

Metabolic Engineering of *Clostridium saccharoperbutylacetonicum* for Improved Solvent Production



Alexander Bones
Department of Plant Sciences
St Edmund Hall

Thesis submitted for the degree of
Doctor of Philosophy
University of Oxford
Trinity Term 2022

Statement of Disruption Due to COVID-19

I request that the examiners take into account the extent of disruption caused by the COVID-19 pandemic. The closure of both the department of plant sciences and Biocleave Ltd prevented me from performing any experimental work during the first national lockdown. In addition, the reduced laboratory hours for several months after reopening severely limited the number of experiments I could perform. To give a brief overview of the work I had planned to complete, the experiments conducted in chapter six were intended to include additional mutants which had to be abandoned due to time constraints. Shuttle vectors to express the glucose uptake system and polyphosphate futile cycling system (described in chapter six) could not be developed. I had also planned on combining the strategies described in chapter six to make various strains with multiple manipulations. This would have required further cloning work, namely subcloning DNA into different shuttle vectors, which was not performed due to time constraints. Additionally, genomic integrations of formate dehydrogenase and pyruvate formate lyase were attempted in this thesis (chapter six), but ultimately were abandoned. In addition, the work presented in chapter five was originally intended to include a flux map of a mutant strain in chapter six, had one of the mutants generated an increased butanol titre compared to *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{1261T}. Given the issues with the cloning work presented in chapter six, it might have been possible to generate an improved strain and subsequently perform ¹³C-metabolic flux analysis of said strain, had the laboratory closures not occurred.

Abstract

In order to avert irreversible damage to the global climate, the global community has committed to reaching net zero carbon emission in the coming years. To meet this ambitious target, substantial changes will be needed. To minimise the disruption to people's lives, there is a need for renewable technologies which are compatible with existing infrastructure, such as biofuels and drop in chemical compounds. A compound which can fulfil both roles is *n*-butanol. *Clostridium* are natural butanol producers but have fallen out of use as they have been unable to compete with fossil fuel-based production methods.

The aim of this thesis was to improve the production of butanol in an asporogenic strain of *Clostridium saccharoperbutylacetonicum* N1-4. A systems scale approach was taken to improve butanol production. Combined analysis of *Clostridium* metabolism using flux balance analysis and ¹³C-metabolic flux analysis was used to guide metabolic engineering strategies, with the aim of increasing butanol production in asporogenic *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}.

Flux balance analysis was used to gain an understanding of *Clostridium* metabolism and to explore manipulations that could lead to increased butanol production in wild type *C. saccharoperbutylacetonicum* N1-4. Key features of *in silico* engineered strains were compared to experimental data and identified an increase in NADH generation and key butanol synthesising genes as targets for increasing butanol production. A second round of flux balance analysis identified further manipulations relevant to *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, mainly that the rate of glucose uptake appeared to be limiting butanol production. Simulations of strains with increased glucose uptake and butanol production suggested that ATP consuming enzymes would have to be engineered into the asporogenic strain to balance ATP.

Additional investigation was performed using ¹³C-metabolic flux analysis, which was able to resolve intracellular fluxes of asporogenic *C. saccharoperbutylacetonicum* N1-4. It also identified a feature of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} metabolism that was unexpected, that ATP was in excess to biomass synthesis requirements, resulting in a futile cycle. The results from this flux analysis confirmed the rationale of the flux balance analysis guided strategies.

In the final chapter, the developed strategies were incorporated into the asporogenic strain of *C. saccharoperbutylacetonicum* N1-4. These mutant strains were analysed in fermentations. While none of the strains produced more butanol than the parent strain, this work incorporated several novel approaches to increase butanol production in *Clostridium* and will serve as a starting point for future metabolic engineering work.

Acknowledgements

I would first and foremost like to thank my supervisors Prof Lee Sweetlove, Prof Mark Howarth, Dr Liz Jenkinson and Dr Mandy Nicolle for their guidance, support, and mentorship throughout this project. It has been a privilege to work with you all.

I would like to thank everyone in the Sweetlove group for their friendship and support throughout my DPhil, particularly Sanu Shameer for his guidance with flux balance analysis and general computational work and Andras Sandor for assistance with various laboratory experiments. I also like to thank Pedro Bota, Rita Ros and Jack Bateman for running the Oxford biochemistry laboratory, and Pedro again for training me on the GC-MS machine. I would also like to thank Nick Kruger, Edward Smith and George Dobson for their assistance planning, processing, and interpreting GC-MS data and ^{13}C -MFA experiments.

I also need to thank everyone at Biocleave Ltd for being so welcoming and sharing their wealth of *Clostridium* knowledge and techniques. I would particularly like to thank Josh Mountford for making and supplying freeze dried stocks of *Clostridium*, and Josh Dormon for training me in *Clostridium* culturing. I would also like to thank Dr Christopher Hills for his guidance on the enzyme assays and Dr Anna Baker for guidance on CLEAVE cloning. A big thank you to Verity Holby for training me on almost every aspect of molecular biology and *Clostridium* fermentations! I also need to thank every member of the analytical team for the huge amount of HPLC analysis they have done for me over the years, and Dr Nathan Fairhurst for training me on the HPLC.

I am grateful for the financial contributions of the BBSRC, Biocleave Ltd, and the University of Oxford, and for the training I received through the Interdisciplinary Bioscience DTP.

Thanks to Thomas Jackson-Soutter for his friendship and support throughout these last few years and the many late night runs to Oxford to pick up reagents. I also need to thank my family for their love and support throughout my DPhil. I couldn't have done this without you. Finally, special thanks go to Kayley Smith for her encouragement, support, and patience throughout these tumultuous few years.

Declaration of Authorship

I, Alexander James Bones, hereby acknowledge the contribution of data in this thesis by the following individuals. The quantification of fermentation products and glucose with HPLC was performed by the following employees of Biocleave/Green Biologics Ltd. The fermentation product data for figure 3.1 was provided by Emily Rose. The fermentation product data for strain *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} in figure 4.1 was provided by James Tyrrel. The fermentation product data for figures 5.1, 6.4, and 6.11 was provided by Dr Nathan Fairhurst. These individuals were not responsible for collecting the cell growth (OD_{600nm}) data presented in these figures. All further analyses besides from the quantification of fermentation products, including statistical analysis and data visualisation, were performed by me. Aside from these contributions, all remaining data collection and analysis presented in this thesis was performed by me.

Abbreviations

ABE – acetone butanol ethanol

ACE - allele coupled exchange

Ack – acetate kinase

Acp – acyl carrier protein

AdhE – bifunctional aldehyde/alcohol dehydrogenase

ADP – adenosine diphosphate

Apy - apyrase

ATP – adenosine triphosphate

Bcd – butyryl-CoA dehydrogenase

bcs – butyryl-CoA synthesis

BCPY - biomass coupled product yield

Bdh – butanol dehydrogenase

BDO – 1,4-butanediol

Buk – butyrate kinase

Cas – CRISPR associated protein

Cat – chloramphenicol acetyltransferase

CCR – carbon catabolite repression

CGM – *Clostridium* growth medium

Ciap - calf intestinal alkaline phosphatase

CiED - cipher of evolutionary design

CoA – Coenzyme A

COBRA – constraint-based reconstruction and analysis

Cre – catabolite response element

CRISPR - clustered regularly interspaced short palindromic repeats

Crt – crotonase

CTAB - cetrimonium bromide

CtfAB - butyrate-acetoacetate CoA-transferase

DNA – deoxyribonucleic acid

DNP - 2,4-dinitrophenol

DTNB - 5,5'-dithiobis-(2-nitrobenzoic acid)

