for 24 h at  $37\text{ }^{\circ}\text{C}$  before colonies were enumerated. The sporulation efficiency was determined as number of heat-resistant CFU/ml of culture, that germinate and grow. All results were confirmed microscopically.

### 2.12. HPLC analysis of glycerol, glucose and metabolites

Samples (2 ml) of *C. pasteurianum* cultures from bottle fermentations were collected, centrifuged (10 min, 16,000g,  $4\text{ }^{\circ}\text{C}$ ) and cell-free supernatants stored at  $-80\text{ }^{\circ}\text{C}$  until analysed. Cell-free supernatants were thawed on ice, mixed with an equal volume (200  $\mu\text{l}$ ) of internal standard solution (80 mM valeric acid [Sigma-Aldrich, Dorset, UK] in 0.005 M  $\text{H}_2\text{SO}_4$ ), filtered through a 0.22  $\mu\text{m}$  HPLC certified syringe filter (Whatman® Spartan® 13/0.2 RC; GE Healthcare Life Sciences, Little Chalfont, UK) and transferred into a HPLC vial with a 100  $\mu\text{l}$  insert. Substrate (glucose, glycerol) and fermentation products (acetate, acetone, butanol, butyrate, lactate, ethanol, PDO) were analysed by the use of a Dionex UltiMate 3000 HPLC system (Thermo Fisher Scientific, Loughborough, UK) equipped with a Bio-Rad Aminex HPX-87H (Hertfordshire, UK) column, a refractive index (RI) and diode array detector (DAD) at UV 210 nm at an isocratic flow rate of 0.5 ml/min of 0.005 M  $\text{H}_2\text{SO}_4$  as mobile phase and a column temperature of  $35\text{ }^{\circ}\text{C}$  for 55 min. The injection volume was 20  $\mu\text{l}$ . If required samples were diluted using reverse osmosis (RO) water. Standard concentrations ranged from 0.98 to 250 mM. For glycerol two additional concentrations of 500 mM and 750 mM were employed. Signal analysis was performed using the Chromeleon 7.2 Chromatography Data System (Thermo Fisher Scientific, Loughborough, UK). Statistical analysis of significance was performed using the PRISM (GraphPad Software, La Jolla, USA) employing Fisher's *t*-test. The significance level ( $\alpha$ -value) was set at 0.05.

### 2.13. Identification of putative Rex boxes in *C. pasteurianum*

The *C. pasteurianum* genome was analysed for the presence of Rex binding boxes using the Rex box consensus sequence (5'-TTGTTAANNNTTAAACAA) reported by Ravcheev et al. (2012) with the 'Virtual Footprint' algorithm (Münch et al., 2005) allowing 2 mismatches in a similar fashion to Wietzke and Bahl (2012). As the *C. pasteurianum* genome was not available in the PRODORIC database

(Münch et al., 2003) to which 'Virtual Footprint' is linked, the results were analysed manually. Each 'hit' was searched against the genome of *C. pasteurianum* DSM 525 (Poehlein et al., 2015) using the Artemis genome browser (Rutherford et al., 2000). The information extracted was the locus tag, gene name if applicable, protein function and location of the target sequence in respect of the genome context. Only target regions in intergenic regions with a putatively regulated gene up- or downstream were considered. For every entry in the list the distance to the closest start codon was measured and a frequency plot with a bin size of 10 bp was considered. The result is a positive skewed distribution with the maximal target site occurrence between 80 bp and 90 bp from the start codon which dropped to near zero occurrences after 250 bp from the start codon. This observation led to further filtering the list with a threshold of 250 bp distance to the start codon. The resulting list comprised 40 sequences which were used in an iterative approach to run 'Virtual Footprint' with a new Position Weight Matrix (PWM). This led to a list of 113 targets which were filtered to exclude hits in coding regions of genes and in intergenic regions with antiodmic genes. The final list comprised 47 targets which putatively regulate downstream genes (Supplementary Table S2).

## 3. Results

### 3.1. Improved transformation of *C. pasteurianum* DSM 525

*C. pasteurianum* ATCC 6013 has been shown to be transformable with several pMTL80000 modular vectors (Heap et al., 2009) at efficiencies of up to  $7.5 \times 10^3$  transformants  $\mu\text{g}^{-1}$  DNA (Pyne et al., 2013). This high frequency was reliant on circumventing the activity of the endogenous *CpaAII* restriction-modification system by *in vivo* methylation of its recognition site (5'-CGCG-3') using the *M.FnuDII* methyltransferase (Pyne et al., 2013). As the *E.coli* host employed (Top10) was *dam+*, the activity of the restriction enzyme *CpaAII*, a *MboI/DpmII*-type restriction endonuclease previously identified as *CpaI* (Richards et al., 1988), was of no consequence. Here we used either *in vitro* or *in vivo* methylation by *M.SssI* or *M.BepI*, respectively. The former methylates all cytosine residues within the double-stranded recognition site 5'-<sup>m</sup>CG-3' (Renbaum et al., 1990), whereas similar to *M.FnuDII*, *M.BepI* methylates the external cytosine (5'-<sup>m</sup>CGCG-3') of the *CpaAII* recognition site (Lunnen et al., 1988). Whilst either methylation procedure allowed transformation of both ATCC 6013 and DSM 525 (equivalent to ATCC 6013) to be obtained with plasmids pMTL83151 (pCB102 replicon) and pMTL85151 (pIM13 replicon) (Heap et al., 2009) using our published method (Cooksley et al., 2012), the success rate and frequencies obtained was very low,  $1.5 \times 10^1$  transformants  $\mu\text{g}^{-1}$  DNA. After adapting the method published by Pyne et al. (2013) with added sucrose in growth medium and extra step of glycine wash, the transformation frequency was improved to  $1.6 \times 10^2$  transformants  $\mu\text{g}^{-1}$  DNA, still two orders of magnitude lower than reported by Pyne et al. (2013) of which up to  $7.5 \times 10^4$  transformants  $\mu\text{g}^{-1}$  DNA was achieved. Plasmids pMTL82151 (pBP1 replicon) and pMTL84151 (pCD6 replicon) could not be transformed.

We hypothesised that one reason for the low number of transformants was that those cells that were competent for the acquisition of plasmid DNA represented rare mutant cells present within the population. This hypothesis was tested by curing the plasmid of a randomly selected transformant of DSM 525 through repeated sub-culture in the absence of the selective antibiotic, and then retesting the cured strain to see if the plasmid-free strain was more amenable to electroporation. The strain was found to transform at frequencies of  $2.6 \times 10^5$  transformants  $\mu\text{g}^{-1}$  DNA. Moreover, similar frequencies were observed with plasmids pMTL85151, pMTL82151 and pMTL84151 (Heap et al., 2009). Whole genome sequencing of the strain with higher frequency *C. pasteurianum* DSM 525-H1 (designated CRG4111) and compared to the parental strain *C. pasteurianum* DSM 525 (Poehlein et al., 2015) indicated that two SNPs had arisen within two distinct

open reading frames, that of CLPA\_c13710 and CLPA\_c33080 (Poehlein et al., 2015). Unexpectedly, neither encoded protein was obviously associated with restriction/ modification. Both SNPs resulted in frame-shifts in the encoded sequence. CLPA\_c13710 encodes a predicted  $\beta$ -lysine N6-acetyltransferase (*ablB*) which was reduced from 283 amino acids to 176 residues in the hypertransformable mutant. CLPA\_c33080 encoded a histidine kinase (ResE9) and leads to a frameshift that reduces the protein from 615 to 160 amino acids in length. To ascertain whether the general phenotypic properties of the CRG4111 mutant strain had been affected by the SNPs, comparative growth experiments were performed on glycerol as the carbon source. The growth and glycerol consumption rates of the mutant were similar to the parental strain. Aside from a small increase in the levels of PDO and a small reduction in *n*-butanol, all other measured metabolites (ethanol, lactate, acetate and butyrate) were essentially the same between the two strains. On this basis, the strain was taken forward for metabolic engineering studies.

### 3.2. Implementation of a gene system roadmap

We have previously described the implementation of allelic exchange, gene KO systems in specifically generated *pyrE* mutants of *Clostridium acetobutylicum* (Ehsaan et al., 2016) and *Clostridium difficile* (Ng et al., 2013), and most recently *Geobacillus thermoglucosidasius* (Sheng et al., 2017). The generation of mutants by allelic exchange is facilitated by the use of replication defective, or pseudo-suicide vectors (Cartman and Minton, 2010), and the incorporation into the vector of a functional copy of a heterologous *pyrE* gene (encoding orotate phosphoribosyl transferase) that serves as a counter/negative selection marker. This approach serves as a general roadmap for the implementation of gene KO systems in clostridia (Minton et al., 2016). Pivotal are the generation of a *pyrE* truncation mutant using Allele-Coupled Exchange (ACE) technology (Heap et al., 2012) and the identification of an appropriately replication defective (pseudo-suicide) vector.

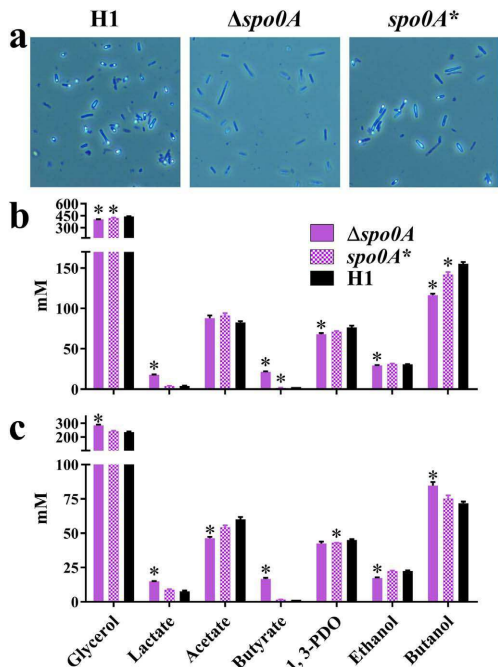
Segregation stability studies (see Supplementary Information) on transformed cells carrying the various pMTL80000 modular vectors established that plasmids based on the pIM13 replicon were suitably defective, with 99% of the cells losing the plasmid after eight 12 h subcultures (approximately 51 generations) in media lacking antibiotic selection. Accordingly, a pIM13-based vector (pMTL-KS01) was constructed (Fig. 2 and Materials and Methods) equivalent to the ACE plasmids pMTL-YN18 (Ng et al., 2013) and pMTL-JH12 (Heap et al., 2012) and a *pyrE* (CLPA\_c26850) truncation mutant generated as previously described (Heap et al., 2012; Ng et al., 2013). In essence, thiamphenicol (Tm) resistant (<sup>R</sup>) transformants in which pMTL-KS01 had integrated into the CRG4111 genome *via* a longer Right Homology Arm (RHA) were identified as larger, faster growing colonies on RCM media supplemented with 15  $\mu$ g/ml Tm (RCM<sub>Tm15</sub>) and purified by restreaking onto the same selective media. Of the 24 selected colonies, all were shown to be pure single crossover mutants through PCR screening using appropriate primers (see Supplementary Table S1). To select for a subsequent double crossover excision event involving the shorter Left Homology Arm (LHA) and the generation of the required 5-fluoroorotic acid (FOA), resistant *pyrE* deletion mutant, single crossover integrants were restreaked onto CBM media supplemented with 5  $\mu$ g/ml uracil (Ura) and FOA at the experimentally determined MIC (600  $\mu$ g/ml) of the parental strain (CBM<sub>Ura5, FOA600</sub>). From a total of eight FOA<sup>R</sup> clones obtained, subsequent restreaking and testing of phenotype on appropriately supplemented media demonstrated that in addition to being FOA<sup>R</sup>, three were uracil auxotrophs (required supplementation of CBM media with exogenous uracil for growth) and sensitive (<sup>S</sup>) to Tm (could not grow if Tm was present). From PCR screening using primers flanking the *pyrE* gene and the subsequent sequencing of the amplified DNA fragment, two of the three clones were shown to carry the desired modification. This equated to a

deletion extending from nt 335–582 of CLPA\_c26850, and the concomitant insertion of the *lacZa*/MCS originating from pMTL-KS01. One strain was selected for further use and designated CRG4273. To demonstrate that CRG4273 could be restored to wildtype (WT) through correction of the *pyrE* allele an ACE, *pyrE* repair vector was made (pMTL-AGH12) equivalent to pMTL-YN1/2 (Ng et al., 2013) and pMTL-ME6 (Ehsaan et al., 2016) (see Fig. 2). The region of homologous DNA in this plasmid essentially comprises a contiguous region of DNA from the *pyrE* locus that includes a complete copy of the *pyrE* gene and downstream region equivalent to the 1200 bp RHA of plasmid pMTL-KS01. This plasmid was transformed into CRG4273 and single crossover integrants identified as larger, faster growing colonies on CBM<sub>Tm15, Ura5</sub> agar plates. Growth from several representative colonies were thereafter restreaked onto CBM agar lacking any supplementation, and single colonies patch plated onto RCM and RCM<sub>Tm15</sub> agar plates to identify those clones in which the excised plasmid had been lost. The *pyrE* locus of three representative clones was amplified by PCR using appropriate PCR primers (KS004\_Cpa\_pyrE\_gen\_Fw and KS004\_Cpa\_pyrE\_gen\_Rev, Supplementary Table S1) and subjected to Sanger sequencing. This data confirmed that the locus had been restored to WT in all three cases. One was chosen for storage and designated CRG5570 (Table 1).

### 3.3. Exemplification of the knock-out vector pMTL-KS15

The availability of the *pyrE* deletion mutant CRG4273 is a prerequisite for the use of a functional heterologous *pyrE* gene as a counter selection marker in gene knock-out (KO) by allelic exchange. Accordingly, a KO vector broadly equivalent to pMTL-YN3/4 (Ng et al., 2013), pMTL-ME3 (Ehsaan et al., 2016) and was made, pMTL-KS15 (Fig. 2), and exemplified through the in-frame deletion of the master regulator of sporulation, *spo0A*. Precise details of the vector are given in Materials and Methods, but is essentially based on the pIM13 replicon and incorporates the *C. perfringens catP* selectable marker and a heterologous *pyrE* gene derived from *C. acetobutylicum* (Liew et al., 2017), as opposed to the *C. sporogenes pyrE* gene used in the previous studies (Ehsaan et al., 2016; Ng et al., 2013).

The KO cassette for *spo0A* (CLPA\_c19180) was generated by SOE PCR as described in Materials and Methods and cloned between the *SbfI* and *NheI* sites of pMTL-KS15. Following transformation of the resultant plasmid (pMTL-KS15::spo0A, Table 1) into CRG4273 and plating on RCM<sub>Tm15</sub> agar plates, putative single crossover integrants were identified as faster growing, larger colonies. Their identity as pure single crossover integrants was confirmed by appropriate PCR (primers KS005\_spo0A\_genome\_Fw, KS005\_rex\_genome\_Rev, Supplementary Table S1) and then one of the twelve identified grown overnight in CBMS<sub>FOA600, Ura5</sub> broth to allow the double crossover event to take place. After the overnight incubation the culture was centrifuged, re-suspended and serial dilutions plated to single colonies on RCM agar plates. Selected colonies were patch plated onto RCM<sub>Tm15</sub> and un-supplemented RCM agar plates and onto CBM media lacking uracil. Growth on RCM alone confirmed loss of the plasmid, and the encoded *catP* and *pyrE* genes, following its excision. PCR screening with flanking primers (KS005\_spo0A\_genome\_Fw and KS005\_spo0A\_genome\_Rev, Supplementary Table S1), was then used to distinguish between those excision events that had generated the desired deletion as opposed to return of the cell to WT. A total of 17 of the 50 putative KO clones screened generated a smaller DNA fragment consistent with the intended deletion in *spo0A* (Supplementary Fig. S1). Through Sanger sequencing of the amplified DNA fragment, the presence of the expected *spo0A* deletion was confirmed in all 17 strains. One of the clones was selected (CRG5514) and restored to uracil prototrophy using the ACE vector pMTL-AGH12 as previously described. The final strain, carrying only the *spo0A* mutation, was designated CRG5516. In parallel, a derivative of pMTL-KS12 was constructed (pMTL-KS12::spo0A\*, Table 1) in which a functional copy of the *C. pasteurianum spo0A* gene, together with its



**Fig. 3.** Comparison of sporulation and fermentation phenotypes of *C. pasteurianum*-H1 (H1), *C. pasteurianum*-H1 $\Delta$ *spo0A* ( $\Delta$ *spo0A*) and *C. pasteurianum*-H1-*spo0A* complementation (*spo0A*\*). a) Spores can be observed in H1 and *spo0A*\* whereas  $\Delta$ *spo0A* as expected is incapable of producing spores. b), c) Pure glycerol fermentation in serum bottles with 60 g/l glycerol in Biebl medium (b) and CGM (c) was carried out for 48 h with fermentation being visibly completed after 24 h which is the time point shown for glucose consumption and product formation. \* indicate statistical significance in *t*-test  $\alpha > 0.05$  of deletion or complementation strain against H1. Error-bars indicate standard error of three replicates.

native promoter, was inserted into the MCS between the LHA and RHA. Use of this strain to restore the *pyrE* mutation of CRG5514 to WT led to insertion of *spo0A* into the chromosome for complementation studies. The genotype of the strain generated, CRG5518, was confirmed by PCR amplification of the *pyrE* locus and inserted *spo0A* gene and nucleotide sequencing of the DNA fragment amplified.

#### 3.4. Phenotypic analysis of *spo0A* deletion mutant

As expected, after 5 days of growth in a pH-controlled batch cultures (pH 5.5), no heat resistant CFUs were obtained after plating on RCM of the *spo0A* mutant strain CRG5516. In contrast, cultures of the WT (CRG4111) and the complemented mutant strain CRG5518 yielded  $1 \times 10^8$  heat resistant CFU/ml. In keeping with these measurements, spores were easily detected using phase contrast microscopy in the cell suspensions of CRG4111 (parent) and CRG5518 (*spo0A*, complemented), whereas none could be detected in cultures of CRG5514 (*spo0A*) (Fig. 3a). To analyse the effects of the *spo0A* deletion on solvent profiles, strains CRG4111 (parent), CRG5516 (*spo0A*) and CRG5518 (*spo0A*, complemented) were grown in 50 ml Biebl medium supplemented with 60 g/l glycerol (Fig. 3b). All strains exhibited a similar growth rate but the pH of the *spo0A* mutant (CRG5516) dropped to a lower level than the parent and complemented strain, presumably due to increased production of lactate and butyrate. Glycerol uptake, acetate production and re-assimilation and

ethanol production are similar in all strains.

Butanol and PDO levels were slightly lower in the mutant in Biebl medium. In contrast, the complemented *spo0A* mutant strain (CRG5518) produced slightly higher levels of these solvents. The data obtained was at some variance from that of Sandoval et al. (2015) who undertook a similar analysis of a *spo0A* mutant of *C. pasteurianum*. However, these authors used rich media (CGM) as opposed to the minimal media employed here. We, therefore, repeated our analysis using CGM (Fig. 3c). In this medium, the production of lactate and butyrate by the *spo0A* mutant was significantly increased compared to the parent strain, CRG4111. Interestingly, the mutant (CRG5516) produced much more acetate in the initial stages of the fermentation, but then re-assimilated the acid from 12 h onwards. No equivalent re-assimilation was seen in either the parental (CRG4111) or complemented (CRG5518) cultures. In common with the observations of Sandoval et al. (2015), compared to the parental strain (CRG4111), glycerol consumption was significantly increased together with the production of higher levels of *n*-butanol. In contrast to Sandoval et al. (2015), however, the PDO levels of the *spo0A* mutant (CRG5516) were not reduced compared to the wildtype but remained unaltered.

