



The University of
Nottingham

UNITED KINGDOM • CHINA • MALAYSIA



BBSRC Doctoral
Training Partnerships

Improving the expression and secretion of GH48 family cellulases in *C. acetobutylicum*

Tom Wilding-Steele BSc (Hons)

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

September 2017

Abstract

Although most energy is currently generated from fossil fuels, their continued use is unsustainable. Biofuels provide an attractive alternative. Unfortunately, such fuels are mainly produced from food crops, causing issues with food security. Lignocellulosic biomass represents the most logical alternative feedstock. However, its recalcitrance means that its use is currently limited by the economic costs involved in its deconstruction. One possible solution would be to undertake consolidated bioprocessing (CBP), a method by which cellulase production, cellulose degradation and fermentation take place in one step. To date, however, no suitable organism(s) have been identified.

Clostridium acetobutylicum is a promising candidate for the generation of a truly CBP organism as it can utilise all sugars present in lignocellulose and it has been used at an industrial scale to produce butanol. It is, however, not cellulolytic. Previous studies have shown that the main barrier to engineering *C. acetobutylicum* to grow on lignocellulose is the low level of production of cellulases from the GH9 and GH48 families. For this reason, the primary aim of this project was to increase the production of these particular cellulases.

To accomplish the project's goal, several attempts were made to increase the level of expression of genes encoding GH48 family hydrolases in *C. acetobutylicum*. These cellulases were shown to be able to bind specifically to N-deacetylated but not acetylated peptidoglycan. *C. acetobutylicum*'s peptidoglycan was determined to be partially N-deacetylated. Proteins binding to peptidoglycan have been shown to be toxic to cells, which may explain the low level of secretion of GH48 family cellulases by *C. acetobutylicum*.

Co-expression of OrfXp, a small protein containing a cohesin domain and believed to be anchored to the cell membrane, was shown to significantly improve the secretion levels of GH48 family cellulases in *C. acetobutylicum*, by an as yet unidentified mechanism. This previously uncharacterised protein, for which homologs are found in all cellulosomal systems, is probably essential for the secretion of GH48 cellulases in all cellulosomal, cellulolytic bacteria.

Acknowledgements

I would like to thank my supervisors Katalin Kovács and Nigel Minton, for all of the invaluable help, support and advice they have given me.

I would like to thank everyone at the SBRC for their friendship and advice during the course of my PhD, with special thanks to members of the Klaustridia and Geobacillus groups. Additionally, I would also like to thank Philippe Soucaille for his help and advice.

I would also like to thank the BBSRC for providing the funding for my PhD.

I would like to thank all members of “The Globe” and friends of the Globe for all the laughter and general good times. Additionally, I would like to thank everyone at the Depot and the University of Nottingham Climbing Club for all the fun times had climbing boulders, cliffs and mountains.

Finally, I would like to thank my family for all of their love, care, encouragement and support.

Table of Contents

Abstract	i
Acknowledgements	iii
Table of Contents	iv
List of Figures	x
List of Tables	xii
Abbreviations	xii
Chapter 1 - General Introduction	1
1.1 Background	1
1.1.1 Need for alternatives to fossil fuels	1
1.2 Composition of lignocellulosic biomass	3
1.2.1 Cellulose	5
1.2.2 Hemicellulose	8
1.2.3 Lignin	9
1.2.4 Model substrates	11
1.2.5 Current methods of lignocellulose utilisation	11
1.3 Consolidated Bioprocessing	13
1.3.1 Native strategy	14
1.3.2 Recombinant strategy	15
1.3.3 Consortia	16
1.4 Cellulases and Cellulosomes	18
1.4.1 Introduction to cellulosomes	18
1.4.2 Scaffoldins	22
1.4.3 Carbohydrate-binding domains	23
1.4.4 Cellulases and hemicellulases	23
1.4.5 Designer cellulosomes	25
1.4.6 <i>R. cellulolyticum</i> 's cellulosome	28
1.5 <i>Clostridium acetobutylicum</i>	30
1.5.1 Introduction to <i>C. acetobutylicum</i>	30
1.5.2 <i>C. acetobutylicum</i> metabolism	32
1.6 <i>Clostridium acetobutylicum</i> 's native cellulosome	33
1.6.1 Potential reasons why <i>C. acetobutylicum</i> cannot utilise cellulose	34
1.6.2 Engineering <i>C. acetobutylicum</i> to grow on lignocellulose	38
1.7 Genetic tools for clostridia	43
1.7.1.1 Plasmids	43
1.7.1.2 Promoters	43

1.7.1.3	Gene inactivation and chromosomal insertions.....	44
1.8	Aims	45
Chapter 2	Materials and Methods	47
2.1	Materials and suppliers	47
2.2	Preparation of growth media	47
2.2.1	Preparation of liquid growth media	47
2.2.2	Preparation of solid growth media.....	48
2.2.3	Preparation of antibiotics	49
2.3	Engineering of <i>E. coli</i>	50
2.3.1	List of <i>E. coli</i> strains.....	50
2.3.1	Storage of <i>E. coli</i> strains.....	50
2.3.2	Culture of <i>E. coli</i> cells.....	50
2.3.3	Preparation of <i>E. coli</i> electrocompetent cells.....	51
2.4	Engineering of <i>C. acetobutylicum</i>	52
2.4.1	Methylation of plasmid DNA	52
2.4.2	<i>C. acetobutylicum</i> cell culture.....	52
2.4.3	Storage of <i>C. acetobutylicum</i> strains	52
2.4.4	Preparation of electrocompetent <i>C. acetobutylicum</i> cells	53
2.4.5	Transformation of <i>C. acetobutylicum</i>	54
2.4.6	Integration of constructs into the genome	54
2.4.7	Creation of in-frame deletion mutants.....	55
	Screening of engineered <i>C. acetobutylicum</i>	55
2.5	General molecular biology techniques	55
2.5.1	Ordering of synthesised genes	55
2.5.2	Purification of plasmid DNA.....	56
2.5.3	PCR amplification.....	56
2.5.4	Restriction analysis of plasmid DNA	56
2.5.5	Agarose gel electrophoresis	57
2.5.6	Ligation of Plasmid DNA	57
2.5.7	Electroporation of <i>E. coli</i> electrocompetent cells	58
2.5.8	Sequencing of DNA constructs	58
2.6	Bioinformatics.....	58
2.6.1	Plasmid map analysis.....	58
2.6.2	Sequence databases	59
2.6.3	BLAST analysis.....	59
2.6.4	Prediction of signal peptides	59
2.6.5	Analysis of RNA structure	59

2.6.6	Statistical analysis	59
2.7	Protein expression analysis	60
2.7.1	Growth conditions for protein analysis	60
2.7.2	Denaturing analysis.....	60
2.7.3	Western blots	61
2.7.4	Coomassie blue staining	62
2.8	Purification of Cel48F from <i>E. coli</i>	62
2.8.1	Growth of <i>E. coli</i>	62
2.8.2	Purification of Cel48F.....	62
2.9	Methods used to assess the interaction between Cel48F and <i>C. acetobutylicum</i>	63
2.9.1	Microscopy.....	63
2.9.2	Autoclaved cells assay	63
2.9.3	Bacteriocin assay	64
2.9.4	Peptidoglycan analysis and peptidoglycan assays.....	65
2.9.4.1	Purification of peptidoglycan.....	65
2.9.4.2	Chemical acetylation of peptidoglycan.....	66
2.9.4.3	Peptidoglycan turbidity assay	66
2.9.4.4	Peptidoglycan binding assay.....	67
2.9.4.5	Preparation of peptidoglycan for LC/MS	67
2.9.4.6	Analysis of peptidoglycan by LC-MS and LC-MS/MS	67
2.10	Analysis of cellulase activity	68
2.10.1	Preparation of Phosphoric acid swollen cellulose (PASC)	68
2.10.2	Growth conditions	69
2.10.3	Cellulase activity assay.....	69
2.10.4	Reducing sugar analysis	69
2.11	Cloning of plasmid constructs.....	70
2.11.1	Cloning of Cel5A, Cel48F and Cel9E downstream the PtcdB promoter. 70	
2.11.2	Cloning the native Cel48F (with its native dockerin domain) and Cel48F without its dockerin downstream the PtcdB promoter.	70
2.11.3	Integration of Cel48F into the chromosome of <i>C. acetobutylicum</i>	71
2.11.4	Cloning of 5 prime UTRs	72
2.11.5	Cloning of 3 prime UTRs	72
2.11.6	Cloning of PrsA.....	73
2.11.7	Cloning of GH48 enzymes.....	73
2.11.8	Construction of <i>E. coli</i> Cel48F overexpression vector	74
2.11.9	Construction of in-frame deletion vectors	74

2.11.10	Integration of Flag-tagged Cel48A into <i>C. acetobutylicum</i>	74
2.11.11	Quick change mutagenesis of Cel48F_N.....	75
2.11.12	Overexpression of <i>R. cellulolyticum</i> 's OrfXp.....	75
2.11.13	Construction of Chimeric Cel48F	75
Chapter 3	Expression of <i>R. cellulolyticum</i> cellulases in <i>C. acetobutylicum</i> using a lactose inducible system.....	76
3.1	Introduction	77
3.1.1	Lactose inducible expression system.....	79
3.2	Results.....	81
3.2.1	Selection of cellulases for testing the lactose inducible system	81
3.2.2	Expression of <i>R. cellulolyticum</i> cellulases using the lactose-inducible system. 82	
3.2.3	Comparison of the lactose inducible system with the facoid promoter 87	
3.2.4	Effect of the addition of protease inhibitors on Cel48F expression	90
3.3	Discussion	92
3.4	Key outcomes	94
Chapter 4	Efforts to increase the expression level of GH48 family cellulases in <i>C. acetobutylicum</i>	95
4.1	Expression of different GH48 family cellulases in <i>C. acetobutylicum</i>	96
4.1.1	Introduction	96
4.1.2	Results.....	98
4.2	mRNA stability	100
4.2.1	Introduction	100
4.2.2	Effect of 5' and 3' stem-loops on the expression of Cel48F	103
4.3	Potential issues related to the secretion of GH48 family cellulases	109
4.3.1	Sec secretion system	110
4.3.2	Factors affecting secretion of protein	111
4.3.3	Effect of expression of PrsA chaperones on Cel48F expression.....	112
4.4	Discussion	114
4.5	Key outcomes	117
Chapter 5	Does Cel48F interact with peptidoglycan during secretion?.....	118
5.1	Introduction	119
5.1.1	Disruption of bacterial cell walls.....	119
5.1.2	Peptidoglycan	119
5.1.3	Peptidoglycan hydrolases	120
5.1.4	Peptidoglycan modifications	121
5.2	Results.....	125

5.2.1	Microscopy analysis	125
5.2.2	Cel48F activity on autoclaved <i>C. acetobutylicum</i> cells	130
5.2.3	Purification of Cel48F.....	133
5.2.4	Purified Cel48F is not active on autoclaved <i>C. acetobutylicum</i> cells or purified peptidoglycan.....	136
5.2.5	Can Cel48F bind to peptidoglycan?	138
5.2.6	Cel48F is not toxic to growing cultures.....	140
5.3	Discussion	142
5.4	Key points	144
Chapter 6 Modification of <i>C. acetobutylicum</i> 's peptidoglycan to enable expression of Cel48F.		145
6.1	Introduction	146
6.2	Results.....	147
6.2.1	Analysis of wild-type <i>C. acetobutylicum</i> peptidoglycan	147
6.2.2	LC-MS/MS analysis.....	153
6.2.3	Knocking out putative N-deacetylase genes.....	158
6.2.3.1	HPLC/MS analysis of peptidoglycan of N-deacetylase mutants	159
6.2.3.2	Expression of Cel48F in N-deacetylase mutants.....	160
6.3	Discussion	162
6.4	Key points	163
Chapter 7 Co-expression of OrfXp results in increased expression of Cel48F		164
7.1	Introduction	164
7.2	Results.....	165
7.2.1	Over-expression of Cel48A in <i>C. acetobutylicum</i>	165
7.2.2	Expression of inactive Cel48F in <i>C. acetobutylicum</i>	167
7.3	OrfXp.....	169
7.3.1	Expression of OrfXp increases Cel48F expression	171
7.3.2	OrfXp localisation.....	174
7.3.3	Enzyme activity	176
7.4	Discussion	177
7.5	Key points	180
General Discussion.....		181
Outcomes of the thesis.....		181
Limitations and proposed future work.....		185
Conclusions.....		189
Professional Internships for PhD Students Reflection Form		190
Appendices		192

Appendix 1	192
Appendix 2	197
Appendix 3.1	199
Appendix 3.2	201
Appendix 4	201
Appendix 5	203
Appendix 6	206
Wild-type	206
Wild-type plus lysozyme	208
WT plus wild-type supernatant	209
WT plus Cel48F supernatant.....	210
References	212

List of Figures

Figure 1.1. Structure of the secondary plant cell wall.	5
Figure 1.2. Basic structure of cellulose.	6
Figure 1.3. Mechanism of breakdown of cellulose by cellulases	7
Figure 1.4. A) Schematic representation of simple and complex cellulosomal systems.	21
Figure 1.5. Assembly of native and designer cellulosomes.	26
In	26
Figure 1.6. A) Light microscope picture of <i>C. acetobutylicum</i> cells.	30
Figure 1.7. Schematic diagram of the ABE pathway of <i>C. acetobutylicum</i>	33
Figure 1.8. Schematic representation of the Cip-Cel cellulosomal operons of <i>R. cellulolyticum</i> and <i>C. acetobutylicum</i>	34
Figure 1.9. Domain structure of cellulases from <i>C. acetobutylicum</i> and <i>R. cellulolyticum</i>	36
Figure 3.1. Schematic diagram of the lactose inducible <i>tcdR</i> system.	81
Figure 3.2. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing cellulases derived from <i>R. cellulolyticum</i> under lactose induction.	83
Figure 3.3. Showing the growth rate of a <i>C. acetobutylicum</i> strain expressing Cel48F induced with 0, 1, 10 and 50mM of lactose.	84
Figure 3.4. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing the Cel5A cellulase from <i>R. cellulolyticum</i> under control of the <i>PtcdB</i> promoter in a <i>Li-tcdR</i> background.	85
Figure 3.5. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing the Cel9E and Cel48F cellulases from <i>R. cellulolyticum</i> under control of the <i>PtcdB</i> promoter in a <i>Li-tcdR</i> background.	86
Figure 3.6. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing the Cel48F cellulase from <i>R. cellulolyticum</i> either under the control of the <i>FacOID</i> promoter or the <i>PtcdB</i> promoter.	87
Figure 3.7. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing the Cel48F cellulase from <i>R. cellulolyticum</i> with different dockerin domains.	89
Figure 3.8. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing the Cel48F cellulase from <i>R. cellulolyticum</i> with protease inhibitors added.	91
Figure 4.1. Comparison of GH48 family cellulases.	98
Figure 4.2. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing different GH48 family cellulases.	99
Figure 4.3. The position of stem-loops in <i>R. cellulolyticum</i> 's Cip-Cel operon.	102
Figure 4.4. The transcription start sites of the <i>PtcdB</i> promoter.	103
Figure 4.5. Predicted RNA structures of <i>PtcdB1</i> and <i>PtcdB2</i> 5'UTR's.	104
Figure 4.6. Predicted 5'UTR structures of <i>SlpA</i> and <i>Cose_Thio</i>	105
Figure 4.7. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing the Cel48F cellulase from <i>R. cellulolyticum</i> under control of the <i>PtcdB</i> promoter with different 5' UTR's.	106
Figure 4.8. Predicted 3'UTR structures of <i>SL2</i> and <i>SlpA_T</i>	108

Figure 4.9. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains expressing the Cel48F cellulase from <i>R. cellulolyticum</i> under control of the PtcdB promoter with different 3' stem-loops.	109
Figure 4.10. Comparison of PrsA proteins.	113
Figure 4.11. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains co-expressing PrsA chaperones with the Cel48F cellulase from <i>R. cellulolyticum</i>	114
Figure 5.1. The basic structure of peptidoglycan.	120
Figure 5.2. Microscope images showing representative images of <i>C. acetobutylicum</i> strains in protoplast buffer.	126
Figure 5.3. Effect of supernatants from <i>C. acetobutylicum</i> strains on protoplast formation.	129
Figure 5.4. Activity of concentrated <i>C. acetobutylicum</i> supernatants towards autoclaved <i>C. acetobutylicum</i>	131
Figure 5.5. Activity of concentrated <i>C. acetobutylicum</i> supernatant towards purified <i>C. acetobutylicum</i> peptidoglycan	132
Figure 5.6. SDS PAGE and Western blot analysis of the soluble fraction of <i>E. coli</i> cells expressing Cel48F.	134
Figure 5.7. SDS PAGE showing His-tag purification of Cel48F.	135
Figure 5.8. Activity of purified Cel48F on PASC.	136
Figure 5.9. Activity of purified Cel48F on autoclaved <i>C. acetobutylicum</i> cells and purified peptidoglycan.	137
Figure 5.10. Cel48F and peptidoglycan co-precipitation assay.	139
Figure 5.11. Toxicity assay of el48F against growing cultures of <i>E. coli</i> , <i>B. subtilis</i> and <i>C. acetobutylicum</i>	141
Figure 6.1. The predicted stem structure of <i>C. acetobutylicum</i> 's peptidoglycan... ..	147
Figure 6.2. Chromatogram of the 656.33 ion.	151
Figure 6.3. 'Zoom' region of Spectrum at 1.44 minutes of ion 614.30.	152
Figure 6.4. MS-MS product ion spectra of precursor ion 698.3103 m/z.	156
Figure 6.5. MS-MS product ion spectrum of precursor ion 656.30 m/z.	157
Figure 6.6. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains overexpressing the Cel48F cellulase with the putative peptidoglycan deacetylase CAC_3377 deleted.	160
Figure 6.7. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains overexpressing the Cel48F cellulase with the putative peptidoglycan deacetylase CAC_2316 deleted.	161
Figure 7.1. SDS-PAGE/Western blot analysis of culture supernatants from <i>C. acetobutylicum</i> strains overexpressing the Cel48A and Cel48F cellulases.	166
Figure 7.2. SDS-PAGE/Western blot analysis of culture supernatants from <i>C. acetobutylicum</i> strains expressing active and inactive Cel48F from <i>R. cellulolyticum</i>	168
Figure 7.4. Domain composition and sequence comparison of OrfXp.	169
Figure 7.5. Composition of the strains used.	172
Figure 7.6. SDS-PAGE/Western blot of supernatants from <i>C. acetobutylicum</i> strains co-expressing GH48 family cellulases with OrfXp.	172
Figure 7.7. SDS-PAGE of supernatants and whole cell fraction from <i>C. acetobutylicum</i> strains expressing Cel48F with and without OrfXp.	175

Figure 7.8. Activity of supernatants from <i>C. acetobutylicum</i> strains expressing GH48 family cellulases on PASC.	176
---	------------

List of Tables

Table 2.1. Antibiotics used in the study.	49
Table 2.2. List of <i>E. coli</i> strains used in this study.	50
Table 6.1. Predicted and measured molecular masses of monomeric mucopeptides and their elution times as measured by RP-LC-MS	149
Table 6.2. Proposed structure, accurate mass, quantity and proposed elemental composition of <i>C. acetobutylicum</i> mucopeptides as determined by LC-MS/MS analysis.....	153
Table 6.3. Showing candidate <i>C. acetobutylicum</i> N-deacetylase genes.	159
Table 7.1. Estimated expression of GH48 family cellulases and their relative fold-change compared to Cel48F.	173

Abbreviations

µg	Microgram
µl	Microlitre
µM	Micromolar
A	Alanine
ABE	Acetone, butanol, and ethanol
ACE	Allele-coupled exchange, a technique where a gene can be integrated into a target site on the genome through the use of homology arms of different lengths thereby controlling the order of recombination
BB2	Biobrick-2 format, describing constructs flanked by upstream <i>EcoRI/NotI/SpeI</i> and downstream <i>NheI/NotI/PstI</i> restriction sites, allowing iterative cloning
BCA	Bicinchoninic acid
BMCC	Bacterial microcrystalline cellulose
bp	Base pairs
BSA	Bovine serum albumin
CAT	Chloramphenicol acetyltransferase
CBD	Cellulose binding domain
CBM	Carbohydrate binding module
CBM	Clostridia basal medium
CBP	Consolidated bioprocessing, the generation of products from biomass by the use of an organism or cohort able to both degrade lignocellulose and ferment the products
CGM	Clostridia growth medium
CMC	Carboxymethyl cellulose

CW	Cell wall
Da	Daltons
DNA	Deoxyribonucleic acid
DOC	Deoxycholate
DTT	Dithiothreitol
E	Glutamic acid
EPB	Electroporation buffer
FOA	Fluoroorotic acid
g	Gram
GH	Glycoside hydrolase
GHG	Green House Gasses
GlcN	Glucosamine
GlcNAc	N-acetylglucosamine
GRAS	Generally Regarded As Safe
L	Litre
LB	Lysogeny broth medium
M	Molar
mg	Milligram
ml	Millilitre
mM	Millimolar
MurN	Muramic acid
MurNAc	<i>N</i> -acetylmuramic acid
ng	nanogram
Oid	'Ideal' symmetrical lac operator
PAGE	Polyacrylamide gel electrophoresis
PASC	Phosphoric acid-swollen cellulose
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pmol	Picomoles
RBS	Ribosome binding site
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
SDS	Sodium dodecyl sulfate
SEM	Scanning electron microscopy
SLH	Surface layer homology
SN	Supernatant
TBS	Tris-buffered saline
TCA	Trichloroacetic acid
TEM	Transmission electron microscopy
TMB	3,3',5,5'-tetramethylbenzidine
tRNA	Transfer RNA
YTG	Yeast, tryptone and glucose medium

Chapter 1 - General Introduction

1.1 Background

1.1.1 Need for alternatives to fossil fuels

Fossil fuels such as coal, oil and natural gas are a limited natural resource. Around 80% of the world's energy comes from fossil fuels with the remainder coming from nuclear and renewables ('Publication: World Energy Balances - 2017 Edition - Overview' 2017).

There is an urgent need for alternatives to fossil fuels as the burning of fossil fuels is causing climate change mainly due to an increase in the atmospheric levels of carbon dioxide (CO₂), methane (CH₄) and nitrous oxide (N₂O) (IPCC, 2014). Currently, global temperatures have already risen by 0.85 °C, and global sea levels have risen by 19 cm since 1850 (Hansen et al. 2016). This increase in temperature is due to an increase in greenhouse gas emissions with atmospheric levels of CO₂ rising from 280 ppm in 1850 to 380 ppm in 2014 (IPCC, 2014). The increase in temperature has had an impact on all ecosystems across continents and oceans, including increased incidences of extreme weather and climate events such as droughts, floods, cyclones and wildfires (Mann et al. 2017).

Current predictions indicate that global surface temperatures are likely to exceed 1.5 °C by 2100 although this is dependent on future GHG emissions as well as natural climate variability. This rise in temperature could result in rising sea levels of around 0.5 m, with more frequent hot and cold temperatures over most land areas on daily and seasonal timescales (IPCC, 2014).

Increasing temperature will have a major impact on ecosystems with a large percentage of species facing extinction due to their inability to adapt to the change in climate (Urban 2015). The impact on the human population could result in decreased

food and water security due to increased extreme weather conditions. The impact this could have on global stability and economics is unquantifiable but could lead to water rationing, starvation and potentially war due to increased competition over scarce resources (Wheeler and Braun 2013).

The Paris climate change agreement has committed to keeping global temperature rise to well below 2 °C of pre-industrial levels with efforts to keep it below 1.5 °C ('United Nations Treaty Collection' 2017). However, global demand for energy is predicted to rise due to an increase in prosperity as well as an increase in the world's population from 7.2 billion to 8.8 billion people by 2035 (Gerland et al. 2014). Currently, 85% of the world's energy is generated from fossil fuels with the remainder coming from nuclear and renewables ('Publication: World Energy Balances - 2017 Edition - Overview' 2017). Although renewables are expecting to grow rapidly with their share of global power expected to increase to 20% by 2035 (Bp-Energy-Outlook-2017, 2017).

One of the main uses of fossil fuels, of around 18% of the total, is in transportation as petrol or diesel. Moreover, the transportation sector is heavily dependent on oil products, providing 93% of the energy for transportation in 2011 (Bp-Energy-Outlook-2017, 2017). Substituting petrol and diesel with biofuels such as butanol could provide an attractive alternative to finite resources, as it can be used in current engines without modifications and use the existing infrastructure for fuel distribution (Tigunova, Shulga, and Blume 2013; Nanda et al. 2017; Green 2011).

Current biofuels, generated from renewable biomass, consist of ethanol, biodiesel and biogas. Bioethanol and biodiesel are currently produced from food-crop derived sugars, by microbial fermentation using microorganisms such as brewer's yeast (*Saccharomyces cerevisiae*). Currently, around 80 million tonnes of oil equivalent of bioethanol and biodiesel are produced annually (BP Global, 2017).

However, the use of food-crop derived sugars has raised ethical concerns, as these should solely be used for human consumption. Additionally, the carbon-neutral nature of biofuels has also been questioned, due to the energy needed to farm and transport these crops (DeCicco et al. 2016).

The use of lignocellulosic biomass as an alternative feedstock is an attractive choice as this is a readily available waste product from forestry, agriculture and the food industry and it does not compete with food supplies.

Unfortunately, the production of biofuels from lignocellulosic biomass is hindered by its recalcitrance. Current methods for releasing the sugars locked within the biomass require expensive pre-treatment methods, in particular the addition of cellulases making the process economically inviable (G. Liu, Zhang, and Bao 2016).

1.2 Composition of lignocellulosic biomass

Lignocellulose is the most abundant polysaccharide in the world, with plants producing almost 200 billion tonnes of cellulose per year. This vast feedstock could potentially be used for the production of biofuels and renewable chemicals (Claassen et al. 1999).

Lignocellulose is composed of three different classes of compound: cellulose, hemicellulose and lignin. The ratio between the three compounds depends on the source of the lignocellulose, and this varies between species, different tissues and the age of the plant. Cellulose is usually present at around 40-50%, followed by hemicellulose (10-40%) and lignin (10-25%) (Pauly and Keegstra 2010). Cellulose is a repeating glucose polymer, which forms fibrils with varying degrees of crystallinity (O'sullivan 1997). Hemicelluloses are typically branched polysaccharides where the backbone is a polymer of xylose, glucose or mannan (Scheller and Ulvskov 2010). Lignin is a polymer of phenolic compounds (Calvo-Flores and Dobado 2010).

As lignocellulose constitutes the cell walls of higher plants, it can constitute up to 30% of a plant's dry weight. Its role is to provide rigidity and protection as well as structure and support to the plant; it also acts as a physical barrier against pathogenic organisms and the environment (Cosgrove and Jarvis 2012).

Although only a relatively small number of monomers are used in the construction of the plant cell wall, the variation of their macromolecular organisation results in a complex structure. An example of this is that genetic studies in Poplar and Arabidopsis have revealed that 1000-2500 genes are involved in cell wall synthesis (Sarkar, Bosneaga, and Auer 2009).

Plant cell walls are complex and can consist of both primary and secondary cell walls. All plant cells have a primary cell wall which is made early on in development. These are usually thin walls, and their main function is to maintain turgor pressure, they also play a role in controlling cell growth. Primary cell walls mainly consist of cellulose microfibrils (15-40%) embedded in a matrix of pectins (30-50%) and hemicellulose (20-30%) (See Figure 1.1) (Cosgrove and Jarvis 2012).

Secondary cell walls are formed after cells have stopped growing and are mainly produced by specialised cells with roles in structural integrity or water transport. They consist of cellulose with lignin and a smaller proportion of hemicellulose, mainly xylans and glucomannans (Cosgrove and Jarvis 2012; Speck and Burgert 2011).

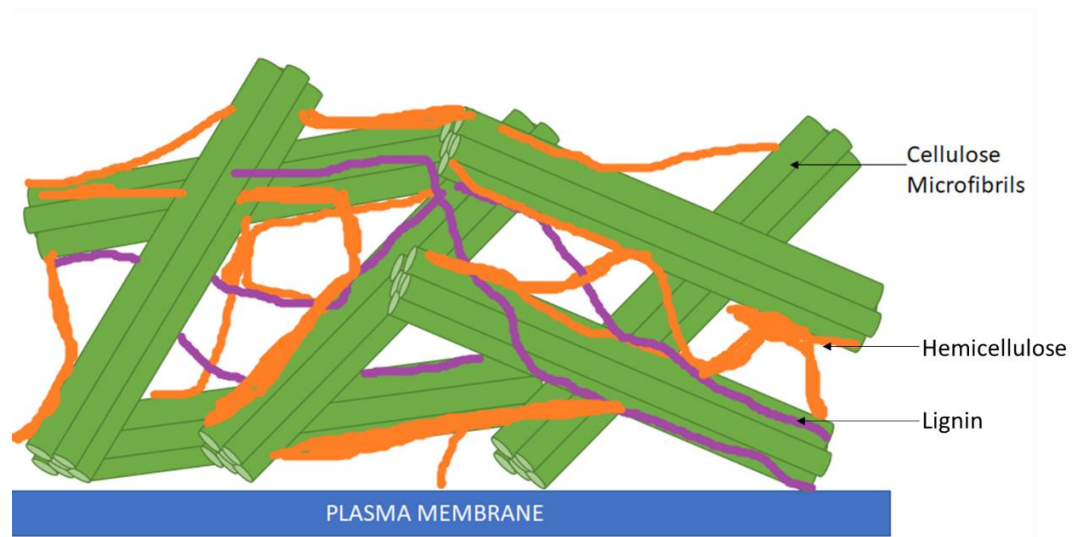


Figure 1.1. Structure of the secondary plant cell wall. The cell wall consists mainly of cellulose microfibrils linked by hemicellulose and lignin.

1.2.1 Cellulose

Cellulose consists of glucose chains linked by β -1,4-glycoside linkages which then form cellulose microfibrils interconnected by hydrogen bonds (Figure 1.2). In cell walls, cellulose chains usually consist of 8000 (in primary cell walls) to 15,000 (in secondary cell walls) units (Klemm et al. 2005). In nature, the arrangement of these chains usually exists in a mix of two different sub polymorphs. Cellulose I α has a triclinic unit cell with a single chain, while cellulose I β consists of a 2-chain monoclinic unit cell. A small proportion of cellulose is also in another polymorph cellulose II. In cellulose I the arrangement of chains is parallel, while in cellulose II they are antiparallel (O'sullivan 1997).

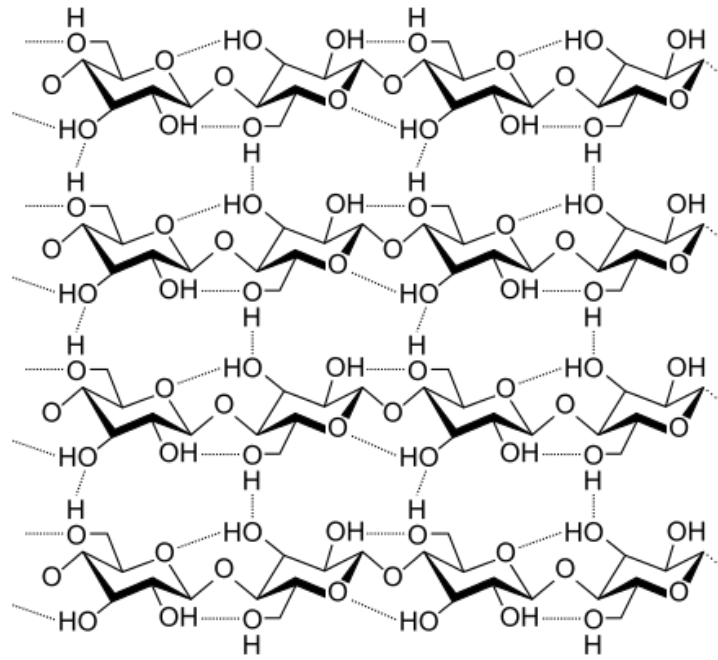


Figure 1.2. Basic structure of cellulose. Glucose molecules are linked by β -1,4-glycoside linkages and individual microfibrils are then interconnected by hydrogen bonds.

In native cellulose, the degree of crystallinity differs based on the species and location, with a higher degree of crystallinity found in secondary cell walls (around 70%) compared to primary cell walls (R. M. Brown 2004). The main reason for the recalcitrance of cellulose is its degree of crystallinity, preventing access of large molecules such as cellulases (Fan, Lee, and Beardmore 1980).

In both primary and secondary cell walls, cellulose is made at the plasma membrane by the cellulose synthase complex (CSC), this extrudes cellulose chains which bind together to form cellulose microfibrils. A triad of proteins called CESA proteins are the core components although other proteins are also associated with the CSC (McFarlane, Annet Doring, and Staffan Persson 2014).

Three types of glycoside hydrolase activities are needed to break down crystalline cellulose: endoglucanases, exoglucanases, and β -glucosidases. Endoglucanases randomly attack cellulose by breaking β -1-4 linkages. They are generally more active on amorphous cellulose and cannot act on short glucose chains (cellobiose).

Exoglucanases are often called cellobiohydrolases, these are processive cellulases and act from the non-reducing end of the cellulose, releasing cellobiose. They are generally more active on crystalline cellulose. Finally, β -glucosidases act on cellobiose and short chains of cellulose releasing glucose (Figure 1.3) (Wilson 2011).

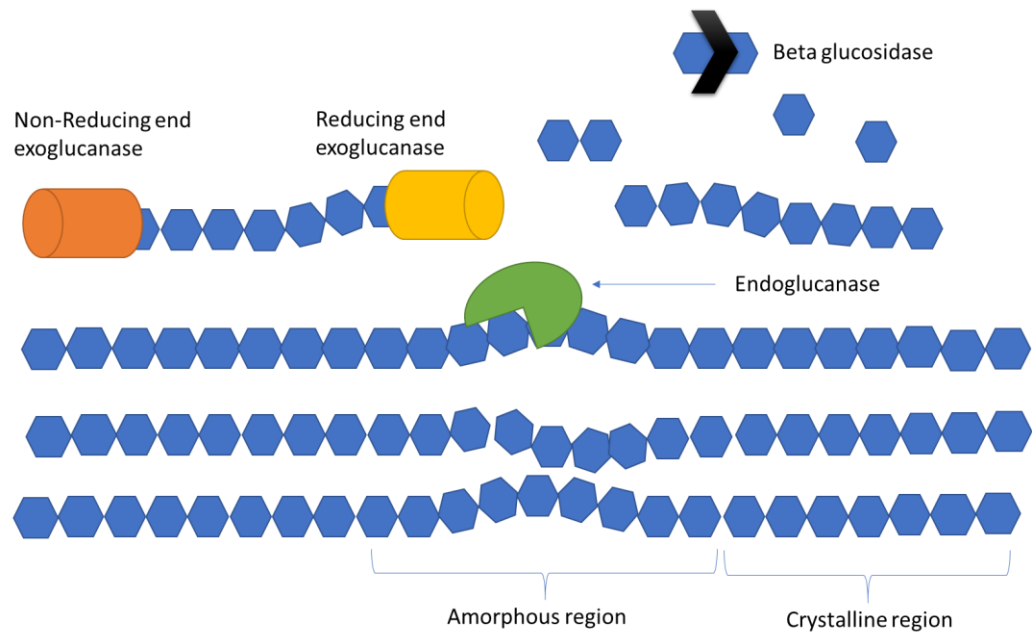


Figure 1.3. Mechanism of breakdown of cellulose by cellulases. At amorphous regions endoglucanases cut in the middle of the cellulose chain. This creates new binding sites for both reducing and non-reducing end exoglucanases. They are often processive releasing cellobiose as the main product. Cellobiose is broken down into glucose by beta glucosidases.

Additionally, non-cellulase enzymes have also been found to increase the rate of cellulose hydrolysis these include lytic polysaccharides; these are selective oxidative enzymes which cleave chains in crystalline regions (Agger et al. 2014). Other accessory proteins, for example, expansins and CBM's have also been found to increase the rate of crystalline cellulose degradation (Cosgrove 2000; Bunterngrsook et al. 2014).

Several studies have looked at modifying cellulose structure by reducing cellulose crystallinity. Several mutants in CESA proteins have been generated in Arabidopsis and tobacco plants. These have been shown to have reduced cellulose crystallinity and to be easier to digest (Harris et al. 2012; Sahoo et al. 2013).

1.2.2 Hemicellulose

Hemicellulose is a term used to describe a very broad range of polysaccharides found in both primary and secondary plant cell walls. The main function of hemicelluloses is to tether cellulose microfibrils, resulting in increased strength of the cell wall (Caffall and Mohnen 2009).

This diverse family is characterised by having a β -1-4 linked backbone with an equatorial configuration at C1 and C4. They consist of xylans, xyloglucan, mannans, glucomannans and β -glucans. These consist of either glucose (glucans and xyloglucans), xylose (xylan), mannose (mannans) or mixed (glucomannan) backbones. Importantly, they can have a branched backbone with additional sugars, for example, xylo-glucans are glucose chains heavily branched with xylose units (Scheller and Ulvskov 2010).

All hemicelluloses are synthesised in the Golgi apparatus from activated nucleotide sugars (Bar-Peled and O'Neill 2011). The backbones of xyloglucans, mannans and β -glucans are then synthesised by glycosyltransferases which are members of the cellulose synthase-like gene family (Cocuron et al. 2007; Dhugga et al. 2004). Xylan backbones, however, are synthesised by glycosyltransferases which have a single type II transmembrane loop (Brown et al. 2007). Hemicelluloses are also often acetylated; this occurs in the Golgi apparatus (Chang and Holtzapple 2000).

In primary cell walls, Angiosperms (hardwood) contain 20-25% xyloglucan and small amounts of arabinoxylan and glucomannans. Grasses contain small amounts of xyloglucan and 20-40% arabinoxylan, traces of glucomannans and small amounts of β -1-4- glucans. Gymnosperms (softwood) contain 10% xyloglucan and small amounts of arabinoxylan (Scheller and Ulvskov 2010; Schädel et al. 2010).

In secondary cell walls, Angiosperms contain 20-30% glucuronoxylan and traces of glucomannans and galactoglucan-mannans. Grasses contain 40-50% arabinoxylan, and trace amounts of glucomannan, while gymnosperms contain 5-15% arabinoxylan and 10-30% galactoglucan-mannans (Scheller and Ulvskov 2010).

Degradation of hemicellulose is more complex than that of cellulose as it is a more diverse substrate requiring the presence of several hemicellulases and other additional enzymes. However, the individual hemicelluloses are themselves easier to degrade than cellulose. Hemicellulose acts to increase lignocellulose recalcitrance by two main mechanisms; acting as a physical barrier by blocking cellulases from reaching cellulose and by hemicellulose directly inhibiting cellulases (B. Yang et al. 2011).

Attempts have been made to modify the hemicellulose content in plants, for example, to decrease xylan feruloylation and to replace xylan with mannan. The replacement of xylan with mannan is important as pentose sugars are poorly utilised by bacteria and yeast, so conversion of these to hexose would be advantageous (Pauly and Keegstra 2008). Additionally, hemicellulose chains are often heavily acetylated, and deacetylation has been shown to triple cellulose digestibility. However, there is a debate on the level of deacetylation that is needed and plants have not yet been engineered with reduced xylan acetylation levels (Kong, Engler, and Soltes 1992; Holzforschung et al. 2006).

1.2.3 Lignin

Lignin is a polymer of aromatic alcohols and is found in secondary cell walls. It is essential for the structural integrity of the cell wall and is responsible for the stiffness of the stem. It also plays a vital role in waterproofing the cell wall, enabling the transport of water. Additionally, it has a crucial role in defence against damage and protection from pathogens (Boerjan, Ralph, and Baucher 2003).

Lignin is composed of monolignols which are polymerised; there are three different main types of monolignols: guaiacyl (G), syringyl (S) and p-hydroxyphenyl (H). These are synthesised from phenylalanine by membrane-bound proteins in the Endoplasmic reticulum (ER). They are subsequently transported to the cell wall, where they are oxidised to monolignol radicals by laccases and peroxidases which are then coupled generating the lignin polymer (Vanholme et al. 2010).

Lignin is a major factor for lignocellulose recalcitrance; this is due to lignin acting as a physical barrier preventing binding and absorption of cellulases and accessory proteins to cellulose. Common pre-treatments of lignocellulose also result in high levels of inhibitory compounds produced from lignin degradation (DeMartini et al. 2013).

Lignin as a source of recalcitrance was first shown in transgenic alfalfa which had variable lignin composition, as transgenic plants with lower levels of lignin were easier to degrade (F. Chen and Dixon 2007). This has since been shown in multiple species including; Tobacco, Sugarcane, Switchgrass, Poplar, Arabidopsis and Eucalyptus (Van Acker et al. 2013). As well as reducing recalcitrance, a reduction in lignin has also been found to affect the composition of cellulose and hemicellulose, usually resulting in an increase in cellulose (Hu et al. 1999).

Lignin is highly recalcitrant to both chemical and biological degradation. However, acid chlorite can effectively remove lignin from lignocellulose, and several organisms have been identified which can grow on lignin. White-rot basidiomycetes were the first organisms identified which could grow on lignin, and several bacteria mainly actinomycetes have since been identified to grow on lignin (Ruiz-Dueñas and Martínez 2009).

The enzymatic degradation of lignin has not been fully elucidated. Four enzymatic activities have been identified: lignin peroxidases (LiPs), manganese peroxidases

(MnPs), versatile peroxidases (VPs) and laccases. Reactive oxygen species are also thought to have a role in lignin degradation (Pollegioni, Tonin, and Rosini 2015).

1.2.4 Model substrates

In the lab model substrates are used for determining enzyme activity. These consist of pure cellulose substrates for example, Avicel, PASC (Phosphoric acid treated cellulose), CMC (carboxymethyl cellulose) and Bacterial cellulose. Avicel consists of highly crystalline cellulose I, and is the most relevant when looking at real substrates. PASC is Avicel treated with Phosphoric acid resulting in an extremely amorphous cellulose (Hall et al. 2010). CMC water soluble due to the presence of carboxymethyl groups, preventing it from forming a crystalline matrix. It is normally only used for determining the rate of endoglucanases. Bacterial cellulose forms a more crystalline structure compared to plant cellulose (Esa, Tasirin, and Rahman 2014).

Natural substrates which resemble substrates likely to be used industrially are also used, commonly used examples are wheat and rice straw. Sometimes pre-treated forms of these natural substrates are also utilised.

1.2.5 Current methods of lignocellulose utilisation

Currently, most lignocellulosic material is hydrolysed to its component sugars before it is used in fermentation by microorganisms. The hydrolysis consists of physical or chemical pre-treatment of the lignocellulosic biomass, often followed by biological or enzymatic treatment that uses a cocktail of different cellulases (Mussatto et al. 2010).

Several different methods for pre-treatment of lignocellulosic biomass have been established, and more efficient processes are constantly being developed. Pre-treatment processes can make use of either chemical, physical, biological or combinations of all these different methods. In addition, the direct conversion of

lignocellulose to compounds other such as synthesis gas and pyrolysis oil are also being explored (Sun and Cheng 2002; Munasinghe and Khanal 2010).

Chemical methods of degradation involve both strong and weak acid treatment, alkaline treatment, the use of ionic liquids as well as organosolv and ozonolysis. Physical treatments involve grinding, milling and microwave treatments. Physical-chemical treatments use steam explosion, liquid hot water, ammonia fibre explosion, wet oxidation and CO₂ explosion. Biological pre-treatments utilise cellulolytic organisms to start to break down the lignocellulose or utilise white rot fungi to degrade the lignin (Salehi Jouzani and Taherzadeh 2015; Kumar and Sharma 2017; X. Zhao, Cheng, and Liu 2009).

All current pre-treatments except biological, result in the production of inhibitory compounds. Acid hydrolysis results in the formation of furfural from xylose and hydroxyfurfural from glucose, these, in turn, can further react to form formic acid. The hydrolysis of lignin can also release phenolic compounds (Palmqvist and Hahn-Hägerdal 2000). Both furfural and hydroxyfurfural inhibit enzymes involved in solvent formation (Modig, Lidén, and Taherzadeh 2002), while lignin-derived phenolic compounds inhibit cellulases (Qin et al. 2016).

Regardless of the pre-treatment strategy, the conversion of lignocellulosic biomass to sugars is expensive. To make lignocellulosic biofuels economically viable, the costs of the pre-treatment has to be significantly reduced, in particular the price of cellulases (Lynd et al. 2005). For cellulosic ethanol and n-butanol from corn stover, current analysis estimates the cost of pre-treatment at 13% of total costs and the cost of cellulases at 16% of total costs (Tao et al. 2014).

1.3 Consolidated Bioprocessing

Consolidated bioprocessing (CBP) is where a single organism (or consortium) is able to simultaneously degrade all the components of the lignocellulosic biomass and efficiently ferment both pentose and hexose sugars into products such as biofuels. CBP would significantly reduce the costs of current processes that use lignocellulose biomass as a carbon source (Lynd et al. 2002). In addition, extending the product range of the CBP organism to produce higher value chemicals, for example, iso-butanol used in making plastics and rubber, could also contribute to the cost-effectiveness of the process (Adsul et al. 2011).

Unfortunately, there is currently no known single organism or consortium of organisms capable of producing biofuels at an industrial scale directly from lignocellulosic biomass. All known cellulolytic bacteria only produce limited amounts of solvents. For example, *R. cellulolyticum* produces solvents at mg/L concentrations compared to solvent producers which can produce over 40 g/L of solvents (Sreekumar et al. 2015; Higashide et al. 2011). The solvent yield required for an CBP process to be economically viable is not clear as it will depend on multiple factors including; the organism(s) utilised, the efficiency of biomass breakdown and the biofuel produced. For butanol a yield in the range of 20 g/L seems a reasonable aim while for ethanol the yield will likely be significantly higher around 60 g/L (Herring et al. 2016; Baral et al. 2016).

There are two main strategies for engineering a CBP organism. The first is known as the native strategy, and this is based on an organism that can grow on lignocellulosic biomass and is engineered to produce industrially relevant levels of solvents and chemicals. The second strategy, also known as the recombinant strategy, is where a well-established solvent producing organism is engineered to grow on lignocellulosic

biomass. In addition to these two strategies, there have been several reports describing the use of consortiums (comprised of known or unidentified bacterial species) for cellulosic solvent production (Lynd et al. 2005).

1.3.1 Native strategy

Several organisms have been considered as a suitable chassis for the application of the native strategy. These include fungal species, bacterial species producing free cellulolytic enzymes and bacterial species producing cellulolytic enzyme complexes attached to the cell, also known as cellulosomes. The model cellulolytic fungi, *Trichoderma reesei*, has been considered for CBP, and some success has been reported in engineering its ability to secrete increased amounts of cellulolytic enzymes, although no progress has been made with respect to engineering it to produce solvents (Q. Xu, Singh, and Himmel 2009).

Clostridium phytofermentans, a free-enzyme secreting organism has good potential to be a CBP organism, as it readily produces 7 g/L of ethanol and 8 g/L of acetate from lignocellulosic biomass under optimised growth conditions (Jin et al. 2011). Genetic tools for the manipulation of this organism have been developed, but no increases in product yield have been reported so far.

Efforts on cellulosome-producing bacteria have focused on *Ruminiclostridium cellulolyticum*, *Clostridium cellulovorans* and *Ruminiclostridium thermocellum*. *R. cellulolyticum* has been engineered to produce butanol, albeit at very low yields (Higashide et al. 2011). Engineering *C. cellulovorans* was more successful, as it resulted in a strain able to produce just over 3 g/L of butanol and ethanol (X. Yang, Xu, and Yang 2015). Similarly, *R. thermocellum* has been successfully engineered to produce 5 g/L of ethanol from cellulose (P. P. Lin et al. 2015; Higashide et al. 2011). However, these yields are still a long way off from being economically viable.

1.3.2 Recombinant strategy

To date, no organism has been engineered to be able to grow on lignocellulose. However, progress has been made in engineering several organisms including: *Saccharomyces cerevisiae*, *Bacillus subtilis* and *C. acetobutylicum*.

S. cerevisiae is a budding yeast currently used industrially for ethanol production. It can efficiently utilise hexose sugars and has been engineered to be able to utilise pentose sugars (which is required for a true CBP organism). Cellulases from both fungi and bacteria have been expressed in this host, including: *T. reesei*, *Ruminococcus flavefaciens*, *Aspergillus aculeatus* and *Ruminoclostridium cellovorans* derived cellulases, and these have been successfully secreted by the host (Den Haan et al. 2007; Hasunuma and Kondo 2012). However, only low levels of cellobiohydrolases were secreted and they also had very low activity. Significant improvements in the expression of these enzymes have since been achieved. However, in these studies the cellulases were expressed aerobically under optimal conditions. While *S. cerevisiae* produces ethanol optimally under anaerobic conditions, and it remains to be seen if the same level of expression can be achieved in anaerobic conditions (Kroukamp et al. 2013).

Additionally, *S. cerevisiae* has been engineered to display a mini-cellulosome, which may reduce the required expression level of cellulases needed (Artzi, Bayer, and Morais 2017). The cellulosome consisted of a scaffoldin and three cellulases resulting in 30% of a 10 g/L solution of PASC being utilised in 73 hours (Tanaka and Kondo 2014).

A strain has since been engineered which can grow on lignocellulosic biomass, however it produced very low levels of solvents. An increase in the level of solvents could be achieved when the cultures were supplemented with cellulases, indicating that the cellulosic capabilities of the strains need to be improved (Zhuo Liu et al. 2016).

Efforts have also been made to engineer *B. subtilis* to grow on lignocellulose, with the most recent study successfully secreting three cellulases from *R. cellulolyticum* (Cel5A, Cel9E and Cel48F) along with an anchored mini-cellulosome. Although it was initially reported to be able to grow on untreated biomass this key finding has since been retracted as the authors could not replicate the key finding of it growing on untreated biomass (Anderson et al. 2013).

Lactobacillus plantarum is a potential CBP candidate as it can produce high levels of ethanol. It has been engineered to express a cellulase and a xylanase from *T. fusca* (Moraïs et al. 2013). A later study expanded on this and used a consortium approach; three different strains were used with each secreting one component (cellulase, xylanase and a scaffoldin). Cells with an anchored scaffoldin showed activity on pre-treated wheat straw. However, the ability of the strains to degrade untreated wheat straw or grow on any cellulose substrate was not tested (Moraïs et al. 2014).

E. coli has also been engineered to express several xylanases and cellulases, resulting in a strain able to grow on xylan, and it has been engineered to grow on pre-treated cellulose (Gao et al. 2015; Zheng et al. 2012). However, *E. coli* has been shown to have very low solvent tolerance and will also need to be engineered to produce solvents (Zingaro and Terry Papoutsakis 2013).

1.3.3 Consortia

Several studies reported the use of co-cultivation technology in which a cellulolytic strain (or consortium) able to degrade lignocellulose, and a solventogenic strain (or consortium) is used to convert the released sugars by the former to produce solvents by the latter.

Co-culture of an established cellulolytic consortium, N3 with the solventogenic *C. acetobutylicum* ATCC 824 lead to a butanol yield of 3.73 g/L with filter paper as the

cellulose source (Wang et al. 2015). A study utilising two thermophilic bacteria species, *R. thermocellum* and *Clostridium thermolactium* generated a similar yield of 3.16 g/L of ethanol on cellulose. An improvement on these yields is needed for them to be suitable for industry.

Zuroff, Xiques, and Curtis developed a co-culture between *C. phytofermentans* and *S. cerevisiae*. However, no improvement of ethanol yield was achieved on cellulose compared to a mono *C. phytofermentans* fermentation. An increase in yield to 20 g/L of ethanol was however achieved when cellulases were added to the co-culture (Zuroff, Xiques, and Curtis 2013).

A promising study used a co-culture between *Aspergillus cellulolyticus* (cellulolytic fungi) and *S. cerevisiae*. This generated an ethanol yield of up to 25 g/L from purified cellulose. However, there are problems with this approach as *A. cellulolyticus* is an obligate aerobe and *S. cerevisiae* optimally produces ethanol anaerobically (Park, Naruse, and Kato 2012).

In another study, genetically engineered *C. cellulovorans* and *C. beijerinckii* was used to produce isobutanol. This co-culture produced 11.5 g/L solvent from 83.2 g/L from alkali-extracted deshelled corn cobs (AECC). A previous study using their wild-type equivalent only produced 4.11 g/L of butanol from 46.8 g/L of AECC. It should be noted that AECC is a form of pre-treated lignocellulose, with *C. beijerinckii* able to grow on it, albeit poorly (Wen et al. 2017, 2014).

Although some progress has been made on identifying suitable co-cultures, no co-culture has been identified which could produce industrial relevant levels of biofuel from untreated biomass.

1.4 Cellulases and Cellulosomes

A wide variety of aerobic and anaerobic organisms (bacteria and fungi) can degrade lignocellulosic biomass. Cellulosic biomass degradation is achieved by secreting cellulases, hemicellulases and assorted accessory proteins. Most aerobic bacteria and fungi secrete free cellulases while others, especially anaerobic bacteria and some fungi, assemble their cellulases into multi-protein complexes called cellulosomes.

1.4.1 Introduction to cellulosomes

The first cellulosome described and characterised was the one found in *R. thermocellum*, a gram-positive thermophilic, cellulolytic bacterium (Lamed et al. 1983). It has since been characterised in several other organisms including: *R. cellulolyticum* (the model mesophilic cellulolytic organism) and *C. acetobutylicum* (although its cellulosome was found to be inactive on cellulose) (Desvaux 2005b; Sabathé, Bélair, and Soucaille 2002).

Cellulosome-producing bacteria have been isolated from a wide variety of environments including: forest and pasture soils, compost and the microbiota of vertebrates and invertebrates. To date, only anaerobic bacteria have been found to produce cellulosomes. This is because under energy limited conditions of anaerobic bacteria they are required to produce a more efficient system (Edward A. Bayer et al. 2004).

Cellulosomes have been shown to be up to five times more efficient than an equivalent free enzyme system (Fierobe et al. 2005). The reasons for this are unclear it is thought that the close proximity of the enzymes results in a synergistic effect. Additionally, the cellulosomes are often attached to the cell wall, allowing the organism to quickly uptake any released sugars (Brás et al. 2016). It is also possible

that cellulosomes can physically separate the cellulose strands thus increasing the surface area for cellulases to attack (G. Resch et al. 2013).

The cellulosome consists of a structural scaffoldin protein which serves as a docking station for the catalytic components of the system. The scaffoldin protein contains several cohesin domains as well as a CBM (carbohydrate binding module). Cellulases and other accessory proteins contain dockerin domains which specifically recognise cohesin domains. The CBM is important as it attaches the cellulosome to cellulose, this is important as cellulose is insoluble while cellulases are soluble (Fontes and Gilbert 2010).

The cohesin-dockerin domains bind together via a strong calcium-dependent protein receptor-ligand bond, the mechanical strength of this pairing has been determined to be up to half the strength of a covalent bond (Stahl et al. 2012). Cohesin/dockerin pairing has been determined to be species-specific, i.e. a cohesin from *R. thermocellum* cannot bind to a dockerin from *R. cellulolyticum* (S. Pagès et al. 1997). However, recently there have been a few exceptions such as a Xyn dockerin from *R. thermocellum* interacting with CipC cohesins from *C. josui* and *R. cellulolyticum* (Haimovitz et al. 2008).

There are two main type of cellulosomal systems: simple and complex (Figure 1.4). Simple cellulosomal systems contain only one scaffoldin unit; this can have up to 9 cohesin domains. They are exclusively found in mesophilic bacteria, including *R. cellulolyticum*, *C. cellulovorans*, *C. josui* and *C. papyrosolvens*.

Complex cellulosomes contain more than one scaffoldin unit. They are generally found in thermophilic bacteria, but have also been identified in mesophiles. Organisms identified with a complex cellulosome include: *R. thermocellum*, *Acetivibrio cellulolyticus*, *Clostridium clariflavum*, *Ruminococcus flacefaciens* and *Ruminococcus*

champanellensis (Artzi, Bayer, and Moraïs 2017). Cellulosomal systems can also exist as cell-bound cellulosomal systems where they are anchored to the cell wall and as cell-free cellulosomal systems (see figure 1.4 B). It was recently discovered that *R. thermocellum* has both anchored and free cellulosomal systems. The cell-free cellulosomal system is thought to act as a long-range cellulosome (Q. Xu et al. 2016). Additionally, several organisms, for example, *R. thermocellum* secrete cellulosomal free cellulases (Vazana et al. 2010).

Fungal cellulosomes are not as well characterised as bacterial cellulosomes. Scaffolding proteins have been identified and they have been shown to assemble via divergent dockerin domains. Genetic evidence shows that they evolved independently from bacterial cellulosomes (Haitjema et al. 2017).

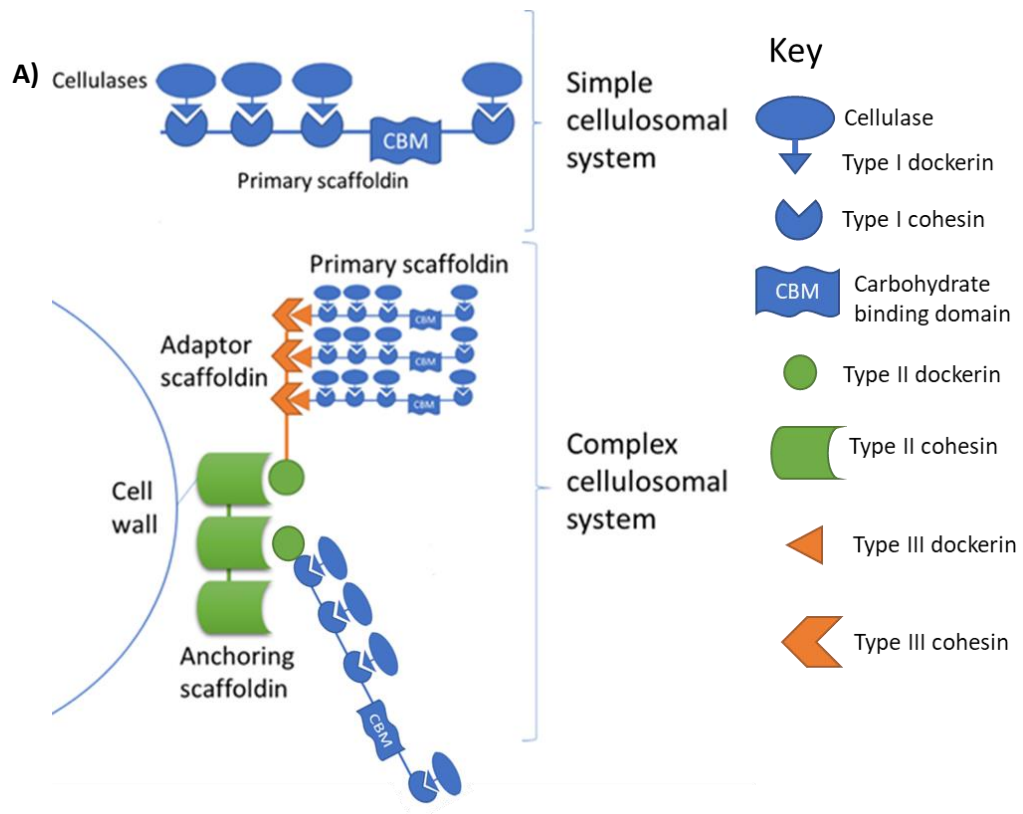
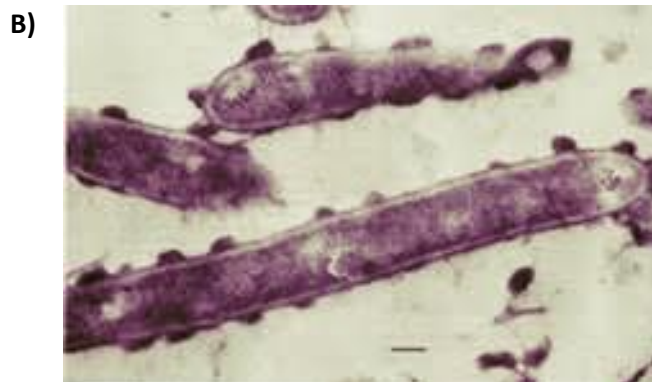


Figure 1.4. A) Schematic representation of simple and complex cellulosomal systems. Simple cellulosomes consist of one main scaffoldin unit (primary scaffoldin), and cellulases are integrated by interactions between type I dockerins of the cellulases and type I cohesins of the primary scaffoldin. They can be attached to the cell wall and are attached to cellulose via a CBM domain. Complex cellulosomes multiply the number of enzymes integrated by two different mechanisms: 1) Anchoring scaffoldins, which attach to the cell wall, can incorporate multiple primary scaffoldins via specialised cohesins-dockerin interactions (normally Type II). 2) Adaptor scaffoldins, which incorporate multiple scaffoldins with their additional enzymes, resulting in a much larger cellulosome. **B)** TEM picture of *R. thermocellum* cells with cellulosomes visible as protrusions attached to the cell surface (E A Bayer and Lamed 1986).

1.4.2 Scaffoldins

Three different types of scaffoldins have been described: primary, anchoring and adaptor scaffoldins. Primary scaffoldins are present in all cellulosomal systems and are usually one of the most abundant cellulosomal components. They contain several cohesin domains which enable the attachment of cellulases, as well as a CBM domain which allows targeting of the scaffoldin to the substrate. The number of cohesin domains is variable, ranging from 8 to 12 (Artzi, Bayer, and Morais 2017).

In simple cellulosomal systems, the mechanism of cellulosome attachment to the cell wall has not been established except in the case of *R. cellulovorans*. Attachment of the cellulosome in *R. cellulovorans* is thought to take place via the enzyme EngE which contains a dockerin domain and an SLH layer domain (Kosugi et al. 2002).

In complex systems, an anchoring scaffoldin is also expressed, the method of the attachment of the scaffoldin to the cell surface varies. However, in most species including in *R. thermocellum*, the anchoring scaffoldin is attached via SLH domains (Q. Xu et al. 2016). An exception to this is in *R. flavefaciens* where the scaffoldin is attached to the cell surface via sortase-mediated attachment (Rincon et al. 2005). Regardless of the method of attachment to the cell wall, the primary scaffoldin attaches to the anchoring scaffoldin via a specialised cohesin dockerin pair.

Adaptor scaffoldins were first described in *R. flavefaciens* (Rincón et al. 2004). Adaptor scaffoldins can either connect two different scaffoldins or a scaffoldin and an enzyme. They are divided into two different classes; 1) Monovalent adaptor scaffoldins act to change the type of cellulases integrated into the cellulosome by changing the cohesin specificity of the primary scaffoldin. 2) Polyvalent scaffoldins act to expand the cellulosome by adding multiple scaffoldins, resulting in an increase in the cellulosome size and the integration of additional enzymes (Morais et al. 2016; Dassa et al. 2012).

1.4.3 Carbohydrate-binding domains

CBM's perform two functions in cellulosomes, the first is the attachment of the cellulosome to the cellulosic material, this is achieved by the scaffoldin CBM. Scaffoldins normally contain a family 3 CBM which binds selectively to cellulose (Morag et al. 1995). However, some cell-free cellulosomes contain family 2 CBM's which are categorised as being able to bind to cellulose or xylan (Simpson et al. 2000).

However, several cellulases in particular GH9 cellulases also contain CBM. The main role of the CBM is to increase the affinity of the cellulases to their substrate (Boraston et al. 2004). They have also been shown to be involved in the processivity of cellulases, for example the removal of the CBM can result in an inactive cellulase (Pan et al. 2016, 22).

1.4.4 Cellulases and hemicellulases

Cellulases and hemicellulases are classed as glycoside hydrolases, enzymes which hydrolyse the glycosidic bond between carbohydrates or a carbohydrate and non-carbohydrate. Hydrolysis can occur *via* overall retention or overall inversion of the anomeric configuration, depending on the spatial position of the catalytic residues (McCarter and Withers 1994). In inverting mechanisms, the reaction occurs in a single-displacement mechanism using acidic acid side chains (normally glutamic and aspartic acid) (McCarter and Withers 1994). In retaining mechanisms, a two-step, double displacement mechanism occurs. Again, acidic acid side chains, normally glutamate and aspartate act as acid/base donors but one also acts as a nucleophile (McIntosh et al. 1996).

Glycoside hydrolases have been classified according to their amino acid sequence similarity. To date, 145 different families have been identified. The CAZy database is

an up to date database of the families based on substrate specificity (Lombard et al. 2014).

Cellulases and hemicellulases have been shown to act synergistically. In addition, some cellulases have different modes of actions on the same substrate which can also result in synergistic effects. For example, endoglucanases cut in the middle of the cellulose chain, this results in two new sites for exoglucanases to act. In lignocellulose, multiple substrates are often linked together, and the action of one enzyme on one substrate could then allow access to the second substrate by a different enzyme (Eriksson, Karlsson, and Tjerneld 2002). Additionally, several cellulases are competitively inhibited by their product (Hsieh et al. 2014). The degradation of this product by another enzyme can remove this inhibition (Gefen et al. 2012).

Cellulolytic organisms secrete many different cellulases and hemicellulases. Genomic analysis has shown *R. thermocellum* encodes 72 dockerin containing genes, of which over half have been classified (Krauss, Zverlov, and Schwarz 2012). This large number of cellulases is needed as lignocellulose is diverse and different cellulases and hemicellulases have different activities on different substrates.

All known cellulosome-producing bacteria have a GH48 family cellulase which is usually expressed at very high levels and is normally the highest expressed component. There are usually only one or two copies of GH48 family cellulase genes in the genome (Koeck et al. 2014). Deletion of the GH48 family cellulase results in a strain which can either no longer grow on crystalline cellulose or has severely impaired growth. For example, deletion of the only GH48 family cellulase Cel48F in *R. cellulolyticum* resulted in a strain unable to grow on untreated biomass (T. Xu et al. 2014). While deletion of CelS, one of two GH48 family enzymes in *R. thermocellum*, lead to a strain with severely impaired growth on crystalline cellulose (Olson et al. 2010).

The GH48 family cellulase is classified as a processive cellobiohydrolase, and it is characterised by activity on amorphous or crystalline cellulose, with very low activity on CMC. Its activity on cellulose is low compared to other cellulases; it is thought to be important due to its synergy with other cellulases, in particular, those from the GH9 family (Sukharnikov et al. 2012). The GH48 family cellulase is equivalent to the processive cellobiohydrolases in the GH7 family produced by fungi and some cellulolytic bacteria (Barr et al. 1996).

GH9 family cellulases are normally also very highly expressed, conversely to GH48 family cellulases there are normally several copies in the genome. For example, in *R. cellulolyticum* 13 GH9 family cellulases are present in the genome, these all have different substrate specificities and activities. Additionally, they all display different levels of synergy with Cel48F, *R. cellulolyticum*'s GH48 family cellulase (Ravachol et al. 2014). It should also be noted that GH9 family cellulases often contain a CBM domain.

In addition to GH48 and GH9 family cellulases, GH5 family cellulases are also highly expressed; these are endoglucanases which often have activity on both cellulose and xylan. GH10, GH11 and GH43 are highly expressed hemicellulases, GH10 and GH11 are endo-xylanases while GH43 are arabinofuranosidases, these remove the arabinose side chains from arabinoxyylan (Scheller and Ulvskov 2010).