ED – Entner-Doudoroff

EDTA - Ethylenediaminetetraacetic acid

EM – ensemble modelling

EMP – Embden-Meyerhof-Parnas

Etf – electron transfer flavoprotein

FBA – flux balance analysis

Fdh – formate dehydrogenase

FDXNRx – ferredoxin NAD(H) oxidoreductase

Flp – flippase

FRT - flippase-recognition target

FVA – flux variability analysis

GC-MS – gas chromatography mass spectrometry

gDW – grams dry weight

Hbd – 3-hydroxybutyryl-CoA dehydrogenase

HMF - hydroxymethylfurfural

Hyd – hydrogenase

HPLC – high-performance liquid chromatography

KEGG – Kyoto encyclopaedia of genes and genomes

KO – knockout

ME – malic enzyme

MES - 2-(*N*-morpholino)ethanesulfonic acid

MFA – metabolic flux analysis

MID – mass isotopomer distribution

MOMA - minimisation of metabolic adjustment

MSTFA - *N*-Methyl-*N*-(trimethylsilyl)trifluoroacetamide

NAD(H) – nicotinamide adenine dinucleotide

NADP(H) – nicotinamide adenine dinucleotide phosphate

NGAM – non-growth associated maintenance

NTG - *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine

OPPP - oxidative pentose phosphate pathway

PCR – polymerase chain reaction

PEP – phospho*eno*/pyruvate
pFBA – parsimonious enzyme-usage flux balance analysis
Pfl – pyruvate formate lyase
Pfl-AE – pyruvate formate lyase activating enzyme
Pfor – pyruvate: ferredoxin oxidoreductase
PHA - polyhydroxyalkanoates
PTA – phosphate acetyltransferase
Ptb – phosphate butyryltransferase
PTS – phospho*eno*/pyruvate dependent phosphotransferase system
qPCR – quantitative polymerase chain reaction
RAM - retrotransposition activated marker
RCM – reinforced *Clostridium* medium
RNA – ribonucleic acid
ROOM – regulatory on/off minimisation
RT-qPCR – quantitative reverse transcription PCR
SNP - single nucleotide polymorphism
SOC – super optimal broth with catabolite repression
sol – solvent
SSR – sum of squared residuals
TAE - Tris-Acetate-EDTA
TBDMSCl - *tert*-Butyldimethylsilyl chloride
TCA – tricarboxylic acid
Thl – thiolase
WT – wild type

Contents

Abstract	3
Acknowledgements	4
Abbreviations	6
Chapter 1: Introduction	14
1.1 Outline of the Thesis	14
1.2 Microbial Production of <i>n</i> -Butanol	16
1.3 ABE Fermentation	21
1.3.1 Metabolism and Physiology of ABE Fermentation	21
1.3.2 Genetics and Biochemistry of ABE Fermentation	26
1.3.3 Overview of Solventogenic Species	33
1.3.4 Substrate Range	35
1.4 Genetic Manipulation Tools in <i>Clostridium</i>	38
1.4.1 Pre-targeted Approaches and Initial Plasmid Systems	38
1.4.2 Group II Intron Systems	40
1.4.3 Allelic Coupled Exchange	40
1.4.4 CRISPR/Cas systems	43
1.4.5 Metabolic Engineering of <i>Clostridium</i> for <i>n</i> -Butanol Production	45
1.5 Computational Modelling of Metabolism	48
1.5.1 Flux Balance Analysis	49
1.5.2 Strain Design Algorithms	52
1.5.3 Kinetic Models of Metabolism	57
1.5.4 ¹³ C-Metabolic Flux Analysis	59
1.4 Objectives of My Research	60
Chapter 2: Materials and Methods	61
2.1 Materials	61
2.2 Bacteria Strains	61
2.3 Bacteria Growth Media	62
2.3.1 Lysogeny Broth (LB)	63
2.3.2 Super Optimal Broth with Catabolite Repression (SOC)	63
2.3.3 <i>Clostriduma</i> Growth Medium (CGM)	63
2.3.4 Reinforced <i>Clostridium</i> Medium (RCM)	63
2.3.5 Defined <i>Clostridium</i> Growth Medium (DCGM)	64
2.3.6 TYIR Medium	64
2.3.7 Sugar Solution	64
2.4 Glycerol Stocks	64

2.5 Freeze Dried Stocks	64
2.6 <i>C. saccharoperbutylacetonicum</i> N1-4 Growth Conditions	64
2.6.1 Growth from Stocks	65
2.6.2 Growth from Agar Plates.....	65
2.6.3 Batch Fermentations	65
2.7 Plasmids and Primers	66
2.8 Molecular Biology Techniques	70
2.8.1 Preparation of Chemically Competent <i>E. coli</i>	70
2.8.2 Heat Shock Transformation of <i>E. coli</i>	70
2.8.3 Plasmid Purification from <i>E. coli</i>	70
2.8.4 Restriction Digests.....	71
2.8.5 Polymerase Chain Reaction.....	72
2.8.6 Colony PCR	72
2.8.7 Agarose Gel Electrophoresis	73
2.8.8 DNA Extraction from Agarose Gel	73
2.8.9 Direct DNA Extraction from Enzymatic Reactions.....	73
2.8.10 Dephosphorylation of Digested Vector	74
2.8.11 DNA Ligation.....	74
2.8.12 Gibson Assembly	74
2.8.13 DNA Sequencing.....	74
2.8.14 <i>C. saccharoperbutylacetonicum</i> N1-4 Genomic DNA Extraction	75
2.8.15 Electrotransformation of <i>C. saccharoperbutylacetonicum</i> N1-4 Strains	75
2.8.16 CLEAVE Genome Editing.....	76
2.8.17 RNA Isolation	77
2.8.18 cDNA Synthesis.....	77
2.8.19 qPCR	78
2.9 HPLC Analysis of Culture Media	78
2.10 Enzyme Assays.....	79
2.10.1 Preparation of Crude Cell Extracts	79
2.10.2 Formate Dehydrogenase Assay.....	79
2.10.3 Chloramphenicol Acetyltransferase Assay	80
2.10.4 Bradford Assay	80
2.11 COBRA Modelling	80
2.11.1 Model Setup	80
2.11.2 External Rate Calculations.....	81
2.11.3 Flux Balance Analysis.....	82

2.11.4 Flux Variability Analysis	82
2.11.5 Strain Design Algorithms	83
2.12 ¹³ C-Metabolic Flux Analysis	85
2.12.1 ¹³ C Labelling Experiment	85
2.12.2 Preparation of Samples for Gas Chromatography-Mass Spectrometry.....	86
2.12.3 GC-MS of Supernatant Glucose and Amino Acid Fragments	86
2.12.4 Analysis of GC-MS Data	87
2.12.5 Flux Estimation from ¹³ C-Dataset	87
2.13 Statistical Analysis	87
Chapter 3 – Exploration of Metabolic Engineering Strategies to Increase Butanol Production in <i>Clostridium</i>	88
3.1 Aims of Work Described in Chapter	88
3.2 Introduction.....	88
3.2.1 Genome-Scale Metabolic Models for Metabolic Engineering	88
3.2.2 Metabolic Models of <i>Clostridium</i>	89
3.3 Results	89
3.3.1 FBA Simulations Optimised for Growth Do Not Reflect Experimental Data	89
3.3.2 Single Reaction Knockouts to Improve Butanol Production	105
3.3.3 Double Reaction Knockouts to Improve Butanol Production	112
3.3.4 Strain-Design Algorithms to Improve Butanol Production.....	118
3.3.5 Insertion of the Entner-Doudoroff Pathway	120
3.3.6 4-Hydroxybutyrate Shunt.....	129
3.4 Discussion	132
3.4.1 Role of Organic Acid Production in <i>Clostridium</i> Metabolism	132
3.4.2 NAD(P)H Formation Drives Butanol Production	134
3.4.3 Consolidated Metabolic Engineering Strategy.....	135
Chapter Four: Analysis of Metabolism of an Asporogenic Mutant Strain of <i>C. saccharoperbutylacetonicum</i> N1-4 using a Genome Scale Metabolic Model	137
4.1 Aims of Work Described in Chapter	137
4.2 Introduction.....	137
4.2.1 Asporogenic Strains of Solvent Producing <i>Clostridium</i>	137
4.3 Results	137
4.3.1 The <i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T} Mutant Produced More Butanol than the Wild Type in Batch Fermentations	137
4.3.2 Comparison of Predicted Flux Distributions of <i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T} and <i>C. saccharoperbutylacetonicum</i> N1-4	139