#### 3.5. Construction of deletion mutants in genes involved in solvent production

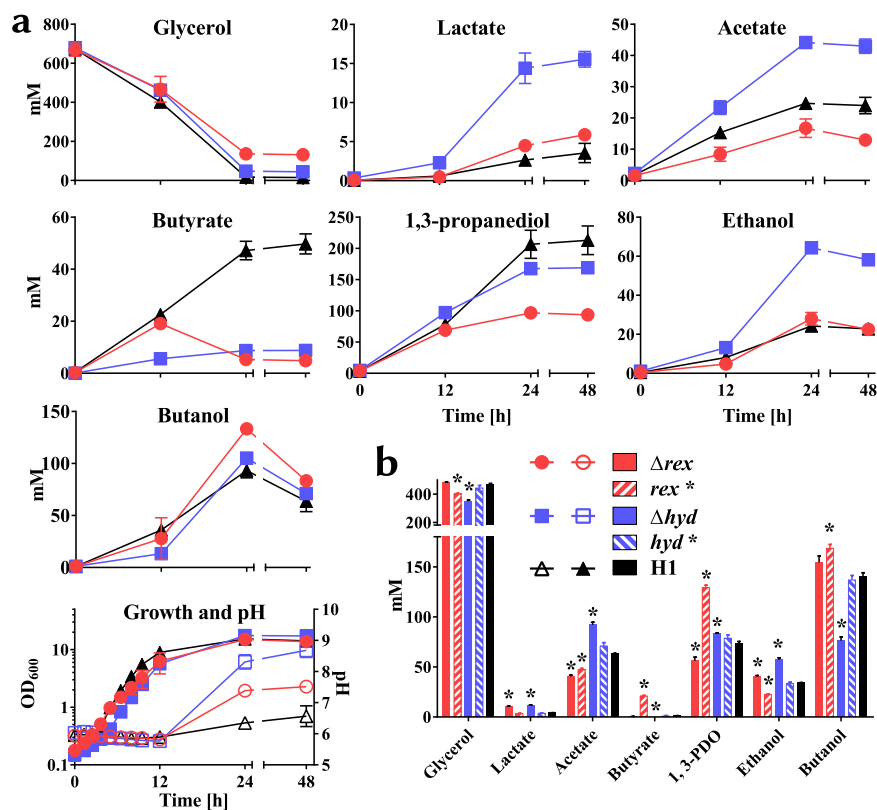
To investigate whether solvent profiles could be altered in favour of *n*-butanol production we targeted genes encoding the redox-responsive regulator Rex, the main hydrogenase (*hydA*) and the glycerol dehydratase (*dhaBCE*).

Rex has been previously shown to control the expression of *n*-butanol biosynthetic genes in response to the cellular NADH/NAD<sup>+</sup> ratio in *C. acetobutylicum* where its disruption led to reduced acid production and increased solvents yields (Wietzke and Bahl, 2012). In *C. acetobutylicum* the *rex* gene resides immediately 5' to the *crt-bcd-etzAB-hbd* operon responsible for butyryl-CoA biosynthesis (Wietzke and Bahl, 2012). An equivalent gene (CLPA\_c28640) is found in the same context in *C. pasteurianum* and encodes a 213 amino acid protein exhibiting 76% amino acid sequence identity to the *C. acetobutylicum* Rex. In the case of hydrogenase, a number of different [FeFe]-hydrogenases and [NiFe]-hydrogenase exist in *C. pasteurianum* but little information is available concerning their specific individual roles in hydrogen formation and redox balance. BLAST analysis showed that the protein encoded by CLPA\_c00280 exhibits 71% identity to the well described *C. acetobutylicum* hydrogenase, *hydA* (Santangelo et al., 1995). Accordingly, KO cassettes to bring about the deletion of CLPA\_c28640 (Rex) and CLPA\_c00280 (Hyd) were generated by SOE PCR, cloned into pMTL-KS15 and the resultant mutants generated by allelic exchange using our standard procedure (Table 1, Supplementary Fig. S1 and Materials and Methods).

Simply, one way to increase butanol titres when growing on glycerol might be to eliminate PDO production. The conversion of glycerol to PDO involves a two-step transformation by glycerol dehydratase (DhaBCE), encoded by *dhaBCE* (CLPA\_c22770-CLPA\_c22760-CLPA\_c22750) and PDO dehydrogenase (DhaT), encoded by *dhaT* (CLPA\_c22740). Given the inability of Pyne et al. (2016) to isolate a clean mutant of *dhaT* in the presence of the toxic intermediate 3-hydroxypropionaldehyde (3-HPA) (Barbirato et al., 1996; Maervoet et al., 2014), we chose to ablate PDO production by eliminating the synthesis of glycerol dehydratase through deletion of the entire *dhaBCE* operon. Accordingly, a KO cassette to achieve this was generated by SOE PCR, cloned into pMTL-KS15 and the resultant KO plasmid used to generate a *dhaBCE* mutant (CRG5532) by allelic exchange using our standard procedure (Table 1, Supplementary Fig. S1 and Materials and Methods).

In all three cases (*rex*, *hyd* and *dhaBCE*), the mutants were restored to uracil prototrophy, through ACE-mediated correction of the *pyrE* mutation, and complemented through the integration of a functional





**Fig. 4.** Pure glycerol fermentations of *C. pasteurianum*-H1 (H1), *C. pasteurianum*-H1 $\Delta$ *rex* ( $\Delta$ *rex*), *C. pasteurianum*-H1-*rex* complementation (*rex*\*), *C. pasteurianum*-H1 $\Delta$ *hyd* ( $\Delta$ *hyd*) and *C. pasteurianum*-H1-*hyd* complementation (*hyd*\*). a) Bioreactor fermentation with 60 g/l glycerol in Biobl medium with 1 g/l yeast extract at pH 6 was carried out for 48 h with fermentation being visibly completed after 24 h. Error-bars indicate range of two fermentations. b) Histogram showing product formation of serum bottle fermentation of deletion strains and complementations. Glycerol usage and product formation is shown after 24 h. \* indicate statistical significance in a *t*-test  $\alpha > 0.05$  of deletion or complementation strain against H1. Error-bars indicate standard error of three replicates.

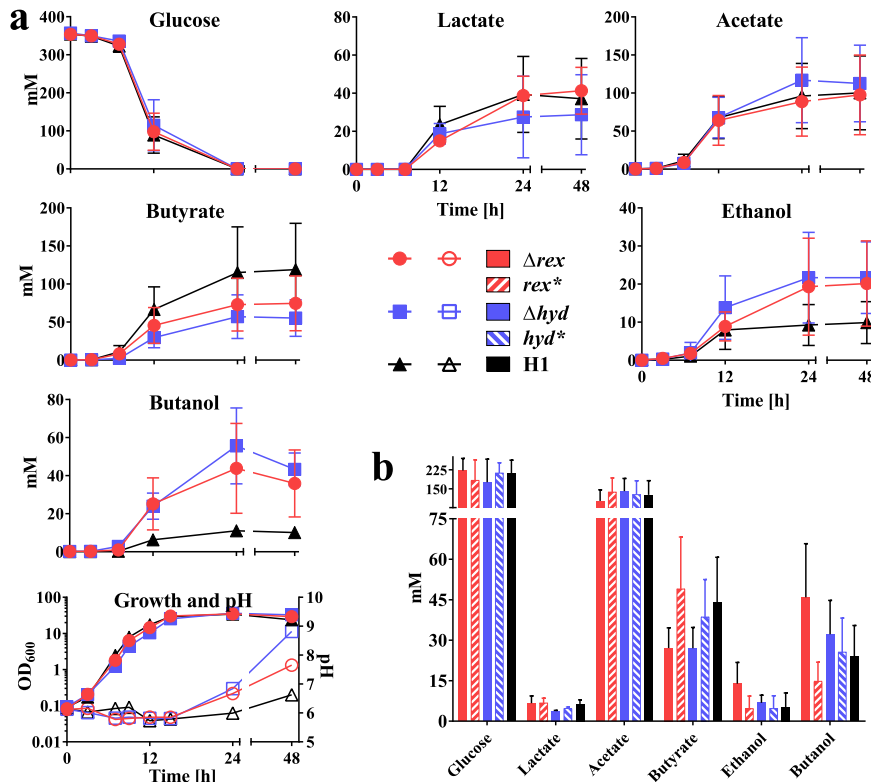
copy of the deleted gene with the native promoter concomitant with restoration of the mutant *pyrE* allele to WT (Materials and Methods). Phenotypic analysis of all mutants, and their complemented equivalents, involved the determination of carbon utilisation as well as the concentration of the following fermentation products: acetate, butyrate, lactate, acetone, ethanol, PDO and *n*-butanol. All studies, with the exception of the glycerol fermentation of the *dhaBCE* mutant (CRG5534), were carried out in batch culture in bioreactors using Biobl medium supplemented with 1 g/l yeast extract and 6% (w/v) glycerol or with glucose omitting yeast extract. Fermentation profiles are shown in Figs. 4–6.

### 3.6. Phenotypic analysis of the *rex* and *hyd* mutants

In glycerol media (Fig. 4), both mutants grew at broadly equivalent growth rates compared to the WT, but achieved higher final pH values, being  $8.7 \pm 0.2$  and  $7.5 \pm 0.1$  after 48 h for the *hydA* and *rex* mutant, respectively compared to  $6.6 \pm 0.3$  in the WT culture. Almost complete glycerol consumption took place in the WT and *hydA* culture (2% and 7% remaining, respectively), with the *rex* mutant culture containing 20% after 48 h. The *hydA* mutant produced the highest titres of both lactate and acetate, but significantly reduced amounts of butyrate. A

slightly greater reduction in butyrate production was observed with the *rex* mutant compared to the *hydA* mutant. In terms of lactate and acetate production, the former were slightly increased compared to the WT, whereas the latter marginally decreased. In the case of solvents, PDO production was reduced in both mutants, with the largest reduction occurring in the *rex* mutant, which decreased by 53% compared to the WT, with only a 19% reduction being evident in the *hydA* mutant. Ethanol formation in the *hydA* mutant was significantly elevated ( $64.3 \pm 3.2$  mM after 24 h) over both the WT and *rex* mutant which, with respective titres of  $24.2 \pm 1.4$  mM and  $28.0 \pm 3.2$  mM, were broadly equivalent. Butanol formation was highest in the *rex* mutant, with the titres achieved after 24 h being  $133.3 \pm 1.8$  mM compared to  $105.1 \pm 0.0$  mM in the *hydA* mutant and  $93.2 \pm 5$  mM in the WT. After 24 h the ethanol and butanol levels decreased, most likely due to solvent extraction by gas stripping from the nitrogen bubbling through the reactor.

On glucose (Fig. 5), the growth rate and glucose utilisation rates were little affected in either mutant compared to the WT. Differences in acid production were also less marked in the two mutants compared to when grown on glycerol. Acetate production in the *hydA* mutant was slightly increased compared to the WT and *rex* mutant, which were essentially the same. On the other hand, lactate production in the *hydA*



**Fig. 5.** Glucose fermentations of *C. pasteurianum*-H1 (H1), *C. pasteurianum*-H1 $\Delta$ rex ( $\Delta$ rex), *C. pasteurianum*-H1-rex complementation ( $rex^*$ ), *C. pasteurianum*-H1 $\Delta$ hyd ( $\Delta$ hyd) and *C. pasteurianum*-H1-hyd complementation ( $hyd^*$ ). a) Bioreactor fermentation with 60 g/l glucose in Biebl medium at pH 6 was carried out for 48 h with fermentation being visibly completed after 24 h. Error-bars indicate range of two fermentations. b) Histogram showing product formation of serum bottle fermentation of deletion strains and complementations. Glucose usage and product formation is shown after 24 h. Error-bars indicate standard error of three replicates.

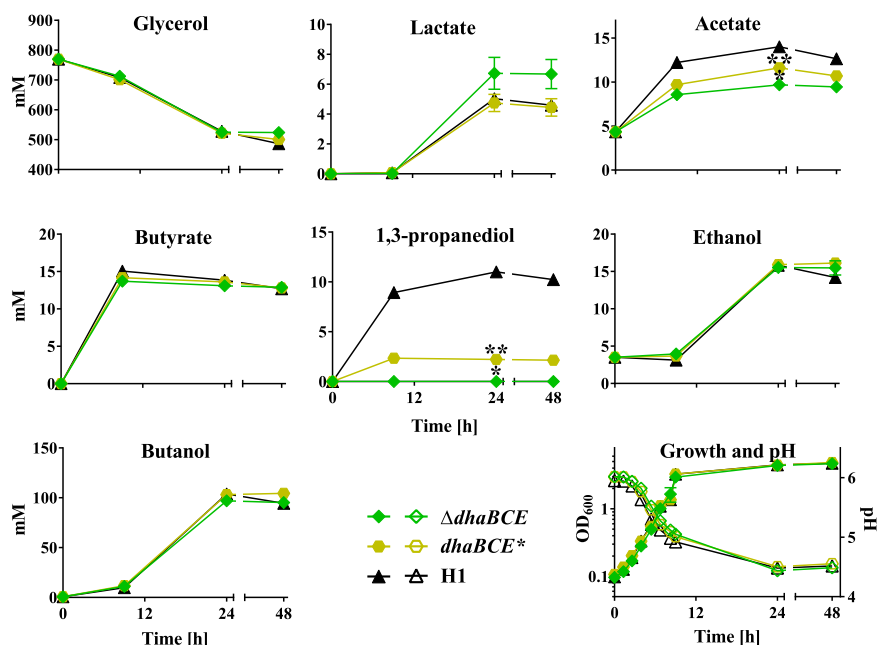
mutant was only slightly reduced compared to the *rex* mutant and WT. Butyrate produced was decreased in both mutants compared to the WT, with the larger reduction being seen in the *hydA* mutant, producing  $57.0 \pm 14.3$  mM compared to  $115.0 \pm 30.0$  mM in WT culture after 24 h. Ethanol and *n*-butanol titres, and in particular the latter solvent, were increased in both mutants, with *hydA* demonstrating a 2.3-fold and 5-fold higher final titre than the WT, respectively, at 24 h.

Whilst the fermentation profile of the complemented *hydA* mutant was essentially restored to that of the WT, complementation of the *rex* mutant was less clear cut. Thus, whereas the PDO levels of the mutant were reduced in comparison to the WT and the *n*-butanol levels increased, the complemented strain produced higher amounts of PDO and reduced titres of *n*-butanol, relative to the WT. Such a phenotype would be indicative of higher than normal expression of *rex* in the complemented strain, however, the *rex* gene was inserted at *pyrE* under the control of its native promoter. In certain instances expression of genes can be affected by the proximity of the gene to the origin of chromosome replication (Sauer et al., 2016). This does not appear to be the case here, as the distance of *rex* and *pyrE* to the origin are relatively similar with *rex* being slightly further away from the origin. In *C. kluyveri*, *rex* is self-regulated through the agency of a Rex box, repressor site immediately upstream of the structural gene (Hu et al., 2016). Whilst such an operator site was not evident in the *C.*

*pasteurianum* *rex* promoter region used to express the complementing copy of *rex*, a Rex box was found 340 bp upstream of the *rex* start codon within the upstream CLPA\_c28650 gene (genome position 3068768–3068786). As the promoter region used in the complemented strain lacks this sequence it is likely that *rex* expression is deregulated, resulting in the observed produced higher amounts of PDO and reduced titres of *n*-butanol, relative to the WT.

### 3.7. Phenotypic analysis of the *dhaBCE* mutant

Growth of *C. pasteurianum* on glycerol is not redox balanced because cell biomass (C.D. 4.3) is more oxidised than glycerol (C.D. 4.67). The requirement to oxidise the excess of reducing equivalents generated is met by the PDO pathway which ensures the net oxidation of 1 mol of NADH per 1 mol of PDO formed. The central role of this pathway in maintaining redox balance suggests that the imbalance caused by its inactivation in *C. pasteurianum* would lead to an inability to ferment glycerol. Indeed, the mutant (CRG5534) was unable to grow in a bioreactor in Biebl medium when the carbon source was 6% (w/v) glycerol. The mutant was, however, able to grow in 2x YT containing glycerol (6%, w/v) at equivalent rates to the WT (Fig. 6). The most notable change in fermentation profile was the production of PDO, which, as expected, was eliminated in the mutant (Fig. 6). In contrast,



**Fig. 6.** Pure glycerol fermentations of *C. pasteurianum*-H1 (H1), *C. pasteurianum*-H1Δ*dhaBCE* (Δ*dhaBCE*) and *C. pasteurianum*-H1-*dhaBCE* complementation (*dhaBCE*\*). Serum bottle fermentation with 60 g/l glycerol in 2x YT medium was carried out for 48 h with fermentation being visibly completed after 24 h. \* indicates statistical significance of Δ*dhaBCE* and \*\* of *dhaBCE*\* in a *t*-test  $\alpha > 0.05$  against H1. Error-bars indicate standard error of three replicates.

the production of metabolites derived from acetyl-CoA (butyrate, ethanol and *n*-butanol) were largely unaffected, apart from a decrease of acetate production to 9.7 mM at time point 24 h compared to 14 mM in the WT, and an increase in lactic acid production (6.7 mM at time point 48 h compared to 4.6 mM in the WT). The solvent profiles of CRG5534 on glucose were as expected essentially the same as the WT (Fig. S2). In the complemented strain, in which *dhaBCE* has been integrated at *pyrE* together with its native promoter, production of PDO was re-instated, although at a lower level than the WT. The lactate and acetate profiles were very similar to that of the WT (Fig. S2).

#### 4. Discussion

The use of *pyrE* alleles for gene knock-out forms part of a general ‘roadmap’ for manipulating bacterial genomes (for a review see Minton et al., 2016), the implementation of which is reliant on effective gene transfer into the target organism. In the case of *C. difficile* (Ng et al., 2013) and *G. thermoglucosidasius* (Sheng et al., 2017) no special measures were required to achieve this, whereas in *C. acetobutylicum* the rational circumvention of host restriction barriers proved to be necessary by appropriate methylation of the plasmid DNA being transferred (Ehsaan et al., 2016). Here we showed that an alternative approach is possible through the isolation of a highly transformable variant of the *C. pasteurianum* strain DSM 525, strain CRG4111.

The reason for the increased competence of CRG4111 is unclear. Neither of the two SNPs present are in genes with any obvious association with restriction/ modification. Both are destructive in nature resulting in frame-shifts and premature termination of the encoded proteins. Whether one or the other, or both, are responsible for the phenotype is not immediately apparent. One gene encodes  $\beta$ -lysine N6-acetyltransferase (*ablB*), which, being involved in lysine

degradation seems to have no connection with improved DNA transfer. CLPA\_c33080 encodes a histidine kinase (ResE9), but does not appear to be part of a two-component system, as no gene encoding a cognate response regulator was located in the immediate vicinity. Interestingly, however, the gene (CLPA\_c33090) immediately adjacent to the ResE9 encodes a homologue of an *agr* quorum sensing AgrB protein, responsible for processing and secretion of the AgrD signal molecule. Indeed, an unannotated ORF is present immediately downstream of CLPA\_c33090 encoding a 57 amino acid peptide that shares significant identity with clostridial AgrD signal peptides, including 58% with those of *Clostridium tetanomorphum* DSM 665 (KAJ50110.1 and KAJ51453.1), *Clostridium scatologenes* (AKA70070.1) and *Clostridium carboxidivorans* P7 (EET84565.1) and 52% identity to homologues in *Clostridium botulinum* (WP\_061325181) and *Clostridium sporogenes* (WP\_053818607.1). All of these clostridial species, and many more besides, share the same structural arrangement of genes, corresponding to *agrB > agrD > resE*. This gene arrangement, was first noted in *C. botulinum*, where it was shown to be involved in the regulation of sporulation and neurotoxin production (Cooksley et al., 2010). The possibility exists that in *C. pasteurianum* at least, this quorum sensing system also regulates a process that can affect the efficiency of DNA uptake.

Regardless of the mechanism responsible for increased competence, our findings have provided a further approach for maximising DNA transfer rates which was used in combination with the developed *pyrE*-based KO system to make a number of mutants (*spoOA*, *hydA*, *rex* and *dhaBCE*) that influenced solvent production. A general workflow and schematic for mutant generation is shown in Supplementary Fig. S3. Significantly we were able, for the first time, to completely eliminate PDO production in *C. pasteurianum*. The formation of PDO from glycerol is mediated by the sequential action of glycerol dehy-

dratase (DhaBCE) and PDO dehydrogenase (DhaT). Here we chose to entirely delete the *dhaBCE* region, and as a consequence generated a strain which no longer produced PDO under all conditions tested. Previous attempts to ablate the production of either glycerol dehydratase or PDO dehydrogenase through the use of ClosTron technology met with little success (Pyne et al., 2013). Intron mutants of *dhaB*, *dhaC* and *dhaE* could not be isolated, whilst an intron insertion in *dhaT* was obtained but the intron delivery plasmid and its encoded LtrA protein could not be cured. Such a phenomenon can occur if the intron insertion is in the sense strand of an essential gene, and LtrA-mediated splicing of the transcribed mRNA provides a route for some production of the encoded protein. However, the intron insertion in the *dhaT* gene was in the antisense orientation. This led to the search for, and apparent discovery of an additional ecotopic, sense strand, intron insertion in *speA* (encoding arginine decarboxylase) which was hypothesised as being essential. One consequence of these difficulties was that Pyne et al. (2016) were unable to eliminate PDO production.