In addition to cellulases, several accessory proteins have also been identified. These include serpin-protease inhibitors which contain a dockerin domain; these have been identified in *R. cellulolyticum*, *R. thermocellum* and *C. cellulovorans*. They are thought to provide protection against both bacterial and plant proteases (T. Xu et al. 2014).

1.4.5 Designer cellulosomes

Designer cellulosomes have been built and used for studying cellulosome biology and to develop strains able to grow on lignocellulosic substrates.

In native cellulosomes, the arrangement of cellulases in a cellulosome is mainly determined by the amount of available enzyme, although a study by Borne et al. 2013 showed that in small trivalent cellulosomes there is a preference for certain combinations. This preference is thought to be caused by the length of the linker regions, with longer linker regions of 40 amino acids compared to the normal 9 eliminating any bias (Borne et al. 2013).

In designer cellulosomes, however, the arrangement of cellulases within the cellulosome is determined, by using a chimeric scaffoldin with divergent cohesin domains and enzymes with unique dockerin domains (Figure 1.5). Due to the modular nature of most cellulases and scaffoldins, as well as the species-specific interaction between cohesin/dockerin domains, this is relatively easy to achieve.

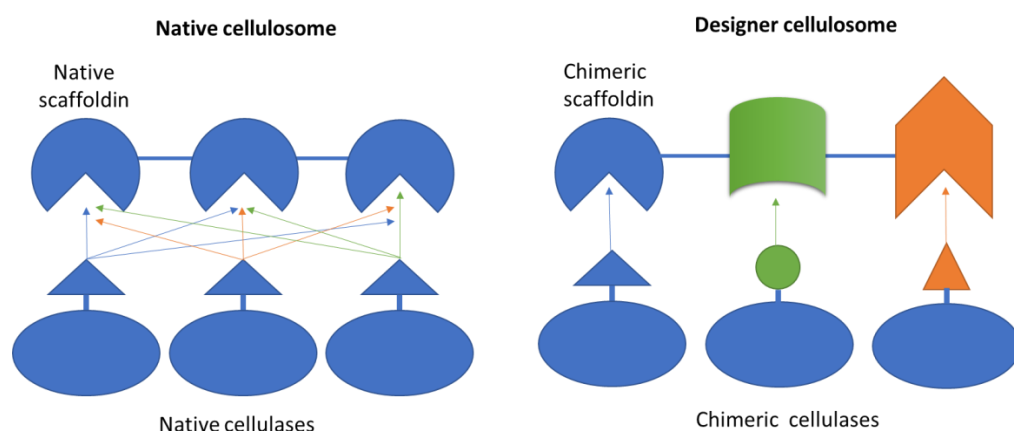


Figure 1.5. Assembly of native and designer cellulosomes. In *native cellulosomes* any dockerin domain-bearing enzyme can bind to any cohesion of the scaffoldin protein, meaning that the composition of the cellulosome is random. In *designer cellulosomes* however, each cellulase has a different dockerin domain which can only bind to a specific cohesion domain of the chimeric scaffoldin protein. This means that the composition of the cellulosome is ordered.

As the placement of cellulases within a cellulosome can have an impact on the degradation of cellulose, this results in higher levels of synergy between cellulases. In a small cellulosome consisting of a mini-scaffoldin and three *T. fusca* enzymes (9A, 48A and 5A), the position of the cellulases within the cellulosome had a significant effect on cellulase activity. With the cellulases in the 5A-48A-9A configuration significantly more activity was observed on crystalline cellulose than in the 48A-9A-5A configuration.

The structure of the scaffoldin can also have an impact on cellulosome efficiency; this was demonstrated by a study which showed that the length of the linkers between the different enzymes had a significant effect on cellulase activity (Vazana et al. 2013). It is thought that the length of the linkers between enzymes allows flexibility of the cellulosomes helping to prevent steric clashes and providing enhanced access to the cellulose substrate (Vazana et al. 2013). This indicates that the choice of scaffoldin can be as important as the cellulases expressed, and needs to be considered and adapted depending on the substrate and cellulases expressed.

Furthermore, studies have integrated enzymes not commonly found in cellulosomes into designer cellulosomes and improved their activity. An example of this is the integration of lytic polysaccharide monooxygenases (LPMOs) into a designer cellulosomes, which showed a 20-50% improvement in reducing sugars compared to an equivalent cellulosome without an LPMO (Agger et al. 2014). LPMO's increase cellulosome efficiency as they are able to cleave glycosidic bonds in crystalline regions of cellulose, while all other cellulases require an amorphous region. This creates a new ends for cellulases to attack (Villares et al. 2017). In another study, the integration of an engineered laccase (oxidising enzyme that acts on lignin) into a cellulosome, increased its activity on wheat straw (Davidi et al. 2016).

Several studies have focused on determining the most efficient combination of cellulases to degrade cellulose and lignocellulose. A recent study had the most efficient designer cellulosome to date with 70% of the activity of that of the *R. thermocellum*'s native cellulosome on wheat straw. It consisted of 8 enzymes, four cellulases and four xylanases. Due to issues with expressing scaffoldins with more than six cohesin domains, an adapter scaffoldin was used in addition to a scaffoldin. The adaptor scaffoldin consisted of a divergent cohesin to attach to the other scaffoldin, a CBM domain and three divergent cohesin domains. While the scaffoldin consisted of a divergent dockerin domain, to attach to the adaptor scaffoldin, a CBM and five divergent cohesins (Stern et al. 2016).

1.4.6 *R. cellulolyticum*'s cellulosome

R. cellulolyticum is a gram-positive mesophilic cellulolytic bacterium. It can efficiently degrade lignocellulose and naturally produces low levels of ethanol. It has been considered for CBP, however, it is currently unsuitable for industrial use as it has a slow growth rate. This is mainly caused by an inefficiently regulated carbon flow (Desvaux 2005a).

Although initially classified in the class Clostridia, it was then reassigned to the family *Ruminococcaceae* due to similarity with *Ruminococcus albus*. However, it has recently been reclassified again into a new genus *Ruminiclostridium* along with other cellulolytic clostridia including *R. thermocellum* (Yutin and Galperin 2013).

In *R. cellulolyticum* the cellulase genes and cellulosomal enzymes are found in operons and clusters of genes. This arrangement is highly conserved in several other mesophilic species including *C. cellulovorans*, *C. papyrosolvans*, *C. josui* and *C. acetobutylicum* (Desvaux 2005b; Tamaru et al. 2000; Sabathé, Bélair, and Soucaille 2002). The main operon is called the Cip-Cel operon, and it contains the scaffoldin gene *cipC* along with

the major cellulase genes. In *R. cellulolyticum* disruption of the Cip-Cel operon results in a strain with severely impaired growth on crystalline cellulose (Maamar et al. 2004). It also contains a xyl-doc cluster of genes involved in hemicellulose degradation. Several other small clusters of cellulase and hemicellulase genes have also been identified (C. Xu et al. 2013).

R. cellulolyticum produces a simple cellulosomal system consisting of a single scaffoldin protein, CipC. This contains eight cohesin domains, a CBM domain and an X domain (Sandrine Pagès et al. 1999). *R. cellulolyticum* has been observed to attach directly to cellulose, however the method of attachment has not yet been elucidated. In contrast to other cellulolytic bacteria, including *R. thermocellum* and *R. cellulovorans*, it is unlikely that *R. cellulolyticum*'s cellulosomes are attached to the cell surface. This is likely, as no method of cellulosome attachment has been identified and cell protrusions have not been observed as is the case for *R. thermocellum* (Ferdinand et al. 2013). Further evidence of a cell-free cellulosome in *R. cellulolyticum* is that the cellulosome is not attached to the cell when grown on xylan (Mohand-Oussaid et al. 1999).

In *R. cellulolyticum*'s genome, 62 dockerin-containing genes have been identified. These consist of cellulases, hemicellulases (including xylanases, mannanases and arabinofuranosidases), pectate lyases and chitinases. Cellulosomal genes are regulated by a carbon catabolite expression system with different sets of genes expressed depending on the carbon source (C. Xu et al. 2013). On crystalline cellulose, Cel48F (GH48 family cellulase) and Cel9E (GH9 family cellulase) are the most abundant cellulases, they are expressed at the same level as the scaffoldin protein CipC. Cel5A, Cel5M, Cel9G, Cel8A, Xyn10A and Cel9M are also expressed at relatively high levels (Fendri et al. 2009).

1.5 *Clostridium acetobutylicum*

1.5.1 Introduction to *C. acetobutylicum*

Solventogenic clostridia are gram-positive, endospore-forming, obligate anaerobes which produce primary metabolites for example, butanol from both simple and complex carbohydrates.

Clostridium acetobutylicum is a gram-positive, spore-forming, saccharolytic, obligate anaerobe. It is able to produce high levels of acids and solvents via the ABE (Acetone-Butanol-Ethanol) fermentation pathway. It is a member of the group 1 of the genus clostridia; this is considered the core genus of clostridia and includes the type strain *C. butyricum* as well as other solventogenic clostridia (Lawson and Rainey 2016).

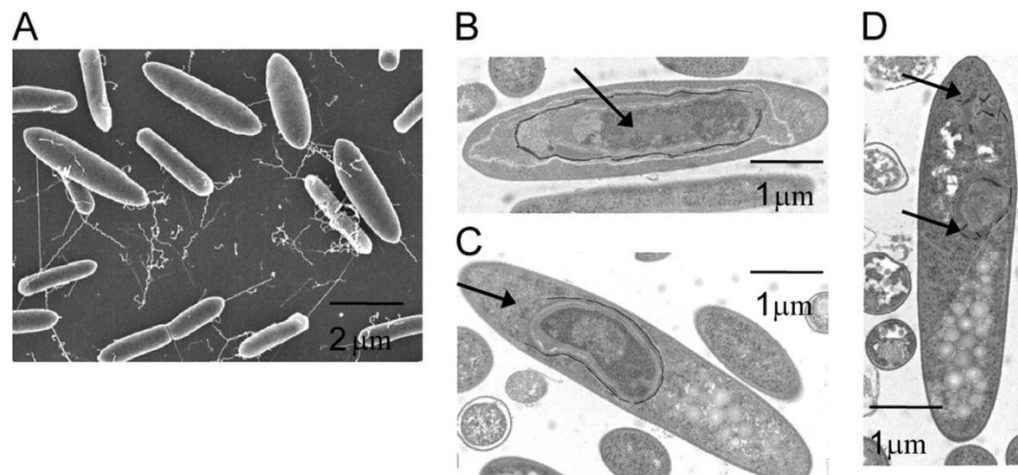


Figure 1.6. **A)** Light microscope picture of *C. acetobutylicum* cells **B)** TEM picture of a *C. acetobutylicum* cell. **C)** TEM picture of a *C. acetobutylicum* cell starting to form an endospore. **D)** TEM pictures of a *C. acetobutylicum* endospores. Diagram adapted from (Tracy, Jones, and Papoutsakis 2011).

C. acetobutylicum was first isolated by Chaim Weizmann, and used for acetone production during the First World War (Weizmann, C 1915). The type strain, *C. acetobutylicum* ATCC 824, was isolated from soil from a Connecticut garden in 1924, and it has been shown to closely resemble the original Weizmann strain, which has

since been lost. *C. acetobutylicum* strains have been found in diverse environments, in soil samples, and the digestive tracts of both animals and humans.

C. acetobutylicum and other related strains including *C. beijerinckii*, *C. saccharoperbutylacetonicum* and *C. saccharobutylicum* were used industrially to produce butanol and acetone into the 1940's, however by the 1950's crude oil provided a cheaper alternative. There was renewed scientific interest in the ABE fermentation process in the 1980's due to an oil crisis, with various strains characterised and the technical side of fermentation investigated (Jones and Woods 1986).

Scientific research received another revival recently when British Petroleum (BP) announced it was going to produce biobutanol *via* the ABE process in 2006. Several biotech companies are also interested in utilising clostridia to produce butanol, for example, Green Biologics (Green, 2011). Currently, there is some production of butanol via ABE fermentation process in China, with over \$200 million dollars invested in 6 plants producing 300,000 tonnes of butanol from cornstarch (Jiang et al. 2015). Green Biologics have also opened a plant in the USA supplying green butanol and acetone from a *Clostridium* chassis ('Bio-Based Chemicals | Renewable Chemicals' n.d.).

C. acetobutylicum is also used as a model for endospore formation and has been compared to *Bacillus subtilis*, which is a well-studied endospore-forming bacterium (Tan and Ramamurthi 2014). The study of endospore formation is important as many pathogens can persist or contaminate through endospores including: *Clostridium difficile* and *Clostridium botulinum* (Barra-Carrasco and Paredes-Sabja 2014).

1.5.2 *C. acetobutylicum* metabolism

C. acetobutylicum is a chemoorganotrophic organism obtaining its energy by substrate phosphorylation during fermentation. This means it requires a carbohydrate source for energy as well as a carbon source. It can utilise all major monosaccharides as well as some polysaccharides. Some strains can grow on cellobiose, and it produces small amounts of xylanase (Cooksley et al., 2012).

C. acetobutylicum has a unique life-cycle, it does not continuously produce solvents, instead during exponential phase it produces acetic and butyric acid from available carbon sources. During acidogenesis, the pH of the media drops, and at a pH of around 4.5 there is a metabolic switch towards solvent production. During the solventogenic phase, the acids produced are re-assimilated and converted to solvents, butanol and acetone. A small amount of ethanol is also produced (S. Y. Lee et al. 2008). During solventogenesis, cells also accumulate granulose which is a prelude to endospore formation (Janssen et al. 2010). A diagram showing the simplified metabolic pathway of *C. acetobutylicum* is shown in Figure 1.7.

There are two major issues with *C. acetobutylicum* fermentation; Firstly, the cultures can acid crash and cells die before they can switch to solventogenesis. This can happen when during acidogenesis the cells produces excess amount of acids and fail to switch to solventogenesis in time (Maddox et al. 2000). Secondly, the strains can degenerate, by losing their pSOL1 plasmid, which contains most of the essential genes for solvent formation. Degenerative strains can be easily identified by their inability to form spores, degrade starch or produce solvents (Cornillot et al. 1997).

type strain to ethyl methanesulfonate could grow on xylan as the sole carbon source (S. F. Lee, Forsberg, and Gibbins 1985b).

In 2001 *C. acetobutylicum*'s genome was sequenced and revealed that *C. acetobutylicum* has a large gene cluster corresponding to cellulosomal genes, similar to that found in *R. cellulolyticum* (Figure 1.8) (Nölling et al. 2001). The operon encodes a gene for the scaffoldin protein CipA, OrfX a membrane-bound protein with a cohesin domain, thought to play a role in cellulosome assembly (Sandrine Pagès et al. 1999) and several cellulases including: a GH48 family cellulase (Cel48A), several GH9 family cellulases (Cel9X, Cel9C, Cel9F, Cel9J and Cel9E), two GH5 family cellulases (Cel5B and Cel5Y) and a mannanase (Man5G). In addition to this operon, several xylanases are present elsewhere in the genome.

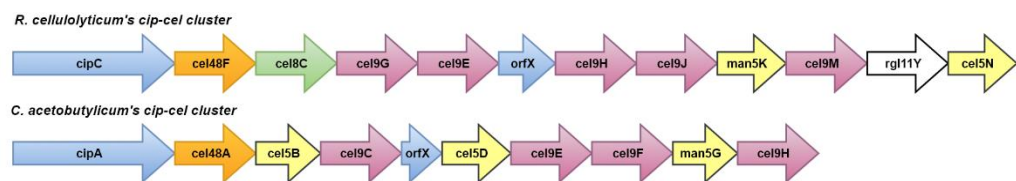


Figure 1.8. Schematic representation of the Cip-Cel cellulosomal operons of *R. cellulolyticum* and *C. acetobutylicum*. Both *R. cellulolyticum* and *C. acetobutylicum* have a similar cip-cel operon structure. The primary scaffoldin protein is the first protein in the operon, they also both have an OrfX gene which consists of only a single cohesin (blue). They both contain a single GH48 family cellulase (orange) as well as several GH9 family cellulases (red) and mannases (yellow). *R. cellulolyticum*'s cip-cel operon also has a GH8 family cellulase (green).

1.6.1 Potential reasons why *C. acetobutylicum* cannot utilise cellulose

Despite the presence of a functional cellulosomal operon, *C. acetobutylicum* cannot grow on cellulose. Western blot analysis utilising antibodies raised against *R. cellulolyticum*'s cellulosomal proteins showed that CipA and several cellulases including Cel9X (homologous to Cel9E), Cel48A and Cel9C/Cel9E (homologous to Cel9G) were expressed (Sabathé, Bélai"ch, and Soucaille 2002). Cel48A has a 58%

amino-acid sequence identity with Cel48F. Cel9X also shows a 52% amino-acid sequence identity with Cel9E from *R. cellulolyticum*, and Cel9C has a 60% amino-acid sequence identity with Cel9G. *C. acetobutylicum* was also shown to form a cellulosome, which included the inactive Cel48A. However, no activity was detected towards crystalline cellulose for either the purified cellulosome or concentrated culture supernatant (Avicel and bacterial cellulose) (Sabathé, Bélai"ch, and Soucaille 2002). Low levels of activity were detected on CMC and PASC for the purified cellulosome and ten times higher values for concentrated supernatant. The activity on CMC and PASC was found to be very low compared to *R. cellulolyticum*'s purified cellulosome (Sabathé, Bélai"ch, and Soucaille 2002).

In another study, it was shown that *C. acetobutylicum* produced the highest level of cellulolytic activity on lichenin followed by xylose, cellobiose then glucose, indicating the presence of a carbon catabolite repression system. Lichenin is a beta-glucan with both β -1-4 and β -1-3 glycosidic bonds (López-Contreras et al. 2004). Additionally, the expression of Cel48A was also determined to be increased in lichenin cultures compared to xylose, cellobiose and glucose. The same study showed that *C. acetobutylicum* was able to grow on Avicel supplemented with a cellulose enzyme cocktail, but was unable to grow on CMC also supplemented with a cellulose cocktail presumably due to an inability to process the carboxymethyl groups (López-Contreras et al. 2004).

Purified Cel9G from *C. acetobutylicum* produced from *E. coli* showed activity towards CMC, PASC and xylan, however Cel9G could not be detected in *C. acetobutylicum*'s supernatant (López-Contreras et al. 2003). Expression and purification of a mini-scaffoldin in *C. acetobutylicum* in which Cel48A was the main component showed no activity towards CMC, PASC, Avicel or Bacterial cellulose (Sabathé and Soucaille 2003).

Several key cellulolytic components of the cip-cel operon have since been shown to be active, including Cel9C and Cel9X equivalent to *R. cellulolyticum*'s Cel9G and Cel9E respectively. However, Cel48A was shown to be completely inactive on all substrates tested (Phillipe Soucaille, personal communication). A schematic representation of the domain structure of GH48 and GH9 families from *R. cellulolyticum* and *C. acetobutylicum* is shown in Figure 1.9. An amino acid sequence alignment between Cel48F and Cel48A is also shown in Figure 1.10.

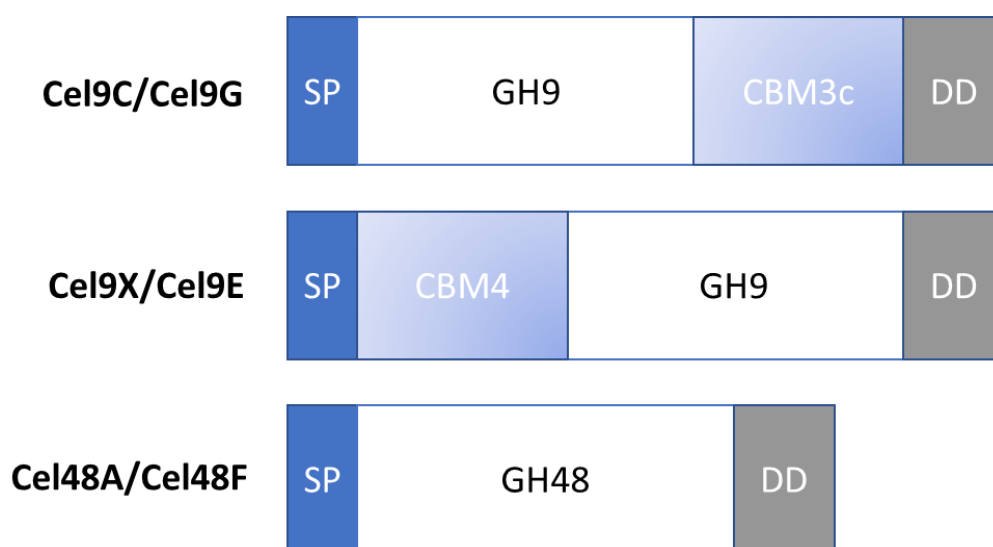


Figure 1.9. Domain structure of cellulases from *C. acetobutylicum* and *R. cellulolyticum*. Cel9C from *C. acetobutylicum* and Cel9G from *R. cellulolyticum* both consist of a signal peptide, followed by a GH9 catalytic domain, CBM3c and dockerin domain. Cel9X from *C. acetobutylicum* and Cel9E from *R. cellulolyticum* both consist of a signal peptide, CBM4, GH9 catalytic domain and a dockerin domain. Cel48A from *C. acetobutylicum* and Cel48F from *R. cellulolyticum* both consist of a signal peptide, GH48 catalytic domain and a dockerin domain.

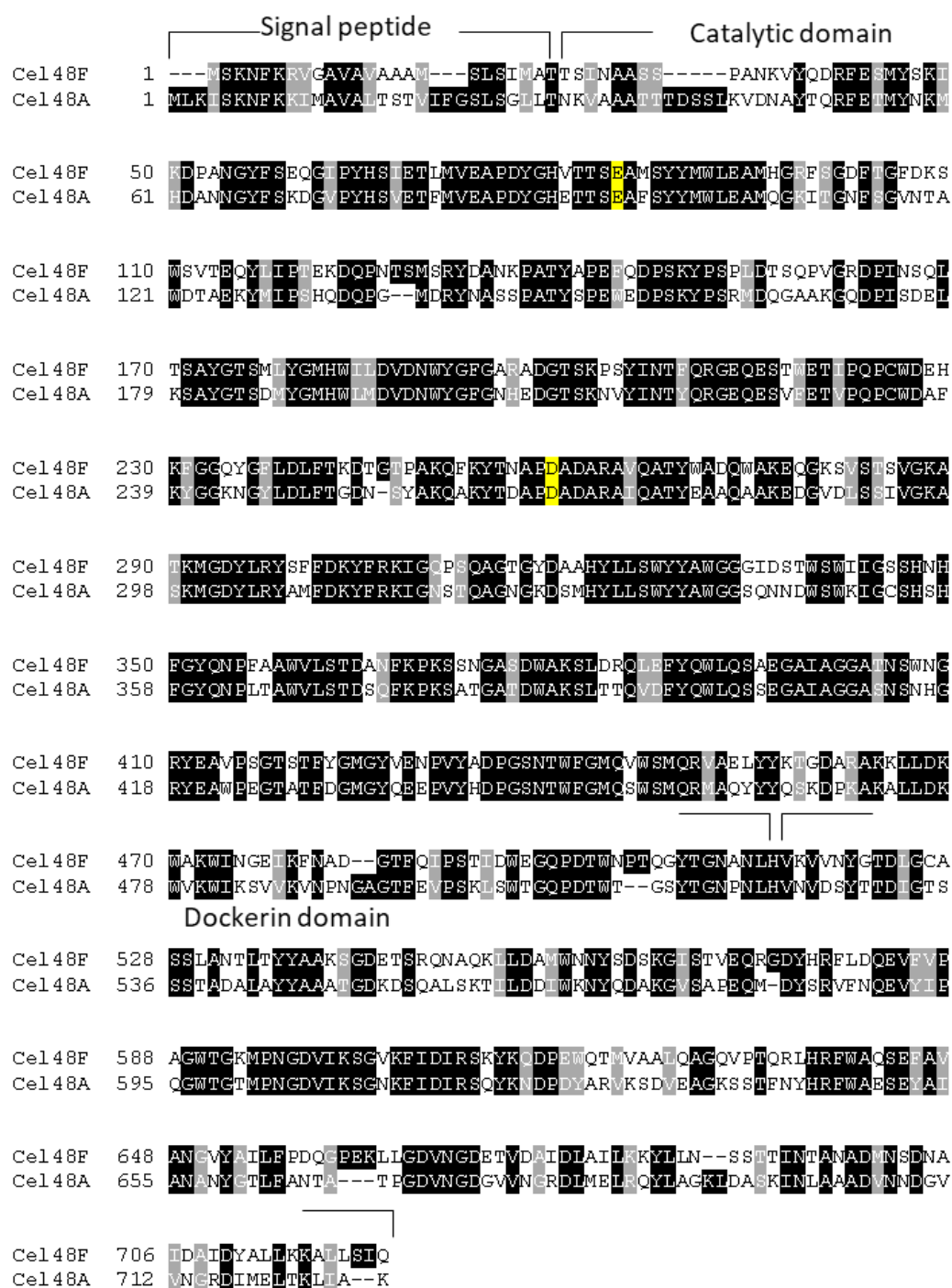


Figure 1.10. Sequence alignment of the GH48 cellulase Cel48A from *R. acetobutylicum* and Cel48F from *R. cellulolyticum*. The two catalytic residues Glutamic acid and Aspartate are highlighted.

The non-functional Cel48A was found to be the main component of *C. acetobutylicum*'s native cellulosome, which could at least partially explain the cellulosomes low activity towards cellulosic substrates. The lack of activity for Cel48A might be the main reason for *C. acetobutylicum*'s inability to grow on cellulosic substrates. It is unclear when or how *C. acetobutylicum* lost its cellulolytic activity. Maintenance of the strain in laboratory conditions, growing the strain in rich media, in the absence of selection pressure might have caused the loss of its cellulolytic activity. It is more likely that *C. acetobutylicum* has lost its cellulolytic ability prior to isolation, as no other *C. acetobutylicum* strain or other solventogenic clostridia have been identified to be able to grow on cellulose (Poehlein et al. 2017).

1.6.2 Engineering *C. acetobutylicum* to grow on lignocellulose

C. acetobutylicum makes an ideal candidate for CBP for several reasons: (1) it has been used industrially to produce butanol, (2) it can utilise all pentose and hexose sugars present in lignocellulosic biomass, (3) already produces a non-functional cellulosome and (4) it has moderate xylanase activity.

Earlier attempts at engineering *C. acetobutylicum* to grow on lignocellulose were focused on the heterologous expression of cellulosomal components derived from known cellulolytic species, such as *R. cellulovorans*, *R. cellulolyticum* and *R. thermocellum*.

Initial studies looked at the expression of scaffoldin components derived from *R. cellulolyticum* in *C. acetobutylicum*. Perret et al described the successful expression of two mini-scaffoldins: miniCipC1, a truncated form of CipC from *R. cellulolyticum* and Scaf3, a designer scaffoldin consisting of mini-CipC1 and a cohesin domain from *R. thermocellum*. Both scaffoldins were readily secreted by *C. acetobutylicum*, and both the cohesin and CBM domains were shown to be functional (Perret et al. 2004).

Other studies describe the expression of individual cellulases. Kim et al. (1994) described the heterologous expression of EngB, a GH5 family cellulase from *C. cellulovorans* in *C. acetobutylicum*. EngB was successfully secreted, and the strain showed increased activity towards CMC and PASC. However, the strain was still unable to grow on crystalline cellulose (Kim et al. 1994).

In another study by Mingardon et al. (2005), the mannase gene *man5K*, from *R. cellulolyticum* was heterologously expressed in *C. acetobutylicum*. A weakened thiolase promoter was used for *man5K* expression, as expression under control of the native thiolase promoter could not be achieved. This could suggest that high level of mannase expression is toxic to the cells. Interestingly, degradation of mannase at the dockerin domain was observed. Furthermore, a mini-operon consisting of miniCipC1 and *man5K* was constructed and showed successful formation with *man5K* almost exclusively in complex with minCipC1 (Mingardon et al. 2005).

A later study looked at the expression of several cellulases from *R. cellulolyticum* in *C. acetobutylicum*, including: Cel5A, Cel8C, Cel9E, Cel9G, Cel9M and Cel48F (Mingardon et al. 2011). Initially, all cellulases were cloned downstream of *C. acetobutylicum*'s *thiolase* promoter. Cel5A, Cel8C and Cel9M were all found to be secreted and were active, although degradation of their dockerin domain was also observed. Constructs carrying the genes for Cel9E, Cel9G and Cel48F under the native, constitutive *thiolase* promoter or the weaker mutated *thiolase* promoter, could not be transformed into *C. acetobutylicum*.

The authors suggested that the toxicity of several cellulases, including Cel48F, might occur during secretion, as removal of the signal peptide resulted in a strain able to heterologously express Cel48F at a relatively high level. Transformation of Cel48F under control of the mutated thiolase promoter was then achieved when the

transformation was performed at 30 °C instead of 37 °C. However, only very low levels of Cel48F could be detected in the supernatant and this phenotype was lost after revival from a spore stock (Mingardon et al. 2011).

To improve the expression and secretion of cel48F, several strategies were tested, including: co-expression of Cel48F with the scaffoldin genes miniCipC1 and co-expression with the OrfXp protein, replacement of Cel48F signal peptide with different signal peptides including CipC and *C. acetobutylicum*'s native Cel48A signal peptide. None of these modifications resulted in viable colonies after transformation, suggesting that they had no impact on relieving Cel48F's toxicity (Mingardon et al. 2011).

Surprisingly, cloning of the *C. acetobutylicum*'s GH48 family cellulase, Cel48A, downstream of the *thiolase* promoter did not result in any colonies, indicating that overexpression of the native GH48 cellulase is also toxic (Mingardon et al. 2011). It should be noted that high levels of Cel48A have been detected in wild-type *C. acetobutylicum* (Sabathé and Soucaille 2003).

A later study looked at using carrier domains to increase the secretion of cellulases, specifically Cel48F in *C. acetobutylicum* (Chanal et al. 2011). In this study, the N-terminal CBM plus the first two X2 domains and a scaffoldin domain from CipC were added downstream of the Cel48F catalytic domain and this enabled the transformation and secretion of Cel48F. The presence of the scaffoldin domain was found to be optional, while the X2 domain was essential. Adding the CBM3a and X2 module to Cel9G also enabled transformation and secretion of Cel9G. It was also found that swapping the X2 module from *R. cellulolyticum* for an X2 module from *C. acetobutylicum* CipC (XA) resulted in higher protein yields, with two X2 domains further increasing the yield. The CBM3a and X2 domain modified Cel48F and Cel9G

activity increasing the activity on Avicel while decreasing it on PASC (Chanal et al. 2011; Fierobe, Mingardon, and Chanal 2012). The expression yields of these fusion proteins were still too low to enable growth on lignocellulose.

Subsequent studies focused on the integration of cellulosomal genes into the genome of *C. acetobutylicum*. Integration of the genes into the hosts genome were thought to at least partially alleviate the toxicity issues seen when multi-copy plasmids were used. In addition, integration into the genome confers strains stability, circumventing the loss of cellulase expression after revival from a spore stock.

The first study which used the genome integration method to create stable cellulolytic *C. acetobutylicum* strains looked at the expression of *R. thermocellum* cellulases, specifically Cel8A, Cel9B, Cel48S and Cel9K as well as truncated forms of CipA (*R. thermocellum*'s scaffoldin protein) (Kovács et al. 2013). Relatively high expression of Cel8A, Cel9B and Cel9K was achieved, however, Cel48A could only be detected when *C. acetobutylicum* cultures were grown in buffered medium. Cel9B, Cel8A and Cel9K were also proven to be active on 4-NPC and cellohexaose. This study confirmed that the integration of cellulases into the genome was a valid strategy for expression of cellulases in *C. acetobutylicum*.

In addition to secretion of all major cellulolytic enzymes by genome integration, a mini-cellulosome, consisting of CipA mini-scaffoldin and two cellulases Cel8A and Cel9B were also successfully secreted and these were shown to be active on cellohexaose. The cellulases were expressed as an operon, and the placement of cellulases/scaffoldins within the operon had a large effect on expression of the components (Kovács et al. 2013). This is probably due to changes in the mRNA structure. Thus optimisation of proteins localisation within the operon may be

necessary, alternatively expression of each cellulase could be driven by their own promoter.

In a more recent study, two cellulases and a xylanase from *R. cellulolyticum* (Cel48F, Cel9G and XynX), along with a cell-wall anchored mini-scaffoldin were successfully secreted and assembled on the surface of *C. acetobutylicum* (Willson et al. 2016). Although the strain showed improved growth on Xylan and released higher levels of reducing sugars than the wild-type strain it was still unable to grow on lignocellulose. The strain showed other improved properties, for example in a Xylan fermentation, the strain produced twice as much butanol as the wild-type control strain.

The strains inability to grow on cellulose was most likely due to the low expression and secretion of cellulases, in particular the GH48 cellulase, Cel48F. Improvement of the expression and secretion of GH48 enzymes will be essential in engineering *C. acetobutylicum* to grow on cellulose (Willson et al. 2016).

In conclusion, significant progress has been made in engineering *C. acetobutylicum* to grow on lignocellulose with complex operons constructed and integrated into the genome. Relatively simple scaffoldins have been expressed, and anchored to the cell wall. Mini-cellulosomes consisting of only three cellulases have been shown to be able to degrade crystalline cellulose. The use of multiple scaffoldins could also be a viable strategy, where one scaffoldin is anchored to the cell wall (anchoring scaffoldin), and secondary scaffoldins are also expressed to attach to the anchored scaffoldin.

Expression of all major cellulases has been achieved, however only low expression of large cellulases, in particular, enzymes from the GH9 and GH48 family has been achieved. These have been shown to be essential for degradation of crystalline cellulose. In order to engineer *C. acetobutylicum* to grow on lignocellulose the reason(s) for the low expression of these cellulases has to be determined.

1.7 Genetic tools for clostridia

1.7.1.1 Plasmids

A set of shuttle plasmid's (pMTL80000) has been developed for use in clostridia by Heap et al (Heap et al. 2009). These plasmids contain four modules: a selection marker, a gram-negative replicon, a gram-positive replicon and an application specific module. They are suitable for cloning of any DNA fragment into the application specific module and shuttling the plasmid from its cloning host (*E.coli*) to the heterologous expression host (Clostridia).

1.7.1.2 Promoters

Genes involved in central metabolic metabolism were hypothesised to be highly expressed, therefore the *adc*, *ptb* and *thiolase* promoters are often used (Feustel, Nakotte, and Dürre 2004). Inducible promoter systems have also been developed and used in clostridia, including, a xylose inducible system, two lactose inducible systems as well as an anhydrotetracycline inducible system.

The xylose inducible system (*xylA*) is based on the *Staphylococcus xylosus*'s, xylose regulatory system. The use of this system is limited as strains have to grow on xylose which is a non-optimal carbon source (Girbal et al. 2003). One of the lactose inducible systems is based on the *lacI*-repressed *fac* promoter, the strong ferredoxin promoter of *C. pastorianum*, and contains the symmetrical operator sequence from *E. coli*'s *lacZ* gene (Heap et al. 2007). The second lactose inducible system is making use of the *PbgaI* and its regulator *BgaR* from *C. perfringens* (Al-Hinai, Fast, and Papoutsakis 2012). The anhydrotetracycline inducible system is based on the Tet-inducible system from *E. coli*'s Tn10 transposon (Dong et al. 2012).

1.7.1.3 Gene inactivation and chromosomal insertions

In Clostridia, techniques have recently been developed which allow both gene deletion and integration onto the chromosome. Integration of heterologous genes into the host's genome confers strain stability, and reduced metabolic burden in the absence of replicating plasmids, making these strains more suitable for industrial applications.

The first gene activation method developed for clostridia was through the use of mobile group 11 introns. These are catalytic RNA's which contain a sequence recognition motif and are self-splicing. Alteration of a few bases can result in sequence recognition of lots of chromosomal sites. The ClosTron system uses an element derived the broad host range LI.LtrB intron of *Lactococcus lactis* and clones can be easily isolated using a special marker called a retrotransposition-activated marker (RAM), which after intron insertion results in resistance to erythromycin (Kuehne and Minton 2012). However, ClosTron knockout can cause polar effects due to the integration of the type 11 intron and may be intrinsically unstable (Steiner et al. 2011). To allow for multiple deletions in the same strain, flanking FRT sites were included adjacent to the erythromycin resistant marker allowing excision of the marker using FLP recombinase (Kuehne and Minton 2012).

Allele-coupled exchange (ACE), used for integrating heterologous genes into the clostridia genome, was developed by our group (Heap et al., 2012). In this, the first recombination event is defined by the use of a larger homology arm, and single crossover recombinants are selected based on the antibiotic resistance of the plasmid, which at this point is integrated into the genome. During the second recombination event, the plasmid backbone is lost by applying a selective pressure, either by switching a second selection marker on after chromosomal insertion, or by restoring an auxotrophic phenotype. For example, the *pyrE* gene which encodes for orotate

phosphoribosyltransferase is essential for pyrimidine synthesis. Cells without *pyrE* can be selected for as they are unable to grow on medium lacking uracil, while cells with *pyrE* are unable to grow on media containing 5-Fluoroorotic Acid (5-FOA), as 5-FOA is a toxic analogue to Uracil. Additionally, the asymmetric homology arm has been shown to favour a double cross-over event in place of a reversion to wild-type.

A marker-less genetic exchange system was also developed by (Christian Croux et al. 2016), similar to ACE this utilises antibiotic selection for single crossover selection. The antibiotic resistance gene is eliminated using the flippase recombinase system from *Saccharomyces cerevisiae*. The UPP gene (which encodes for uracil phosphoribosyltransferase) was used as a counter-selection marker.

1.8 Aims

The overall aim of this study is to engineer a strain of *C. acetobutylicum* that is suitable for use as a CBP organism, i.e. capable of utilising lignocellulose and producing solvents.

Previous studies have shown that the main barrier to engineering *C. acetobutylicum* to grow on lignocellulose is the low expression of cellulases from the GH9 and GH48 families. For this reason, the primary aim of this project was to increase the expression of GH48 and GH9 family cellulases.

In this work, several attempts were made to increase the expression level of GH48 family cellulases, in particular Cel48F, the mesophilic GH48 family cellulase from *R. cellulolyticum* in *C. acetobutylicum*. The strategies used included: the use of an inducible promoter system, addition of five and three prime UTR's to stabilise mRNA transcripts and the expression of chaperones.

Interaction of active and inactive Cel48F with *C. acetobutylicum* cell wall components was assessed, and we have found that Cel48F is able to specifically bind to N-deacetylated peptidoglycan but not N-acetylated. During this study, the structure of *C. acetobutylicum*'s peptidoglycan was determined using LC-MS/MS, and we found that both MurNAc and GlcNAc residues are at least partially N-deacetylated. Two potential N-deacetylase genes were identified and inactivated, however the N-deacetylation levels of the deletions strains remained unchanged, suggesting that either other genes are involved in the de-acetylation of *C. acetobutylicum*'s peptidoglycan or multiple gene deletions are necessary to achieve a higher acetylation levels of both residues.

Co-expression of OrfXp, a small protein containing a cohesin domain and believed to be anchored to the cell membrane, was shown to significantly improve the secretion levels of GH48 family cellulases in *C. acetobutylicum*, by a yet unidentified mechanism. This previously uncharacterised protein, for which homologs were found in all cellulosomal systems, is probably essential for the secretion of GH48 cellulases in all cellulosomal, cellulolytic bacteria.

Chapter 2 Materials and Methods

2.1 Materials and suppliers

A full list of materials and reagents can be found in Appendix 1.1.

2.2 Preparation of growth media

2.2.1 Preparation of liquid growth media

Standard lysogeny broth (LB) medium was prepared with a concentration of 10 g/L sodium chloride, 5 g/L yeast extract, and 10 g/L tryptone.

2xYTG medium was prepared with a final concentration of 15 g/L tryptone, 10 g/L yeast extract, 4 g/L sodium chloride. Prior to autoclaving, solution was pH balanced to 5.2 and adjusted to 90% of the desired final volume. Sterile 50% glucose was added subsequent to autoclaving to the desired concentration (normally 0.5%)

Clostridia basal medium (CBM) was prepared with a final concentration of 4 g/L casein enzymatic hydrolysate (N-Z-Amine® A), 0.2 g/L magnesium sulfate heptahydrate, 0.01 g/L manganese sulfate tetrahydrate, 0.01 g/L iron (II) sulfate heptahydrate, 1 mg/L p-aminobenzoic acid, 2 µg/L biotin, 1mg/L thiamine, and 10 g/L glucose. After autoclaving, CBM was buffered by addition of potassium phosphate monobasic and potassium phosphate dibasic, both to a final concentration of 500 mg/L. CBM-MOPS was prepared by adding MOPS to a concentration of 40 mM and adjusting the pH to 6.9 prior to autoclaving. Sterile 50% glucose was added subsequent to autoclaving to the desired concentration (normally 0.5%)

2.2.2 Preparation of solid growth media

CGM-agar was prepared with a concentration of 2 g/L ammonium sulfate, 1 g/L potassium phosphate dibasic, 0.5 g/L potassium phosphate monobasic, 0.1 g/L magnesium sulfate heptahydrate, 30 mg/L iron sulfate heptahydrate, 20 mg/L calcium chloride, 20 mg/L manganese sulfate monohydrate, 0.01 g/L cobalt chloride, 0.002 g/L zinc sulfate, 2 g/L tryptone, 1 g/L yeast extract, 50 g/L glucose and 12 g/L agar no. 1.

LB-agar (prepared as LB medium but containing 15 g/L agar no. 3), CBM-agar (prepared as liquid CBM but containing 10 g/L agar no. 1) and CGM-agar were melted by microwaving at low power and allowed to cool to 50°C in a water bath. Potassium phosphate and glucose solutions were added to molten CBM-agar as for liquid CBM. Antibiotic supplements were added to the indicated working concentrations, and the melted agar immediately poured into 9 cm Petri dishes, with approximately 20 ml per plate. Plates were dried at 37°C for 1 hour before using.

2.2.3 Preparation of antibiotics

The following antibiotics were used in this study

Antibiotic	Stock solution	Concentration (<i>E. coli</i>)	Concentration (<i>C. acetobutylicum</i>)
Ampicillin	50 mg/ml in dH ₂ O	100 µg/ml	N/A
Kanamycin	50 mg/ml in dH ₂ O	50 µg/ml	N/A
Zeocin	100 mg/ml as provided	100 µg/ml	N/A
Tetracycline	5 mg/ml in 100% ethanol	5 µg/ml	N/A
Chloramphenicol	25 mg/ml in 100% ethanol	25 µg/ml in plates, 12.5 µg/ml in liquid	N/A
Erythromycin	50 mg/ml in 100% ethanol	500 µg/ml	50 µg/ml
Thiamphenicol	15 mg/ml in 50% ethanol	N/A	15 µg/ml
Fluoroorotic acid	40 mg/ml in DMSO	N/A	400 µg/ml

Table 2.1. Antibiotics used in the study. All antibiotics were prepared as stock solutions and added to the growth medium to give the final concentrations as described above.

2.3 Engineering of *E. coli*

2.3.1 List of *E. coli* strains

This section contains a description of the *E. coli* strains used in this study, with a description of their genotypes.

Strain	Characteristics	Source
Top10	F- mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80lacZ Δ M15 Δ lacX74 recA1 araD139 Δ (ara-leu)7697 galU galK rpsL endA1 nupG	Life Technologies
pAN-2	As Top10 but bearing pAN-2 methylation plasmid	Dr Steve Cartman
BL21 (DE3)	fhuA2 [lon] ompT gal (λ DE3) [dcm] Δ hsdS λ DE3 = λ sBamHIo Δ EcoRI-B int::(<i>lac</i> ::PlacUV5::T7 gene1) i21 Δ nin5	NEB
BL21(DE3) - pET48Fc	As BL21 (DE3) but bearing the pET48Fc plasmid	This study
BL21(DE3) -pET16B	As BL21 (DE3) but bearing the pET16B plasmid	This study

Table 2.2. List of *E. coli* strains used in this study.

2.3.2 Storage of *E. coli* strains

All *E. coli* strains were stored at -80°C through the use of Microbank™ vials or glycerol stocks. A liquid culture of the strain of interest was streaked onto a plate of LB-agar with the relevant antibiotics and incubated overnight at 37°C unless otherwise stated. Cells were scraped from the plate and used to inoculate three Microbank™ vials, or a solution of sterile 20% glycerol.

2.3.3 Culture of *E. coli* cells

Liquid cultures of *E. coli* were prepared by inoculation of 8 ml LB medium supplemented with antibiotics with either a single bead from a Microbank™ stock or with single colonies grown on LB-agar plates. Cultures were incubated overnight (approx. 16 hours) at 37 °C with shaking at 200 rpm, if cultures contained a toxic gene they were incubated at 30 °C or room temperature, for a longer period of time instead.

E. coli cultures on solid medium were prepared by a variety of methods. For generation of biomass (e.g. for generation of -80°C stocks), a 10 µl loop of growth medium from a monoclonal liquid culture of the strain was struck onto LB-agar medium as prepared above. For isolation of single colonies (e.g. after following electroporation), cell suspension was pipetted onto an LB-agar plate and immediately spread using a sterile plastic spreader until dry. Inoculated plates were typically incubated overnight (approx. 16 hours), although as above, certain strains required incubation at 30°C or room temperature for a longer period of time.

2.3.4 Preparation of *E. coli* electrocompetent cells

An overnight culture of *E. coli* was prepared by addition of a single bead from a Microbank™ vial to a 20 ml universal tube containing 8 ml LB medium supplemented with the relevant antibiotics. The culture was incubated overnight (approx. 16 hours) at 37°C with shaking at 200 rpm. 2 ml overnight culture was added to a conical flask containing 200 ml LB medium supplemented with the relevant antibiotics and pre-warmed to 37°C. Flasks were incubated at 37°C with shaking at 200 rpm until the cultures had reached an OD₆₀₀ of 0.6-0.8 (not more than 4 hours).

Each culture was equally divided into 50 ml centrifuge tubes and centrifuged for 10 minutes at 3000 xg. Supernatants were discarded and pellets were washed twice with 30 ml sterile dH₂O prechilled to 4°C. Following washing, pellets were resuspended in 10 ml sterile 10% glycerol prechilled to 4°C. Cell suspensions were centrifuged for 10 minutes at 4000 xg, and supernatants were discarded. Pellets were resuspended and combined in 0.5 ml sterile 10% glycerol. This suspension was divided into 50 µl aliquots which were immediately frozen at -80°C.

2.4 Engineering of *C. acetobutylicum*

For a full list of *C. acetobutylicum* strains used in this study, please refer to Appendix 2.

2.4.1 Methylation of plasmid DNA

Plasmids intended for transformation into *C. acetobutylicum* were methylated by transformation into *E. coli* pAN-2. Transformations were carried out as previously described. Methylated plasmid DNA was purified from cells isolated from a plate as previously described.

2.4.2 *C. acetobutylicum* cell culture

All work involving living *C. acetobutylicum* cultures was carried out in an anaerobic cabinet (MACS-MG-1000 anaerobic workstation, Don Whitley) with an atmosphere of 80% nitrogen, 10% hydrogen, and 10% carbon dioxide, a humidity of 70% and a temperature of 37°C. All liquid media and reagents were reduced prior to use via storage in the cabinet overnight. All plates were similarly reduced for 3-4 hours. If cultures needed to be temporarily removed from the cabinet e.g. for centrifugation, culture tubes were tightly sealed and transported on ice to a pre-cooled centrifuge.

2.4.3 Storage of *C. acetobutylicum* strains

All *C. acetobutylicum* strains were stored either as glycerol stocks frozen at -80°C or as spore suspensions stored at 4°C. Glycerol stocks were prepared either by addition of sterile, reduced 100% glycerol to an early exponential ($OD_{600} \sim 0.8$) culture of *C. acetobutylicum*, with a final concentration of 20% glycerol. Alternatively, cells were scraped directly from CGM-agar plates and resuspended in 20% glycerol. The cell suspensions were transferred to sterile 2 ml screw-cap microtubes and immediately

frozen at -80°C. Spore preparations were prepared by streaking a *C. acetobutylicum* liquid culture onto CBM-agar (5-6% Glucose) and incubating in the anaerobic cabinet for 5-7 days. Alternatively, a strain was grown in liquid culture for 5-7 days in CBM (5-6% Glucose). Spores were harvested from the plate by resuspension in 1 ml PBS, heated at 80°C for 10 minutes in order to kill any vegetative cells, and immediately stored at 4°C. Spores were harvested from the liquid culture by centrifuging the culture (10,000 RPM, 2 minutes), and resuspending the pellet in 1ml PBS, heated at 80°C for 10 minutes in order to kill any vegetative cells, and immediately stored at 4°C.

Strains were recovered from glycerol stocks by streaking of a loop of frozen suspension onto a CGM plate supplemented with antibiotics and incubating in the anaerobic cabinet overnight. Strains were recovered from spore preparations by incubating resuspended spore preparations at 80°C for 10 minutes, and immediately spreading onto CGM-agar.

2.4.4 Preparation of electrocompetent *C. acetobutylicum* cells

In order to produce *C. acetobutylicum* electrocompetent cells, the relevant strains were streaked on CGM-agar and incubated in an anaerobic cabinet overnight; 10 ml anaerobic 2xYTG medium was inoculated with a loop of material from the plate and three 10-fold dilutions were made. Electroporation buffer (EPB) (284mM Sucrose, 5mM sodium phosphate) was reduced in the cabinet overnight. Overnight cultures were incubated overnight, at which point the most diluted culture that still exhibited growth was used to inoculate a flask containing 60 ml anaerobic 2xYTG medium. At an OD₆₀₀ of between 0.6 and 0.8, this culture was divided into four 15 ml aliquots in degassed 50 ml centrifuge tubes and centrifuged at 6500 RPM for 10 minutes at 4°C; pellets were washed by resuspension in 5 ml cold anaerobic EPB and repetition of the

centrifugation step. Both pellets were resuspended and combined in a total volume of 2.3 ml cold EPB and used immediately or stored at -80°C in 10% DMSO.

2.4.5 Transformation of *C. acetobutylicum*

20 µl methylated plasmid DNA was introduced to a pre-chilled electroporation cuvette with a channel width of 4 mm. 570 µl of electrocompetent *C. acetobutylicum* cells were added and incubated on ice for 2 minutes. Cells were then electroporated using a BioRad GenePulser XCell™ electroporator with conditions of 2.0 kV, 25 µF and $\infty\Omega$ with an expected time constant of approx. 8-12 ms. 1 ml warm 2xYTG medium was immediately added and the suspension was transferred to a tube containing 5 ml 2xYTG medium. After a recovery period of 3-4 hours, the cultures were centrifuged at 6500 RPM for 10 minutes at room temperature, pellets were resuspended in warm 2xYTG medium and were spread onto CGM plates.

2.4.6 Integration of constructs into the genome

Constructs were integrated into the genome at two different loci using Allele-Coupled Exchange.

pMTL-JH14 and pMTL-ME6 based vectors were used to integrate at the *pyrE* locus in a host which had previously had the *pyrE* gene deleted (Heap et al. 2012; Ehsaan et al. 2016). To select for integrants, after transformation using the pMTL-JH14 vector, cells were streaked onto both antibiotic/uracil and unsupplemented CGM-agar plates; colonies were repeatedly passaged on both until colonies exhibited complete uracil autotrophy and antibiotic sensitivity.

pMTL-JH16 (Heap et al. 2012) based vectors were used to integrate downstream of the strong genomic *thl* promoter, leading to expression of an erythromycin resistance gene after the second crossover event. For transformations with these vectors, cells

were streaked on thiamphenicol CGM-agar plates; cells were then restreaked on thiamphenicol/erythromycin and on erythromycin alone until colonies exhibited complete erythromycin resistance and thiamphenicol sensitivity.

2.4.7 Creation of in-frame deletion mutants

pMTL-ME3 (Ehsaan et al. 2016) based vectors were used to create in-frame deletion mutants, vectors carrying the required KO cassette were electroporated into the *pyrE* mutant strain and plated on thiamphenicol CGM-agar plates. Single crossover mutants were selected for by streaking the fastest growing colonies on the same medium. Single-crossover mutants were screened by PCR using the appropriate primers.

Pure single crossover clones were then plated onto uracil CGM-agar plates, then on Uracil and FOA CGM-agar plates. Loss of the plasmid was confirmed by plating on CGM and thiamphenicol CGM-agar plates. Colonies were screened using appropriate primers.

Screening of engineered *C. acetobutylicum*

For genomic DNA preparation, 3 ml *C. acetobutylicum* liquid culture at an OD₆₀₀ of approximately 0.8 was centrifuged at full speed with a benchtop microcentrifuge. DNA was prepared from the pellet through use of a Qiagen Blood and Tissue Kit in accordance with the protocol for Gram-positive bacteria.

2.5 General molecular biology techniques

2.5.1 Ordering of synthesised genes

Synthesised genes used in this project were ordered from either Biomatik or Thermofisher. Genes from Biomatik were provided in the pBMH vector as lyophilized DNA. DNA was resuspended in sterile dH₂O to a concentration of 100 ng/μl and 0.5 μl

was used to transform *E. coli* by electroporation as described later in this section. Genes from Thermofisher were provided as GeneArt® Strings™ format; these were digested with the appropriate restriction enzymes and cloned as described later in this section.

2.5.2 Purification of plasmid DNA

Plasmid DNA was purified by the use of the GenElute HP Plasmid MiniPrep kits in accordance with the manufacturer's protocol. DNA was purified from either monoclonal liquid culture or from solid agar culture.

2.5.3 PCR amplification

Oligonucleotides designed as part of this study were ordered from either Eurofins MWG Operon or Sigma. Oligonucleotides were provided as lyophilised DNA and resuspended with sterile dH₂O to a final concentration of 100 pmol/ul.

PCR was conducted using Phusion®, or Taq polymerases according to the manufacturer's instructions. GreenTaq polymerase was used for screening of *C. acetobutylicum* genomic DNA, while Phusion was used for amplification of sequences for subsequent cloning or sequencing. A full list of plasmids used is given in Appendix 3.

2.5.4 Restriction analysis of plasmid DNA

In this study, FastDigest enzymes from Fermentas were used. Reaction mixtures for FastDigest enzymes were prepared with 0.5-1 µg DNA solution, 2 µl 10x reaction buffer, 1 µl of each enzyme, and diluted to a final volume of 20 µl. Volume of plasmid DNA was adjusted for subsequent digests if necessary according to the concentration of the DNA. Reactions were incubated at 37°C for 1 hour and subjected to gel electrophoresis in order to separate the resulting bands

2.5.5 Agarose gel electrophoresis

A 0.8% agarose gel solution was prepared by heating 0.8 g agarose in 100 ml 1x TAE buffer (40 mM Tris-acetate, 1 mM EDTA) with a microwave until all agarose had melted. The molten gel was allowed to cool to 50°C in a water bath before pouring into a gel casting tray; 10 µl SybrSafe (10000X concentrate) was immediately added per 100 ml gel and thoroughly mixed in before solidification. The gel was allowed to set and immediately used for electrophoresis.

20 µl restriction digest mixture was loaded into each well. For reaction mixtures requiring the addition of loading dye, 4 µl gel loading dye (Gel Loading Dye, Blue (6X)) was added prior to loading; samples digested with Fermentas enzymes and containing 10X FastDigest Green Buffer were able to be loaded directly onto the gel. The gel was subjected to electrophoresis in 1x TAE buffer at 100V using a Biometra Compact XS/S Horizontal Gel Electrophoresis Apparatus until an appropriate migration had been reached, upon which point the gel was examined under blue light.

2.5.6 Ligation of Plasmid DNA

Restricted DNA was subjected to electrophoresis as described above, imaged, and the relevant bands excised using a scalpel. The DNA was then purified using a Zymoclean™ Gel DNA Recovery Kit in accordance with manufacturer's protocol. A ligation mixture was prepared containing 1 µl 'backbone' DNA fragment, either 3 µl of each insert fragment if two fragments were used or 6 µl insert fragment if only one was used, 1 µl 10x ligation buffer, and 1 µl T4 ligase; the mixture was diluted to 10 µl with sterile distilled water. The mixture was incubated either overnight at 4°C or for 1 hour at room temperature and subsequently used for electroporation.

2.5.7 Electroporation of *E. coli* electrocompetent cells

0.5 µl purified plasmid DNA, or 5 µl dialysed ligation reaction mixture, was pipetted into a pre-chilled 2 mm electroporation cuvette, and 50 µl *E. coli* electrocompetent cells added. Cells were subjected to electroporation with a BioRad Micropulser electroporator according to the pre-programmed Ec2 protocol (25kV, 1 pulse). After electroporation, 500 µl LB medium was immediately added; the cell suspension was transferred to a sterile 1.5ml microcentrifuge tube and incubated at 37°C for 1 hour with shaking at 200 RPM. Immediately after incubation, cells were plated onto LB-agar plates supplemented with the relevant antibiotics, and incubated at 37°C unless otherwise specified.

2.5.8 Sequencing of DNA constructs

All sequencing was carried out by Source Bioscience Lifesciences. 20 µl DNA solution and 100 µl primer DNA at a concentration of 3.2 pmol/µl was sent for each sequencing reaction. A full list of sequencing primers is given in Appendix 3.2.

2.6 Bioinformatics

2.6.1 Plasmid map analysis

Benchling was used to conduct sequence alignment, analyse sequences for restriction sites, and design plasmid maps (<http://benchling.com>). NEB's online design tool (<http://nebbuilder.neb.com/>) was used to design primers for use with NEB Hi-Fi assembly and Quick-change mutagenesis.

2.6.2 Sequence databases

The databases at the NCBI website (<http://www.ncbi.nlm.nih.gov/>) were used to identify nucleotide sequences and peptide sequences of genes of interest. These databases were also used to identify and examine different protein families.

2.6.3 BLAST analysis

BLAST software, available on the NCBI website, (<http://blast.ncbi.nlm.nih.gov/Blast.cgi#>) was used to analyse sequences and identify homology. Nucleotide BLAST optimised for highly similar sequences (megablast) was used to identify homologous genes based on DNA sequence; blastp (protein-protein blast) was used to identify homologous proteins and calculate degrees of amino acid similarity.

2.6.4 Prediction of signal peptides

PSORT <http://psort.hgc.jp/form.html> was used to predict the cellular localisation of proteins of interest.

2.6.5 Analysis of RNA structure

Mfold (<http://mfold.rna.albany.edu/?q=mfold/rna-folding-form>) was used to determine the structure of 5 prime and 3 prime UTR's as well as determining the initial free energies.

2.6.6 Statistical analysis

Graphical depictions of numerical data were prepared using GraphPad Prism software. This software was used for analysis of confidence margins and statistical significance.

2.7 Protein expression analysis

2.7.1 Growth conditions for protein analysis

C. acetobutylicum strains were streaked onto CGM plates with the relevant supplements, incubated overnight, and serial dilutions prepared as above. Strains were first inoculated into 5 ml CBM, after several hours 5 ml of CBM MOPS was added, after 1-hour serial dilutions in CBM-MOPS were set up and incubated overnight. For each strain, the most diluted tube still exhibiting growth was used to inoculate a 50 ml centrifuge tube containing the same medium to an OD₆₀₀ of 0.05. This culture was incubated until the OD₆₀₀ had reached 1.0. At this point, cultures were centrifuged at 4°C for 10 minutes at 6500 RPM and supernatants concentrated for further analysis. When the lactose inducible strains were used they were induced at OD 0.3 with the appropriate concentration of lactose (500 mM stock).

Whole cell fractions, were prepared by centrifuging at 4°C for 10 minutes at 6500 RPM cultures at the same point as supernatants were collected. The pellet was collected and frozen at -20°C.

2.7.2 Denaturing analysis

For denaturing analysis, supernatants were concentrated via trichloroacetic acid (TCA) precipitation. A solution of 10% deoxycholine sodium salt was added to the supernatants to a concentration of 0.03%; supernatants were incubated at room temperature for 10 minutes prior to addition of 100% TCA to a concentration of 10%. Solutions were frozen at -20°C overnight, thawed and centrifuged at 10000 RPM for 10 minutes, at which point the supernatants were discarded. The pellets were then washed in 10 mM Tris buffer pH 8 and allowed to dry. 4x loading buffer was prepared by addition of dithiothreitol to NuPAGE® LDS Sample Buffer (4X) to a final

concentration of approx. 0.4 M (123 mg added to 2 ml); this was diluted in a 1:1 ratio with 1 M Tris buffer pH 8 to give 2x loading buffer. Each pellet was resuspended in 100 µl 2x loading buffer and solutions heated at 100°C for 10 minutes before storage at -20°C until required.

Whole cell pellets were stored at -20°C overnight, thawed and resuspended in 2x loading buffer and heated at 100°C for 10 minutes before storage at -20°C until required.

5 µl of Ladder (ColorPlus™ Prestained Protein Ladder, Broad Range (10-230 kDa) was loaded directly. (Amino-terminal FLAG-BAP™ Fusion Protein) was prepared by adding 1 µl protein stock to 200 µl of 2x loading buffer. Before loading, samples, ladder and control were heated at 100°C for 10 minutes; 10 µl of each sample was loaded onto a gel (NuPAGE® Novex 4-12% Bis-Tris Gel 1.0 mm) and run for approximately 2 hours at 120V in NuPAGE® MES SDS Running Buffer (1X produced by dilution of 20X stock with dH₂O).

2.7.3 Western blots

SDS-PAGE gels were subjected to Western blot using the BioRad TransBlot® Turbo™ transfer system and Trans-Blot® Turbo™ Mini PVDF Transfer Packs in accordance with the manufacturer's protocol; the pre-programmed 'mixed MW' program (1.3A, 25V, 7 minutes) was used.

Membranes were then blocked using 30 ml TBS (50 mM Tris, 150mM NaCl, pH 8) containing 5% milk powder for 1 hour at room temperature on a shaker. Blocking buffer was discarded; membranes were labelled by addition of 30 ml of the same blocking buffer containing 0.5 µg/ml monoclonal ANTI-FLAG M2-peroxidase antibody and incubating overnight on a shaker at 4°C. Blots were washed six times with TBS +

0.1% Tween 20 and immediately developed via addition of 3, 3', 5, 5'-tetramethylbenzidine.

2.7.4 Coomassie blue staining

SDS-PAGE gels as prepared in section 2.7.2 were rinsed three times with water for 5 minutes each time in order to remove SDS from the gel. The gel was subsequently immersed in staining solution (0.1% Coomassie Blue R, 40% ethanol, 10% acetic acid) for 1 hour on a shaker at 60 RPM. The staining solution was decanted and a destaining solution (40% ethanol, 10% acetic acid) added; destaining solution was changed regularly in order to obtain efficient destaining. After a suitable level of destaining had been reached, gels were rehydrated by immersion in water for 1 hour on a shaker at 60 RPM and subsequently imaged.

2.8 Purification of Cel48F from *E. coli*

2.8.1 Growth of *E. coli*

E. coli strain BL21(DE3) containing the plasmid pET48Fc was inoculated into 500 ml of LB containing 1% glycerol and 100 µg/ml Ampicillin and grown until OD₆₀₀ of 1.8, shaking at 37°C. The temperature was then reduced to 16°C and cells were induced with 40 µM of IPTG and incubated for 16 hours.

The cells were then pelleted by centrifugation for 30 minutes at 4°C at 10000 RPM, and the cell pellet frozen at -20°C until use.

2.8.2 Purification of Cel48F

Cells were lysed using Bugbuster according to Manufactures instructions, R-lysozyme and Benzonase were also added. Cel48F was then purified using a HisPur™ Ni-NTA Spin Purification Kit according to manufactures instructions. Purity of Cel48F was assessed

by SDS-Page analysis followed by Coomassie blue staining. The concentration of protein was determined using a Bicinchoninic Acid Kit according to manufactures instructions. Activity was confirmed by incubating with PASC, followed by reducing sugar quantification (see 2.11).

2.9 Methods used to assess the interaction between Cel48F and *C. acetobutylicum*

2.9.1 Microscopy

C. acetobutylicum was grown to OD 1.0 in CBM, centrifuged at 6500 RPM for 10 minutes, and resuspended in protoplast buffer (CBM, 0.3 M Sucrose, 25 mM CaCl₂, 25 mM MgCl₂) to OD 10.0. 100 µl of either buffer control (100 mM Tris-HCL pH 5.5), concentrated supernatant or lysozyme (1 mg/ml stock) was added to 900 µl of *C. acetobutylicum* cells. If appropriate they were induced with 10 mM lactose. Time points were taken every hour. 10 µl of culture was pipetted onto a glass slide, and a coverslip added. 10 photos were then taken at random for each sample using a Nikon Model Eclipse microscope.

For analysis, the number of protoplasts and intact cells were counted, and the percentage of protoplasts calculated. Statistical significance was determined using students t-test (Non-parametric).

2.9.2 Autoclaved cells assay

The relevant strains were streaked on CGM-agar and incubated in an anaerobic cabinet overnight; 10 ml anaerobic 2xYTG medium was inoculated with a loop of material from the plate and three 10-fold dilutions were made. The lowest dilution which showed some growth was used to inoculate the appropriate volume of 2xYTG, at OD₆₀₀ 0.6 the culture was centrifuged at 6500 RPM for 10 minutes at 4°C, the pellet was then washed

twice in dH₂O before being resuspended in 300 µl of dH₂O and autoclaved. The autoclaved cells were resuspended to OD₆₀₀ 0.6 in 100 mM Tris-HCL (pH 5.5) and stored at 4°C until needed.

Strains were grown in CBM-MOPS (40 mM) to OD 1.0 as per western blot. Protease inhibitors (EDTA free) were added and supernatant concentrated 100 times using SpinX-UF columns (10 kDa cutoff), all steps were carried out at 4°C and concentrated S/N stored at 4°C.

30 µl of concentrated supernatants, lysozyme or purified Cel48F were added to 270 µl of autoclaved *C. acetobutylicum* cells in a 96-well flat-bottomed microplate. OD₆₀₀ was measured every 10 minutes using a Clariostar plate reader with the temperature set at 37°C. Final results were expressed as a percent of the initial OD and were the average of three technical replicates. Standard error of the mean and statistical significance were assessed using GraphPad Prism software.

2.9.3 Bacteriocin assay

To assess the bacteriolytic activity of Cel48F *E. coli*, *B. subtilis* and *C. acetobutylicum* were used. For *E. coli* and *B. subtilis*, 270 µl of a growing culture with an OD₆₀₀ of 0.1 was added to a 96-well flat-bottomed microplate with either 30 µl of a buffer control (100 mM Tris-HCL pH 5.5), lysozyme or purified Cel48F. OD₆₀₀ was measured every 10 minutes using a Clariostar plate reader with the temperature set at 37°C. Final results were expressed as a percent of the initial OD and were the average of three technical replicates; standard error of the mean and statistical significance were assessed using GraphPad Prism software.

For *C. acetobutylicum* cell growth was assessed using a Biolector, in an anaerobic cabinet 900 µl of a growing culture with an OD₆₀₀ of 0.1 was added to the Biolector

plate, and 100 µl of either buffer control (100 mM Tris-HCL pH 5.5), lysozyme or purified Cel48F was added. The plate was transferred into the Biolector and incubated at 37°C anaerobically. Biomass was measured every 30 minutes. Final results were expressed as a percent of the initial OD and were the average of three technical replicates; standard error of the mean and statistical significance were assessed using GraphPad Prism software.