4.3.3 <i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T} is Pareto Optimised for Butanol Production.....	144
4.3.4 <i>In silico</i> Investigation of Increased Glucose Uptake.....	146
4.3.5 Viability of Proposed Engineering Strategy in <i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T}	152
4.4 Discussion.....	154
4.4.1 Metabolic Engineering Approaches to Increase Glucose Uptake	154
4.4.2 Increased ATP Consumption for Metabolic Engineering Applications.....	155
4.4.3 Implications for Engineering <i>C. saccharoperbutylacetonicum spo0A</i> ^{I261T}	157
4.5 Conclusion	157
Chapter 5: Metabolic Flux Analysis of an Asporogenic strain of <i>C. saccharoperbutylacetonicum</i> N1-4.....	159
5.1 Aim of Work in this Chapter.....	159
5.2 Introduction.....	159
5.2.1 Metabolic Flux Analysis	159
5.2.2 Synergy of Metabolic Flux Analysis and Flux Balance Analysis	159
5.3 Results	160
5.3.1 Parallel Labelling Experiments	160
5.3.2 Metabolic Flux Map of Asporogenic <i>C. saccharoperbutylacetonicum spo0A</i> ^{I264T}	164
5.3.3 pFBA Simulation with Fermentation Product Constraints	169
5.3.4 pFBA Simulations with Futile Cycle Constraints.....	177
5.4 Discussion.....	186
5.4.1 Metabolic Flux Map of <i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T}	186
5.4.2 Role of the Futile Cycle.....	187
5.4.3 Impact of ¹³ C-MFA Flux Map on the Proposed Metabolic Engineering Strategy.....	189
5.5 Conclusions.....	190
Chapter 6 - Experimental Implementation of Metabolic Engineering Interventions Suggested by Computational Modelling	191
6.1 Introduction.....	191
6.1.1 Aim of Work Described in this Chapter	191
6.1.2 Engineering Solventogenic <i>Clostridium</i> for <i>n</i> -Butanol Production.....	191
6.1.3 Engineering Approaches Implemented in This Chapter.....	192
6.2 Results	196
6.2.1 Development and Characterisation of <i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T} Δ <i>rex</i>	196
6.2.2 <i>pflA</i> , <i>pflB</i> & <i>fdh</i> Overexpressing Strain.....	200
6.2.3 Characterisation of Cac_3451 Promoter	205

6.2.4 <i>apy</i> Expressing Strain	206
6.2.5 Unsuccessful Genetic Manipulation Attempts.....	212
6.3 Discussion	214
6.4 Conclusion	219
7 General Discussion	220
7.1 Summary of Results.....	220
7.2 Evaluating the Metabolic Engineering Strategy in this Thesis	221
7.2.1 <i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{1261T} as a Chassis for Butanol Production .	221
7.2.2 Evaluating the Metabolic Engineering Attempts in this Thesis.....	223
7.2.3 Limitations of Constraint Based Modelling of <i>Clostridium</i>	225
7.3 Future Directions for Metabolic Engineering <i>Clostridium</i> for Butanol Production.....	228
References.....	231
Appendices	267
Appendix A	267
Appendix B	268
Appendix C	271
Appendix D	273
Appendix E.....	276
Appendix F.....	283
Appendix G	286
Appendix H	289
Appendix I	292

Chapter 1: Introduction

1.1 Outline of the Thesis

In an era when the environmental effects of human activity are being felt, there is a pressing need for a shift to more sustainable technologies. Sustainable energy technologies such as solar, wind and hydroelectric, have been a growing part of the UK energy mix (Egbumokei, 2021), but there is still a need for sustainable chemical and fuel production. While first generation biofuels, such as bioethanol, are available in many countries (Rass-Hansen *et al.*, 2007) and have the advantages of requiring little to no change in existing infrastructure, production of these compounds often involves competition for land which could otherwise be used for food production (Mohr and Raman, 2015). Second generation biofuels, produced from inedible cellulosic biomass, have been touted as the solution. These materials, which are often disposed of through burning, offer a ready supply of feedstock for microbial fermentations.

Biobutanol is an important compound in this context, in that it has uses both as a fuel and in chemical synthesis. It can be used as a drop-in chemical, replacing existing petrochemically derived sources, and can also be utilised as a fuel in combustion engines by blending with conventional petroleum, and has advantages compared to ethanol blending (Pugazhendhi *et al.*, 2019; Zhen, Wang and Liu, 2020). A benefit of butanol is an existing bacterial fermentation process using organisms in the genus *Clostridium*, which have historically been used to produce butanol. The use of these organisms for butanol production declined as petrochemical methods advanced and could produce butanol at a much lower cost (Jones and Woods, 1986). In recent years, volatile oil prices, pressure to reduce greenhouse gas emissions and interest in increased fuel and energy security have meant that *Clostridium* have received renewed interest for commercial scale butanol production (Jin *et al.*, 2011).

My project was a partnership with Biocleave Ltd, formerly Green Biologics Ltd, who produced sustainable butanol and acetone using these organisms. Manufacturing chemicals from fermentation at scale generally benefits from continuous fermentation, which can achieve higher productivity and reduced product inhibition due to removal of culture media (Jang and Choi, 2018; Rahimi *et al.*, 2019; Veza, Muhamad Said and Latiff, 2021; Mancini *et al.*, 2022). However, *Clostridium* form endospores during solvent production (Allcock and Reid, 1982; Long, Jones and Woods, 1984b, 1984a), which reduces culture productivity and leads to washout (Diallo, Kengen and López-Contreras, 2021). Asporogenic strains are therefore desirable. In *Clostridium*, however, the production of acetone and butanol is linked to the initiation of endospore formation at the gene-regulatory level (Ravagnani *et al.*, 2000), and many asporogenic mutants lack the ability to

form solvents (Clark, Bennett and Rudolph, 1989; Cornillot *et al.*, 1997; Harris, Welker and Papoutsakis, 2002; Atmadjaja *et al.*, 2019). Biocleave Ltd. have made a recent breakthrough, by developing an asporogenic strain of *Clostridium saccharoperbutylacetonicum* N1-4 which continues to produce solvents (Atmadjaja *et al.*, 2019), strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. The aim of my project was to develop and test metabolic engineering strategies to improve the butanol yield of the asporogenic *Clostridium* strain developed by BioCleave Ltd, following the design-build-test-improve model (**Fig. 1.1**). I first used *in silico* stoichiometric simulations of *C. saccharoperbutylacetonicum* N1-4 with constraints from batch fermentations to explore the metabolic network of this organism. This was combined with systematic searches for knockout targets and the strain design algorithms OptKnock (Burgard, Pharkya and Maranas, 2003) and OptForce (Ranganathan, Suthers and Maranas, 2010). I used the results of these simulations to develop a metabolic engineering strategy. Next, the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was characterised in batch fermentations, and the strain was subject to further computational modelling. This led to the generation of further metabolic engineering interventions. Next, I generated a ¹³C-flux map of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, which was further analysed using stoichiometric simulations, which confirmed the importance of the previously suggested engineering strategies. Finally, the metabolic engineering strategies were implemented in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, and the behaviour of the mutants was assessed in batch fermentations.

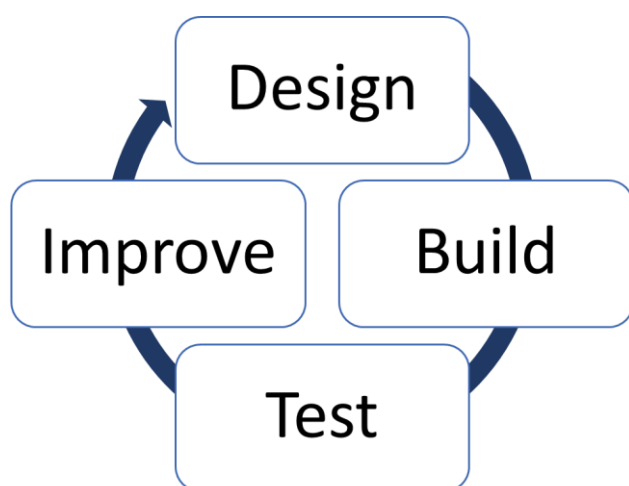


Figure 1.1 – Representation of metabolic engineering design-build-test-improve cycle. This is a cyclical process, generally involving multiple cycles before the production targets are achieved.