Glycerol is metabolized both oxidatively and reductively in *Clostridium* (Malaviya et al., 2012). The pyruvate-generating oxidative pathway leads to the production of CO<sub>2</sub>, H<sub>2</sub>, and electrons in the form of the reducing equivalent NADH, which is also released during biomass formation. Glycerol, being a highly reduced substrate, results in twice the amount of reducing equivalents compared to using glucose as fermentation substrate. These electrons require acceptors for redox homeostasis, which is precisely the purpose of glycerol reducing pathway producing the highly reduced product PDO. By deleting *dhaBCE* in *C. pasteurianum*, the pathway for funneling extra reducing power during glycerol fermentation is absent, resulting in the poor growth we observed in the mutant in defined medium. The mutant was, however, able to grow in a rich medium containing yeast extract. The latter ingredient supports growth of the mutant by both providing an alternative source of carbon and acting as a reducing agent. Re-introducing *dhaBCE* back into the deletion mutant reversed the phenotype, confirming the production of PDO in *C. pasteurianum* is important for biomass production and redox balance. Our findings support the hypothesis of Johnson and Rehmann (2016), who found that unlike *C. acetobutylicum*, *C. pasteurianum* glycerol fermentation does not display strong biphasic behaviour, suggesting that PDO production is regulated and further that its regulation may serve to function in redox homeostasis and allow *C. pasteurianum* to behave in a non-biphasic manner.

The Rex response regulator directly senses changes in the redox status of the cell, having a much higher affinity for NADH than for the oxidised nucleotide NAD<sup>+</sup> (Wang et al., 2008). Under oxic conditions (i.e., when the NADH/NAD<sup>+</sup> ratio is low) it binds to its target sites to inhibit gene transcription but dissociates from these sites when the NADH/NAD<sup>+</sup> ratio increases (Brekasis and Paget, 2003). Thus, it was expected that inactivation of Rex would lead to increased levels of those enzymes whose genes were under Rex-mediated control, leading to a consequent change in metabolic profiles. To predict such regulated genes we used the procedure outlined in Materials and Methods to search for potential Rex target sites in the genome. We identified a list of 47 putative Rex box sequences which potentially regulate the adjacent downstream genes (Supplementary Table S2) and defined a consensus sequence for *C. pasteurianum*, 5'-TTGTTAAWNNNTTAACAA. Of the non-fermentative genes previously reported, none were found using our approach (Ravcheev et al., 2012; Zhang et al., 2014). However, other notable putative Rex targets found include CLPA\_c05450, a nitroreductase related to *narAB* or *narK* in *C. carboxidivorans* (Zhang et al., 2014), CLPA\_c09900, a nicotinamidase family protein, CLPA\_c17480 (*spoVD2*) encoding a stage V sporulation protein and CLPA\_c39090, a NADPH dehydrogenase.

Given that a high ratio of NADH/NAD<sup>+</sup> inhibits binding of Rex to DNA, it is expected that Rex-binding sites should be located upstream of genes encoding NADH-dependent enzymes. This is true for the alcohol dehydrogenases putatively involved in butanol and ethanol

production (*adhE2* and its two alleles *adhE1* and *adhE4*), indirectly the thiolases that channel carbon into the reductive part of the fermentative pathway (*thlA1* and *thlA2*), the *crt-bcd-efBA-hbd* operon (crotonase-butyryl-CoA dehydrogenase-electron transfer flavoprotein B/A-hydroxybutyryl-CoA dehydrogenase) which uses 2 NADH to reduce acetyl-CoA to butyryl-CoA, and the reductive part of the *dha* (glycerol dehydratase) operon. In keeping with these observations, production of ethanol and *n*-butanol was elevated in the *rex* mutant when grown on either glycerol or glucose, as was demonstrated in *C. acetobutylicum* (Wietzke and Bahl, 2012).

By the same reasoning, and based on observations made in *C. acetobutylicum* (Ravcheev et al., 2012), we would have expected to find Rex sequences upstream of *ldh* (lactate dehydrogenase), *hyaA* (hydrogenase) and *ptb-buk* (phosphotransbutyrylase-butyrate kinase), as was found in other species (Wietzke and Bahl, 2012; Hu et al., 2016), but none were found, even with relaxed algorithm parameters. Nevertheless, contrary to *in silico* expectations and the observations made in *C. acetobutylicum*, the *C. pasteurianum rex* mutant produced more lactate in glycerol fermentation compared to WT (Fig. 4). On the other hand, as expected, butyrate production was significantly reduced in the absence of Rex, and a clear butyrate reassimilation in glycerol fermentation was observed in the *rex* deletion mutant. Butyrate assimilation has been shown in *C. pasteurianum* when feeding external butyrate which led to increased butanol titres (Kao et al., 2013; Lin et al., 2015; Sabra et al., 2014) but reassimilation of butyrate in a standard fermentation was to our knowledge never observed before. Some of the higher *n*-butanol titres observed in the mutant strain could be explained by reassimilation and carbon recycling of butyrate.

Acetone formation was never observed in fermentations using the *C. pasteurianum* WT, in agreement with Biebl (2001), Sandoval et al. (2015) and Pyne et al. (2016). However, in the *rex* deletion mutant butyrate is visibly reassimilated which in the pathway proposed by Malaviya et al. (2012) is coupled to acetone formation via the acetoacetyl-CoA: butyrate:CoA transferase in a similar fashion to *C. acetobutylicum* (Jones and Woods, 1986). These findings suggest that *C. pasteurianum* may be able to reassimilate butyrate without the need of acetone production either through the reverse reaction of the butyrate kinase (*buk*) and the phosphotransbutyrylase (*ptb*), as previously suggested (Hüsemann and Papoutsakis, 1989; Millat et al., 2014), or via a third, as yet undescribed, mechanism proposed by Millat et al. (2014). Reassimilation of butyrate without acetone production has been described in several *C. acetobutylicum* inactivational mutants, most notably mutants of *ctfA* (Millat et al., 2014), *ptb* (phosphotransbutyrylase) and *adc* (acetoacetate decarboxylase) (Lehmann et al., 2012) and butyrate kinase (*buk*) gene (Desai et al., 1999). The fact that *C. pasteurianum* does not produce acetone under any condition tested here, or elsewhere (Biebl, 2001; Pyne et al., 2016; Sandoval et al., 2015), makes this organism an ideal target for investigations of the butyrate and acetate uptake mechanisms not reliant on CtfAB.

Since alcohol production competes for electrons with the formation of molecular hydrogen, reducing the flux of electrons towards H<sub>2</sub> synthesis might be expected to increase *n*-butanol titres. Various approaches have been reported that were designed to inhibit *in vivo* hydrogenase activity, for example, by the application of artificial electron carriers or sparging the culture with carbon monoxide (Datta and Zeikus, 1985; Hüsemann and Papoutsakis, 1989; Kim et al., 1984; Peguin et al., 1994). The resultant increased NAD(P)H availability, led to more *n*-butanol and ethanol and less acetone being produced and a consequently improved alcohol: acetone ratio in *C. acetobutylicum*. However, previous attempts to inactivate the gene (*hyaA*) encoding the major hydrogenase of *C. acetobutylicum* using the ClosTron were unsuccessful (Cooksley et al., 2012), suggesting that the gene is essential under the conditions tested. Whilst the *C. pasteurianum hyaA* gene (CLPA\_c00280) sharing the closest similarity to the *C. acetobutylicum hyaA* gene was recently knocked-down by Pyne et al.



(2015) using an anti-sense approach, here we were able to generate a null mutant through the deletion of the entire CLPA\_c00280 encoding gene.

In keeping with predictions, deletion of *hydA* led to increases in the products of NADH-consuming pathways, with the mutant producing 2.7-fold higher levels of ethanol and 5-fold higher levels of lactate (Fig. 4). In contrast, only a slight increase in *n*-butanol was observed with the mutant generating only 1.1-fold higher levels than the WT. Similar results were obtained with the knock-down mutant of Pyne et al. (2015). Although not a NADH-consuming pathway, a significant increase (1.8-fold) in acetate levels was also seen. Both the acetate and butyrate pathways generate ATP *via* substrate level phosphorylation. One consequence of which was an increase in biomass attained by the mutant (the mutant OD<sub>600</sub> was 114% higher than the WT) and was characterised by a shorter fermentation time as adjudged by the earlier onset of increase in pH (after 18.7 h as opposed to 22.0 h in the WT). The same observation was made by Pyne et al. (2015). This increase in acetate production was not evident in the study of Dabrock et al. (1992), who used CO to inhibit hydrogenase activity on glucose grown *C. pasteurianum*. Here acetate production was reduced by almost 50%. In this instance, however, all hydrogenase activity would have been affected. The mutant strain utilised in our study, and that of Pyne et al. (2015), is only affected in the production of HydA (CLPA\_c00280). Strain *C. pasteurianum* DSM 525 contains 3 more genes/ enzymes (CLPA\_c07060-70, CLPA\_c33960 and CLPA\_c37830) (Poehlein et al., 2015) with homology to hydrogenase enzymes. With CLPA\_c07060 and CLPA\_c07070 *C. pasteurianum* seems to possess a rare [NiFe]-hydrogenase with the genes encoding the small and large subunit, respectively. The other gene encodes an iron only hydrogenase. All of these proteins might compensate for the loss of HydA (CLPA\_c00280).

Unexpectedly, PDO production was significantly reduced in the *hydA* mutant, by 19% compared to the WT (Fig. 4). This is counter intuitive as the PDO pathway is NADH-consuming. The study of Pyne et al. (2016) also reported such a reduction in levels of PDO (in this case 30%) in their knock-down mutant, while noting that PDO levels can vary widely between fermentations – an observation also made in other studies (Biebl, 2001, Tacconi et al., 2009). Our data reinforces the view that the reduction in PDO production that results from ablation of *hydA* function is a real phenomenon. It should also be noted that the reduction in PDO and accompanying increase in biomass is contrary to suggestion of Biebl (2001) that the production of reducing equivalents by biomass build up has to be equalled by production of PDO to regenerate the pool of NAD<sup>+</sup> to NADH.

## 5. Conclusion

In the current report we have developed and exemplified a robust tool kit for the metabolic engineering of *C. pasteurianum* based on the use of *pyrE* alleles. To maximise gene transfer we deployed a novel approach based on the isolation of a highly transformable variant within the host population. In keeping with previous work on other Gram-positive chassis, we once again showed the advantage of using a *pyrE* minus strain as the host through the rapid complementation of the deletion mutants by the integration of the complementing gene into the genome at the *pyrE* locus concomitant with ACE-mediated restoration of prototrophy. The system was used to make in-frame deletion mutants of pivotal genes involved in solvent production, namely *hydA* (hydrogenase), *rex* (Rex response regulator) and *dhaBCE* (glycerol dehydratase). We were, for the first time in *C. pasteurianum*, able to eliminate PDO synthesis and demonstrate its production was essential for growth on glycerol as a sole carbon source. Inactivation of both *rex* and *hydA* resulted in increase in *n*-butanol titres, representing the first steps towards improving the utilisation of *C. pasteurianum* as a chassis for the industrial production of this important chemical.

## Competing interests

All of the authors declare that they have no competing interests.

## Acknowledgements

This work was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) [grant number BB/L004356/1] as part of the TSB/BBSRC “Advancing the Industrial Application of Synthetic Biology Feasibility Study Competition”. CR acknowledges the financial support of the European Community’s Seventh Framework Program “CLOSTNET” [Grant number PEOPLE-ITN-2008–237942] while AGH acknowledges the support of the Forman Hardy Trust and the University of Nottingham, Vice-Chancellors Scholarship. The authors thank Klaus Winzer and Wouter Kuit for fruitful discussions and Matthew Abbott for his help with the HPLC methodology and analysis.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jymben.2017.01.009.

## References

- Barbtrato, F., Grivet, J.P., Soucaille, P., Bories, A., 1996. 3-Hydroxypropionaldehyde, an inhibitory metabolite of glycerol fermentation to 1,3-propanediol by enterobacterial species. *Appl. Environ. Microbiol.* 62, 1448–1451.
- Biebl, H., 2001. Fermentation of glycerol by *Clostridium pasteurianum*—batch and continuous culture studies. *J. Indust. Microbiol. Biotechnol.* 27, 18–26.
- Brekasis, D., Paget, M.S., 2003. A novel sensor of NADH/NAD<sup>+</sup> redox poise in *Streptomyces coelicolor* A3. *EMBO J.* 22 (2), 4856–4865.
- 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.
- 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 *tdcC* genotype and toxin production. *Appl. Environ. Microbiol.* 78, 4683–4690.
- Chen, C.K., Blaschek, H.P., 1999. Acetate enhances solvent production and prevents degeneration in *Clostridium beijerinckii* BA101. *Appl. Microbiol. Biotechnol.* 52 (2), 170–173.
- Cherubini, F., Jungmeier, G., Mandl, M., Philips, C., Wellisch, M., Jørgensen, H., Skiadas, I., Boniface, L., Dohy, M., Pouet, J.C. and Willke, T., 2009. IEA Bioenergy Task 42: Report on Participating Countries. Document of IEA Bioenergy Task 42 “Bioferries”.
- Clomburg, J.M., Gonzalez, R., 2013. Anaerobic fermentation of glycerol: a platform for renewable fuels and chemicals. *Trends Biotechnol.* 31 (1), 20–28.
- Cooksley, C.M., Davis, L.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 (13), 4448–4460.
- 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 (6), 630–641.
- Dabrock, B., Bahl, H., Gottschalk, G., 1992. Parameters affecting solvent production by *Clostridium pasteurianum*. *Appl. Environ. Microbiol.* 58, 1233–1239.
- Datta, R., Zeikus, J.G., 1985. Modulation of acetone–butanol–ethanol fermentation by carbon monoxide and organic acids. *Appl. Environ. Microbiol.* 49, 522–529.
- Desai, R.P., Harris, L.M., Welker, N.E., Papoutsakis, E.T., 1999. Metabolic flux analysis elucidates the importance of the acid-formation pathways in regulating solvent production by *Clostridium acetobutylicum*. *Metab. Eng.* 1, 206–213.
- Dobson, R., Gray, V., Rumbold, K., 2012. Microbial utilization of crude glycerol for the production of value-added products. *J. Indust. Microbiol. Biotechnol.* 39, 217–226.
- Dürre, P., 2007. Biobutanol: an attractive biofuel. *Biotechnol. J.* 2, 1525–1534.
- 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.
- Heap, J.T., Ehsaan, M., Cooksley, C.M., Ng, Y.K., Cartman, S.T., Winzer, K., Minton, N.P., 2012. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucl. Acids Res.* 40, e59.
- 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.
- Hu, L., Huang, H., Yuan, H., Tao, F., Xie, H., Wang, S., 2016. Rex in *Clostridium kluyveri* is a global redox-sensing transcriptional regulator. *J. Biotechnol.* 233, 17–25.
- Hüsemann, M.H., Papoutsakis, E.T., 1989. Comparison between *in vivo* and *in vitro* enzyme activities in continuous and batch fermentations of *Clostridium*