2.9.4 Peptidoglycan analysis and peptidoglycan assays

2.9.4.1 Purification of peptidoglycan

Peptidoglycan was purified from *C. acetobutylicum* by the same process but in two different volumes depending if it was for HPLC/MS analysis (mini) or peptidoglycan turbidity assay (Maxi). First, the *C. acetobutylicum* strain was grown to OD₆₀₀ 0.8 in 2xYTG. The cells were collected by centrifugation (7000 xg 10 minutes) either 15 ml were used for the mini protocol or 500 ml for the maxi protocol. Cells were resuspended in 1 ml (mini) or 30 ml (maxi) of 100 mM Tris-HCL pH 6.8 with 0.25% SDS and boiled at 100°C for 10 minutes. The samples were then centrifuged at 10,000RPM for 5 mins and washed 6 times with dH₂O (1.5 ml mini and 30 ml maxi).

Samples were sonicated for 30 minutes using a Bioruptor, after sonication they were incubated with Benzonase (1 µl/ml) for 1 hour at 37°C shaking followed by adding (50 µl (mini) or 1.5 ml (maxi) of a 50 µg/ml solution of Trypsin in 100 mM Tris-HCL pH 6.8 and incubating for 1 hour at 37°C. The enzymes were then inactivated by boiling at 100°C for 3 minutes.

The peptidoglycan was then optionally incubated with 1 M HCL for 4h at 37°C shaking (500 µl mini, 15 ml (Maxi) to remove cell wall bound teichoic acids. The suspension

was centrifuged at 10000 RPM for 10 minutes and washed with ddH₂O until the pH reached 5-6.

2.9.4.2 Chemical acetylation of peptidoglycan

Peptidoglycan was acetylated by adding 5% Acetic anhydride to a 1 mg/ml solution of peptidoglycan, the pH was then adjusted to pH 6.8 by adding NaCO₃. The solution was then incubated for 30 minutes at 4°C followed by 1 hour at room temperature. The peptidoglycan was then washed in PBS three times.

2.9.4.3 Peptidoglycan turbidity assay

The activity of supernatants of strains to peptidoglycan was determined by growing the indicated strains to an OD₆₀₀ of 1.0 in CBM (as MOPS appeared to inhibit all muramidase activity). The supernatants were concentrated 100 times using SpinX-UF columns (10 kDa cutoff).

Peptidoglycan purified using the maxi method up to the point of removal of wall teichoic acids were first resuspended in 50 mM Tris-Acetate buffer (pH 4) and sonicated for 15 minutes using the BioRuptor to resolubilise the peptidoglycan.

30 µl of either 100 times concentrated supernatant, mutanolysin, lysozyme or purified protein were added to 270 µl of peptidoglycan in a 96-well flat-bottomed microplate. OD₆₀₀ was measured every 10 minutes using a Clariostar plate reader with the temperature set at 37°C. Final results were expressed as a percent of the initial OD and were the average of three technical replicates; standard error of the mean and statistical significance was assessed using GraphPad Prism software.

2.9.4.4 *Peptidoglycan binding assay*

The peptidoglycan binding assay was adapted from (Atilano et al. 2011), 100 µg of purified protein was added to 0.2 mg of peptidoglycan from *B. subtilis* PBS was added to a total volume of 200 µl. If appropriate peptidoglycan was first chemically acetylated (2.9.4.2). The samples were then incubated for 30 minutes at 37°C. The peptidoglycan and any bound protein was then pelleted by centrifugation (10,000 RPM, 5 minutes), washed in 100 µl of PBS three times. The pellet was then resuspended in 200 µl of loading buffer and loading buffer added to the supernatant fractions. The samples were then loaded on an SDS-PAGE gel and analysed by western blot.

2.9.4.5 *Preparation of peptidoglycan for LC/MS*

Peptidoglycan (with cell wall teichoic acids removed) was resuspended in 100-250 µl of digestion buffer (12.5 mM sodium di-hydrogen phosphate pH 5.5) to an OD₅₇₈ of less than 3.0, sonicated for 30 minutes to resolubilise the peptidoglycan. Followed by digestion with mutanolysin by adding 1/10 volume of mutanolysin (5 U/ml of mutanolysin in dH₂O), then incubated for 16h at 37°C (150 RPM shaking).

Mutanolysin was inactivated by boiling for 3 minutes. The sample centrifuged (5 mins 10,000 RPM), supernatant collected and MurNAc reduced by adding 50 µl of (10 mg/ml sodium borohydrate in 0.5 M borax in dH₂O (pH 9), incubating for 20 minutes at room temperature. The reaction was stopped by adding 10 µl of phosphoric acid (98%).

2.9.4.6 *Analysis of peptidoglycan by LC-MS and LC-MS/MS*

Reversed-phase high pressure liquid chromatography-full scan mass-spectrometry (RP-LC-MS) measurements were made using an inlet comprising binary pumps and auto-sampled system (Perkin-Elmer 200 series) delivering a gradient of acetonitrile in water, 0.1% (v/v) formic acid at 0.2 ml.min⁻¹ with a polar C18 column (T3 Cortecs,

2.1x50 mm, 2.7 m, Waters) to a hybrid quadrupole-Time of flight MS instrument (Q-ToF Premier; Waters) equipped with a conventional geometry electrospray ion-source, controlled using Waters MassLynx 4.1 software.

Structural analyses were performed using LC-MS/MS experiments using a high-resolution instrument equipped with an Orbitrap analyser (Q-Exactive-Plus; Thermo) with an electrospray source in data-dependent MS2 mode; controlled using Thermo Xcalibur software. The inlet comprised a capillary-scale RP-LC (Ultimate 3000; Thermo) with a C18 column (self-pack 0.5x50 mm, 3 m, Luna C18; Phenomenex) chromatographed at 5 ml.min⁻¹ using a gradient of acetonitrile in water 0.1% (v/v) formic acid.

2.10 Analysis of cellulase activity

2.10.1 Preparation of Phosphoric acid swollen cellulose (PASC)

0.2 g of Avicel was added to a 50 mL centrifuge tube and 0.6 mL distilled water was added to wet the cellulose powder to form a cellulose-suspended slurry. 10 ml of ice-cold 86.2% H₃PO₄ was slowly added to the slurry with vigorous stirring so that the final phosphoric acid concentration was ca. 83.2%. The cellulose mixture turned transparent within several minutes, and stood for about an hour on ice with occasional stirring. Approximately 40 mL of ice-cold water was added at a rate of approximately 10 mL per addition with vigorous stirring between additions, resulting in a white cloudy precipitate. The precipitated cellulose was then centrifuged at 5000 xg and 4 °C for 20 minutes, the pellet was resuspended by ice-cold water; this step was repeated 4 times. Approximately 0.5 mL of 2 M Na₂CO₃ was added to neutralize the residual phosphoric acid, and the pellet was resuspended in 45 ml of distilled water. After centrifugation, the pellet was resuspended in distilled water and centrifuged twice or until the pH reached 5.5.

2.10.2 Growth conditions

Strains were grown up in CBM-MOPS to OD₆₀₀ 1 as for western blot. At this point, cultures were centrifuged at 4°C for 10 minutes at 7000 xg. Protease inhibitor cocktail (Roche EDTA free) were added to supernatant followed by concentration 10 times using SpinX-UF columns. As residual sugars interfered with reducing sugar analysis, the supernatants were dialysed three times using the SpinX-UF columns by adding the original volume of 100 mM Tris-HCL (pH 5.5) plus protease inhibitors. All centrifuge steps were carried out at 4°C and the supernatants stored at 4°C.

2.10.3 Cellulase activity assay

90 µl of 10 times concentrated supernatant from each strain was incubated with 90 µl of 2% substrate (PASC) in 100 mM Tris-HCL, 1 mM CaCl₂ pH 5.5 and incubated at 37°C for 72 hours. Three biological replicates of each strain were used.

2.10.4 Reducing sugar analysis

Reducing sugar concentration was analysed by dinitrosalicylic acid (DNS) assay carried out after the method described by Anderson *et al.* [171]. A DNS assay solution, consisting of 1% dinitrosalicylic acid, 0.2% phenol, 0.05% sodium sulphite, 1% sodium hydroxide, and 10% sodium potassium tartrate, was prepared. Samples from the digestion were centrifuged and supernatants were transferred to a fresh tube. 60 µl supernatant and 120 µl DNS reaction mixture were mixed in a 96 well PCR plate and heated to 95°C for 5 minutes using a PCR thermocycler, followed by cooling to 4°C for 1 minute to stop the reaction. 36 µl reaction mixture was diluted with 160 µl dH₂O in a 96-well flat-bottomed microplate and the absorbance at 540 nm was measured using a Clariostar plate reader. All absorbance readings were taken as an average of three readings. Reducing sugar concentration was determined by a glucose standard curve

(0, 0.5, 1, 2, 3 and 4 mg/ml glucose). Final results were expressed as the average of three technical replicates; standard error of the mean and statistical significance were assessed using GraphPad Prism software.

2.11 Cloning of plasmid constructs

A full list of plasmids used is given in Appendix 4.

2.11.1 Cloning of *Cel5A*, *Cel48F* and *Cel9E* downstream the *PtcdB* promoter.

Previous work in our group had constructed hybrid cellulases consisting of the cellulase domain from *C. cellulolyticum* and dockerin domain from *C. thermocellum* (**pJ206_BB2_Cel48F_CelS_Flag**, **pJ206_BB2_Cel9E_CelE_Flag** and **pJ206_BB2_Cel5A_CelA_Flag**). The promoter *PtcdB* was also in BB2 format (**pJ204_BB2_PtcdB**). To clone the cellulases downstream the *PtcdB* promoter, **pJ204_BB2_PtcdB** was digested with *NheI* and *PstI* and the cellulases with *SpeI* and *PstI* and ligated together, resulting the following plasmids: **pJ204_BB2_PtcdB_Cel48F_CelS_Flag**, **pJ204_BB2PtcdB_Cel9E_CelE_Flag** and **pJ204_BB2PtcdB_Cel5A_CelA_Flag**. Gene of interest along with the *PtcdB* promoter were sub-cloned into the **pMTL82151** shuttle vector by digesting all plasmids with *NotI* and *NheI* restriction enzymes, and ligating together creating the **pMTL82151_BB2_PtcdB_Cel48F_CelS_Flag**, **pMTL82151_BB2PtcdB_Cel9E_CelE_Flag** and **pMTL82151_BB2PtcdB_Cel5A_CelA_Flag** constructs.

2.11.2 Cloning the native *Cel48F* (with its native dockerin domain) and *Cel48F* without its dockerin downstream the *PtcdB* promoter.

pMTL82151-BB2_PtcdB_Cel48F_N (*Cel48F* with its native dockerin domain) and **pMTL82151-BB2_PtcdB_Cel48F_ND** (without its native dockerin domain) were cloned by first adding a Flag tag onto **pJ204_BB2_cel48F_ND** (present in CRG Culture

collection) by cutting **pJ204_BB2_cel48F_ND** with *EcoRI* and *SpeI* and cutting **pJ202_BB2_Flag_Tag** with *EcoRI* and *NheI*, ligating them together, creating **pJ202_BB2_Cel48F_ND_F**, **pJ202_BB2_Cel48F_ND_F** and **pJ204_BB2_Cel48F_N_Flag** constructs, which were subsequently cloned by restriction digest as *NdeI* and *NotI* fragments into the pMTL82151 shuttle vector, creating **pMTL82151_PtcdB_Cel48F_ND_Flag** and **pMTL82151_PtcdB_Cel48F_N_Flag** constructs.

2.11.3 Integration of Cel48F into the chromosome of *C. acetobutylicum*

Cellulases were integrated into the chromosome at the *thiolase* locus using the pMTL-JH16 vector (Heap et al., 2012). To avoid a possible interference from the chromosomal PthI promoter, a repression a terminator (T1T2) was placed before the *PtcdB* promoter. To create the **pJ204_BB2_T1T2_PtcdB** construct the **pJ204_BB2_T1T2 construct was digested** with *NheI* and *PstI* and **pJ204_BB2_PtcdB construct** with *SpeI* and *PstI* then ligated.

pJ204_BB2_T1T2_PtcdB_Cel48F_N_Flag construct was then created by cutting **pJ204_Cel48F_N plasmid** with *SpeI* and *PstI* and **pJ204_BB2_T1T2_PtcdB plasmid** with *NheI* and *PstI* followed by ligation.

The **pMTL-JH16_T1T2_PtcdB_Cel48F_N_Flag integration plasmid** was created by cutting the **pMTL-JH16** and the **pJ204_BB2_T1T2_PtcdB_Cel48F_N_Flag** with *NotI* and *NheI* followed by ligation.

The **pMTL-JH16_T1T2_PtcdB_Cel48F_ND_Flag** and the **pMTL-JH16_T1T2_PtcdB_Cel48F_CelS_Flag** plasmids were created by cutting the **pMTL82151_PtcdB_Cel48F_ND_Flag** and the **pMTL82151_PtcdB_Cel48F_CelS_Flag** plasmids with the same enzymes (*NdeI* and *NheI*) and **pMTL-**

JH16_T1T2_PtcdB_Cel48F_N_Flag also with *NdeI* and *NheI* followed by ligation, resulting in the **pMTL-JH16_T1T2_PtcdB_Cel48F_ND_Flag** and **pMTL-JH16_T1T2_PtcdB_Cel48F_CelS_Flag** integration plasmids.

2.11.4 Cloning of 5 prime UTRs

Several 5 prime UTR's were ordered from Biomatik in BB2 format along with the *PtcdB* promoter. These were first cloned upstream the *Cel48F_ND* by digesting **pBMH_BB2_PtcdB_SlpA**, **pBMH_BB2_PtcdB_CoseGsib** and **pBMH_BB2_PtcdB_CoseThio** with *SpeI* and *PstI* and digesting **pJ202_BB2_Cel48F_ND_F** with *NheI* and *PstI* and ligating them together creating the **pJ204_BB2_PtcdB_SlpA_Cel48F_ND_Flag**, **pJ204_BB2_PtcdB_CoseGsib_Cel48F_ND_Flag** and **pJ204_BB2_PtcdB_CoseGsib_Cel48F_ND_Flag** constructs.

These were subsequently cloned into the **pMTL82151** vector by digesting all constructs with *NotI/NheI* followed by ligation resulting in; **pMTL82151_PtcdB_SlpA_Cel48F_ND_Flag**, **pMTL82151_PtcdB_CoseGsib_Cel48F_ND_Flag** and **pMTL82151_PtcdB_CoseThio_Cel48F_ND_Flag**.

In addition, a *PtcdB* promoter with its native RBS instead of a thiolase RBS was constructed by digesting **pMTL-82151_PtcdB(N)** with *NotI* and *NdeI* and **pMTL82151_PtcdB_Cel48F_ND_Flag** with *NotI/NdeI* followed by ligation resulting in; **pMTL82151_PtcdB(N)-Cel48F_ND**.

2.11.5 Cloning of 3 prime UTRs

Two 3 prime UTR's were ordered from Biomatik in BB2 format: **pBMH-BB2_SL2** and **pBMH-BB2_SL4**, in addition another 3 prime UTR in the CRG culture collection **pJ202-**

BB2_SlpA(term) was used. These were digested with Not1 and Nde1 and **pMTL82151_PtcdB_CoseThio_Cel48F_ND_Flag** was digested with Not1 and Nde1 followed by ligation resulting in; **pMTL82151_PtcdB(Coe_Thio)_Cel48F_ND_SL2**, **pMTL82151_PtcdB(Coe_Thio)_Cel48F_ND_SL4** and **pMTL82151_PtcdB(Coe_Thio)_Cel48F_ND_SlpA(term)**.

2.11.6 Cloning of PrsA

Two PrsA genes from *R. cellulolyticum* CCell0751 and Ccel_3056 were ordered from Thermofisher as GeneArt® Strings™ format in BB2 format. The PrsA genes from *R. cellulolyticum* were cloned into **pMTL-82151-TcpF** by digesting **pMTL-82151-TcpF** with Nde1 and Nhe1 creating: **pMTL-82151-TcpF-PrsA-CCell0751** and **pMTL-82151-TcpF-PrsA-Ccel3056**.

The PrsA genes from *C. acetobutylicum* and *B. subtilis* were constructed using NEB's NEBuilder® HiFi DNA Assembly according to manufacturer's instructions. The plasmid **pMTL-82151-TcpF** was digested with *NdeI* and *NheI* and the PrsA genes amplified from genomic DNA using primers PrsA_CAC_Fw, PrsA_CAC_Rev and PrsA_BAC_Fw, PrsA_BAC-Rev respectively creating plasmids **pMTL-82151-TcpF-PrsA_CAC** and **pMTL-82151-TcpF-PrsA_Bac**,

2.11.7 Cloning of GH48 enzymes

GH48 enzymes were ordered from Biomatik in BB2 format with the thiolase RBS and Flag-tagged at their C terminus; **pBMH_BB2_ExgS_Flag**, **pBMH_BB2_C748_Flag**, **pBMH_BB2_Cphy68_Flag**. They were cloned into pMTL82151 with a *PtcdB* promoter by digesting them and **pMTL82151_PtcdB_Cel5A_CelA_Flag** with *NdeI/NheI* creating; **pMTL82151_PtcdB_ExgS_Flag**, **pMTL82151_PtcdB_C748_Flag** and **pMTL82151_PtcdB_Cphy68_Flag**.

2.11.8 Construction of *E. coli* Cel48F overexpression vector

The vector **pET-48Fc** was constructed using NEB's NEBuilder® HiFi DNA Assembly according to manufacturer's instructions, the plasmid **pET-16B** was amplified by PCR using the primers 14Cel48F_F and 14Cel48F_R and Cel48F was amplified from **pJ204_BB2_Cel48F_N_Flag** using the primers Cel48A_F and Cel48A_R.

2.11.9 Construction of in-frame deletion vectors

All in-frame deletion vectors were constructed using NEB's NEBuilder® HiFi DNA Assembly according to manufacturer's instructions. The in-frame deletion vector **pMTL-ME3** was used and 750bp left and right homology arms were used.

To create the **CAC_2316** deletion vector, **pMTL-ME3** was digested with *NotI*/*NheI* and the primers 2316_RHA_F and 2316_RHA_R, 2316_LHA_F and 2316_LHA_R were used respectively to PCR the right and left homology arms from *C. acetobutylicum*'s genomic DNA. NEB's NEBuilder® HiFi DNA Assembly kit was used to assemble the vector **pMTL-ME3-2316**.

To create the **CAC_3377** deletion vector, **pMTL-ME3** was digested with *NotI*/*NheI* and the primers 3377_RHA_F and 3377_RHA_R, 3377_LHA_F and 3377_LHA_R were used respectively to PCR the right and left homology arms from *C. acetobutylicum*'s genomic DNA. NEB's NEBuilder® HiFi DNA Assembly kit was used to assemble the vector **pMTL-ME3-3377**.

2.11.10 Integration of Flag-tagged Cel48A into *C. acetobutylicum*

The integration vector **pMTL-JH14_Facoid_Cel48A_Flag** were constructed using NEB's NEBuilder® HiFi DNA Assembly according to manufacturer's instructions, the plasmid **pMTL-JH14-Facoid_Cel48F(N)_Flag** was amplified by PCR using the primers

14Cel48F_F and 14Cel48F_R and *Cel48A* was amplified from *C. acetobutylicum*'s genomic DNA using the primers Cel48A_F and Cel48A_R.

2.11.11 Quick change mutagenesis of Cel48F_N

To create an E55Q substitution in Cel48F_N the cloning vector **pJ204_BB2_Cel48F_N_Flag** was used as QuickChange of the final vector **pMTL-JH14-Facoid_Cel48F(N)_Flag**, resulted in truncations most likely due to the presence of a terminator. The substitution was performed with NEB's Q5® Site-Directed Mutagenesis Kit according to manufacturer's instructions using the primers E55Q_F and E55Q_R creating the construct **pJ204_BB2_Cel48F_N-E55Q_Flag**.

This was then cloned into the vector **pMTL-JH14-Facoid** by digestion of all constructs with *NdeI/NheI* followed by ligation resulting in; **pMTL-JH14-Facoid_Cel48F(N)_Flag**.

2.11.12 Overexpression of *R. cellulolyticum*'s OrfXp

pMTL-82151-OrfX_Cell was constructed using NEB's NEBuilder® HiFi DNA Assembly according to manufacturer's instructions, the plasmid **pMTL-82151-TcpF** was linearised with *NotI/NheI* and *OrfX* was amplified from *R. cellulolyticum*'s genomic DNA using the primers OrFX_F and OrFX_R.

2.11.13 Construction of Chimeric Cel48F

pMTL-JH14-Facoid-Cel48F-Chimera was constructed using NEB's NEBuilder® HiFi DNA Assembly according to manufacturer's instructions, the plasmid backbone **pMTL-JH14-Facoid_Cel48A_Flag** was amplified by PCR using the primers 14Cel48F_F and 14Cel48F_R and Cel48F'S catalytic domain was amplified from **pMTL-JH14-Facoid_Cel48F(N)_Flag** using the primers Cel48A_F and Cel48A_R.

Chapter 3 Expression of *R. cellulolyticum* cellulases in *C. acetobutylicum* using a lactose inducible system

Heterologous expression of cellulases from *Ruminiclostridium cellulolyticum* in *C. acetobutylicum* is an attractive strategy for engineering *C. acetobutylicum* to grow on lignocellulosic biomass.

However, the expression of cellulases from *R. cellulolyticum* in *C. acetobutylicum* has previously been reported to be difficult, perhaps due to their toxicity to the host (Mingardon et al. 2011). In particular, cellulases belonging to the GH48 and GH9 family have been reported to be expressed and secreted at significantly lower levels than other cellulases (Willson et al. 2016).

In this chapter, the use of an inducible promoter system for the expression of cellulases in *C. acetobutylicum* was explored. Additionally, the effect of different dockerin domains on the level of degradation of Cel48F, *R. cellulolyticum*'s GH48 family cellulase was also assessed.

3.1 Introduction

The expression of cellulosomal components from *R. cellulolyticum* in *C. acetobutylicum* is an attractive strategy towards an engineered *C. acetobutylicum* strain able to grow on lignocellulosic biomass. Despite being distantly related, *C. acetobutylicum* and *R. cellulolyticum* have a remarkably similar cellulosome, as their main cip-cel operon has a very similar genetic structure. Additionally, there is a high amino-acid sequence homology between all cellulosomal proteins. For example, Cel48A of *C. acetobutylicum* has a 58% amino-acid sequence identity to Cel48F from *R. cellulolyticum*. Cel9X also shows a 52% amino-acid sequence identity with Cel9E from *R. cellulolyticum*, and Cel9C has a 60% amino-acid sequence identity with Cel9G. Due to their high sequence homology, several cellulases could be identified in *C. acetobutylicum* using antibodies raised against *R. cellulolyticum* cellulases (López-Contreras et al. 2004). It has been proposed that due to the phylogenetic distance between *C. acetobutylicum* and *R. cellulolyticum*, horizontal gene transfer rather than a common ancestor is the more likely hypothesis for the close taxonomic relatedness of the cip-cel operon (Desvaux 2005b).

However, the expression of cellulases from *R. cellulolyticum* in *C. acetobutylicum* has previously been reported to be difficult, perhaps due to their toxicity to the host (Mingardon et al. 2011). In these studies, several key cellulases from *R. cellulolyticum*, such as Cel48F, Cel9E and Cel9G could not be transferred to *C. acetobutylicum*.

Previous work in our lab overcame these issues, and successfully expressed several cellulases from *R. cellulolyticum* in *C. acetobutylicum*, including Cel48F and Cel9G; this was achieved by integrating the genes into the genome (Willson et al. 2016). The utilisation of Lac-repressible promoters was required for the cloning of cellulases in *E. coli* as cloning downstream of strong promoters could not be achieved. This is thought

to be due to the accumulation of proteins in inclusion bodies (C. Reverbel Leroy et al. 1997).

The promoter used for the expression of Cel9G and Cel48F was the *Clostridium sporogenes* ferredoxin promoter (P_{fdx}) which carries the 'ideal' symmetrical repression sequence, it is referred to as the FacOID promoter (Müller-Hill and Gilbert W 1968). CatP, a gene expression reporter assay has shown that FacOID is a strong promoter, retaining 80% of the activity when compared to the native thiolase promoter (Tummala, Welker, and Papoutsakis 1999, Ben Willson, 2014).

Secretion of Cel48F and Cel9G was achieved and the strains were shown to be stable, however, the expression levels of Cel9G and Cel48F were significantly reduced compared to the scaffolding protein, CipC (expressed from the *thiolase* promoter). Additionally, significantly lower expression levels of Cel48F were observed compared to Cel9G even when expressed using the same promoter.

The low expression level of Cel48F could be due to several reasons; it is possible that Cel48F is poorly expressed or secreted in *C. acetobutylicum*. It has been proposed that the poor expression and toxicity of Cel48F occurs during secretion (Mingardon et al. 2011). The evidence supporting this was that Cel48F was highly expressed intracellularly, and the addition of X2 domains led to non-toxic secretion of Cel48F (Chanal et al. 2011). X2 domains are found in scaffoldins and some cellulases, their function is unclear, however they are thought to play a role in helping proteins bind to cellulose or the cell surface (Kosugi et al. 2004). Work in our laboratory found that the addition of X2 domains did not result in an increase in the expression level of Cel48F (Katrin Schwarz, personal communication). Wilson et al. 2016 looked at the different cell fractions of a strain expressing Cel48F and only found Cel48F in the supernatant fraction, indicating that all Cel48F produced is successfully secreted.

It is also possible that Cel48F is degraded in the culture supernatant after secretion. The degradation of Cel48F at the dockerin domain has been observed before, even when it was purified from *E. coli* (C. Reverbel-Leroy et al. 1997). In our laboratory, a small amount of degradation of Cel48F was observed with a band of around 10 kDa corresponding to the size of the dockerin domain. Intact Cel48F was detected using Flag-antibody, as the protein carried a C-terminus Flag-tag. When the degradation occurred at the C-terminus, cleaving the Flag-tag, these degradation products would not be detected using the Flag antibody. Therefore, it is possible that there was a higher level of degradation that could not be observed.

3.1.1 Lactose inducible expression system

The use of an inducible promoter system for expression of cellulases in *C. acetobutylicum* has several advantages over their constitutive expression. It should allow for easy cloning of cellulases in *E. coli*, as repression should circumvent toxicity issues. Additionally, the level of induction can also be adjusted resulting in the highest level of expression possible without any toxic effects. Previously the expression of Cel48F under the constitutive *facoid* promoter did not result in any visible toxic effect or impact on growth (Willson et al. 2016). This implies that the expression of Cel48F could be increased when expressed under a stronger promoter.

A lactose inducible system has recently been developed in our laboratory. This is based on a system first developed in *C. perfringens* (Hartman, Liu, and Melville 2011). This inducible system consists of a transcriptional regulator *BgaR*, which regulates the expression of any gene under control of the *PbgaI* promoter in a lactose dependent manner. Lactose was utilised as an inducer as IPTG was shown to be ineffective at inducing the system.

Unfortunately, *PbgaI* is not a strong promoter so to improve the expression, the system was coupled with the *tcdR/PtcdB* orthogonal expression system, derived from *C. difficile* (Minton et al. 2015). In *C. difficile*, the alternative *tcdR* sigma factor has been shown to control the expression of the two main toxin genes, *tcdB* and *tcdA* via the respective promoters *PtcdB* and *PtcdA*. The sigma factor has been assigned to group 5 of the sigma (70) sigma factors, which includes sigma factors from pathogenic organisms including: *C. botulinum*, *C. tetani* and *C. perfringens*. They have been shown to be sufficiently similar, and they could be substituted for each other (Hundsberger et al. 1997).

The *tcdR/PtcdB* system has been shown to result in very high expression of genes in *C. acetobutylicum*; CatP assay data showed it is stronger than the native *thiolase* promoter. Importantly, the *PtcdB* promoter is not functional in the absence of the alternative sigma factor *tcdR* allowing for cloning in *E. coli*.

In order to maintain a high induction level without the inducer being consumed by the host, a $\Delta lacG$ strain was created, via in-frame deletion and the lactose inducible system integrated into the genome at the *pyrE* locus via ACE (Heap et al. 2012). The lactose inducible system consisted of *BgaR* and *tcdR* under control of *PbgaI*. Any gene under control of the *PtcdB* promoter can then be introduced into this strain, either by integration into the genome or on a plasmid, with gene expression controlled in a lactose-dependent manner. The *PbgaI* promoter was used to drive the expression of *tcdR*, and high level of protein expression was observed when genes were under the control of the *PtcdB* promoter induced with lactose. The system was called the lactose-inducible *tcdR* system (LI-*tcdR*) as shown in figure 3.1. The system was developed in our laboratory by Hengzhang Wong.

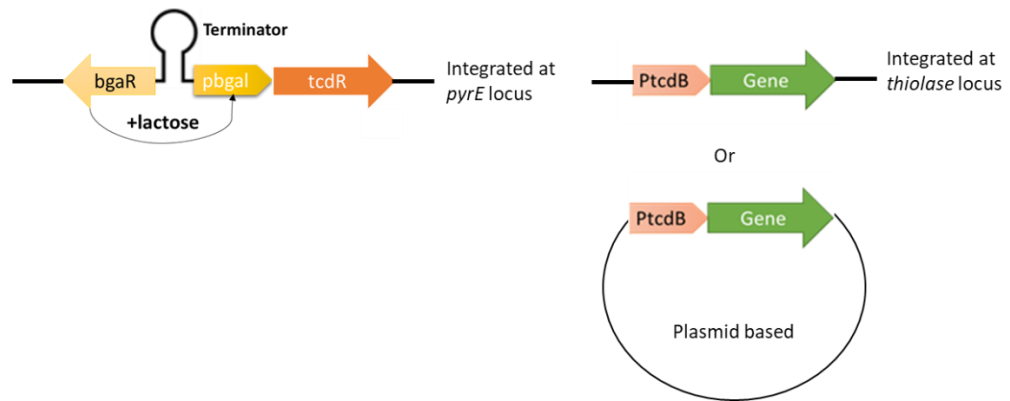


Figure 3.1. Schematic diagram of the lactose inducible *tcdR* system. The lactose-inducible cassette consisting of *bgaR*, *pbgal* and *tcdR* is integrated at the *pyrE* locus in a $\Delta lacG$ background. Any gene can be cloned downstream the *PtcdB* and expression controlled in a lactose dependent manner. The construct can be integrated into the genome at the thiolase locus or be plasmid-based. *BgaR* acts as a repressor of *pbgal* in the absence of lactose, which results in no expression of *tcdR*. In the presence of lactose, *pbgal* is active resulting in *tcdR* being transcribed, *tcdR* can then activate any gene under control of the *PtcdB* promoter.

3.2 Results

3.2.1 Selection of cellulases for testing the lactose inducible system

To exemplify the lactose inducible system, three cellulase genes from *R. cellulolyticum* Cel5A, Cel9E and Cel48F were selected. These cellulases were chosen for several reasons: Cel48F and Cel9E are the highest expressed cellulases in *R. cellulolyticum*'s cellulosome. Additionally, deletion of Cel48F results in a strain unable to grow on untreated biomass and deletion of Cel9E results in a strain with severely impaired growth (T. Xu et al. 2014). These results indicate that they are essential for the degradation of cellulose in *R. cellulolyticum*. Cel5A is also an important cellulase as it has both endocellulase and xylanase activity. *In-vitro* studies have shown that the incorporation of these cellulases into mini-cellulosomes results in a highly active mini-cellulosome (Fierobe et al. 2005). Additionally, they have all been successfully expressed in *C. acetobutylicum*, albeit at different levels (Willson et al. 2016; 2014).

Chimeric forms of the enzymes were used, with the dockerin domain replaced with different dockerin domains from *R. thermocellum*. Cel5A has the dockerin domain from CelA (GH5 family cellulase), Cel9E from CelF (GH9 family cellulase) and Cel48F from Cel48S (GH48 family cellulase). The chimeric forms of the enzymes were made during previous work in our group, they are in BB2 format and consisted of: a *thiolase* RBS followed by the catalytic domain, a BB2 scar followed by the dockerin domain and the protein was Flag-tagged at the C-terminus.

3.2.2 Expression of *R. cellulolyticum* cellulases using the lactose-inducible system.

For the initial testing of the inducible system, a plasmid-based strategy has been chosen, as this can result in higher expression of genes compared to the integration of the genes into the genome due to a higher copy number. The genes were cloned into the pMTL82151 shuttle vector, under control of the *PtcdB* promoter. All cellulases were successfully cloned and transformed into the *C. acetobutylicum* LI-tcdR strain. Cel48F and Cel9E had to be cloned at 30 °C in *E. coli*, showing that the *PtcdB* promoter still had some activity in *E. coli*. Incubation of transformants at 30°C was not required for transformation of these constructs into *C. acetobutylicum*, and all strains were stable when stored as both glycerol and spore stocks. This is in contrast to what was reported by Mingardon et al., where incubation at 30 °C was required, and expression of Cel48F was lost after the *C. acetobutylicum* strains were revived from a spore stock.

To determine the expression levels of the cellulases, the engineered strains were grown in CBM buffered with MOPS (40 mM MOPS, pH 6.8) to an OD of 1.0. We used these growth conditions, as previous work in our group showed that these were the optimum conditions for detecting cellulases (Ben Willson, 2014). The strains were induced with either 0, 0.1 or 1 mM of lactose at OD 0.3. The proteins in the

supernatants were concentrated 40-fold by TCA precipitation, then subjected to SDS-PAGE and western blot analysis (Figure 3.2).

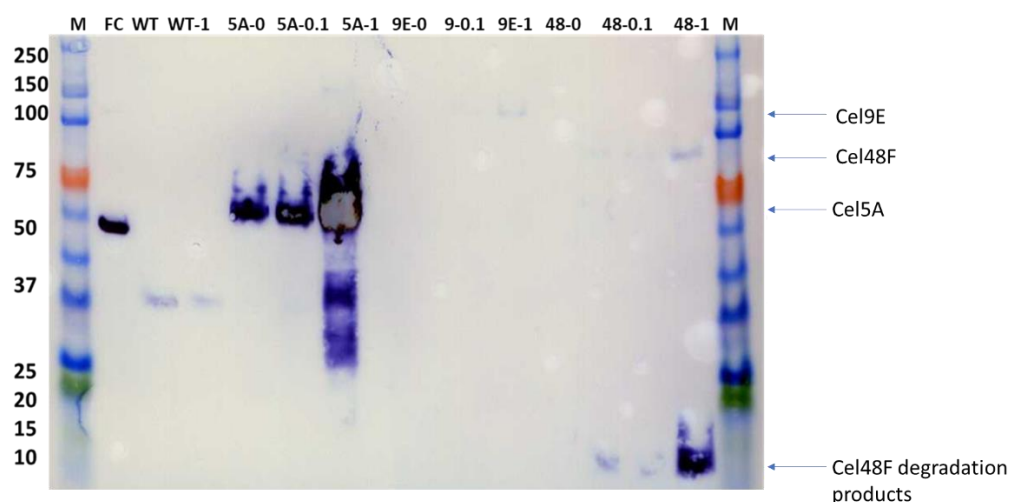


Figure 3.2. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing cellulases derived from *R. cellulolyticum* under lactose induction. Supernatants were prepared as described in 2.4 and concentrated 40-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; 5A: strain expressing Cel5A; 9E: strain expressing Cel9E; 48: strain expressing Cel48F; 0, 0.1 and 1 indicate the level of induction by lactose (mM).

For Cel5A, a single band of the expected size, 60 kDa can be seen in all samples, even in the non-induced sample, indicating that the system is leaky. There is no difference in expression in the 0.1 mM induced sample compared to the non-induced. However, there is a clear increase in expression of Cel5A when induced with 1 mM lactose.

For Cel9E, a faint band can be detected at the expected size of 100 kDa in the sample induced with 1 mM lactose, while no expression can be detected in the 0 and 0.1 mM samples.

For Cel48F, a very faint band at 90 kDa can be detected in the 0 and 0.1 mM induced samples and a clear band for the 1 mM sample. There is also a strong band at around 10 kDa for the sample induced with 1 mM lactose, with smaller bands for the 0 and 0.1

mM samples. This indicates the degradation of Cel48F at the dockerin domain which has been observed previously in *C. acetobutylicum* (Willson et al. 2016), as well as and when it was purified from *E. coli* (C. Reverbel-Leroy et al. 1997).

Higher levels of induction were later tested to attempt to increase the expression levels of Cel48F and Cel9E. All three strains were grown as before and induced with either 0, 1, 10 and 50 mM of lactose. Interestingly, at inductions above 1 mM of lactose, the strains expressing Cel9E and Cel48F showed a marked decrease in growth rate. (see Figure 3.3). This was shown most clearly when induced with 50 mM of lactose resulting in a strain which stopped growing after 1 hour of induction. For these strains, instead of growing to OD 1.0, samples were taken when the 10 mM sample had reached OD 1.0. Strains expressing Cel5A showed no difference in growth rate, regardless of the level of induction (Data not shown).

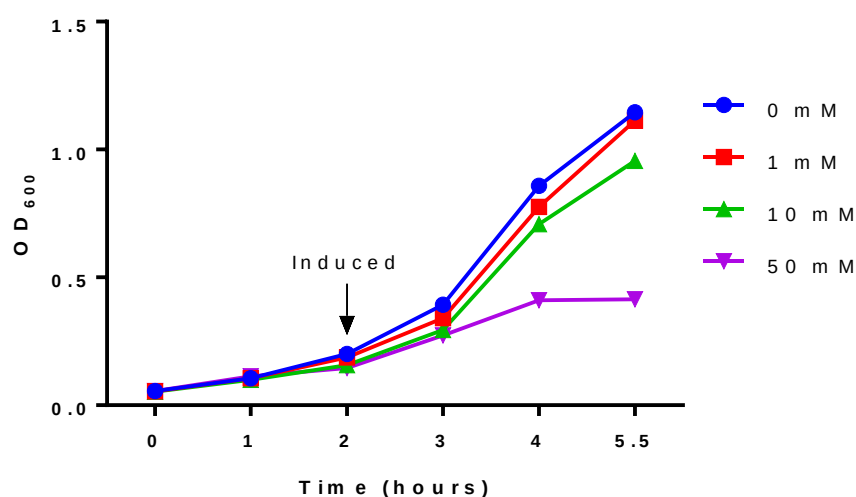


Figure 3.3. Showing the growth rate of a *C. acetobutylicum* strain expressing *Cel48F* induced with 0, 1, 10 and 50mM of lactose.

To allow for better comparison of expression, Cel5A samples were loaded in a separate blot. No difference in expression of Cel5A was detected between the 1, 10 and 50 mM samples, with no expression detected in the non-induced sample (Figure 3.4). It should also be noted that Cel5A is also the highest expressed protein in the supernatant fraction as determined by a Coomassie-stained gel (Data not shown).

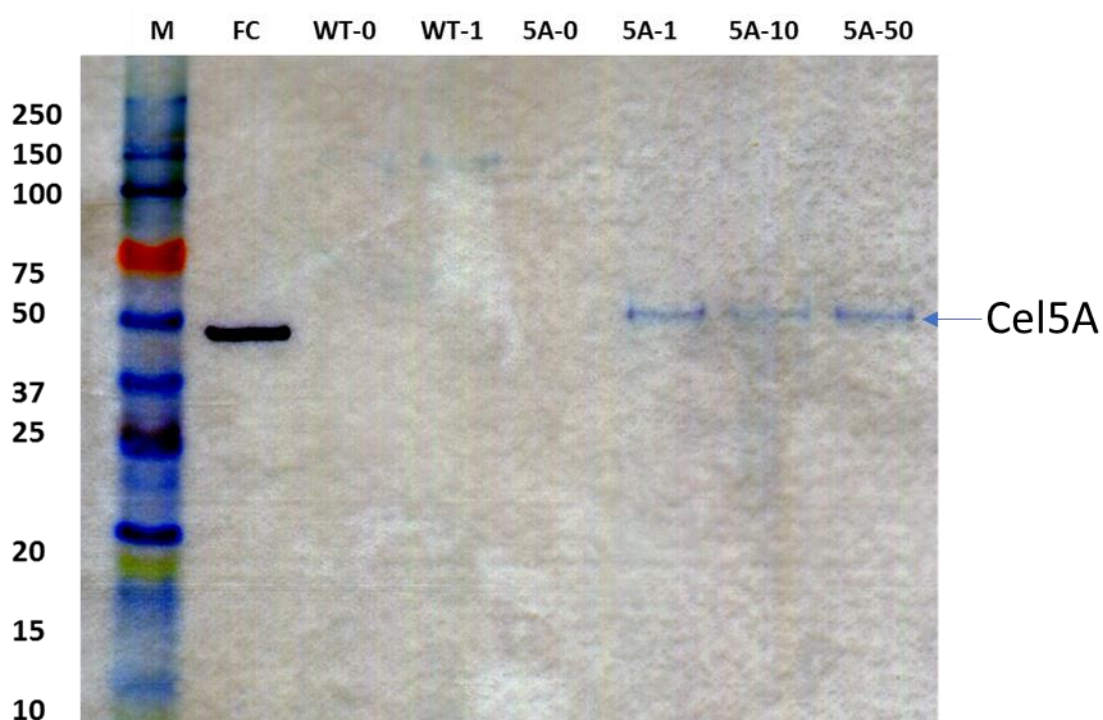


Figure 3.4. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing the Cel5A cellulase from *R. cellulolyticum* under control of the *PtcdB* promoter in a *Li-tcdR* background. Supernatants were loaded non-concentrated. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; 5A: strain expressing Cel5A; 0, 1, 10 and 50 indicate the level of induction by lactose (mM).

For both Cel9E and Cel48F induction at 10 mM of lactose had the highest level of expression, with induction at 1 mM being significantly lower. The 50 mM induced sample had a slightly decreased level of expression compared to the 10 mM induced sample; a potential consequence of the slow growth of these cultures. Degradation of Cel48F was again observed, the amount of degradation correlating with the level of expression. Degradation products for Cel9E were also observed, although unlike Cel48F degradation, occurring at the dockerin domain (Figure 3.5).

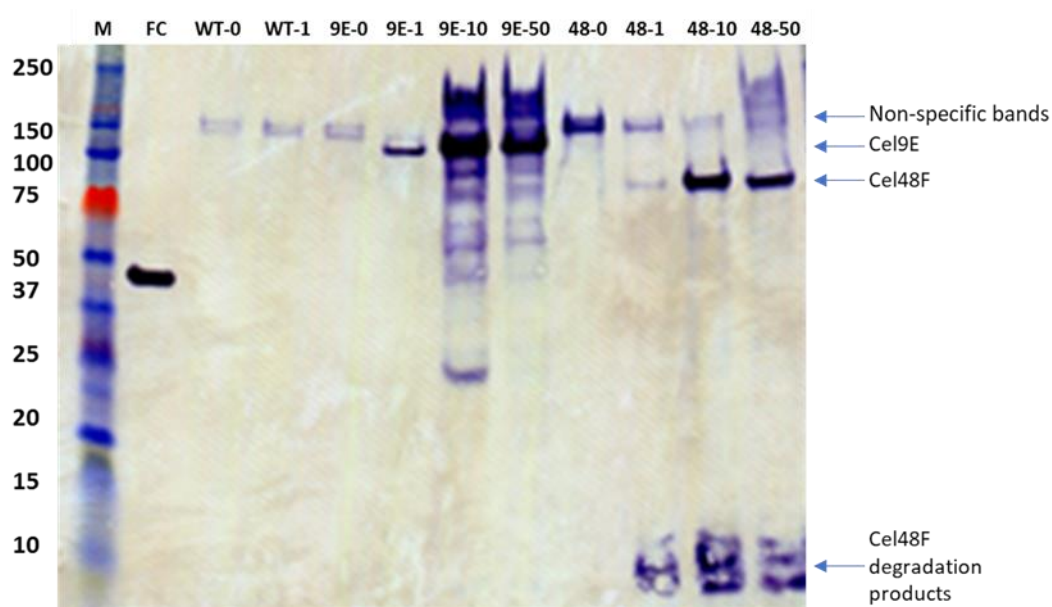


Figure 3.5. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing the Cel9E and Cel48F cellulases from *R. cellulolyticum* under control of the *PtcdB* promoter in a *Li-tcdR* background. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; 9E: strain expressing Cel9E; 48: strain expressing Cel48F; 0, 1, 10 and 50 indicate the level of induction by lactose (mM).

3.2.3 Comparison of the lactose inducible system with the *facoid* promoter

Previously, the highest expression of Cel48F was achieved by integrating Cel48F into the genome and placing it under the control of the *facoid* promoter (Willson, 2014) (*pyrE::facoid-Cel48F_N*). Here, we aimed to compare the two expression systems. Strains were grown as before, in CBM (40 mM MOPS, pH 6.8) to an OD of 1.0 and if appropriate induced at OD₆₀₀ 0.3.

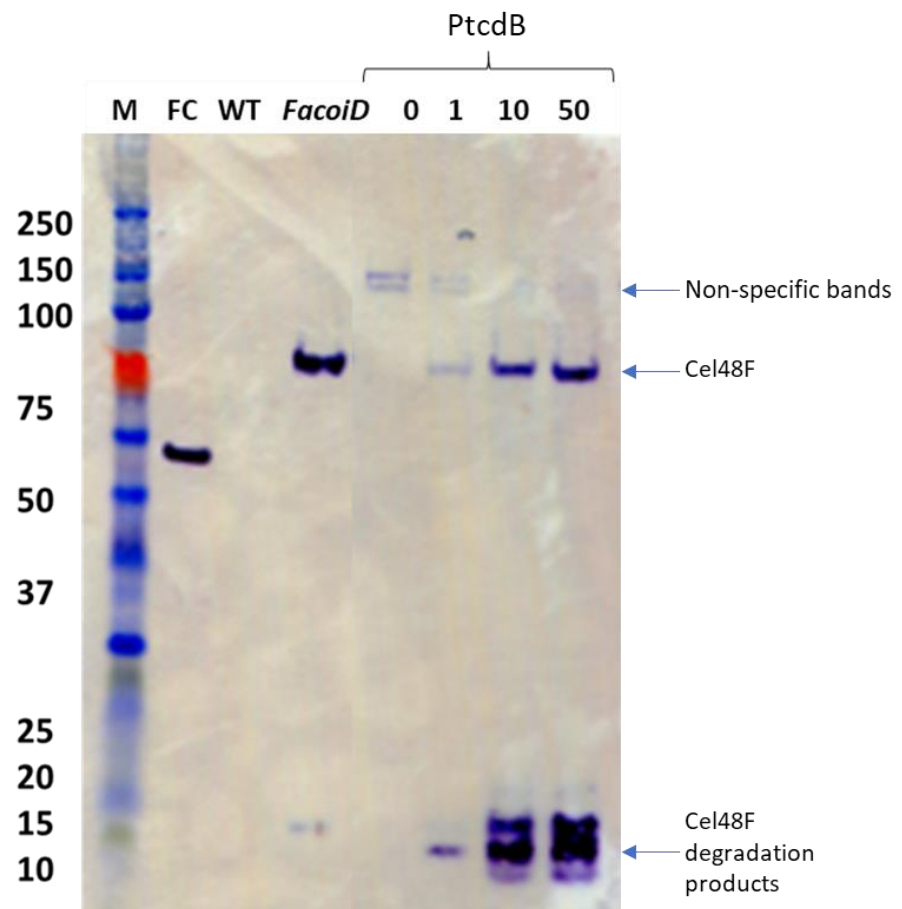


Figure 3.6. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing the Cel48F cellulase from *R. cellulolyticum* either under the control of the *FacOID* promoter or the *PtcdB* promoter. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; *FacOID*: Cel48F expressed under control of the *facOID* promoter; *PtcdB*: Cel48F expressed under control of the *PtcdB* promoter in the *Li-tcdR* background; 0, 1, 10 and 50 indicate the level of induction by lactose (mM).

Surprisingly, the highest level of expression was achieved with Cel48F under control of the *facOID* promoter, when compared to the highest level of induction of the Li-tcdR strain. (Figure 3.6). Interestingly, the amount of degradation is significantly increased in all samples expressed from the lactose inducible system compared to the expression from the *facOID* promoter.

The degradation product is hypothesised to be the dockerin domain. In the *facoid* expression system, Cel48F with its native dockerin domain was used, while in the lactose inducible system, the dockerin domain from *R. thermocellum*'s Cel48S was used. This implies that the replacement of the dockerin domain results in significantly increased levels of degradation, potentially explaining the difference in expression levels between the lactose-inducible system and the *facOID* promoter.

Based on these observations, we set out to test the lactose inducible system to compare the level degradation of Cel48F with either; no dockerin domain, Cel48F's native dockerin domain or the dockerin domain from *R. thermocellum*'s Cel48S cellulase.

Constructs with either no dockerin domain or the native dockerin domain were cloned into the pMTL82151 vector downstream the *PtcdB* promoter, as before. However, Cel48F with its native dockerin domain was not stably expressed, with the phenotype often lost after revival from glycerol stocks. To alleviate these issues, the constructs were instead integrated into the genome, as this has previously been demonstrated to result in stable expression. Constructs were cloned into the pMTL-JH16 vector under control of the *PtcdB* promoter and integrated into the *thiolase* locus. The strong T1T2 terminator was also included upstream of the *PtcdB* promoter to prevent read through from the *thiolase* promoter. Strains were grown as before, in CBM (40 mM MOPS, pH 6.8) to an OD of 1.0 and induced with 1 mM lactose at OD 0.3.

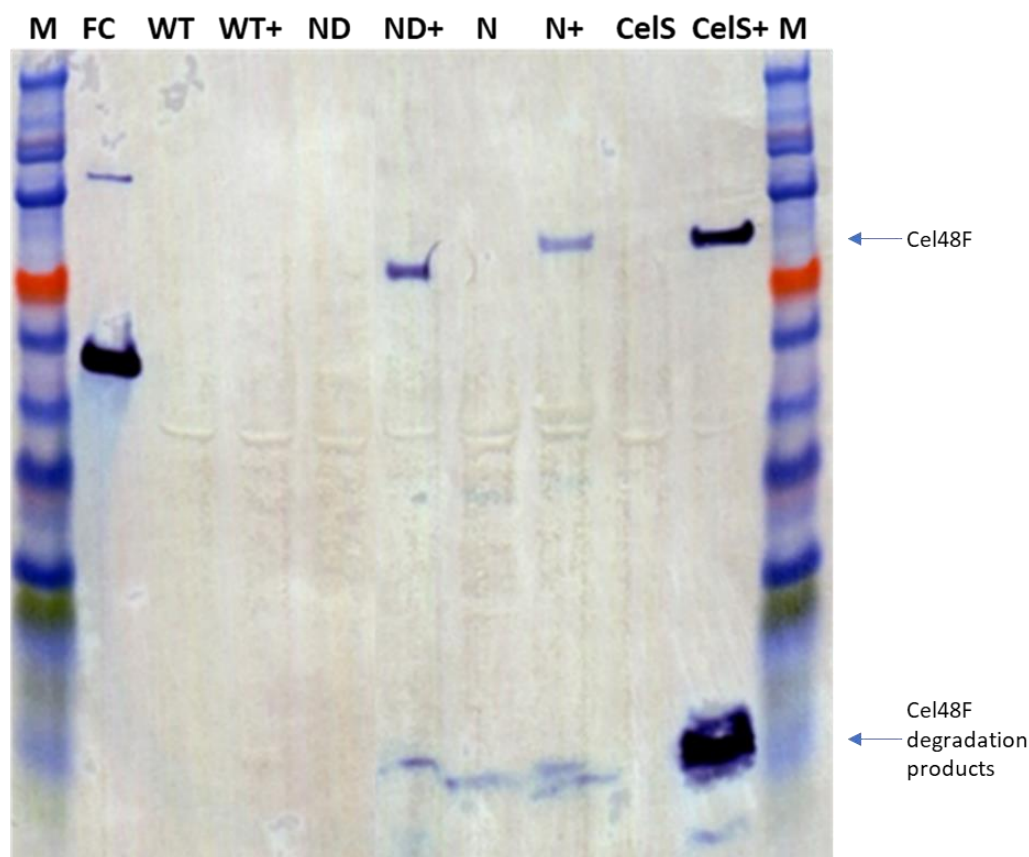


Figure 3.7. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing the Cel48F cellulase from *R. cellulolyticum* with different dockerin domains. Cel48F was under control of the PtcdB promoter in a Li-tcdR background. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; ND: strain expressing Cel48F without a dockerin domain; N: strain expressing Cel48F with its native dockerin domain; CelS: Strain expressing Cel48F with the CelS dockerin domain from *R. thermocellum*; +: indicates induction with 1 mM of lactose.

The expression level of full-length Cel48F was similar for all variants of Cel48F (Figure 3.7), with slightly higher expression of Cel48F with the Cel48A dockerin from *R. thermocellum*, compared to Cel48F with its native dockerin. Cel48F without a dockerin domain appeared to have an intermediate level of expression. However, there was significantly increased levels of degradation of Cel48F with the Cel48S dockerin from *R. thermocellum*, compared to Cel48F with a native dockerin domain.

3.2.4 Effect of the addition of protease inhibitors on Cel48F expression

Secreted Cel48F variants are potentially degraded in the culture supernatant by proteases secreted by the host. Inactivation of proteases which act in the cell wall or secreted into the growth medium, have been shown to result in an increase in protein expression in *B. subtilis* (Ye, R. and S. L. Wong, 1996). The addition of protease inhibitors to the growth medium has also been shown to have the same effect (Westers et al. 2008).

In *B. subtilis* the expression of cellulases was increased in a strain with all feeding proteases inactivated (Strain WB800). It is therefore possible that feeding protease genes are present in the genome of *C. acetobutylicum* and might be responsible for the low levels of full-length Cel48F secretion and its degradation.

Expression of two GH9 family cellulases in *B. subtilis* was different, with Cel9E being expressed at significantly lower levels than Cel9G (Anderson et al. 2011). The same trend was observed in *C. acetobutylicum* with significantly higher levels of Cel9G expressed compared to Cel9E (Willson, 2014). This suggests that there might be a common factor, potentially related to feeding proteases, influencing the secretion of some cellulases in *B. subtilis* and *C. acetobutylicum*. Multiple proteases with a putative signal peptide indicating that they are secreted have been identified in *C. acetobutylicum*'s genome (Desvaux et al. 2005). Creation of a protease deficient strain similar to *B. subtilis*'s WB800 would have required the deletion of multiple putative proteases. Instead the addition of protease inhibitors to the growth media was investigated, as this has been shown to have a similar effect.

To determine if the addition of protease inhibitors will result in an increased level of Cel48F in the supernatant, strains were grown as before in CBM (40 mM MOPS, pH

6.8) to an OD of 1.0 and induced at OD₆₀₀ 0.3. Roche complete EDTA free protease inhibitor cocktail was added at the time of the induction.

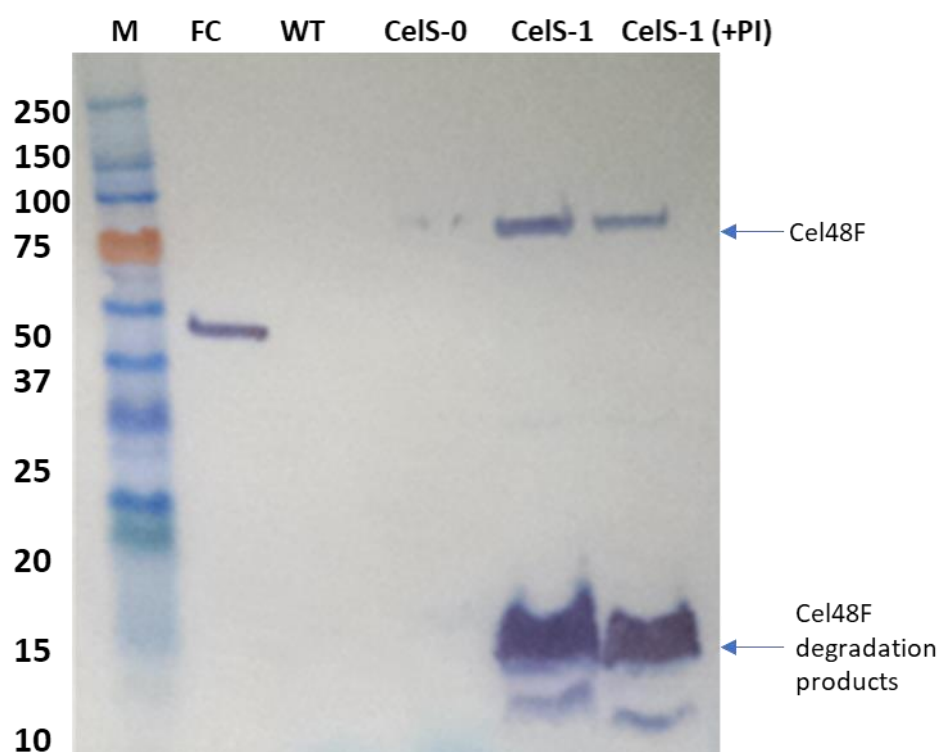


Figure 3.8. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing the Cel48F cellulase from *R. cellulolyticum* with protease inhibitors added. Cel48F was under control of the PtcdB promoter in a Li-tcdR background. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; CelS: Strain expressing Cel48F with the CelS domain from *R. thermocellum*; 0 and 1 indicates the level of induction by lactose (mM); +PI: The addition of protease inhibitors.

The stability of Cel48F in the presence of protease inhibitors was the same as the control with none added. This indicating that the addition of the protease inhibitors did not improve the stability of Cel48F and high levels of degradation products were still present (Figure 3.8).

3.3 Discussion

By utilising an inducible expression system, stable plasmid-based expression of Cel48F and Cel9E was achieved. However, the expression of Cel5A was significantly higher than the expression of Cel48F and Cel9E despite utilising the same expression system. Induction at levels above 1 mM lactose for Cel48F and Cel9E resulted in slow growth, and inhibition of growth at 50 mM lactose, clearly indicating that high levels of expression of these cellulases is toxic to *C. acetobutylicum*. This has also been observed in *B. subtilis* previously, with a high level of induction of cellulases from *R. thermocellum*, resulting in a decrease in cell growth and a reduction in cellulase expression (C.-C. Lin et al. 2017).

The reason for the difference in the expression or perhaps secretion of different cellulases using the same expression system is unclear. It is possible that differences occur already at the transcription or at the translation levels, some transcripts being more stable than others. Some of the cellulases could be more toxic than others to the cell. This hypothesis is supported by the observation that higher induction levels resulted in growth inhibition. As Cel48F has been reported to be expressed at fairly high levels intracellularly, it is likely that secretion of this and other cellulases require additional factors for their secretion to be optimal in *C. acetobutylicum* (Mingardon et al. 2011). A reduced growth rate has been observed in *B. subtilis* when proteins are overexpressed and the authors suggested that this might be caused by secretion stress, due to the accumulation of misfolded proteins (Darmon et al. 2002).

The expression of Cel48F was slightly lower using the lactose inducible system compared to expression from the *facoid* promoter. This was thought to be due to high levels of degradation, and the use of the Cel48S dockerin domain from *R. thermocellum*.

The increased level of degradation was shown to be caused by the replacement of the dockerin domain, with significantly higher levels of degradation seen when Cel48F's native dockerin domain was substituted for a dockerin domain from *R. thermocellum*. However, there were only minor differences in the level of full-length Cel48F expression, and that the use of the *facoid* promoter results in slightly higher levels of Cel48F in the culture supernatant.

It is unclear why the use of different dockerin domains resulted in significant differences in the levels of degradation. It is possible that the dockerin domain CelS from *R. thermocellum* is more stable than Cel48F's native dockerin domain as *R. thermocellum* is thermophilic and proteins from thermophiles are generally more stable (Szilágyi and Závodszky 2000). Additionally, thermophilic proteins have also been shown to be more resistant to proteolysis than mesophilic homologues (Daniel et al. 1982). Alternatively, differences in the expression or secretion of the two different proteins may be responsible for the noted differences. Cel48F with a CelS dockerin domain may derive from a more stable mRNA resulting in higher level of protein (Gaspar et al. 2013), or it may fold more efficiently, again, resulting in higher levels of secreted protein.

Surprisingly, the addition of protease inhibitors did not affect the level of the full-length protein or the level of degradation products seen. Indicating that it is unlikely that the degradation of Cel48F by feeding proteases is the main reason for the low level of mature Cel48F in the culture supernatant.

It is possible that the cellulases are degraded during secretion by cell-wall bound proteases, possibly activated by high levels of misfolded Cel48F (Darmon et al. 2002). This was suggested by other studies where intracellular expression of Cel48F was not

toxic while the secretion of Cel48F was toxic to *C. acetobutylicum* (Mingardon et al. 2011).

3.4 Key outcomes

- Successful expression and secretion of *R. cellulolyticum* cellulases Cel5A, Cel48F and Cel9E was demonstrated using a lactose inducible system. The inducible system allowed, for the first time, the creation of stable strains expressing the Cel9E and Cel48F cellulases using a plasmid-based system.
- High induction levels resulted in growth impairment of *C. acetobutylicum* strains expressing Cel9E and Cel48F, but not Cel5A.
- The addition of a protease inhibitor cocktail to the actively growing *C. acetobutylicum* cultures expressing Cel48F did not reduce the amount of Cel48F degradation products seen in the culture supernatant, indicating that feeding proteases are probably not responsible for the low levels of secreted Cel48F.

Chapter 4 Efforts to increase the expression level of GH48 family cellulases in *C. acetobutylicum*.

In the previous chapter, a lactose-inducible system was used to express cellulases from the GH5, GH9 and GH48 families. There was a marked difference in the level of expression of the cellulases, with the expression of GH9 and GH48 family cellulases significantly lower than that of GH5 cellulases. This has been observed previously both in our lab and by other groups (Mingardon et al. 2011; Willson et al. 2016).

A difference in the level of expression and secretion of different proteins under control of the same promoter is common. There are several reasons for this, and they include, issues in the transcription and/or translation of the proteins; for example, mRNA stability, codon usage as well as issues related to the secretion of the proteins (Komar 2016; Lambertz et al. 2014).

In this chapter, several attempts were made to increase the expression level of GH48 family cellulases. Firstly, the expression and secretion level of different GH48 family cellulases in *C. acetobutylicum* was assessed. This was to determine if there was a specific issue with the secretion of Cel48F (*R. cellulolyticum*'s GH48 family cellulase) in particular or secretion of GH48 family cellulases in general.

Secondly, several attempts were made to improve the expression of Cel48F, *R. cellulolyticum*'s GH48 family cellulase. 5' and 3'-stem-loops were utilised to increase the stability of Cel48F mRNA transcripts, potentially resulting in improved expression (Phan, Nguyen, and Schumann 2013). PrsA chaperones were also overexpressed as it was hypothesised that the poor expression level might be due to the poor folding of Cel48F during secretion.

4.1 Expression of different GH48 family cellulases in *C. acetobutylicum*

4.1.1 Introduction

Family GH48 cellulases are key enzymes in many lignocellulose degrading systems and have been found in all the cellulosomal systems described to date (Artzi, Bayer, and Morais 2017). They are found in clostridia, bacilli and actinobacteria as well as in fungi. In most bacteria, they are one of the most abundant cellulases, expressed at approximately a 1:1 ratio with the major structural scaffoldin protein (Osiro et al. 2017; Badalato et al. 2017). In several bacteria, it has been shown that deletion of the GH48 gene results in a strain with a reduced ability to grow on crystalline cellulose, or in the case of *R. cellulolyticum*, it resulted in a strain unable to grow on crystalline cellulose (T. Xu et al. 2014). Most organisms contain only one GH48 family cellulase gene, although some do contain two or more, for example, *R. thermocellum* (Izquierdo, Sizova, and Lynd 2010).

GH48 family cellulases normally consist of a single catalytic subunit along with a dockerin or in non-cellulosomal systems a CBM (Sukharnikov et al. 2012). Some also have fibronectin type II (FN3) domains, the function of which is unclear, however, they may enhance cellulose hydrolysis by modifying the cellulose surface (Kataeva et al. 2002). Additionally, in some organisms, they are present in multi-domain enzymes, for example, *Caldocellulsiruptor besci* expresses an enzyme consisting of a GH48 domain, a GH9 domain and a CBM domain (Yi et al. 2013).

GH48 family cellulases are usually endo-processive cellulases although GH48 enzymes with chitinase activity have also been reported (Fujita et al. 2006). They are normally active on amorphous or crystalline cellulose, with very low activity on CMC; some are multifunctional, showing activity on xylan (C.-C. Liu and Doi 1998; Zhang et al. 2010; Corinne Reverbel-Leroy et al. 1996). GH48 family cellulases usually have very low

activity on their own, but have very high synergy with other enzymes including cellulases from the GH9 and GH5 families (Vazana et al. 2010; Fierobe et al. 2005).

As *R. cellulolyticum*'s GH48 family cellulase, Cel48F, is poorly expressed in *C. acetobutylicum*, we decided to screen GH48 family cellulases from other organisms to determine if they could be expressed at higher levels. This would also determine if there is a specific issue in the expression of *R. cellulolyticum*'s GH48 cellulase or a more general issue in the expression of GH48 family cellulases.

GH48 family cellulases from mesophilic, cellulolytic bacteria were selected. These were cellulases either from bacteria which use free-enzyme systems to degrade cellulose or species which use a cellulosomal system. For this reason, the GH48 family cellulases from *C. cellulovorans*, *C. papyrosolvens* (C7) and *C. phytofermentans* were chosen.

C. cellulovorans is a mesophilic, cellulolytic organism which has been shown to form a functional cellulosome; its cip-cel operon has a high homology to *R. cellulolyticum*'s (Doi and Tamaru 2001). Its GH48 family cellulase (ExgS) has been previously characterised; it can digest Avicel, Crystalline cellulose and Xylan but not CMC (C.-C. Liu and Doi 1998). ExgS's GH48 domain is quite similar to Cel48F with a 63% amino acid similarity. *C. papyrosolvens* is another mesophilic, cellulolytic bacteria which forms a cellulosome (Pohlschröder, Leschine, and Canale-Parola 1994), its GH48 family cellulase has not been characterised. It has the same structure as *R. cellulolyticum*'s Cel48F with a high amino acid similarity (88%).

C. phytofermentans is also a mesophilic, cellulolytic bacterium. However, it does not produce a cellulosome, instead secreting free enzymes (Warnick, Methé, and Leschine 2002). *C. phytofermentans* GH48 cellulase has been previously characterised; its GH48 domain shares only 48% amino acid similarity with Cel48F, it also contains a CBM3 and CBM X2 domain. It has been shown to have activity on Avicel and BAC, with low very

low activity on CMC. It was also successfully expressed and secreted in *B. subtilis* at high levels (Zhang et al. 2010).