1.2 Microbial Production of *n*-Butanol

Biobutanol is an attractive alternative to bioethanol as a fuel and a chemical. As a fuel, this chemical provides superior properties when blended with gasoline, with a high calorific value and octane number (Zhen, Wang and Liu, 2020). Biological production of bulk chemicals such as ethanol and butanol can be achieved through microbial fermentations (Keasling *et al.*, 2021). To produce a given chemical, a host strain is selected, a pathway to produce the desired compound is constructed, and metabolic engineering is used to optimise production of the target compound (Nielsen, 2001). The well understood physiology, consistent and fast growth, as well as the numerous manipulation tools have made *E. coli* a popular host for metabolic engineering, particularly as an alternative to native butanol producers (Pontrelli *et al.*, 2018). *E. coli* has previously been used as a host for biobutanol production for these reasons (**Table 1.1**). Early efforts introduced the butanol forming pathway into *E. coli* from the native producer *C. acetobutylicum* (**Fig. 1.2**) (Atsumi *et al.*, 2008). Although this initially produced only low levels of butanol (13.9 mg L^{-1}) replacement of the *Clostridium* thiolase, which condenses two acetyl-CoA molecules to form acetoacetyl-CoA, with the native *E. coli* acetyl-CoA acetyltransferase increased biobutanol titres over three-fold (Atsumi *et al.*, 2008). Deletion of the enzymes lactate dehydrogenase, the aldehyde-alcohol dehydrogenase and fumarate reductase reduced the number of competing NADH consuming pathways and deleting phosphate acetyltransferase removed the unwanted side product acetate, but these interventions reduced biobutanol titres. The anaerobic regulator of the pyruvate dehydrogenase complex, *fnr*, was removed as well, which increased butanol titres considerably (Atsumi *et al.*, 2008).

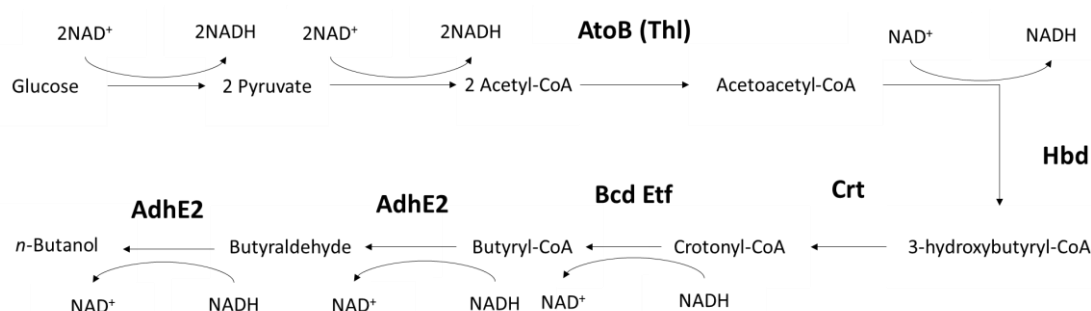


Figure 1.2 - Schematic representation of 1-butanol production in engineered *E. coli*. AtoB, acetyl-CoA acetyltransferase; Thl, acetoacetyl-CoA thiolase; Hbd, 3-hydroxybutyryl-CoA dehydrogenase; Crt, crotonase; Bcd, butyryl-CoA dehydrogenase; Etf, electron transfer flavoprotein; AdhE2, bifunctional aldehyde/alcohol dehydrogenase. Adapted from (Atsumi *et al.*, 2008).

<i>E. coli</i> Strain	Genes Expressed	Genes Knocked Out	Butanol Titre	Butanol Yield	Reference
BW25113	<i>atoB</i> (<i>E. coli</i>), <i>hbd</i> , <i>crt</i> , <i>bcd</i> , <i>etfAB</i> , <i>adhE2</i> (all <i>C. acetobutylicum</i>)	<i>adhE</i> , <i>ldhA</i> , <i>frdBC</i> , <i>fnr</i> , <i>pta</i>	552 mg L ⁻¹	0.028 g g ⁻¹ glucose (note, additional glycerol was added to the culture)	(Atsumi <i>et al.</i> , 2008)
BW25113	<i>atoB</i> (<i>E. coli</i>), <i>adhE2</i> , <i>crt</i> , <i>hbd</i> (all <i>C. acetobutylicum</i>), <i>fdh</i> (<i>C. boidinii</i>), <i>ter</i> (<i>T. denticola</i>)	<i>adhE</i> , <i>ldhA</i> , <i>frd</i>	30 g L ⁻¹	0.28 g g ⁻¹ glucose (in continuous culture with gas stripping)	(Shen <i>et al.</i> , 2011)
BW25113	<i>atoB</i> (<i>E. coli</i>), <i>adhE2</i> , <i>crt</i> , <i>hbd</i> (all <i>C. acetobutylicum</i>), <i>fdh</i> (<i>C. boidinii</i>), <i>ter</i> (<i>T. denticola</i>), <i>panK</i> (<i>E. coli</i>)	<i>ldhA</i> , <i>adhE</i> , <i>frdBC</i> , <i>pta</i>	18.3 g L ⁻¹	0.31 g g ⁻¹ glucose (with cysteine supplementation)	(Ohtake <i>et al.</i> , 2017)
BW25113	<i>atoB</i> , <i>hbd</i> , <i>crt</i> , <i>ter</i> , <i>adhE2</i> , <i>fdh</i> (<i>C. boidinii</i>)	<i>yqhD</i> , <i>ldhA</i> , <i>ackA-pta</i> , <i>adhE</i> , <i>frdBC</i> , <i>eutE</i> , <i>fdhF</i> , <i>hyc-hyp</i> , <i>mdh</i> , <i>maeB</i> , <i>yieP</i> , <i>stpA</i> , <i>yqeG</i> , <i>yagM</i>	20.08 g L ⁻¹	0.29 g g ⁻¹ glucose	(Dong <i>et al.</i> , 2017)
MG1655	<i>tesBT</i> (<i>Bacteroides thetaiotaomicron</i>), <i>car</i> (<i>Mycobacterium marinum</i>), <i>sfp</i> (<i>Bacillus subtilis</i>), <i>adh2</i> (<i>Saccharomyces cerevisiae</i>)	-	2.9 g L ⁻¹	0.31 g g ⁻¹ glucose	(Jawed <i>et al.</i> , 2020)

Table 1.1 – Summary of published work for *n*-butanol production in *E. coli*. Gram per gram yields of butanol are given, where possible, using information from the respective publication. In cases where amount of glucose consumed was not reported, the original yield as a percentage is given, which were given by the authors of the respective publications.

Other groups have achieved success with a slightly modified approach, by replacing the butyryl-CoA dehydrogenase enzyme complex with *Ter* from *Treponema denticola* and expressing formate dehydrogenase (Shen *et al.*, 2011). The latter is an enzyme which oxidises formate to form CO₂ with subsequent reduction of NAD⁺, which was used to increase NADH generation. As in the Atsumi *et al.* study, Shen *et al.* removed competing NADH consuming reactions to channel the additional NADH more effectively towards butanol production. This strategy produced butanol titres of 30 g L⁻¹ in fed-batch culture with gas stripping (Shen *et al.*, 2011).

Subsequent efforts involved integration of target genes into the chromosome, removing the need for antibiotics to maintain plasmid stability (Dong *et al.*, 2017). Integrated butanol producing genes were similar to those used by Shen *et al.*, which generated 2.62 g L⁻¹ butanol (Dong *et al.*, 2017). Through substantial metabolic engineering, involving further gene knockouts, both by deleting competing pathways and Tn5 transposon mutagenesis, and screening promoter libraries for stronger expression of introduced genes, the final strain was able to produce 20 g L⁻¹ butanol in fermentations (Dong *et al.*, 2017).

An alternative route to butanol synthesis in *E. coli* was reported using the native fatty acid synthesis pathway. Thioesterases were first used to convert butyryl-acyl carrier protein and butyryl-CoA into butyrate (Jawed *et al.*, 2016). In a subsequent study, carboxylic acid reductase (*car*), phosphopantetheinyl transferase (*sfp*) and alcohol dehydrogenase (*adh2*) were also introduced into the strain carrying thioesterase, which enabled butanol production from exogenous butyric acid (Jawed *et al.*, 2020). These studies exemplify the utility of *E. coli* as a host organism for metabolic engineering, and the wide range of manipulation tools and the speed at which genes can be inserted and knocked out has enabled substantial improvements in butanol titre to be achieved.