- acetobutylicum*. Appl. Microbiol. Biotechnol. 30, 585–595.
- 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., Lee, S.Y., 2012. Enhanced butanol production obtained by reinforcing the direct butanol-forming route in *Clostridium acetobutylicum*. MBio 3, e00314–12.
- Jensen, T.O., Kvist, T., Mikkelsen, M.J., Christensen, P.V., Westermann, P., 2012. Fermentation of crude glycerol from biodiesel production by *Clostridium pasteurianum*. J. Indust. Microbiol. Biotechnol. 39, 709–717.
- Johnson, D.T., Taconi, K.A., 2007. The glycerin glut: options for the value-added conversion of crude glycerol resulting from biodiesel production. Environ. Prog. 26, 338–348.
- Johnson, E.E., Rehmann, L., 2016. The role of 1,3-propanediol production in fermentation of glycerol by *Clostridium pasteurianum*. Bioresour. Technol. 209, 1–7.
- Jones, D.T., Woods, D.R., 1986. Acetone-butanol fermentation revisited. Microbiol. Rev. 50, 484.
- Kao, W.C., Lin, D.S., Cheng, C.L., Chen, B.Y., Lin, C.Y., Chang, J.S., 2013. Enhancing butanol production with *Clostridium pasteurianum* CH4 using sequential glucose-glycerol addition and simultaneous dual-substrate cultivation strategies. Bioresour. Technol. 135, 324–330.
- Kerr, B.J., Shurson, G.C., Johnston, L.J., Dozier, W.A., 2011. Utilization of crude glycerin in nonruminants. INTECH Open Access Publisher.
- Kim, B.H., Bellows, P., Datta, R., Zeikus, J.G., 1984. Control of carbon and electron flow in *Clostridium acetobutylicum* fermentations: utilization of carbon monoxide to inhibit hydrogen production and to enhance butanol yields. Appl. Environ. Microbiol. 48, 764–770.
- Lehmann, D., Hönigke, D., Ehrenreich, A., Schmidt, M., Weuster-Botz, D., Bahl, H., Lütke-Eversloh, T., 2012. Modifying the product pattern of *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. 94, 743–754.
- 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. <http://dx.doi.org/10.1016/j.ymben.2017.01.007>, (in press), pii: S1096-7176(17)30031-9.
- Lin, D.S., Yen, H.W., Kao, W.C., Cheng, C.L., Chen, W.M., Huang, C.C., Chang, J.S., 2015. Bio-butanol production from glycerol with *Clostridium pasteurianum* CH4: the effects of butyrate addition and in situ butanol removal via membrane distillation. Biotechnol. Biofuels 8, 1.
- Liu, H., Xu, Y., Zheng, Z., Liu, D., 2010. 1,3-Propanediol and its copolymers: research, development and industrialization. Biotechnol. J. 5, 1137–1148.
- Lunnen, K.D., Barsomian, J.M., Camp, R.R., Card, C.O., Chen, S.Z., Croft, R., Looney, M.C., Meda, M.M., Moran, L.S., Nwankwo, D.O., Slatko, B.E., 1988. Cloning type-II restriction and modification genes. Gene 74, 25–32.
- Maervoet, V.E., De Maeseneire, S.L., Avci, F.G., Beauprez, J., Soetaert, W.K., De Mey, M., 2014. 1,3-propanediol production with *Citrobacter werkmanii* DSM17579: effect of a *dhaD* knock-out. Microb. Cell. Fact. 13, 1.
- Malaviya, A., Jang, Y.S., Lee, S.Y., 2012. Continuous butanol production with reduced byproducts formation from glycerol by a hyper producing mutant of *Clostridium pasteurianum*. Appl. Microbiol. Biotechnol. 93, 1485–1494.
- Market and Market, 2015. 1,3-Propanediol (PDO) Market by Applications (PTT, Polyurethane, Cosmetic, Personal Care & Home Cleaning & Others) & Geography - Global Market Trends & Forecasts to 2021, Report code: CH3091, available online (<http://www.marketsandmarkets.com/Market-Reports/1-3-propanediol-pdo-market-760.html>).
- Millat, T., Voigt, C., Janssen, H., Cooksley, C.M., Winzer, K., Minton, N.P., Bahl, H., Fischer, R.J., Wolkenhauer, O., 2014. Coenzyme A-transferase-independent butyrate re-assimilation in *Clostridium acetobutylicum*—evidence from a mathematical model. Appl. Microbiol. Biotechnol. 98, 9059–9072.
- 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.
- Moon, C., Lee, C.H., Sang, B.I., Um, Y., 2011. Optimization of medium compositions favoring butanol and 1,3-propanediol production from glycerol by *Clostridium pasteurianum*. Bioresour. Technol. 102, 10561–10568.
- Münch, R., Hiller, K., Barg, H., Heldt, D., Linz, S., Wingender, E., Jahn, D., 2003. PRODORIC: prokaryotic database of gene regulation. Nucl. Acids Res. 31, 266–269.
- Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., Jahn, D., 2005. Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. Bioinformatics 21, 4187–4189.
- Ng, Y.K., Ehsaan, M., Philip, S., Colley, 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, e56051.
- O'Brien, R.W., Morris, J.G., 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. Microbiol. 68, 307–318.
- Papoutsakis, E.T., Meyer, C.L., 1985. Fermentation equations for propionic-acid bacteria and production of assorted oxychemicals from various sugars. Biotechnol. Bioeng. 27 (1), 67–80.
- Peguín, S., Goma, G., Delorme, P., Soucaille, P., 1994. Metabolic flexibility of *Clostridium acetobutylicum* in response to methyl viologen addition. Appl. Microbiol. Biotechnol. 42, 611–616.
- Percheron, G., Brechignac, F., Soucaille, P., Condoret, J.S., 1995. Carbon dioxide desorption from fermentation broth by use of oxygen vents. Bioprocess Eng. 12 (1–2), 11–16.
- Poehlein, A., Grosse-Honebrink, A., Zhang, Y., Minton, N.P., Daniel, R., 2015. Complete genome sequence of the nitrogen-fixing and solvent-producing *Clostridium pasteurianum* DSM 525. Genome Announc. 3, (e01591-14).
- Pyne, M.E., Moo-Young, M., Chung, D.A., Chou, C.P., 2013. Development of an electrotransformation protocol for genetic manipulation of *Clostridium pasteurianum*. Biotechnol. Biofuels 6, 1.
- Pyne, M.E., Moo-Young, M., Chung, D.A., Chou, C.P., 2015. Antisense-RNA-mediated gene downregulation in *Clostridium pasteurianum*. Fermentation 1 (1), 113–126.
- Pyne, M.E., Sokolenko, S., Liu, X., Srirangan, K., Bruder, M.R., Aucoin, M.G., Moo-Young, M., Chung, D.A., Chou, C.P., 2016. Disruption of the reductive 1,3-propanediol pathway triggers production of 1,2-propanediol for sustained glycerol fermentation by *Clostridium pasteurianum*. Appl. Environ. Microbiol. 82, 5375–5388.
- Quispe, C.A., Coronado, C.J., Carvalho, J.A., Jr., 2013. Glycerol: production, consumption, prices, characterization and new trends in combustion. Renew. Sustain. Energy Rev. 27, 475–493.
- Ravcheev, D.A., Li, X., Lafi, H., Zengler, K., Leyn, S.A., Korostelev, Y.D., Kazakov, A.E., Novichkov, P.S., Osterman, A.L., Rodionov, D.A., 2012. Transcriptional regulation of central carbon and energy metabolism in bacteria by redox-responsive repressor Rex. J. Bacteriol. 194, 1145–1157.
- REN21, 2015. Renewables Global Status Report, REN21Secretariat, Editor. 2015. p. ISBN 978-3-9815934-6-4.
- Renbaum, P., Abrahamov, D., Fainsod, A., Wilson, G.G., Rottem, S., Razin, A., 1990. Cloning, characterization, and expression in *Escherichia coli* of the gene coding for the CpG DNA methylase from *Spiroplasma sp.* strain MQ1 (M Sssl). Nucl. Acids Res. 18, 1145–1152.
- Richards, D.F., Linnett, P.E., Oultram, J.D., Young, M., 1988. Restriction endonucleases in *Clostridium pasteurianum* ATCC 6013 and *C. thermoautotrophicum* DSM 568. Microbiol. 134, 3151–3157.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell, B., 2000. Artemis: sequence visualization and annotation. Bioinformatics 16, 944–945.
- Sabra, W., Groeger, C., Sharma, P.N., Zeng, A.P., 2014. Improved n-butanol production by a non-acetone producing *Clostridium pasteurianum* DSMZ 525 in mixed substrate fermentation. Appl. Microbiol. Biotechnol. 98, 4267–4276.
- Sarchami, T., Johnson, E., Rehmann, L., 2016. Optimization of fermentation condition favoring butanol production from glycerol by *Clostridium pasteurianum* DSM 525. Bioresour. Technol. 208, 73–80.
- Sandoval, N.R., Venkataramanan, K.P., Groth, T.S., Papoutsakis, E.T., 2015. Whole-genome sequence of an evolved *Clostridium pasteurianum* strain reveals Spo0A deficiency responsible for increased butanol production and superior growth. Biotechnol. Biofuels 8, 1.
- Santangelo, J.D., Dürre, P., Woods, D.R., 1995. Characterization and expression of the hydrogenase-encoding gene from *Clostridium acetobutylicum* P262. Microbiol. 141, 171–180.
- Sauer, C., Sverrtsson, S., Bohorquez, L., Cruz, R., Harwood, C., Van Rij, T., Hamoen, L.W., 2016. Effect of genome position on heterologous gene expression in *Bacillus subtilis*: an unbiased analysis. ACS Synth. Biol. 5, 942–947.
- Sheng, L., Kovács, K., Winzer, K., Zhang, Y., Minton, N.P., 2017. Development and implementation of rapid metabolic engineering tools for chemical and fuel production in *Geobacillus thermoglucosidarius* NCIMB 11955. Biotechnol. Biofuels. 10, 5.
- Som, T., Tomizawa, J.I., 1982. Origin of replication of *Escherichia coli* plasmid RSF 1030. Mol. Gen. Genet. 187, 375–383.
- Steiner, E., Scott, J., Minton, N.P., Winzer, K., 2012. An *agr* quorum sensing system that regulates granulose formation and sporulation in *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 78, 1113–1122.
- Taconi, K.A., Venkataramanan, K.P., Johnson, D.T., 2009. Growth and solvent production by *Clostridium pasteurianum* ATCC® 6013™ utilizing biodiesel-derived crude glycerol as the sole carbon source. Environ. Prog. Sust. Eng. 28 (1), 100–110.
- Wang, E., Bauer, M.C., Rogstam, A., Linse, S., Logan, D.T., Von Wachenfeldt, C., 2008. Structure and functional properties of the *Bacillus subtilis* transcriptional repressor rex. Mol. Microbiol. 69, 466–478.
- Werp, T., Todd, Petersen, Gene, Aden, A., Bozell, J., Holladay, J., White, J., Manheim, Amy, Eliot, D., Lasure, L., Jones, S., 2004. Top value added chemicals from biomass. volume 1—results of screening for potential candidates from sugars and synthesis gas. No. DOE/GO-102004-1992. Dep. Energy Wash. DC.
- Wietzke, M., Bahl, H., 2012. The redox-sensing protein rex, a transcriptional regulator of solventogenesis in *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. 96, 749–761.
- Yang, F., Hanna, M.A., Sun, R., 2012. Value-added uses for crude glycerol—a byproduct of biodiesel production. Biotechnol. Biofuels 5, 1.
- Yazdani, S.S., Gonzalez, R., 2007. Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry. Curr. Opin. Biotechnol. 18, 213–219.
- Zhang, J., 2004. Study of poly(trimethylene terephthalate) as an engineering thermoplastics material. J. Appl. Polym. Sci. 91, 1657–1666.
- Zhang, L., Nie, X., Ravcheev, D.A., Rodionov, D.A., Sheng, J., Gu, Y., Yang, S., Jiang, W., Yang, C., 2014. Redox-responsive repressor rex modulates alcohol production and oxidative stress tolerance in *Clostridium acetobutylicum*. J. Bact. 196, 3949–3963.

RESEARCH ARTICLE

# A Universal *Mariner* Transposon System for Forward Genetic Studies in the Genus *Clostridium*

Ying Zhang\*, Alexander Grosse-Honebrink, Nigel P. Minton\*

Clostridia Research Group, BBSRC/EPSC Synthetic Biology Research Centre (SBRC), School of Life Sciences, University of Nottingham, Nottingham, United Kingdom

\* [ying.zhang@nottingham.ac.uk](mailto:ying.zhang@nottingham.ac.uk) (YZ); [nigel.minton@nottingham.ac.uk](mailto:nigel.minton@nottingham.ac.uk) (NPM)



 OPEN ACCESS

**Citation:** Zhang Y, Grosse-Honebrink A, Minton NP (2015) A Universal *Mariner* Transposon System for Forward Genetic Studies in the Genus *Clostridium*. PLoS ONE 10(4): e0122411. doi:10.1371/journal.pone.0122411

**Academic Editor:** Daniel Paredes-Sabja, Universidad Andres Bello, CHILE

**Received:** October 24, 2014

**Accepted:** February 20, 2015

**Published:** April 2, 2015

**Copyright:** © 2015 Zhang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC), UK and the European Union, through grant codes BB/I004475/1, BB/G016224/1 and PEOPLE-ITN-2008-237942. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** I have read the journal's policy and the authors of this manuscript have the following competing interests: The University of Nottingham

## Abstract

DNA transposons represent an essential tool in the armoury of the molecular microbiologist. We previously developed a *catP*-based mini transposon system for *Clostridium difficile* in which the expression of the transposase gene was dependent on a sigma factor unique to *C. difficile*, TcdR. Here we have shown that the host range of the transposon is easily extended through the rapid chromosomal insertion of the *tcdR* gene at the *pyrE* locus of the intended clostridial target using Allele-Coupled Exchange (ACE). To increase the effectiveness of the system, a novel replicon conditional for plasmid maintenance was developed, which no longer supports the effective retention of the transposon delivery vehicle in the presence of the inducer isopropyl β-D-1-thiogalactopyranoside (IPTG). As a consequence, those thiamphenicol resistant colonies that arise in clostridial recipients, following plating on agar medium supplemented with IPTG, are almost exclusively due to insertion of the mini transposon into the genome. The system has been exemplified in both *Clostridium acetobutylicum* and *Clostridium sporogenes*, where transposon insertion has been shown to be entirely random. Moreover, appropriate screening of both libraries resulted in the isolation of auxotrophic mutants as well as cells deficient in spore formation/germination. This strategy is capable of being implemented in any *Clostridium* species.

## Introduction

*Clostridium* is an ancient genus that comprises a large number of Gram-positive, rod-shaped bacteria. Obligate anaerobes, they are capable of producing endospores and are both of medical and industrial importance. The more notorious members of the genus are responsible for devastating infectious diseases (*Clostridium difficile*) and intoxications (*Clostridium botulinum*, *Clostridium perfringens*, and *Clostridium tetani*) of humans and animals [1,2]. The majority of its members are, however, non-pathogenic and are being investigated as the chassis for the sustainable production of chemicals and fuels, most notably solventogenic species, such as *Clostridium acetobutylicum*, *Clostridium beijerinckii*, *Clostridium thermocellum* and *Clostridium autoethanogenum* [3,4]. Yet other species, most notably *Clostridium sporogenes* and

has filed two patents (WO2013/144653 and WO2013/144647) on which Ying Zhang and Nigel Minton are coinventors which describes the use of the developed transposon system. Any net revenue after covering patent costs is shared equitably between the University and the inventors. This does not alter the authors' adherence to all PLOS ONE policies on sharing data and materials.

*Clostridium novyi*, have potential medical applications as the delivery vehicle of novel anti-tumour agents [5,6].

The growing importance of the genus has in recent times stimulated the development of genetic tools which may be deployed to both better understand clostridial biology and implement key process improvements. Accordingly, substantive progress has been made in the derivation of methods for directed gene knock-out and knock-in, including group II intron retargeting methods [7,8] as well as various procedures for undertaking allelic exchange using recombination-based strategies [9–14]. These methods may be deployed in classical reverse genetic approaches, where random mutants with the desired phenotype are isolated, and the nature of the genotype responsible is determined. The most useful tool in this context is a transposon. In this respect, *mariner*-based transposon mutagenesis systems offer the greatest utility [15,16], having been used in many different bacterial species, including low-GC-content Gram-positive species [17–21]. They insert into a 'TA' target site through a "cut-and-paste" mechanism [22], which makes them an ideal mutagen for clostridia.

To date, the successful use of a *mariner*-based mutagenesis system has been reported in *C. difficile* [15] and *C. perfringens* [16]. Both employed a similar plasmid-based strategy, in which the transposase gene was placed under the control of a conditional promoter which mitigated against expression of the transposase in the *E. coli* donor. This strategy ensures stability of the element in *E. coli*, both during the construction of the plasmid and prior to its transfer to the clostridial recipient. In both cases, the transposase was supplied in trans, and mediated the transfer of a mini-transposon element in a *catP* gene encoding resistance to chloramphenicol (Cm) / thiamphenicol (Tm), flanked by the two ITRs.

In the system deployed in *C. difficile* R20291, expression of the *Himar1 C9* transposase gene was under the control of the promoter of the *C. difficile* toxin B gene, *tcdB*. This promoter ( $P_{tcdB}$ ) is exclusively recognised by a specialised class of sigma factor, TcdR, which belongs to a family that is unique to toxinogenic clostridial species [15]. The utility of the system for forward genetic studies was demonstrated through the isolation of mutants defective in germination, (*cspBA*), and uracil metabolism, (*pyrB*). One drawback of the method was that whilst the delivery vehicle employed was segregationally unstable pseudo-suicide vector, it required a minimum of two passages of the recipient bacteria to eradicate the *Himar1 C9* transposase encoding plasmid. This limits the systems utility in high-throughput mutagenesis strategies. In the case of the system developed for *C. perfringens* strain 13, a counter-selection marker (the *galK* gene) was incorporated into the backbone of the transposon delivery vector and used in a host (strain HN13) in which the *galK* (galactokinase) and *galT* (Gal-1-phosphate uridylyltransferase) genes were deleted [16]. In the presence of 3% galactose, cells carrying a functional *galK* gene are unable to grow. Thus, by plating cells carrying the delivery vehicle on Tm-containing agar media supplemented with galactose, only transposon insertions that have lost the plasmid are able to grow. The system was successfully used to undertake a large scale screen for mutants lacking in gliding motility on agar plates. Whilst this experiment demonstrated the utility of this approach, the requirement for specific mutants lacking functional *galK* and *galT* genes is not ideal.

In the present study we have improved the utility of the system through the derivation of a transposon delivery vehicle that is conditional for replication, providing the facility for rapid plasmid loss after transfer into the clostridial host. Moreover, we have devised a simple strategy whereby the conditional expression of the transposase in the target clostridial host can be effected through the genomic incorporation of the *tcdR* gene. The system has been exemplified in two different clostridial species, *C. acetobutylicum* and *C. sporogenes*, and is potentially universally applicable to many different *Clostridium* species.

## Materials and Methods

### Bacterial strains and media

Bacterial strains utilised in this study are listed in [Table 1](#). *E. coli* strains were grown in Luria-Bertani medium at 37°C. *Clostridium spp* were cultured under anaerobic condition in an anaerobic cabinet (MG1000 Anaerobic Work Station, Don Whitley Scientific Ltd) containing an atmosphere of 80% nitrogen, 10% hydrogen and 10% carbon dioxide. Antibiotics were used at the following concentrations: erythromycin (Em), 20 µg/ml, thiamphenicol (Tm), 15 µg/ml for *Clostridium spp*; Em, 500 µg/ml, chloramphenicol (Cm), 25 µg/ml, tetracycline (Tc), 10 µg/ml for *E. coli*.

*C. acetobutylicum* ATCC 824 and its mutant derivatives were grown in Clostridial Growth Medium (CGM) [23] for routine manipulations, or Supplemented Clostridium Basal Medium (CBMS) for fermentation assays. CBMS was based on Clostridium Basal Medium (CBM) as previously described [24] with two modifications. Glucose was added at a final concentration of 50 g/l, and 5 g/l calcium carbonate was added as a buffering agent. P2 medium [25] with 20 g/l glucose was used to screen for auxotrophic mutants.

*C. sporogenes* NCIMB 10956 and its mutant derivatives were grown in TYG medium (2% trypticase, 0.5% peptone, 0.1% glucose, 0.5% yeast extract and 0.1% cysteine-HCl) [26] for routine manipulations, or P2Y medium for fermentation assays. P2Y medium is P2 medium with some modifications, in which glucose was added at a final concentration of 50 g/l, 5 g/l calcium carbonate was added as a buffering agent and 2 g/l yeast extract was added for extra amino acids. MACC medium [27] with 20 g/l glucose was used to screen for auxotrophic mutants.

### Plasmids, primers, DNA techniques

Plasmids and primers used in this study are listed in [Tables 1](#) and [2](#), respectively. Chromosomal DNA preparation, plasmid isolation and purification of DNA fragments from agarose gels were carried out using the DNeasy Tissue kit, the QIAprep Miniprep kit and the QIAquick Gel Extraction kit, respectively (Qiagen, UK). Restriction enzymes were supplied by New England Biolabs and were used according to the manufacturer's instructions. *E. coli* strains were transformed by electroporation using a Gene-Pulser (Bio-Rad), as recommended by the manufacturer. PCR amplifications were carried out using the KOD Hot Start Master Mix (Merck). Oligonucleotides used in this study are detailed in [Table 2](#), which were synthesised by Eurofins MWG Operon, Germany.

To establish whether transposition had occurred, inverse PCR (INV) was performed according to the procedure of Cartman and Minton. [15]. All DNA sequencing was carried out by Source BioScience, United Kingdom.

To identify the genomic location of transposon insertions, sequence data were analyzed using DNASTAR ([www.dnastar.com](http://www.dnastar.com)) and compared to the genome sequences using Artemis ([www.sanger.ac.uk/resources/software/artemis/](http://www.sanger.ac.uk/resources/software/artemis/)). The sequences of *C. acetobutylicum* ATCC 824 genome and megaplasmid pSOL1 are published (Refseq number NC\_003030.1 and NC\_001988.2; GenBank accession number AE001437 and AE001438) [28]. The GenBank accession number of *C. sporogenes* NCIMB 10956 is CP009225.

### Plasmid transfer in *Clostridium* species

*C. acetobutylicum* and *C. sporogenes* were transformed as described previously [29]. For *C. acetobutylicum*, prior to transformation, plasmid DNA was purified from *E. coli* TOP10 cells containing plasmid pAN2. This plasmid contains the  $\phi$ 3TI methyltransferase gene of *B. subtilis*



**Table 1. Bacterial strains and plasmids used in this study.**

Strain or Plasmid	Relevant characteristics <sup>a</sup>	Source <sup>b</sup>
<b>Bacterial Strains</b>		
<i>Clostridium acetobutylicum</i> ATCC 824	Wild type	ATCC
<i>Clostridium sporogenes</i> NCIMB 10696	Wild type	NCIMB
<i>E. coli</i> TOP10		Invitrogen
CRG3011	<i>C. acetobutylicum</i> ATCC 824 with <i>tcdR</i> inserted at <i>pyrE</i> locus	This study
CRG3817	<i>C. sporogenes</i> NCIMB 10696 with <i>tcdR</i> inserted at <i>pyrE</i> locus	This study
<b>Plasmids</b>		
pMTL83251	<i>Clostridium</i> modular plasmid used for construction of IPTG inducible promoter system. Em <sup>R</sup>	[32]
pMTL-ME6C	ACE vector for DNA integration at <i>pyrE</i> locus in <i>C. acetobutylicum</i> 824. Cm <sup>R</sup>	[10]
pMTL-YZ29	ACE vector for DNA integration at <i>pyrE</i> locus in <i>C. sporogenes</i> NCIMB 10696. Cm <sup>R</sup>	This study
pMTL-ME6C-tcdR	ACE plasmid for <i>tcdR</i> integration at <i>pyrE</i> locus in <i>C. acetobutylicum</i> 824. Cm <sup>R</sup>	This study
pMTL-YZ29-tcdR	ACE plasmid for <i>tcdR</i> integration at <i>pyrE</i> locus in <i>C. sporogenes</i> NCIMB 10696. Cm <sup>R</sup>	This study
pMTL83251-lacI	<i>Clostridium</i> modular plasmid with conditional replicon	This study
pMTL83251-lacI-T	<i>Clostridium</i> modular plasmid with conditional replicon and a <i>fdx</i> terminator	This study
pMTL87250	<i>Clostridium</i> modular plasmid with conditional replicon	This study
pMTL-YZ13	Transposon plasmid with pCB102 replicon	This study
pMTL-YZ14	Transposon plasmid with conditional replicon	This study
pMTL82254-PtcdB	<i>Clostridium</i> modular plasmid with <i>catP</i> reporter expressed by promoter <i>tcdB</i>	This study
pMTL82254-Pfdx	<i>Clostridium</i> modular plasmid with <i>catP</i> reporter expressed by promoter <i>fdx</i>	This study

<sup>a</sup> Cm<sup>R</sup>, chloramphenicol/thiamphenicol resistance gene; Em<sup>R</sup>, erythromycin resistance gene

<sup>b</sup> ATCC, American Type Culture Collection; NCIMB, The UK's National Collection of Industrial, Food and Marine Bacteria

doi:10.1371/journal.pone.0122411.t001

phage  $\phi$ 3tI, which protects DNA from *C. acetobutylicum* Cac824I DNA restriction activity [30].