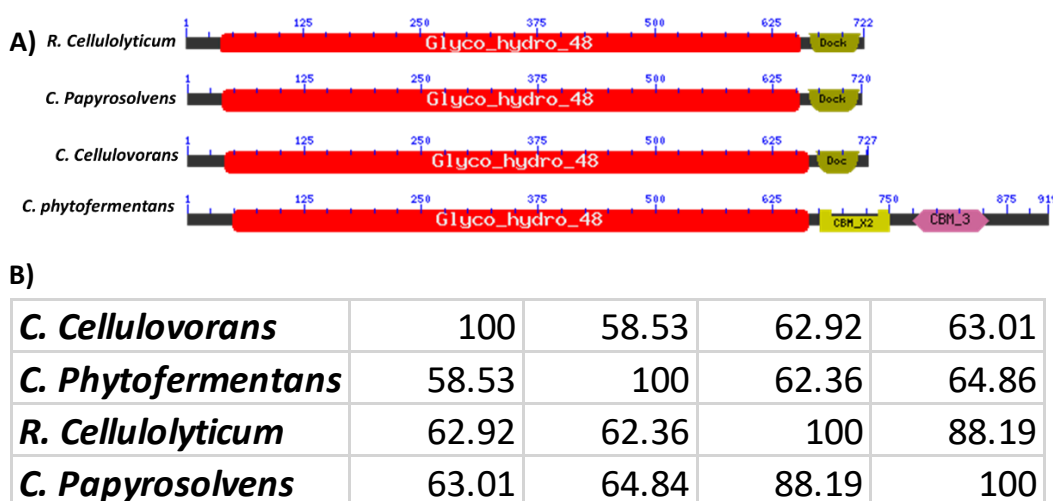


Figure 4.1. Comparison of GH48 family cellulases. A) Domain composition of the GH48 family cellulases from *R. cellulolyticum*, *C. papyrosolvens*, *C. cellulovorans* and *C. phytofermentans*. B) Percent identity matrix of the GH48 family cellulases from *R. cellulolyticum*, *C. papyrosolvens*, *C. cellulovorans* and *C. phytofermentans*.

4.1.2 Results

All GH48 family cellulases were synthesised by Biomatik, in the case of *C. cellulovorans* and *C. papyrosolvens* the dockerin domain was removed. The genes were cloned downstream of the *PtcdB* promoter in the pMTL82151 plasmid and transformed into *C. acetobutylicum* strain Li-tcdR. Strains were grown as before in CBM (40 mM MOPS, pH 6.8) to an OD of 1.0 and induced at OD₆₀₀ 0.3.

The expression of all GH48 cellulases was roughly equivalent to the expression of Cel48F, with a slightly increased level of expression for ExgS from *C. cellulovorans* and a slightly decreased level of expression for Cphy3368 from *C. phytofermentans* (Figure 4.2). To determine if there was increased enzyme activity compared to Cel48F an enzyme activity assay on PASC was carried out. Supernatants were concentrated 10-fold and dialysed to remove any residual sugar. They were then incubated with 2% PASC (100mM Tris-HCL, 1mM CaCl₂ pH 5.5) for 72 hours at 37 °C and the amount of

reducing sugars determined using DNS assay. These conditions were previously found to be optimal for a *C. acetobutylicum* strain expressing a mini-scaffoldin consisting of Cel48F, Cel9G and Xyn10 (Willson et al. 2016). Unfortunately, no difference in enzyme activity could be detected for any recombinant strain compared to wild-type *C. acetobutylicum* (Data not shown).

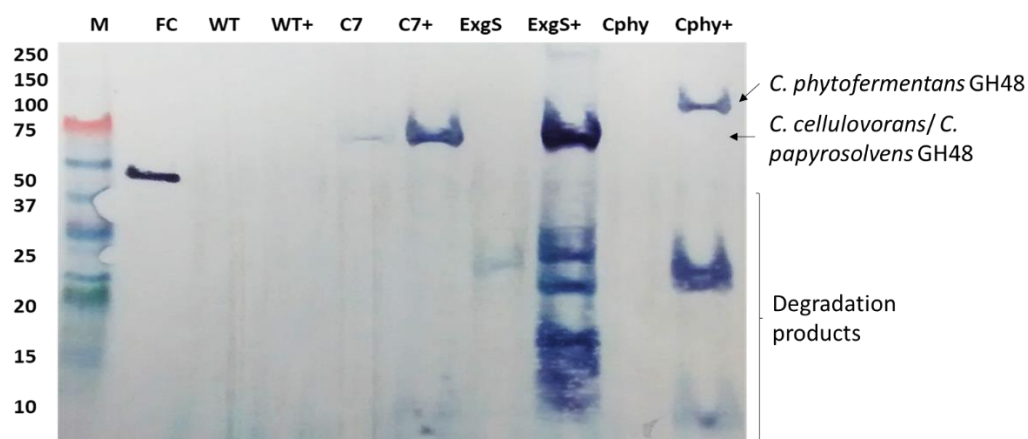


Figure 4.2. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing different GH48 family cellulases.

All cellulases were under control of the *PtcdB* promoter in a *Li-tcdR* background. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; C7: GH48 family cellulase from *C. papyrosolvens*; ExgS: GH48 family cellulase from *C. cellulovorans*; Cphy: GH48 family cellulase from *C. phytofermentans*. + indicates induction with 1 mM lactose

4.2 mRNA stability

4.2.1 Introduction

It was hypothesised that the low expression levels of GH48 family cellulases could be due to mRNA instability, this is for several reasons: (1) the expression level of Cel48F is lower than other cellulases even when expressed from the same promoter, (2) the expression level of CipA scaffoldin variants with 2, 3 and up to 9 cohesin domains is influenced by the length of the transcript, perhaps due to mRNA stability (Kovács et al. 2013) and (3) the regulation of the cip-cel operon in all mesophilic cellulosomal systems is controlled by the presence of 3' and 5'-stem-loops. For example, a 3'-stem-loop is located directly after all the highly expressed components in *R. cellulolyticum* (C. Xu et al. 2015).

Increasing the stability of mRNA transcripts can lead to higher expression levels due to a longer mRNA half-life resulting in a higher translation rate per mRNA (Grunberg-Manago 1999). The presence of a stem-loop in the 5' and 3' UTR (Untranslated Region) is one the major factors affecting the stability of mRNA (Régner and Arraiano 2000). In general, the lower the ΔG of the stem-loop, the higher the stability. The main reason that stem loops increase stability is that they help prevent degradation of the mRNA by endoribonucleases (Gu, Zhou, and Wilke 2010; Bugaut and Balasubramanian 2012).

However, the stability of mRNA can be affected by several other factors including; the strength of the RBS (Ribosomal binding site) with a stronger RBS resulting in a longer half-life (Jürgen, Schweder, and Hecker 1998). The overall structure of the mRNA can also have an impact on mRNA stability, as improperly folded mRNA can result in the translation initiation site being blocked or the movement of ribosomes impeded (Gaspar et al. 2013). It was also found that the length of the linker region between the 5'UTR and RBS also had an impact on expression. Importantly, the effect of free

nucleotides before the stem-loop significantly reduced the stability of the mRNA (Phan, Nguyen, and Schumann 2013). Additionally, codon usage can have an impact on mRNA stability as well as the translation rate (Presnyak et al. 2015); previously Cel48F has been codon optimised for *C. acetobutylicum*, and no increase in expression was observed (Ben Willson, personal communication).

Numerous studies have linked specific 5'UTR's with an increase in protein expression due to increased mRNA half-life. An example of this is the 5'UTR from the *slpA* gene from *Lactobacillus acidophilus* was used to increase expression of α -amylase in *Lactobacillus casei* and *Lactococcus lactis*. It was shown to improve mRNA half-life due to its 5' stem-loop structure (Narita et al. 2006).

Another study in *B. subtilis* was able to increase the mRNA half-life of a reporter gene *BgaB* from 8 minutes to over an hour resulting in significantly increased expression and activity of *BgaB*. This was achieved via the addition of a synthetic 5' UTR (Cose) which formed a stem-loop in combination with a strong artificial RBS (Gsib) (Phan, Nguyen, and Schumann 2013).

Additionally, during this study, a paper was published which utilised 5' stem-loops and strong terminators (3'UTR's) to increase gene expression in *C. acetobutylicum*. The addition of a strong terminator increased expression by 36%. While the addition of a 5' stem-loop was shown to further increase gene expression by 4.6 fold, this correlated with an increase in mRNA levels (J. Lee et al. 2016).

R. cellulolyticum and other cellulosome-producing organisms arrange their cellulosome genes into operons (cip-cel cluster). A recent paper showed that the cellulosome stoichiometry and the stability of the mRNA transcripts of the cip-cel cluster in *R. cellulolyticum* is regulated by stem-loops (C. Xu et al. 2015). In the cip-cel operon, several stem-loops are found, one in the 5'UTR upstream of CipC, and four

3'UTR stem-loops, after each of the highest expressed components (CipC, Cel48F, Cel9G and Cel9E) (Figure 4.3). The cip-cel operon was confirmed to be transcribed as one unit. However, it was found to be processed, with several monocistronic transcripts consisting of CipC, Cel48F and Cel9E and several polycistronic transcripts (C. Xu et al. 2015).

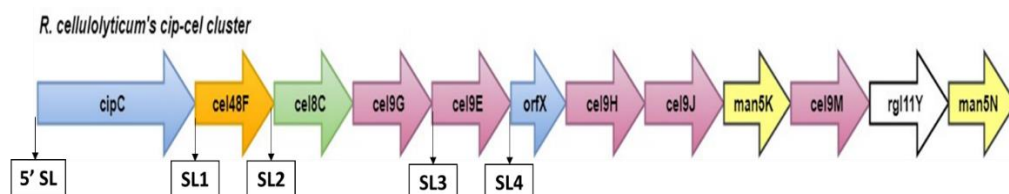


Figure 4.3. The position of stem-loops in *R. cellulolyticum*'s Cip-Cel operon. 5' SL: 5' stem-loop at the start of the transcript; SL 1-4: 3' stem-loops. Stem-loop structures were identified by C. Xu et al. 2015.

The full-length transcript is processed by endonucleases, this is directed by the presence of cleavage signals present in the intergenic regions. The stability and abundance of these transcripts is due to the presence of stem-loops at the 3' region, preventing degradation by exonucleases. This is surprising as normally the presence of stem-loops in the 5' region is more important. Additionally, different stem-loops resulted in a difference in the abundance of transcripts, as stem-loops with the lowest ΔG resulted in the highest transcript levels (C. Xu et al. 2015).

Additionally, a stem-loop was found to be present after GH48 family enzymes in all other organisms with a cip-cel operon except *C. acetobutylicum* (C. Xu et al. 2015). This shows that the use of stem-loops in the 5' and 3'UTR may be necessary for the high expression of Cel48F and other cellulases in *R. cellulolyticum*, so the low expression of cellulases in *C. acetobutylicum* may be due to the mRNA transcript being unstable resulting in low expression.

4.2.2 Effect of 5' and 3' stem-loops on the expression of Cel48F

To test the influence of the 5' and 3' stem-loops on the expression levels of cellulases, and in particular of that from the GH48 family, several 5'-stem-loops were added to *PtcdB* promoter upstream of the Cel48F cellulase.

Free nucleotides that are not part of the stem-loop in 5'UTR's have been shown to reduce mRNA stability (Phan, Nguyen, and Schumann 2013), so it was necessary to add the 5' stem-loop directly after the transcription start site of the *PtcdB* promoter. The transcription start sites of *PtcdB* has been determined previously; it contains two sites, as shown in Figure 4.4 (Hundsberger et al. 1997). Here, the 5'UTR's were added immediately after the first transcription start site.

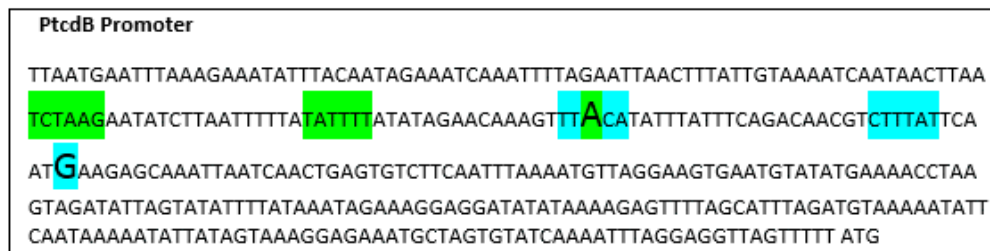


Figure 4.4. The transcription start sites of the *PtcdB* promoter. *PtcdB1* and *PtcdB2* -10 and -35 regions are highlighted green and blue respectively with the transcription start sites capitalised and in bold.

In silico analysis of the *PtcdB* 5' UTR's revealed that both transcripts have a long 5'UTR with stem-loop structures. For *PtcdB1* there are several possible structures with a ΔG ranging from -34.8 to -36.5, while *PtcdB2* has only one possible structure with a ΔG of -28.87. Both structures have a negative ΔG comparable to other 5'UTR's and a complicated secondary structure which could act as a possible 5'-stem-loop (Figure 4.5). Previously, we used a *PtcdB* promoter variant in which the RBS sequence of the native *PtcdB* was replaced with the RBS sequence from *C. acetobutylicum*'s thiolase gene, as this RBS is known to be very strong in its native host. However, replacement

of the RBS sequences resulted in disruption of the secondary structure of the mRNA, which might have had a negative effect on mRNA stability. To address this issue, the native PtcdB promoter with its native RBS was tested.

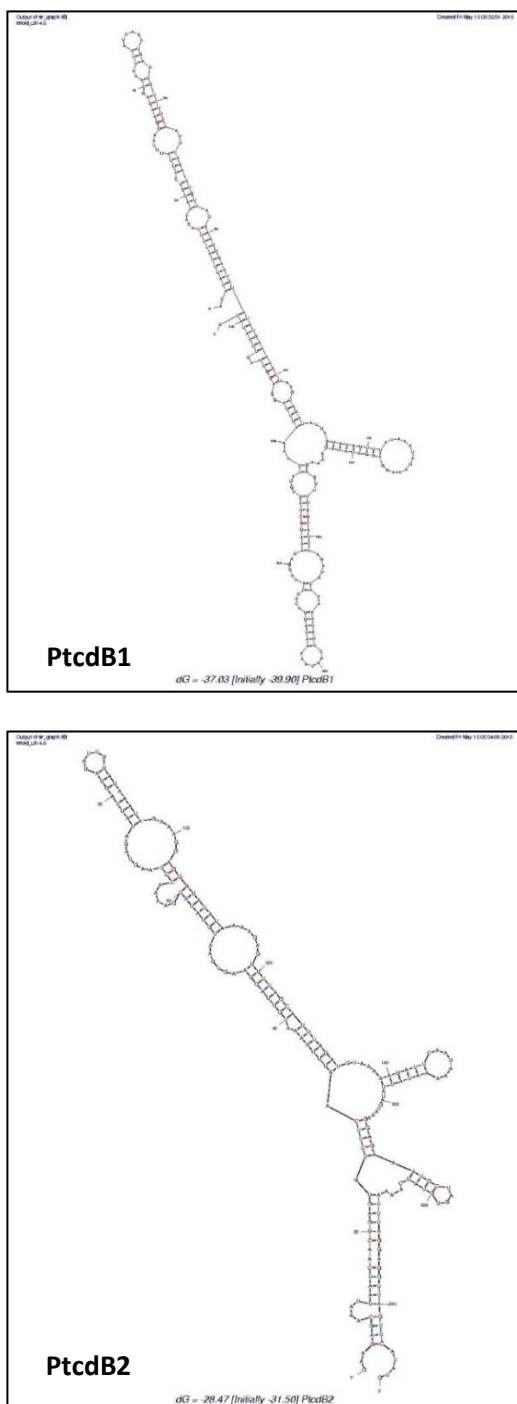


Figure 4.5. Predicted RNA structures of PtcdB1 and PtcdB2 5'UTR's. The mRNA structures were predicted using mFold.

Additionally, two proven 5' stem-loop structures were also chosen, SlpA as it is a natural 5'UTR and COSE which is an artificial 5'UTR. As a strong RBS has also been shown to increase mRNA stability, we used the *C. acetobutylicum*'s native strong *thiolase* RBS. However, in SlpA 5'UTR the *thiolase* RBS could not be substituted without disrupting the stem-loop, so the natural RBS was used instead. For the artificial COSE 5'UTR, both the Gsib RBS and the *thiolase* RBS was used (Figure 4.6).

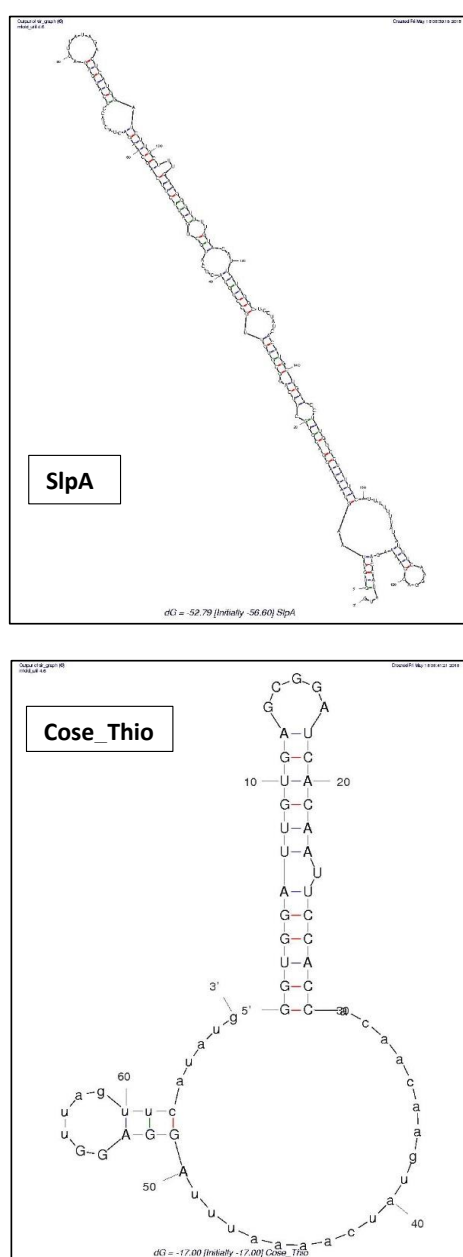


Figure 4.6. Predicted 5'UTR structures of SlpA and Cose_Thio. The structure for Cose_Gsib is extremely similar to Cose_Thio (not shown). The mRNA structures were predicted using mFold.

All 5'UTR's were synthesised together with the *PtcdB* promoter by Biomatik in BB2 format and cloned into the pMTL-82151 plasmid upstream of Cel48F without a dockerin domain. The cloning of constructs, in particular, the construct with the Cose_Thio 5'UTR were extremely toxic to *E. coli* with incubation having to take place at room temperature. Strains were transformed into *C. acetobutylicum* and grown as before in CBM (40 mM MOPS, pH 6.8) to an OD of 1.0 and induced at OD₆₀₀ 0.3. 1 mM of lactose was used for induction, as inhibition of growth has been observed previously when induced at 10 mM.

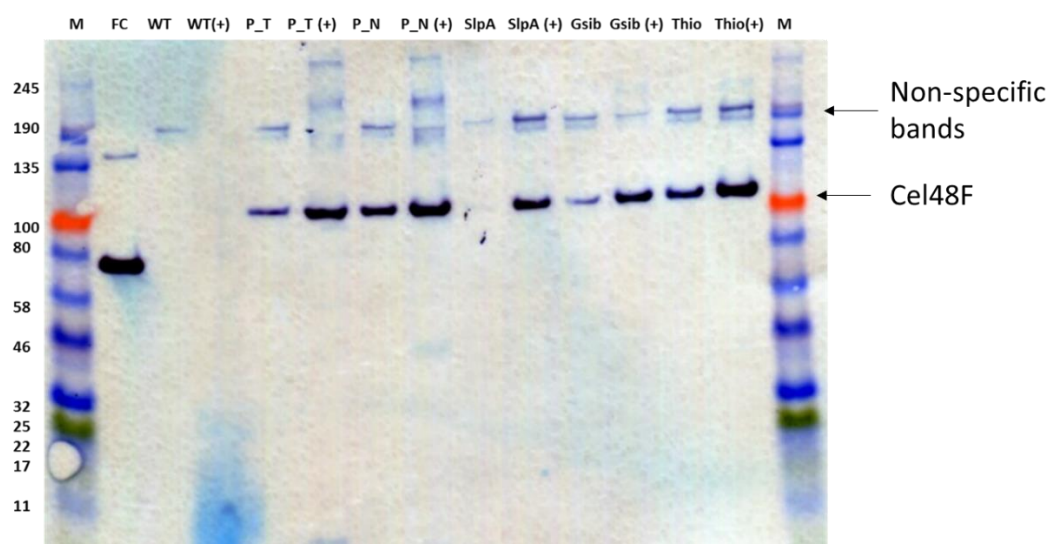


Figure 4.7. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing the Cel48F cellulase from *R. cellulolyticum* under control of the *PtcdB* promoter with different 5' UTR's. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; **P_T**:: Strain expressing Cel48F from the *PtcdB* promoter with a thiolase RBS (used previously); **P_N**: Strain expressing Cel48F from the *PtcdB* promoter with *PtcdB*'s native RBS; **SlpA**: Strain expressing Cel48F from the *PtcdB* promoter followed by *SlpA* 5'UTR and thiolase RBS; **Gsib**: Strain expressing Cel48F from the *PtcdB* promoter with COSE 5'UTR and *Gsib* RBS; **Thio**: Strain expressing Cel48F from the *PtcdB* promoter with COSE 5'UTR and thiolase RBS, +; induced with 1 mM lactose.

The expression of Cel48F was similar for all constructs, with only a small increase in expression for constructs with the PtcdB_N and Cose_thiolase 5'UTR compared to PtcdB_thiolase used previously (Figure 4.8).

To test whether the addition of a 3' stem-loop could further increase the expression of cellulases, 3' stem-loop structure were also tested, as these were shown to be important for the native expression of GH48 family cellulases (C. Xu et al. 2015). The 3' stem-loops from Cel48F (SL2) as well as the 3' stem-loop from *slpA* (*Lactococcus lactis*) were chosen, as they have both been shown to increase the expression of proteins previously. The 3' stem-loops were cloned downstream of Cel48F with the best 5'UTR (CoseThio 5'UTR). Strains were transformed into *C. acetobutylicum* and grown as before in CBM (40 mM MOPS, pH 6.8) to an OD of 1.0 and induced at OD₆₀₀ 0.3.

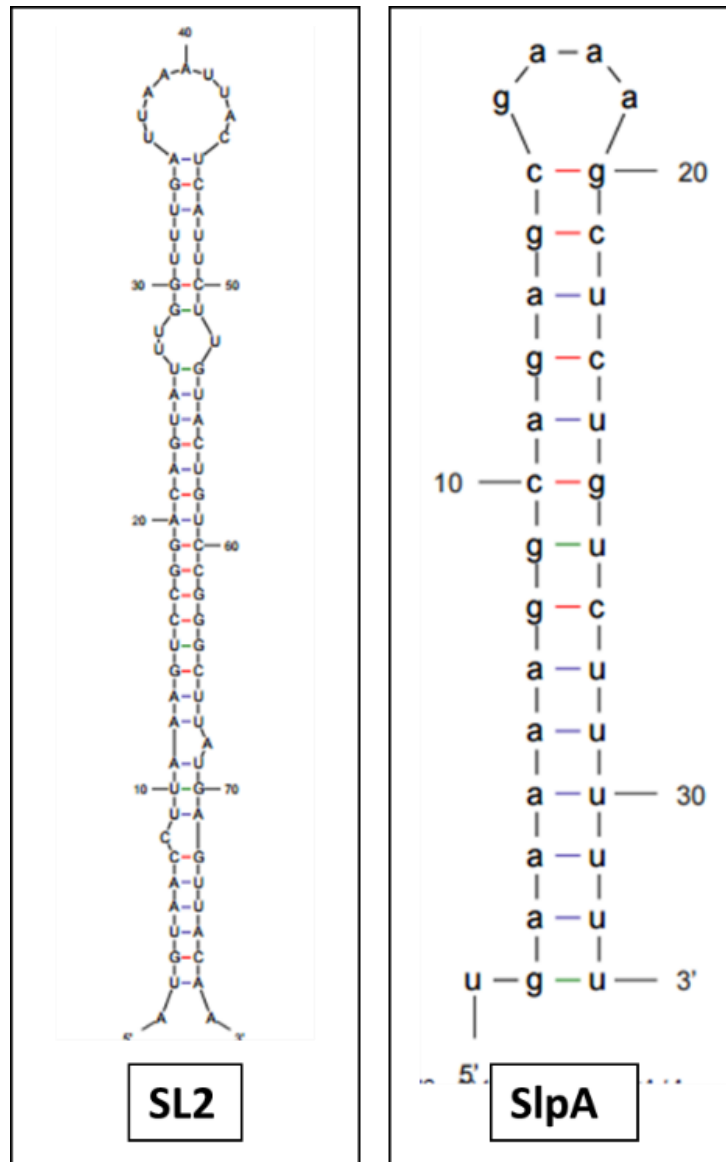


Figure 4.8. Predicted 3'UTR structures of SL2 and SlpA_T. The mRNA structures were predicted using mFold.

The addition of a 3' stem-loop structures did not significantly increase the expression of Cel48F in *C. acetobutylicum*. (Figure 4.9).

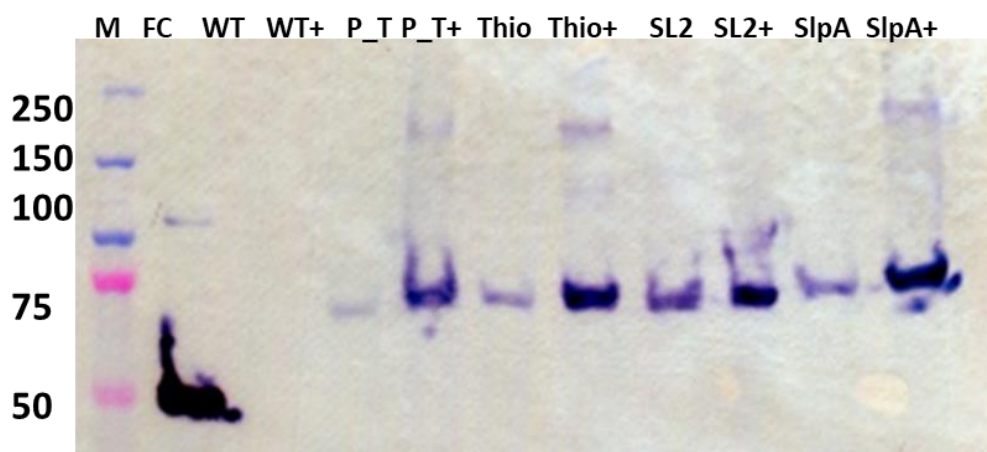


Figure 4.9. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains expressing the *Cel48F* cellulase from *R. cellulolyticum* under control of the *PtcdB* promoter with different 3' stem-loops. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; **P_T**:: Strain expressing *Cel48F* from the *PtcdB* promoter with a thiolase RBS **Thio**: Strain expressing *Cel48F* from the *PtcdB* promoter with COSE 5'UTR and thiolase RBS; **SL2**: Strain expressing *Cel48F* from the *PtcdB* promoter with COSE 5'UTR and thiolase RBS as well as *Cel48F*'s native 3' stem-loop, SL2; **SlpA**: Strain expressing *Cel48F* from the *PtcdB* promoter with COSE 5'UTR and thiolase RBS as well as the 3' stem-loop from *SlpA*; +: induced with 1 mM lactose.

4.3 Potential issues related to the secretion of GH48 family cellulases

The expression of GH48 family cellulases, in particular, *Cel48F* (*R. cellulolyticum*'s GH48 family cellulase), results in a toxic effect when overexpressed in *C. acetobutylicum*. For example, Mingardon *et al.* failed to transform *Cel48F* into *C. acetobutylicum* under control of the strong *thiolase* promoter. It could be transformed under control of the significantly weaker mutated *thiolase* promoter, although incubation of transformants at 30°C was necessary to achieve this (Mingardon *et al.* 2011). Additionally, it resulted in an unstable phenotype. Removal of the signal peptide resulted in the stable transformation of *Cel48F* even under control of the strong *thiolase* promoter (Mingardon *et al.* 2011). Additionally, in Chapter 3 of this study, an inhibition of growth

was observed when strains containing Cel48F under control of the *PtcdB* promoter were induced at a high level, again indicating a toxic effect.

A further study described the addition of scaffoldin modules (X2 domains) to Cel48F to promote the secretion of cel48F and proposed that these can act as carrier domains to promote the secretion of cellulases. The impact on the activity of Cel48F was not determined, although, they did show some activity towards amorphous cellulose (Chanal et al. 2011). Additionally, unpublished work in our lab showed that there was no difference in the expression level of Cel48F and chimeric Cel48F with these X2 domains.

This indicates that there is specifically, an issue in the secretion of GH48 family cellulases, in the next section, the overexpression of a chaperone protein, PrsA is explored.

4.3.1 Sec secretion system

Most proteins destined for secretion including cellulases are secreted through the sec secretion pathway (Harwood and Cranenburgh 2008). When synthesised, they are stabilised by intracellular chaperones (in Gram-positives GroES and DnaK). Secreted proteins contain a signal peptide which is recognised by a Signal recognition particle (SRP). In *B. subtilis* and other Gram-positive bacteria, this SRP complex consists of a scRNA (small cytosolic RNA), a GTPase called Ffh and a histone-like protein Hbsu. Hbsu is bound to scRNA, and Ffh targets the complex to SecA. Once recognised, proteins are translocated through the cytoplasmic membrane via the Sec translocase; this consists of SecA, secYEG and SecDF. SecA is an ATPase motor which binds the SRP and pore complex. SecYEG acts as the pore, while SecDF is involved in releasing the protein from the translocon (Wely et al. 2001). Homologs of all proteins are found in *C. acetobutylicum* and *R. cellulolyticum*.

4.3.2 Factors affecting secretion of protein

Many factors can affect the efficiency of secretion of proteins, and heterologous proteins can often misfold or fold slowly after secretion, this results in them being degraded. A high level of misfolded proteins can block the sec secretion system resulting in secretion stress; this can result in inhibition of cell growth and cell lysis (Hyryläinen et al. 2001). Improperly folded proteins are quickly degraded by proteases, in *B. subtilis* there are several, which are upregulated in response to secretion stress (Hyryläinen et al. 2001).

The *B. subtilis* secretion stress response is a two-component system called CssR-CssS it is homologous to the CpxR-CpxA system from *E. coli* (Darmon et al. 2002; Pogliano et al. 1997). Homologs of the CsxR-CsxS are present in both *C. acetobutylicum* and *R. cellulolyticum*. CsxS is a histidine kinase which is regulated by the CsxR, response regulator. CsxR-CsxS upregulates the expression of two serine proteases *HtrA* and *HtrB* in response to secretion stress (Darmon et al. 2002). *HtrA* and *HtrB* have been shown to act as protein chaperones but their main role is the degradation of misfolded extracellular proteins (Noone et al. 2001). The cell wall bound protease WtrA is also upregulated in response to secretion stress but is not regulated by the CsxR-CsxS system (Stephenson and Harwood 1998).

Proteins are transported across the cytoplasmic membrane in an unfolded, or loosely folded state. Immediately after translocation proteins fold into their native configuration at the membrane-cell wall interface, after which they diffuse through the cell wall. Several factors have been shown to influence the rate of folding, including; PrsA (a protein chaperone), the concentration of metal ions, the composition of the cytoplasmic membrane and other cell wall components. The signal peptide can also have an effect (S. Park and Schumann 2015).

PrsA is a lipoprotein which is essential in *B. subtilis*, deletion of PrsA results in gross morphological alterations and ultimately cell death (Jacobs et al. 1993). A knockdown of PrsA expression results in an increase in the level of degradation of proteins, as a higher proportion misfold (Kontinen and Sarvas 1993). In *B. subtilis* depletion of PrsA results in secretion stress when proteins are overexpressed, for example, amylase (Hyyryläinen et al. 2001). Several studies have demonstrated that overexpression of PrsA results in a significant increase in the secretion yields of proteins in several organisms including *B. subtilis* and *L. lactis* (J. Chen et al. 2015; Lindholm et al. 2006).

In summary, the toxic effect observed when Cel48F is overexpressed could be due to Cel48F misfolding, resulting in blockage of the sec secretion system which can result in cell death. The reason for Cel48F misfolding is most likely due to the low expression or absence of a protein chaperone. To date the only extracellular protein chaperone identified in gram-positive bacteria is PrsA.

It should also be noted that to our knowledge, no protein has been expressed in any organism resulting in a similar phenotype to that observed when Cel48F is overexpressed in high levels in *C. acetobutylicum*. For example, amylases expressed in *B. subtilis*, only have a similar phenotype of cessation of cell growth and cell lysis when induced at high levels in a background when a component of the sec secretion system had been deleted or knocked down. For example; the *PrsA3* mutation or deletion of the secretion stress regulator *cssR*, resulting in reduced expression of *htrA* or *htrB*, which degrade misfolded proteins (Antelmann et al. 2003). This implies that very high levels of Cel48F are misfolding in *C. acetobutylicum*.

4.3.3 Effect of expression of PrsA chaperones on Cel48F expression

PrsA genes from *C. acetobutylicum*, *B. subtilis* and the two from *R. cellulolyticum* were chosen to be expressed in *C. acetobutylicum*. All PrsA candidate genes share a low

sequence identity of around 30% (Figure 4.10). They all contain a rotamase domain which increase the rate of protein folding by catalysing the interconversion of cis-proline and trans-proline. Both *C. acetobutylicum*'s PrsA and *R. cellulolyticum*'s 3079 also contain a SurA N-terminal domain. This domain is found at the N-terminus of the chaperone SurA. It is a helical domain of unknown function.

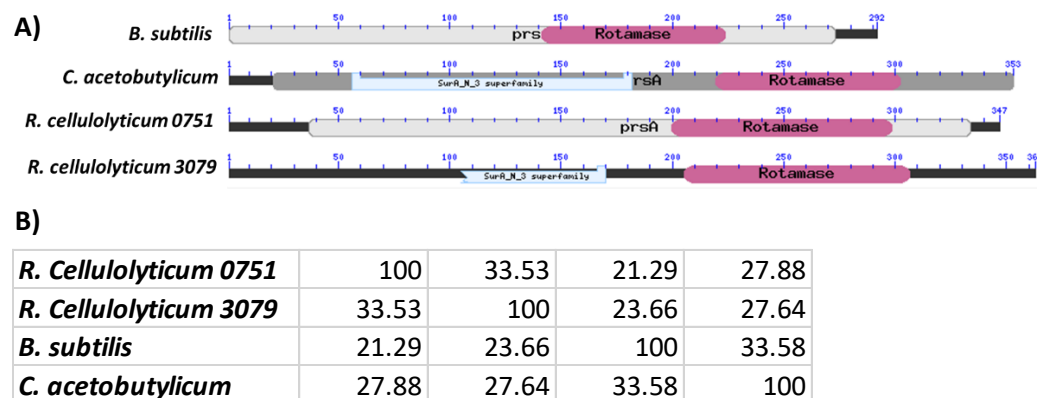


Figure 4.10. Comparison of PrsA proteins. A) Domain organisation of the PrsA proteins from *B. subtilis*, *C. acetobutylicum* and *R. cellulolyticum*'s (locus 0751 and 3079). B) Percent identity matrix of the PrsA proteins from *B. subtilis*, *C. acetobutylicum* and *R. cellulolyticum*'s (locus 0751 and 3079).

The constructs were cloned downstream of the *TcpF* promoter in the pMTL-82151 backbone and transformed into a strain expressing Cel48F integrated into the genome (thl:: T1T2-PtcdB-Cel48F(ND), pyrE:: LI-tcdR). Strains were grown as previously described, in CBM (40 mM MOPS, pH 6.8) to an OD of 1.0 and induced at OD₆₀₀ 0.3.

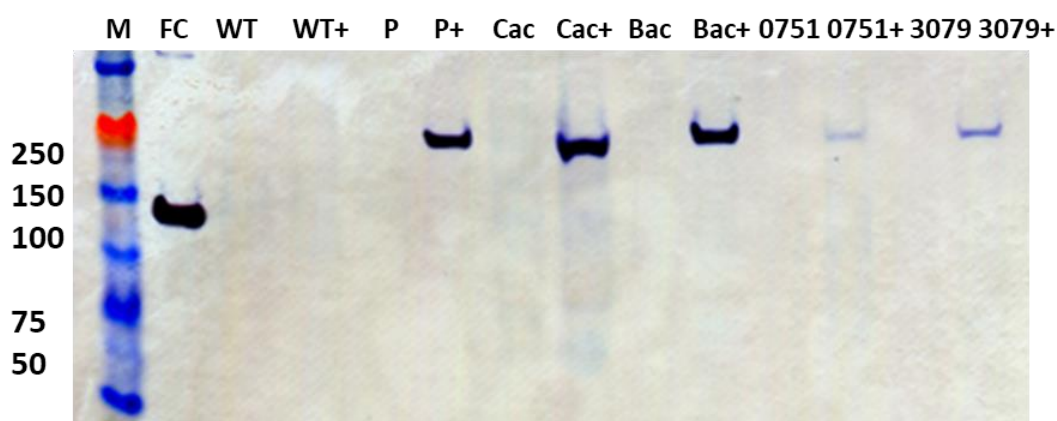


Figure 4.11. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains co-expressing *PrsA* chaperones with the *Cel48F* cellulase from *R. cellulolyticum*. *Cel48F* was under control of the *PtcdB* promoter in a *Li-tcdR* background and *PrsA* chaperones were under control of the *TcpF* promoter. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. *M*: Prestained Protein Ladder; *Flag*: FLAG-BAP positive control; *WT*: wild-type *C. acetobutylicum*; *P*: Strain expressing *Cel48F* with a plasmid control; *Cac*: Strain expressing *Cel48F* as well as overexpressing *C. acetobutylicum*'s native *PrsA* protein; *Bac*: Strain expressing *Cel48F* as well as *B. subtilis*'s *PrsA* protein; *0751*: Strain expressing *Cel48F* as well as the *PrsA* gene 0751 from *R. cellulolyticum*; *3079*: Strain expressing *Cel48F* as well as the *PrsA* gene 3079 from *R. cellulolyticum*; +: Induction with 1 mM lactose.

There was a slight increase in the level of *Cel48F* overexpressing *C. acetobutylicum*'s native *PrsA* gene, the same level of expression in the strain expressing *B. subtilis*'s *PrsA* gene and surprisingly a large decrease in the level of expression of *Cel48F* in both strains overexpressing *R. cellulolyticum*'s *PrsA* genes (Figure 4.11).

4.4 Discussion

Firstly, the low expression of GH48 family cellulases from several different organisms indicates that there is a general issue with the expression and secretion of GH48 family cellulases in *C. acetobutylicum*. For this reason, we focused our attention on potential issues related to the expression of GH48 family cellulases, specifically *Cel48F*.

It was hypothesised that the low expression of GH48 family enzymes may be due to mRNA instability. 5' and 3' stem-loops have been shown to be the main factors affecting mRNA stability (Phan, Nguyen, and Schumann 2013). Several 5' and 3' stem-loops were tested to increase mRNA stability. However, only small increases in expression were achieved, indicating that mRNA stability is not the main reason for low expression of GH48 family cellulases in this host.

Similarly, the low expression of GH48 family cellulases may be due to poor secretion, likely due to misfolding during secretion, resulting in blockage of the Sec-secretion pathway. This could explain the toxicity observed when Cel48F is overexpressed. However, the overexpression of PrsA chaperones, both *C. acetobutylicum*'s native PrsA protein and PrsA proteins from *B. subtilis* and *R. cellulolyticum* did not result in a significant increase in Cel48F expression. Although, it should be noted that the expression of the PrsA proteins were not confirmed, in future expression could be confirmed by performing western blot or proteomic analysis. Surprisingly, the expression of both PrsA chaperones from *R. cellulolyticum* resulted in a decrease in expression. An increase in the expression of proteins in a PrsA deficient *B. subtilis* strain has been reported previously, although it is strange that only the expression of PrsA proteins from *R. cellulolyticum* resulted in a decrease in expression (Vitikainen et al. 2005). For these reasons, the low expression of a protein chaperone is unlikely to be the reason for the low expression of GH48 family cellulases in *C. acetobutylicum*. Although, the expression of a specific chaperone for Cel48F secretion, not yet identified, cannot be ruled out.

Several other factors have also shown to have an impact on the secretion of proteins in *B. subtilis*. The concentration of metal ions (calcium, magnesium and iron) have been shown to have an impact on the rate that proteins fold as well as the stability of

proteins. An example of this is the model protein levansucrase, where the presence of iron ions has been shown to be essential for the folding of it in vitro (Chambert and Petit-Glatron 1988).

An α -amylase from *B. licheniformis* which was designed to be more calcium-dependent was also shown to be more susceptible to cell-associated proteolysis in *B. subtilis* than the native protein (Stephenson et al. 1998). GH48 family cellulases, including Cel48F, have been shown to contain calcium ions and magnesium ions (magnesium ions are present in the active site and calcium in the dockerin domain). For these reasons, the concentration of metal ions will likely affect the folding of them (Parsiegla et al. 2008).

The concentration of available metal ions is mainly dependent on the overall net negative charge of the cell wall. In *B. subtilis* teichoic acids have been shown to be d-alanylated, this is controlled by the *dlt* operon, inactivation of the *dlt* operon results in a more negative environment and a significant increase in the secretion of several proteins (Hyyrylainen et al. 2000). Analysis of *C. acetobutylicum* and *R. cellulolyticum* genomes show that neither have a functional *dlt* operon, this could imply that this is not the reason for the low expression of GH48 family cellulases in *C. acetobutylicum*.

A recent paper also demonstrated that changing the composition of lipids in the cytoplasmic membrane has an effect on the expression of amylases, with a more negative charge resulting in higher expression. Interestingly, the charge of the protein itself appeared to have an effect on expression, as amylases with a more negative pI were expressed at higher levels (Cao et al. 2017).

The predicted pI's of Cel48F and Cel5A are very similar, so the pI is unlikely to be the reason for the low expression of GH48 family cellulases. However, the charge of the membrane is unknown in both *C. acetobutylicum* and *R. cellulolyticum*, so the charge of the membrane could have an influence on Cel48F secretion.

Overall, it seems that none of the usual reasons for the low expression of proteins are responsible for the low expression of GH48 family cellulases in *C. acetobutylicum*. At this point, there are two possible explanations: (1) a yet unidentified protein chaperone is required for the secretion of Cel48F (2) or modification to *C. acetobutylicum*'s cell membrane or wall is necessary to facilitate secretion of Cel48F.

4.5 Key outcomes

- The low expression and secretion levels of all GH48 family cellulases tested in *C. acetobutylicum* indicates that there is a general problem with the expression and secretion of GH48 family cellulases.
- mRNA instability is unlikely to be the main reason for the low expression levels of GH48 family cellulases, as the addition of 5' and 3' stem loops did not significantly improve their expression.
- The overexpression of the only extracellular protein chaperone, PrsA did not result in significant increases in expression of GH48 family cellulases, indicating that low expression of the known protein chaperone is not the reason for the low expression of GH48 family cellulases.

Chapter 5 Does Cel48F interact with peptidoglycan during secretion?

As described in the previous chapters, we have shown that the expression and secretion of GH48 family cellulases in *C. acetobutylicum* is significantly lower than other cellulase families. The potential factors that might be responsible for the low expression levels of GH48 family cellulases were also explored. None of the commonly known factors, such as mRNA stability, promoter strength or the presence or absence of protein chaperones were found to be responsible for the low expression levels observed for GH48 family cellulases in *C. acetobutylicum*.

Another hypothesis that we set out to test was that some modification to *C. acetobutylicum*'s cell membrane or cell wall is necessary to enable high-level secretion of GH48 family cellulases. In this chapter, the possibility that GH48 family cellulases are toxic to *C. acetobutylicum* due to their ability to either bind to or degrade peptidoglycan, thus resulting in cell lysis is explored.

5.1 Introduction

5.1.1 Disruption of bacterial cell walls

The cell wall and membrane(s) in bacteria are essential, and disruption of either of them results in cell lysis and death; for this reason, they are a common target for antimicrobials. Antimicrobials which disrupt cell integrity can be split into two main categories: peptidoglycan degrading hydrolases and bacteriocins which normally form pores in the cytoplasmic membrane (Alcorlo et al. 2017).

There are two main categories of cell walls: Gram-negative and Gram-positive. Gram-negative cells have a thin layer of peptidoglycan in the periplasmic space between the inner and outer lipid membranes, with peptidoglycan attached to both membranes via lipoproteins. The outer membrane also contains lipopolysaccharides (Costerton, Ingram, and Cheng 1974). Gram-positive cells have a thick layer of peptidoglycan and lipoteichoic acids attached to a single lipid membrane by lipoproteins (Beeby et al. 2013).

5.1.2 Peptidoglycan

Peptidoglycan is composed of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc), linked by β -1-4 glycosidic bonds (Figure 5.1). A peptide stem is attached to every MurNAc residue at the D-lactyl group (Waldemar Vollmer, Blanot, and De Pedro 2008). The composition of the peptide stem varies, usually, it is composed of D-amino acids and other modified amino acids, for example, m-DAP (2-6-diaminopimelic acid). In gram-negatives such as *E. coli*, the peptide stem consists of L-Ala- γ -D-Glu-m-DAP-D-Ala-D-Ala while in most gram-positives it consists of L-Ala- γ -D-Glu-L-Lys-D-Ala-D-Ala. Individual peptidoglycan strands are crosslinked via the peptide stem (Waldemar Vollmer, Blanot, and De Pedro 2008).

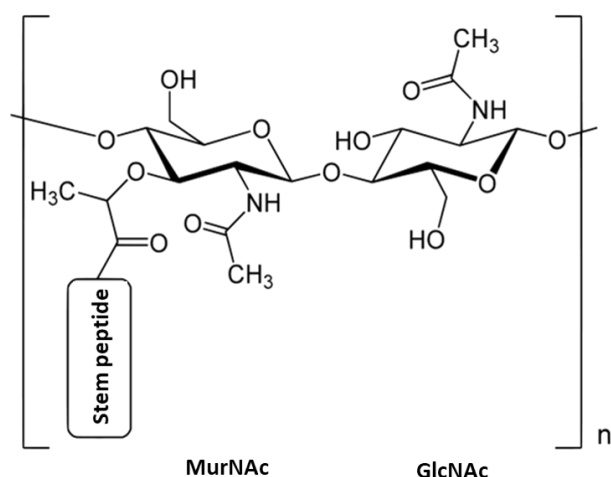


Figure 5.1. The basic structure of peptidoglycan. Peptidoglycan consists of MurNAc and GlcNAc residues linked by a β -1-4 glycosidic bonds. A stem peptide is attached to MurNAc at the D-lactyl group.

5.1.3 Peptidoglycan hydrolases

Peptidoglycan hydrolases can be split into two main groups: glycosidases and peptidases. Glycosidases cleave the β -1-4 glycosidic bond between GlcNAc and MurNAc; these can be split into the following subgroups: (1) glucosaminidases, which hydrolytically cleave the β -1-4 glycosidic bond, (2) lysozymes and (3) lytic transglycosylases, which cleave the β -1-4 glycosidic bond sequentially (Alcorlo et al. 2017). Peptidases can be split into amidases and peptidases; amidases cleave the peptide bond between MurNAc and the peptide stem, while peptidases cleave the peptide bonds within the peptide stem (Alcorlo et al. 2017).

Peptidoglycan hydrolases are essential for several normal cellular functions including: peptidoglycan maturation, peptidoglycan turnover, peptidoglycan recycling and autolysis. They are also required to cleave the septum during cell division (T. K. Lee and Huang 2013). However, they are also secreted by a large number of organisms for protection against invasion, the most common example being lysozyme which is secreted by most vertebrates (Bilej 2015).

5.1.4 Peptidoglycan modifications

Bacteria can modify their peptidoglycan to both provide protection from peptidoglycan hydrolases and to regulate the activity of their own hydrolases. The most common modifications are N-deacetylation and O-acetylation of both the MurNAc and GlcNAc sugar of the peptidoglycan backbone (Moynihan, Sychantha, and Clarke 2014). Other peptidoglycan modifications only provide an indirect protection for example, the 1,6-anhydro ring commonly found in gram-negative bacteria (Atrih et al. 1999; W. Vollmer and Höltje 2001). In gram-negative bacteria, 1-6-anhydroMurNAc residues have been found to induce the expression of β -lactamase. β -Lactamase breaks down β -lactam antibiotics which inhibit peptidoglycan synthesis (Höltje et al. 1994). Additionally, MurNAc can also be N-glycosylated, and during spore formation, Muramic δ -lactam is also formed (Popham 2002). However, none of these modification provide protection from hydrolases (Moynihan, Sychantha, and Clarke 2014).

In addition to the modifications of the peptidoglycan backbone, the peptide stem can also be modified. For example, in Vancomycin-Resistant Enterococci, D-alanine is replaced with D-lactate to provide resistance to vancomycin (Bugg et al. 1991). Additionally, the cell wall can be modified to prevent the binding of peptides to the cell wall, a common example of this type of modification is the D-alanylation of teichoic acids (Herbert et al. 2007).

N-deacetylation of peptidoglycan is generally found in gram-positive bacteria, although it has been observed in some gram-negatives. N-deacetylation was first observed in the lysozyme resistant organism, *Streptococcus pneumonia*, where the GlcNAc residue was found to be N-deacetylated. The *pgdA* gene was later found to be responsible for this modification, and homologs have since been found in several

organisms including: *Bacillus anthracis*, *Enterococcus faecalis*, *Helicobacter pylori* and *Listeria monocytogenes* (Waldemar Vollmer and Tomasz 2000; Moynihan, Sychantha, and Clarke 2014). Additionally, a similar gene has been found in *B. subtilis* and several other organisms that deacetylate MurNAc instead of GlcNAc. Both N-deacetylation of GlcNAc and MurNAc contribute to lysozyme resistance (Kobayashi et al. 2012).

O-acetylation of the peptidoglycan is found in both Gram-positive and Gram-negative bacteria. It normally occurs on MurNAc, but several organisms have also been reported to O-acetylate the GlcNAc residue. Only O-acetylation of MurNAc contributes to lysozyme resistance, while O-acetylation of GlcNAc is thought to modulate the control of native autolysins (Bernard et al. 2011).

In Gram-positive bacteria, the *oatA* gene has been identified as being responsible for O-acetylation of MurNAc (Bera et al. 2006), while a homolog named *oatB* has been found to be responsible for O-acetylation of GlcNAc in *L. plantarum* (Bernard et al. 2011). In gram-negative bacteria, a different set of enzymes have been found to be responsible, with *patA* responsible for O-acetylation of MurNAc and *patB* responsible for O-acetylation of GlcNAc (Weadge, Pfeffer, and Clarke 2005).

It is possible that GH48 family cellulases are able to bind or even cleave peptidoglycan, as the structure of cellulose and peptidoglycan is relatively similar. They both contain sugars linked by beta 1-4 linkages however, the sugars in cellulose and peptidoglycan are different. With cellulose consisting of glucose, while peptidoglycan has MurNAc and GlcNAc sugars. Several enzyme families have been shown to have cross-specificity, for example, some chitinases have been shown to be able to degrade peptidoglycan (Terwisscha van Scheltinga et al. 1995). Conversely, some lysozymes have also been shown to be able to degrade chitin but not cellulose. Some members of the GH48 family have been shown to be chitinases, although they were unable to degrade

cellulose, and their activity on peptidoglycan was not tested (Fujita et al. 2006). Additionally, both lysozymes and cellulases have been shown to bind to their substrate via tryptophan residues (Inoue et al. 1992; Parsiegla et al. 2008). Both lysozymes and cellulases also utilise the same catalytic mechanism and residues to hydrolyse the β -1-4 glycosidic bond (Malcolm et al. 1989; Vital de Oliveira 2014).

GH48 family cellulases could potentially cause cell lysis when they bind to peptidoglycan, even if they are unable to cleave it. This is supported by the fact that the antimicrobial activity of lysozyme has been shown to be independent of its ability to hydrolyse peptidoglycan (Ibrahim, Matsuzaki, and Aoki 2001). Although the exact mechanism by which the inactivated lysozyme causes cell lysis has not been determined, it is thought to induce autolysis by binding to peptidoglycan and activating the native autolysins. Another theory is that the lysozyme may impair membrane functions by binding to the cell surface (Ibrahim, Matsuzaki, and Aoki 2001; Herbert et al. 2007).

Mammalian peptidoglycan recognition proteins (PGRP's) are lethal to bacteria (Lu et al. 2006). The mode of action of PGRP's has recently been determined. They were found to bind to peptidoglycan and activate the misfolded protein response. For example, in *B. subtilis* they activate the CsxR-CsxS system which activates the misfolded protein response. The misfolded protein response results in cell death due to membrane depolarisation, inhibition of cellular processes (DNA, RNA, protein and peptidoglycan synthesis) and the production of hydroxyl radicals (for more information see Chapter 4). PGRP's activate the CsxR-CsxS response as a conformational change occurs following PGRP's binding peptidoglycan which results in hydrophobic regions being exposed, these are then recognised by the CsxS sensor protein (Kashyap et al. 2011).

There is strong evidence that *C. acetobutylicum* N-deacetylates its peptidoglycan, as one of its autolysins acts only on N-deacetylated peptidoglycan (C. Croux et al. 1992a). In addition, *C. acetobutylicum*'s genome contains several putative N-deacetylase genes. *C. acetobutylicum* has been isolated from several different mammal faecal samples, as well as the digestive tract of humans (Cato, E.P, W.L. George, and S.M. Finegold. 1986). N-deacetylation of its peptidoglycan would provide protection from lysozyme.

R. cellulolyticum's peptidoglycan is most likely N-acetylated, as although its genome does contain three putative N-deacetylases. Xylan deacetylases, which are structurally similar to peptidoglycan deacetylases (Taylor et al. 2006), have been identified in several cellulolytic organisms as deacetylation of xylan has been shown to make it less recalcitrant. This as xylan groups hinder access to endo-xylanases, removal of the xylan groups results in more accessible sites for xylanases to act (X. Chen et al. 2012). All three putative *R. cellulolyticum* peptidoglycan deacetylases are expressed at low levels and are upregulated when grown on xylose or xylan (Blouzard et al. 2010). This indicates that they are probably xylan deactylases, however, the possibility that *R. cellulolyticum* N-deacetylates its peptidoglycan cannot be discounted.

Additionally, the stem peptide is likely to be the same for *C. acetobutylicum* and *R. cellulolyticum*. *C. acetobutylicum* and most clostridia have been determined to have A1y type peptidoglycan (Schleifer and Kandler 1972). Although *R. cellulolyticum*'s peptidoglycan type has not been determined, analysis of the pathway using KEGG indicates it also has A1y type peptidoglycan.

It is possible that GH48 family cellulases can bind to or degrade N-deacetylated peptidoglycan but not N-acetylated peptidoglycan. This could potentially explain the low expression of GH48 family cellulases observed in *C. acetobutylicum*. GH48 family

cellulases could degrade peptidoglycan, which would result in higher levels of cell lysis. Alternatively, if GH48 cellulases could bind to peptidoglycan, this may initiate the CssR-CssS response, or result in an increase in the activity of autolysin both of which would result in cell lysis.

5.2 Results

5.2.1 Microscopy analysis

To determine if the expression of Cel48F results in the degradation of the cell wall, we investigated the protoplast formation rate of the Cel48F overexpressing strains compared to the wild-type strain. Protoplasts are cells without a cell wall, they can be produced by treating cells with lysozyme in media with a high sugar content to prevent osmotic shock and cell lysis (Allcock et al. 1982).

Initial experiments looked at the rate of protoplast formation, both in the wild-type and a strain expressing Cel48F under control of the lactose inducible promoter ($\Delta LacG$ *pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_N_Flag*), referred to as Li-48F for the rest of this study. An inducible system was utilised as it was hypothesised that it would be easier to determine differences between induced and uninduced samples, then between different strains.

Lysozyme was added to a final concentration of 100 $\mu\text{g}/\text{ml}$ when indicated as a control. The Li-48F strain was induced with either 0 or 100 mM of lactose. After 1 hour of incubation, pictures of the samples were taken using a bright-field microscope.

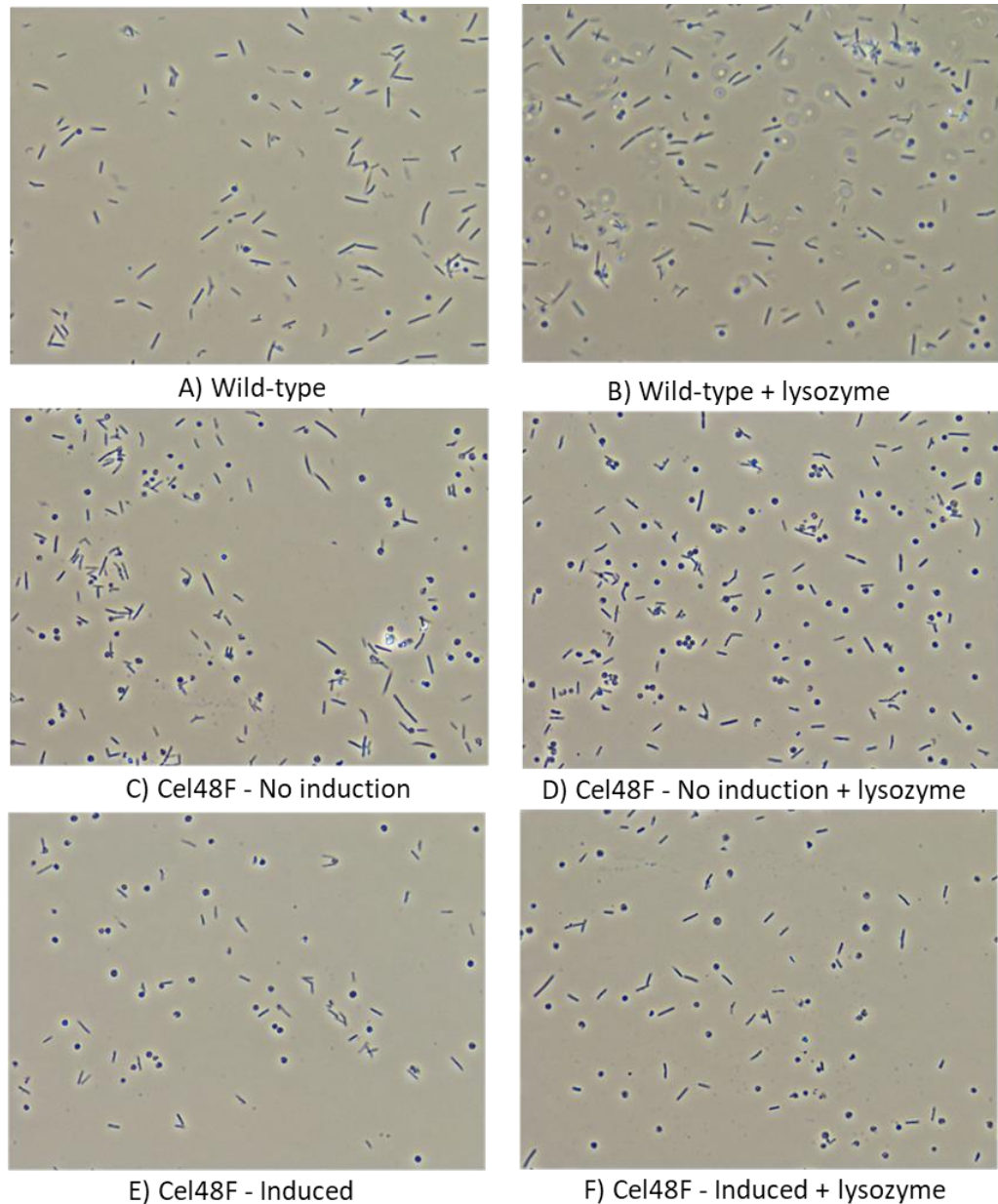


Figure 5.2. Microscope images showing representative images of *C. acetobutylicum* strains in protoplast buffer. Cells were grown up to OD_{600} 1.0 in CBM then re-suspended in protoplast buffer to an OD_{600} of 10.0. A) Wild-type B) Wild-type, incubated with a final concentration of 100 $\mu\text{g/ml}$ Lysozyme, C) Strain expressing Cel48F under control of the PtcdB promoter in the Li-tcdR background, D) Strain expressing Cel48F under control of the PtcdB promoter in the Li-tcdR background incubated with a final concentration of 100 $\mu\text{g/ml}$ lysozyme, E) Strain expressing Cel48F under control of the PtcdB promoter in the Li-tcdR background induced with a final concentration of 100 mM lactose, F) Strain expressing Cel48F under control of the PtcdB promoter in the Li-tcdR background incubated with a final concentration of 100 $\mu\text{g/ml}$ lysozyme and induced with a final concentration of 100 mM lactose. The images were taken after 1-hour incubation.

Protoplasts and intact cells can be easily distinguished as intact cells are seen as bright phase rods, while protoplasts are bright phase circles. In all samples, protoplasts were formed after 1-hour incubation and remained stable.

In the wild-type samples with no added lysozyme, there was a small proportion of cells which formed protoplasts, probably caused by the activity of *C. acetobutylicum*'s native autolysins (Figure 5.2, picture A). The addition of lysozyme to wild-type cells resulted in a high level of protoplast formation (Figure 5.2, picture B). The protoplast formation was not 100% this was probably due to the relatively low level of lysozyme used.

Samples containing the Li-48F strain, in both the induced and non-induced samples, a similar proportion of protoplasts were formed (Figure 5.2, pictures C and E). The level of protoplast formation was relatively high, and similar to the number of protoplasts observed in the wild-type plus lysozyme samples. The addition of lysozyme did not significantly increase the number of protoplasts formed (Figure 5.2, pictures D and F).

Surprisingly, there was no difference in protoplast formation between the induced and non-induced samples. A low level of expression of Cel48F is observed in non-induced samples (see Figure 3.4), so it is possible that this low level of expression is sufficient to cause degradation of the cell wall. As protoplasts seemed stable in both the wild-type and the strain expressing Cel48F, it is unlikely that Cel48F can disrupt the cell membrane.

Due to low levels of protoplast formation in the wild-type and no difference between the induced and non-induced samples, we then tested the influence of concentrated supernatants derived from the wild-type strain or a strain expressing Cel48F on wild-type cells. The strain expressing Cel48F constitutively from the thiolase promoter was

used (*pyrE*:: Facoid_Cel48F_N strain (F_48)), as it has been shown to result in a higher expression level of Cel48F (see Figure 3.5).

The experiment was repeated as before, this time adding concentrated supernatants from either the wild-type or the Cel48F expressing strain, to wild-type cells. To quantify the level of protoplast formation, ten pictures were randomly taken from each sample and the percentage of protoplasts determined.

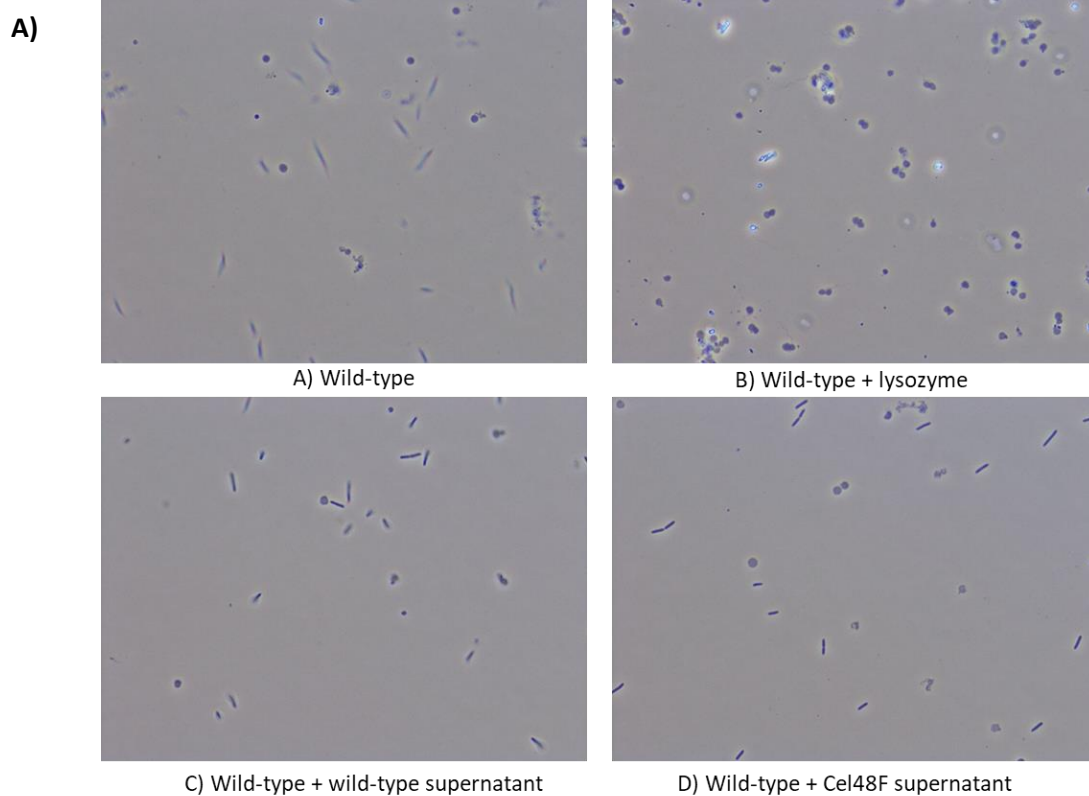
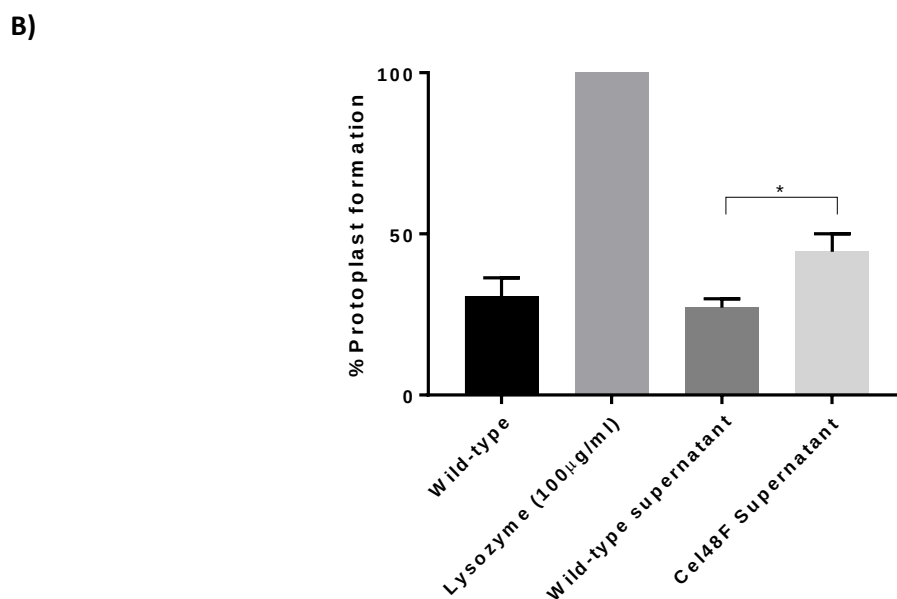


Figure 5.3. Effect of supernatants from *C. acetobutylicum* strains on protoplast formation.

A) Microscope images of wild-type *C. acetobutylicum*. Cells were grown to OD_{600} 1.0 then resuspended to OD_{600} 10.0 in protoplast media and incubated for 1 hour with either a final concentration of Lysozyme (100 ug/ml) or 10x concentrated S/N from wild-type or strain expressing Cel48F under control of the *facOID* promoter. A) wild-type, B) Wild-type plus lysozyme, C) Wild-type plus wild-type supernatant, D) Wild-type plus Cel48F supernatant. **B)** The percent of protoplasts formed after 1-hour incubation. 12 pictures were taken randomly, and the number of protoplasts and intact cells counted (See Appendix). Mean values $\pm$ S.E.M. ($n=10$).