Aside from *E. coli*, investigators have also tried to produce butanol in other organisms. Butanol has been produced in the organism *Klebsiella pneumoniae* through expression of trans-2-enoyl-CoA reductase, NADH dependent butanol dehydrogenase A and NADH dependent butanol dehydrogenase B, producing *n*-butanol titres of 15 mg L⁻¹ with glycerol as substrate (Wang, Fan and Tan, 2014). An alternative pathway was also utilised, the 2-keto acid pathway, which produced *n*-butanol from 2-ketovalerate using the enzyme α -ketoisovalerate decarboxylase. When this enzyme was introduced, higher *n*-butanol titres of 28.7 mg L⁻¹ were reported (Wang, Fan and Tan, 2014). Decreasing the expression of genes involved in 1,3-propanediol production, which competes with butanol production for NADH, increased butanol titres by 83 % compared to the parental strain (M. Wang *et al.*, 2015). It should be noted that in all these studies, the exact culture and fermentation conditions vary, and the titres and yields reported are not easily comparable. In general, when

comparing different yields and titres, it is important to consider the medium used and the type of feedstock. The overall production process also needs to be considered, as the overall economics of the production process are influenced by factors such as reactor size and running volume, feedstock cost, and energy costs (Kumar *et al.*, 2012).

The ease of growth and manipulation have led to bacteria, particularly *E. coli*, being preferred hosts for metabolic engineering (Choi *et al.*, 2016), but there are downsides. Bacterial fermentations often utilise expensive sugars as substrates (Zeikus, 1980), which are estimated to contribute between 40 – 80 % of the total operating costs (Choi and Lee, 1999; Ahmad, Banat and Taher, 2020). These sugars can also compete with arable crops for land use (Desai and Atsumi, 2013). Glucose, for example, is one of the most widely used substrates for fermentation and is produced by hydrolysing starch sourced from various plant material (Nigam and Singh, 1995). To reduce the land use and reduce the cost of fermentations, photosynthetic production of chemicals using microalgae or cyanobacteria has also been explored, as these organisms can grow autotrophically (Gangl *et al.*, 2015).

The genes responsible for butanol synthesis from *C. acetobutylicum* were introduced into *Anabaena* sp. PCC 7120, using a heterocyst specific riboswitch system to generate butanol under the microaerobic conditions formed in heterocysts (Higo and Ehira, 2019). This strain produced low levels of butanol (1.14 mg L^{-1}) after ten days of cultivation. Replacing the *C. acetobutylicum* genes *bcd-etfAB* with *ter*, both of which catalyse formation of butyryl-CoA from crotonyl-CoA, but the former requires ferredoxin as an electron acceptor (Demmer *et al.*, 2017), while the trans-enoyl CoA-reductase encoded by *ter* only uses NADH (Shen *et al.*, 2011), doubled the *n*-butanol titre (Higo and Ehira, 2019).

Another cyanobacterium, *Synechococcus elongatus* PCC 7942, has been engineered to produce butanol through the introduction of a synthetic pathway (Lan and Liao, 2012). Acetyl-CoA was first converted to malonyl-CoA via acetyl-CoA carboxylase, followed by conversion to acetoacetyl-CoA catalysed by ketoacyl-ACP synthase III. Conversion of acetoacetyl-CoA to 3-hydroxybutyryl-CoA was mediated by acetoacetyl-CoA reductase, followed by reduction by (R)-specific enoyl-CoA hydratase. Trans-2-enoyl-CoA reductase was responsible for converting crotonyl-CoA to butyryl-CoA, and two NADPH utilising enzymes (NADP-dependent alcohol dehydrogenase and butyraldehyde dehydrogenase) converted butyryl-CoA to butanol. This system produced 29.9 mg L^{-1} butanol in 20 days (Lan and Liao, 2012). Further engineering by replacing the NADP-dependent alcohol dehydrogenase with a propanediol utilising CoA-dependent propionaldehyde dehydrogenase, increased butanol titres to 317 mg L^{-1} (Lan, Ro and Liao, 2013).

Further engineering efforts in *Synechocystis* PCC 6803 gave a substantial improvement in butanol titres, reaching cumulative titres of 4.8 g L⁻¹ (X. Liu *et al.*, 2019). To achieve this feat, the authors first constructed two separate butanol synthesis pathways, the *Clostridium* pathway and the reversed β -oxidation pathway and found that the former pathway produced higher butanol titres. Next, different enzymes responsible for butyraldehyde and butanol formation were screened, with the former improving butanol titres by use of an oxygen tolerant enzyme and the latter through selecting an NADPH specific dehydrogenase. Strains with different acetoacetyl-CoA synthesising pathways were next compared, with the malonyl-CoA route producing more butanol than the thiolase route. Mutated enzymes responsible for acetoacetyl-CoA reduction with higher affinity for NADPH and acetoacetyl-CoA were screened. Next, expression of genes was modulated by altering the 5'-regions and promoters of the introduced genes. Different genomic integration sites were also trialled, with integration of the genes *phaJ* and *ter* (whose products convert 3-hydroxybutyryl-CoA into crotonyl-CoA and the latter into butyryl-CoA, respectively) into the acetyl-CoA hydrolase (*ach*) site (which also knocked out the *ach* gene) producing the highest butanol titre. In addition, a phosphoketolase pathway was implemented for the generation of acetyl-phosphate, and a phosphate transacetylase was also expressed to aid conversion of acetyl-phosphate to acetyl-CoA. These strategies were combined into a single strain, which produced *n*-butanol titres of 836 mg L⁻¹ (X. Liu *et al.*, 2019). By optimising the light and NaHCO₃ feeding conditions, the in-flask titres of butanol reached a peak of 2.13 g L⁻¹ (X. Liu *et al.*, 2019).

The advantages of photosynthetic organisms are clear; however, they are still hampered by their low growth rates and low titres of butanol, the latter increasing the separation costs (Huang, Ramaswamy and Liu, 2014) which makes these organisms less cost competitive than traditional fermenting microorganisms. These issues have meant that commercial scale fuel production from algae has yet to be demonstrated (Foley, Beach and Zimmerman, 2011). Even with substantial engineering the titres achieved in the above studies are far from commercially viable levels (Kumar *et al.*, 2012). Therefore, these organisms are better suited to production of native, high value complex compounds such as carotenoids and certain polyunsaturated fatty acids (Gangl *et al.*, 2015).

The examples of butanol production outlined in this section highlight the standard approaches to metabolic engineering and common traits of successful approaches. Initial insertion of the desired pathway alone often produces only low titres of the target compound. Rounds of improvement are necessary by deleting enzymes encoding competing pathways, possible repressors of gene expression and altering expression of enzymes by promoter engineering. Metabolic engineering is

often an iterative process, requiring numerous changes to be made, with the phenotype being used to guide further rounds of improvement.

1.3 ABE Fermentation

1.3.1 Metabolism and Physiology of ABE Fermentation

Using heterologous production of a given compound is often advantageous, as the above examples have shown, and is a commonly used method when native producers of a compound generate only low titres. Certain bacteria in the anaerobic, spore-forming genus *Clostridium* are capable of butanol production, and were used historically for the production of butanol (Jones and Woods, 1986), with the most widely used organism being *C. acetobutylicum* (Yoo, Nguyen and Soucaille, 2020). In batch culture, these organisms undergo a biphasic pattern of growth; during exponential growth the main fermentation products are the organic acids acetate and butyrate (Allcock and Reid, 1982; Long, Jones and Woods, 1984a; Maddox et al., 2000), termed the ‘acidogenic’ phase of culture (**Fig. 1.3a**). The production of these acids is associated with a decrease in the pH of the culture media. With pH control or an appropriate pH buffer system, the bacteria then undergo a switch to produce the solvents acetone, butanol, and ethanol (Allcock and Reid, 1982; Holt, Stephens and Morris, 1984; Long, Jones and Woods, 1984a; Maddox et al., 2000), giving this fermentation process (ABE fermentation) its name (**Fig. 1.3b**). During the switch to solventogenesis, the bacteria stop dividing and undergo changes in morphology, from the characteristic rod-shaped vegetative cells to swollen ‘cigar-like’ clostridial forms (Long, Jones and Woods, 1984a), and become less motile (Allcock and Reid, 1982). Another feature of certain species is the accumulation of the carbohydrate storage molecule granulose during the solventogenic phase (Robson, Robson and Morris, 1974; Allcock and Reid, 1982; Long, Jones and Woods, 1984a, 1984b; Woolley and Morris, 1990). The bacteria also reassimilate the acids produced during the acidogenic phase, which coincides with an increase in the media pH (Allcock and Reid, 1982; Holt, Stephens and Morris, 1984; Long, Jones and Woods, 1984a). While the native production of butanol makes these organisms ideal hosts, a considerable downside to ABE fermentation is that production of solvents coincides with the formation of endospores (Allcock and Reid, 1982; Long, Jones and Woods, 1984b, 1984a). These are inactive forms of the bacteria, which reduce the culture productivity. Productivity is the concentration of butanol produced over time ($\text{g L}^{-1} \text{h}^{-1}$). Another downside of *Clostridium* is that they are obligate anaerobes, requiring specialist equipment for their cultivation.