### Construction of plasmids

Plasmid pMTL5401Fcat is a previously constructed plasmid [30] which mediates the IPTG inducible expression of the *catP* gene. To simplify and construct an IPTG inducible promoter

**Table 2. Oligonucleotide primers used in this study.**

	Sequence (5'-3')	Description
Del-traJ	AAAAAAGCTTATAATTATCCTTATTGGACTTTCAAGTGCGCCAGATAGGGTG	Delete traJ in pMTL5401Fcat
Del-traJ-antisense	CAGATTGTACAAATGTGGTGATAACAGATAAGTCTTTCAATTTAACTTACCTTTCTTTGT	Delete traJ in pMTL5401Fcat
YZ4	TGAACGCAAGTTTCTAATTTTCGATTTCCAATCGATAGAGGAAAGTGCT	Amplify IPTG cassette
YZ35	GCTGATATGGTAATGAAGGG	Amplify IPTG cassette
NotI/Ascl	AACAGCTATGACCGCGCGCCGCTCACTGCCCGC	Change NotI site to Ascl in pMTL83251-lacI
NotI/Ascl-antisense	GCGGGCAGTGAGCGCGCGCCGGTCATAGCTGTT	Change NotI site to Ascl in pMTL83251-lacI
tcdR-F	GCTATATCAAGTGCTAAAGGTC	Amplify tcdR gene
tcdR-R	AAAAAAGCTTATAATTATCCTTATGCGGTCCAAGACGTGCGCCAGATAGGGTG	Amplify tcdR gene
catP-INV-F1	CAGATTGTACAAATGTGGTGATAACAGATAAGTCCAAGACTCTAACTTACCTTTCTTTGT	Inverse PCR [15]
catP-INV-R1	TGAACGCAAGTTTCTAATTTTCGATTTACCGCTCGATAGAGGAAAGTGCT	Inverse PCR [15]
catP-INV-R2	GGATCTGTAATAACCCATAAAG	Sequencing of inverse PCR products [15]

doi:10.1371/journal.pone.0122411.t002

cassette, a 982 bp DNA intervening region encompassing *oriT/traJ* was deleted from plasmid pMTL5401Fcat using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) with primers Del-traJ and Del-traJ-antisense. Having deleted the *traJ* region, the 2728 bp *lac* inducible promoter cassette containing *lacI*,  $P_{ptb}$ ,  $P_{fac}$  and *cat* was then PCR amplified using primers YZ4 and YZ5 and cloned into pMTL83251 via NotI/HindIII, resulting in plasmid pMTL83251-YZ2 (Fig 1A).

Having established the functionality of the  $P_{fac}$  inducible promoter, the *cat* gene was deleted from plasmid pMTL83251-YZ2, in order to bring the pCB102 replicon [31] under the transcriptional control of  $P_{fac}$ . Two plasmids were created. In the one (pMTL83251-lacI-T see Fig 2A), pMTL83251-YZ2 was cleaved with BamHI and HindIII, the sticky-ends created blunt-ended by treatment with T4 polymerase, and the resultant linear fragment subjected to self-ligation. In a second plasmid (pMTL83251-lacI, see Fig 2B), pMTL83251-YZ2 was digested with BamHI and AscI, the sticky-ends created blunt-ended by treatment with T4 polymerase, and the resultant linear fragment subjected to self-ligation.

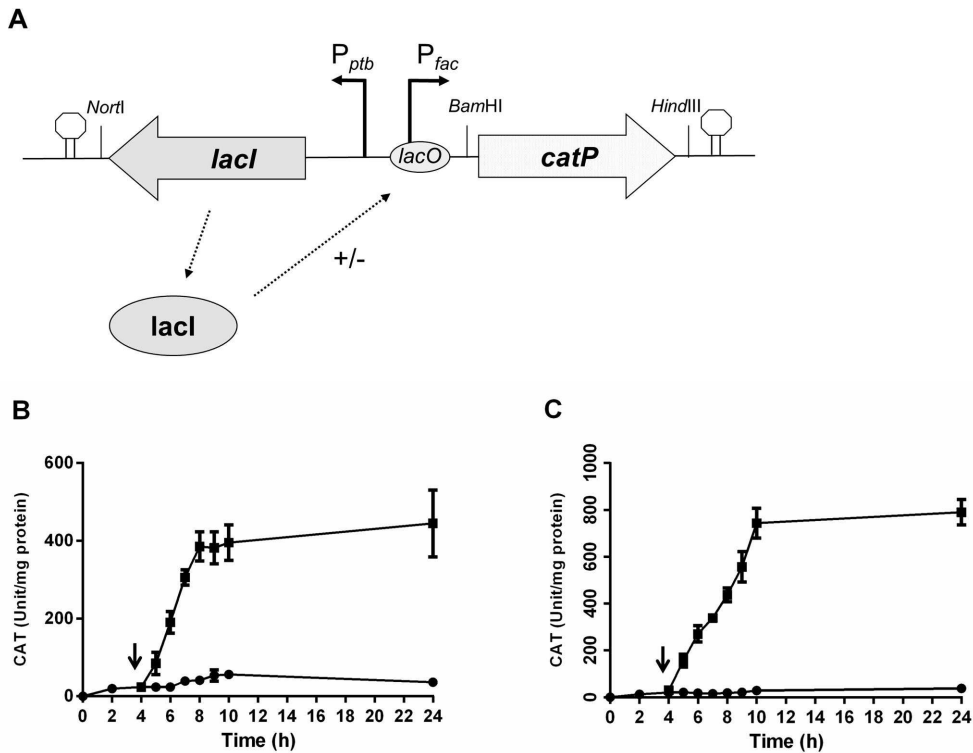
To confine the conditional replicon to a cassette compatible with the pMTL80000 modular vectors series [32], the DNA fragment encompassing the  $P_{fac}/P_{ptb}::lacI$  cassette, plus the pCB102 replicon, was localised to a portable AscI-FseI fragment. This was accomplished by the following steps. In the first instance, the NotI site preceding the  $P_{fac}/P_{ptb}::lacI$  cassette of pMTL83251-lacI was changed to an AscI site using a QuikChange Site-Directed Mutagenesis Kit (Stratagene) and primers NotI/AscI and NotI/AscI-antisense, to yield the plasmid pMTL87250. This now allowed the excision of the  $P_{fac}/P_{ptb}::lacI$  cassette, together with the pCB102 replicon, as a 3537 bp AscI and FseI fragment. This was substituted in place of the equivalent restriction fragment encoding the pBP1 replicon in plasmid pMTL-SC1. The plasmid obtained was designated pMTL-YZ14. In parallel, pMTL-YZ13 was made replacing the pBP1 replicon in pMTL-SC1 with replicon pCB102.

Plasmid pMTL82254- $P_{tcdB}$  as shown in Fig 3A, was constructed by excising a ca. 334 bp NotI/NdeI fragment from the plasmid pMTL-SC1 [15] and inserting it between the NotI and NdeI sites of plasmid pMTL82254 [32]. For comparative purposes, a second plasmid was constructed identical to pMTL82254- $P_{tcdB}$ , but in which the fragment encompassing the  $P_{tcdB}$  promoter was replaced with an equivalent NotI/NdeI fragment encompassing the  $P_{fdx}$  promoter. This plasmid was designated pMTL82254- $P_{fdx}$  as illustrated in Fig 3B.

### Integration of the *tcdR* gene into the clostridial host genome using ACE

A DNA fragment encompassing the *tcdR* structural gene encoding TcdR and its ribosome binding (RBS) site was PCR-amplified from the chromosome of the *C. difficile* strain 630 with primers tcdR-F and tcdR-R, and cloned into the vector pGEM-T. This localised the *tcdR* gene and RBS to a NotI-BamHI fragment which was excised and inserted between the equivalent sites of the plasmid pMTL-ME6C to yield the plasmid pMTL-ME6C-tcdR. Plasmid pMTL-ME6C is essentially plasmid pMTL-JH14, but which lacks the *lacZ'*- encoding region residing between the shorter Left Homology Arm (LHA, encoding a *pyrE* allele) and the longer Right Homology Arm (RHA, encoding CAC0028) and in which a region of DNA encompassing the transcriptional terminator of the *Clostridium pasteurianum* ferredoxin gene has been inserted 3' to the *pyrE* allele, and preceding CAC0028.

To introduce the *tcdR* gene into the chromosome of *C. acetobutylicum*, Allele-Coupled Exchange (ACE) was employed as described in Heap et al. [10]. *In vivo* methylated pMTL-ME6C-tcdR plasmid DNA was transformed into the ACE-generated *C. acetobutylicum pyrE* mutant essentially as described in Heap JT et al., 2012. Transformed cells were plated onto CGM agar supplemented with 15  $\mu$ g/ml Tm and 20  $\mu$ g/ml uracil. After 24 h, fast growing



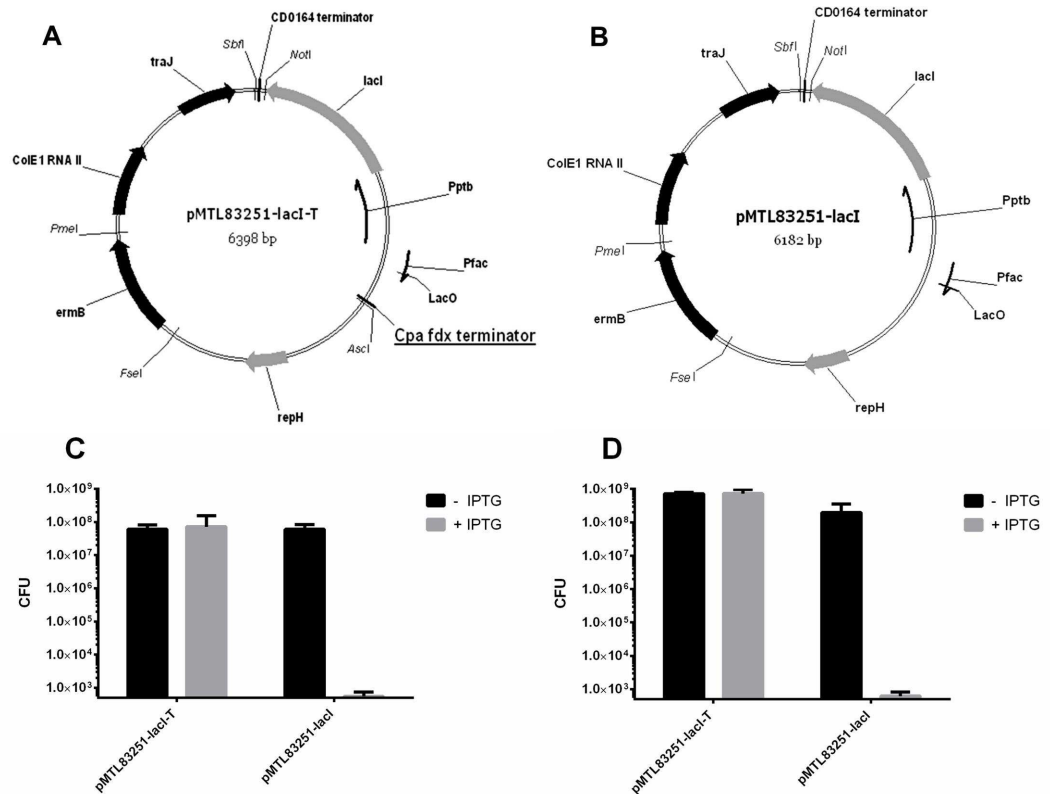
**Fig 1. Schematic diagram of the *lac*-based, IPTG inducible expression cassette in pMTL-YZ2 (A), and the demonstration of its function in *C. acetobutylicum* (B) and *C. sporogenes* (C).** A: Schematic diagram of the *lac*-based, IPTG inducible expression cassette. Key: *lacI* is the *lacI* repressor protein gene. *lacI* binds to the indicated *lacO* region, blocking transcription from the  $P_{lac}$  promoter. The  $P_{ptb}$  promoter (derived from the *C. beijerinckii* gene encoding phosphotransbutyrylase) directs the transcription of the *lacI* gene. B and C: IPTG induction of CAT production in *C. acetobutylicum* (B) and *C. sporogenes* (C) harbouring pMTL-YZ2. Circles equate to cells which received no IPTG, squares represents samples from cells that were induced with IPTG. The arrow indicates the time of adding IPTG. Activity is expressed as units of CAT activity per mg of soluble protein.

doi:10.1371/journal.pone.0122411.g001

single colonies were picked and re-streaked twice onto CGM agar containing 15 µg/ml Tm and 20 µg/ml uracil. These cells represented those in which the plasmid had integrated into the genome by single cross-over recombination between the RHA. Thereafter, cells were streaked onto CBM agar to select for cells able to grow in the absence of exogenous uracil as a consequence of plasmid excision (through recombination between the duplicated LHA) and restoration of a functional *pyrE* allele. The final construct has the *tcdR* gene, together with its RBS, inserted immediately downstream of the *pyrE* gene, residing upstream of the *Clostridium pasteurianum* transcriptional terminator. The *tcdR* gene is transcribed from the promoter upstream of CAC0025 (as demonstrated in Fig 4A). The *C. acetobutylicum* strain generated was designated CRG3011.

To generate a strain producing functional TcdR protein in *C. sporogenes* NCIMB 10696, the following steps were undertaken. A DNA fragment encoding TcdR and its ribosome binding

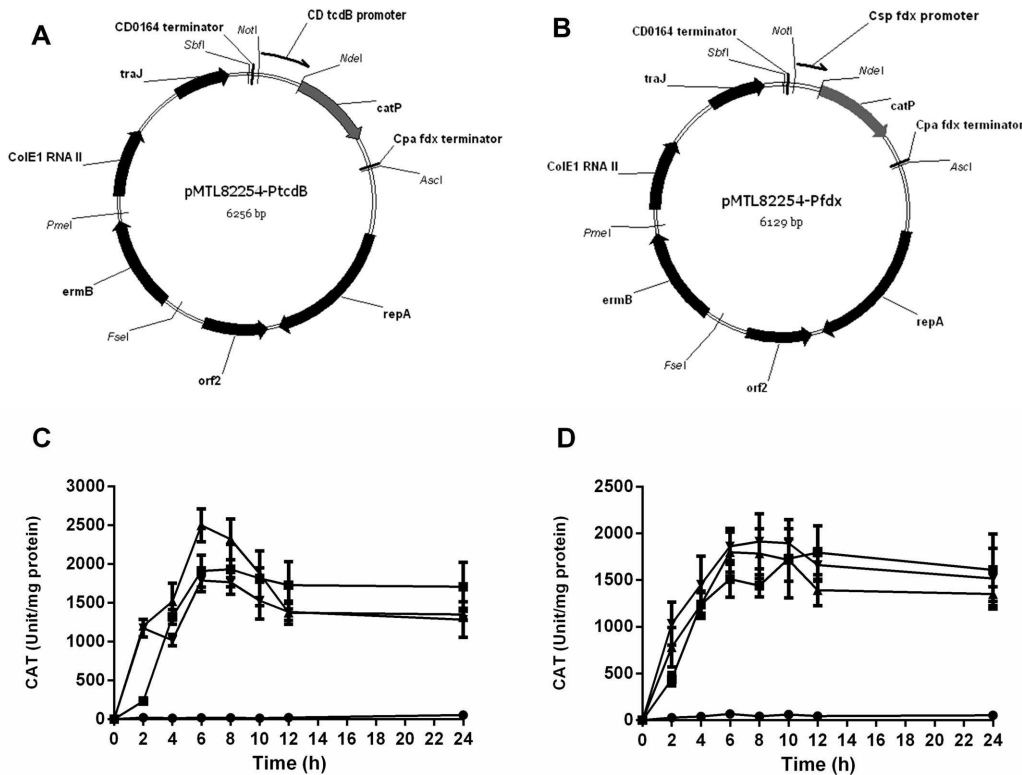




**Fig 2. Plasmid maps and the conditionality analysis of the conditional replicon.** Plasmid maps of the non-conditional control plasmid pMTL83251-lacI-T (A) and conditional plasmid pMTL83251-lacI (B). Key: CD0164 terminator, a transcriptional terminator isolated from downstream of the *Clostridium difficile* strain 630 CD0164 gene; *lacI*, the *E. coli* gene encoding LacI repressor;  $P_{ptb}$ , the promoter of the *C. beijerinckii* gene encoding phosphotransbutyrylase;  $P_{tac}$ , the promoter of the *C. pasteurianum* ferredoxin gene derivatised to include an *E. coli lac* operator; *repH*, replication region of the *Clostridium butyricum* plasmid pCB102; *ermB*, the macrolide-lincosamide-streptogramin B antibiotic resistance gene of plasmid pAMB31; ColE1, the replication origin of plasmid ColE1, and; *traJ*, transfer function of the RP4 *oriT* region. *Cpa fdx* terminator, transcriptional terminator of the ferredoxin gene of *C. pasteurianum* (this feature is underlined due to its existence only in A: pMTL83251-lacI-T, but not in B: pMTL83251-lacI). C and D: the ability to replicate of pMTL83251-lacI-T and pMTL83251-lacI in of *C. acetobutylicum* (C) and *C. sporogenes* (D) in the absence (CFU account in black) and presence (CFU account in grey) of IPTG.

doi:10.1371/journal.pone.0122411.g002

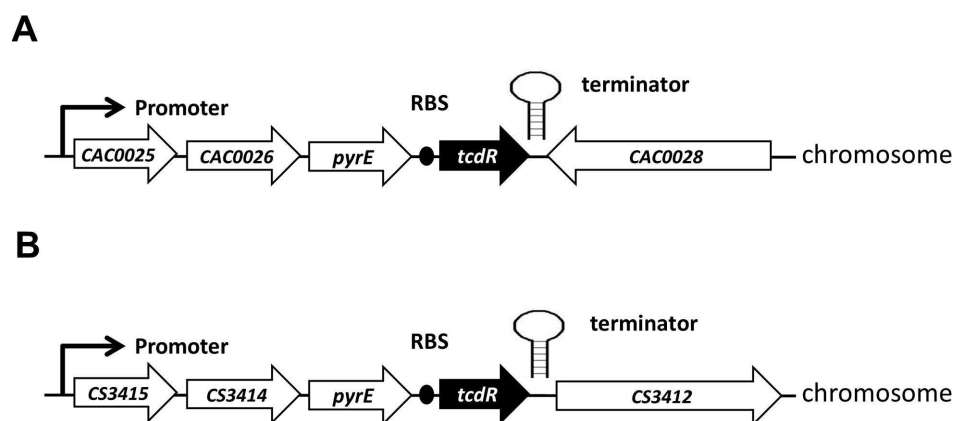
(RBS) site was excised as a NotI-BamHI fragment from plasmid pMTL-ME6C-tcdR, and inserted between the equivalent sites of the plasmid pMTL-YZ29 to yield the plasmid pMTL-YZ29-tcdR. ACE vector pMTL-YZ29 is as same as pMTL-JH29 with the Gram-positive replicon being changed to pCD6 instead of pIM13. Plasmid pMTL-YZ29-tcdR plasmid DNA was transformed into the ACE-generated *C. sporogenes pyrE* mutant described in Heap et al. [10]. Transformed cells were plated onto TYG agar supplemented with 15 µg/ml Tm. After 24 h, fast growing single colonies were picked and re-streaked twice onto TYG agar containing 15 µg/ml Tm. These cells represented those in which the plasmid had integrated into the genome by single cross-over recombination between the RHA. Thereafter, cells were streaked



**Fig 3. Analysis of the functionality of the exogenous TcdR in two clostridial strains (CRG3011 and CRG3817).** Plasmid maps of the pMTL82254-P<sub>tcdB</sub> (A) and pMTL82254-P<sub>fdx</sub> (B). Key: CD0164 terminator, a transcriptional terminator isolated from downstream of the *C. difficile* strain 630 CD0164 gene; *catP*, a *C. perfringens*-derived gene encoding chloramphenicol acetyltransferase; *Cpa fdx* terminator, transcriptional terminator of the ferredoxin gene of *C. pasteurianum*; *repA* and *orf2*, replication region of the *C. botulinum* plasmid pBP1; *ermB*, the macrolide-lincosamide-streptogramin B antibiotic resistance gene of plasmid pAMB1; *ColE1*, the replication origin of plasmid *ColE1*, and; *traJ*, transfer function of the RP4 *oriT* region. *CD tcdB* promoter, the promoter region of the *C. difficile tcdB* gene; *Csp fdx* promoter: the promoter region of the *C. sporogenes fdx* gene. (C): CAT activity of either *C. acetobutylicum* ATCC 824 wild type or CRG3011 (*tcdR* containing *C. acetobutylicum* ATCC 824) carrying plasmids pMTL82254-P<sub>tcdB</sub> and pMTL82254-P<sub>fdx</sub>. (D): CAT activity of either *C. sporogenes* NCIMB 10969 wild type or CRG3817 (*tcdR* containing *C. sporogenes* NCIMB 10969) carrying plasmids pMTL82254-P<sub>tcdB</sub> and pMTL82254-P<sub>fdx</sub>. Black circles ●, wild type with pMTL82254-P<sub>tcdB</sub>; black squares ■, wild type with pMTL82254-P<sub>fdx</sub>; black triangles ▲, CRG3011/CRG3817 with pMTL82254-P<sub>tcdB</sub>; black triangles ▼, CRG3011/CRG3817 with pMTL82254-P<sub>fdx</sub>.

doi:10.1371/journal.pone.0122411.g003

onto MACC minimal medium agar to select for cells able to grow in the absence of exogenous uracil as a consequence of plasmid excision (through recombination between the duplicated LHA) and restoration of a functional *pyrE* allele. The final construct has the *tcdR* gene, together with its RBS, inserted immediately downstream of the *pyrE* gene from where it is transcribed from the promoter upstream of CS3415 (Fig 4B). The strain generated was designated CRG3817.