As expected, a small number of protoplasts were formed in the wild-type sample without lysozyme (Figure 5.3A, picture A), while most cells in the samples where lysozyme was added formed protoplasts (Figure 5.3A, picture B). The addition of wild-type supernatant to wild-type cells did not increase the formation of protoplasts (Figure 5.3A, picture C), however significantly higher levels of protoplasts formation was observed when supernatant from a strain expressing Cel48F was added (45.5% compared to 29%), ($p < 0.05$) (Figure 5.3A, picture D).

The increase in protoplast formation when supernatants from a strain expressing Cel48F is added is indicative that Cel48F might be able to degrade *C. acetobutylicum*'s cell wall.

5.2.2 Cel48F activity on autoclaved *C. acetobutylicum* cells

To assess the ability of Cel48F to degrade *C. acetobutylicum*'s cell wall, an assay commonly used to determine autolysin activity was used (Meyrand et al. 2007). In this assay, autoclaved cells are used and a decrease in the absorbance levels indicates enzyme activity.

To perform this assay, cells were grown to OD₆₀₀ 0.6, autoclaved and re-suspended in 100 mM Tris-HCL (pH 5.5) to the same starting OD₆₀₀. Supernatants were concentrated 100-fold from either wild-type *C. acetobutylicum* or strain F.48 expressing Cel48F. 30 µl of concentrated supernatant were then added to 270 µl of autoclaved cells. The OD₆₀₀ was measured every 10 minutes and presented as a percentage of initial OD. This experiment was repeated three separate times, with the same trend repeatedly observed; a representative experiment is shown below.

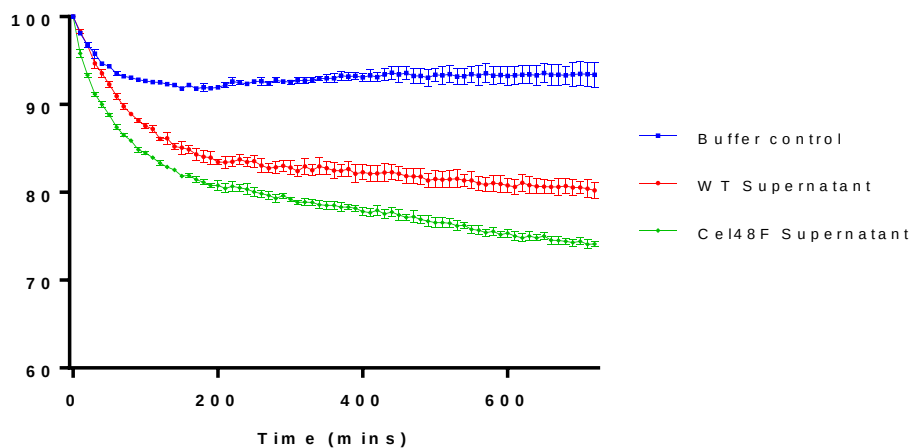


Figure 5.4. Activity of concentrated *C. acetobutylicum* supernatants towards autoclaved *C. acetobutylicum*. The graph shows the percent of starting OD over time of autoclaved *C. acetobutylicum* cells after the addition of concentrated supernatants. Blue: Buffer control (100 mM Tris-HCL pH 5.5); Red: Supernatant from wild-type *C. acetobutylicum*; Green: Supernatant from a *C. acetobutylicum* strain expressing Cel48F under control of the *facOID* promoter. Mean values $\pm$ S.E.M. ($n=3$).

As expected, there was a large decrease in OD when the wild-type supernatant was added due to the presence of *C. acetobutylicum*'s native autolysins. However, in the strain expressing Cel48F, there was a statistically significantly further decrease in OD (400 minutes = $P<0.0001$). This again indicates that Cel48F is able to degrade the cell wall of *C. acetobutylicum* cells, most likely by degrading its peptidoglycan. Activity of Cel48F on purified *C. acetobutylicum* peptidoglycan

A peptidoglycan turbidity assay was performed to determine the ability of secreted Cel48F to specifically degrade *C. acetobutylicum*'s peptidoglycan. Peptidoglycan was purified from *C. acetobutylicum*, and when appropriate, chemically N-acetylated. Supernatants were concentrated 100-fold from *C. acetobutylicum* strains as before, then added to a final concentration of 10-fold. A buffer control, Mutanolysin (a N-acetylmuramidase) and hen egg white lysozyme were used as a control. Mutanolysin is not affected by the acetylation state of peptidoglycan, while lysozyme is inhibited by N-deacetylated peptidoglycan.

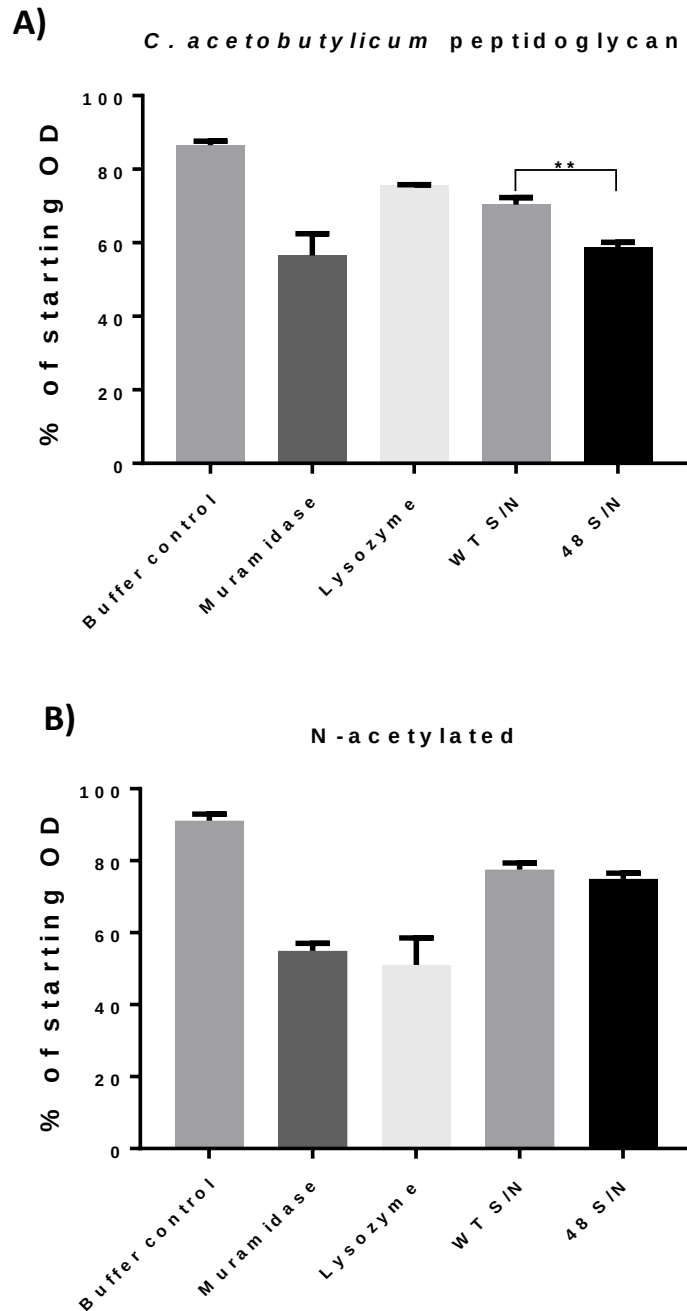


Figure 5.5. Activity of concentrated *C. acetobutylicum* supernatant towards purified *C. acetobutylicum* peptidoglycan. **A)** The activity of *C. acetobutylicum* supernatants against unmodified purified *C. acetobutylicum* peptidoglycan. **B)** The activity of concentrated *C. acetobutylicum* supernatants against chemically N-acetylated *C. acetobutylicum* peptidoglycan. Buffer control: 100 mM Bis-Tris HCL pH 5.5; Muramidase: Final concentration of 30 units/ml of Muramidase; Lysozyme: Final concentration of 100ug/ml of lysozyme; WT S/N: Concentrated supernatant from wild-type *C. acetobutylicum*; 48 S/N: Concentrated supernatants from a *C. acetobutylicum* strain expressing Cel48F under control of the *facOID* promoter. The graphs show the percent of starting OD after 3 hours incubation. Mean values $\pm$ S.E.M. ($n=3$).

For wild-type *C. acetobutylicum* peptidoglycan, as expected, a large decrease in OD was observed for both the muramidase and lysozyme. The OD600 only decreased to 50% this is consistent with other studies, this is most likely as purified peptidoglycan is insoluble. A sonication step to solubilise the peptidoglycan was crucial in for efficient degradation of peptidoglycan. A large decrease in OD was observed for wild-type supernatant (70%), as expected, due to the activity of native autolysins. For supernatant from a strain expressing Cel48F the OD600 decreased even further when compared to supernatant from the wild-type strain (60% compared to 70%) (Figure 5.5).

For chemically N-acetylated peptidoglycan, the activity of muramidase remains the same (Figure 5.5B), as this enzyme has been shown to be equally active on both N-deacetylated and N-acetylated peptidoglycan (Laaberki et al. 2011). Lysozyme was more active on N-acetylated peptidoglycan (Figure 5.5B), this was expected as lysozyme has been shown to be more active on N-acetylated peptidoglycan. Interestingly, the decrease in OD was the same for both supernatants, from both the wild-type strain and from the strain expressing Cel48F (Figure 5.5B). This could potentially mean that Cel48F is active on peptidoglycan, but is less active on N-acetylated peptidoglycan. In these preliminary experiments, concentrated supernatants were utilised instead of purified Cel48F enzyme. To confirm that Cel48F enzyme activity was responsible for cell wall lysis, the experiment was repeated using purified Cel48F.

5.2.3 Purification of Cel48F

Cel48F has been previously overexpressed and purified from *E. coli* cells, by expressing it intracellularly under control of the IPTG inducible T7 promoter (C. Reverbel-Leroy et al. 1997). To purify Cel48F, the published protocol was followed, plasmid pET-48Fc was

transformed into, BL21 (DE3) cells and grown to OD₆₀₀ 1.8-2.0 in LB (1.2% glycerol), before being induced with 40nM of IPTG and incubating for 16 hours at 16°C. After purification using His-tag columns, the samples were run on SDS-PAGE gels and stained with Coomassie blue to determine purity (Figure 5.7). Interestingly, Cel48F did not appear to be toxic to *E. coli* when expressed intracellularly. This is the same as in *C. acetobutylicum*, where it is only when Cel48F is secreted that it is toxic.

The elute fractions were pooled together and the concentration of Cel48F determined by BCA protein assay. The activity of Cel48F was also confirmed by carrying out an activity assay on PASC (Figure 5.8).

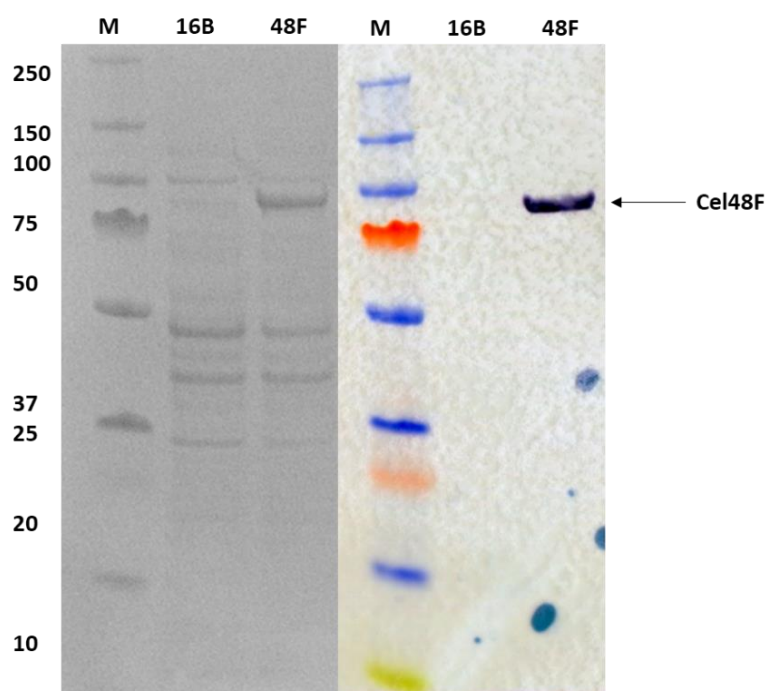


Figure 5.6. SDS PAGE and Western blot analysis of the soluble fraction of *E. coli* cells expressing Cel48F. The strains were induced with 40 nM IPTG and prepared as described in the materials and methods section. SDS/PAGE gels were stained with Coomassie Blue and Western blot was carried out using monoclonal ANTI-HIS antibody. M: Prestained Protein Ladder; 16B: Vector control strain - BL21(DE3)-pET16B; 48F: Strain expressing Cel48F- BL21(DE3) -pET48Fc.

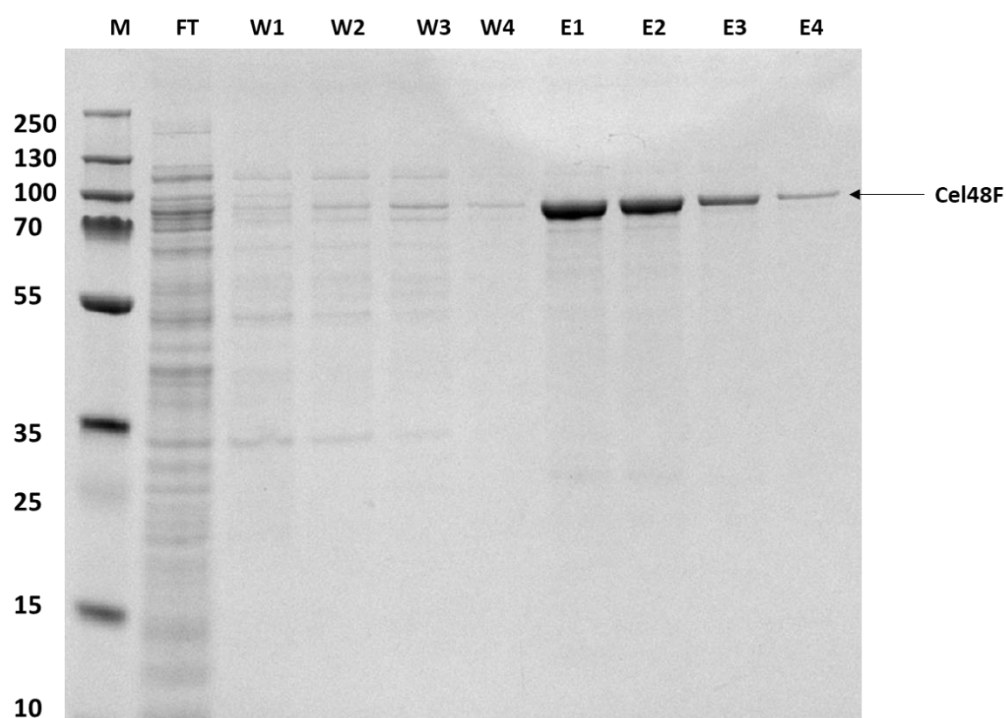


Figure 5.7. SDS PAGE showing His-tag purification of Cel48F. Cel48F was expressed under control of the T7 promoter with plasmid pET48Fc in BL21 cells and was induced with 40 nM IPTG. Cel48F was purified using Nickel His-tag purification columns as described in the materials and methods section. SDS/PAGE gels were stained with Coomassie Blue M: Prestained Protein Ladder; FT: Flow-through; W1-4: Wash steps 1-4 and E1-4: Elute fractions 1-4.

To confirm activity of Cel48F, purified Cel48F (final concentration 50 µg/ml) was incubated with 1% PASC, for 72 hours at 37°C, and releasing sugars were determined using a DNS assay (Figure 5.8). The low activity of Cel48F on PASC is similar to published values (Corinne Reverbel-Leroy et al. 1996)

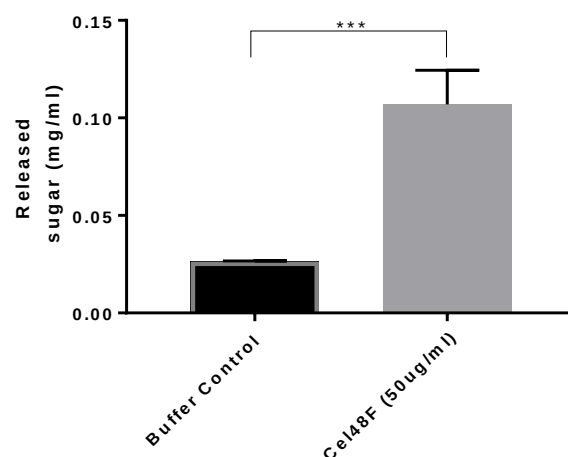
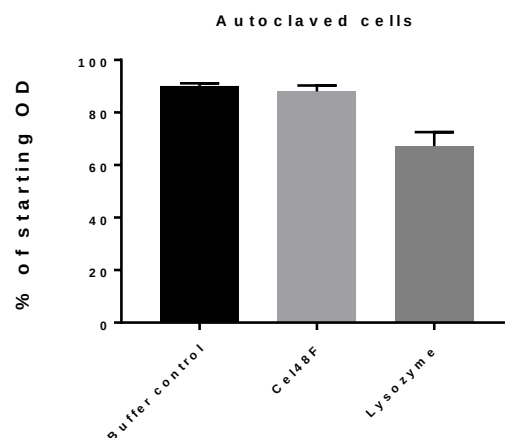


Figure 5.8. Activity of purified Cel48F on PASC. 50 µg/ml of purified Cel48F was incubated with PASC for 72 hours at 37°C and the levels of reducing sugars released were determined using the sugar release assay as described in the General Materials and Methods section. Buffer Control: 100 mM Tris-HCL pH 5.5; Cel48F: 50 µg/ml. Mean values $\pm$ S.E.M. (n=3). $P > 0.001$ between Buffer control and Cel48F.

5.2.4 Purified Cel48F is not active on autoclaved *C. acetobutylicum* cells or purified peptidoglycan

To confirm that Cel48F shows activity against *C. acetobutylicum* autoclaved cells and purified wild-type *C. acetobutylicum* peptidoglycan, the respective assays were repeated as before, using a final concentration of 100 µg/ml of Cel48F and 100 µg/ml of lysozyme as a control.

A)



B)

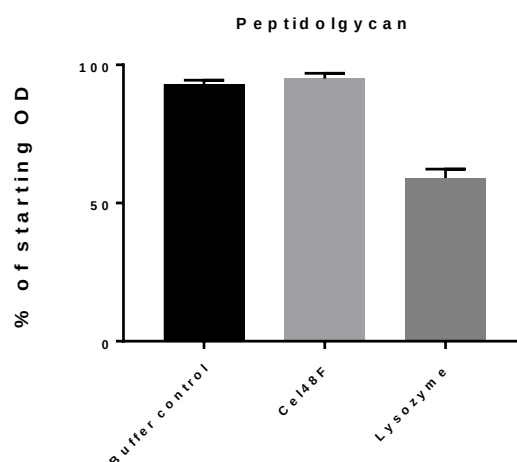


Figure 5.9. Activity of purified Cel48F on autoclaved *C. acetobutylicum* cells and purified peptidoglycan. A) The activity of purified Cel48F on autoclaved *C. acetobutylicum* cells. **B)** The activity of purified Cel48F *C. acetobutylicum* peptidoglycan. Buffer control: 100 mM Bis-Tris HCL pH 5.5; Cel48F: Final concentration of 100 $\mu\text{g/ml}$ of purified Cel48F; Lysozyme: Final concentration of 100 $\mu\text{g/ml}$ of lysozyme. The graphs show the percent of starting OD after 3 hours incubation. Mean values $\pm$ S.E.M. ($n=3$).

No decrease in OD compared to the buffer control was observed using purified Cel48F with either autoclaved cells or peptidoglycan as a substrate. The addition of lysozyme (100 $\mu\text{g/ml}$) resulted in a large decrease in OD, of both autoclaved cells and *C. acetobutylicum* peptidoglycan as expected (Figure 5.9). This shows that Cel48F purified from *E. coli* cells is not able to degrade peptidoglycan. This was somewhat unexpected, as the concentrated supernatants from the *C. acetobutylicum* strain expressing Cel48F

had significantly higher activity on autoclaved cells and purified peptidoglycan when compared to wild-type supernatant.

There are several explanations for this. One explanation could be that the intracellularly expressed Cel48F in *E. coli* has different folding and/or binding properties compared to the secreted, mature Cel48F from *C. acetobutylicum*. Another theory for the increased activity of the secreted Cel48F on the levels of cell lysis, might be explained by the presence of autolysins in the supernatant. Although unclear, this cell lysis may be caused by either by Cel48F directly binding to peptidoglycan, or recruiting the autolysin, or the misfolded Cel48F triggering a cell response, resulting in increased levels of cell lysis.

5.2.5 Can Cel48F bind to peptidoglycan?

To determine if Cel48F can bind to peptidoglycan a binding assay was carried out, based on a published assay by Atilano et al. 2011. Briefly, Cel48F purified from *E. coli* was incubated with peptidoglycan purified from *C. acetobutylicum* (N-deacetylated) or chemically N-acetylated peptidoglycan for 30 minutes at 37°C. The sample was centrifuged to pellet the peptidoglycan and any bound protein, before being washed three times. Samples representing the peptidoglycan pellet and the supernatant from each of the wash steps were loaded on an SDS-PAGE gel and western blot carried out.

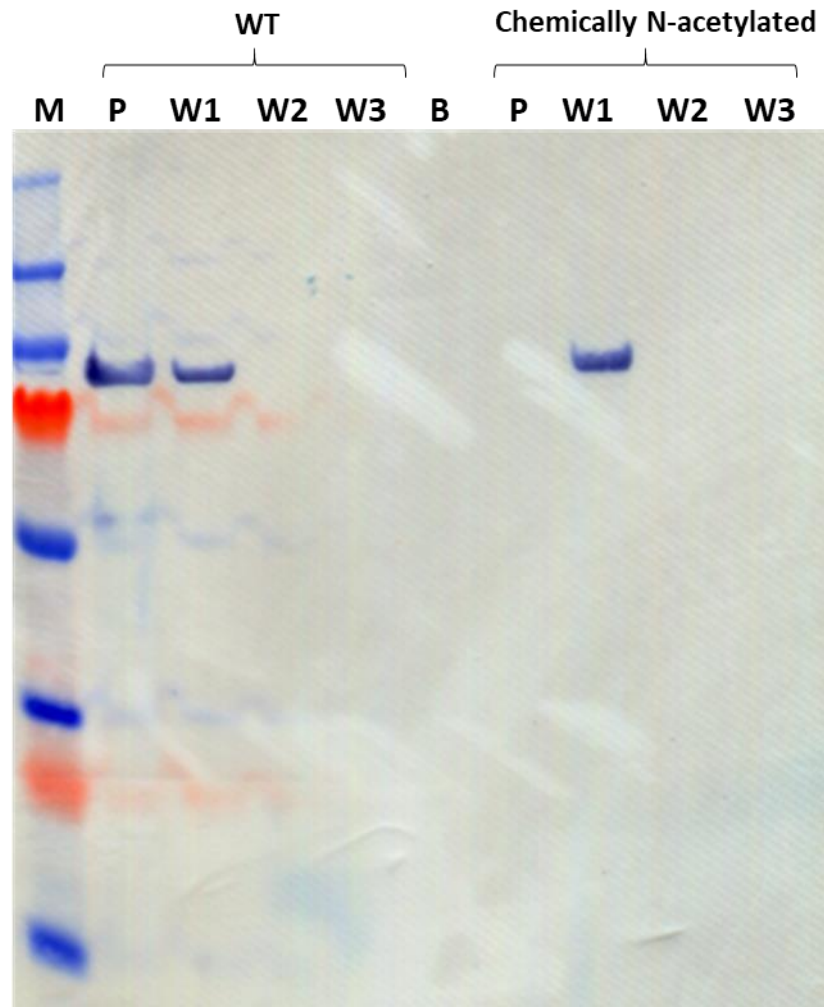


Figure 5.10. Cel48F and peptidoglycan co-precipitation assay. Purified Cel48F was incubated with peptidoglycan for 30 minutes. Cel48F bound to the peptidoglycan was found in the pellet fraction, while unbound Cel48F was found in the supernatant fraction. Samples were loaded on an SDS-PAGE and western blot was carried out using monoclonal ANTI-HIS antibody. **M**: Prestained Protein Ladder; **P**: Pellet fraction containing peptidoglycan and associated proteins; **W1-3**: Sample of supernatant from each wash step; **B**: Blank lane; **WT**: wild-type *C. acetobutylicum* peptidoglycan; **Chemically N-acetylated**: Chemically N-acetylated *C. acetobutylicum* peptidoglycan.

The assay confirmed that Cel48F can bind to wild-type *C. acetobutylicum* peptidoglycan, as the majority of Cel48F is in the pellet fraction. A smaller amount is located in the first wash, indicating either that the binding efficiency is not 100% or more likely the peptidoglycan was saturated, as no Cel48F can be detected in the second and third washes. Acetylation of peptidoglycan resulted in no binding of Cel48F to peptidoglycan, with Cel48F only being detected in the first wash fraction (Figure 10).

5.2.6 Cel48F is not toxic to growing cultures

Purified Cel48F, lysozyme or a buffer control were added to exponentially growing cultures. For *E. coli* and *B. subtilis* the OD₆₀₀ was measured using a plate reader while for *C. acetobutylicum* biomass was measured using a Biolector Microfermentation system.

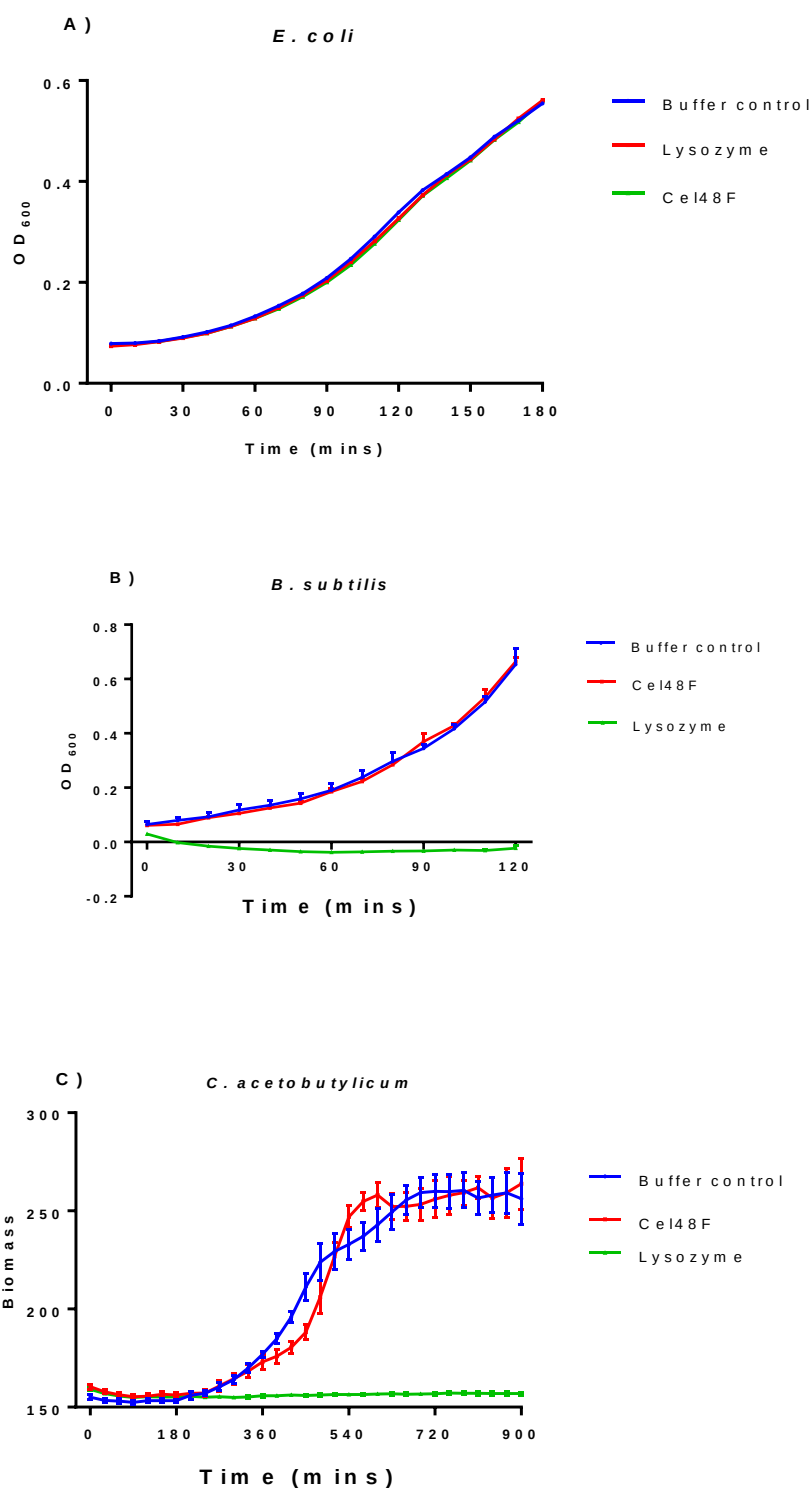


Figure 5.11. Toxicity assay of el48F against growing cultures of *E. coli*, *B. subtilis* and *C. acetobutylicum*.

A) Growth curve of *E. coli* after addition of Cel48F **B)** Growth curve of *B. subtilis* after addition of Cel48F **C)** Growth curve of *C. acetobutylicum* after the addition of Cel48F. *E. coli* and *B. subtilis* were grown in LB media and *C. acetobutylicum* was grown in CBM media (with no calcium carbonate added). Buffer control: 100 mM Tris-HCL pH 5.5; Cel48F: Purified Cel48F at a final concentration of 100 µg/ml; Lysozyme: Lysozyme at a final concentration of 100 µg/ml.

The addition of purified Cel48F to growing cultures of *E. coli*, *B. subtilis* and *C. acetobutylicum* did not result in any inhibition of growth (Figure 5.11). The concentration of Cel48F secreted by *C. acetobutylicum* in this and other studies have been reported to be between 0.1 to 1.5 µg/ml. This shows that the addition of Cel48F is not toxic to growing cultures even at significantly higher concentrations than would normally be encountered.

5.3 Discussion

The addition of supernatant from a *C. acetobutylicum* strain expressing Cel48F resulted in a significantly higher level of protoplast formation, activity on autoclaved cells and purified peptidoglycan when compared to the addition of supernatant from the wild-type strain. However, the addition of purified Cel48F did not result in any activity on either autoclaved cells or peptidoglycan.

This raises the question of why the addition of supernatant from a strain expressing Cel48F resulted in a higher proportion of protoplast formation and activity on autoclaved cells and peptidoglycan. The most likely explanation is an increase in the concentration of autolysins in the supernatant. Several autolysins have been isolated from *C. acetobutylicum*, with several having muramidase activity (Allcock et al. 1981; C. Croux et al. 1992b; Zhen Liu et al. 2015). This suggests that the expression of Cel48F, results in an increase in autolysin concentration in the supernatant, possibly due to an increase in the rate of cell lysis. Cell lysis would increase the concentration of autolysin in the supernatant as the majority of autolysins are present in the cytoplasm (C. Croux et al. 1992b). Additionally, cell lysis also results in an increase in the activity of a metalloprotease, potentially explaining the high levels of degradation of Cel48F (C. Croux et al. 1992b). How the expression of Cel48F can result in an increase in the rate of cell lysis, remains unclear.

A reduction in growth when high levels of Cel48F are produced was previously thought to be due to high levels of Cel48F misfolding, resulting in activation of the protein misfolding response (CssR-CssS). Activation of the protein misfolding response has been shown to result in cell lysis in *B. subtilis* (Hyryläinen et al. 2001) however, increased expression of the extracellular protein chaperone PrsA had no effect on Cel48F expression.

Proteins binding to peptidoglycan have been shown to be toxic as they either result in an increase in the activity of autolysins or activate the CssR-CssS misfolding protein system. (Ibrahim, Matsuzaki, and Aoki 2001; Kashyap et al. 2011). Cel48F was shown to bind to wild-type *C. acetobutylicum* peptidoglycan which is N-deacetylated, but not chemically N-acetylated peptidoglycan. *C. acetobutylicum*'s peptidoglycan is widely believed to be N-deacetylated, while *R. cellulolyticum*'s is likely to be N-acetylated.

For these reasons, the most likely explanation for the low expression and secretion of Cel48F, is that Cel48F binding to peptidoglycan during secretion, resulting either in activation of the CssR-CssS protein misfolding response or triggering autolysis.

Additionally, the addition of purified Cel48F to growing cultures of *E. coli*, *B. subtilis* or *C. acetobutylicum* did not appear to have any inhibitory effect on growth. This was unexpected as Cel48F would be expected to be able to bind to peptidoglycan, and trigger cell lysis. It is possible that the intracellularly expressed Cel48F by *E. coli* has somewhat different properties to the secreted Cel48F. Another possible explanation is the observation that PGRP's have been found to specifically enter the cell through cell separation sites and bind to the cell wall in the vicinity of the cell membrane. Protein-bound elsewhere significantly reduced the CssR-CssS misfolded protein response (Kashyap et al. 2011). Cel48F may not be able to bind to peptidoglycan near the cell membrane when added to the media. However, it would be able to bind to

peptidoglycan at the cell membrane while being secreted. It is unclear how this hypothesis could be tested, as upregulation of the CssR-CssS could be due to either Cel48F binding to peptidoglycan, or high concentrations of Cel48F misfolding.

Modifying *C. acetobutylicum*'s peptidoglycan to be N-acetylated may result in Cel48F being unable to bind to peptidoglycan during secretion, thus resulting in higher levels of Cel48F expression. This was investigated further in Chapter 6.

5.4 Key points

- The expression of Cel48F in *C. acetobutylicum* results in a supernatant which shows increased activity against *C. acetobutylicum* cells and purified peptidoglycan. This is possibly caused by an increase in the level of autolysins in the supernatant due to the expression of Cel48F.
- Purified Cel48F from *E. coli* was shown to be unable to degrade purified peptidoglycan.
- Purified Cel48F from *E. coli* was shown to be able to bind to N-deacetylated but not N-acetylated peptidoglycan, such as *C. acetobutylicum*'s.
- Cel48F binding to peptidoglycan near the cell membrane during secretion is therefore the most likely reason for the increased cell lysis observed in cells expressing Cel48F. This would also explain the low expression of Cel48F and other GH48 family cellulases in *C. acetobutylicum*.

Chapter 6 Modification of *C. acetobutylicum*'s peptidoglycan to enable expression of Cel48F.

In the last chapter, the possibility that GH48 cellulases could bind to or even degrade peptidoglycan was discussed. We found that GH48 family cellulases can bind specifically to N-deacetylated but not N-acetylated peptidoglycan. Binding of proteins to peptidoglycan can result in cell death, due to activation of the misfolded protein response (Kashyap et al. 2011). It is possible that *C. acetobutylicum*'s peptidoglycan is at least partially N-deacetylated, while *R. cellulolyticum*'s is fully or partially N-acetylated giving resistance to its hydrolases. We hypothesised that N-acetylation of *C. acetobutylicum*'s peptidoglycan could result in increased resistance to GH48 family cellulases.

Since the acetylation state of the *C. acetobutylicum*'s peptidoglycan is still unknown, in this chapter, we focused on analysis and characterisation of the acetylation state of both GlcNAc and MurNAc residues in wild-type *C. acetobutylicum* and in putative N-deacetylase deletion mutants. In addition, *cel48F* was overexpressed in these deletion strains and the level of secreted GH48 family cellulase compared to that in the wild-type strain.

6.1 Introduction

In chapter 5, we found that GH48 family cellulases can bind specifically to N-deacetylated but not N-acetylated peptidoglycan. Binding of proteins to peptidoglycan can result in cell death, due to activation of the misfolded protein response (Kashyap et al. 2011). If *C. acetobutylicum*'s peptidoglycan is N-deacetylated, while *R. cellulolyticum*'s is N-acetylated, deletion of *C. acetobutylicum*'s putative N-deacetylase gene(s) to create a mutant with N-acetylated peptidoglycan, may result in increased expression of GH48 family cellulases.

Peptidoglycan is synthesised as a disaccharide with a pentapeptide stem and transported to the cell membrane, where they are integrated into the existing peptidoglycan forming long glycans. These long glycan strands can form cross-links via the peptide stem. In organisms with an mDAP this is achieved either by the cross-linking of two mDAP residues or by the cross-linking a mDAP and D-alanine residue (Waldemar Vollmer, Blanot, and De Pedro 2008). Peptidoglycan deacetylases and O-acetylases are responsible for the chemical modifications of the peptidoglycan. These are usually attached to the cell membrane.

C. acetobutylicum has been predicted to have a type A1 γ cell wall, with the pentapeptide stem consisting of L-alanine linked to Muramic acid, followed by glutamic acid, m-DAP and two D-alanine's (Figure 6.1) (Schleifer and Kandler 1972). It is also predicted to be N-deacetylated, as one of its autolysins acts only on N-deacetylated peptidoglycan (C. Croux et al. 1992a). In addition, *C. acetobutylicum*'s genome contains several putative N-deacetylase genes.

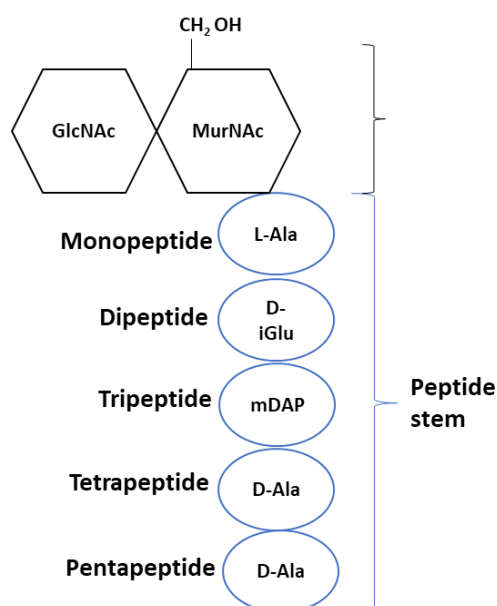


Figure 6.1. The predicted stem structure of *C. acetobutylicum*'s peptidoglycan.

The chemical composition of peptidoglycan was first determined using HPLC (Markiewicz, Glauner, and Schwarz 1983), although recently MALDI-TOF-MS and LC-MS/MS analysis has been preferred as it is a faster and more specific method (Meyrand et al. 2007; Bern, Beniston, and Mesnage 2017; Kuhner et al. 2014).

6.2 Results

6.2.1 Analysis of wild-type *C. acetobutylicum* peptidoglycan

To characterise the composition of *C. acetobutylicum*'s peptidoglycan, peptidoglycan was purified from *C. acetobutylicum*, digested by Mutanolysin (an N-acetylmuramidase) and subjected to reverse-phase high-pressure liquid chromatography (RP-LC-MS) using an electrospray ion source and both Q-TOF-MS and Orbitrap MS-MS analyses.

Analysis of the Q-TOF MS data showed that the predicted composition of the peptide stem is correct, with the exception that Glutamic acid appears to be replaced with

glutamine, as there is a difference of 1 Dalton between the predicted and experimental values for all samples (Table 6.1).

N-deacetylation of peptidoglycan was shown to occur as a loss of an acetyl group can be determined for several species. For example, GlcNAc-MurNAc with a dipeptide stem(Di-AE) molecular ion has a m/z of 698.34, an ion with m/z of 656.33 can also be detected corresponding to a loss of an acetyl group (neutral mass of 41.98) (Table 1.1). O-acetylation of peptidoglycan does not appear to occur as no peaks corresponding to an additional acetyl group can be detected. For example, if Di-AE was O-acetylated an ion with a m/z of 740.31 would be detected.

Interestingly, ions corresponding to muropeptides which are N-deacetylated for example Di-AE (656.33) have two separate peaks. (Figure 6.2). This indicates that there are two different muropeptides with the same molecular mass but two different structures, as molecules with the same molecular mass but different structures may be chromatographically resolved. The most likely explanation for this is that both MurNAc and GlcNAc are N-deacetylated, resulting in two different structures with the same mass. Additionally, very small peaks can be detected for the loss of two acetyl groups for Di-AE (614.30) (Figure 6.3), implying N-deacetylation occurring on both MurNAc and GlcNAc.

Table 6.1. Predicted and measured molecular masses of monomeric mucopeptides and their elution times as measured by RP-LC-MS.

Composition of peptide	Exact Mass	Charge state	Predicted m/z	Measured m/z	Time eluted (min)
di (AE)					
GlcNAc-MurNAc	698.3018	$[M+1H]^{1+}$	699.31	698.34	1.44
GlcNAc-MurNAc (minus one acetyl group)	656.2674	$[M+1H]^{1+}$	657.33	656.33	1.44 and 1.94
MurNAc	495.2612	$[M+1H]^{1+}$	496.26	495.23	1.44
GlcN-MurN	614.04	$[M+1H]^{1+}$	615.04	614.30	1.44
tri(AeM)					
GlcNAc-MurNAc	870.39	$[M+1H]^{1+}$	871.39	870.42	1.31
		$[M+2H]^{2+}$	436.69	435.72	1.31
GlcNAc-MurNAc (minus one acetyl group)	828.41	$[M+1H]^{1+}$	829.41	828.43	1.31
MurNAc	667.3412	$[M+1H]^{1+}$	668.34	667.34	1.33
tetra (AEmA)					
GlcNAc-MurNAc	941.43	$[M+1H]^{1+}$	942.43	941.46	1.40
		$[M+2H]^{2+}$	472.21	471.23	1.40
GlcNAc-MurNAc (minus one acetyl group)	899.45	$[M+1H]^{1+}$	900.45	899.45	1.40
		$[M+2H]^{2+}$	451.22	450.23	1.40

MurNAc	738.3812	$[M+1H]^{1+}$	739.38	738.38	1.40
Penta (AEmAA)					
GlcNAc- MurNAc	1012.47	$[M+1H]^{1+}$	1013.50	1012.50	1.49
GlcNAc- MurNAc (minus one acetyl group)	970.49	$[M+1H]^{1+}$	971.49	970.49	1.49
MurNAc	809.38	$[M+1H]^{1+}$	810.38	809.41	1.49

di; GlcNAc-MurNAc A; Alanine, E; Glutamic acid, m; m-DAP AE; Contains an Alanine and Glutamic acid peptide stem, AeM; Contains a peptide stem consisting of Alanine, Glutamic acid and m-DAP, AEmA; Contains a peptide stem consisting of Alanine, Glutamic acid, m-DAP and a D-alanine; AEmAA; Contains a peptide stem consisting of Alanine, Glutamic acid, m-DAP and two D-alanine's.

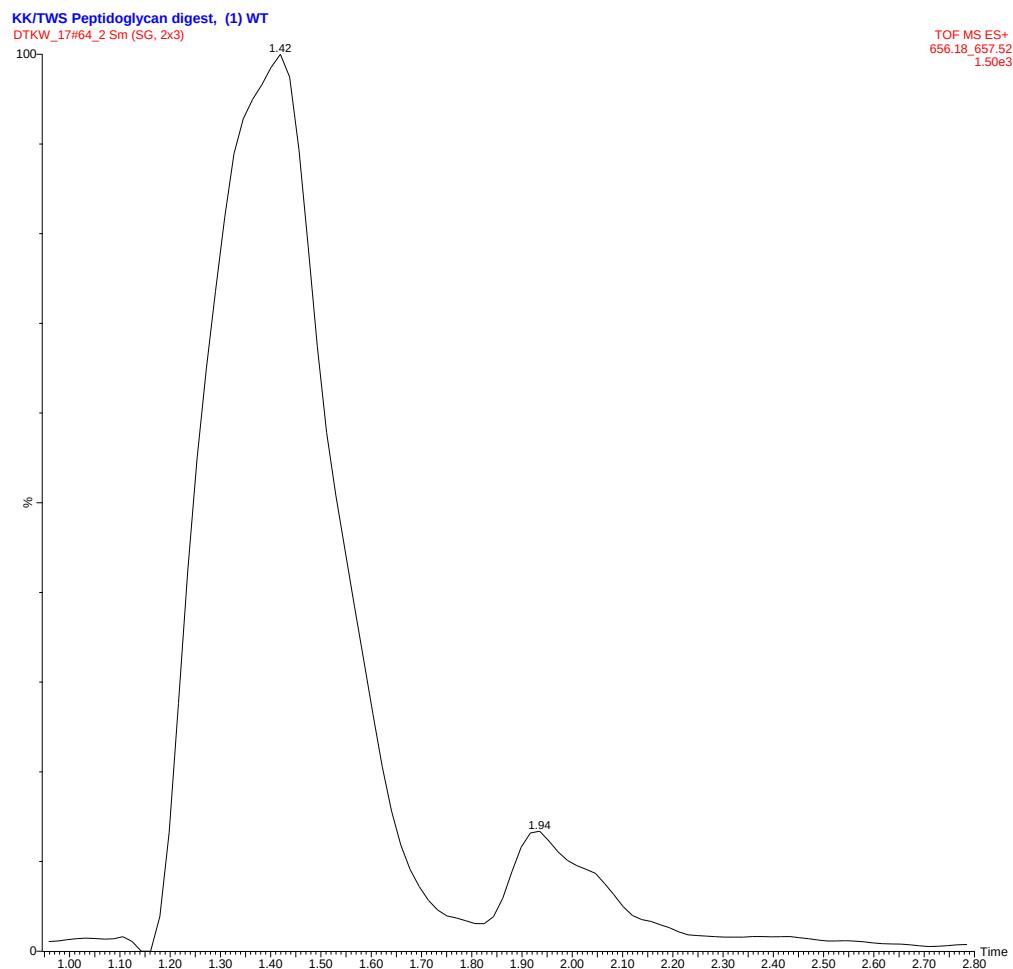


Figure 6.2. Chromatogram of the 656.33 ion. This mass corresponds to the mucopeptide MurNAc-GlcNAc-AE (one of the residues lacking an acetyl group). Unusually, it eluted at two different times (1.42 and 1.94 minutes)

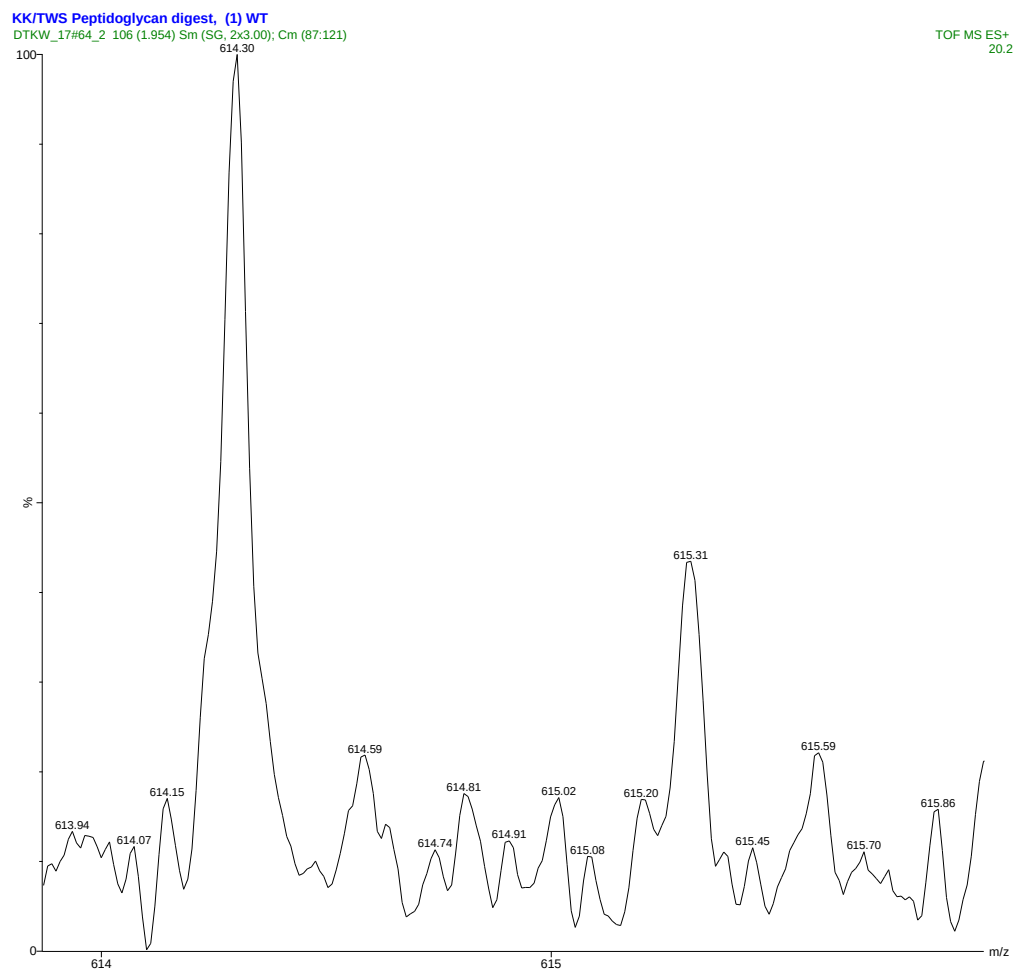


Figure 6.3. 'Zoom' region of Spectrum at 1.44 minutes of ion 614.30. The m/z 614.30 could correspond to the muropeptide MurN-GlcN di(Ae).

6.2.2 LC-MS/MS analysis

To confirm the that N-deacetylation occurs on both MurNAc and GlcNAc residues, the muropeptides were run using a more precise LC-MS/MS using a high-resolution Orbitrap analyser. Several muropeptides were selected and structural analysis of the muropeptides was carried out using MS/MS.

Table 6.2. Proposed structure, accurate mass, quantity and proposed elemental composition of *C. acetobutylicum* muropeptides as determined by LC-MS/MS analysis.

Peptide	Exact mass	m/z [M+1H] ¹⁺	Observed m/z	Peak Area	Elemental composition
Di (AE) GlcNAc- MurNAc	698.2858 697.3018	699.2936 698.3096	699.2939 698.3104	3082344 and 6822239 (2PKS) 5041649 and 7978859 (2PKS)	C27H46N4O17 C27H47N5O16*
Di (AE) GlcNAc- MurNAc minus one acetyl group	655.2674 654.2834	656.2752 655.2912	656.2995 655.3110	32375830 1771338	C25H43N4O16 C25H44N5O15*
Di (AE) GlcN-MurN minus two acetyl groups	612.2490 611.2650	613.2568 612.2728	ND ND		C23H40N4O15 C23H41N5O14*

ND: Not detected; *: Stem peptide consisting of Glutamic acid at position two of the stem peptide instead of Glutamine.

LC-MS/MS using an Orbitrap analyser confirmed the presence of acetylated and N-deacetylated mucopeptides as both masses corresponding to Di (AE) GlcNAC-MurNAC (698.3104) and Di (AE) GlcNAC-MurNAC minus one acetyl group (655.3110) could be observed.

Surprisingly, analysis by LC-MS/MS did not detect Di-AE minus two acetyl groups, and only one peak could be detected for Di-Ae minus one acetyl group. However, two peaks could be detected for Di-Ae no deacetylation; this suggests two different structures with the same elemental composition. One possibility is that *C. acetobutylicum*, O-acetylates its peptidoglycan on either MurNac or GlcNac and so there are three possible structures consistent with m/z 698.28; GlcNac-MurNac with no acetylation and GlcNac-MurNac with one O-acetylation (either on MurNac or GlcNac) and one deacetylation event again occurring either on MurNac or GlcNac. No peak corresponding to O-acetylated GlcNAC -MurNac-AE (740.3) could be identified in either the Q-TOF or Orbitrap LC-MS data. Additionally, peaks corresponding to both Glutamic acid and Glutamine at position two of the stem peptide can be identified (Table 6.2). This is probably due to amidation of glutamine to glutamic acid (Bern, Beniston, and Mesnage 2017).

As mentioned previously for Di-AE (698.3103), there are two peaks at 1.12 and 1.63 minutes. Analysis of the MS/MS spectra of these peaks showed that masses corresponding to MurNac-AE (495.2276) and MurN-AE (435.2017) can be detected in both peaks. These correspond to a loss of GlcNac and GlcNac-OAc respectively. Ions corresponding to GlcNac (204.0860 m/z) can also be identified in both chromatogram peaks, corresponding to the loss of MurNac-AE. However, no ion corresponding to GlcNac-OAc (246.0978 m/z), corresponding to the loss of MurN-AE can be found in either peak (Figure 6.4). Additionally, no ions corresponding to GlcN (162.0762 m/z) or

O-acetylated MurNAc (537.2184) could be identified. Common ions can be identified in both peaks, however, there are differences in the ratio of components. This might be rationalised in that the chromatographic peaks were not fully separated.

The ion corresponding to MurN-AE can be identified, the most likely explanation is that GlcNAc is O-acetylated, however, as no ion corresponding to GlcNAc-OAc can be identified this cannot be confirmed.

Interpretation of the MS/MS spectra of the 656.30 m/z ion, showed fragments corresponding to both N-deacetylated muropeptides MurN and GlcN-AE (162.0762 and 453.2182 m/z) as well as peaks corresponding to the acetylated muropeptides MurNAc and GlcNAc (204.0860 and 495.2276 m/z). It is likely that component 656.2295 probably contains two different structures of the same molecular mass which could not be separated chromatographically. One structure would be GlcNAc N-deacetylated and one MurNAc N-deacetylated.

This data shows that *C. acetobutylicum*'s peptidoglycan is N-deacetylated at both MurNAc and GlcNAc residues, and is most likely O-acetylated on the GlcNAc but not MurNAc residue. The percentage of N-deacetylation is hard to determine due to the possibility of GlcNAc being O-acetylated. However, an estimated 60% of MurNAc and GlcNAc muropeptides are N-deacetylated assuming no O-acetylation and similar molecular ionisation (calculated using peak area, Table 6.2). Individual percentages of GlcNAc/MurNAc N-deacetylation cannot be determined as they eluted as one peak.

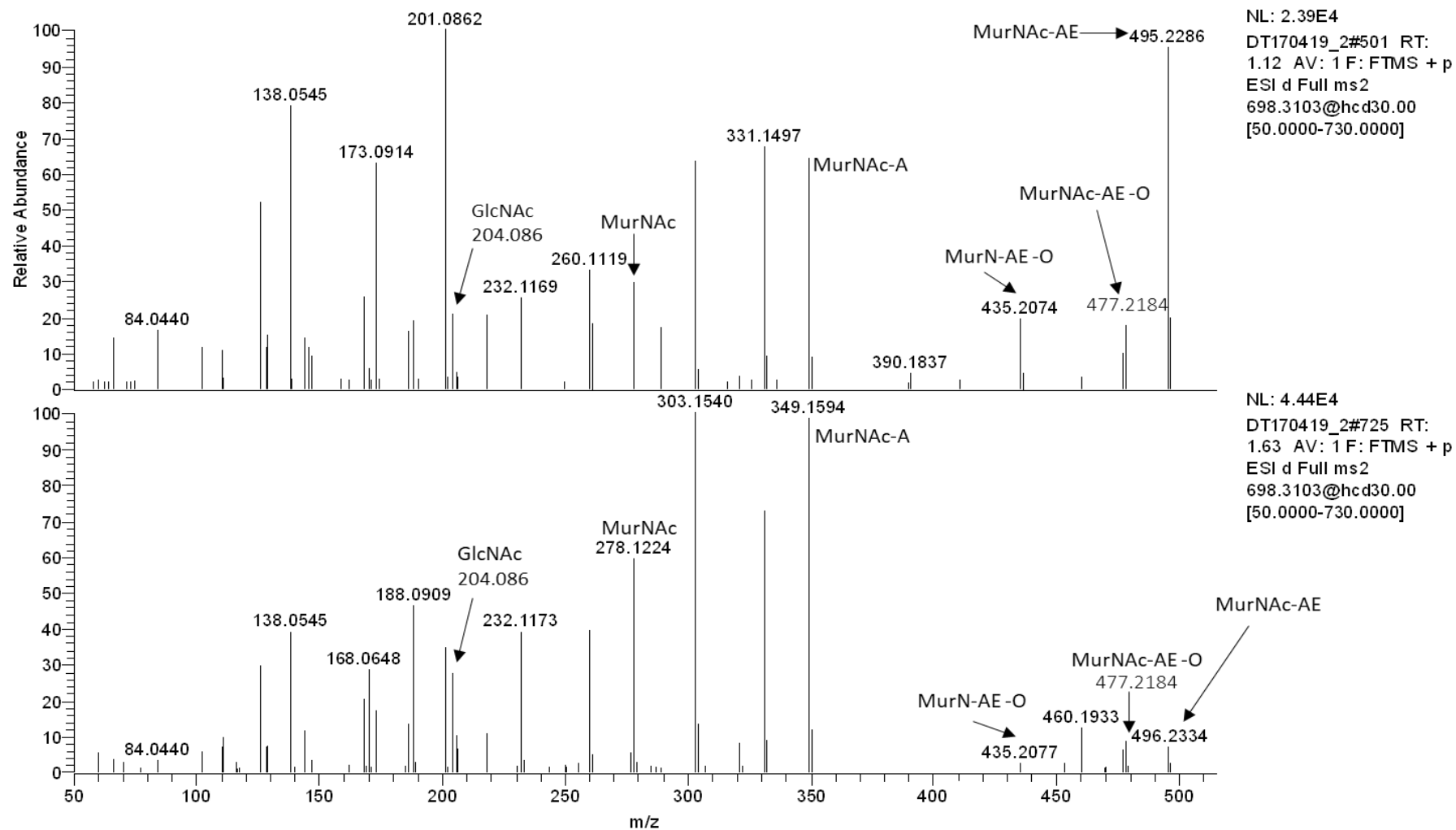


Figure 6.4. MS-MS product ion spectra of precursor ion 698.3103 m/z. Both 1.12 and 1.63 min peaks yield the same fragments. Ions corresponding to the muropeptides GlcNac (204.0860 m/z), MurNac-AE (495.2276 and 477.2184 m/z), MurN-AE (435.2017 m/z), MurNac-A (349.1594 m/z) and MurNac (278.1224 m/z) can be identified.

DT170419_2 #322-1956 RT: 0.72-4.48 AV: 5 NL: 1.41E5
T: Average spectrum MS2 656.30 (322-1956)

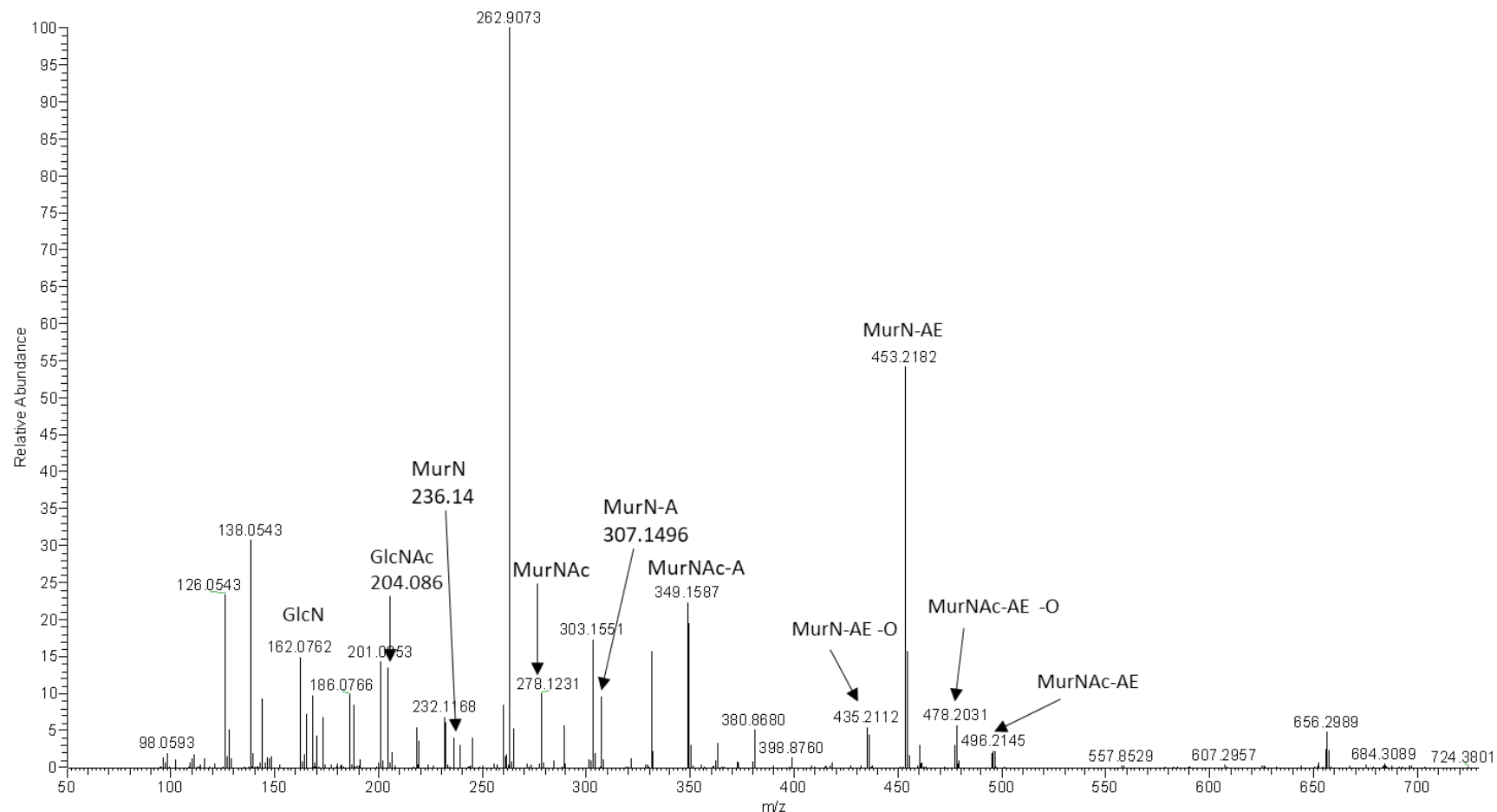


Figure 6.5. MS-MS product ion spectrum of precursor ion 656.30 m/z. Ions corresponding to the mucopeptides GlcNAc (204.0860 m/z), GlcN (162.0762 m/z) MurNAc-AE (495.2276 and 477.2184 m/z), MurN-AE (453.2182 and 435.2112 m/z), MurNAc-A (349.1594 m/z), MurN-A (307.1496 m/z), MurN (236.14 m/z) and MurNAc (278.1224 m/z) can be identified.

6.2.3 Knocking out putative N-deacetylase genes

N-deacetylases are known to carry a non-cleavable signal peptide to direct them to the cell wall and a conserved zinc-binding triad (Asp-184, His-234 and His-238) (Meyrand et al. 2007). Based on the sequence similarity of the *Lactococcus lactis* Pgda gene, 11 putative peptidoglycan N-deacetylases were identified in *C. acetobutylicum*'s genome. From the identified 11 candidates, three lacked the non-cleavable signal peptides indicating that they are secreted and probably xylanase or xylan deacetylases. Additionally, CAC_1037 was found to lack the conserved zinc-binding triad. For the remaining seven candidates, transcriptomic data was used to verify their expression level (average transcript level) (Janssen et al. 2010). Two candidate genes, locus tag CAC_2316 and CAC_3377, with the highest level of expression were selected (Table 6.3) (Janssen et al. 2010).

Table 6.3. Showing candidate *C. acetobutylicum* N-deacetylase genes.

Gene	Classification	Non-cleavable signal peptide	Zinc binding triad	Expression level
CAC_2316	Polysaccharide deacetylase-like protein; Xylanase/chitin deacetylase family enzyme	Yes	Yes	1000
CAC_3377	Xylanase/chitin deacetylase family enzyme	Yes	Yes	300
CAC_0436	Predicted xylanase/chitin deacetylase, similar to yxkH <i>B.subtilis</i>	Yes	Yes	230
CAC_2396	Predicted xylanase/chitin deacetylase	Yes	Yes	150
CAC_0359	Predicted xylanase/chitin deacetylase	Yes	Yes	No data
CAC_1037	Predicted xylanase/chitin deacetylase	Yes	No	2500
CAC_2383	Predicted xylanase/chitin deacetylase	Yes	Yes	170
CAC_3009	Xylanase/chitin deacetylase family protein	Yes	Yes	125
CAC_0090	Xylanase/chitin deacetylase, NodB family	No	No	250
CAC_3017	Predicted xylanase/chitin deacetylase	No	No	5000
CAC_0960	Predicted xylanase related protein	No	Yes	No data

Cac_3377 and CAC_2316 were knocked out using in-frame deletion in a *pyrE* negative strain and *pyrE* was repaired using either pMTL-ME6 or pMTL-JH14-Facoid_Cel48F.

6.2.3.1 HPLC/MS analysis of peptidoglycan of N-deacetylase mutants

Peptidoglycan from both the CAC_3377 and CAC_2316 mutants was purified and characterised as described before. Both GlcNAc and MurNAc appeared to be N-deacetylated in both strains with no difference from wild-type peptidoglycan (data not shown). This indicates that neither CAC_3377 nor CAC_2316 are involved in N-deacetylation of peptidoglycan in *C. acetobutylicum*.

6.2.3.2 Expression of Cel48F in N-deacetylase mutants

Strains were grown as before in CBM (40 mM MOPS, pH 6.8) to an OD of 1.0. The supernatants were concentrated 100 and 50-fold by TCA ppt. Then subjected to SDS-page and western blot analysis.

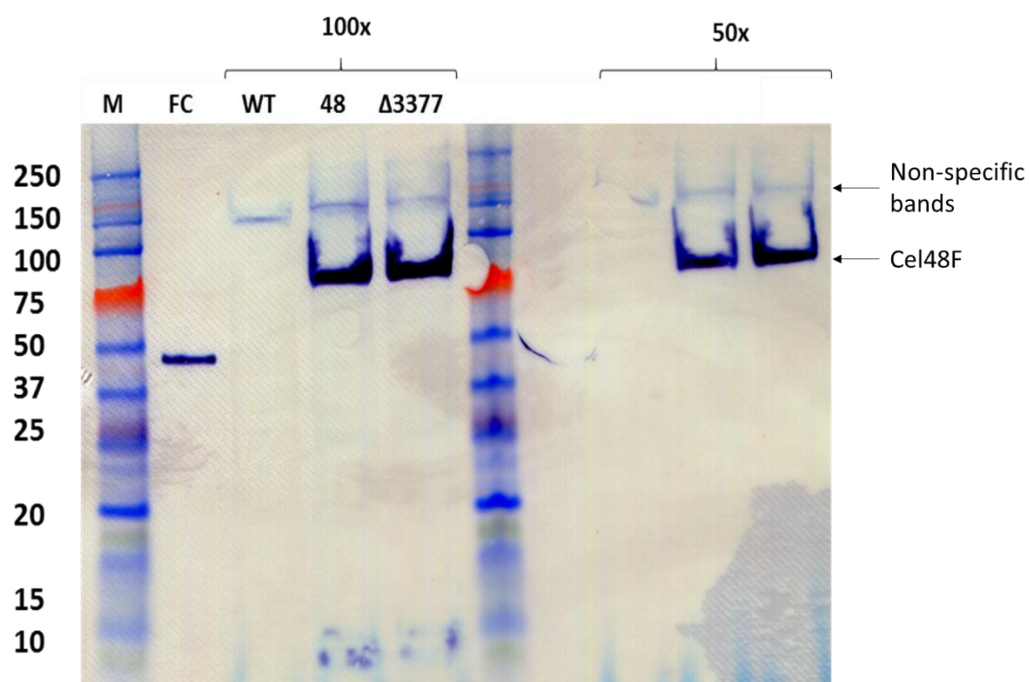


Figure 6.6. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains overexpressing the Cel48F cellulase with the putative peptidoglycan deacetylase CAC_3377 deleted. Cel48F was under control of the *facOID* promoter. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100 and 50-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; FC: FLAG-BAP positive control; WT: Wild-type *C. acetobutylicum*; 48: Wild-type *C. acetobutylicum* expressing Cel48F; Δ3377: *C. acetobutylicum* strain Δ3377 expressing Cel48F.