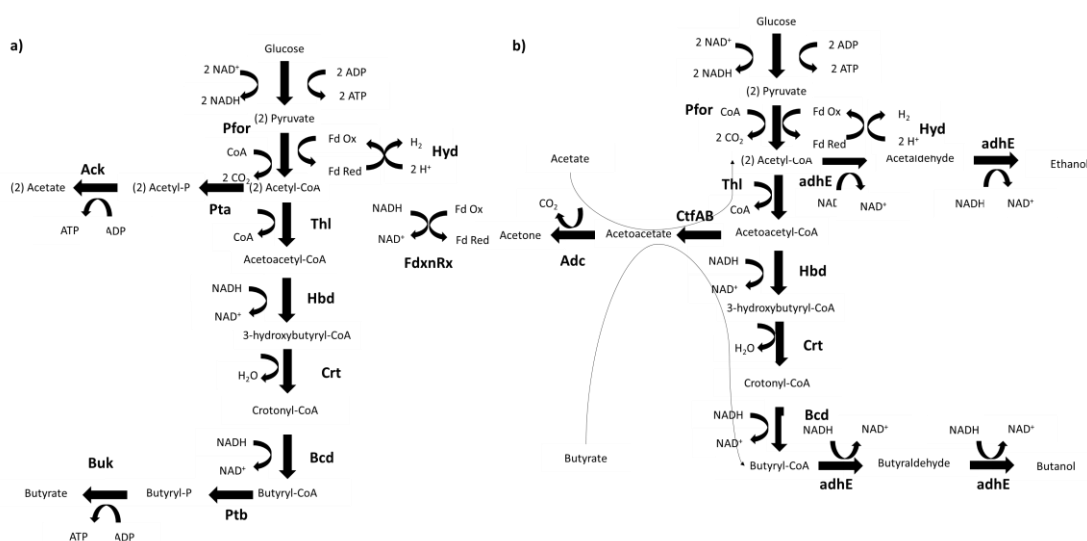


Figure 1.3 – Schematic of acetone-butanol-ethanol fermentation in solventogenic *Clostridia*. a) Acidogenic product profile occurs during exponential growth. Glucose is converted to the acids acetate and butyrate, which generate ATP through substrate level phosphorylation. b) Solventogenic production profile occurs during stationary phase of culture. Glycolytic reactions are not shown for brevity. Acetate and butyrate are taken up from the culture media and a coenzyme A moiety is transferred from acetoacetyl-CoA to acetate or butyrate. Butanol and ethanol are formed in this stage. The enzyme abbreviations refer to the following enzymes: Pfor – pyruvate ferredoxin oxidoreductase, Hyd – hydrogenase, Pta – phosphate acetyltransferase, Ack – acetate kinase, Thl – thiolase, Hbd – 3-hydroxybutyryl-CoA dehydrogenase, Crt – crotonase, Bcd – butyryl-CoA dehydrogenase, Ptb – phosphate butyryltransferase, Buk – butyrate kinase, adhE – aldehyde/alcohol dehydrogenase, CtfAB – butyrate-acetoacetyl CoA-transferase, FdxnRx – ferredoxin:NAD(H) oxidoreductase, CoA – coenzyme A. Adapted from (Jones and Woods, 1986).

Culture Conditions Impacting ABE Fermentation

There has been substantial effort to investigate conditions which trigger the onset of solventogenesis in *Clostridium* cultures. As discussed, ABE fermentation occurs in a two-stage process during batch culture, however this is not the case in continuous culture. The latter is arguably more advantageous for bulk chemical production, as it enables stable conditions to be maintained, more uniform product and reduced processing time (Maxon, 1955), therefore efforts have been made to find ways to induce this state in continuous fermentations.

The effect of pH on *Clostridium* cultures has been widely studied, with pH greater than 5.0 causing butyrate and acetate to be the major fermentation products in continuous (Bahl, Andersch and Gottschalk, 1982) and batch (Holt, Stephens and Morris, 1984) fermentations. Continuous cultures maintained between pH 4.3 and 4.5 had high titres of acetone and butanol (Bahl, Andersch and Gottschalk, 1982). Cultures maintained at pH 6.0 had virtually no solvent production (Monot, Engasser and Petitdemange, 1984; Girbal, Croux, *et al.*, 1995), and lactic acid accumulation was reported when the pH was maintained above 5.0 (Diez-Gonzalez, Russell and Hunter, 1995), although others have reported this only under iron limitation (Bahl *et al.*, 1986). Without pH control, cultures can undergo a phenomenon known as ‘acid crash’, where the production of acids

is continuous without solvent production (Maddox, Steiner, Hirsch, Wessner, Gutierrez, et al., 2000).

In batch culture, the decrease in pH is associated with the accumulation of the organic acids acetate and butyrate, followed by solventogenesis, and the impact of these compounds has also been studied. In the species *C. acetobutylicum*, supplementing the culture media with butyric acid while maintaining the pH at 4.3 (Bahl *et al.*, 1982) or above 4.8 (Martin *et al.*, 1983) increased the concentration of acetone and butanol proportionally to increasing butyrate concentrations. This was also reported in *C. saccharoperbutylacetonicum*, in non-controlled pH (Al-Shorgani *et al.*, 2012) and pH-stat fed batch culture (Tashiro *et al.*, 2004). The proposed trigger of solventogenesis was the concentration of undissociated butyric acid in the cytoplasm, which increased as cells entered solventogenesis (Terracciano and Kashket, 1986; Hüsemann and Papoutsakis, 1988).

In addition to butyrate, the effects of acetate on ABE fermentation have also been investigated. In *C. beijerinckii*, supplementing the culture media with sodium acetate increased total solvent titres (C. K. Chen and Blaschek, 1999), although another study reported no effect on acetone or butanol titres (Cho, Shin and Kim, 2012). In *C. saccharoperbutylacetonicum*, acetate addition did not cause changes in butanol production but did increase the specific production rate of acetone (Tashiro *et al.*, 2004). This was also reported in *C. acetobutylicum*, with acetone titres increasing with increasing concentrations of acetate (Martin *et al.*, 1983; Cho, Shin and Kim, 2012). Adding high concentrations of acetate and butyrate (100 mM each) induced solvent production in *C. beijerinckii*, even when the culture was maintained at neutral pH (Holt, Stephens and Morris, 1984) or pH 6.8 (George and Chen, 1983). The concentration of butyryl-phosphate increased prior to solventogenesis, which was not seen in the asporogenic M5 strain (Zhao *et al.*, 2005). Butyryl-phosphate has been proposed to act as a phosphodonator for regulatory proteins (Desai, Harris and Welker, 1999; Zhao *et al.*, 2005).

The minerals and salts in the culture media also have an impact on the production profile of ABE fermentations. In cultures of *C. acetobutylicum*, cell growth was dependent on the presence of Mg^{2+} , Fe^{2+} , K^+ , *p*-Aminobenzoic acid and biotin (Monot *et al.*, 1982). Increasing the amount of the latter two compounds in media increased both the total solvent titre and biomass by 62 % and 243 % respectively (Soni, Soucaille and Goma, 1987). Certain compounds, such as ammonium acetate, had a detrimental effect on solvent titres (Monot *et al.*, 1982). Increasing concentration of Fe^{2+} in *C. beijerinckii* cultures reduced butanol productivity, while hydrogen productivity increased (Wu *et al.*, 2017). In addition, the availability of nitrogen as ammonium ions appears to be important, as cultures of *C. beijerinckii* (Gottschal and Morris, 1981) and *C. acetobutylicum* (Andersch, Bahl and Gottschalk, 1982) grown under nitrogen limitation showed low butanol production in batch

(Gottschal and Morris, 1981) and continuous (Andersch, Bahl and Gottschalk, 1982) culture, although reducing culture pH from 6.0 to 5.2 and lower showed a slight increase in butanol titres (Andersch, Bahl and Gottschalk, 1982). In contrast, phosphate limitation has enabled solvent formation in continuous and batch cultures of *C. acetobutylicum* (Bahl, Andersch and Gottschalk, 1982).