**Fig 4. Schematic representations of the genome of *tcdR* strain of *C. acetobutylicum* (A) and *C. sporogenes* (B).** A promoter-less copy of the *tcdR* gene (including its ribosome binding site, RBS) of *C. difficile* strain 630 has been inserted into the genome using ACE technology immediately downstream of the *pyrE* gene (A: CAC0027, B: CS3413). Illustrated are the surrounding genes, and the position of the promoter responsible for expression of *tcdR*, and the position of the *tcdR* RBS. The illustrated terminator is the T1 transcriptional terminator of the *C. pasteurianum* ferredoxin gene.

doi:10.1371/journal.pone.0122411.g004

### Chloramphenicol acetyltransferase (CAT) assay

CAT activity was determined according to the method of Shaw. [33]. A quartz cuvette was prepared containing 540  $\mu$ l of 100 mM Tris buffer (pH 7.8), 200  $\mu$ l of 2.5 mM DTNB (5,5'-dithio-bis-2-nitrobenzoic acid) solution in 100 mM Tris buffer (pH 7.8), 200  $\mu$ l of 5.0 mM freshly prepared acetyl Coenzyme A solution in deionized water and 10  $\mu$ l of cell lysate. The cuvette was pre-warmed to 25°C, and the reaction initiated by adding 10  $\mu$ l of 0.3% w/v Cm solution in deionized water. The initial rate of increase of absorption at 412 nm was measured using an Analytik Jena SPECORD 250 PLUS spectrophotometer.

### Fermentation

For *C. acetobutylicum*, batch fermentation experiments were carried out in 100 ml volumes of CBMS broth in 250 ml conical flasks. For *C. sporogenes*, batch fermentation experiments were carried out in 100 ml volumes of P2Y broth in 250 ml conical flasks. Cultures were inoculated to a starting OD<sub>600</sub> of approximately 0.05 with an overnight starter culture, and incubated in an anaerobic cabinet at 37°C for 96 h. At each sample point, 1 ml samples were removed, placed on ice and centrifuged at 16,000 x g for 1.5 minutes. Supernatants were stored at -80°C before analysis by high performance liquid chromatography. OD<sub>600</sub> was also measured at each time point. Samples subjected to heat shock (80°C, 10 min) were taken at time points 11, 24, 48, 72 and 96 h for sporulation assay.

### Analytical techniques

The supernatant samples described above were subjected to HPLC analysis using a Varian Prostar HPLC (model 240) equipped with an Aminex HPX-87H Ion Exclusion Column Column (300 x 7.8 mm Biorad), a refractive index detector (Varian 356LC) and a photo diode array (UV) detector (Varian). The mobile phase used was 0.005 M H<sub>2</sub>SO<sub>4</sub> and the column was operated at 30°C with a flow rate of 0.6 ml/min. The injection volume was 20  $\mu$ l and elution

was isocratic. Supernatants were diluted appropriately in Milli-Q-water and filtered through a 0.2  $\mu\text{m}$  pore size membrane filter prior to analysis.

### Isolation of transposon mutants

For *C. acetobutylicum*, plasmid pMTL-YZ14 was *in vivo* methylated and transformed into *C. acetobutylicum* CRG3011 as described previously [30]. The transformed cells were selected on CGM agar containing Em and were incubated for 48 to 72 h in order to expand the population of cells containing the transposon plasmid. All growth was harvested by flushing the whole plate with CGM broth with 10% Glycerol, and serial dilutions were made and plated onto CGM agars in duplicate: CGM agars without any antibiotics to calculate transposition frequency, CGM agars containing Tm (15  $\mu\text{g}$  per ml) and IPTG (1 mM) to select for the transposon-based *catP* antibiotic marker and induce plasmid loss at the same time. Cell suspensions were stored at  $-80^{\circ}\text{C}$  for future use. Individual colonies, visible after 12 to 16 h, were picked and patch plated onto CGM agars containing Tm only and CGM agars containing Em, respectively, to check for plasmid loss and for further analysis and/or phenotypic screening.

For *C. sporogenes*, plasmid pMTL-YZ14 was transformed into *C. sporogenes* CRG3817 strain as described previously without methylation. The transformed cells were selected on TYG agar containing Em and were incubated for 24 to 48 h. The plating and selections were performed essentially as described for *C. acetobutylicum* except for the growth medium employed.

## Results

### Transcription can affect maintenance of pCB102 replicon-based plasmids

During the construction of the ClosTron plasmid pMTL007 [30], it was noted that one of the derivative plasmids, pMTL540F, was unable to transform *C. acetobutylicum*. In contrast, a previously constructed plasmid pMTL540FT, transformed the same clostridial host at normal frequencies [34]. The only difference between the two plasmids is the presence of a transcriptional terminator, positioned between the strong  $P_{\text{fac}}$  promoter and the pCB102 plasmid replicon. We hypothesised that transcription into the replication region was interfering with the ability of the plasmid to stably establish itself in a transformed cell. It follows, that were the transcription from  $P_{\text{fac}}$  to be placed under regulatory control, and repressed, then the plasmid should be successfully installed in the transformed cell. Moreover, by adding the requisite inducer to a transformed cell, then the activation of transcription into the replicon region should subsequently lead to plasmid loss. To test this system, an inducible promoter system was required.

### Modularisation of an IPTG inducible promoter cassette

We have previously demonstrated that  $P_{\text{fac}}$  controlled expression of a *catP* gene in *C. acetobutylicum* and *C. sporogenes* could be repressed by the incorporation of an *E. coli lacI* gene onto a plasmid backbone in which expression of the repressor gene was placed under the transcriptional control of the  $P_{\text{ptb}}$  promoter of the *C. beijerinckii ptb* (phosphotransbutyrylase) gene. Moreover, repression could be relieved by the addition of exogenous IPTG, resulting in production of *cat*-encoded chloramphenicol acetyltransferase (CAT). For convenience, we elected to localise the  $P_{\text{ptb}}::lacI$  element and  $P_{\text{fac}}$  promoter to a single portable restriction fragment. As the region encompassing these elements on the previously made plasmid, pMTL5401Fcat [30], were separated by a region of DNA encompassing *oriT*, we elected to delete this element during

this process—*oriT* is superfluous as it is not needed for clostridia that can be transformed, and is in any case already part of the modular vector series (31) that we later planned to use. Accordingly, we used QuikChange to delete *oriT*, converted the  $P_{\text{ptb}}::\text{lacI} + P_{\text{fac}} + \text{cat}$  region to a NotI/HindIII restriction fragment (Materials and Methods), and inserted it between the equivalent sites of plasmid pMTL83251 to yield plasmid pMTL-YZ2 (Fig 1A).

To confirm that expression of the *cat* gene remained under IPTG-inducible control, we transformed pMTL-YZ2 into *C. acetobutylicum* or *C. sporogenes*, and then grew the transformants in the presence or absence of added inducer. IPTG-mediated induction of CAT activity was observed in both organisms, but a significant basal level of expression was observed in *C. acetobutylicum* (Fig 1B) in the absence of IPTG, indicating the repression by LacI is less stringent than in *C. sporogenes*, where tight repression was evident (Fig 1C). This is consistent with a decrease in expression of *lacI* due to a reduced activity of the  $P_{\text{ptb}}$  promoter as cells entered the solventogenesis phase in *C. acetobutylicum* [35].

### Development of a conditional replicon

In order to test our hypothesis, that the establishment of a pCB102 replicon-based plasmid could be controlled by the IPTG inducible promoters system, we constructed two new plasmids, pMTL83251-lacI-T and pMTL83251-lacI (Fig 2A and 2B). The essential difference between the two plasmids is that pMTL83251-lacI-T carries a transcriptional terminator (that of the ferredoxin gene of *Clostridium pasteurianum*) between the  $P_{\text{fac}}$  promoter and the *repH* replication gene of the pCB102 replicon. Plasmid pMTL83251-lacI is broadly equivalent to pMTL-YZ2, except it lacks the *cat* gene.

Transformed cells of *C. acetobutylicum* ATCC 824 carrying either plasmid were cultivated and the effect of adding IPTG on plasmid retention estimated. Specifically, *C. acetobutylicum* transformants were cultured for 12 h in CGM broth supplemented with Em to select for the plasmid. The culture was then washed twice in PBS to remove the antibiotic and then used to inoculate fresh, CGM broth with and without IPTG (1mM) at 1% (vol/vol). After 12 h of growth, the culture was plated to enumerate Em<sup>R</sup> Colony Forming Units (CFU). As predicted, plasmid pMTL83251-lacI was found to only be stably maintained in the absence of IPTG (Fig 2C). In the presence of IPTG, the plasmid was rapidly lost from the cell population (Fig 2C), as evidenced by an almost complete loss of CFU on plates supplemented with Em. A similar loss was not evident in cells harbouring pMTL83251-lacI-T (Fig 2C). It was concluded that transcriptional read-through into the pCB102 replication region was interfering with plasmid replication/ maintenance. One explanation for this phenomenon might have been that the RepH protein was being overexpressed, and that over production of RepH was detrimental to the cell. However, no difference in growth rate was observed between cells carrying pMTL83251-lacI cultivated in medium in the presence or absence of IPTG (data not shown). To test whether a similar phenomenon was observed in a different clostridial species, the equivalent experiment was replicated in *C. sporogenes* NCIMB 10696, in this case using TYG broth. A similar result to that seen in *C. acetobutylicum* was obtained. Thus, in the presence of IPTG, plasmid pMTL83251-lacI was rapidly lost from the population (Fig 2D), whereas retention of plasmid pMTL83251-lacI-T was unaffected by addition of the inducer (Fig 2D). These data indicate that the system has broader application than just a single *Clostridium* species.

### Endowment of the genome with *tcdR*

Having derived a prototype conditional delivery vehicle for use in *C. acetobutylicum* and *C. sporogenes* we sought to implement our *mariner* transposon system in these clostridia by endowing them with a genomic copy of the *C. difficile* *tcdR* gene. The presence of TcdR in either

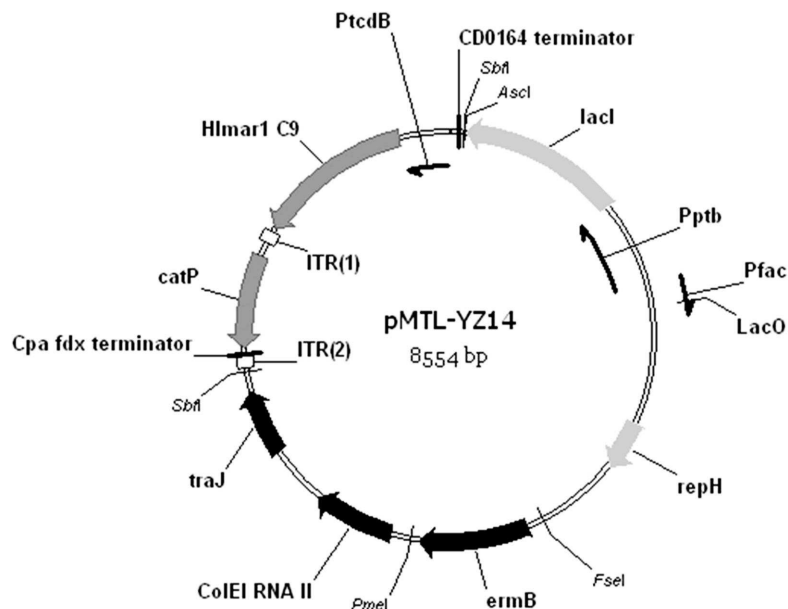
clostridia would direct the expression of the transposase gene under the control of the  $P_{tcdB}$  promoter when introduced on a plasmid. Accordingly, a promoter-less copy of the *tcdR* gene was cloned into the requisite ACE vectors needed to both convert the *pyrE* allele of the clostridial *pyrE* mutant hosts used back to wild type using the ACE protocol (Materials and Methods), while simultaneously inserting the *tcdR* gene into the genome immediately downstream of the restored *pyrE* gene (Fig 4A). This placed the *tcdR* gene under the control of the promoter responsible for *pyrE* transcription. The strain of *C. acetobutylicum* carrying *tcdR* in its genome was designated CRG3011, whereas the equivalent strain of *C. sporogenes* was designated CRG3817 (Fig 4B).

To test the functionality of TcdR in the two clostridial strains (CRG3011 and CRG3817), the  $P_{tcdB}$  or the  $P_{fdx}$  promoter were inserted upstream of the promoter-less *catP* gene of the reporter plasmid pMTL82254, to yield the respective plasmids pMTL82254- $P_{tcdB}$  and pMTL82254- $P_{fdx}$  (Fig 3A and 3B). Each plasmid was introduced into both the wild type strains of *C. acetobutylicum* and *C. sporogenes*, and their respective derivatives (CRG3011 and CRG3817) carrying *tcdR*. Subsequently, the four cell lines were cultivated in CGM or TYG broth, cells harvested at designated time points and lysates prepared for assays of CAT activity (Fig 3C and 3D). In both organisms it was evident from the low CAT activity observed in the wild type cell lines that the  $P_{tcdB}$  promoter was essentially inactive in the absence of TcdR. In contrast the level of CAT expression seen in CRG3011 and CRG3817 cells carrying pMTL82254- $P_{tcdB}$  was broadly equivalent to that seen in cells carrying plasmid pMTL82254- $P_{fdx}$ , where expression was due to the strong  $P_{fdx}$  promoter.

It was of interest to determine whether the presence of *tcdR* in the genome of either organism had any discernible effect. It was evident that in either case, there were no detectable alterations to either clostridial species in terms of appearance under the microscope, colony morphology, sensitivity to heat shock, growth rate in liquid media or the solvent/acid supernatant profiles observed after four days growth of the cells (S1, S2 and S3 Figs).

### Use of the conditional vector for transposon delivery

By endowing the two clostridial hosts with *tcdR*, effective transposition of the mini-transposon from plasmid pMTL-SC1 should now occur. Before testing this assumption, the plasmid was first modified to incorporate the pCB102 conditional replicon in place of the pBP1 replication region. This was accomplished by substituting the modular pBP1 replicon, excised as an *AscI* and *FseI* restriction fragment, with an equivalent fragment encompassing the modular, conditional pCB102 replicon of pMTL87250. The plasmid generated was designated pMTL-YZ14 (Fig 5). To test its effectiveness as a transposon delivery system, the plasmid was transformed into the *tcdR*-containing strains of *C. acetobutylicum* and *C. sporogenes*, CRG3011 and CRG3817. Transfer of the plasmid was initially selected on solidified media supplemented with Em and left on plates for 24 (CRG3817) to 72 (CRG3011) hours, before being plated on media containing Tm and IPTG to select for transposon events and promote plasmid loss, respectively. After 12 to 16 h of incubation, Tm<sup>R</sup> colonies were visible at a frequency of  $2.6 (\pm 0.6) \times 10^{-4}$  and  $3.2 (\pm 0.5) \times 10^{-4}$  (calculated as the ratio of Tm<sup>R</sup> CFU to total CFU) in *C. acetobutylicum* and *C. sporogenes*, respectively. A total of 100 Tm<sup>R</sup> colonies of each *Clostridium* species were picked and patch plated onto appropriate solidified rich media containing either Tm or Em to estimate the percentage of cells that had lost the plasmid. In total, 80% of *C. acetobutylicum* and 100% of *C. sporogenes* colonies were Tm resistant (<sup>R</sup>) and Em sensitive (<sup>S</sup>), indicative of successful insertion of the *catP* mini-transposon into chromosome and substantive plasmid loss.



**Fig 5. Vector map of plasmid pMTL-YZ14.** Expression of the hyperactive *mariner* transposase gene *Himar1 C9* was driven by the *C. difficile* toxin B promoter,  $P_{tcdB}$ . The plasmid backbone consisted of the conditional replicon between restriction sites *AscI* and *FseI*, the macrolide-lincosamide-streptogramin B antibiotic resistance gene *ermB*, and the Gram-negative replicon, *ColE1*. The whole *mariner* element (i.e., transposase gene and *catP* mini-transposon) can be excised as a *SbfI* fragment. The control plasmid pMTL-YZ13 was identical, except that the Gram-positive replicon is the pCB102 replicon from *C. butyricum*. This plasmid conforms to the pMTL80000 modular system for *Clostridium* shuttle plasmids [32].

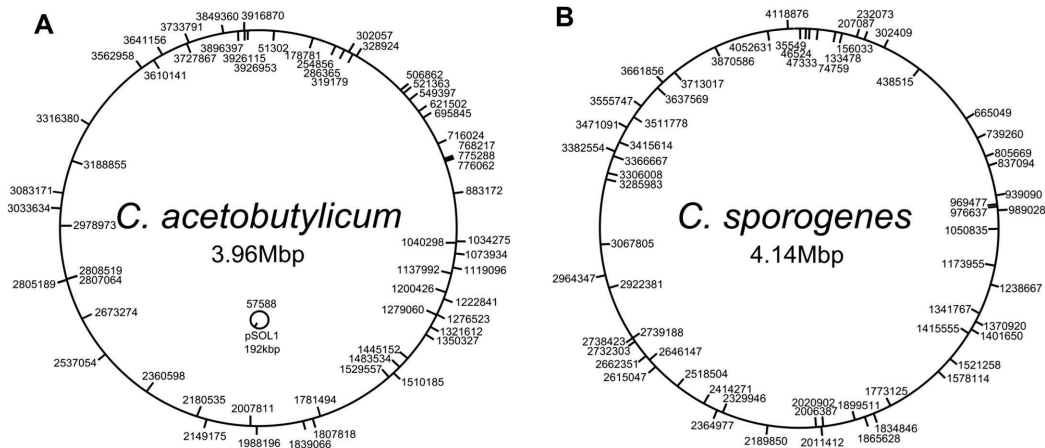
doi:10.1371/journal.pone.0122411.g005

To establish whether transposition had indeed occurred, a more thorough molecular characterisation was undertaken. Genomic DNA of 60 randomly selected  $Tm^R$  and  $Em^S$  clones was prepared, digested with *HindIII* and subjected to Inverse PCR after self-ligation. The gel-purified inverse PCR products were subjected to nucleotide sequence analysis using the primer *catP-INV-R2* to determine the sites of insertion. The data demonstrated that each transposon insertion had taken place at a different position around the genome as illustrated in Fig 6.