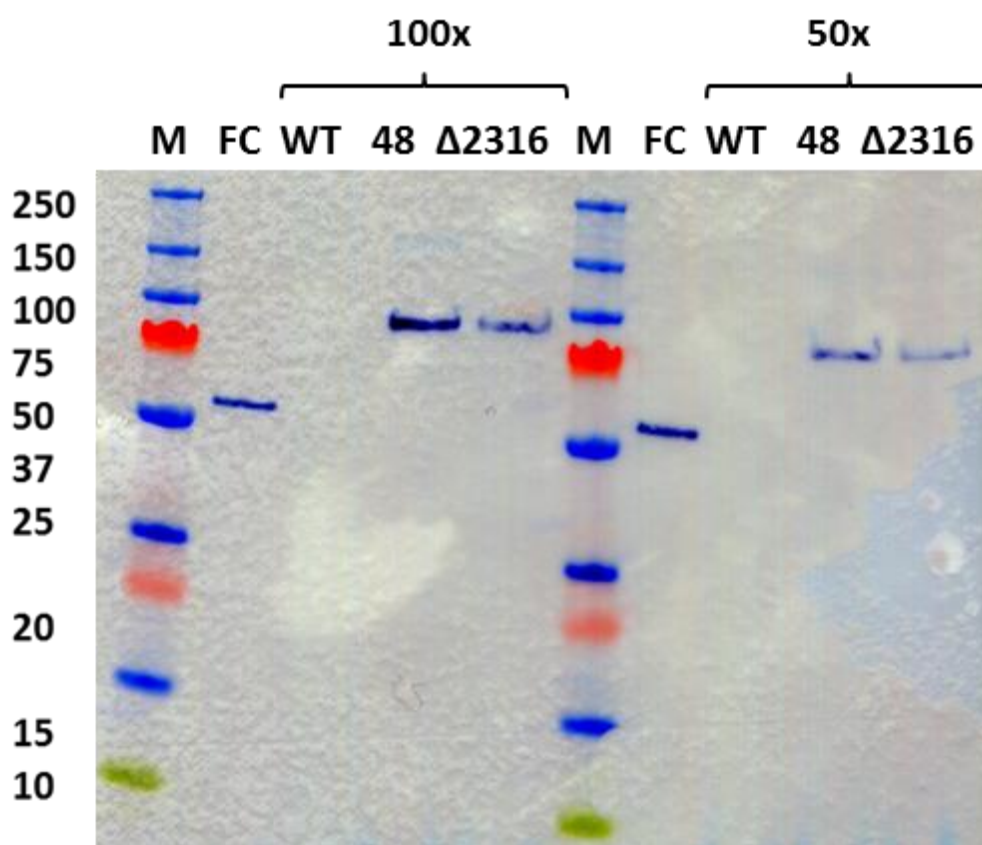


Figure 6.7. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains overexpressing the Cel48F cellulase with the putative peptidoglycan deacetylase CAC_2316 deleted. Cel48F was under control of the *facOID* promoter. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100 and 50-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; FC: FLAG-BAP positive control; WT: Wild-type *C. acetobutylicum*; 48: Wild-type *C. acetobutylicum* expressing Cel48F; Δ 2316: *C. acetobutylicum* strain Δ 2316 expressing Cel48F.

In both Δ 2316 F.48F and Δ 3377 F.48F, the expression of Cel48F was the same as in a wild-type background (Figures 6.6 and 6.6.7). This implies that deletion of CAC_2316 and CAC_3377 does not affect the expression of Cel48F.

6.3 Discussion

Analysis of *C. acetobutylicum*'s peptidoglycan using LC-MS/MS with systems having Q-ToF and Orbitrap analysers revealed that both the MurNAc and GlcNAc residues are N-deacetylated. In addition, the presence of multiple peaks for Di-AE (698.3103 m/z) and the presence of MurN-AE in MS/MS analysis of the Di-AE peaks also suggests that some of the GlcNAc residues are likely to be O-acetylated.

N-deacetylation on both MurNAc and GlcNAc has been previously reported to impart resistance to lysozyme (Moynihan, Sychantha, and Clarke 2014). Approximately 60% of MurNAc and GlcNAc residues were found to be N-deacetylated; this proportion is likely to be higher if some of the GlcNAc residues are O-acetylated. This level of acetylation is similar to the levels observed in other *Bacillus* species (Zipperle, Ezzell, and Doyle 1984; Moynihan, Sychantha, and Clarke 2014).

O-acetylation of GlcNAc residues has been reported in *L. plantarum* and *B. anthracis*. O-acetylation of GlcNAc has been shown to have no role in lysozyme resistance, and has instead been shown to modulate the activity of native autolysins; it probably plays a similar role in *C. acetobutylicum* (Bernard et al. 2011). A protein sequence homology comparison (BLAST) revealed no homologs of *oatB* (*L. plantarum* O-acetyl transferase) or any other *oatA/oatB* genes in *C. acetobutylicum*. A BLAST search of *B. anthracis* *PatA* and *PatB* genes, however, showed several potential homologs in *C. acetobutylicum*. *PatA/B* genes are usually found in Gram-negative bacteria, however, they have been previously identified in *B. anthracis* (Laaberki et al. 2011).

Potential homologs of *PatA/B* genes were also found in *R. cellulolyticum* genome. Therefore, it is possible that the peptidoglycan of *R. cellulolyticum* is also O-acetylated, but this has to be confirmed by detailed peptidoglycan analysis. Taken together, it is

likely that O-acetylation of peptidoglycan by *PatA/B* in Gram-positives it is more common than previously thought.

Deletion of two putative peptidoglycan N-deacetylase genes, individually did not alter the acetylation state of the peptidoglycan residues in these mutants. These results can be explained by;

1. Neither of these two putative deacetylases are involved in the deacetylation of the peptidoglycan and other deacetylases are responsible for deacetylation.
2. Both deacetylases might be able to deacetylate and compensate in the absence of the other.
3. Potential double or multiple mutants have to be created in order to achieve some degree of acetylation and potential resistance to hydrolases.
4. Future work, would involve deletion and characterisation of the remaining putative peptidoglycan N-deacetylase genes and determining if Cel48F expression is increased in an N-acetylated background.

6.4 Key points

- Peptidoglycan from *C. acetobutylicum* is partially N-deacetylated on both MurNAc and GlcNAc.
- Peptidoglycan is also probably O-acetylated on GlcNAc but not MurNAc, and unusually *PatA/B* homologs are probably responsible for this modification. To confirm this, *PatA/B* knockout strains have to be constructed and their peptidoglycan residues characterised.
- Deletion of two putative peptidoglycan N-deacetylase genes individually, did not alter the acetylation state of the peptidoglycan of the single deletion mutants and did not improve the secretion levels of the GH48 family cellulases.

Chapter 7 Co-expression of OrfXp results in increased expression of Cel48F

In Chapter 5, we have identified that Cel48F can bind to N-deacetylated peptidoglycan, but not N-acetylated peptidoglycan. As lysozyme and PGRP's have been shown to be able to kill bacteria through binding to peptidoglycan, this is still the most likely explanation for the low expression level of GH48 family cellulases in *C. acetobutylicum*. In Chapter 6, efforts were made to modify *C. acetobutylicum*'s peptidoglycan by inactivating putative N-deacetylase genes to increase the acetylation levels and thus resistance to GH48 family cellulases. Single deletion of two putative N-deacetylases did not alter the acetylation levels of *C. acetobutylicum*'s peptidoglycan. There are five other potential N-deacetylases in the genome and multiple gene-knockouts will have to be performed, to create a strain with fully acetylated peptidoglycan. Due to the difficulty and time restrictions limiting the generation of a strain displaying fully acetylated peptidoglycan, we set out to investigate the role of other potential factors, such as auxiliary proteins/chaperones present in the cip-cel operon.

7.1 Introduction

Under certain conditions, Cel48A, *C. acetobutylicum*'s native GH48 family cellulase has been reported to be expressed at high levels, although it has no activity towards crystalline cellulose (Sabathé and Soucaille 2003). It is unclear why Cel48A is completely inactive, as all the key residues which have been shown to affect the activity of other Cel48 family cellulases, including the key catalytic residues, are conserved in Cel48A.

It is possible that Cel48A is expressed at high levels in *C. acetobutylicum* due to other mutations, outside of the catalytic domain which might result in its inability to bind to peptidoglycan. These mutations could also affect its ability to bind to cellulose,

explaining why it is inactive. It is also possible, that its secretion is facilitated by an uncharacterised protein, OrfXp, encoded by a gene located downstream of Cel48A, on the same cellulosomal operon.

To test these, hypotheses, overexpression of Cel48A and OrfXp in *C. acetobutylicum* was performed and the resulting strains were compared with those expressing the Cel48F cellulase.

7.2 Results

7.2.1 Over-expression of Cel48A in *C. acetobutylicum*

Cel48A was overexpressed in *C. acetobutylicum* by integrating Cel48A into the *pyrE* locus under control of the *facoid* promoter. A Flag-Tag was added on the C-terminal end of Cel48A. The pMTL-JH14-Cel48A plasmid was used to create the F_48A strain. The F_48F strain and F_48A strain were grown as before to OD₆₀₀ 1.0, in CBM (40 mM MOPS) and the culture supernatants concentrated 100-fold by TCA precipitation. Samples were loaded on an SDS-PAGE gel and analysed by western blot.

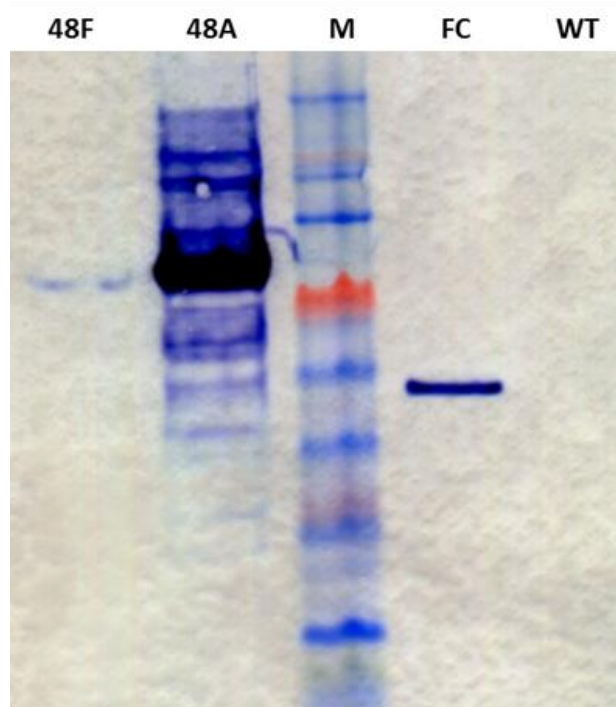


Figure 7.1. SDS-PAGE/Western blot analysis of culture supernatants from *C. acetobutylicum* strains overexpressing the Cel48A and Cel48F cellulases. All strains were under control of the *facOID* promoter. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: Wild-type *C. acetobutylicum*; 48F: Strain expressing Cel48F under control of the *facOID* promoter; 48A: Strain expressing Cel48A under control of the *facOID* promoter.

The expression of Cel48A was significantly higher than that of Cel48F expressed from the same promoter (Figure 7.1), and the gene integrated into the same locus into the genome. This is a very interesting result, and can be explained by the following: 1) Cel48A is inactive and it does not bind or degrade the peptidoglycan, and/or might not trigger the production of the host autolysins in the same way as Cel48F does; 2) Cel48A's dockerin domain may bind to *C. acetobutylicum*'s OrfXp protein via the cohesion-dockerin interaction, and the OrfXp-Cel48A complex allows for higher levels of protein secretion (discussed further below).

7.2.2 Expression of inactive Cel48F in *C. acetobutylicum*

To test whether the activity of Cel48F was the reason for its low expression, Cel48F was mutated at its active site. Several mutations have been shown to completely inactivate Cel48F and lose its activity towards cellulose. For example, substituting the Glutamic acid for Glutamine at position 55 (E55Q) results an inactive Cel48F. (Parsieglia et al. 2008). Quick change mutagenesis was used to generate the mutant in the pMTL-JH14_48F plasmid, which was integrated into the *pyrE* locus of *C. acetobutylicum* to create strain F.48F-E55Q.

The F_48F strain and F_48F-E55Q strain were grown as before to OD 1.0, in CBM (40 mM MOPS), and the supernatant concentrated 100-fold by TCA precipitation. Samples were loaded on an SDS-PAGE gel and analysed by western blot.

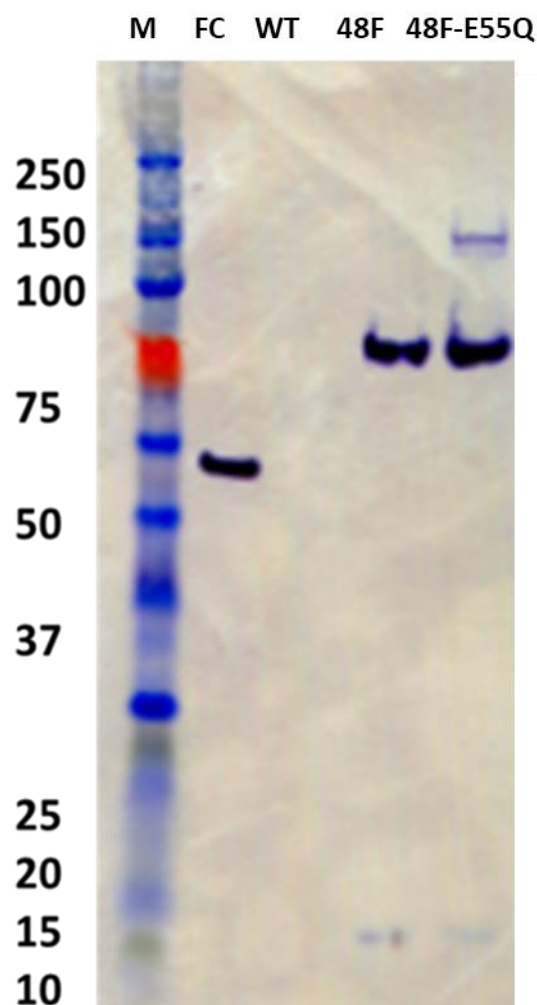


Figure 7.2. SDS-PAGE/Western blot analysis of culture supernatants from *C. acetobutylicum* strains expressing active and inactive Cel48F from *R. cellulolyticum*. All strains were under control of the *facOID* promoter. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; 48F: Strain expressing Cel48F under control of the *facOID* promoter; 48F-E55Q: Strain expressing Cel48F-E55Q under control of the *facOID* promoter.

Surprisingly, Cel48F-E55Q was expressed at the same level as Cel48F (Figure 7.2). This indicates that the high expression of Cel48A is not solely due to its lack of activity.

7.3 OrfXp

In *R. cellulolyticum* a small protein, which consists of a cohesin domain and a transmembrane domain called OrfXp has been identified. There is a homolog of OrfXp in all simple cellulosomal systems and possible homologs in all complex cellulosomal systems. A diagram showing the location of OrfXp in both *C. acetobutylicum* and *R. cellulolyticum*'s cip-cel operons is shown in Figure 1.8. (Pinheiro et al. 2009; Sandrine Pagès et al. 1999). *C. acetobutylicum*'s OrfXp protein shares 43% similarity with that of *R. cellulolyticum*'s, importantly both contain a single cohesin and a transmembrane domain. *C. acetobutylicum*'s OrfXp contains 1 putative transmembrane domain while *R. cellulolyticum*'s OrfXp contains 3 putative transmembrane domains. The basic structure and a sequence comparison is shown in Figure 7.4.

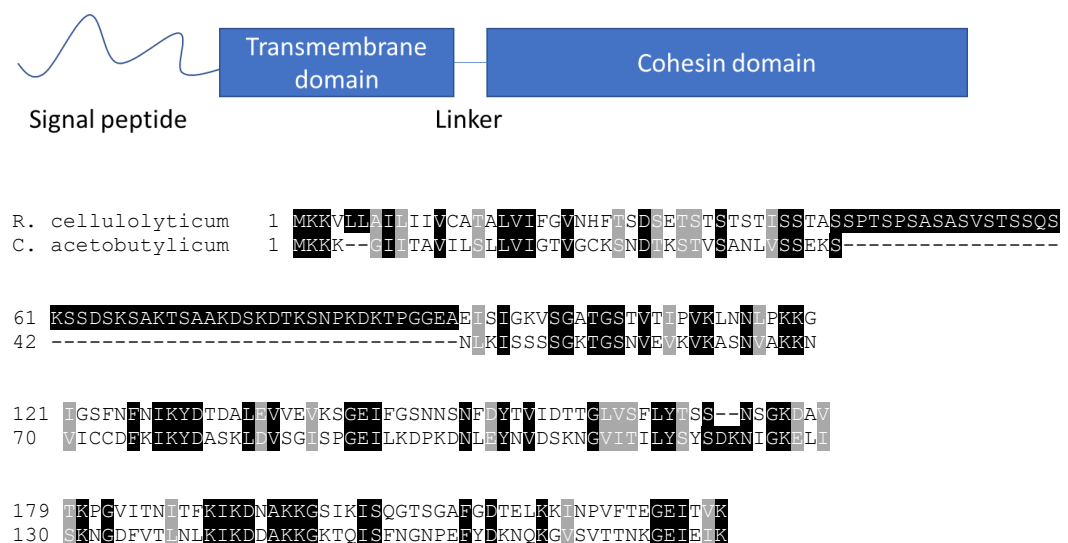


Figure 7.4. Domain composition and sequence comparison of OrfXp A) The domain composition of the OrfXp protein from *R. cellulolyticum* and *C. acetobutylicum*. B) Sequence comparison of the OrfXp proteins from *R. cellulolyticum* and *C. acetobutylicum*.

In *R. cellulolyticum* OrfXp is located in the cell membrane, and it has been shown to bind to dockerin domains. Interestingly, it bound to dockerin domains with a significantly lower affinity than cohesins from CipC the main scaffoldin protein, disassociating from the dockerin domain at a faster rate (Sandrine Pagès et al. 1999).

For this reason, it was hypothesised that OrfXp has a role in cellulosomal assembly by acting as a shuttle, carrying free cellulases to the scaffoldin CipC protein (Sandrine Pagès et al. 1999). The presence of an OrfXp protein has however been shown to not be required for the assembly of cellulosomes either *in vitro* or *in vivo* (Anderson et al. 2011; Fierobe et al. 2001).

Two homologs of OrfXp, OlpC and OlpA have been identified in *R. thermocellum*. They both consist of a type I cohesin domain and an SLH domain and have been found to be attached to the cell wall via the SLH domains (Pinheiro et al. 2009). OlpC is more likely to play a similar role to *R. cellulolyticum*'s OrfXp protein. OlpA binds specifically to several cellulases, while OlpC binds generally to most cellulases, additionally OlpC is relatively abundant, while OlpA is found only at low levels (G. Zhao et al. 2006). OlpC and OlpA SLH domains were both found to bind specifically to SCWP (Secondary cell wall polymer) instead of peptidoglycan (G. Zhao et al. 2006). There is a general trend that scaffoldins, which attach cellulosomes to the cell wall, bind to peptidoglycan, while OlpC and other homologs with a similar function bind to SCWP, it is unclear why this is the case.

Co-expression of OrfXp and a mini-scaffoldin protein along with Cel48F in *C. acetobutylicum* was previously unsuccessful. This was as the construct could not be transferred to *C. acetobutylicum*, even when the expression of Cel48F was under the control of a weakened *thiolase* promoter. For this reason, the potential role of OrfXp in the expression and secretion of cellulases could not be determined (Mingardon et al. 2011).

Since the publication of the Mingardon et al. 2011 study, controllable expression systems have been developed and successfully applied for the expression of cellulosomal components by our group; these were presented in previous chapters

and published (Willson et al. 2016; Kovács et al. 2013). As the hypothesis regarding the potential role of OrfXp is plausible, and might provide evidence for the high level of Cel48A observed (Cel48A was cloned with its native dockerin domain, which has a cognate dockerin domain to *C. acetobutylicum*'s OrfXp's cohesin domain), we set out to re-investigate the role of OrfXp in cellulase expression. All other cellosomal proteins expressed so far in *C. acetobutylicum* have contained either *R. cellulolyticum*'s or *R. thermocellum*'s dockerin domains, which will not be able to bind to *C. acetobutylicum*'s OrfXp cohesin domain.

7.3.1 Expression of OrfXp increases Cel48F expression

To determine if an OrfXp protein can result in increased expression of Cel48F, two strains were constructed. In the first strain, the catalytic domain of Cel48A was replaced with the catalytic domain from Cel48F. Resulting in a chimeric GH48 cellulase, consisting of Cel48F's catalytic domain and Cel48A's dockerin domain. This chimeric protein was integrated into *C. acetobutylicum*'s *pyrE* locus under control of the *facoid* promoter creating the strain F.48F-Chimeric. As the chimeric Cel48F contained Cel48A's dockerin domain, it will be able to bind to *C. acetobutylicum*'s native OrfXp protein. In the second strain, the OrfXp protein from *R. cellulolyticum* was co-expressed with Cel48F containing its native dockerin domain. *R. cellulolyticum*'s OrfXp protein was cloned into the plasmid pMTL82151 under control of the *TcpF* promoter. This construct (pMTL82151-TcpF-OrfXp) was then transformed into the strain F.48F to create the strain F.48F+OrfXp.

In this strain Cel48F will be able to bind to *R. cellulolyticum*'s OrfXp protein via its native dockerin domain. The strains were grown as before to OD 1.0, in CBM (40 mM MOPS), then concentrated 50 and 10-fold by TCA precipitation. Samples were loaded on an SDS-PAGE gel and analysed by western blot.

Strain Name	GH48 Structure		OrfXP?
Cel48F	Cel48F Catalytic domain	Cel48F dockerin	No OrfXP
Cel48A	Cel48A Catalytic domain	Cel48A dockerin	C. acetobutylicum's OrfXP
OrfX	Cel48F Catalytic domain	Cel48F dockerin	R. cellulolyticum's OrfXP
Chimeric	Cel48F Catalytic domain	Cel48A dockerin	C. acetobutylicum's OrfXP

Figure 7.5. Composition of the strains used. Cel48F consists of the catalytic and dockerin domain from *R. cellulolyticum*'s GH48 family cellulase with no cognate OrfX protein. Cel48A consists of the catalytic and dockerin domain from *C. acetobutylicum*'s native GH48 family cellulase and can bind to *C. acetobutylicum*'s native OrfXp protein. OrfX consists of the catalytic and dockerin domain from *R. cellulolyticum*'s GH48 family cellulase and can bind to *R. cellulolyticum*'s OrfXp protein. Chimeric consists of the catalytic domain from *R. cellulolyticum*'s GH48 family cellulase and the dockerin domain from *C. acetobutylicum*'s GH48 cellulase and can bind to *C. acetobutylicum*'s native OrfXp protein.

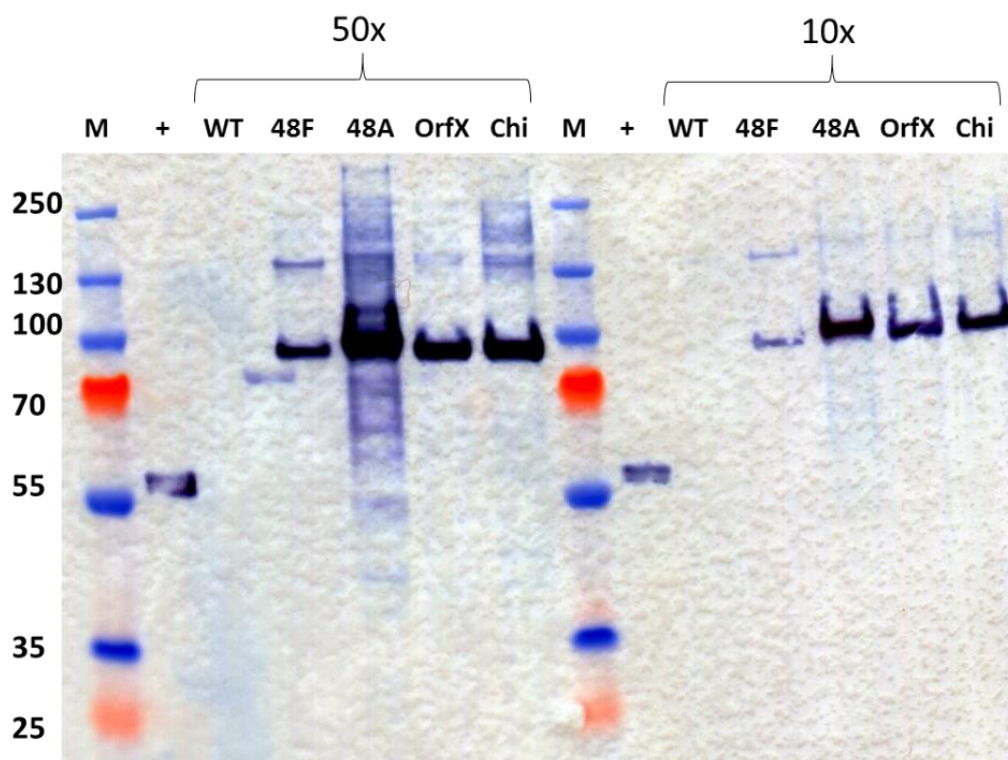


Figure 7.6. SDS-PAGE/Western blot of supernatants from *C. acetobutylicum* strains co-expressing GH48 family cellulases with OrfXp. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold and 50-fold by TCA precipitation. Western blot was carried out using monoclonal ANTI-FLAG M2-peroxidase antibody. M: Prestained Protein Ladder; Flag: FLAG-BAP positive control; WT: wild-type *C. acetobutylicum*; 48F: Strain expressing Cel48F with its native dockerin domain; 48A: Strain expressing Cel48A with its native dockerin domain; OrfX: Strain co-expressing Cel48F with its native dockerin domain and the OrfXp protein from *R. cellulolyticum*; Chi: Strain expressing Cel48F with the dockerin domain from Cel48A.

The expression of Cel48F in the chimeric strain and the strain co-expressing OrfXp was increased compared to Cel48F on its own, with the expression of both only slightly less than Cel48A (Figure 7.6). This indicates that expression of an OrfXp protein with a cognate cohesin domain to Cel48F results in significantly increased expression of Cel48F. Additionally, this might at least partially explain the high level of Cel48A expression (Figure 7.1).

Strain	Concentration of protein ($\mu\text{g/ml}$ of Supernatant)	Fold change compared to Cel48F
Cel48F	0.088	N/A
Cel48A	1.56	17-fold
Cel48F + OrfXp	0.784	8.6-fold
Cel48F – chimeric	1.140	12.5-fold

Table 7.1. Estimated expression of GH48 family cellulases and their relative fold-change compared to Cel48F. The estimated expression level of protein expressed in each strain was calculated using GelQuant, based on the known concentration of FLAG-BAP control protein. *Cel48F*: Cel48F with its native dockerin domain; *Cel48A*: Cel48A with its native dockerin domain; *Cel48F + OrfXp*: Cel48F with its native dockerin domain with the OrfXp protein from *R. cellulolyticum*; *Cel48F - Chimeric*: Cel48F with the dockerin domain from Cel48A.

The approximate fold-change in expression and an estimate of the amount of Cel48F expressed was calculated, using GelQuant. Cel48A was expressed at approximately 17-fold higher levels than Cel48F on its own. Strain, co-expressing Cel48F and *R. cellulolyticum*'s OrfXp protein showed an increased expression of Cel48F by 8.4-fold. Strains expressing the chimeric Cel48F showed an increased expression of Cel48F

12.5-fold when compared to Cel48F on its own. These ratios could be calculated as a FLAG-BAP control protein with a known concentration (50 ng/lane) was loaded onto the SDS-PAGE gel. Cel48F was expressed at approximately 0.088 µg/ml, Cel48A at 1.5 µg/ml, Cel48F plus OrfXp at 0.78 µg/ml and chimeric Cel48F at 1.14 µg/ml (Table 7.1).

7.3.2 OrfXp localisation

To determine the localisation of OrfXp, the wild-type, F.48F and F.48F+OrfX strains were grown as before to OD 1.0, in CBM (40 mM MOPS). The supernatant was concentrated 100-fold as before, a whole cell fraction was also taken and concentrated 10-fold. Samples were loaded on an SDS-PAGE gel, and Coomassie stained, as OrfXp was not tagged, this was as previously adding a tag onto a scaffoldin has been shown to affect its expression (Willson, 2014).

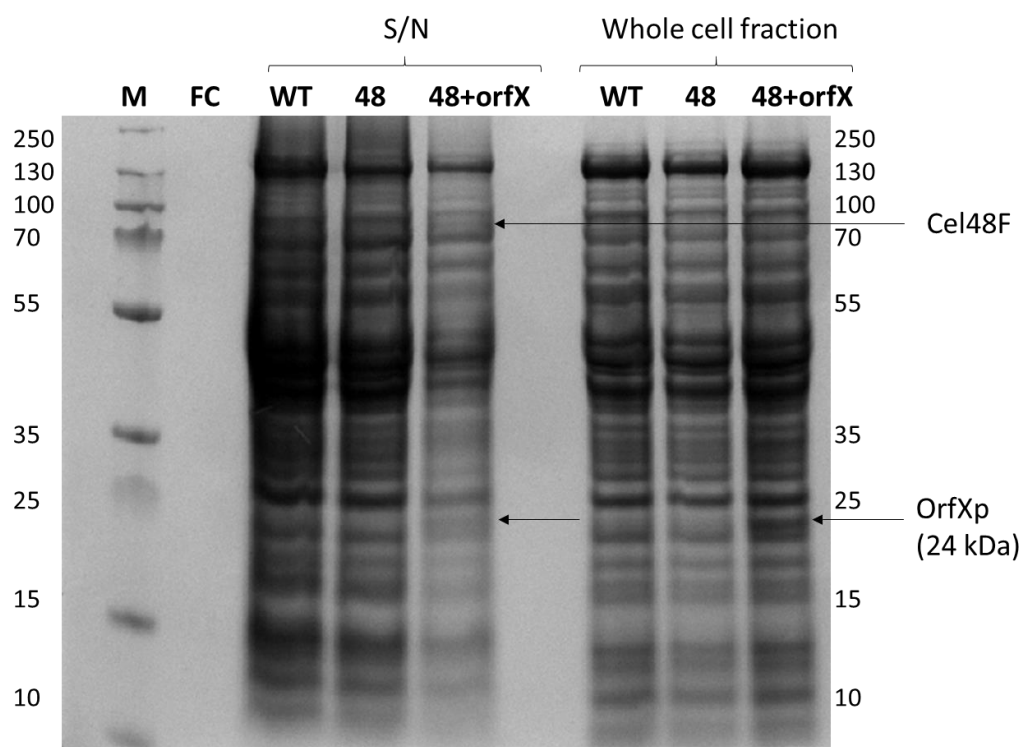


Figure 7.7. SDS-PAGE of supernatants and whole cell fraction from *C. acetobutylicum* strains expressing Cel48F with and without OrfXp. Supernatants were prepared as described in the General Materials and Methods chapter and concentrated 100-fold. Whole cell fractions were prepared as described in the General Materials and Methods section and concentrated 10-fold. Gels were stained with Coomassie blue. M: Prestained Protein Ladder; FC: FLAG-BAP positive control; WT: wild-type; 48: Strain expressing Cel48F under control of the *facOID* promoter; 48+OrfX: Strain co-expressing Cel48F under control of the *facOID* promoter and OrfXp from *R. cellulolyticum* under control of the *TcpF* promoter.

An extra band, most likely corresponding to OrfXp can be seen in the whole cell fraction at approximately 24 kDa, with a significantly weaker band of the same size in the supernatant fraction. Interestingly, the amount of Cel48F secreted is approximately the same as the amount of OrfXp in the supernatant. It is possible that OrfXp is bound to Cel48F in the supernatant (Figure 7.7).

7.3.3 Enzyme activity

An enzyme assay to determine the activity of Cel48F of the strains described above was performed. Strains were grown to an OD of 1.0, in CBM (40 mM MOPS), and the supernatant dialysed using ultra-filtration columns to remove residual sugars whilst concentrating it 40-fold. The supernatants were then incubated in a 1:1 ratio with 2% PASC for 72 hours, and the amount of reducing sugars quantified using DNS assay.

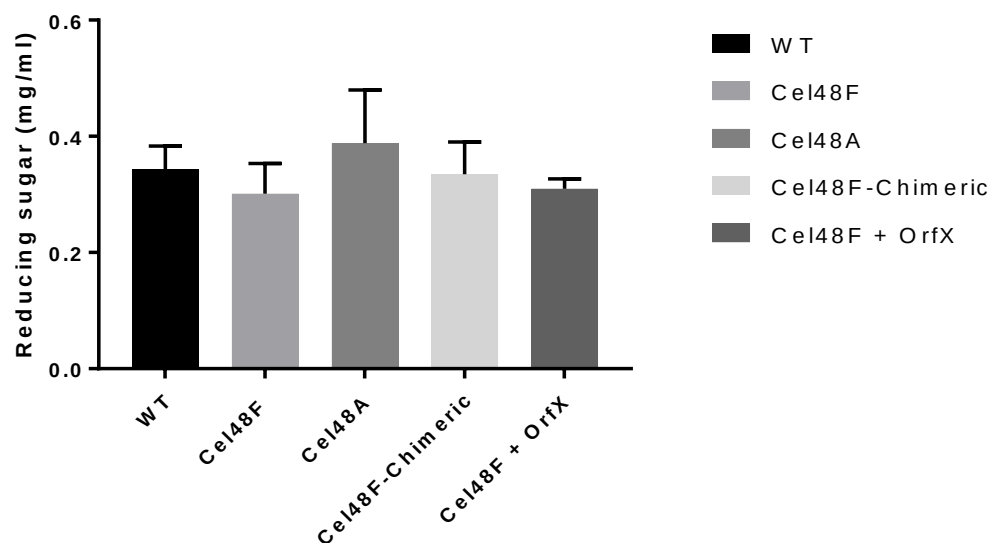


Figure 7.8. Activity of supernatants from *C. acetobutylicum* strains expressing GH48 family cellulases on PASC. 20-fold concentrated supernatants from *C. acetobutylicum* strains were incubated with PASC for 72 hours at 37°C and the levels of reducing sugars released were determined using the sugar release assay as described in the General Materials and Methods section. WT: wild-type *C. acetobutylicum*; Cel48F: Cel48F with its native dockerin domain; Cel48A: Cel48A with its native dockerin domain; Cel48F + OrfXp: Cel48F with its native dockerin domain with the OrfXp protein from *R. cellulolyticum*; Cel48F - Chimeric: Cel48F with the dockerin domain from Cel48A.

All strains released very similar amounts of sugars, regardless of the amount of Cel48F they produced (Figure 7.8). This is probably as Cel48F has very low activity on its own, combined with the relatively high background activity of wild-type *C. acetobutylicum* supernatant on PASC.

7.4 Discussion

OrfXp was found to be essential for the high expression of GH48 family cellulases in *C. acetobutylicum* and presumably in all other organisms which use a cellulosomal system. It is also likely that OrfXp will improve the expression of other cellulases. For example, the GH9 family cellulases: Cel9E and Cel9G, as these have previously been shown to be toxic when expressed in *C. acetobutylicum* (Mingardon et al. 2011). Determining whether OrfXp will increase the level of secretion of these cellulases in *C. acetobutylicum* and other organisms will be an essential step towards characterising OrfXp and engineering *C. acetobutylicum* to be able to grow on lignocellulose.

How OrfXp results in increased expression of Cel48F remains unclear. In *R. cellulolyticum* OrfXp has previously been shown to be localised to the cell wall and it also contains a transmembrane domain, which suggests that it is probably anchored to the cell membrane (Sandrine Pagès et al. 1999). Other OrfXp homologs have also been shown to be attached to the cell wall, for example *R. thermocellum*'s OrfXp homolog OlpC (G. Zhao et al. 2006). In this chapter, the majority of OrfXp was found in the whole cell fraction, presumably attached to the cell membrane, but a small proportion of OrfXp was also found in the supernatant. This is probably as it was expressed at higher than native levels and was secreted from the cell either on its own or bound to Cel48F.

As Cel48F with a cognate dockerin domain to OrfXp's cohesin domain can almost certainly bind to OrfXp via the cohesin-dockerin domain, this would result in Cel48F

bound to the cell membrane via OrfXp. This is supported by the fact that an OrfXp protein had to be expressed with a cognate cohesin domain to result in increased expression of Cel48F.

It was originally proposed that OrfXp acted as a molecular shuttle, temporarily binding cellulases to the cell surface before transferring them to the scaffoldin protein (Sandrine Pagès et al. 1999). However, it is unclear how this would result in increased expression of Cel48F, especially as a scaffoldin protein was not required to increase the expression of Cel48F. This questions the possibility of OrfXp transferring cellulases to a scaffoldin at the cell membrane. Although it should be noted that the expression of chimeric Cel48F (*C. acetobutylicum*'s native OrfXp protein and scaffoldin) was higher than Cel48F plus OrfXp (*R. cellulolyticum*'s OrfXp protein and no scaffoldin), implying that the scaffoldin could have an effect on Cel48F expression. Determining the optimum expression level of OrfXp and if co-expression of a scaffoldin further increases the level of expression of Cel48F, will be needed to ascertain the mechanism of OrfXp.

The transfer of dockerin bound proteins from OrfXp to CipC is possible as dockerins have been shown to be able to be transferred to different cohesins (M. Voronov-Goldman et al. 2015). OrfXp's cohesin domain has been shown to have a lower binding affinity to dockerin domains than CipC cohesins (Sandrine Pagès et al. 1999). In *R. flavefaciens*, a membrane-bound type II cohesin, CohG, was proposed to act as a molecular shuttle for the delivery of scaffoldins to the cell surface (Milana Voronov-Goldman et al. 2014). It binds to dockerins using an alternative binding mode, binding by CohE (the anchoring cohesin) results in displacement of CohG (M. Voronov-Goldman et al. 2015). It would be interesting to determine if a CipC cohesin domain

could displace OrfXp from dockerins, as well as determining if OrfXp also uses an alternative binding mode similar to CohG.

In previous chapters, we showed that Cel48F can bind to N-deacetylated peptidoglycan, so the most likely mechanism by which OrfXp increases expression of Cel48F is by preventing it from binding to peptidoglycan. This can happen either at the cell membrane or once secreted and at a distance from the cell wall, Cel48F is transferred from OrfXp to a scaffoldin.

Cel48F purified from *E. coli* was found to bind to N-deacetylated peptidoglycan but not N-acetylated peptidoglycan. It would be interesting to determine if the expression of Cel48F could be further increased in *C. acetobutylicum* with N-acetylated peptidoglycan. This is supported by the fact that Cel48A is expressed at a significantly higher level than Cel48F-Chimeric, even though the only difference between the two proteins is the catalytic domain. As the expression level of inactive Cel48F was the same as active Cel48F this is unlikely to be solely due to its lack of activity. One possibility is that Cel48A is expressed at higher levels as it is unable to bind to N-deacetylated peptidoglycan. If this is the case, this would again support the theory that expression of Cel48F would be further increased in a strain with N-acetylated peptidoglycan.

If Cel48A is unable to bind to peptidoglycan, this would imply that the level of Cel48F could be further increased in a *C. acetobutylicum* strain with N-acetylated peptidoglycan. Additionally, determining if OrfXp bound to Cel48F can prevent Cel48F from binding to peptidoglycan would also indicate that OrfXp increases the expression of Cel48F by preventing it from binding to peptidoglycan.

OrfXp could alternatively support the expression of Cel48F, by acting as a protein chaperone. If this is the case, OrfXp would allow for: 1) Cel48F to fold properly, 2) OrfXp

could protect Cel48F from degradation by proteases during folding, 3) OrfXp could prevent Cel48F from binding to components of the cell wall, for example, peptidoglycan during folding.

More experimental evidence is needed to determine the role of OrfXp in cellulase expression. Firstly, it will be necessary to confirm whether OrfXp can transfer cellulases to a scaffoldin protein. It will be interesting to determine if this is by a similar mechanism to CohG described in *R. flavefaciens*. Secondly, it will be necessary to determine if the co-expression of OrfXp and a scaffoldin result in a further increase in expression of Cel48F. Thirdly, the exact mechanism by which OrfXp improved the expression of Cel48F will need to be determined. Additionally, a knockout of OrfXp in *R. cellulolyticum*, and characterisation of the mutant would be helpful. It would also be interesting to determine if OrfXp can increase the expression of other cellulases.

7.5 Key points

- *C. acetobutylicum*'s native GH48 family cellulase, Cel48A is expressed at high levels in *C. acetobutylicum*.
- Inactive Cel48F is expressed at the same level as active Cel48F in *C. acetobutylicum*.
- The co-expression of OrfXp with Cel48F results in significantly increased levels of secreted Cel48F in *C. acetobutylicum*.
- OrfXp and homologs of OrfXp are likely to play a crucial role in enabling high levels of cellulase expression in *R. cellulolyticum* and other cellulosomal organisms.

General Discussion

Outcomes of the thesis

The overall aim of this study was to engineer a strain of *C. acetobutylicum* that is suitable for use as a CBP organism, i.e. capable of utilising lignocellulose and producing solvents. Previous studies have shown that the main barrier to engineering *C. acetobutylicum* to grow on lignocellulose is the low expression and secretion of cellulases from the GH9 and GH48 families (Mingardon et al. 2011; Willson et al. 2016). For this reason, the primary aim of this project was to increase the expression of GH48 and GH9 family cellulases in the solvent producing host, *C. acetobutylicum*.

A novel lactose inducible system was used to test the expression levels of several cellulases belonging to several GH families. Using this system, a high level of secretion was observed for cellulases belonging to the GH5 family. However, the secretion levels of GH9 and GH48 family cellulases were very low under all tested induction and growth conditions. Additionally, when high levels of induction were used, this caused a reduction in growth rate, which suggests that the expression and secretion of GH48 and GH9 cellulases is toxic to *C. acetobutylicum*. The high level of secretion observed for the GH5 family cellulases using the lactose-inducible system indicates that this expression system might be the preferred for the production of cellulosomal components and other proteins in the future.

Toxicity of the GH48 and GH9 enzymes to *C. acetobutylicum* has been observed in other studies (Mingardon et al. 2011; Chanal et al. 2011; Willson et al. 2016). In the Mingardon et al. study, GH5, GH8 and some GH9 cellulases were successfully expressed under control of a weakened *thiolase* promoter and were shown to be secreted. However, cellulases from the GH48 family and some from the GH9 family could not be transformed into *C. acetobutylicum* even under the control of a weakened

promoter. These cellulases were shown to be expressed at high levels intracellularly, indicating that might be an issue during secretion. Willson et al 2016. applied improved genomic integration techniques, which allowed for the expression and secretion of GH48 family cellulases. However, secretion levels were significantly lower when compared to cellulases belonging to GH5 and GH10 families.

We have selected four GH48 family cellulases from several well-known cellulolytic organisms for heterologous expression in *C. acetobutylicum*. All selected GH48 family cellulases were found to be secreted at very low levels in *C. acetobutylicum* when compared to cellulases belonging to GH5 and GH10 families. This finding indicates that there is a common factor impeding the expression and secretion of GH48 family cellulases in this host.

In an attempt to increase the expression levels of cellulases belonging to GH48 family, we aimed to improve the stability of mRNA by adding 5' and 3' UTR's to the coding region of Cel48F (*R. cellulolyticum*'s GH48 family cellulase). These UTRs have been shown to result in increased levels of expression of proteins in several different organisms including *C. acetobutylicum* (Phan, Nguyen, and Schumann 2013; J. Lee et al. 2016). Unfortunately, only minor increases in expression and secretion levels of Cel48F was achieved, indicating that mRNA stability is not the main reason for the low expression of GH48 family cellulases.

In *B. subtilis*, high levels of misfolded proteins have been shown to activate the CssR-CssS protein misfolding system (Hyyryläinen et al. 2001; Darmon et al. 2002). PrsA chaperones, the only extracellular chaperones identified in Gram-positive bacteria have been shown to be essential for the folding of some proteins (Jacobs et al. 1993). Increasing the expression of the PrsA protein chaperone has been shown to increase the expression of several proteins (Lindholm et al. 2006; J. Chen et al. 2015). In *B.*

subtilis depletion of PrsA results in secretion stress when proteins are overexpressed, for example, amylase (Hyryläinen et al. 2001). The misfolding of GH48 family cellulases during secretion could partially explain the toxicity of GH48 family cellulases, as this could activate the CssR-CssS protein misfolding system. This observation is further supported by the Mingardon et al. 2011 study, where it was shown that the removal of the signal-peptide allowed for relatively high levels of intracellular Cel48F accumulation.

PrsA chaperones from *B. subtilis*, *C. acetobutylicum* and *R. cellulolyticum* were overexpressed in *C. acetobutylicum*. However, overexpression of these chaperones did not improve the secretion level of Cel48F. This could indicate that PrsA chaperones might not be involved in the correct folding or secretion of this particular cellulase in *C. acetobutylicum*. If misfolding of Cel48F is the explanation for the low level of active protein, it is likely that other, not yet identified, chaperones might be responsible for the misfolding of GH48 family cellulases.

Significantly, Cel48F was found to be able to bind to N-deacetylated but not N-acetylated peptidoglycan. As we found that *C. acetobutylicum*'s peptidoglycan is at least partially deacetylated, it is possible that the expression of Cel48F is toxic to *C. acetobutylicum* during secretion by binding to its peptidoglycan. The binding of proteins to peptidoglycan has been shown to result in cell death by two possible mechanisms: (1) proteins binding to peptidoglycan activating the misfolded protein response, or by (2) recruiting or activating native autolysins (Herbert et al. 2007; Kashyap et al. 2011). The fact that purified Cel48F was unable to hydrolyse the isolated peptidoglycan or any other components of *C. acetobutylicum*'s cell wall, suggests that either or both of the above-mentioned mechanisms could be responsible for the toxicity of Cel48F to *C. acetobutylicum*.

The acetylation levels of *C. acetobutylicum*'s peptidoglycan were determined by LC-MS and LC-MS/MS. This analysis confirmed that the peptidoglycan is at least partially N-deacetylated on both GlcNAc and MurNAc residues, and possibly O-acetylated on the GlcNAc residue. This may also help explain why *C. acetobutylicum* is unable to grow on lignocellulose, as it is likely that N-acetylation of peptidoglycan is required for high levels of secretion of key cellulases. *C. acetobutylicum* has been isolated from the digestive tracts of mammals (Cato, E.P, W.L. George, and S.M. Finegold. 1986), where N-deacetylation of peptidoglycan is probably essential to protect *C. acetobutylicum* from host defences, for example, lysozyme.

A total of 11 putative peptidoglycan deacetylases were identified in *C. acetobutylicum*'s genome. Two of these putative peptidoglycan N-deacetylases were shown to be expressed at the highest levels and were therefore chosen for inactivation experiments (Janssen et al. 2010). Unfortunately, single deletion of these genes did not affect the acetylation levels of the peptidoglycan. These results can be explained by: 1) Neither of these two putative deacetylases are involved in the deacetylation of the peptidoglycan and other deacetylases are responsible for deacetylation. 2) Both deacetylases might be able to deacetylate and compensate in the absence of the other. 3) Potential double or multiple deletion mutants have to be created in order to achieve some degree of acetylation and potential resistance to hydrolysis. Future work, would involve deletion and characterisation of the remaining putative peptidoglycan N-deacetylase genes and determining if Cel48F secretion levels are increased in an N-acetylated background.

Cel48A, *C. acetobutylicum*'s native GH48 family cellulase which has been shown to be completely inactive on cellulose (Phillipe Soucaille, personal communication), was shown to be expressed at significantly higher levels in *C. acetobutylicum* than Cel48F.

This was shown to be unrelated to its activity, as inactivation of the Cel48F did not result in higher levels of secreted inactive Cel48F. This could indicate that the inactive GH48 family cellulases might still be able to bind to the peptidoglycan and cause cell death by one of the proposed mechanisms described above.

In *R. cellulolyticum*, a small protein which consists of a cohesin domain and a transmembrane domain called OrfXp has been identified. There is a homolog of OrfXp in all simple cellulosomal systems and possible homologs in most complex cellulosomal systems (Pinheiro et al. 2009; Sandrine Pagès et al. 1999). It was hypothesised to help in cellulosomal assembly by acting as a shuttle, carrying free cellulases to the scaffoldin CipC (Sandrine Pagès et al. 1999).

Co-expression of OrfXp and Cel48F in *C. acetobutylicum* resulted in significantly higher levels of secreted Cel48F. *C. acetobutylicum* also has an OrfXp homolog. Replacement of Cel48F's dockerin domain with the dockerin domain from Cel48A, *C. acetobutylicum*'s native GH48 family cellulase also resulted in significantly increased secretion levels of the chimeric cellulase. Replacement of the dockerin domain would allow the chimeric cellulase to bind to *C. acetobutylicum*'s native OrfXp protein as they have cognate cohesin-dockerin domains.

However, the expression of the chimeric Cel48F consisting of Cel48A's dockerin domain and signal peptide was still expressed at lower levels than Cel48A. This may be because Cel48A is inactivated in such a way that it cannot bind to N-deacetylated peptidoglycan, supporting the theory that the expression of Cel48F could be increased further in a strain with N-acetylated peptidoglycan.

Limitations and proposed future work

As OrfXp has been shown to result in increased expression of Cel48F in *C. acetobutylicum*, it is also likely to be able to increase the expression of other cellulases

which are expressed at low levels. Potential candidates are GH9 family cellulases, for example, Cel9E and Cel9G from *R. cellulolyticum*, as these have also been shown to be secreted at low levels in *C. acetobutylicum*.

The elucidation of the role of OrfXp adds another paradigm to the continuing complexity of the cellulosome, as it is essential for the expression of GH48 family cellulases in *C. acetobutylicum*. As an OrfXp homolog is found in all cellulosomal bacteria, it is likely that an OrfXp homolog would be essential for the secretion of GH48 family cellulases and potentially other cellulases in all cellulosomal bacteria.

OrfXp was successfully inactivated in *R. cellulolyticum*, implying that it is not essential. However, the strain was not characterised for its growth on cellulosic substrates (C. Xu et al. 2015). A reduction in the level of Cel48F, and possibly several other cellulases would be expected. Deletion of Cel48F has been reported to result in a strain which is unable to grow on crystalline cellulose (T. Xu et al. 2014); it would be interesting to determine whether an OrfXp negative strain would be able to grow on crystalline cellulose.

As Cel48F has been shown to bind to N-deacetylated peptidoglycan, it would also be interesting to determine the acetylation state of peptidoglycan in native cellulolytic organisms, for example, *R. cellulolyticum* as well as determining if Cel48F can bind to *R. cellulolyticum*'s peptidoglycan.

The mechanism by which OrfXp aids the secretion of Cel48F is still unclear. Elucidation of the mechanism by which OrfXp aids the secretion of cellulases and the assembly of cellulosomal complexes would expand our understanding of cellulolytic organisms.

We propose two potential mechanisms by which OrfXp can aid the secretion of Cel48F; Firstly, it might prevent Cel48F from binding to peptidoglycan, and secondly by acting as a protein chaperone.

It is my view, OrfXp is probably preventing the GH48 cellulases from binding to peptidoglycan; this is as Cel48F has been shown to be able to bind to N-deacetylated but not N-acetylated peptidoglycan. Determining if OrfXp prevents Cel48F from binding to peptidoglycan *in vitro* would be relatively straight forward. It would also be interesting to determine if OrfXp would be required in a host with N-acetylated peptidoglycan. Although purified Cel48F has been shown to be unable to bind to N-acetylated peptidoglycan, Cel48F may be able to bind to N-acetylated peptidoglycan before folding or if it misfolds. If this is the case, further increases in the secretion levels of Cel48F and other cellulases could be achieved in *C. acetobutylicum* without the need to modify the acetylation level of its peptidoglycan.

One possibility for determining if OrfXp increases the secretion levels of Cel48F by acting as a protein chaperone would be to express Cel48F in a strain with the control proteases deleted. If OrfXp acts as a protein chaperone, Cel48F would be misfolded and so inactive in a strain without OrfXp and active in a strain with OrfXp co-expressed. This method has been used to confirm the role of PrsA as a protein chaperone previously (Hyryläinen et al. 2001).

Previous work in our laboratory and the wider literature has shown the successful assembly of mini-cellulosomes in *C. acetobutylicum*. Crucially, mini-cellulosomes have also been successfully anchored to the cell of *C. acetobutylicum* (Willson et al. 2016). The low expression of cellulases in particular GH48 family cellulases was thought to be the main reason why the recombinant strains were unable to grow on the lignocellulosic material (Willson et al. 2016).

The increase in the expression of GH48 family cellulases and probably other cellulases raises the prospect of engineering *C. acetobutylicum* to be able to grow on lignocellulose. The only other foreseeable hurdle is the expression of scaffoldins, as

the expression of recombinant scaffoldins with a large number of cohesins resulted in low expression. This has been thought to be due to mRNA stability issues. However, the use of a more complex system involving adaptor scaffoldins could be utilised instead circumventing the issue (Stern et al. 2016).

Previously, the use of designer cellulosomes was attractive for the construction of a cellulolytic species as they have several advantages over native cellulosomes. However, as OrfXp is likely to be required for the expression of multiple cellulases, this would require a different OrfXp protein for each different dockerin domain used. This would greatly increase the number of components needed to be expressed, possibly making designer cellulosomes impractical.

For these reason(s) the easiest route to engineering *C. acetobutylicum* to grow on cellulose would be to repair its native cellulosome. Previous work has shown that the only cellulase shown to be inactive in *C. acetobutylicum*'s native Cip-Cel operon is Cel48A (Philippe Soucaille, personal communication). So, repair of Cel48A by swapping Cel48A's inactive catalytic domain with Cel48F's catalytic domain would result in a functional GH48 cellulase. This should result in a functional cellulosome. Additionally, as *C. acetobutylicum*'s Cip-Cel operon is not expressed at high levels, replacing the native promoter for a stronger promoter will probably be necessary. This is feasible using established genetic tools.

Alternatively, as *C. acetobutylicum*'s native cellulosome is relatively simple, it may be necessary to utilise a cellulosome from an established cellulolytic organism, for example, *R. cellulolyticum*. A more complex cellulosome could also be designed and integrated into *C. acetobutylicum*. For growth on lignocellulose, xylanases and other hemicellulases would also need to be expressed.

Conclusions

GH48 family cellulases were shown to be able to bind specifically to N-deacetylated but not acetylated peptidoglycan. *C. acetobutylicum*'s peptidoglycan was also determined to be partially N-deacetylated. Proteins binding to peptidoglycan have been shown to be toxic to cells, this may explain the toxicity of GH48 family cellulases to *C. acetobutylicum*.

Co-expression of OrfXp, a small protein with a cohesin domain and believed to be anchored to the cell membrane, was shown to significantly improve secretion levels of GH48 family cellulases in *C. acetobutylicum*. This shows the importance of the OrfXp protein for the functionality of cellulosomal systems. However, the mechanism by which it aids the secretion of individual cellulases remains unclear and requires further studies.

This increase in the secretion level of GH48 family cellulases in *C. acetobutylicum* is an essential step towards being able to engineer *C. acetobutylicum* to grow on lignocellulose. Engineering *C. acetobutylicum* to be able to grow on lignocellulose is now extremely promising.

Professional Internships for PhD Students Reflection Form

Name of Organization

BBSRC

Details of Placement

Please describe your main activities during the placement.

The main role of my placement was to develop a series of case studies for the UK Synthetic Biology Strategic Plan 2016. In this role, I identified potential case studies and carried out interviews.

During the course of my placement I attended several synthetic biology conferences including; UK Synthetic Biology conference, OpenPlant conference and Bristol SynBio as well as the Synthetic Biology Leadership Council. Networking events during these conferences were valuable in identifying potential case studies.

I was also responsible for mapping the UK synthetic biology landscape, this involved assessing the RCUK awarded grants. I was also involved in organising a DNA synthesis mission, this was to facilitate collaboration between the UK and the US synthetic biology communities. I was also involved in the process of assessing grant applications.

Placement Achievements

Please detail all outcomes from the placement, including any publications, presentations given and reports written etc.

I was responsible for all case studies in the Synthetic Biology roadmap refresh: Biodesign for the Bioeconomy: UK Synthetic Biology Strategic Plan 2016 is published by the SBLC. SBLC Members, Observers, Policy Advisors and Secretariat, December 2015.

I was also responsible for compiling a portfolio on all UK government funding on synthetic biology. Additionally, I wrote an internal BBSRC report on the current UK Synthetic Biology landscape in comparison to the US and Europe.

Skill development

Has this Placement helped you developed any new skills or enhanced your previous skill set?

My placement helped develop several important skill sets, my main role was developing case studies for the synthetic biology strategic plan. This helped develop several skills, including; general writing skills, as I had to be able to communicate complex scientific ideas to be suitable for a lay audience.

Appendices

Appendix 1

Table showing the list of materials and reagents along with supplier.