Several other factors relating to culture conditions can affect solvent production. For example, accumulation of carbon monoxide (Thauer *et al.*, 1974), caused a reduction in biomass, acetone, acetate, and butyrate yields, while increasing the butanol yields in *C. acetobutylicum* (Kim *et al.*, 1984; Meyer, McLaughlin and Papoutsakis, 1985). In continuous culture, transient sparging with carbon monoxide triggered butanol production and a reduction in biomass, which was reversed after sparging (Meyer, Roos and Papoutsakis, 1986). A similar response to increasing carbon monoxide concentration in fermenter headspace has been reported in *C. pasteurianum*, which showed growth inhibition at CO concentrations > 5 % and increasing amounts of butanol, ethanol, and lactate (Dabrock, Bahl and Gottschalk, 1992). Adding butyrate and acetate to act as electron acceptors while gassing also increased acetone and butanol titres (Datta and Zeikus, 1985). These studies found also found that, in *C. acetobutylicum*, butyrate uptake can proceed via an acetone independent route under CO sparging (Datta and Zeikus, 1985; Meyer, Roos and Papoutsakis, 1986). The reason for increased butanol yield during carbon monoxide sparging is carbon monoxide is an inhibitor of ferredoxin hydrogenase activity. As shown in **Fig. 1.3**, the oxidation of NADH can be achieved through production of ethanol and butanol, but the ferredoxin hydrogenase is also a possible route. Here, NADH is oxidised with ferredoxin as the electron acceptor through a ferredoxin: NAD(H) oxidoreductase enzyme (Petitdemange *et al.*, 1976), with the reduced ferredoxin then being oxidised by a hydrogenase reaction, which generates molecular hydrogen (Adams, Mortenson and Chen, 1980).

The partial pressure of hydrogen in the fermenter also causes changes in the product distribution, with increasing pressure increasing the proportion of butanol in the total solvents (Yerushalmi, Volesky and Szczesny, 1985), and higher butanol titres have been reported (Doremus, Linden and Moreira, 1985). In *C. saccharoperbutylacetonicum*, the accumulation of biologically evolved gas in the fermenter headspace was important for solvent formation, as allowing the gas to escape prevented acetone and butanol production (Brosseau, Yan and Lo, 1986). The addition of viologen dyes, which can replace ferredoxin as an electron acceptor in the pyruvate-ferredoxin oxidoreductase reaction (Meyer, McLaughlin and Papoutsakis, 1985) and the hydrogenase reactions (Adams and Mortenson, 1984), increased the titre and yield of butanol when added to fermentation media, while reducing growth and biomass formation at both pH 4.5 and pH 6.5

controlled fermentations (Peguin *et al.*, 1994). In continuous culture, viologen addition caused production of butanol, while also increasing cellular ATP and NAD(P)H levels (Grupe and Gottschalk, 1992). Addition of neutral red, another electron carrier, induced a similar shift in product pattern to butanol and ethanol at pH 6.5 (Girbal, Vasconcelos, *et al.*, 1995), pH 6.4 and pH 6.3 (Rao and Mutharasan, 1987) controlled experiments. The hydrogenase has been proposed to contribute to maintaining the higher internal pH of the cell (Girbal, Vasconcelos and Soucaille, 1994), in addition to the NADH regenerating role. The ability of viologen dyes to increase the NAD(P)H levels may be due to stoichiometric limitations of intracellular electron acceptors, i.e., ferredoxin, to sufficiently balance redox. It has been proposed that the addition of viologen dyes allows the cells to bypass hydrogen formation, transferring electrons from pyruvate to methyl viologen, then from methyl viologen to NAD(H)⁺ (Peguin *et al.*, 1994), which is consistent with this hypothesis.

Carbon substrates vary in their degree of reduction, that is, the number of electrons transferred to oxygen when the substrate is fully oxidised to H₂O, CO₂ and (if nitrogen is present) NH₃ (Vallino, Hopkinson and Hobbie, 1996). This impacts the intracellular ATP levels and the NADH/NAD⁺ ratio, which has effects on the product pattern. For example, glucose-limited cultures tend to show high acetate and butyrate titres, with very low solvent titres (Monot *et al.*, 1982; Meyer, Roos and Papoutsakis, 1986; Charles L. Meyer and Papoutsakis, 1989; Girbal, Croux, *et al.*, 1995). In continuous culture, higher ATP/ADP and NADH/NAD⁺ ratios were associated with solventogenesis (Charles L. Meyer and Papoutsakis, 1989; C. L. Meyer and Papoutsakis, 1989), although when grown on mixtures of glucose and glycerol, solvent production occurred even at low ATP levels (Girbal and Soucaille, 1994). In the latter case, production of the alcohols butanol and ethanol was associated with a reduction in the specific production rate of hydrogen (Vasconcelos, Girbal and Soucaille, 1994). Interestingly, the specific activities of NADH-ferredoxin reductase and ferredoxin-NAD reductase were 8.8-fold lower and 9-fold higher, respectively, compared to cultures grown on glucose only (Vasconcelos, Girbal and Soucaille, 1994). Understanding the culture conditions which enable high butanol production provides a starting point for improving butanol production.

Strain Degeneration

During prolonged culture, *Clostridium* can reduce or lose their ability to produce solvents and have reduced ability to take up substrate, a phenomenon referred to as strain degeneration (Kutzenok and Aschner, 1952; Finn and Nowrey, 1959). Other triggers of degeneration include maintaining cultures at high pH (> 6.0), high dilution rates (0.16 h⁻¹) and maintaining feeding with ammonium sulphate (Barbeau, Marchal and Vandecasteele, 1988). In *C. acetobutylicum*, the pSOL1 megaplasmid contains the *sol*-operon, whose genes encode butyrate-acetoacetate CoA-

transferase (*ctfAB*), acetoacetate decarboxylase (*adc*) and aldehyde/alcohol dehydrogenase (*adhE1*), the first two being responsible for acetone production and the latter synthesising butanol and ethanol from butyryl-CoA and acetyl-CoA, respectively (Cornillot *et al.*, 1997; Nölling *et al.*, 2001; Fontaine *et al.*, 2002). Strains such as M5 have lost the pSOL1 megaplasmid and do not produce appreciable levels of acetone or butanol (Cornillot *et al.*, 1997). Unlike the M5 strain, degenerate *C. beijerinckii* NCIMB 8052 has not lost the acetone forming butyrate-acetoacetate CoA-transferase genes (Chih Kuang Chen and Blaschek, 1999a). This has also been observed in *C. saccharoperbutylacetonicum* (Kosaka *et al.*, 2007). Degenerate strains typically show high levels of hydrogen production in addition to acid production, likely a method of oxidising NADH in the absence of solvent formation (Guerrero *et al.*, 2021). Degenerate strains of *C. acetobutylicum* can regain their solvent producing capabilities by transformation with plasmids containing the *sol*-operon genes (Cornillot *et al.*, 1997), or by increasing the volume of culture transferred during subcultures (Finn and Nowrey, 1959). A degenerate strain of *C. saccharoperbutylacetonicum* was able to produce solvents when fractions from solid phase extractions isolated from solvent producing cultures were added to the culture, indicating a quorum-sensing mechanism may induce solventogenesis (Kosaka *et al.*, 2007). A strain of *C. beijerinckii* was isolated which was able to undergo nearly three times as many subcultures before culture death, while still producing solvents (Kashket and Cao, 1993). Strain degeneration is one of the more poorly understood aspects of *Clostridium* culture but presents a risk when deploying continuous culture.