For *C. acetobutylicum* (Fig 6A), all of the  $Tm^R$  clones sequenced had a single transposon insertion, with the exception of one which had a double insertion revealed by Inverse PCR (S4 Fig). In one clone the mini-transposon had an insertion into the megaplasmid pSOL1. The remainder had all inserted at random around the chromosome. In total, there were 34 insertions in the plus strand and 27 in the minus strand, with 51 of the insertions located within open reading frames. In the case of *C. sporogenes* (Fig 6B), two out of the sixty clones contained double insertions of the mini-transposon. The remaining 58 clones all contained single insertions scattered at random around the genome. In total, there were 30 insertions in the plus strand and 32 in the minus strand. The number of insertions within open reading frames was 49, representing a slightly higher ratio of insertions into intergenic regions compared to *C. acetobutylicum*.

As a control, both clostridial strains were also transformed with pMTL-YZ13, which is identical to pMTL-YZ14 except that the pCB102 replicon is not preceded by the IPTG inducible





**Fig 6. Genetic map of *mariner* transposon insertions in *C. acetobutylicum* (A) and in *C. sporogenes* (B).** Sixty independent transposon mutants were identified and the insertions were sequenced. Insertions in the plus orientation are marked on the circle exterior. Insertions in the minus orientation are marked on the circle interior. Numbers indicate the precise point of insertion according to genome sequence data for (A) *C. acetobutylicum* ATCC 824 genome and megaplasmid pSOL1 (Refseq number NC\_003030.1 and NC\_001988.2; GenBank accession number AE001437 and AE001438) [28] and (B) *C. sporogenes* NCIMB 10956 (GenBank accession number CP009225).

doi:10.1371/journal.pone.0122411.g006

cassette ( $P_{pb::lacI::P_{fac}}$ ). In this case, following growth on plates supplemented with Em, cells were plated on CGM agar containing Tm, but no IPTG. A total of 24  $Tm^R$  colonies of each strain were picked and re-streaked 3 times, one more passage than described in Cartman and Minton. [15] Colonies were then patch plated onto solidified media with and without Em to estimate plasmid loss. Only 4 of clones had become  $Em^S$  in the case of *C. acetobutylicum* and just 3 out of 24 in the case of *C. sporogenes*. These data underlined the advantages of using the conditional pCB102 replicon.

### Isolation of phenotypic mutants affected in sporulation/ spore germination

To demonstrate the utility of the developed system for forward genetic studies, we screened representative transposon libraries of the two clostridial species for mutants with defects in the ability to form colony forming units (CFUs) after heat shock. These were presumed to be defective in either the process of sporulation or spore germination, i.e., *spo/ger* mutants. The number of  $Tm^R$  clones screened for *C. acetobutylicum* and *C. sporogenes* were approximately 3000 and 6000, respectively. Clones were inoculated into fresh CBM or YTG liquid broth (*C. acetobutylicum* and *C. sporogenes*, respectively) in replica 96-well microtitre plates and anaerobically incubated for one week to allow sporulation to take place. The microtitre plates were subjected to heat treatment (80°C for 10 min), while the other was held at room temperature. The contents of each microtitre well were individually plated onto either CGM (*C. acetobutylicum*) or YTG (*C. sporogenes*) agar media supplemented with Tm. Those cells unable to form colonies after heating 80°C for 10 min, but still able to produce colonies when not subjected to heat shock, were deemed possible sporulation or germination mutants.

In *C. acetobutylicum* a single transposon mutant was isolated with the desired *spo/ger* phenotype, whereas in *C. sporogenes*, two such mutants were obtained. Inverse PCR was carried



out on the isolated genomic DNA of the three mutants to identify the genes which had been inactivated (Table 3). In the case of *C. acetobutylicum*, the mini-transposon was found to have inserted in a gene (CAC2631) encoding a hypothetical protein that exhibited homology (70% similarity in DNA sequence) to the *Bacillus subtilis* spore protein YkuD reported to be involved in assembly of proteins on the forespore [36]. The two genes affecting the ability of *C. sporogenes* to form CFUs after heat shock were CLSPO\_c23320, a flavodoxin oxidoreductase, and CLSPO\_c31790, annotated as encoding the stage V sporulation protein SpoVAD. The *spoVA* operon in *B. subtilis* encodes six proteins, each of which has several putative membrane-spanning domains, and together they have been suggested to be involved in dipicolinic acid DPA transport into the developing forespore and the release of  $\text{Ca}^{2+}$ -DPA and other small molecules during spore germination [37,38]. The defect in this *C. sporogenes* gene is therefore consistent with the observed phenotype.

The phenotype associated with the mutation in CLSPO\_c23320 is less clear. Oxidoreductase is commonly related to aerotolerance in anaerobic organisms, providing protection against oxidative stress [39]. Since spores are important in the dissemination of the clostridial species it is likely that proteins with ability to scavenge the oxidative radicals are present in the outermost layers of the spores, and play a dual role—one in structural built-up of spores coat layers by mediating dityrosine cross-linking among proteins and/or second in resistance against oxidative stress [40,41] or they may have some other hitherto unknown function.

### Isolation of auxotrophic phenotypic mutants

To screen for auxotrophic mutants, the two libraries of  $\text{Tm}^R$  colonies were simply replica plated onto either solidified rich media (CGM for *C. acetobutylicum* and YTG for *C. sporogenes*) or an appropriate minimal medium agar medium (P2 for *C. acetobutylicum* and MACC for *C. sporogenes*), in all cases supplemented with  $\text{Tm}$ .

A single auxotrophic mutant of *C. acetobutylicum* was isolated that failed to grow on P2 agar medium after 7 days of incubation at 37°C. Nucleotide sequencing of isolated genomic DNA demonstrated that the min-transposon insertion had inserted into CAC0971 (CitB), which is annotated as encoding aconitase (aconitate hydratase). In aerobic bacteria this enzyme is responsible for the conversion of citrate to iso-citrate as part of the tricarboxylic acid (TCA) cycle. Iso-citrate is further converted into the TCA cycle intermediates oxalosuccinate and  $\alpha$ -ketoglutarate, and the later into the non-TCA cycle metabolite glutamate, which can subsequently be converted into arginine and glutamine. Whilst strict anaerobes are not generally regarded as being in the possession of a TCA cycle, two independent research groups have recently suggested that *C. acetobutylicum* possesses a complete, albeit bifurcated TCA cycle in which oxaloacetate flows to succinate both through citrate/ $\alpha$ -ketoglutarate and via malate/fumarate [42,43].

The gene coding for citrate synthase has been tentatively identified as CAC0970 and is the first gene of a putative tricistronic operon, comprising citrate synthase (CAC0970), aconitase (CAC0971) and isocitrate dehydrogenase (CAC0972), which is expressed at higher levels at stationary phase [43]. Accordingly, the mutant, together with *C. acetobutylicum* CRG3011, was plated onto P2 minimal agar medium, individually supplemented (1 mM) with the tricarboxylic and amino acids, citrate, iso-citrate,  $\alpha$ -ketoglutarate, glutamate, glutamine, proline and arginine (Table 4). Qualitative growth of the mutant was only fully restored to wild type by exogenous glutamate, whereas growth was partially restored by glutamine. Intriguingly, the mutant remained unable to grow on the minimal media when supplemented with  $\alpha$ -ketoglutarate, which is suggested to be the product of the conversion of iso-citrate by isocitrate dehydrogenase (CAC0971) and the substrate for the synthesis of glutamate [42]. Our results clearly

**Table 3. Transposon insertion sites.**

Strain	Phenotype	Transposon insertion site <sup>a</sup>	Location in genome (nt) <sup>b</sup>	ORF interrupted <sup>c</sup>	Description
<i>C. acetobutylicum</i> ATCC 824	spo/ger	ATACTAACTTGATATTATA—Tn— <b>TAAATATAACTTTCTTCTTT</b>	2740454	CAC2631	hypothetical protein(L,D-transpeptidase catalytic domain protein; YkuD (spore protein of <i>Bacillus subtilis</i> ))
<i>C. acetobutylicum</i> ATCC 824	Auxotroph	CTAATCATTTGCAAGAATA—Tn— <b>TACACAAGGCTAATCTAATC</b>	1119096	citB (CAC0971)	aconitate hydratase / dehydrogenase Catalyzes the conversion of citrate to isocitrate
<i>C. sporogenes</i> NCIMB 10696	spo/ger	ATTGCACTCTAATGGAAATA—Tn— <b>TAGAATAATATCTATAATAG</b>	2615047	CLSPO_c23320	flavodoxin oxidoreductase
<i>C. sporogenes</i> NCIMB 10696	spo/ger	CAGCGCATAGCATATCTATA—Tn— <b>TATCTGTATCCTTTAGATTA</b>	3471091	spoVAD (CLSPO_c31790)	stage V sporulation protein AD
<i>C. sporogenes</i> NCIMB 10696	Auxotroph	ACTGTAAGTTTTACTATGTA—Tn— <b>TAAGTACCAGGATCTTTATA</b>	1834846	colA (CLSPO_c16750)	microbial collagenase
<i>C. sporogenes</i> NCIMB 10696	Auxotroph	CTACTACTAAATTATTTCTA—Tn— <b>TATATCTACTAGTAACTGGA</b>	2364977	CLSPO_c21170	hypothetical protein
<i>C. sporogenes</i> NCIMB 10696	Auxotroph	ATTGTACTGCCATAGAAATA—Tn— <b>TAAATATAACTCTAAACTA</b>	3285983	CLSPO_c29860	ABC-2 family transporter protein

<sup>a</sup> The Tn insertion is indicated by dashes on either side, and the target site duplication is shown in boldface.

<sup>b</sup> nt, nucleotide.

<sup>c</sup> ORF, open reading frame.

doi:10.1371/journal.pone.0122411.t003

implicate CAC0971 in glutamate biosynthesis, but may suggest that glutamate is derived from a source other than  $\alpha$ -ketoglutarate, assuming that this ketone can be taken up by *C. acetobutylicum*. Whilst glutamine is most likely derived from glutamate, the fact that it could not entirely restore the growth of the *citB* mutant to wild type levels suggest that the reversal of the glutamate to glutamine reaction is not thermodynamically favourable under normal laboratory conditions.

As a proteolytic clostridia, *C. sporogenes* is not able to undertake the biosynthesis of as many amino acids as the saccharolytic *C. acetobutylicum*. It was for this reason that a larger library was screened. In total, three mutants were isolate which were compromised in their ability to grow on minimal media compared to the wild type control. The inactivated genes were CLSPO\_c16750, encoding a collagenase (annotated as ColA), CLSPO\_c21170, coding for a

**Table 4. Characterization of the found auxotroph mutant with different additives in P2 minimal medium.**

Supplementation	citB (CAC0972) mutant	<i>C. acetobutylicum</i> CRG3011
Citrate	--	++
Iso-citrate	--	++
$\alpha$ -Ketoglutarate	--	++
Glutamate	++	++
Glutamine	+	++
Proline	--	++
Arginine	--	++
Without supplementation	--	++

No growth(--), slight growth(+) and normal growth(++).

doi:10.1371/journal.pone.0122411.t004

hypothetical protein and CLSPO\_c29860, annotated as an ABC-2 family transporter protein. A BLAST search of the protein sequence of CLSPO\_c21170 revealed no putative conserved domains. The role of this protein in supporting growth on the minimal media employed, therefore, remains unknown. Similarly, the function of the protein encoded by CLSPO\_c29860 is unclear, although the annotated name of the ABC-2 family of transporter proteins indicates it is likely involved in the translocation of an as yet unknown substrate(s)/ nutrient(s) into the cell. In the case of ColA, whilst there is no direct evidence to suggest bacterial *colA* genes are essential for survival and growth in defined medium, it is interesting to note that some of the organisms shown to elaborate collagenolytic enzymes are asaccharolytic [44]. Since their metabolism is dependent on the uptake of small peptides and amino acids, the production of collagenolytic enzymes by these organisms may be essential for survival and growth. Closely related with *C. botulinum*, *C. sporogenes* is proteolytic and asaccharolytic, and extra yeast extract is known to be essential for growth in fermentation compared to other saccharolytic clostridia. Hence, ColA could play an important role for the uptake of small peptides and amino acids in *C. sporogenes*.

## Discussion

In the current study, we have implemented two fundamental refinements to an existing *mariner* transposon system that greatly increase its utility for use in *Clostridium* spp. On the one hand we have devised an improved transposon delivery vehicle that is conditional for plasmid maintenance. This significantly reduces the background of antibiotic resistant transposon colonies that still retain the plasmid after plating under the non-permissive condition. On the other hand, we have devised a strategy whereby the benefits of using the  $P_{\text{tcdB}}$  promoter, which does not function in the *E.coli* donor but is confined to the target *Clostridium* cell, can be implemented in potentially any *Clostridium* species. This is simply achieved through the introduction of the cognate sigma factor gene, *tcdR*, into the target host genome using the well established gene integration technology, ACE.

The non-permissive condition for the maintenance of the plasmid carrying the conditional pCB102-based replicon in the clostridial cell is the presence of IPTG. The absence of IPTG represents the permissive condition. The molecular basis of plasmid loss in the presence of IPTG is currently unknown. The fact that transcriptional terminators positioned between the strong  $P_{\text{tac}}$  promoter and the pCB102 replicon prevents plasmid loss clearly indicated that the cause of this loss is increased transcription of the pCB102 replicon region. The replicative mechanisms used by the pCB102 replicon is currently unknown, save that the minimal replication region carries a single putative ORF designed RepH. It seems unlikely, however, that this is a consequence of overproduced RepH protein causing a metabolic burden on cells carrying the plasmid, as no growth rate difference is evident in cells carrying the plasmid in the presence or absence of IPTG. Moreover, it is still not certain that RepH actually encodes a protein. The predicted amino acid sequence has no functional homologue in current databases, and its primary sequence is highly unusual in being composed of 7 Arg and 22 Lys residues as well as 10 Cys residues from a total of only 104 amino acids. It may be that the replicative mechanism is RNA based, and that overproduction of RNA components leads to a breakdown in the ability of the plasmid to maintain itself in the cell, either by impeding replication or causing some form of catastrophic instability. Further work would be required to clarify the situation. In the meantime, in the absence of cause, the effect can be used to our advantage.

It should be noted that this conditional system as it stands cannot be used in an organism in which the IPTG inducible systems does not function. Thus, while we have shown that the conditional replicon functions effectively in *C. beijerinckii* (G. Little, unpublished), *C. botulinum*

(C. Humphreys, unpublished) and *C. autoethanogenum* (C. woods, unpublished), it cannot be used in *C. difficile* (M. Fit, unpublished). The latter observation is because *C. difficile* apparently does not take up IPTG. However, it is clear that the IPTG inducible promoter system developed here can be substituted for with analogous inducible system shown to work in the target organism. Examples include the anhydrotetracycline inducible system exemplified in *C. difficile* [45], as well as the lactose inducible system of *C. perfringens* [46].

The second technical innovation described here is to capitalise on the conditional nature of the *C. difficile*  $P_{tcdB}$ , in respect of being active only in the target host, and not the *E. coli* donor. This desirable feature prevents transposon activity in *E. coli* prior to transfer of the vector. To researchers unfamiliar with ACE technology, transferring *tcdR* into the clostridial host may seem not seem at first glance a trivial undertaking. However, the ACE method is extremely efficient, requiring the initial selection of the required *pyrE* uracil auxotrophic mutant on the basis of resistance to fluoroorotic acid (FOA), and thereafter, the integration of the *tcdR* gene concomitant with restoration of the *pyrE* allele to wild type, selecting for growth on minimal media lacking uracil. The interconversion of wild type and mutant *pyrE* alleles has now been accomplished in our laboratory using ACE in eight different clostridial species to date (unpublished data). Our data also shows, that at least in *C. acetobutylicum* and *C. sporogenes*, the presence of the *tcdR* gene has no discernible effect on the organisms' phenotype. In any case, transposon mutagenesis should be very much viewed as a screening tool. Thus, it would be expected that once the preliminary phenotype of a specific mutant has been determined, that a definitive mutation would be made in a clean background lacking *tcdR* using a directed method, such as the Clostron or alternative allelic exchange procedures.

In the current study, the system was exemplified in *C. acetobutylicum* and *C. sporogenes*. Transposition into clostridial genome occurred at a frequency of  $2.6 (\pm 0.6) \times 10^{-4}$  and  $3.2 (\pm 0.5) \times 10^{-4}$  in *C. acetobutylicum* and *C. sporogenes* respectively, in a random fashion, with simultaneously plasmid loss at 80% and 100%, generating transposon mutants with just single insertion in an overwhelming majority of cases (98.3% and 96.7% in this study). Moreover, 51 of the 61 insertions sequenced (83.6%) in *C. acetobutylicum* and 49 of the 62 (79%) in *C. sporogenes* were located within protein coding sequences. This is within the range that would be expected for a random mutagen in *Clostridium*, considering that around 80% of the clostridial genome is protein coding.

We further demonstrated that the tool could be used to isolate specific mutants with a particular phenotype, by screening for clones affected in sporulation/germination as well as auxotrophic strains which could no longer grow on minimal media. Our intention was merely to exemplify that the method could be used for such a purpose, rather than gather fundamental information on the nature of the cellular components identified. Nevertheless, some preliminary conclusions can be made on the nature of the mutants identified. Thus, in the main, the sporulation/germination mutants obtained were in genes that encoded proteins associated with *B. subtilis* spores (spore protein YkuD and SpoVA). The reason for the absence of CFU after heat shock when the *C. sporogenes* CLSPO\_c23320 gene was inactivated (encoding a putative flavodoxin oxidoreductase) was less clear. Auxotrophic mutant screening has clearly implicated the *C. acetobutylicum* gene CAC0971 (encoding CitB, aconitase) in glutamate biosynthesis, and most likely glutamine, but has called in to question whether glutamate is formed from  $\alpha$ -ketoglutarate. In the case of *C. sporogenes*, the reason why inactivation of the three genes (CLSPO\_c16750, a collagenase; CLSPO\_c21170, a hypothetical protein, and; CLSPO\_c29860, a protein annotated as an ABC-2 family transporter protein) prevented growth on a minimal media were largely unclear, with the possible exception of the collagenase where a possible role in the uptake of small peptides and amino acids may exist.

## Conclusions

We have developed a novel *mariner* transposon system and exemplified it in *C. acetobutylicum* and *C. sporogenes*, where a number of auxotrophic and spore-related mutants were isolated. The system makes use of a conditional plasmid delivery vehicle, which may be deployed to significantly reduce the background of cells that still retain autonomous plasmids, as well as a conditional expression system that confines production of the *mariner* transposase to the target clostridial species, and not the *E. coli* donor. The system is able to generate completely random mutant libraries at high frequencies and should prove widely applicable to members of the genus *Clostridium*.

## Supporting Information

**S1 Fig. Sporulation profiles of (A) *C. acetobutylicum* wild type and CRG3011; (B) *C. sporogenes* wild type and CRG3817 over 96 hours.** ClosTron mutants Cac-spoOA::CTermB and Cspo-spoOA::CTermB were included as a negative control.  
(TIF)

**S2 Fig. Growth curves and fermentation profiles of *C. acetobutylicum* wild type (black circles) and CRG3011 (black squares).**  
(TIF)

**S3 Fig. Growth curves and fermentation profiles of *C. sporogenes* wild type (black circles) and CRG3817 (black squares).**  
(TIF)

**S4 Fig. Inverse PCR screens of nine random selected pMTL-YZ14 derived Tm<sup>R</sup> and Em<sup>S</sup> clones of CRG3011.** Genomic DNA prepared from each clone was screened for the transposon based insertion. Lane M, 1kb ladder (NEB); lane-, negative control (genomic DNA of CRG3011); lane 1–9, pMTL-YZ14 derived Tm<sup>R</sup> and Em<sup>S</sup> clones 1 to 9, clone 3 shows double transposon insertion while the other clones have single insertion.  
(TIF)

## Acknowledgments

We thank Drs Stephen T. Carman and John Heap for useful discussions.