Materials	Supplier
Agar technical	Oxoid (LP0013)
Agarose	Fisher Scientific (BP1356)
Amino-terminal Flag-BAP control protein	Sigma-Aldrich (P7582)
Ampicillin	Sigma-Aldrich (A9393)
Avicel® PH-101	Sigma-Aldrich (11363)
Benzonase® Nuclease	E1014 SIGMA
Bicinchoninic Acid Kit for Protein Determination	Sigma (BCA1)
Bicinchoninic Acid Protein Assay Kit	Sigma-Aldrich (BCA1)
BioLector	M2p labs
BioLector Round Well Plate	M2p labs (MTP-R48-B)
BioLector Sealing foil reduced evaporation	M2p labs
BioRad GenePulser Xcell™	BioRad (165-2662)
BioRad Micropulser™	BioRad (165-2100)
BioRad TransBlot® Turbo™ Transfer System	BioRad (170-4155)
Bioruptor® Plus	Diegenode
Biotin	Sigma-Aldrich (B4501)
Bovine serum albumin (BSA)	Sigma-Aldrich
BugBuster® Protein Extraction Reagent	Sigma
Calcium chloride dihydrate	Sigma-Aldrich (31307)
Chloramphenicol	Sigma-Aldrich (C0378)
CLARIOstar	BMG LABTECH

ColorPlus™ Prestained Protein Ladder, Broad Range (10-230 kDa)	New England Biolabs (P7711)
Compact XS/S Horizontal Gel Electrophoresis Apparatus	Biometra (846-025-000)
Coomassie Brilliant Blue R	Sigma-Aldrich (B0149)
Corning® 96 well plates, clear bottom	CLS3603 SIGMA
Corning® Spin-X® UF Concentrators	cls431483
COVER SLIPS 18 X 18 MM NO. 0	VWR International Ltd.631-0119
DNS (3, 5-dinitrosalicylic acid)	Sigma-Aldrich (D0550)
DOC (deoxycholine), sodium salt	Sigma-Aldrich (D6750)
DTT (dithiothreitol)	Sigma-Aldrich (43819)
EDTA (ethylenediaminetetraacetic acid disodium salt)	Sigma-Aldrich (ED2SS)
Electroporation cuvette, 4mm	Sigma-Aldrich (Z706094)
Electroporation cuvettes, 2mm	Flowgen Bioscience (FBR-102)
Erythromycin	Sigma-Aldrich (E5389)
Ethanol	Fisher Scientific (G/0650DF/17)
Ethidium bromide	Life Technologies (15585011)
Fluoroorotic acid	Europa Biosciences (ML501A)
Gel Loading Dye, Blue (6X)	New England Biolabs (B7021S)
GenElute™ Bacterial Genomic DNA Kit	Sigma-Aldrich (NA2120)
GenElute™ HP Plasmid MiniPrep Kit	Sigma-Aldrich (NA0160)
Glucose	Sigma-Aldrich (G8270)
Glycerol	Sigma-Aldrich (G5516)
Glycine	Sigma-Aldrich (G8898)

HisPur™ Ni-NTA Spin Purification Kit	Thermo (88228)
Hydrochloric acid	Fisher Scientific (A144)
Iron (II) sulfate heptahydrate	Sigma-Aldrich (215422)
Kanamycin	Sigma-Aldrich (K1377)
Lysozyme	Sigma-Aldrich (L6876)
Magnesium sulfate heptahydrate	Sigma-Aldrich (230391)
Manganese sulfate monohydrate	Sigma-Aldrich (221287)
Marvel Original Dried Skimmed Milk Powder	Marvel
Methanol	Fisher Scientific (M/4056/17)
Microbank™ vials	Pro-Lab Diagnostics (PL.170)
Microscope slide Super Premium plain extra white glass 76mm x 26mm x 1.0 to 1.2mm Thermo Scientific Gerhard Menzel	12332098 VWR International Ltd.
Microtube 2ml PP	Sarstedt (72.694.006)
Monoclonal ANTI-FLAG M2-peroxidase	Sigma-Aldrich (A8592)
MOPS (3-(N-morpholino)propanesulfonic acid)	Alfa Aesar (A12914)
Mutanolysin from Streptomyces globisporus ATCC 21553	M9901 SIGMA
NEBuilder® HiFi DNA Assembly Master Mix	NEB (E2621S)
Nikon Model Eclipse CI-L	Nikon
Novex® 4-12% Tris-Glycine Mini Protein Gels	Life Technologies (EC6035)
Novex® Tris-Glycine Native Running Buffer	Life Technologies (LC2672)
NuPAGE® LDS Sample Buffer (4X)	Life Technologies (NP0008)
NuPAGE® MES SDS running buffer (20x)	Life Technologies (NP0002)
PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa	Thermo 26619
Phosphate-buffered saline tablets	Oxoid (BR0014G)

Phusion® HF Buffer	New England Biolabs(B0518)
Phusion® polymerase	New England Biolabs (M0530)
Potassium phosphate dibasic	Sigma-Aldrich (P3786)
Potassium phosphate monobasic	Sigma-Aldrich (P0662)
Proteinase inhibitor cocktail VII	Calbiochem (539138)
Q5® Site-Directed Mutagenesis Kit	NEB (E0554S)
QIAprep Spin Miniprep Kit	Qiagen (27106)
Restriction enzymes	New England Biolabs, Fermentas
rLysozyme™ Solution	Merck Millipore (71110)
Sodium azide	Sigma-Aldrich (71290)
SODIUM BORATE, 0.5M BUFFER SOLN., PH 8.0	J60068.AK VWR International Ltd.
Sodium borohydride	452882 ALDRICH
Sodium chloride	Sigma-Aldrich (S7653)
Sodium hydroxide	VWR Chemicals (28244-495)
Sucrose	Sigma-Aldrich (S7903)
Sulfuric acid	Sigma-Aldrich (320521)
SYBR® Safe	Life Technologies (S33102)
T4 DNA Ligase Reaction Buffer	New England Biolabs (B0202)
T4 ligase	New England Biolabs (M0202)
Taq polymerase	New England Biolabs (M0267)
TCA (trichloroacetic acid)	Sigma-Aldrich (T6399)

Tetracycline	Sigma-Aldrich (T3258)
Thiamine HCl	Sigma-Aldrich (T4625)
Thiamphenicol	Sigma-Aldrich (T0261)
TMB (3,3',5,5'-Tetramethylbenzidine)	Sigma-Aldrich (T0565)
Trans-Blot® Turbo™ Mini PVDF Transfer Packs	BioRad (170-4156)
Tris base	Calbiochem (648310)
Trypsin from bovine pancreas	T1005 SIGMA
Tryptone	Bacto (211705)
Tween® 20	Fluka (P5927)
Uracil	Sigma-Aldrich (U1128)
Wheat straw (ball-milled)	Prof. Greg Tucker, University of Nottingham
X-Cell II™ Blot Module	Life Technologies (EI9051)
X-Cell SureLock™ MiniCell	Life Technologies (EI0001)
Xylan from beechwood	Sigma-Aldrich (X4252)
Xylose	Sigma-Aldrich (X1500)
Yeast extract	Oxoid (LP0021)
Zeocin	Life Technologies (ANT-ZN-1)
Zinc sulfate heptahydrate	Alfa Aesar (Z4750)

Appendix 2

Table of *C. acetobutylicum* strains used in this study

Strain Name	Description	Source
Wild-type	WT	ATCC culture
Li-tcdR	Δ LacG pyrE::Li-tcdR	Dr. Hengzhang Wang
ΔpyrE	Δ pyrE	Dr. Muhammad Ehsaan
F_Cel48F	pyrE::Facoid_Cel48F_N	Dr. Benjamin Willson
Li-tcdR Cel48F-CelS	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_Cel48F_CelS_Flag	Transformation of pMTL82151_PtcdB_Cel48F_CelS_Flag into Δ lacG pyrE::Li-tcdR
Li-tcdR Cel9E	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_Cel9E_CelE_Flag	Transformation of pMTL82151_PtcdB_Cel9E_CelE_Flag into Δ lacG pyrE::Li-tcdR
Li-tcdR Cel5A	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_Cel5A_CelA_Flag	Transformation of pMTL82151_PtcdB_Cel5A_CelA_Flag into Δ lacG pyrE::Li-tcdR
Integrated Li-tcdR Cel48F_N	Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_N_Flag	Integration of pMTL-JH16-PtcdB_Cel48F_N into Δ LacG pyrE::Li-tcdR
Integrated Li-tcdR Cel48F_N D	Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND_Flag	Integration of pMTL-JH16-PtcdB_Cel48F_ND into Δ LacG pyrE::Li-tcdR
Integrated Li-tcdR Cel48F_CelS	Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_CelS_Flag	Integration of pMTL-JH16-PtcdB_Cel48F_CelS into Δ LacG pyrE::Li-tcdR
Li-tcdR ExgS	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_ExgS_Flag	Transformation of pMTL82151_PtcdB_ExgS_Flag into Δ lacG pyrE::Li-tcdR
Li-tcdR C748	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_C748_Flag	Transformation of pMTL82151_PtcdB_C748_Flag into Δ lacG pyrE::Li-tcdR
Li-tcdR Cphy68	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_Cphy68_Flag	Transformation of pMTL82151_PtcdB_Cphy68_Flag into Δ lacG pyrE::Li-tcdR
PtcdB_SlpA	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_SlpA_Cel48F_ND_Flag	Transformation of pMTL82151_PtcdB_SlpA_Cel48F_ND_Flag into Δ lacG pyrE::Li-tcdR

Cel48F_N D		
PtcdB CoseGsiB Cel48F_N D	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_CoseGsiB_Cel48F_ND_Flag	Transformation of pMTL82151_PtcdB_CoseGsiB_Cel48F_ND_Flag into Δ LacG pyrE::Li-tcdR
PtcdB CoseThio Cel48F_N D	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB_CoseThio_Cel48F_ND_Flag	Transformation of pMTL82151_PtcdB_CoseThio_Cel48F_ND_Flag into Δ LacG pyrE::Li-tcdR
PtcdB(N) Cel48F_N D	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB(N)_Cel48F_ND_Flag	Transformation of pMTL82151_PtcdB(N)_Cel48F_ND_Flag into Δ LacG pyrE::Li-tcdR
PtcdB CoseThio Cel48F_N D SL2	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB(Cose_Thio)_Cel48F_ND_SL2	Transformation of pMTL82151_PtcdB_CoseThio_Cel48F_ND_SL2 into Δ LacG pyrE::Li-tcdR
PtcdB CoseThio Cel48F_N D SlpA_T	Δ LacG pyrE::Li-tcdR and pMTL82151_PtcdB(Cose_Thio)_Cel48F_ND_SlpA_T	Transformation of pMTL82151_PtcdB_CoseThio_Cel48F_ND_SlpA_T into Δ LacG pyrE::Li-tcdR
Cel48F + PrsA_CAC	Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND and pMTL82151_TcpF_PrsA_CAC	Transformation of pMTL82151_TcpF_PrsA_CAC into Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND_Flag
Cel48F + PrsA_BAC	Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND and pMTL82151_TcpF_PrsA_Bac	Transformation of pMTL82151_TcpF_PrsA_BAC into Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND_Flag
Cel48F + PrsA_Ccel 0751	Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND and pMTL82151_TcpF_PrsA_Ccel0751	Transformation of pMTL82151_TcpF_PrsA_Ccel0751 into Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND_Flag
Cel48F + PrsA_Ccel 3056	Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND and pMTL82151_TcpF_PrsA_Ccel3056	Transformation of pMTL82151_TcpF_PrsA_Ccel3056 into Δ LacG pyrE::Li-tcdR thl::T1T2_PtcdB_Cel48F_ND_Flag
ΔpyrE Δ2316	Δ pyrE Δ 2316	Deletion of CAC_2316 using pMTL-ME3_2316 in Δ pyrE
ΔpyrE Δ3377	Δ pyrE Δ 3377	Deletion of CAC_3377 using pMTL-ME3_3377 in Δ pyrE
Δ2316 Cel48F	pyrE::Facoid_Cel48F_N Δ pyrE Δ 2316	Integration of pMTL-JH14-Facoid_Cel48F(N) into Δ pyrE Δ 2316
Δ3377 Cel48F	pyrE::Facoid_Cel48F_N Δ 3377	Integration of pMTL-JH14-Facoid_Cel48F(N) into Δ pyrE Δ 3377
Δ2316	Δ 2316	Integration of pMTL-ME6 into Δ pyrE Δ 2316

Δ3377	Δ3377	Integration of pMTL-ME6 into ΔpyrE Δ3377
Cel48A	pyrE::Facoid_Cel48A	Integration of pMTL-JH14-Cel48A into ΔpyrE
Cel48F-E55Q	pyrE::Facoid_Cel48F_E55Q	Integration of pMTL-JH14-Cel48F-E55Q into ΔpyrE
Cel48F + Ccel OrfXp	pyrE::Facoid_Cel48F_N and pMTL-82151_TcpF_OrfX	Transformation of pMTL-82151-TcpF-OrfX-Cell into pyrE::Facoid_Cel48F_N
Cel48F-Chimera	pyrE::Facoid_Cel48F_Chimera	Integration of pMTL-JH14-Cel48F-chimera into ΔpyrE

Appendix 3.1

Table showing the primers used in this study

Primer Name	Primer sequence
Cel48A	
pMTL_JH14-lacoid_Cel48F FWD	ttactaaattaattgcaaaagattacaaggatgacgacga
pMTL_JH14-lacoid_Cel48F_REV	ttcttacttatctttaacatatgaaatacacctccttaaaatttaatacataagttg
NC_003030_Cac_ATCC_824_reseq_ref_V4 FWD	tttaaggaggtgtatttcatatgttaaagataagtaagaattttaaaaaataatggc
NC_003030_Cac_ATCC_824_reseq_ref_V4 REV	tcgtcgtcatccttgtaatcttttgcaattaatttagtaagttcca
pET48Fc	
16_fw	CATGGTATATCTCCTTCTTAAAG
16_rev	caccaccaccaccactgagCATATGCTCGAGGATCCG
48_Fw	Gccggatcctcgagcatatgctcagtggtggtggtggtggtg TTGGATAGAAAGAAGTGCTTTC
48_rev	taagaaggagatataccatgGCTTCAAGTCCTGCAAC
Orfx	
orfX_Cell_Fw	ctgaataggaggttagttcaATGAAAAAAGTATTATTGGCAATATTAATAATTG
orfX_Cell_rev	gcaggcttctatttttatgTTATTTAACTGTTATCTCACCC
Cel48F- E55Q	
E55Q_fw	GCTTCAAGTCCTGCAAAC

E55Q_rev	CATATGAACTAACCTCCTAAATTTTG
Δ3377	
3377_LHA_Fw	tccatatgaccatgattacgTACTGATATAATAAC TCATACTATGTCTC
3377_LHA_Rev	ctttatataattaCAAAAGATTCACCTTCTAATT GTATTTTATAAG
3377_RHA_Fw	gaatcttttgTAATTATATAAAGAAAAAGCAA GAGAAC
3377_RHA_Rev	gcaggcttcttatttttatgTAATACAATCAAAAT GCGGC
Δ2316	
2316_LHA_Fw	tccatatgaccatgattacgTTTAGCTGGTCTTGG AAATTC
2316_LHA_Rev	tcactaaatgTAATATGGAGGGTTTAGCATAT G
2316_RHA_Fw	ctccatattaCATTTAGTGAAGTGTCACTAAA C
2316_RHA_Rev	gcaggcttcttatttttatgATGGAAAATATCACT ATGTGGTTAC
48F - 48A Chimera	
48A_BB_Fw	AATTATGGAACCTTTATTCGCTAATAC
48A_BB_Rev	TGCAACTTTATTAGTTAATTAATCC
48Fcat_Fw	TATTAACATAATAAAGTTGCAGCTTCAAGTC CTGCAAA
48Fcat_Rev	GCGAATAAAGTTCCATAATTACCTTGATCT GGGAAGAG
OatB	
OatB_Fw	TTAGTTCatATGCGGATTAAATGGTTTAGT TTA
OatB_Rev	gccgcgctagcTTAGGCTGGTTTAGCCGGAGT C
PrsA- C. acetobutylicum	
PrsA_CAC_Fw	TAGTTAGGAGGTTAGTTCATATGGTGAAA TCTAAACTAAATACCT
PrsA_CAC_Rev	ATGAAAAAACCTTATGTAAGCTAGCGCG GCCGCCTGCAG
PrsA_B.subtilis	
PrsA_BAC_Fw	gttaggaggttagttcatATGAAGAAAATCGCA ATAGC
PrsA_BAC_Rev	gctgcagcgccgcgctagcTTATTTAGAATTGC TTGAAGATGAAG

Appendix 3.2

Table showing the sequencing primers used in this study

Sequencing primer Name	Primer sequence
M13R – For forward sequencing pMTL-x plasmids	AACAGCTATGACCATG
Rev_8 – For reverse sequencing pMTL-x plasmids	atccagggtgctatcttcgt
Rev_3 – for sequencing pj20x plasmids	GAATAAATTTTGTGTCGCCC
Fw_5 – for sequencing pj20x plasmids	GACTGGGCCTTTCGCCCCGGGGTA
JH14_Fw - for sequencing pMTL-JH14 plasmids	CATGTTCTATTGTTACTATGGAAGA GG
JH14_Rev - for sequencing pMTL-JH14 plasmids	GAAAACCAGGTGAAGGACTTGC

Appendix 4

Table of plasmids used in this study

Plasmid Number	Description	Source
1	pJ206_BB2Cel48F_CelS_Flag	CRG Culture collection
2	pJ206_BB2Cel9E_CelE_Flag	CRG Culture collection
3	pJ206_BB2Cel5A_CelA_Flag	CRG Culture collection
4	PJ204_BB2_PtcdB	CRG Culture collection
5	pMTL82151	CRG Culture collection
6	pMTL82151_PtcdB_Cel48F_CelS_Flag	This study
7	pMTL82151_PtcdB_Cel9E_CelE_Flag	This study
8	pMTL82151_PtcdB_Cel5A_CelA_Flag	This study
9	pJ204_BB2_Cel48F_ND	This study
10	pJ204_BB2_2xStop	CRG Culture collection
11	pJ202_BB2_Flag_Tag	CRG Culture collection
12	pJ202_BB2_Cel48F_ND_F	This study
13	pMTL82151_PtcdB_Cel48F_ND_Flag	This study
14	pJ204_BB2_Cel48F_N_Flag	This study
15	pMTL82151_PtcdB_Cel48F_N_Flag	This study
16	pJ204_BB2_T1T2	CRG Culture collection
17	pJ204_BB2_T1T2_PtcdB	This study
18	pJ204_BB2_T1T2_PtcdB_Cel48F_N_Flag	This study

19	pMTL-JH16	CRG Culture collection
20	pMTL-JH16_T1T2_PtcdB_Cel48F_N_Flag	This study
21	pMTL-JH16_T1T2_PtcdB_Cel48F_ND_Flag	This study
22	pMTL-JH16_T1T2_PtcdB_Cel48F_CelS_Flag	This study
	GH48's	
23	pBMH_BB2_ExgS_Flag	Biomatik
24	pBMH_BB2_C748_Flag	Biomatik
25	pBMH_BB2_Cphy68_Flag	Biomatik
26	pMTL82151_PtcdB_ExgS_Flag	This study
27	pMTL82151_PtcdB_C748_Flag	This study
28	pMTL82151_PtcdB_Cphy68_Flag	This study
29	pMTL82151_PtcdB_Cphy67_Flag	This study
	5'UTRS	
30	pBMH_BB2_PtcdB_SlpA	Biomatik
31	pBMH_BB2_PtcdB_CoseGsib	Biomatik
32	pBMH_BB2_PtcdB_CoseThio	Biomatik
33	pJ204_BB2_PtcdB_SlpA_Cel48F_ND_Flag	This study
34	pJ204_BB2_PtcdB_CoseGsib_Cel48F_ND_Flag	This study
35	pJ204_BB2_PtcdB_CoseThio_Cel48F_ND_Flag	This study
36	pMTL82151_PtcdB_SlpA_Cel48F_ND_Flag	This study
37	pMTL82151_PtcdB_CoseGsib_Cel48F_ND_Flag	This study
38	pMTL82151_PtcdB_CoseThio_Cel48F_ND_Flag	This study
39	pMTL82151_PtcdB(N)_Cel48F_ND_Flag	This study
	3'UTRS	
40	pBMH-BB2_SL2	This study
41	pBMH-BB2_SL4	This study
42	pJ202-BB2_SlpA(term)	CRG Culture Collection
43	pJ202_BB2_Cel48F_ND_SL2	This study
44	pJ202_BB2_Cel48F_ND_SL4	This study
45	pJ202_BB2_Cel48F_ND_SlpA_T	This study
46	pMTL82151_PtcdB(Coe_Thio)_Cel48F_ND_SL2	This study
47	pMTL82151_PtcdB(Coe_Thio)_Cel48F_ND_SL4	This study
48	pMTL82151_PtcdB(Coe_Thio)_Cel48F_ND_SlpA_T	This study
	PrsA overexpression	
49	pJ202_BB2_PrsA_CAC	This study
50	pJ202_BB2_PrsA_bac	This Study

51	pMK_RQ_BB2_Ccel_0751	Biomatik
52	pMK_RQ_BB2_Ccel_3056	Biomatik
53	pMTL-82151-TcpF	This study
54	pMTL82151_TcpF_PrsA_CAC	This study
55	pMTL82151_TcpF_PrsA_Bac	This study
56	pMTL82151_TcpF_PrsA_Ccel0751	This study
57	pMTL82151_TcpF_PrsA_Ccel3056	This study
	N-deacetylase	
58	pMTL-ME3	CRG Culture collection
59	pMTL-ME3-2316	This study
60	pMTL-ME3-2316	This study
62	pMTL-JH14-FacoiD-Cel48F_N	CRG Culture Collection
63	pMTL-JH14-Cel48A_Flag	This study
64	pMTL-JH14-Cel48F_N-E55Q	This study
	OrfX	
65	pMTL-82151-TcpF-OrfX-Cell	
66	pMTL-JH14-Cel48F-Chimera	This study

Appendix 5

Predicted and measured molecular masses of dimeric mucopeptides from wild-type C. acetobutylicum and their elution times. tri: GlcNAc-MurNAc with a tripeptide stem peptide, tetra: GlcNAc-MurNAc with a tetrapeptide stem peptide; A: Alanine; E: Glutamic acid; m: m-DAP.

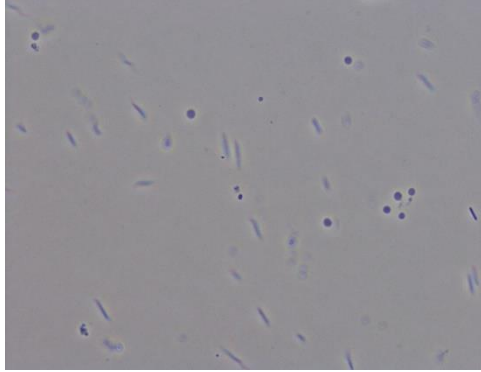
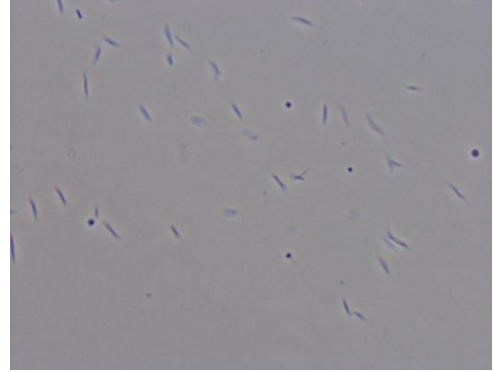
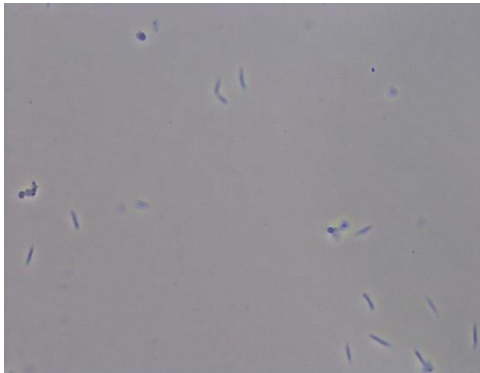
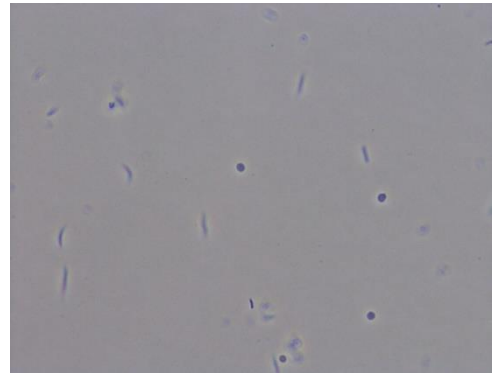
Dimers					
tri-tetra (Aem, AEmA)					
Two deacetylated					
Muropeptide	Neutral mass	Charge	Predicted M/Z	Measured M/Z	Time eluted
GlcNAc-MurNAc, GlcNAc-MurNAc	1708.767	2	855.3835	855.43	1.51 and 1.81
		3	570.589	570.29	1.53
MurNAc-GlcNAc, MurNAc	1547.698	2	774.8491	774.39	1.51 and 1.97
MurNAc, MurNAc	1505.718	2	753.8591	753.35	1.51
No deacetylation					

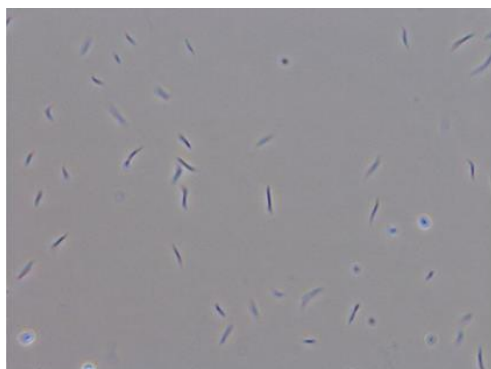
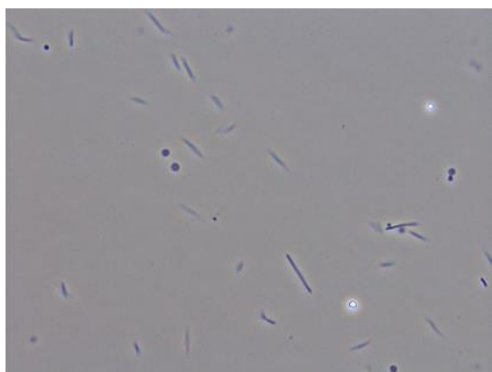
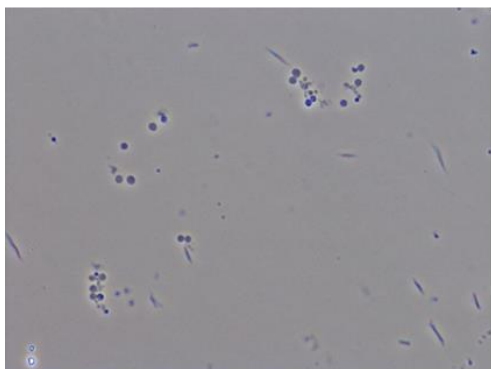
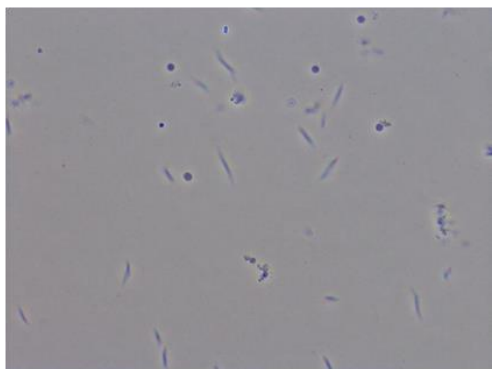
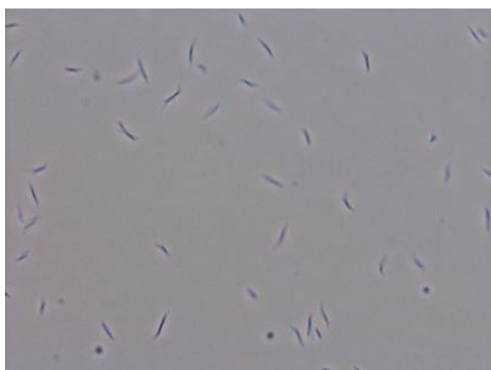
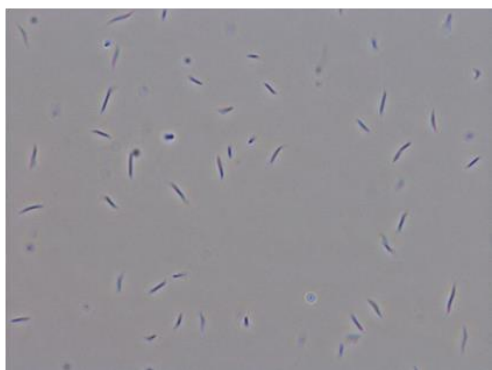
GlcNaC-MurNAc, GlcNaC-MurNAc	1792.757	2		897.3785	897.44	2.58
		3		598.5857	598.28	2.636
GlcNAc-MurNAc, MurNAc	1631.688	2		816.8441	816.41	
MurNAc, MurNAc	1470.619	2		736.3097	555.31	1.972
One deacetylation						
GlcNaC-MurNAc, GlcNAc-MurNAc	1750.777	2		876.3885	876.41	1.95, 2.36
		3		584.5923	584.3	1.95
GlcNAc-MurNAc, MurNAc	1589.708	2		795.8541	795.38	2.54
GlcNAc-MurNAc, MurNAc	1270.582	2		636.291	636.3	1.95
			3	424.5273	424.73	1.972
tetra-tetra (AemA-AEmA)						
Fully deacetylated						
GlcNAc-MurNAc, GlcNAc-MurNAc	1779.804	2		890.902	890.44	1.86 and 2.67
		3		594.268	593.97	1.86
MurNAc- GlcNAc, MurNAc	1618.735	2		810.3676	809.89	1.86 and 2.87
		3		540.5784	540.25	1.86
No deacetylation						
GlcNaC-MurNAc, GlcNaC-MurNAc	1863.764	2		932.882	932.45	3.7
		3		622.2547	NF	
GlcNAc-MurNAc, MurNAc	1702.695	2		852.3476	NF	
MurNAc, MurNAc	1541.626	2		771.8132	NF	
One deacetylation						
GlcNaC-MurNAc, GlcNAc-MurNAc	1821.784	2		911.892	911.95	2.67
		3		608.2613	608.31	2.67
GlcNAc-MurNAc, MurNAc	1660.715	2		831.3576	830.91	3.72
		3		554.5717	NF	
MurNAc, MurNAc	1499.646	2		750.8232	750.38	?
tri-tri (Aem, Aem)						
Fully deacetylated						
GlcN-MurNAc, GlcN- MurNAc	1637.729	2		819.8645	819.42	1.44
		3		546.9097	546.6	1.44
MurNAc- GlcNAc, MurNAc	1476.66	2		739.3301	738.88	1.42

		3	493.2201	NF	
MurNAc, MurNAc	1315.591	2	658.7957	nf	
		3	439.5305	nf	
One deacetylation					
GlcNAc-MurNAc GlcNAc-MurNAc	1821.814	2	911.907	911.44	2.63
		3	608.2713	607.97	2.63
GlcNAc-MurNAc, MurNAc	1660.745	2	831.3726	830.91	3.72
MurNAc-MurNAc	1499.676	2	750.8382	nf	
No deacetylation					
GlcNAc-MurNAc, GlcNAc-MurNAc	1721.689	2	861.8445	861.43	2.1
		3	574.8963	nf	
GlcNAc-MurNAc, MurNAc	1560.62	2	781.3101	nf	
MurNAc, MurNAc	1399.551	2	700.7757	nf	

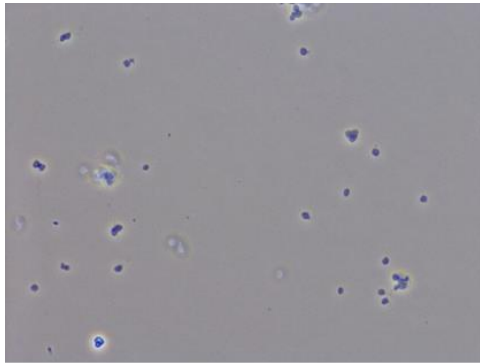
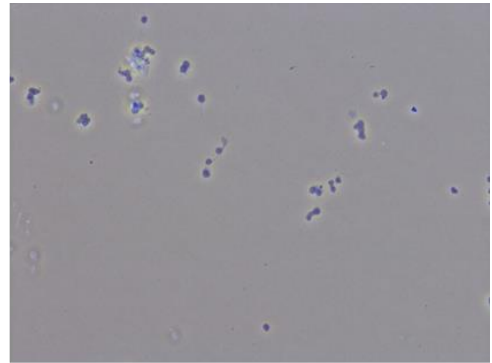
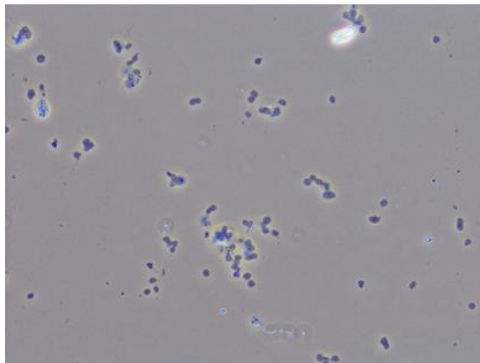
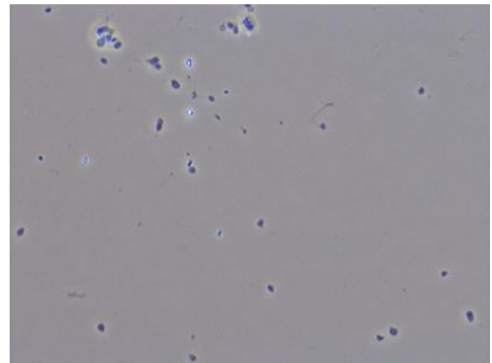
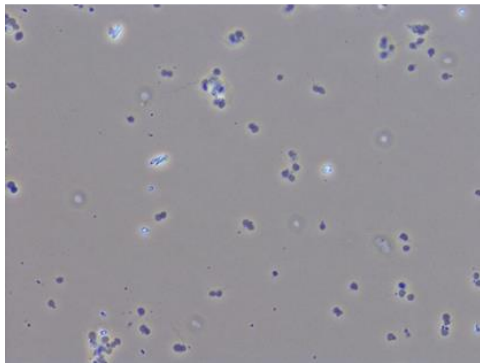
Appendix 6

Wild-type

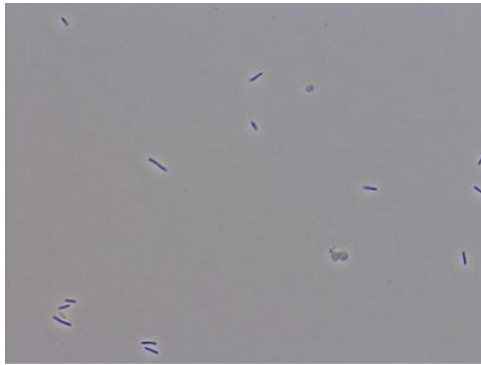
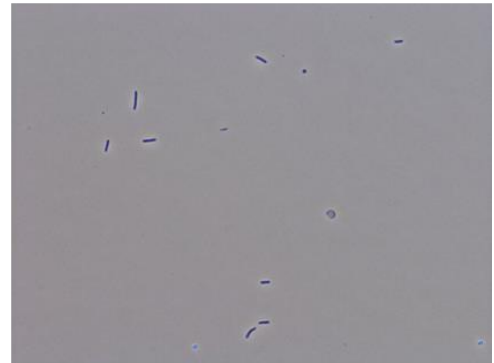
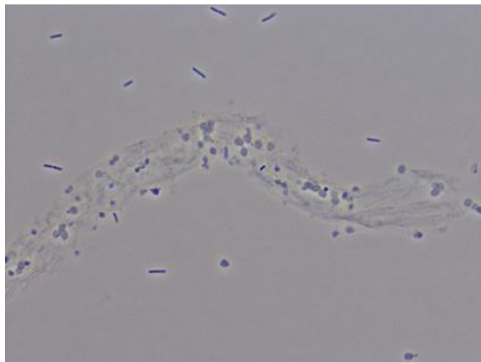


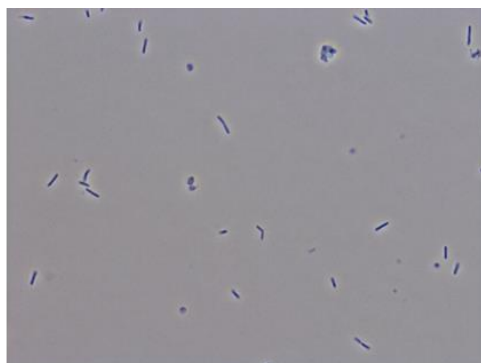
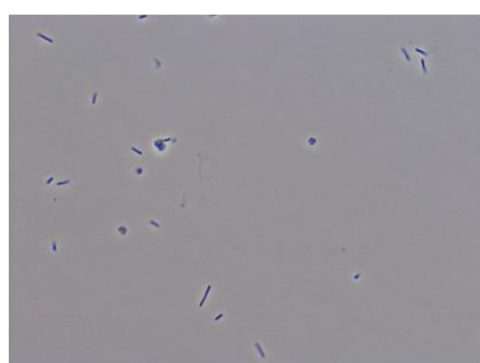


Wild-type plus lysozyme

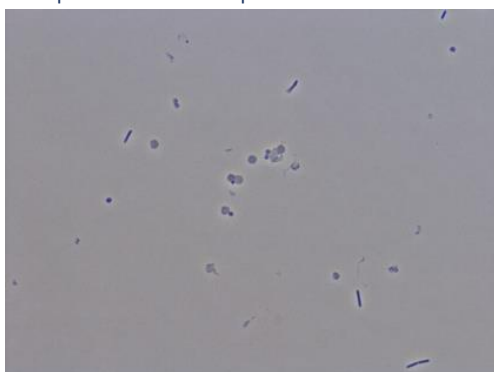
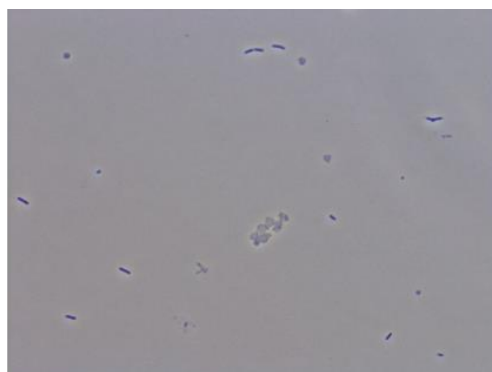
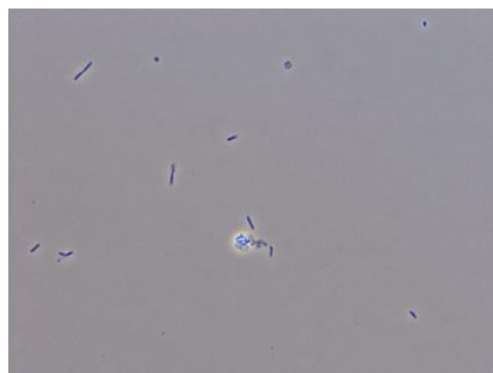


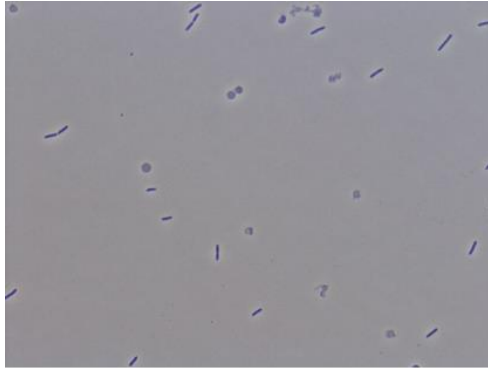
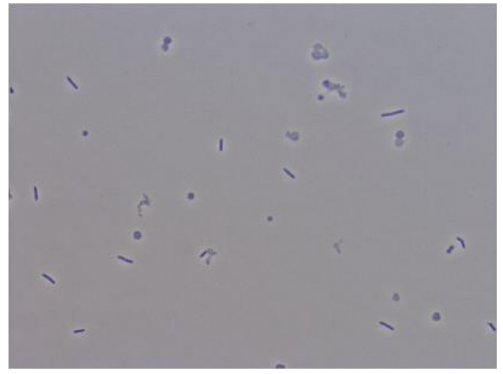
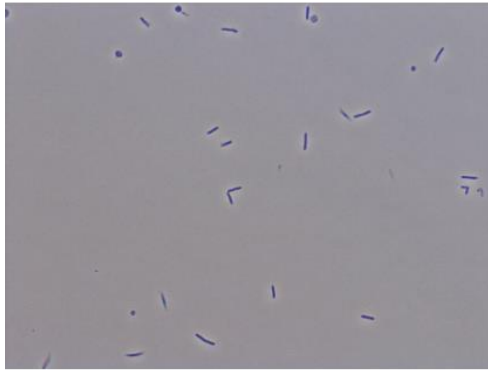
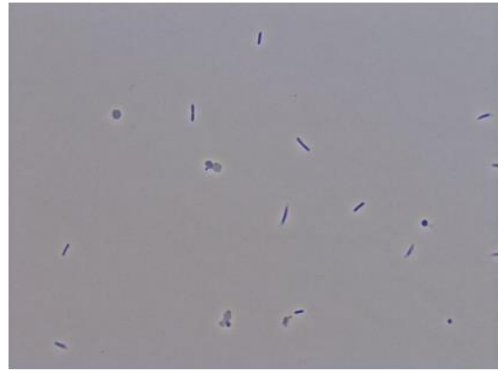
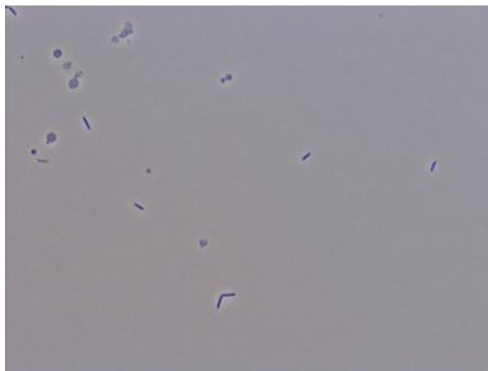
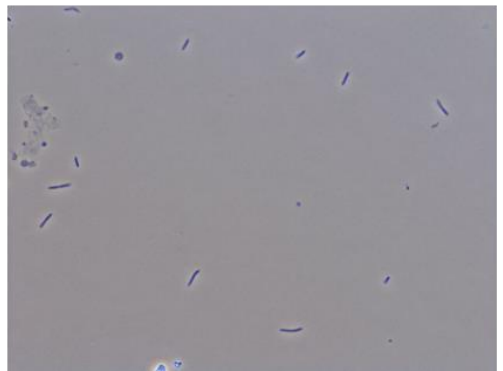
WT plus wild-type supernatant





WT plus Cel48F supernatant





References

- Adsul, M. G., M. S. Singhvi, S. A. Gaikawai, and D. V. Gokhale. 2011. 'Development of Biocatalysts for Production of Commodity Chemicals from Lignocellulosic Biomass'. *Bioresource Technology* 102 (6):4304–12. <https://doi.org/10.1016/j.biortech.2011.01.002>.
- Agger, Jane W., Trine Isaksen, Anikó Várnai, Silvia Vidal-Melgosa, William G. T. Willats, Roland Ludwig, Svein J. Horn, Vincent G. H. Eijsink, and Bjørge Westereng. 2014. 'Discovery of LPMO Activity on Hemicelluloses Shows the Importance of Oxidative Processes in Plant Cell Wall Degradation'. *Proceedings of the National Academy of Sciences of the United States of America* 111 (17):6287–92. <https://doi.org/10.1073/pnas.1323629111>.
- Alcorlo, Martín, Siseth Martínez-Caballero, Rafael Molina, and Juan A Hermoso. 2017. 'Carbohydrate Recognition and Lysis by Bacterial Peptidoglycan Hydrolases'. *Current Opinion in Structural Biology, Carbohydrates • Sequences and topology*, 44 (June):87–100. <https://doi.org/10.1016/j.sbi.2017.01.001>.
- Al-Hinai, Mohab A., Alan G. Fast, and Eleftherios T. Papoutsakis. 2012. 'Novel System for Efficient Isolation of Clostridium Double-Crossover Allelic Exchange Mutants Enabling Markerless Chromosomal Gene Deletions and DNA Integration'. *Applied and Environmental Microbiology* 78 (22):8112–21. <https://doi.org/10.1128/AEM.02214-12>.
- Allcock, Errol R., Sharon J. Reid, David T. Jones, and David R. Woods. 1981. 'Autolytic Activity and an Autolysis-Deficient Mutant of Clostridium Acetobutylicum'. *Applied and Environmental Microbiology* 42 (6):929–35.
- Allcock, Errol R., Sharon J. Reid, David T. Jones, and David R. Woods. 1982. 'Clostridium Acetobutylicum Protoplast Formation and Regeneration'. *Applied and Environmental Microbiology* 43 (3):719–21.
- Anderson, Timothy D., J. Izaak Miller, Henri-Pierre Fierobe, and Robert T. Clubb. 2013. 'Recombinant Bacillus Subtilis That Grows on Untreated Plant Biomass'. *Applied and Environmental Microbiology* 79 (3):867–76. <https://doi.org/10.1128/AEM.02433-12>.
- Anderson, Timothy D., Scott A. Robson, Xiao Wen Jiang, G. Reza Malmirchegini, Henri-Pierre Fierobe, Beth A. Lazazzera, and Robert T. Clubb. 2011. 'Assembly of Minicellulosomes on the Surface of Bacillus Subtilis'. *Applied and Environmental Microbiology* 77 (14):4849–58. <https://doi.org/10.1128/AEM.02599-10>.
- Antelmann, Haike, Elise Darmon, David Noone, Jan-Willem Veening, Helga Westers, Sierd Bron, Oscar P. Kuipers, Kevin M. Devine, Michael Hecker, and Jan Maarten Van Dijl. 2003. 'The Extracellular Proteome of Bacillus Subtilis under Secretion Stress Conditions'. *Molecular Microbiology* 49 (1):143–56. <https://doi.org/10.1046/j.1365-2958.2003.03565.x>.
- Artzi, Lior, Edward A. Bayer, and Sarah Morais. 2017. 'Cellulosomes: Bacterial Nanomachines for Dismantling Plant Polysaccharides'. *Nature Reviews Microbiology* 15 (2):83–95. <https://doi.org/10.1038/nrmicro.2016.164>.

- Atilano, Magda L., James Yates, Marcus Glittenberg, Sergio R. Filipe, and Petros Ligoxygakis. 2011. 'Wall Teichoic Acids of *Staphylococcus Aureus* Limit Recognition by the *Drosophila* Peptidoglycan Recognition Protein-SA to Promote Pathogenicity'. *PLOS Pathogens* 7 (12):e1002421. <https://doi.org/10.1371/journal.ppat.1002421>.
- Atrih, Abdelmadjid, Gerold Bacher, Günter Allmaier, Michael P. Williamson, and Simon J. Foster. 1999. 'Analysis of Peptidoglycan Structure from Vegetative Cells of *Bacillus Subtilis* 168 and Role of PBP 5 in Peptidoglycan Maturation'. *Journal of Bacteriology* 181 (13):3956–66.
- Badalato, Nelly, Alain Guillot, Victor Sabarly, Marc Dubois, Nina Pourette, Bruno Pontoire, Paul Robert, et al. 2017. 'Whole Proteome Analyses on *Ruminiclostridium Cellulolyticum* Show a Modulation of the Cellulolysis Machinery in Response to Cellulosic Materials with Subtle Differences in Chemical and Structural Properties'. *PLOS ONE* 12 (1):e0170524. <https://doi.org/10.1371/journal.pone.0170524>.
- Baral, Nawa R., Lauren Slutzky, Ajay Shah, Thaddeus C. Ezeji, Katrina Cornish, and Ann Christy. 2016. 'Acetone-Butanol-Ethanol Fermentation of Corn Stover: Current Production Methods, Economic Viability and Commercial Use'. *FEMS Microbiology Letters* 363 (6). <https://doi.org/10.1093/femsle/fnw033>.
- Bar-Peled, Maor, and Malcolm A. O'Neill. 2011. 'Plant Nucleotide Sugar Formation, Interconversion, and Salvage by Sugar Recycling*'. *Annual Review of Plant Biology* 62 (1):127–55. <https://doi.org/10.1146/annurev-arplant-042110-103918>.
- Barr, B. K., Y. L. Hsieh, B. Ganem, and D. B. Wilson. 1996. 'Identification of Two Functionally Different Classes of Exocellulases'. *Biochemistry* 35 (2):586–92. <https://doi.org/10.1021/bi9520388>.
- Barra-Carrasco, Jonathan, and Daniel Paredes-Sabja. 2014. 'Clostridium Difficile Spores: A Major Threat to the Hospital Environment'. *Future Microbiology* 9 (4):475–86. <https://doi.org/10.2217/fmb.14.2>.
- Bayer, E A, and R Lamed. 1986. 'Ultrastructure of the Cell Surface Cellulosome of *Clostridium Thermocellum* and Its Interaction with Cellulose.' *Journal of Bacteriology* 167 (3):828–36.
- Bayer, Edward A., Jean-Pierre Belaich, Yuval Shoham, and Raphael Lamed. 2004. 'The Cellulosomes: Multienzyme Machines for Degradation of Plant Cell Wall Polysaccharides'. Review-article. <http://Dx.doi.org/10.1146/annurev.micro.57.030502.091022>. 13 September 2004. <http://www.annualreviews.org/doi/10.1146/annurev.micro.57.030502.091022>.
- Beeby, Morgan, James C. Gumbart, Benoît Roux, and Grant J. Jensen. 2013. 'Architecture and Assembly of the Gram-Positive Cell Wall'. *Molecular Microbiology* 88 (4):664–72. <https://doi.org/10.1111/mmi.12203>.
- Ben Willson. 2014. 'Genetic Modification of *Clostridium Acetobutylicum* for Deconstruction of Lignocellulose'. University of Nottingham.
- Bera, Agnieszka, Raja Biswas, Silvia Herbert, and Friedrich Götz. 2006. 'The Presence of Peptidoglycan O-Acetyltransferase in Various Staphylococcal Species Correlates with Lysozyme Resistance and

- Pathogenicity'. *Infection and Immunity* 74 (8):4598–4604.
<https://doi.org/10.1128/IAI.00301-06>.
- Bern, Marshall, Richard Beniston, and Stéphane Mesnage. 2017. 'Towards an Automated Analysis of Bacterial Peptidoglycan Structure'. *Analytical and Bioanalytical Chemistry* 409 (2):551–60.
<https://doi.org/10.1007/s00216-016-9857-5>.
- Bernard, Elvis, Thomas Rolain, Pascal Courtin, Alain Guillot, Philippe Langella, Pascal Hols, and Marie-Pierre Chapot-Chartier. 2011. 'Characterization of O-Acetylation of N-Acetylglucosamine'. *The Journal of Biological Chemistry* 286 (27):23950–58.
<https://doi.org/10.1074/jbc.M111.241414>.
- Bilej, Martin. 2015. 'Chapter 9 - Mucosal Immunity in Invertebrates'. In *Mucosal Immunology (Fourth Edition)*, edited by Jiri Mestecky, Warren Strober, Michael W. Russell, Brian L. Kelsall, Hilde Cheroutre, and Bart N. Lambrecht, 135–44. Boston: Academic Press.
<https://doi.org/10.1016/B978-0-12-415847-4.00009-4>.
- 'Bio-Based Chemicals | Renewable Chemicals'. n.d. Accessed 9 May 2017.
<http://www.greenbiologics.com/index.php>.
- 'Biofuels Production | Renewable Energy - 2015 in Review | Statistical Review of World Energy | Energy Economics | BP Global'. n.d. Bp.com. Accessed 27 January 2017.
<http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/renewable-energy/biofuels-production.html>.
- Blouzard, Jean-Charles, Pedro M. Coutinho, Henri-Pierre Fierobe, Bernard Henrissat, Sabrina Lignon, Chantal Tardif, Sandrine Pagès, and Pascale de Philip. 2010. 'Modulation of Cellulosome Composition in *Clostridium Cellulolyticum*: Adaptation to the Polysaccharide Environment Revealed by Proteomic and Carbohydrate-Active Enzyme Analyses'. *PROTEOMICS* 10 (3):541–54.
<https://doi.org/10.1002/pmic.200900311>.
- Boerjan, Wout, John Ralph, and Marie Baucher. 2003. 'Lignin Biosynthesis'. *Annual Review of Plant Biology* 54 (1):519–46.
<https://doi.org/10.1146/annurev.arplant.54.031902.134938>.
- Boraston, Alisdair B., David N. Bolam, Harry J. Gilbert, and Gideon J. Davies. 2004. 'Carbohydrate-Binding Modules: Fine-Tuning Polysaccharide Recognition'. *The Biochemical Journal* 382 (Pt 3):769–81.
<https://doi.org/10.1042/BJ20040892>.
- Borne, Romain, Edward A. Bayer, Sandrine Pagès, Stéphanie Perret, and Henri-Pierre Fierobe. 2013. 'Unraveling Enzyme Discrimination during Cellulosome Assembly Independent of Cohesin–dockerin Affinity'. *FEBS Journal* 280 (22):5764–79. <https://doi.org/10.1111/febs.12497>.
- 'Bp-Energy-Outlook-2017.pdf'. n.d. Accessed 27 January 2017.
<http://www.bp.com/content/dam/bp/pdf/energy-economics/energy-outlook-2017/bp-energy-outlook-2017.pdf>.
- Brás, Joana L. A., Benedita A. Pinheiro, Kate Cameron, Fiona Cuskin, Aldino Viegas, Shabir Najmudin, Pedro Bule, et al. 2016. 'Diverse Specificity of Cellulosome Attachment to the Bacterial Cell Surface'. *Scientific Reports* 6 (December):38292. <https://doi.org/10.1038/srep38292>.
- Brown, David M., Florence Goubet, Vicky W. Wong, Royston Goodacre, Elaine Stephens, Paul Dupree, and Simon R. Turner. 2007.

- ‘Comparison of Five Xylan Synthesis Mutants Reveals New Insight into the Mechanisms of Xylan Synthesis’. *The Plant Journal: For Cell and Molecular Biology* 52 (6):1154–68. <https://doi.org/10.1111/j.1365-313X.2007.03307.x>.
- Brown, R. Malcolm. 2004. ‘Cellulose Structure and Biosynthesis: What Is in Store for the 21st Century?’ *Journal of Polymer Science Part A: Polymer Chemistry* 42 (3):487–95. <https://doi.org/10.1002/pola.10877>.
- Bugaut, Anthony, and Shankar Balasubramanian. 2012. ‘5'-UTR RNA G-Quadruplexes: Translation Regulation and Targeting’. *Nucleic Acids Research* 40 (11):4727–41. <https://doi.org/10.1093/nar/gks068>.
- Bugg, T. D., G. D. Wright, S. Dutka-Malen, M. Arthur, P. Courvalin, and C. T. Walsh. 1991. ‘Molecular Basis for Vancomycin Resistance in *Enterococcus Faecium* BM4147: Biosynthesis of a Depsipeptide Peptidoglycan Precursor by Vancomycin Resistance Proteins VanH and VanA’. *Biochemistry* 30 (43):10408–15.
- Bunternngsook, Benjarat, Wuttichai Mhuantong, Verawat Champreda, Arinthip Thamchaipenet, and Lily Eurwilaichitr. 2014. ‘Identification of Novel Bacterial Expansins and Their Synergistic Actions on Cellulose Degradation’. *Bioresource Technology* 159 (May):64–71. <https://doi.org/10.1016/j.biortech.2014.02.004>.
- Caffall, Kerry Hosmer, and Debra Mohnen. 2009. ‘The Structure, Function, and Biosynthesis of Plant Cell Wall Pectic Polysaccharides’. *Carbohydrate Research*, Pectin: Structure and FunctionA special issue dedicated to Professor Walter Pilnik, 344 (14):1879–1900. <https://doi.org/10.1016/j.carres.2009.05.021>.
- Calvo-Flores, Francisco García, and José A. Dobado. 2010. ‘Lignin as Renewable Raw Material’. *ChemSusChem* 3 (11):1227–35. <https://doi.org/10.1002/cssc.201000157>.
- Cao, Haojie, Auke J. van Heel, Hifza Ahmed, Maarten Mols, and Oscar P. Kuipers. 2017. ‘Cell Surface Engineering of *Bacillus Subtilis* Improves Production Yields of Heterologously Expressed α -Amylases’. *Microbial Cell Factories* 16 (April):56. <https://doi.org/10.1186/s12934-017-0674-0>.
- Cato, E.P, W.L. George, and S.M. Finegold. 1986. *Bergey's Manual of Systematic Bacteriology*, Vol. 2.
- Chambert, R., and M. F. Petit-Glatron. 1988. ‘Secretion Mechanism of *Bacillus Subtilis* Levansucrase: Characterization of the Second Step’. *Journal of General Microbiology* 134 (5):1205–14. <https://doi.org/10.1099/00221287-134-5-1205>.
- Chanal, Angélique, Florence Mingardon, Marielle Bauzan, Chantal Tardif, and Henri-Pierre Fierobe. 2011. ‘Scaffoldin Modules Serving as “Cargo” Domains To Promote the Secretion of Heterologous Cellulosomal Cellulases by *Clostridium Acetobutylicum*’. *Applied and Environmental Microbiology* 77 (17):6277–80. <https://doi.org/10.1128/AEM.00758-11>.
- Chang, V. S., and M. T. Holtzapple. 2000. ‘Fundamental Factors Affecting Biomass Enzymatic Reactivity’. *Applied Biochemistry and Biotechnology* 84–86:5–37.

- Chen, Fang, and Richard A. Dixon. 2007. 'Lignin Modification Improves Fermentable Sugar Yields for Biofuel Production'. *Nature Biotechnology* 25 (7):759–61. <https://doi.org/10.1038/nbt1316>.
- Chen, Jingqi, Gang Fu, Yuanming Gai, Ping Zheng, Dawei Zhang, and Jianping Wen. 2015. 'Combinatorial Sec Pathway Analysis for Improved Heterologous Protein Secretion in *Bacillus Subtilis*: Identification of Bottlenecks by Systematic Gene Overexpression'. *Microbial Cell Factories* 14 (1):92. <https://doi.org/10.1186/s12934-015-0282-9>.
- Chen, Xiaowen, Joseph Shekiri, Rick Elander, and Melvin Tucker. 2012. 'Improved Xylan Hydrolysis of Corn Stover by Deacetylation with High Solids Dilute Acid Pretreatment'. *Industrial & Engineering Chemistry Research* 51 (1):70–76. <https://doi.org/10.1021/ie201493g>.
- Claassen, P. a. M., J. B. van Lier, A. M. Lopez Contreras, E. W. J. van Niel, L. Sijtsma, A. J. M. Stams, S. S. de Vries, and R. A. Weusthuis. 1999. 'Utilisation of Biomass for the Supply of Energy Carriers'. *Applied Microbiology and Biotechnology* 52 (6):741–55. <https://doi.org/10.1007/s002530051586>.
- Cocuron, Jean-Christophe, Olivier Lerouxel, Georgia Drakakaki, Ana P. Alonso, Aaron H. Liepman, Kenneth Keegstra, Natasha Raikhel, and Curtis G. Wilkerson. 2007. 'A Gene from the Cellulose Synthase-like C Family Encodes a β -1,4 Glucan Synthase'. *Proceedings of the National Academy of Sciences* 104 (20):8550–55. <https://doi.org/10.1073/pnas.0703133104>.
- Cornillot, E., R. V. Nair, E. T. Papoutsakis, and P. Soucaille. 1997. 'The Genes for Butanol and Acetone Formation in *Clostridium Acetobutylicum* ATCC 824 Reside on a Large Plasmid Whose Loss Leads to Degeneration of the Strain'. *Journal of Bacteriology* 179 (17):5442–47.
- Cosgrove, Daniel J. 2000. 'Loosening of Plant Cell Walls by Expansins'. *Nature* 407 (6802):321–26. <https://doi.org/10.1038/35030000>.
- Cosgrove, Daniel, and Michael Jarvis. 2012. 'Comparative Structure and Biomechanics of Plant Primary and Secondary Cell Walls'. *Frontiers in Plant Science* 3. <https://doi.org/10.3389/fpls.2012.00204>.
- Costerton, J W, J M Ingram, and K J Cheng. 1974. 'Structure and Function of the Cell Envelope of Gram-Negative Bacteria.' *Bacteriological Reviews* 38 (1):87–110.
- Croux, C., B. Canard, G. Goma, and P. Soucaille. 1992a. 'Purification and Characterization of an Extracellular Muramidase of *Clostridium Acetobutylicum* ATCC 824 That Acts on Non-N-Acetylated Peptidoglycan.' *Applied and Environmental Microbiology* 58 (4):1075–81.
- . 1992b. 'Autolysis of *Clostridium Acetobutylicum* ATCC 824'. *Journal of General Microbiology* 138 (5):861–69. <https://doi.org/10.1099/00221287-138-5-861>.
- Croux, Christian, Ngoc-Phuong-Thao Nguyen, Jieun Lee, Céline Raynaud, Florence Saint-Prix, Maria Gonzalez-Pajuelo, Isabelle Meynial-Salles, and Philippe Soucaille. 2016. 'Construction of a Restriction-Less, Marker-Less Mutant Useful for Functional Genomic and Metabolic Engineering of the Biofuel Producer *Clostridium Acetobutylicum*'.

- Biotechnology for Biofuels* 9:23. <https://doi.org/10.1186/s13068-016-0432-2>.
- Daniel, R. M., D. A. Cowan, H. W. Morgan, and M. P. Curran. 1982. 'A Correlation between Protein Thermostability and Resistance to Proteolysis'. *Biochemical Journal* 207 (3):641–44. <https://doi.org/10.1042/bj2070641>.
- Darmon, Elise, David Noone, Anne Masson, Sierd Bron, Oscar P. Kuipers, Kevin M. Devine, and Jan Maarten van Dijk. 2002. 'A Novel Class of Heat and Secretion Stress-Responsive Genes Is Controlled by the Autoregulated CsrRS Two-Component System of *Bacillus Subtilis*'. *Journal of Bacteriology* 184 (20):5661–71. <https://doi.org/10.1128/JB.184.20.5661-5671.2002>.
- Dassa, Bareket, Ilya Borovok, Raphael Lamed, Bernard Henrissat, Pedro Coutinho, Christopher L. Hemme, Yue Huang, Jizhong Zhou, and Edward A. Bayer. 2012. 'Genome-Wide Analysis of *Acetivibrio Cellulolyticus* Provides a Blueprint of an Elaborate Cellulosome System'. *BMC Genomics* 13:210. <https://doi.org/10.1186/1471-2164-13-210>.
- Davidi, Lital, Sarah Morais, Lior Artzi, Doriv Knop, Yitzhak Hadar, Yonathan Arfi, and Edward A. Bayer. 2016. 'Toward Combined Delignification and Saccharification of Wheat Straw by a Laccase-Containing Designer Cellulosome'. *Proceedings of the National Academy of Sciences* 113 (39):10854–59. <https://doi.org/10.1073/pnas.1608012113>.
- DeCicco, John M., Danielle Yuqiao Liu, Joonghyeok Heo, Rashmi Krishnan, Angelika Kurthen, and Louise Wang. 2016. 'Carbon Balance Effects of U.S. Biofuel Production and Use'. *Climatic Change* 138 (3–4):667–80. <https://doi.org/10.1007/s10584-016-1764-4>.
- DeMartini, Jaclyn D., Sivakumar Pattathil, Jeffrey S. Miller, Hongjia Li, Michael G. Hahn, and Charles E. Wyman. 2013. 'Investigating Plant Cell Wall Components That Affect Biomass Recalcitrance in Poplar and Switchgrass' 6 (3):898–909. <https://doi.org/10.1039/C3EE23801F>.
- Den Haan, Riaan, Shaunita H. Rose, Lee R. Lynd, and Willem H. van Zyl. 2007. 'Hydrolysis and Fermentation of Amorphous Cellulose by Recombinant *Saccharomyces Cerevisiae*'. *Metabolic Engineering* 9 (1):87–94. <https://doi.org/10.1016/j.ymben.2006.08.005>.
- Desvaux, Mickaël. 2005a. 'Clostridium Cellulolyticum: Model Organism of Mesophilic Cellulolytic Clostridia'. *FEMS Microbiology Reviews* 29 (4):741–64. <https://doi.org/10.1016/j.femsre.2004.11.003>.
- . 2005b. 'The Cellulosome of *Clostridium Cellulolyticum*'. *Enzyme and Microbial Technology* 37 (4):373–85. <https://doi.org/10.1016/j.enzmictec.2004.04.025>.
- Desvaux, Mickaël, Arshad Khan, Anthony Scott-Tucker, Roy R. Chaudhuri, Mark J. Pallen, and Ian R. Henderson. 2005. 'Genomic Analysis of the Protein Secretion Systems in *Clostridium Acetobutylicum* ATCC 824'. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1745 (2):223–53. <https://doi.org/10.1016/j.bbamcr.2005.04.006>.
- Dhugga, Kanwarpal S., Roberto Barreiro, Brad Whitten, Kevin Stecca, Jan Hazebroek, Gursharn S. Randhawa, Maureen Dolan, et al. 2004. 'Guar Seed β -Mannan Synthase Is a Member of the Cellulose Synthase Super

- Gene Family'. *Science* 303 (5656):363–66.
<https://doi.org/10.1126/science.1090908>.
- Doi, R. H., and Y. Tamaru. 2001. 'The Clostridium Cellulovorans Cellulosome: An Enzyme Complex with Plant Cell Wall Degrading Activity'. *Chemical Record (New York, N.Y.)* 1 (1):24–32.
- Dong, Hongjun, Wenwen Tao, Yanping Zhang, and Yin Li. 2012. 'Development of an Anhydrotetracycline-Inducible Gene Expression System for Solvent-Producing Clostridium Acetobutylicum: A Useful Tool for Strain Engineering'. *Metabolic Engineering* 14 (1):59–67.
<https://doi.org/10.1016/j.ymben.2011.10.004>.
- 'Effect of Acetyl Groups on Enzymatic Hydrolysis of Cellulosic Substrates : Holzforschung'. 2006. 1 July 2006.
<https://www.degruyter.com/view/j/hfsg.2006.60.issue-4/hf.2006.062/hf.2006.062.xml>.
- Ehsaan, M., W. Kuit, Y. Zhang, S.T. Cartman, J.T. Heap, K. Winzer, and N.P. Minton. 2016. 'Mutant Generation by Allelic Exchange and Genome Resequencing of the Biobutanol Organism Clostridium Acetobutylicum ATCC 824'. *Biotechnology for Biofuels* 9 (1).
<https://doi.org/10.1186/s13068-015-0410-0>.
- Eriksson, Torny, Johan Karlsson, and Folke Tjerneld. 2002. 'A Model Explaining Declining Rate in Hydrolysis of Lignocellulose Substrates with Cellobiohydrolase I (Cel7A) and Endoglucanase I (Cel7B) of Trichoderma Reesei'. *Applied Biochemistry and Biotechnology* 101 (1):41–60. <https://doi.org/10.1385/ABAB:101:1:41>.
- Esa, Faezah, Siti Masrinda Tasirin, and Norliza Abd Rahman. 2014. 'Overview of Bacterial Cellulose Production and Application'. *Agriculture and Agricultural Science Procedia*, 2nd International Conference on Agricultural and Food Engineering (CAFE 2014) - New Trends Forward, 2 (Supplement C):113–19.
<https://doi.org/10.1016/j.aaspro.2014.11.017>.
- Fan, L. T., Yong-Hyun Lee, and David H. Beardmore. 1980. 'Mechanism of the Enzymatic Hydrolysis of Cellulose: Effects of Major Structural Features of Cellulose on Enzymatic Hydrolysis'. *Biotechnology and Bioengineering* 22 (1):177–99. <https://doi.org/10.1002/bit.260220113>.
- Fendri, Imen, Chantal Tardif, Henri-Pierre Fierobe, Sabrina Lignon, Odile Valette, Sandrine Pagès, and Stéphanie Perret. 2009. 'The Cellulosomes from Clostridium Cellulolyticum'. *FEBS Journal* 276 (11):3076–86.
<https://doi.org/10.1111/j.1742-4658.2009.07025.x>.
- Ferdinand, Pierre-Henri, Romain Borne, Valentine Trotter, Sandrine Pagès, Chantal Tardif, Henri-Pierre Fierobe, and Stéphanie Perret. 2013. 'Are Cellulosome Scaffolding Protein CipC and CBM3-Containing Protein HycP, Involved in Adherence of Clostridium Cellulolyticum to Cellulose?' *PLOS ONE* 8 (7):e69360.
<https://doi.org/10.1371/journal.pone.0069360>.
- Feustel, Lothar, Stephan Nakotte, and Peter Dürre. 2004. 'Characterization and Development of Two Reporter Gene Systems for Clostridium Acetobutylicum'. *Applied and Environmental Microbiology* 70 (2):798–803.
- [IPCC, 2014). "Summary for Policymakers. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects.

- Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.” IPCC 2014.
- Fierobe, Henri-Pierre, Adva Mechaly, Chantal Tardif, Anne Belaich, Raphael Lamed, Yuval Shoham, Jean-Pierre Belaich, and Edward A. Bayer. 2001. ‘Design and Production of Active Cellulosome Chimeras SELECTIVE INCORPORATION OF DOCKERIN-CONTAINING ENZYMES INTO DEFINED FUNCTIONAL COMPLEXES’. *Journal of Biological Chemistry* 276 (24):21257–61. <https://doi.org/10.1074/jbc.M102082200>.
- Fierobe, Henri-Pierre, Florence Mingardon, and Angélique Chanal. 2012. ‘Chapter Sixteen - Engineering Cellulase Activity into *Clostridium Acetobutylicum*’. In *Methods in Enzymology*, edited by Harry J. Gilbert, Volume 510:301–16. Cellulases. Academic Press. <http://www.sciencedirect.com/science/article/pii/B9780124159310000161>.
- Fierobe, Henri-Pierre, Florence Mingardon, Adva Mechaly, Anne Bélaïch, Marco T. Rincon, Sandrine Pagès, Raphael Lamed, Chantal Tardif, Jean-Pierre Bélaïch, and Edward A. Bayer. 2005. ‘Action of Designer Cellulosomes on Homogeneous Versus Complex Substrates CONTROLLED INCORPORATION OF THREE DISTINCT ENZYMES INTO A DEFINED TRIFUNCTIONAL SCAFFOLDIN’. *Journal of Biological Chemistry* 280 (16):16325–34. <https://doi.org/10.1074/jbc.M414449200>.
- Fontes, Carlos M.G.A., and Harry J. Gilbert. 2010. ‘Cellulosomes: Highly Efficient Nanomachines Designed to Deconstruct Plant Cell Wall Complex Carbohydrates’. *Annual Review of Biochemistry* 79 (1):655–81. <https://doi.org/10.1146/annurev-biochem-091208-085603>.
- ‘From Cells to Muropeptide Structures in 24 H: Peptidoglycan Mapping by UPLC-MS : Scientific Reports’. n.d. Accessed 17 May 2016. <http://www.nature.com/articles/srep07494#methods>.
- Fujita, Kosuke, Kumiko Shimomura, Kei-ichiro Yamamoto, Tetsuro Yamashita, and Koichi Suzuki. 2006. ‘A Chitinase Structurally Related to the Glycoside Hydrolase Family 48 Is Indispensable for the Hormonally Induced Diapause Termination in a Beetle’. *Biochemical and Biophysical Research Communications* 345 (1):502–7. <https://doi.org/10.1016/j.bbrc.2006.04.126>.
- Gao, Dongfang, Yaqi Luan, Qian Wang, Quanfeng Liang, and Qingsheng Qi. 2015. ‘Construction of Cellulose-Utilizing *Escherichia Coli* Based on a Secretable Cellulase’. *Microbial Cell Factories* 14 (October). <https://doi.org/10.1186/s12934-015-0349-7>.
- Gaspar, Paulo, Gabriela Moura, Manuel A. S. Santos, and José Luís Oliveira. 2013. ‘mRNA Secondary Structure Optimization Using a Correlated Stem–loop Prediction’. *Nucleic Acids Research* 41 (6):e73. <https://doi.org/10.1093/nar/gks1473>.
- Gefen, Gilad, Michael Anbar, Ely Morag, Raphael Lamed, and Edward A. Bayer. 2012. ‘Enhanced Cellulose Degradation by Targeted Integration of a Cohesin-Fused β -Glucosidase into the *Clostridium Thermocellum* Cellulosome’. *Proceedings of the National Academy of Sciences* 109 (26):10298–303. <https://doi.org/10.1073/pnas.1202747109>.

- Gerland, Patrick, Adrian E. Raftery, Hana Ševčíková, Nan Li, Danan Gu, Thomas Spoorenberg, Leontine Alkema, et al. 2014. 'World Population Stabilization Unlikely This Century'. *Science* 346 (6206):234–37. <https://doi.org/10.1126/science.1257469>.
- Girbal, Laurence, Isabelle Mortier-Barrière, Frédéric Raynaud, Céline Rouanet, Christian Croux, and Philippe Soucaille. 2003. 'Development of a Sensitive Gene Expression Reporter System and an Inducible Promoter-Repressor System for *Clostridium Acetobutylicum*'. *Applied and Environmental Microbiology* 69 (8):4985–88. <https://doi.org/10.1128/AEM.69.8.4985-4988.2003>.
- Green, Edward M. 2011. 'Fermentative Production of Butanol—the Industrial Perspective'. *Current Opinion in Biotechnology*, Energy biotechnology – Environmental biotechnology, 22 (3):337–43. <https://doi.org/10.1016/j.copbio.2011.02.004>.
- G. Resch, Michael, Bryon S. Donohoe, John O. Baker, Stephen R. Decker, Edward A. Bayer, Gregg T. Beckham, and Michael E. Himmel. 2013. 'Fungal Cellulases and Complexed Cellulosomal Enzymes Exhibit Synergistic Mechanisms in Cellulose Deconstruction'. *Energy & Environmental Science* 6 (6):1858–67. <https://doi.org/10.1039/C3EE00019B>.
- Grunberg-Manago, Marianne. 1999. 'Messenger RNA Stability and Its Role in Control of Gene Expression in Bacteria and Phages'. *Annual Review of Genetics* 33 (1):193–227. <https://doi.org/10.1146/annurev.genet.33.1.193>.
- Gu, Wanjun, Tong Zhou, and Claus O. Wilke. 2010. 'A Universal Trend of Reduced mRNA Stability near the Translation-Initiation Site in Prokaryotes and Eukaryotes'. *PLoS Comput Biol* 6 (2):e1000664. <https://doi.org/10.1371/journal.pcbi.1000664>.
- Haimovitz, Rachel, Yoav Barak, Ely Morag, Milana Voronov-Goldman, Yuval Shoham, Raphael Lamed, and Edward A. Bayer. 2008. 'Cohesin-Dockerin Microarray: Diverse Specificities between Two Complementary Families of Interacting Protein Modules'. *PROTEOMICS* 8 (5):968–79. <https://doi.org/10.1002/pmic.200700486>.
- Haitjema, Charles H., Sean P. Gilmore, John K. Henske, Kevin V. Solomon, Randall de Groot, Alan Kuo, Stephen J. Mondo, et al. 2017. 'A Parts List for Fungal Cellulosomes Revealed by Comparative Genomics'. *Nature Microbiology* 2 (8):nmicrobiol201787. <https://doi.org/10.1038/nmicrobiol.2017.87>.
- Hall, Mélanie, Prabuddha Bansal, Jay H. Lee, Matthew J. Realff, and Andreas S. Bommarius. 2010. 'Cellulose Crystallinity – a Key Predictor of the Enzymatic Hydrolysis Rate'. *FEBS Journal* 277 (6):1571–82. <https://doi.org/10.1111/j.1742-4658.2010.07585.x>.
- Harris, Darby M., Kendall Corbin, Tuo Wang, Ryan Gutierrez, Ana L. Bertolo, Carloalberto Petti, Detlef-M. Smilgies, et al. 2012. 'Cellulose Microfibril Crystallinity Is Reduced by Mutating C-Terminal Transmembrane Region Residues CESA1A903V and CESA3T942I of Cellulose Synthase'. *Proceedings of the National Academy of Sciences* 109 (11):4098–4103. <https://doi.org/10.1073/pnas.1200352109>.
- Hartman, A. H., H. Liu, and S. B. Melville. 2011. 'Construction and Characterization of a Lactose-Inducible Promoter System for

- Controlled Gene Expression in *Clostridium Perfringens*'. *Applied and Environmental Microbiology* 77 (2):471–78.
<https://doi.org/10.1128/AEM.01536-10>.
- Harwood, Colin R., and Rocky Cranenburgh. 2008. 'Bacillus Protein Secretion: An Unfolding Story'. *Trends in Microbiology* 16 (2):73–79.
<https://doi.org/10.1016/j.tim.2007.12.001>.
- Hasunuma, Tomohisa, and Akihiko Kondo. 2012. 'Development of Yeast Cell Factories for Consolidated Bioprocessing of Lignocellulose to Bioethanol through Cell Surface Engineering'. *Biotechnology Advances*, Special issue on ACB 2011, 30 (6):1207–18.
<https://doi.org/10.1016/j.biotechadv.2011.10.011>.
- Heap, John T., Muhammad Ehsaan, Clare M. Cooksley, Yen-Kuan Ng, Stephen T. Cartman, Klaus Winzer, and Nigel P. Minton. 2012. 'Integration of DNA into Bacterial Chromosomes from Plasmids without a Counter-Selection Marker'. *Nucleic Acids Research* 40 (8):e59–e59. <https://doi.org/10.1093/nar/gkr1321>.
- Heap, John T., Oliver J. Pennington, Stephen T. Cartman, Glen P. Carter, and Nigel P. Minton. 2007. 'The ClosTron: A Universal Gene Knock-out System for the Genus *Clostridium*'. *Journal of Microbiological Methods* 70 (3):452–64. <https://doi.org/10.1016/j.mimet.2007.05.021>.
- Heap, John T., Oliver J. Pennington, Stephen T. Cartman, and Nigel P. Minton. 2009. 'A Modular System for *Clostridium* Shuttle Plasmids'. *Journal of Microbiological Methods* 78 (1):79–85.
<https://doi.org/10.1016/j.mimet.2009.05.004>.
- Heather E. McFarlane, Annet Doring, and Staffan Persson. 2014. 'The Cell Biology of Cellulose Synthesis'. *Annual Review of Plant Biology* 65 (1):69–94. <https://doi.org/10.1146/annurev-arplant-050213-040240>.
- Herbert, Silvia, Agnieszka Bera, Christiane Nerz, Dirk Kraus, Andreas Peschel, Christiane Goerke, Michael Meehl, Ambrose Cheung, and Friedrich Götz. 2007. 'Molecular Basis of Resistance to Muramidase and Cationic Antimicrobial Peptide Activity of Lysozyme in *Staphylococci*'. *PLOS Pathogens* 3 (7):e102.
<https://doi.org/10.1371/journal.ppat.0030102>.
- Herring, Christopher D., William R. Kenealy, A. Joe Shaw, Sean F. Covalla, Daniel G. Olson, Jiayi Zhang, W. Ryan Sillers, et al. 2016. 'Strain and Bioprocess Improvement of a Thermophilic Anaerobe for the Production of Ethanol from Wood'. *Biotechnology for Biofuels* 9 (June):125. <https://doi.org/10.1186/s13068-016-0536-8>.
- Higashide, Wendy, Yongchao Li, Yunfeng Yang, and James C. Liao. 2011. 'Metabolic Engineering of *Clostridium Cellulolyticum* for Production of Isobutanol from Cellulose'. *Applied and Environmental Microbiology* 77 (8):2727–33. <https://doi.org/10.1128/AEM.02454-10>.
- Höltje, Joachim-Volker, Ursula Kopp, Astrid Ursinus, and Bernd Wiedemann. 1994. 'The Negative Regulator of β -Lactamase Induction AmpD Is a N-Acetyl-Anhydromuramyl-L-Alanine Amidase'. *FEMS Microbiology Letters* 122 (1):159–64.
- Hsieh, Chia-wen C., David Cannella, Henning Jørgensen, Claus Felby, and Lisbeth G. Thygesen. 2014. 'Cellulase Inhibition by High Concentrations of Monosaccharides'. *Journal of Agricultural and Food Chemistry* 62 (17):3800–3805. <https://doi.org/10.1021/jf5012962>.