## Author Contributions

Conceived and designed the experiments: NPM YZ. Performed the experiments: YZ AGH. Analyzed the data: NPM YZ AGH. Contributed reagents/materials/analysis tools: NPM YZ AGH. Wrote the paper: NPM YZ.

## References

1. Hatheway CL (1990) Toxigenic clostridia. *Clin Microbiol Rev* 3: 66–98. PMID: [2404569](#)
2. Brazier JS (2008) *Clostridium difficile*: from obscurity to superbug. *Br J Biomed Sci* 65: 39–44. PMID: [18476496](#)
3. Dürre P (2005) Handbook on clostridia. Boca Raton: Taylor & Francis. 903 p. p.
4. Papoutsakis ET (2008) Engineering solventogenic clostridia. *Curr Opin Biotechnol* 19: 420–429. doi: [10.1016/j.copbio.2008.08.003](#) PMID: [18760360](#)
5. Heap JT, Theys J, Ehsaan M, Kubiak AM, Dubois L, Paesmans K, et al. (2014) Spores of *Clostridium* engineered for clinical efficacy and safety cause regression and cure of tumors in vivo. *Oncotarget* 5: 1761–1769. PMID: [24732092](#)

6. Agrawal N, Bettgeowda C, Cheong I, Geschwind JF, Drake CG, Hipkiss EL, et al. (2004) Bacteriolytic therapy can generate a potent immune response against experimental tumors. *Proc Natl Acad Sci U S A* 101: 15172–15177. PMID: [15471990](#)
7. Heap JT, Cartman ST, Kuehne SA, Cooksley C, Minton NP (2010) Clostron-targeted mutagenesis. *Methods Mol Biol* 646: 165–182. doi: [10.1007/978-1-60327-365-7\\_11](#) PMID: [20597009](#)
8. Chen Y, Helmus R, McClane B, Hoffman R, Watkins S, Wehrli T, et al. (2004) Use of a *Clostridium perfringens* vector to express high levels of SIV p27 protein for the development of an oral SIV vaccine. *Virology* 329: 226–233. PMID: [15518803](#)
9. Cartman ST, Kelly ML, Heeg D, Heap JT, Minton NP (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. doi: [10.1128/AEM.00249-12](#) PMID: [22522680](#)
10. Heap JT, Ehsaan M, Cooksley CM, Ng YK, Cartman ST, Winzer K, et al. (2012) Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic Acids Res* 40: e59. doi: [10.1093/nar/gkr1321](#) PMID: [22259038](#)
11. Ng YK, Ehsaan M, Philip S, Collyer MM, Janoir C, Cartman ST, et al. (2013) Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using pyrE alleles. *PLoS One* 8: e56051. doi: [10.1371/journal.pone.0056051](#) PMID: [23405251](#)
12. Leang C, Ueki T, Nevin KP, Lovley DR (2013) A genetic system for *Clostridium ljungdahlii*: a chassis for autotrophic production of biocommodities and a model homoacetogen. *Appl Environ Microbiol* 79: 1102–1109. doi: [10.1128/AEM.02891-12](#) PMID: [23204413](#)
13. Al-Hinai MA, Fast AG, Papoutsakis ET (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: 8112–8121. doi: [10.1128/AEM.02214-12](#) PMID: [22983967](#)
14. Nariya H, Miyata S, Suzuki M, Tamai E, Okabe A (2011) Development and application of a method for counterselectable in-frame deletion in *Clostridium perfringens*. *Appl Environ Microbiol* 77: 1375–1382. doi: [10.1128/AEM.01572-10](#) PMID: [21183644](#)
15. Cartman ST, Minton NP (2010) A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile*. *Appl Environ Microbiol* 76: 1103–1109. doi: [10.1128/AEM.02525-09](#) PMID: [20023081](#)
16. Liu H, Bouillaut L, Sonenshein AL, Melville SB (2013) Use of a mariner-based transposon mutagenesis system to isolate *Clostridium perfringens* mutants deficient in gliding motility. *J Bacteriol* 195: 629–636. doi: [10.1128/JB.01288-12](#) PMID: [23204460](#)
17. Gao LY, Groger R, Cox JS, Beverley SM, Lawson EH, Brown EJ. (2003) Transposon mutagenesis of *Mycobacterium marinum* identifies a locus linking pigmentation and intracellular survival. *Infect Immun* 71: 922–929. PMID: [12540574](#)
18. Le Breton Y, Mohapatra NP, Haldenwang WG (2006) In vivo random mutagenesis of *Bacillus subtilis* by use of TnYLB-1, a mariner-based transposon. *Appl Environ Microbiol* 72: 327–333. PMID: [16391061](#)
19. Maier TM, Pechous R, Casey M, Zahrt TC, Frank DW (2006) In vivo *Himar1*-based transposon mutagenesis of *Francisella tularensis*. *Appl Environ Microbiol* 72: 1878–1885. PMID: [16517634](#)
20. Cao M, Bitar AP, Marquis H (2007) A mariner-based transposition system for *Listeria monocytogenes*. *Appl Environ Microbiol* 73: 2758–2761. PMID: [17308180](#)
21. Wilson AC, Perego M, Hoch JA (2007) New transposon delivery plasmids for insertional mutagenesis in *Bacillus anthracis*. *J Microbiol Methods* 71: 332–335. PMID: [17931726](#)
22. Muñoz-López M, García-Pérez JL (2010) DNA transposons: nature and applications in genomics. *Curr Genomics* 11: 115–128. doi: [10.2174/138920210790886871](#) PMID: [20885819](#)
23. Hartmanis MG, Gatenbeck S (1984) Intermediary Metabolism in *Clostridium acetobutylicum*: Levels of Enzymes Involved in the Formation of Acetate and Butyrate. *Appl Environ Microbiol* 47: 1277–1283. PMID: [16346566](#)
24. O'Brien RW, Morris JG (1971) Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *J Gen Microbiol* 68: 307–318. PMID: [4332793](#)
25. Baer SH, Blaschek HP, Smith TL (1987) Effect of Butanol Challenge and Temperature on Lipid Composition and Membrane Fluidity of Butanol-Tolerant *Clostridium acetobutylicum*. *Appl Environ Microbiol* 53: 2854–2861. PMID: [16347502](#)
26. Purdy D, O'Keefe TA, Elmore M, Herbert M, McLeod A, Bokori-Brown M, et al. (2002) Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. *Mol Microbiol* 46: 439–452. PMID: [12406220](#)
27. Lovitt RW, Morris JG, Kell DB (1987) The Growth and Nutrition of *Clostridium-Sporogenes* Ncib-8053 in Defined Media. *Journal of Applied Bacteriology* 62: 71–80. PMID: [3571034](#)

28. Nolling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, et al. (2001) Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol* 183: 4823–4838. PMID: [11466286](#)
29. Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET (1992) Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. *Biotechnology (N Y)* 10: 190–195. PMID: [1368230](#)
30. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP (2007) The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J Microbiol Methods* 70: 452–464. PMID: [17658189](#)
31. Minton N, Morris JG (1981) Isolation and Partial Characterization of 3 Cryptic Plasmids from Strains of *Clostridium-Butyricum*. *Journal of General Microbiology* 127: 325–331.
32. Heap JT, Pennington OJ, Cartman ST, Minton NP (2009) A modular system for *Clostridium* shuttle plasmids. *J Microbiol Methods* 78: 79–85. doi: [10.1016/j.mimet.2009.05.004](#) PMID: [19445976](#)
33. Shaw WV (1975) Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria. *Methods Enzymol* 43: 737–755. PMID: [1094240](#)
34. Fox ME, Lemmon MJ, Mauchline ML, Davis TO, Giaccia AJ, Minton NP, et al. (1996) Anaerobic bacteria as a delivery system for cancer gene therapy: in vitro activation of 5-fluorocytosine by genetically engineered clostridia. *Gene Ther* 3: 173–178. PMID: [8867865](#)
35. Feustel L, Nakotte S, Durre P (2004) Characterization and development of two reporter gene systems for *Clostridium acetobutylicum*. *Appl Environ Microbiol* 70: 798–803. PMID: [14766557](#)
36. Kodama T, Takamatsu H, Asai K, Ogasawara N, Sadaie Y, Watabe K. (2000) Synthesis and characterization of the spore proteins of *Bacillus subtilis* YdhD, YkuD, and YkvP, which carry a motif conserved among cell wall binding proteins. *J Biochem* 128: 655–663. PMID: [11011148](#)
37. Tovar-Rojo F, Chander M, Setlow B, Setlow P (2002) The products of the spoVA operon are involved in dipicolinic acid uptake into developing spores of *Bacillus subtilis*. *J Bacteriol* 184: 584–587. PMID: [11751839](#)
38. Vepachedu VR, Setlow P (2007) Role of SpoVA proteins in release of dipicolinic acid during germination of *Bacillus subtilis* spores triggered by dodecylamine or lysozyme. *J Bacteriol* 189: 1565–1572. PMID: [17158659](#)
39. Hillmann F, Doring C, Riebe O, Ehrenreich A, Fischer RJ, Bahl H. (2009) The role of PerR in O<sub>2</sub>-affected gene expression of *Clostridium acetobutylicum*. *J Bacteriol* 191: 6082–6093. doi: [10.1128/JB.00351-09](#) PMID: [19648241](#)
40. Lehmann Y, Meile L, Teuber M (1996) Rubrerythrin from *Clostridium perfringens*: cloning of the gene, purification of the protein, and characterization of its superoxide dismutase function. *J Bacteriol* 178: 7152–7158. PMID: [8955396](#)
41. Riebe O, Fischer RJ, Wampler DA, Kurtz DM Jr., Bahl H (2009) Pathway for H<sub>2</sub>O<sub>2</sub> and O<sub>2</sub> detoxification in *Clostridium acetobutylicum*. *Microbiology* 155: 16–24. doi: [10.1099/mic.0.022756-0](#) PMID: [19118342](#)
42. Amador-Noguez D, Feng XJ, Fan J, Roquet N, Rabitz H, Rabinowitz JD. (2010) Systems-level metabolic flux profiling elucidates a complete, bifurcated tricarboxylic acid cycle in *Clostridium acetobutylicum*. *J Bacteriol* 192: 4452–4461. doi: [10.1128/JB.00490-10](#) PMID: [20622067](#)
43. Crown SB, Indurthi DC, Ahn WS, Choi J, Papoutsakis ET, Antoniewicz MR. (2011) Resolving the TCA cycle and pentose-phosphate pathway of *Clostridium acetobutylicum* ATCC 824: Isotopomer analysis, in vitro activities and expression analysis. *Biotechnol J* 6: 300–305. doi: [10.1002/biot.201000282](#) PMID: [21370473](#)
44. Harrington DJ (1996) Bacterial collagenases and collagen-degrading enzymes and their potential role in human disease. *Infect Immun* 64: 1885–1891. PMID: [8675283](#)
45. Fagan RP, Fairweather NF (2011) *Clostridium difficile* has two parallel and essential Sec secretion systems. *J Biol Chem* 286: 27483–27493. doi: [10.1074/jbc.M111.263889](#) PMID: [21659510](#)
46. Hartman AH, Liu H, Melville SB (2011) Construction and characterization of a lactose-inducible promoter system for controlled gene expression in *Clostridium perfringens*. *Appl Environ Microbiol* 77: 471–478. doi: [10.1128/AEM.01536-10](#) PMID: [21097603](#)



## Complete Genome Sequence of the Nitrogen-Fixing and Solvent-Producing *Clostridium pasteurianum* DSM 525

Anja Poehle, Alexander Grosse-Honebrink, Ying Zhang, Nigel P. Minton, Rolf Daniel

Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Georg-August University, Göttingen, Germany; The Clostridia Research Group, BBSRC/EPSRC Synthetic Biology Research Centre, School of Life Sciences, Centre for Biomolecular Sciences, University of Nottingham, Nottingham, United Kingdom

Here, we report on the closed genome sequence of *Clostridium pasteurianum* DSM 525, which is an anaerobic, Gram-positive and endospore-forming organism. *C. pasteurianum* can fix N<sub>2</sub> and produce solvents such as butanol and 1,3-propanediol from carbohydrates. The genome consists of a single 4,350,673-bp replicon.

Received 30 December 2014 Accepted 5 January 2015 Published 19 February 2015

Citation Poehle A, Grosse-Honebrink A, Zhang Y, Minton NP, Daniel R. 2015. Complete genome sequence of the nitrogen-fixing and solvent-producing *Clostridium pasteurianum* DSM 525. *Genome Announc* 3(1):e01591-14. doi:10.1128/genomeA.01591-14.

Copyright © 2015 Poehle et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Rolf Daniel, rdaniel@gwdg.de.

The Gram-positive anaerobic spore-forming bacterium *Clostridium pasteurianum* DSM 525, is able to produce butanol from carbohydrates (1). In contrast to most other solventogenic clostridia, *C. pasteurianum* is able to grow with glycerol as sole carbon and energy source (1, 2) and couple glycerol breakdown with a highly active butanol-producing pathway. The major products during glycerol degradation are 1,3-propanediol, ethanol, and butanol (1, 3).

Strain DSM 525 was derived from the DSMZ (Braunschweig, Germany). Chromosomal DNA of *C. pasteurianum* DSM 525 was isolated using the MasterPure complete DNA purification kit (Epicentre, Madison, WI, USA). Subsequently, 454-shotgun and Illumina paired-end libraries were generated from the isolated DNA as described by the manufacturers. The libraries were sequenced using a 454 GS-FLX system (Titanium GS70 Chemistry, Roche Life Science, Mannheim, Germany) and MiSeq Illumina system (Illumina, San Diego, CA, USA), respectively. Sequencing yielded 201,156 454-shotgun and 1,215,244 paired-end Illumina reads. Assembly of the reads using the Roche Newbler assembly software 2.9 and the MIRA software (4) resulted in 139 contigs. For scaffolding and contig ordering, the move contigs tool of the Mauve genome alignment software (5) was used. The closed genome of *C. pasteurianum* ATCC 6013 (CP009267) served as the reference. Remaining gaps were closed by PCR-based techniques and Sanger sequencing of the products using BigDye 3.0 chemistry and an ABI3730XL capillary sequencer (Applied Biosystems, Life Technologies GmbH, Darmstadt, Germany). For this purpose, the Gap4 (v4.11) software of the Staden package (6) was employed. The complete genome of *C. pasteurianum* DSM 525 consists of a single chromosome of 4,350,673 bp with an overall G+C content of 30%. Automatic gene prediction was performed with the software tool prodigal (Prokaryotic Dynamic Programming Gene-finding Algorithm) (7). Identification of rRNA and tRNA genes was done with RNAMmer (8) and tRNAscan (9), respectively. An integrated microbial genomes/expert review (IMG/ER) system (10, 11) was used for automatic annotation, which was subsequently manually curated by using the Swiss-Prot, TrEMBL, and InterPro

databases (12). We identified 10 rRNA operons, 81 tRNA genes, 3,220 protein-encoding genes with function prediction, and 768 genes coding for hypothetical proteins. Genes coding for key enzymes of butanol fermentation such as butyryl-CoA dehydrogenase (*bad*), electron transfer flavoprotein (*eftAB*), 3-hydroxybutyryl-CoA dehydrogenase (*hbd*), and 3-hydroxybutyryl-CoA dehydratase (*crt*) form a cluster that is identical to those identified in other solventogenic clostridia, such as *C. acetobutylicum*, *C. saccharoperbutylacetonicum*, or *C. saccharobutylicum* (13–15). In addition, the genome of *C. pasteurianum* DSM 25 harbors a cluster coding for CoA transferase (*ctfAB*), acetoacetate decarboxylase (*adc*), and alcohol/aldehyde dehydrogenase (*adhE*), which showed the identical arrangement as the *sol* operon of *C. acetobutylicum* (16). We also encountered genes encoding acetate kinase (*ackA*), phosphate acetyltransferase (*pta*), butyrate kinase (*buk*), and phosphate butyryltransferase (*ptb*). In addition, the presence of the previously described genes encoding key enzymes for 1,3-propanediol production such as B<sub>12</sub>-dependent glycerol dehydratase (17) and 1,3-propanediol dehydrogenase (2) was confirmed.

**Nucleotide sequence accession number.** The complete genome sequence has been deposited in GenBank under the accession no. CP009268.

### ACKNOWLEDGMENTS

A.P. and R.D. thank the “Bundesministerium für Ernährung und Landwirtschaft (BMEL)” for support. A.G.H. acknowledges the financial support of the Forman Hardy Charitable Trust. Y.Z. and N.P.M. acknowledge the financial support of the United Kingdom Biotechnology and Biological Sciences Research Council (BBSRC), grant BB/L004356/1.

We thank Kathleen Gollnow and Frauke-Dorothee Meyer for technical support.

### REFERENCES

- Dabrock B, Bahl H, Gottschalk G. 1992. Parameters affecting solvent production by *Clostridium pasteurianum*. *Appl Environ Microbiol* 58: 1233–1239.
- Luers F, Seyfried M, Daniel R, Gottschalk G. 1997. Glycerol conversion to 1,3-propanediol by *Clostridium pasteurianum*: cloning and expression



- of the gene encoding 1,3-propanediol dehydrogenase. *FEMS Microbiol Lett* 154:337–345. <http://dx.doi.org/10.1111/j.1574-6968.1997.tb12665.x>.
3. Biebl H. 2001. Fermentation of glycerol by *Clostridium pasteurianum*—batch and continuous culture studies. *J Ind Microbiol Biotechnol* 27: 18–26. <http://dx.doi.org/10.1038/sj/jim/7000155>.
  4. Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. *In* Wingender E (ed), *Computer science and biology: proceedings of the German conference on bioinformatics (GCB) 1999*. GBF-Braunschweig, Hannover, Germany. Department of Bioinformatics, Braunschweig, Germany.
  5. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. <http://dx.doi.org/10.1371/journal.pone.0011147>.
  6. Staden R, Beal KF, Bonfield JK. 2000. The Staden package, 1998. *Methods Mol Biol* 132:115–130.
  7. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <http://dx.doi.org/10.1186/1471-2105-11-119>.
  8. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <http://dx.doi.org/10.1093/nar/gkm160>.
  9. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25: 955–964. <http://dx.doi.org/10.1093/nar/25.5.0955>.
  10. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. 2009. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 25:2271–2278. <http://dx.doi.org/10.1093/bioinformatics/btp393>.
  11. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res* 40:D115–D122. <http://dx.doi.org/10.1093/nar/gkr1044>.
  12. Zdobnov EM, Apweiler R. 2001. InterProScan—an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17: 847–848. <http://dx.doi.org/10.1093/bioinformatics/17.9.847>.
  13. Nölling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, GTC Sequencing Center Production Finishing, and Bioinformatics Teams, Wolf YI, Tatusov RL, Sabathe F, Doucette-Stamm L, Soucaille P, Daly MJ, Bennett GN, Koonin EV, Smith DR. 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol* 183: 4823–4838. <http://dx.doi.org/10.1128/JB.183.16.4823-4838.2001>.
  14. Poehlein A, Krabben P, Dürre P, Daniel R. 2014. Complete genome sequence of the solvent producer *Clostridium saccharoperbutylacetonicum* strain DSM 14923. *Genome Announc* 2(5):e01056-14. <http://dx.doi.org/10.1128/genomeA.01056-14>.
  15. Poehlein A, Hartwich K, Krabben P, Ehrenreich A, Liebl W, Dürre P, Gottschalk G, Daniel R. 2013. Complete genome sequence of the solvent producer *Clostridium saccharobutylicum* NCP262 (DSM 13864). *Genome Announc* 1(6):e00997-13. <http://dx.doi.org/10.1128/genomeA.00997-13>.
  16. Fischer RJ, Helms J, Dürre P. 1993. Cloning, sequencing, and molecular analysis of the *sol* operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. *J Bacteriol* 175:6959–6969.
  17. Macis L, Daniel R, Gottschalk G. 1998. Properties and sequence of the coenzyme B<sub>12</sub>-dependent glycerol dehydratase of *Clostridium pasteurianum*. *FEMS Microbiol Lett* 164:21–28. <http://dx.doi.org/10.1111/j.1574-6968.1998.tb13062.x>.