- Hu, Wen-Jing, Scott A. Harding, Jrhau Lung, Jacqueline L. Popko, John Ralph, Douglas D. Stokke, Chung-Jui Tsai, and Vincent L. Chiang. 1999. 'Repression of Lignin Biosynthesis Promotes Cellulose Accumulation and Growth in Transgenic Trees'. *Nature Biotechnology* 17 (8):808–12. <https://doi.org/10.1038/11758>.
- Hundsberger, Thomas, Veil Braun, Manfred Weidmann, Petra Leukel, Markus Sauerborn, and Christoph Von Eichel-Streiber. 1997. 'Transcription Analysis of the Genes tcdA-E of the Pathogenicity Locus of *Clostridium Difficile*'. *European Journal of Biochemistry* 244 (3):735–42. <https://doi.org/10.1111/j.1432-1033.1997.t01-1-00735.x>.
- Hyrylainen, H. L., M. Vitikainen, J. Thwaite, H. Wu, M. Sarvas, C. R. Harwood, V. P. Kontinen, and K. Stephenson. 2000. 'D-Alanine Substitution of Teichoic Acids as a Modulator of Protein Folding and Stability at the Cytoplasmic Membrane/Cell Wall Interface of *Bacillus Subtilis*'. *The Journal of Biological Chemistry* 275 (35):26696–703. <https://doi.org/10.1074/jbc.M003804200>.
- Hyryläinen, Hanne-Leena, Albert Bolhuis, Elise Darmon, Leila Muukkonen, Pertti Koski, Marika Vitikainen, Matti Sarvas, et al. 2001. 'A Novel Two-Component Regulatory System in *Bacillus Subtilis* for the Survival of Severe Secretion Stress'. *Molecular Microbiology* 41 (5):1159–72. <https://doi.org/10.1046/j.1365-2958.2001.02576.x>.
- Ibrahim, Hisham R, Tetsuji Matsuzaki, and Takayoshi Aoki. 2001. 'Genetic Evidence That Antibacterial Activity of Lysozyme Is Independent of Its Catalytic Function'. *FEBS Letters* 506 (1):27–32. [https://doi.org/10.1016/S0014-5793\(01\)02872-1](https://doi.org/10.1016/S0014-5793(01)02872-1).
- Inoue, Makoto, Hidenori Yamada, Takanori Yasukochi, Ryota Kuroki, Takeyoshi Miki, Tadao Horiuchi, and Taiji Imoto. 1992. 'Multiple Role of Hydrophobicity of Tryptophan-108 in Chicken Lysozyme: Structural Stability, Saccharide Binding Ability, and Abnormal pKa of Glutamic Acid-35'. *Biochemistry* 31 (24):5545–53. <https://doi.org/10.1021/bi00139a017>.
- Izquierdo, Javier A., Maria V. Sizova, and Lee R. Lynd. 2010. 'Diversity of Bacteria and Glycosyl Hydrolase Family 48 Genes in Cellulolytic Consortia Enriched from Thermophilic Biocompost'. *Applied and Environmental Microbiology* 76 (11):3545–53. <https://doi.org/10.1128/AEM.02689-09>.
- Jacobs, M., J. B. Andersen, V. Kontinen, and M. Sarvas. 1993. 'Bacillus Subtilis PrsA Is Required in Vivo as an Extracytoplasmic Chaperone for Secretion of Active Enzymes Synthesized Either with or without pro-Sequences'. *Molecular Microbiology* 8 (5):957–66. <https://doi.org/10.1111/j.1365-2958.1993.tb01640.x>.
- Janssen, Holger, Christina Döring, Armin Ehrenreich, Birgit Voigt, Michael Hecker, Hubert Bahl, and Ralf-Jörg Fischer. 2010. 'A Proteomic and Transcriptional View of Acidogenic and Solventogenic Steady-State Cells of *Clostridium Acetobutylicum* in a Chemostat Culture'. *Applied Microbiology and Biotechnology* 87 (6):2209–26. <https://doi.org/10.1007/s00253-010-2741-x>.
- Jiang, Yu, Jinle Liu, Weihong Jiang, Yunliu Yang, and Sheng Yang. 2015. 'Current Status and Prospects of Industrial Bio-Production of N-Butanol in China'. *Biotechnology Advances, Industrial Biotechnology*:

- Tools and Applications, 33 (7):1493–1501.
<https://doi.org/10.1016/j.biotechadv.2014.10.007>.
- Jin, Mingjie, Venkatesh Balan, Christa Gunawan, and Bruce E. Dale. 2011. 'Consolidated Bioprocessing (CBP) Performance of Clostridium Phytofermentans on AFEX-Treated Corn Stover for Ethanol Production'. *Biotechnology and Bioengineering* 108 (6):1290–97.
<https://doi.org/10.1002/bit.23059>.
- Jones, D T, and D R Woods. 1986. 'Acetone-Butanol Fermentation Revisited.' *Microbiological Reviews* 50 (4):484–524.
- Jürgen, B., T. Schweder, and M. Hecker. 1998. 'The Stability of mRNA from the *gsiB* Gene of *Bacillus Subtilis* Is Dependent on the Presence of a Strong Ribosome Binding Site'. *Molecular & General Genetics: MGG* 258 (5):538–45.
- Kashyap, Des Raj, Minhui Wang, Li-Hui Liu, Geert-Jan Boons, Dipika Gupta, and Roman Dziarski. 2011. 'Peptidoglycan Recognition Proteins Kill Bacteria by Activating Protein-Sensing Two-Component Systems'. *Nature Medicine* 17 (6):676–83. <https://doi.org/10.1038/nm.2357>.
- Kataeva, Irina A., Ronald D. Seidel, Ashit Shah, Larry T. West, Xin-Liang Li, and Lars G. Ljungdahl. 2002. 'The Fibronectin Type 3-Like Repeat from the Clostridium Thermocellum Cellobiohydrolase CbhA Promotes Hydrolysis of Cellulose by Modifying Its Surface'. *Applied and Environmental Microbiology* 68 (9):4292–4300.
<https://doi.org/10.1128/AEM.68.9.4292-4300.2002>.
- Kim, A. Y., G. T. Attwood, S. M. Holt, B. A. White, and H. P. Blaschek. 1994. 'Heterologous Expression of Endo-Beta-1,4-D-Glucanase from Clostridium Cellulovorans in Clostridium Acetobutylicum ATCC 824 Following Transformation of the *engB* Gene.' *Applied and Environmental Microbiology* 60 (1):337–40.
- Klemm, Dieter, Brigitte Heublein, Hans-Peter Fink, and Andreas Bohn. 2005. 'Cellulose: Fascinating Biopolymer and Sustainable Raw Material'. *Angewandte Chemie International Edition* 44 (22):3358–93.
<https://doi.org/10.1002/anie.200460587>.
- Kobayashi, Kaori, I Putu Sudiarta, Takeko Kodama, Tatsuya Fukushima, Katsutoshi Ara, Katsuya Ozaki, and Junichi Sekiguchi. 2012. 'Identification and Characterization of a Novel Polysaccharide Deacetylase C (PdaC) from *Bacillus Subtilis*'. *The Journal of Biological Chemistry* 287 (13):9765–76.
<https://doi.org/10.1074/jbc.M111.329490>.
- Koeck, Daniela E, Alexander Pechtl, Vladimir V Zverlov, and Wolfgang H Schwarz. 2014. 'Genomics of Cellulolytic Bacteria'. *Current Opinion in Biotechnology, Cell and Pathway Engineering*, 29 (October):171–83.
<https://doi.org/10.1016/j.copbio.2014.07.002>.
- Komar, Anton A. 2016. 'The Art of Gene Redesign and Recombinant Protein Production: Approaches and Perspectives'. In *Protein Therapeutics*, 161–77. Topics in Medicinal Chemistry. Springer, Cham.
https://doi.org/10.1007/7355_2016_2.
- Kong, Fanran, Cady R. Engler, and Ed J. Soltes. 1992. 'Effects of Cell-Wall Acetate, Xylan Backbone, and Lignin on Enzymatic Hydrolysis of Aspen Wood'. *Applied Biochemistry and Biotechnology* 34–35 (1):23–35. <https://doi.org/10.1007/BF02920531>.

- Kontinen, V. P., and M. Sarvas. 1993. 'The PrsA Lipoprotein Is Essential for Protein Secretion in *Bacillus Subtilis* and Sets a Limit for High-Level Secretion'. *Molecular Microbiology* 8 (4):727–37.
- Kosugi, Akihiko, Yoshihiko Amano, Koichiro Murashima, and Roy H. Doi. 2004. 'Hydrophilic Domains of Scaffolding Protein CbpA Promote Glycosyl Hydrolase Activity and Localization of Cellulosomes to the Cell Surface of *Clostridium Cellulovorans*'. *Journal of Bacteriology* 186 (19):6351–59. <https://doi.org/10.1128/JB.186.19.6351-6359.2004>.
- Kosugi, Akihiko, Koichiro Murashima, Yutaka Tamaru, and Roy H. Doi. 2002. 'Cell-Surface-Anchoring Role of N-Terminal Surface Layer Homology Domains of *Clostridium Cellulovorans* EngE'. *Journal of Bacteriology* 184 (4):884–88.
- Kovács, Katalin, Benjamin J. Willson, Katrin Schwarz, John T. Heap, Adam Jackson, David N. Bolam, Klaus Winzer, and Nigel P. Minton. 2013. 'Secretion and Assembly of Functional Mini-Cellulosomes from Synthetic Chromosomal Operons in *Clostridium Acetobutylicum* ATCC 824'. *Biotechnology for Biofuels* 6 (1):117. <https://doi.org/10.1186/1754-6834-6-117>.
- Krauss, Jan, Vladimir V. Zverlov, and Wolfgang H. Schwarz. 2012. 'In Vitro Reconstitution of the Complete *Clostridium Thermocellum* Cellulosome and Synergistic Activity on Crystalline Cellulose'. *Applied and Environmental Microbiology* 78 (12):4301–7. <https://doi.org/10.1128/AEM.07959-11>.
- Kroukamp, Heinrich, Riaan den Haan, Niël van Wyk, and Willem H. van Zyl. 2013. 'Overexpression of Native PSE1 and SOD1 in *Saccharomyces Cerevisiae* Improved Heterologous Cellulase Secretion'. *Applied Energy*, Special Issue on Advances in sustainable biofuel production and use - XIX International Symposium on Alcohol Fuels - ISAF, 102 (February):150–56. <https://doi.org/10.1016/j.apenergy.2012.05.062>.
- Kuehne, Sarah A., and Nigel P. Minton. 2012. 'ClosTron-Mediated Engineering of *Clostridium*'. *Bioengineered* 3 (4):247–54. <https://doi.org/10.4161/bioe.21004>.
- Kumar, Adepu Kiran, and Shaishav Sharma. 2017. 'Recent Updates on Different Methods of Pretreatment of Lignocellulosic Feedstocks: A Review'. *Bioresources and Bioprocessing* 4 (1). <https://doi.org/10.1186/s40643-017-0137-9>.
- Laaberki, Maria-Halima, John Pfeffer, Anthony J. Clarke, and Jonathan Dworkin. 2011. 'O-Acetylation of Peptidoglycan Is Required for Proper Cell Separation and S-Layer Anchoring in *Bacillus Anthracis*'. *The Journal of Biological Chemistry* 286 (7):5278–88. <https://doi.org/10.1074/jbc.M110.183236>.
- Lambertz, Camilla, Megan Garvey, Johannes Klinger, Dirk Heesel, Holger Klose, Rainer Fischer, and Ulrich Commandeur. 2014. 'Challenges and Advances in the Heterologous Expression of Cellulolytic Enzymes: A Review'. *Biotechnology for Biofuels* 7 (October). <https://doi.org/10.1186/s13068-014-0135-5>.
- Lamed, R., E. Setter, R. Kenig, and E. A. Bayer. 1983. 'The Cellulosome: A Discrete Cell Surface Organelle of *Clostridium Thermocellum* Which Exhibits Separate Antigenic, Cellulose-Binding and Various Cellulolytic Activities'. *ResearchGate* 13 (13).

- https://www.researchgate.net/publication/236400605_The_cellulosome_A_discrete_cell_surface_organelle_of_Clostridium_thermocellum_which_exhibits_separate_antigenic_cellulose-binding_and_various_cellulolytic_activities.
- Lawson, Paul A., and Fred A. Rainey. 2016. 'Proposal to Restrict the Genus *Clostridium* Prazmowski to *Clostridium* Butyricum and Related Species'. *International Journal of Systematic and Evolutionary Microbiology* 66 (2):1009–16. <https://doi.org/10.1099/ijsem.0.000824>.
- Lee, Joungmin, Yu-Sin Jang, Eleftherios T. Papoutsakis, and Sang Yup Lee. 2016. 'Stable and Enhanced Gene Expression in *Clostridium* Acetobutylicum Using Synthetic Untranslated Regions with a Stem-Loop'. *Journal of Biotechnology* 230 (July):40–43. <https://doi.org/10.1016/j.jbiotec.2016.05.020>.
- Lee, Sang Yup, Jin Hwan Park, Seh Hee Jang, Lars K. Nielsen, Jaehyun Kim, and Kwang S. Jung. 2008. 'Fermentative Butanol Production by Clostridia'. *Biotechnology and Bioengineering* 101 (2):209–28. <https://doi.org/10.1002/bit.22003>.
- Lee, Song F., Cecil W. Forsberg, and L. N. Gibbins. 1985a. 'Cellulolytic Activity of *Clostridium* Acetobutylicum'. *Applied and Environmental Microbiology* 50 (2):220–28.
- Lee, Song F., Cecil W. Forsberg, and L. N. Gibbins. 1985b. "Xylanolytic Activity of *Clostridium* Acetobutylicum." *Applied and Environmental Microbiology* 50 (4): 1068–76.
- Lee, Timothy K., and Kerwyn Casey Huang. 2013. 'The Role of Hydrolases in Bacterial Cell-Wall Growth'. *Current Opinion in Microbiology*, Growth and development: eukaryotes/prokaryotes, 16 (6):760–66. <https://doi.org/10.1016/j.mib.2013.08.005>.
- Lin, Chia-Chi, Cally Joe San Yap, Shu-Chen Kan, Nai-Chi Hsueh, Liang-Yu Yang, Chwen-Jen Shieh, Chieh-Chen Huang, and Yung-Chuan Liu. 2017. 'Deciphering Characteristics of the Designer Cellulosome from *Bacillus Subtilis* WB800N via Enzymatic Analysis'. *Biochemical Engineering Journal* 117 (January):147–55. <https://doi.org/10.1016/j.bej.2016.10.010>.
- Lin, Paul P., Luo Mi, Amy H. Morioka, Kouki M. Yoshino, Sawako Konishi, Sharon C. Xu, Beth A. Papanek, Lauren A. Riley, Adam M. Guss, and James C. Liao. 2015. 'Consolidated Bioprocessing of Cellulose to Isobutanol Using *Clostridium* Thermocellum'. *Metabolic Engineering* 31 (September):44–52. <https://doi.org/10.1016/j.ymben.2015.07.001>.
- Lindholm, Agneta, Ulla Ellmén, Marja Tolonen-Martikainen, and Airi Palva. 2006. 'Heterologous Protein Secretion in *Lactococcus Lactis* Is Enhanced by the *Bacillus Subtilis* Chaperone-like Protein PrsA'. *Applied Microbiology and Biotechnology*. http://agris.fao.org/agris-search/search.do;jsessionid=AE33635115F3697125CF7068E47D5890?request_locale=es&recordID=US201301118339&sourceQuery=&query=&sortField=&sortOrder=&agrovocString=&advQuery=&centerString=&enableField=.
- Liu, Chi-Chi, and Roy H Doi. 1998. 'Properties of *exgS*, a Gene for a Major Subunit of the *Clostridium* Cellulovorans Cellulosome'. *Gene* 211 (1):39–47. [https://doi.org/10.1016/S0378-1119\(98\)00081-X](https://doi.org/10.1016/S0378-1119(98)00081-X).

- Liu, Gang, Jian Zhang, and Jie Bao. 2016. 'Cost Evaluation of Cellulase Enzyme for Industrial-Scale Cellulosic Ethanol Production Based on Rigorous Aspen Plus Modeling'. *Bioprocess and Biosystems Engineering* 39 (1):133–40. <https://doi.org/10.1007/s00449-015-1497-1>.
- Liu, Zhen, Kai Qiao, Lei Tian, Quan Zhang, Zi-Yong Liu, and Fu-Li Li. 2015. 'Spontaneous Large-Scale Autolysis in *Clostridium Acetobutylicum* Contributes to Generation of More Spores'. *Microbial Physiology and Metabolism*, 950. <https://doi.org/10.3389/fmicb.2015.00950>.
- Liu, Zhuo, Shih-Hsin Ho, Kengo Sasaki, Riaan den Haan, Kentaro Inokuma, Chiaki Ogino, Willem H. van Zyl, Tomohisa Hasunuma, and Akihiko Kondo. 2016. 'Engineering of a Novel Cellulose-Adherent Cellulolytic *Saccharomyces Cerevisiae* for Cellulosic Biofuel Production'. *Scientific Reports* 6 (April). <https://doi.org/10.1038/srep24550>.
- Lombard, Vincent, Hemalatha Golaconda Ramulu, Elodie Drula, Pedro M Coutinho, and Bernard Henrissat. 2014. 'The Carbohydrate-Active Enzymes Database (CAZy) in 2013'. *Nucleic Acids Research* 42 (1):D490-495. <https://doi.org/10.1093/nar/gkt1178>.
- López-Contreras, Ana M., Krisztina Gabor, Aernout A. Martens, Bernadet A. M. Renckens, Pieter A. M. Claassen, John van der Oost, and Willem M. de Vos. 2004. 'Substrate-Induced Production and Secretion of Cellulases by *Clostridium Acetobutylicum*'. *Applied and Environmental Microbiology* 70 (9):5238–43. <https://doi.org/10.1128/AEM.70.9.5238-5243.2004>.
- López-Contreras, Ana M., Aernout A. Martens, Nora Szijarto, Hans Mooibroek, Pieter A. M. Claassen, John van der Oost, and Willem M. de Vos. 2003. 'Production by *Clostridium Acetobutylicum* ATCC 824 of CelG, a Cellulosomal Glycoside Hydrolase Belonging to Family 9'. *Applied and Environmental Microbiology* 69 (2):869–77. <https://doi.org/10.1128/AEM.69.2.869-877.2003>.
- Lu, Xiaofeng, Minhui Wang, Jin Qi, Haitao Wang, Xinna Li, Dipika Gupta, and Roman Dziarski. 2006. 'Peptidoglycan Recognition Proteins Are a New Class of Human Bactericidal Proteins'. *Journal of Biological Chemistry* 281 (9):5895–5907. <https://doi.org/10.1074/jbc.M511631200>.
- Lynd, Lee R., Paul J. Weimer, Willem H. van Zyl, and Isak S. Pretorius. 2002. 'Microbial Cellulose Utilization: Fundamentals and Biotechnology'. *Microbiology and Molecular Biology Reviews* 66 (3):506–77. <https://doi.org/10.1128/MMBR.66.3.506-577.2002>.
- Lynd, Lee R., Willem H van Zyl, John E McBride, and Mark Laser. 2005. 'Consolidated Bioprocessing of Cellulosic Biomass: An Update'. *Current Opinion in Biotechnology, Tissue and cell engineering/Biochemical engineering*, 16 (5):577–83. <https://doi.org/10.1016/j.copbio.2005.08.009>.
- Maamar, Hédia, Odile Valette, Henri-Pierre Fierobe, Anne Bélaich, Jean-Pierre Bélaich, and Chantal Tardif. 2004. 'Cellulolysis Is Severely Affected in *Clostridium Cellulolyticum* Strain cipCMut1'. *Molecular Microbiology* 51 (2):589–98. <https://doi.org/10.1046/j.1365-2958.2003.03859.x>.
- Maddox, I. S., E. Steiner, S. Hirsch, S. Wessner, N. A. Gutierrez, J. R. Gapes, and K. C. Schuster. 2000. 'The Cause Of “acid-Crash” and “acidogenic

- Fermentations” during the Batch Acetone-Butanol-Ethanol (ABE-) Fermentation Process’. *Journal of Molecular Microbiology and Biotechnology* 2 (1):95–100.
- Malcolm, B. A., S. Rosenberg, M. J. Corey, J. S. Allen, A. de Baetselier, and J. F. Kirsch. 1989. ‘Site-Directed Mutagenesis of the Catalytic Residues Asp-52 and Glu-35 of Chicken Egg White Lysozyme’. *Proceedings of the National Academy of Sciences* 86 (1):133–37.
- Mann, Michael E., Stefan Rahmstorf, Kai Kornhuber, Byron A. Steinman, Sonya K. Miller, and Dim Coumou. 2017. ‘Influence of Anthropogenic Climate Change on Planetary Wave Resonance and Extreme Weather Events’. *Scientific Reports* 7 (March).
<https://doi.org/10.1038/srep45242>.
- Markiewicz, Z., B. Glauner, and U. Schwarz. 1983. ‘Murein Structure and Lack of DD- and LD-Carboxypeptidase Activities in *Caulobacter Crescentus*’. *Journal of Bacteriology* 156 (2):649–55.
- McCarter, J. D., and S. G. Withers. 1994. ‘Mechanisms of Enzymatic Glycoside Hydrolysis’. *Current Opinion in Structural Biology* 4 (6):885–92.
- McIntosh, Lawrence P., Greg Hand, Philip E. Johnson, Manish D. Joshi, Michael Körner, Leigh A. Plesniak, Lothar Ziser, Warren W. Wakarchuk, and Stephen G. Withers. 1996. ‘The pKa of the General Acid/Base Carboxyl Group of a Glycosidase Cycles during Catalysis: A ¹³C-NMR Study of *Bacillus Circulans* Xylanase’. *Biochemistry* 35 (31):9958–66. <https://doi.org/10.1021/bi9613234>.
- Meyrand, Mickael, Aïda Boughammoura, Pascal Courtin, Christine Mézange, Alain Guillot, and Marie-Pierre Chapot-Chartier. 2007. ‘Peptidoglycan N-Acetylglucosamine Deacetylation Decreases Autolysis in *Lactococcus Lactis*’. *Microbiology* 153 (10):3275–85.
<https://doi.org/10.1099/mic.0.2007/005835-0>.
- Mingardon, Florence, Angelique Chanal, Chantal Tardif, and Henri-Pierre Fierobe. 2011. ‘The Issue of Secretion in Heterologous Expression of *Clostridium Cellulolyticum* Cellulase-Encoding Genes in *Clostridium Acetobutylicum* ATCC 824⁺’. *Applied and Environmental Microbiology* 77 (9):2831–38. <https://doi.org/10.1128/AEM.03012-10>.
- Mingardon, Florence, Stéphanie Perret, Anne Bélaïch, Chantal Tardif, Jean-Pierre Bélaïch, and Henri-Pierre Fierobe. 2005. ‘Heterologous Production, Assembly, and Secretion of a Minicellulosome by *Clostridium Acetobutylicum* ATCC 824’. *Applied and Environmental Microbiology* 71 (3):1215–22. <https://doi.org/10.1128/AEM.71.3.1215-1222.2005>.
- Minton, Nigel Peter, Ying Zhang, Katrin Schwarz, and John Heap. 2015. Bacterial expression system. US20150056168 A1, filed 28 March 2013, and issued 26 February 2015.
<http://www.google.com/patents/US20150056168>.
- Modig, Tobias, Gunnar Lidén, and Mohammad J Taherzadeh. 2002. ‘Inhibition Effects of Furfural on Alcohol Dehydrogenase, Aldehyde Dehydrogenase and Pyruvate Dehydrogenase’. *Biochemical Journal* 363 (Pt 3):769–76.
- Mohand-Oussaid, O., S. Payot, E. Guedon, E. Gelhay, A. Youyou, and H. Petitdemange. 1999. ‘The Extracellular Xylan Degradative System

- in *Clostridium Cellulolyticum* Cultivated on Xylan: Evidence for Cell-Free Cellulosome Production'. *Journal of Bacteriology* 181 (13):4035–40.
- Morag, E., A. Lapidot, D. Govorko, R. Lamed, M. Wilchek, E. A. Bayer, and Y. Shoham. 1995. 'Expression, Purification, and Characterization of the Cellulose-Binding Domain of the Scaffoldin Subunit from the Cellulosome of *Clostridium Thermocellum*.' *Applied and Environmental Microbiology* 61 (5):1980–86.
- Moraïs, Sarah, Yonit Ben David, Lizi Bensoussan, Sylvia H. Duncan, Nicole M. Koropatkin, Eric C. Martens, Harry J. Flint, and Edward A. Bayer. 2016. 'Enzymatic Profiling of Cellulosomal Enzymes from the Human Gut Bacterium, *Ruminococcus Champanellensis*, Reveals a Fine-Tuned System for Cohesin-Dockerin Recognition'. *Environmental Microbiology* 18 (2):542–56. <https://doi.org/10.1111/1462-2920.13047>.
- Moraïs, Sarah, Naama Shterzer, Inna Rozman Grinberg, Geir Mathiesen, Vincent G. H. Eijsink, Lars Axelsson, Raphael Lamed, Edward A. Bayer, and Itzhak Mizrahi. 2013. 'Establishment of a Simple *Lactobacillus Plantarum* Cell Consortium for Cellulase-Xylanase Synergistic Interactions'. *Applied and Environmental Microbiology* 79 (17):5242–49. <https://doi.org/10.1128/AEM.01211-13>.
- Moraïs, Sarah, Naama Shterzer, Raphael Lamed, Edward A. Bayer, and Itzhak Mizrahi. 2014. 'A Combined Cell-Consortium Approach for Lignocellulose Degradation by Specialized *Lactobacillus Plantarum* Cells'. *Biotechnology for Biofuels* 7 (1):112. <https://doi.org/10.1186/1754-6834-7-112>.
- Moynihan, Patrick J., David Sychantha, and Anthony J. Clarke. 2014. 'Chemical Biology of Peptidoglycan Acetylation and Deacetylation'. *Bioorganic Chemistry* 54 (June):44–50. <https://doi.org/10.1016/j.bioorg.2014.03.010>.
- Müller-Hill, B., Crapo, and Gilbert W. 1968. 'Mutants That Make More Lac Repressor.' *Proceedings of the National Academy of Sciences of the United States of America*, no. 59:1259.
- Munasinghe, Pradeep Chaminda, and Samir Kumar Khanal. 2010. 'Biomass-Derived Syngas Fermentation into Biofuels: Opportunities and Challenges'. *Bioresource Technology*, Special Issue on Lignocellulosic Bioethanol: Current Status and Perspectives, 101 (13):5013–22. <https://doi.org/10.1016/j.biortech.2009.12.098>.
- Mussatto, Solange I., Giuliano Dragone, Pedro M. R. Guimarães, João Paulo A. Silva, Livia M. Carneiro, Inês C. Roberto, António Vicente, Lucília Domingues, and José A. Teixeira. 2010. 'Technological Trends, Global Market, and Challenges of Bio-Ethanol Production'. *Biotechnology Advances* 28 (6):817–30. <https://doi.org/10.1016/j.biotechadv.2010.07.001>.
- Nanda, Sonil, Dasantila Golemi-Kotra, John C. McDermott, Ajay K. Dalai, Iskender Gökalp, and Janusz A. Kozinski. 2017. 'Fermentative Production of Butanol: Perspectives on Synthetic Biology'. *New Biotechnology* 37 (Part B):210–21. <https://doi.org/10.1016/j.nbt.2017.02.006>.
- Narita, Junya, Sori Ishida, Kenji Okano, Sakurako Kimura, Hideki Fukuda, and Akihiko Kondo. 2006. 'Improvement of Protein Production in

- Lactic Acid Bacteria Using 5'-Untranslated Leader Sequence of slpA from *Lactobacillus Acidophilus*: Improvement in Protein Production Using UTLS'. *Applied Microbiology and Biotechnology* 73 (2):366–73. <https://doi.org/10.1007/s00253-006-0477-4>.
- Nölling, Jörk, Gary Breton, Marina V. Omelchenko, Kira S. Makarova, Qiandong Zeng, Rene Gibson, Hong Mei Lee, et al. 2001. 'Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium Acetobutylicum*'. *Journal of Bacteriology* 183 (16):4823–38. <https://doi.org/10.1128/JB.183.16.4823-4838.2001>.
- Noone, D., A. Howell, R. Collery, and K. M. Devine. 2001. 'YkdA and YvtA, HtrA-like Serine Proteases in *Bacillus Subtilis*, Engage in Negative Autoregulation and Reciprocal Cross-Regulation of ykdA and yvtA Gene Expression'. *Journal of Bacteriology* 183 (2):654–63. <https://doi.org/10.1128/JB.183.2.654-663.2001>.
- Olson, Daniel G., Shital A. Tripathi, Richard J. Giannone, Jonathan Lo, Nicky C. Caiazza, David A. Hogsett, Robert L. Hettich, Adam M. Guss, Genia Dubrovsky, and Lee R. Lynd. 2010. 'Deletion of the Cel48S Cellulase from *Clostridium Thermocellum*'. *Proceedings of the National Academy of Sciences* 107 (41):17727–32. <https://doi.org/10.1073/pnas.1003584107>.
- Osiro, Karen O., Brenda R. de Camargo, Rachel Satomi, Pedro Ricardo V. Hamann, Jéssica Pinheiro Silva, Marcelo Valle de Sousa, Betania F. Quirino, et al. 2017. 'Characterization of *Clostridium Thermocellum* (B8) Secretome and Purified Cellulosomes for Lignocellulosic Biomass Degradation'. *Enzyme and Microbial Technology* 97 (Supplement C):43–54. <https://doi.org/10.1016/j.enzmictec.2016.11.002>.
- O'sullivan, Antoinette C. 1997. 'Cellulose: The Structure Slowly Unravels'. *Cellulose* 4 (3):173–207. <https://doi.org/10.1023/A:1018431705579>.
- Pagès, S., A. Bélaïch, J. P. Bélaïch, E. Morag, R. Lamed, Y. Shoham, and E. A. Bayer. 1997. 'Species-Specificity of the Cohesin-Dockerin Interaction between *Clostridium Thermocellum* and *Clostridium Cellulolyticum*: Prediction of Specificity Determinants of the Dockerin Domain'. *Proteins* 29 (4):517–27.
- Pagès, Sandrine, Anne Bélaïch, Henri-Pierre Fierobe, Chantal Tardif, Christian Gaudin, and Jean-Pierre Bélaïch. 1999. 'Sequence Analysis of Scaffolding Protein CipC and ORFXp, a New Cohesin-Containing Protein in *Clostridium Cellulolyticum*: Comparison of Various Cohesin Domains and Subcellular Localization of ORFXp'. *Journal of Bacteriology* 181 (6):1801–10.
- Palmqvist, Eva, and Bärbel Hahn-Hägerdal. 2000. 'Fermentation of Lignocellulosic Hydrolysates. II: Inhibitors and Mechanisms of Inhibition'. *Bioresource Technology* 74 (1):25–33. [https://doi.org/10.1016/S0960-8524\(99\)00161-3](https://doi.org/10.1016/S0960-8524(99)00161-3).
- Pan, Ronghua, Yimei Hu, Liangkun Long, Jing Wang, and Shaojun Ding. 2016. 'Extra Carbohydrate Binding Module Contributes to the Processivity and Catalytic Activity of a Non-Modular Hydrolase Family 5 Endoglucanase from *Fomitiporia Mediterranea* MF3/22'. *Enzyme and Microbial Technology* 91 (Supplement C):42–51. <https://doi.org/10.1016/j.enzmictec.2016.06.001>.

- Park, Enoch Y., Kazuya Naruse, and Tatsuya Kato. 2012. 'One-Pot Bioethanol Production from Cellulose by Co-Culture of *Acremonium Cellulolyticus* and *Saccharomyces Cerevisiae*'. *Biotechnology for Biofuels* 5:64. <https://doi.org/10.1186/1754-6834-5-64>.
- Park, Sunghoon, and Wolfgang Schumann. 2015. 'Optimization of the Secretion Pathway for Heterologous Proteins in *Bacillus Subtilis*'. *Biotechnology and Bioprocess Engineering* 20 (4):623–33. <https://doi.org/10.1007/s12257-014-0843-5>.
- Parsieglä, Goetz, Corinne Reverbel, Chantal Tardif, Hugues Driguez, and Richard Haser. 2008. 'Structures of Mutants of Cellulase Cel48F of *Clostridium Cellulolyticum* in Complex with Long Hemithiocellooligosaccharides Give Rise to a New View of the Substrate Pathway during Processive Action'. *Journal of Molecular Biology* 375 (2):499–510. <https://doi.org/10.1016/j.jmb.2007.10.039>.
- Pauly, Markus, and Kenneth Keegstra. 2008. 'Cell-Wall Carbohydrates and Their Modification as a Resource for Biofuels'. *The Plant Journal* 54 (4):559–68. <https://doi.org/10.1111/j.1365-313X.2008.03463.x>.
- Pauly, Markus, and Kenneth Keegstra. 2010. 'Plant Cell Wall Polymers as Precursors for Biofuels'. *Current Opinion in Plant Biology* 13 (3):304–11. <https://doi.org/10.1016/j.pbi.2009.12.009>.
- Perret, S., L. Casalot, H.-P. Fierobe, C. Tardif, F. Sabathe, J.-P. Belaich, and A. Belaich. 2004. 'Production of Heterologous and Chimeric Scaffoldins by *Clostridium Acetobutylicum* ATCC 824'. *Journal of Bacteriology* 186 (1):253–57. <https://doi.org/10.1128/JB.186.1.253-257.2004>.
- Phan, Trang Thi Phuong, Hoang Duc Nguyen, and Wolfgang Schumann. 2013. 'Construction of a 5'-controllable Stabilizing Element (CoSE) for over-Production of Heterologous Proteins at High Levels in *Bacillus Subtilis*'. *Journal of Biotechnology* 168 (1):32–39. <https://doi.org/10.1016/j.jbiotec.2013.07.031>.
- Pinheiro, Benedita A., Harry J. Gilbert, Kazutaka Sakka, Kazuo Sakka, Vânia O. Fernandes, José A. M. Prates, Victor D. Alves, David N. Bolam, Luís M. A. Ferreira, and Carlos M. G. A. Fontes. 2009. 'Functional Insights into the Role of Novel type I Cohesin and Dockerin Domains from *Clostridium Thermocellum*'. *Biochemical Journal* 424 (3):375–84. <https://doi.org/10.1042/BJ20091152>.
- Poehlein, Anja, José David Montoya Solano, Stefanie K. Flitsch, Preben Krabben, Klaus Winzer, Sharon J. Reid, David T. Jones, et al. 2017. 'Microbial Solvent Formation Revisited by Comparative Genome Analysis'. *Biotechnology for Biofuels* 10 (March):58. <https://doi.org/10.1186/s13068-017-0742-z>.
- Pogliano, J., A. S. Lynch, D. Belin, E. C. Lin, and J. Beckwith. 1997. 'Regulation of *Escherichia Coli* Cell Envelope Proteins Involved in Protein Folding and Degradation by the Cpx Two-Component System.' *Genes & Development* 11 (9):1169–82. <https://doi.org/10.1101/gad.11.9.1169>.
- Pohlschröder, M., S. B. Leschine, and E. Canale-Parola. 1994. 'Multicomplex Cellulase-Xylanase System of *Clostridium Papyrosolvans* C7'. *Journal of Bacteriology* 176 (1):70–76.

- Pollegioni, Loredano, Fabio Tonin, and Elena Rosini. 2015. 'Lignin-Degrading Enzymes'. *FEBS Journal* 282 (7):1190–1213. <https://doi.org/10.1111/febs.13224>.
- Popham, D. L. 2002. 'Specialized Peptidoglycan of the Bacterial Endospore: The Inner Wall of the Lockbox'. *Cellular and Molecular Life Sciences CMLS* 59 (3):426–33. <https://doi.org/10.1007/s00018-002-8435-5>.
- Presnyak, Vladimir, Najwa Alhusaini, Ying-Hsin Chen, Sophie Martin, Nathan Morris, Nicholas Kline, Sara Olson, et al. 2015. 'Codon Optimality Is a Major Determinant of mRNA Stability'. *Cell* 160 (6):1111–24. <https://doi.org/10.1016/j.cell.2015.02.029>.
- 'Publication: World Energy Balances - 2017 Edition - Overview'. n.d. Accessed 14 September 2017. <http://www.iea.org/publications/freepublications/publication/world-energy-balances---2017-edition---overview.html>.
- Qin, Lei, Wen-Chao Li, Li Liu, Jia-Qing Zhu, Xia Li, Bing-Zhi Li, and Ying-Jin Yuan. 2016. 'Inhibition of Lignin-Derived Phenolic Compounds to Cellulase'. *Biotechnology for Biofuels* 9 (March):70. <https://doi.org/10.1186/s13068-016-0485-2>.
- Ravachol, J., R. Borne, C. Tardif, P. de Philip, and H.-P. Fierobe. 2014. 'Characterization of All Family-9 Glycoside Hydrolases Synthesized by the Cellulosome-Producing Bacterium *Clostridium Cellulolyticum*'. *Journal of Biological Chemistry* 289 (11):7335–48. <https://doi.org/10.1074/jbc.M113.545046>.
- Régnier, P., and C. M. Arraiano. 2000. 'Degradation of mRNA in Bacteria: Emergence of Ubiquitous Features'. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 22 (3):235–44. [https://doi.org/10.1002/\(SICI\)1521-1878\(200003\)22:3<235::AID-BIES5>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1521-1878(200003)22:3<235::AID-BIES5>3.0.CO;2-2).
- Reverbel-Leroy, C., S. Pages, A. Belaich, J. P. Belaich, and C. Tardif. 1997. 'The Processive Endocellulase CelF, a Major Component of the *Clostridium Cellulolyticum* Cellulosome: Purification and Characterization of the Recombinant Form.' *Journal of Bacteriology* 179 (1):46–52.
- Reverbel-Leroy, Corinne, Anne Belaich, Alain Bernadac, Christian Gaudin, Jean-Pierre Belaich, and Chantal Tardif. 1996. 'Molecular Study and Overexpression of the *Clostridium Cellulolyticum* celF Cellulase Gene in *Escherichia Coli*'. *Microbiology* 142 (4):1013–23. <https://doi.org/10.1099/00221287-142-4-1013>.
- Rincon, Marco T., Tadej Čepeljnik, Jennifer C. Martin, Raphael Lamed, Yoav Barak, Edward A. Bayer, and Harry J. Flint. 2005. 'Unconventional Mode of Attachment of the *Ruminococcus Flavefaciens* Cellulosome to the Cell Surface'. *Journal of Bacteriology* 187 (22):7569–78. <https://doi.org/10.1128/JB.187.22.7569-7578.2005>.
- Rincón, Marco T., Jennifer C. Martin, Vincenzo Aurilia, Sheila I. McCrae, Garry J. Rucklidge, Martin D. Reid, Edward A. Bayer, Raphael Lamed, and Harry J. Flint. 2004. 'ScaC, an Adaptor Protein Carrying a Novel Cohesin That Expands the Dockerin-Binding Repertoire of the *Ruminococcus Flavefaciens* 17 Cellulosome'. *Journal of Bacteriology* 186 (9):2576–85. <https://doi.org/10.1128/JB.186.9.2576-2585.2004>.

- Ruiz-Dueñas, Francisco J., and Ángel T. Martínez. 2009. 'Microbial Degradation of Lignin: How a Bulky Recalcitrant Polymer Is Efficiently Recycled in Nature and How We Can Take Advantage of This'. *Microbial Biotechnology* 2 (2):164–77. <https://doi.org/10.1111/j.1751-7915.2008.00078.x>.
- Sabathé, Fabrice, Anne Bélaïch, and Philippe Soucaille. 2002. 'Characterization of the Cellulolytic Complex (Cellulosome) of *Clostridium Acetobutylicum*'. *FEMS Microbiology Letters* 217 (1):15–22. <https://doi.org/10.1111/j.1574-6968.2002.tb11450.x>.
- Sabathé, Fabrice, and Philippe Soucaille. 2003. 'Characterization of the CipA Scaffolding Protein and In Vivo Production of a Minicellulosome in *Clostridium Acetobutylicum*'. *Journal of Bacteriology* 185 (3):1092–96. <https://doi.org/10.1128/JB.185.3.1092-1096.2003>.
- Sahoo, Dipak K., Jozsef Stork, Seth DeBolt, and Indu B. Maiti. 2013. 'Manipulating Cellulose Biosynthesis by Expression of Mutant Arabidopsis proM24::CESA3ixr1-2 Gene in Transgenic Tobacco'. *Plant Biotechnology Journal* 11 (3):362–72. <https://doi.org/10.1111/pbi.12024>.
- Salehi Jouzani, Gholamreza, and Mohammad J. Taherzadeh. 2015. 'Advances in Consolidated Bioprocessing Systems for Bioethanol and Butanol Production from Biomass : A Comprehensive Review'. *Biofuel Research Journal* 2 (1):152–95.
- Sarkar, Purbasha, Elena Bosneaga, and Manfred Auer. 2009. 'Plant Cell Walls throughout Evolution: Towards a Molecular Understanding of Their Design Principles'. *Journal of Experimental Botany* 60 (13):3615–35. <https://doi.org/10.1093/jxb/erp245>.
- Schädel, Christina, Andreas Blöchl, Andreas Richter, and Günter Hoch. 2010. 'Quantification and Monosaccharide Composition of Hemicelluloses from Different Plant Functional Types'. *Plant Physiology and Biochemistry* 48 (1):1–8. <https://doi.org/10.1016/j.plaphy.2009.09.008>.
- Scheller, Henrik Vibe, and Peter Ulvskov. 2010. 'Hemicelluloses'. *Annual Review of Plant Biology* 61 (1):263–89. <https://doi.org/10.1146/annurev-arplant-042809-112315>.
- Schleifer, K H, and O Kandler. 1972. 'Peptidoglycan Types of Bacterial Cell Walls and Their Taxonomic Implications.' *Bacteriological Reviews* 36 (4):407–77.
- Simpson, Peter J., Hefang Xie, David N. Bolam, Harry J. Gilbert, and Michael P. Williamson. 2000. 'The Structural Basis for the Ligand Specificity of Family 2 Carbohydrate-Binding Modules'. *Journal of Biological Chemistry* 275 (52):41137–42. <https://doi.org/10.1074/jbc.M006948200>.
- Speck, Thomas, and Ingo Burgert. 2011. 'Plant Stems: Functional Design and Mechanics'. *Annual Review of Materials Research* 41 (1):169–93. <https://doi.org/10.1146/annurev-matsci-062910-100425>.
- Sreekumar, Sanil, Zachary C. Baer, Anbarasan Pazhamalai, Gorkem Gunbas, Adam Grippo, Harvey W. Blanch, Douglas S. Clark, and F. Dean Toste. 2015. 'Production of an Acetone-Butanol-Ethanol Mixture from *Clostridium Acetobutylicum* and Its Conversion to High-Value Biofuels'. *Nature Protocols* 10 (3):528–37. <https://doi.org/10.1038/nprot.2015.029>.

- Stahl, Stefan W., Michael A. Nash, Daniel B. Fried, Michal Slutzki, Yoav Barak, Edward A. Bayer, and Hermann E. Gaub. 2012. 'Single-Molecule Dissection of the High-Affinity Cohesin–dockerin Complex'. *Proceedings of the National Academy of Sciences* 109 (50):20431–36. <https://doi.org/10.1073/pnas.1211929109>.
- Steiner, Elisabeth, Angel E. Dago, Danielle I. Young, John T. Heap, Nigel P. Minton, James A. Hoch, and Michael Young. 2011. 'Multiple Orphan Histidine Kinases Interact Directly with Spo0A to Control the Initiation of Endospore Formation in *Clostridium Acetobutylicum*'. *Molecular Microbiology* 80 (3):641–54. <https://doi.org/10.1111/j.1365-2958.2011.07608.x>.
- Stephenson, Keith, Noel M Carter, Colin R Harwood, Marie-Françoise Petit-Glatron, and Régis Chambert. 1998. 'The Influence of Protein Folding on Late Stages of the Secretion of α -Amylases from *Bacillus Subtilis*'. *FEBS Letters* 430 (3):385–89. [https://doi.org/10.1016/S0014-5793\(98\)00698-X](https://doi.org/10.1016/S0014-5793(98)00698-X).
- Stephenson, Keith, and Colin R. Harwood. 1998. 'Influence of a Cell-Wall-Associated Protease on Production of α -Amylase by *Bacillus Subtilis*'. *Applied and Environmental Microbiology* 64 (8):2875–81.
- Stern, Johanna, Sarah Moraïs, Raphael Lamed, and Edward A. Bayer. 2016. 'Adaptor Scaffoldins: An Original Strategy for Extended Designer Cellulosomes, Inspired from Nature'. *mBio* 7 (2):e00083-16. <https://doi.org/10.1128/mBio.00083-16>.
- Sukharnikov, Leonid O., Markus Alahuhta, Roman Brunecky, Amit Upadhyay, Michael E. Himmel, Vladimir V. Lunin, and Igor B. Zhulin. 2012. 'Sequence, Structure, and Evolution of Cellulases in Glycoside Hydrolase Family 48'. *Journal of Biological Chemistry* 287 (49):41068–77. <https://doi.org/10.1074/jbc.M112.405720>.
- Sun, Ye, and Jiayang Cheng. 2002. 'Hydrolysis of Lignocellulosic Materials for Ethanol Production: A Review'. *Bioresource Technology*, Reviews Issue, 83 (1):1–11. [https://doi.org/10.1016/S0960-8524\(01\)00212-7](https://doi.org/10.1016/S0960-8524(01)00212-7).
- Szilágyi, András, and Péter Závodszy. 2000. 'Structural Differences between Mesophilic, Moderately Thermophilic and Extremely Thermophilic Protein Subunits: Results of a Comprehensive Survey'. *Structure* 8 (5):493–504. [https://doi.org/10.1016/S0969-2126\(00\)00133-7](https://doi.org/10.1016/S0969-2126(00)00133-7).
- Tamaru, Yutaka, Shuichi Karita, Atef Ibrahim, Helen Chan, and Roy H. Doi. 2000. 'A Large Gene Cluster for the *Clostridium Cellulovorans* Cellulosome'. *Journal of Bacteriology* 182 (20):5906–10. <https://doi.org/10.1128/JB.182.20.5906-5910.2000>.
- Tan, Irene S., and Kumaran S. Ramamurthi. 2014. 'Spore Formation in *Bacillus Subtilis*'. *Environmental Microbiology Reports* 6 (3):212–25. <https://doi.org/10.1111/1758-2229.12130>.
- Tanaka, Tsutomu, and Akihiko Kondo. 2014. 'Cell-Surface Display of Enzymes by the Yeast *Saccharomyces Cerevisiae* for Synthetic Biology'. *FEMS Yeast Research*, October, n/a-n/a. <https://doi.org/10.1111/1567-1364.12212>.
- Tao, Ling, Eric C. D. Tan, Robert McCormick, Min Zhang, Andy Aden, Xin He, and Bradley T. Zigler. 2014. 'Techno-Economic Analysis and Life-Cycle Assessment of Cellulosic Isobutanol and Comparison with

- Cellulosic Ethanol and N-Butanol'. *Biofuels, Bioproducts and Biorefining* 8 (1):30–48. <https://doi.org/10.1002/bbb.1431>.
- Taylor, Edward J., Tracey M. Gloster, Johan P. Turkenburg, Florence Vincent, A. Marek Brzozowski, Claude Dupont, François Shareck, et al. 2006. 'Structure and Activity of Two Metal Ion-Dependent Acetylxyln Esterases Involved in Plant Cell Wall Degradation Reveals a Close Similarity to Peptidoglycan Deacetylases'. *The Journal of Biological Chemistry* 281 (16):10968–75. <https://doi.org/10.1074/jbc.M513066200>.
- Terwisscha van Scheltinga, A. C., S. Armand, K. H. Kalk, A. Isogai, B. Henrissat, and B. W. Dijkstra. 1995. 'Stereochemistry of Chitin Hydrolysis by a Plant Chitinase/Lysozyme and X-Ray Structure of a Complex with Allosamidin: Evidence for Substrate Assisted Catalysis'. *Biochemistry* 34 (48):15619–23.
- Tigunova, O. A., S. M. Shulga, and Y. B. Blume. 2013. 'Biobutanol as an Alternative Type of Fuel'. *Cytology and Genetics* 47 (6):366–82. <https://doi.org/10.3103/S0095452713060042>.
- Tracy, Bryan P., Shawn W. Jones, and Eleftherios T. Papoutsakis. 2011. 'Inactivation of σ E and σ G in *Clostridium Acetobutylicum* Illuminates Their Roles in Clostridial-Cell-Form Biogenesis, Granulose Synthesis, Solventogenesis, and Spore Morphogenesis'. *Journal of Bacteriology* 193 (6):1414–26. <https://doi.org/10.1128/JB.01380-10>.
- Tummala, Seshu B., Neil E. Welker, and Eleftherios T. Papoutsakis. 1999. 'Development and Characterization of a Gene Expression Reporter System for *Clostridium Acetobutylicum* ATCC 824'. *Applied and Environmental Microbiology* 65 (9):3793–99.
- 'United Nations Treaty Collection'. n.d. Accessed 15 September 2017. https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_n o=XXVII-7-d&chapter=27&clang=_en.
- Urban, Mark C. 2015. 'Accelerating Extinction Risk from Climate Change'. *Science* 348 (6234):571–73. <https://doi.org/10.1126/science.aaa4984>.
- Van Acker, Rebecca, Ruben Vanholme, Véronique Storme, Jennifer C. Mortimer, Paul Dupree, and Wout Boerjan. 2013. 'Lignin Biosynthesis Perturbations Affect Secondary Cell Wall Composition and Saccharification Yield in *Arabidopsis Thaliana*'. *Biotechnology for Biofuels* 6:46. <https://doi.org/10.1186/1754-6834-6-46>.
- Vanholme, Ruben, Brecht Demedts, Kris Morreel, John Ralph, and Wout Boerjan. 2010. 'Lignin Biosynthesis and Structure'. *Plant Physiology* 153 (3):895–905. <https://doi.org/10.1104/pp.110.155119>.
- Vazana, Yael, Yoav Barak, Tamar Unger, Yoav Peleg, Melina Shamshoum, Tuval Ben-Yehezkel, Yair Mazor, Ehud Shapiro, Raphael Lamed, and Edward A. Bayer. 2013. 'A Synthetic Biology Approach for Evaluating the Functional Contribution of Designer Cellulosome Components to Deconstruction of Cellulosic Substrates'. *Biotechnology for Biofuels* 6 (December):182. <https://doi.org/10.1186/1754-6834-6-182>.
- Vazana, Yael, Sarah Morais, Yoav Barak, Raphael Lamed, and Edward A. Bayer. 2010. 'Interplay between *Clostridium Thermocellum* Family 48 and Family 9 Cellulases in Cellulosomal versus Noncellulosomal States'. *Applied and Environmental Microbiology* 76 (10):3236–43. <https://doi.org/10.1128/AEM.00009-10>.

- Villares, Ana, Céline Moreau, Chloé Bennati-Granier, Sona Garajova, Loïc Foucat, Xavier Falourd, Bodo Saake, Jean-Guy Berrin, and Bernard Cathala. 2017. 'Lytic Polysaccharide Monooxygenases Disrupt the Cellulose Fibers Structure'. *Scientific Reports* 7 (January):40262. <https://doi.org/10.1038/srep40262>.
- Vital de Oliveira, Osmair. 2014. 'Molecular Dynamics and Metadynamics Simulations of the Cellulase Cel48F'. *Enzyme Research* 2014 (May):e692738. <https://doi.org/10.1155/2014/692738>.
- Vitikainen, M., H.-L. Hyyryläinen, A. Kivimäki, V.p. Kontinen, and M. Sarvas. 2005. 'Secretion of Heterologous Proteins in *Bacillus Subtilis* Can Be Improved by Engineering Cell Components Affecting Post-Translocational Protein Folding and Degradation'. *Journal of Applied Microbiology* 99 (2):363–75. <https://doi.org/10.1111/j.1365-2672.2005.02572.x>.
- Vollmer, W., and J. V. Höltje. 2001. 'Morphogenesis of *Escherichia Coli*'. *Current Opinion in Microbiology* 4 (6):625–33.
- Vollmer, Waldemar, Didier Blanot, and Miguel A. De Pedro. 2008. 'Peptidoglycan Structure and Architecture'. *FEMS Microbiology Reviews* 32 (2):149–67. <https://doi.org/10.1111/j.1574-6976.2007.00094.x>.
- Vollmer, Waldemar, and Alexander Tomasz. 2000. 'The pgdA Gene Encodes for a PeptidoglycanN-Acetylglucosamine Deacetylase in *Streptococcus Pneumoniae*'. *Journal of Biological Chemistry* 275 (27):20496–501. <https://doi.org/10.1074/jbc.M910189199>.
- Voronov-Goldman, M., O. Yaniv, O. Gul, H. Yoffe, O. Salama-Alber, M. Slutzki, M. Levy-Assaraf, et al. 2015. 'Standalone Cohesin as a Molecular Shuttle in Cellulosome Assembly'. *FEBS Letters* 589 (14):1569–76. <https://doi.org/10.1016/j.febslet.2015.04.013>.
- Voronov-Goldman, Milana, Maly Levy-Assaraf, Oren Yaniv, Gloria Wisserman, Sadanari Jindou, Ilya Borovok, Edward A. Bayer, Raphael Lamed, Linda J. W. Shimon, and Felix Frolow. 2014. 'Structural Characterization of a Novel Autonomous Cohesin from *Ruminococcus Flavefaciens*'. *Acta Crystallographica. Section F, Structural Biology Communications* 70 (Pt 4):450–56. <https://doi.org/10.1107/S2053230X14004051>.
- Wang, Zhenyu, Guangli Cao, Ju Zheng, Defeng Fu, Jinzhu Song, Junzheng Zhang, Lei Zhao, and Qian Yang. 2015. 'Developing a Mesophilic Co-Culture for Direct Conversion of Cellulose to Butanol in Consolidated Bioprocess'. *Biotechnology for Biofuels* 8 (June). <https://doi.org/10.1186/s13068-015-0266-3>.
- Warnick, Thomas A., Barbara A. Methé, and Susan B. Leschine. 2002. '*Clostridium Phytofermentans* Sp. Nov., a Cellulolytic Mesophile from Forest Soil'. *International Journal of Systematic and Evolutionary Microbiology* 52 (Pt 4):1155–60.
- Weadge, Joel T., John M. Pfeffer, and Anthony J. Clarke. 2005. 'Identification of a New Family of Enzymes with Potential O-Acetylpeptidoglycan Esterase Activity in Both Gram-Positive and Gram-Negative Bacteria'. *BMC Microbiology* 5 (August):49. <https://doi.org/10.1186/1471-2180-5-49>.

- Weizmann, C. 1915. Improvements in the bacterial fermentation of carbohydrates and in bacterial cultures for the same. 4845, issued 1915.
- Wely, Van, Karel H.m, Jelto Swaving, Roland Freudl, and Arnold J. M. Driessen. 2001. 'Translocation of Proteins across the Cell Envelope of Gram-Positive Bacteria'. *FEMS Microbiology Reviews* 25 (4):437–54. <https://doi.org/10.1111/j.1574-6976.2001.tb00586.x>.
- Wen, Zhiqiang, Nigel P. Minton, Ying Zhang, Qi Li, Jinle Liu, Yu Jiang, and Sheng Yang. 2017. 'Enhanced Solvent Production by Metabolic Engineering of a Twin-Clostridial Consortium'. *Metabolic Engineering* 39 (January):38–48. <https://doi.org/10.1016/j.ymben.2016.10.013>.
- Wen, Zhiqiang, Mianbin Wu, Yijun Lin, Lirong Yang, Jianping Lin, and Peilin Cen. 2014. 'Artificial Symbiosis for Acetone-Butanol-Ethanol (ABE) Fermentation from Alkali Extracted Deshelled Corn Cobs by Co-Culture of Clostridium Beijerinckii and Clostridium Cellulovorans'. *Microbial Cell Factories* 13:92. <https://doi.org/10.1186/s12934-014-0092-5>.
- Westers, Lidia, Helga Westers, Geeske Zanen, Haike Antelmann, Michael Hecker, David Noone, Kevin M. Devine, Jan Maarten van Dijl, and Wim J. Quax. 2008. 'Genetic or Chemical Protease Inhibition Causes Significant Changes in the Bacillus Subtilis Exoproteome'. *PROTEOMICS* 8 (13):2704–13. <https://doi.org/10.1002/pmic.200800009>.
- Wheeler, Tim, and Joachim von Braun. 2013. 'Climate Change Impacts on Global Food Security'. *Science* 341 (6145):508–13. <https://doi.org/10.1126/science.1239402>.
- Willson, Benjamin J., Katalin Kovács, Tom Wilding-Steele, Robert Markus, Klaus Winzer, and Nigel P. Minton. 2016. 'Production of a Functional Cell Wall-Anchored Minicellulosome by Recombinant Clostridium Acetobutylicum ATCC 824'. *Biotechnology for Biofuels* 9:109. <https://doi.org/10.1186/s13068-016-0526-x>.
- Wilson, David B. 2011. 'Microbial Diversity of Cellulose Hydrolysis'. *Current Opinion in Microbiology, Ecology and industrial microbiology* / Special section: Archaea, 14 (3):259–63. <https://doi.org/10.1016/j.mib.2011.04.004>.
- Xu, Chenggang, Ranran Huang, Lin Teng, Xiaoyan Jing, Jianqiang Hu, Guzhen Cui, Yilin Wang, Qiu Cui, and Jian Xu. 2015. 'Cellulosome Stoichiometry in Clostridium Cellulolyticum Is Regulated by Selective RNA Processing and Stabilization'. *Nature Communications* 6 (April). <https://doi.org/10.1038/ncomms7900>.
- Xu, Chenggang, Ranran Huang, Lin Teng, Dongmei Wang, Christopher L. Hemme, Ilya Borovok, Qiang He, et al. 2013. 'Structure and Regulation of the Cellulose Degradome in Clostridium Cellulolyticum'. *Biotechnology for Biofuels* 6:73. <https://doi.org/10.1186/1754-6834-6-73>.
- Xu, Qi, Michael G. Resch, Kara Podkaminer, Shihui Yang, John O. Baker, Bryon S. Donohoe, Charlotte Wilson, et al. 2016. 'Dramatic Performance of Clostridium Thermocellum Explained by Its Wide Range of Cellulase Modalities'. *Science Advances* 2 (2):e1501254. <https://doi.org/10.1126/sciadv.1501254>.

- Xu, Qi, Arjun Singh, and Michael E Himmel. 2009. 'Perspectives and New Directions for the Production of Bioethanol Using Consolidated Bioprocessing of Lignocellulose'. *Current Opinion in Biotechnology*, Energy biotechnology / Environmental biotechnology, 20 (3):364–71. <https://doi.org/10.1016/j.copbio.2009.05.006>.
- Xu, Tao, Yongchao Li, Zhili He, and Jizhong Zhou. 2014. 'Dockerin-Containing Protease Inhibitor Protects Key Cellulosomal Cellulases from Proteolysis in *Clostridium Cellulolyticum*'. *Molecular Microbiology* 91 (4):694–705. <https://doi.org/10.1111/mmi.12488>.
- Yang, Bin, Ziyu Dai, Shi-You Ding, and Charles E. Wyman. 2011. 'Enzymatic Hydrolysis of Cellulosic Biomass'. *Biofuels* 2 (4):421–49. <https://doi.org/10.4155/bfs.11.116>.
- Yang, Xiaorui, Mengmeng Xu, and Shang-Tian Yang. 2015. 'Metabolic and Process Engineering of *Clostridium Cellulovorans* for Biofuel Production from Cellulose'. *Metabolic Engineering* 32 (November):39–48. <https://doi.org/10.1016/j.ymben.2015.09.001>.
- Ye, R., L. P. Yang, and S. L. Wong. 1996. 'Construction of Protease Deficient *Bacillus Subtilis* Strains for Expression Studies: Inactivation of Seven Extracellular Proteases and the Intracellular LonA Protease'. *Proceedings of the International Symposium on Recent Advances in Bioindustry. The Korean Society for Applied Microbiology, Seoul, Korea*.
- Yi, Zhuolin, Xiaoyun Su, Vanessa Revindran, Roderick I. Mackie, and Isaac Cann. 2013. 'Molecular and Biochemical Analyses of CbCel9A/Cel48A, a Highly Secreted Multi-Modular Cellulase by *Caldicellulosiruptor Bescii* during Growth on Crystalline Cellulose'. *PLOS ONE* 8 (12):e84172. <https://doi.org/10.1371/journal.pone.0084172>.
- Yutin, Natalya, and Michael Y. Galperin. 2013. 'A Genomic Update on Clostridial Phylogeny: Gram-Negative Spore-Formers and Other Misplaced Clostridia'. *Environmental Microbiology* 15 (10):2631–41. <https://doi.org/10.1111/1462-2920.12173>.
- Zhang, Xiao-Zhou, Zuoming Zhang, Zhiguang Zhu, Noppadon Sathitsuksanoh, Yunfeng Yang, and Y.-H. Percival Zhang. 2010. 'The Noncellulosomal Family 48 Cellobiohydrolase from *Clostridium Phytofermentans* ISDg: Heterologous Expression, Characterization, and Processivity'. *Applied Microbiology and Biotechnology* 86 (2):525–33. <https://doi.org/10.1007/s00253-009-2231-1>.
- Zhao, Guangshan, Huazhong Li, Benson Wamalwa, Makiko Sakka, Tetsuya Kimura, and Kazuo Sakka. 2006. 'Different Binding Specificities of S-Layer Homology Modules from *Clostridium Thermocellum* AnCA, Slp1, and Slp2'. *Bioscience, Biotechnology, and Biochemistry* 70 (7):1636–41. <https://doi.org/10.1271/bbb.50699>.
- Zhao, Xuebing, Keke Cheng, and Dehua Liu. 2009. 'Organosolv Pretreatment of Lignocellulosic Biomass for Enzymatic Hydrolysis'. *Applied Microbiology and Biotechnology* 82 (5):815–27. <https://doi.org/10.1007/s00253-009-1883-1>.
- Zheng, Zongbao, Tao Chen, Meina Zhao, Zhiwen Wang, and Xueming Zhao. 2012. 'Engineering *Escherichia Coli* for Succinate Production from

- Hemicellulose via Consolidated Bioprocessing'. *Microbial Cell Factories* 11:37. <https://doi.org/10.1186/1475-2859-11-37>.
- Zingaro, Kyle A., and Eleftherios Terry Papoutsakis. 2013. 'GroESL Overexpression Imparts Escherichia Coli Tolerance to I-, N-, and 2-Butanol, 1,2,4-Butanetriol and Ethanol with Complex and Unpredictable Patterns'. *Metabolic Engineering* 15 (Supplement C):196–205. <https://doi.org/10.1016/j.ymben.2012.07.009>.
- Zipperle, G. F., J. W. Ezzell, and R. J. Doyle. 1984. 'Glucosamine Substitution and Muramidase Susceptibility in Bacillus Anthracis'. *Canadian Journal of Microbiology* 30 (5):553–59.
- Zuroff, Trevor R., Salvador Barri Xiques, and Wayne R. Curtis. 2013. 'Consortia-Mediated Bioprocessing of Cellulose to Ethanol with a Symbiotic Clostridium Phytofermentans/Yeast Co-Culture'. *Biotechnology for Biofuels* 6:59. <https://doi.org/10.1186/1754-6834-6-59>.