1.3.2 Genetics and Biochemistry of ABE Fermentation

The biphasic nature of growth in *Clostridium* culture and sporulation are controlled by complex regulatory networks. To effectively engineer solvent production, an understanding of native genetic organisation, enzyme expression and activity is needed. As I have already discussed, alterations to the culture environment causes changes in the fermentation product pattern, cellular motility, and spore formation. As genetic tools available for *Clostridium* have improved in the last few decades, it has been possible to study the regulation of solvent formation in detail.

Acid Forming Pathways

The key role of biosynthesis of acids such as butyric and acetic acid in solventogenic *Clostridium* (Jones and Woods, 1986) is related to energy metabolism: the biosynthetic pathways for these acids generate ATP by substrate level phosphorylation. However, the biosynthesis of acids diverts carbon away from production of the desired solvents. This has made these acid biosynthesis pathways targets for downregulation or knockouts. The enzymes phosphotransacetylase (Pta), acetate kinase (Ack), phosphotransbutyrylase (Ptb) and butyrate kinase (Buk) are responsible for acetate (former two) and butyrate (latter two) production in a two-step process, involving

phosphorylation of the respective acyl-CoA, followed by transfer of the phosphate to ADP through substrate level phosphorylation (**Fig. 1.3**). In *C. acetobutylicum*, the genes responsible for acetate synthesis are on the chromosome and arranged as an operon (Boynton, Bennett and Rudolph, 1996), as are the genes responsible for butyrate synthesis (Walter *et al.*, 1993). The operonic organisation of *ptb-buk* has also been observed in *C. beijerinckii* (Chih Kuang Chen and Blaschek, 1999a).

In addition, to their genetic organisation, the acid forming enzymes have been purified which has enabled their enzymatic properties to be elucidated. In *C. acetobutylicum*, the acid forming enzymes are highly specific for their respective substrates, i.e., Ack has negligible activity with butyrate as a substrate (Winzer, Lorenz and Dürre, 1997), and Ptb and Buk have negligible activity with acetyl-CoA (Valentine and Wolfe, 1960) and acetate (Twarog and Wolfe, 1962; Hartmanis, 1987), as substrates, respectively. In batch culture of *C. acetobutylicum* and *C. beijerinckii*, the acetate forming operon is highly expressed during the early acidogenic phase of growth and the transition stage (Papoutsakis and Alsaker, 2005; Wang *et al.*, 2012). In contrast, this operon is expressed at similar levels throughout fermentation under both acidogenic and solventogenic conditions in continuous culture (Winzer, Lorenz and Dürre, 1997). The butyrate forming operon is expressed during the late exponential and early stationary phase in batch culture (Papoutsakis and Alsaker, 2005), while in continuous culture the butyrate forming operon shows the same expression pattern as the acetate forming operon (Janssen *et al.*, 2012).

The activity levels of these enzymes match the appearance of acids in fermentations, which all have higher specific activity in cultures of *C. acetobutylicum* maintained at pH 6.0 (acidogenic culture) than at pH 4.3 (solventogenic culture) (Andersch, Bahl and Gottschalk, 1983). The *C. beijerinckii* PTB enzyme has an optimum pH of 7.5, with decreasing specific activity as the pH of the culture decreases (Colby and Chen, 1992). ACK has a similar regulation mechanism, with a lower specific activity during solventogenesis in batch culture (Hartmanis, Klason and Gatenbeck, 1984), despite the encoding gene being transcribed at a constant level throughout fermentation (Winzer, Lorenz and Dürre, 1997). Aside from culture phase, the butyrate forming enzymes have reduced specific activities when intracellular ATP concentrations increase (Girbal and Soucaille, 1994), suggesting when ATP levels are already sufficient in the cell the production of further ATP is avoided.

The development of gene knockout systems in *Clostridium* has enabled further study of the function of the acid forming genes. Knockout of the *ack* genes in *C. acetobutylicum* led to delayed acetate production relative to the wild type, but peak acetate levels were similar, suggesting alternative enzymes exist (Green *et al.*, 1996; Kuit *et al.*, 2012). In contrast, a *ptb* knockout strain

of *C. acetobutylicum* failed to produce butyrate, suggesting a single copy of this enzyme was present (Lehmann, Radomski and Lütke-Eversloh, 2012a).

Solvent Forming Enzymes

Oxidation of carbon sources in living systems involves the formation of reducing equivalents, such as NAD(P)H. These reducing equivalents then themselves require oxidation to allow continued oxidation of carbon sources. In anaerobic organisms, reducing equivalents are oxidised by reducing pathway intermediates, such as pyruvate, into reduced end-products, such as ethanol (Müller, 2008). These end products tend to be excreted from the cell. In solventogenic *Clostridium*, there is an additional role of removing extracellular acids acetate and butyrate, which accumulate during acidogenesis. As discussed, the acid forming enzyme are transcribed and have higher specific activities during acidogenesis, while most of the solvent forming enzymes have the opposite pattern of transcription and activities, i.e., being higher at low pH. The solvent forming genes are also generally transcribed as operons and are grouped close together on the chromosome or megaplasmid, depending on the species.

The major operons involved in solvent formation are the *sol*-operon and the *bcs*-operon. The former varies between species, consisting of *ctfA*, *ctfB*, and *adhE1* in *C. acetobutylicum* (Cornillot *et al.*, 1997). In *C. saccharoperbutylacetonicum* and *C. beijerinckii*, the *sol*-operon contains *ald* (aldehyde dehydrogenase), *ctfA* and *ctfB* (CoA transferase subunits) and *adc* (acetoacetate decarboxylase) and is found on the genome rather than the megaplasmid (Chih Kuang Chen and Blaschek, 1999b; Y. Wang *et al.*, 2011; Poehlein *et al.*, 2014). The *bcs*-operon consists of the genes responsible for butyryl-CoA synthesis: *bhbd* (3-hydroxybutyryl-CoA dehydrogenase), *crt* (crotonase) and *bcd* (butyryl-CoA dehydrogenase) (Boynton, Bennett and Rudolph, 1996). The electron transfer flavoprotein complex (*etfA/B*) genes are not part of the *bcs*-operon but are located downstream of it (Boynton, Bennett and Rudolph, 1996). The EtfAB complex is involved in butyryl-CoA synthesis (Boynton, Bennett and Rudolph, 1996). Despite being responsible for the first step of acetone and butanol formation (acetoacetyl-CoA formation), thiolase is not part of either the *bcs* or *sol*-operons. In *C. acetobutylicum*, one thiolase gene (*thlA*) is transcribed as a monocistronic operon (Stim-Herndon, Petersen and Bennett, 1995), whereas the other (*thlB*) is transcribed in a polycistronic operon with *thlR* and *thlC* (Winzer *et al.*, 2000).

In ABE fermentation, acetyl-CoA and butyryl-CoA are reduced via acetaldehyde and butyraldehyde into ethanol and butanol, respectively. These reactions are catalysed in a two-step process by an aldehyde dehydrogenase then an alcohol dehydrogenase (**Fig. 1.3b**). In many solventogenic *Clostridium* there are multiple enzymes which possess alcohol dehydrogenase activity, aldehyde

dehydrogenase activity, or both. In *C. acetobutylicum*, the aldehyde/alcohol dehydrogenases encoded by *adhE1* and *adhE2* can catalyse the reduction of acetyl-CoA to ethanol and butyryl-CoA to butanol (Fischer, Helms and Durre, 1993; R. v. Nair, Bennett and Papoutsakis, 1994; Fontaine et al., 2002). In addition, there are two genes encoding NADH dependent butanol dehydrogenase on the chromosome (*bdhA*, *bdhB*) (Walter, Bennett and Papoutsakis, 1992). These two genes are adjacent on the chromosome but are not transcribed as an operon (K. A. Walter, Bennett and Papoutsakis, 1992). Both enzymes have higher specific activities for butyraldehyde than acetaldehyde, by 2.2-fold and 3.6-fold for BdhA and BdhB, respectively (Petersen et al., 1991).

In terms of *in vivo* contribution to product formation, *adhE1* is the major contributor to butanol and ethanol formation (Dai et al., 2016), followed by *bdhB* and *bdhA*, with *adhE2* showing the lowest contribution (Dai et al., 2016). Knockouts of both *bdh* genes did not show appreciable changes in solvent profiles (Cooksley et al., 2012), suggesting they do not greatly contribute to butanol formation. In terms of activity, both *adhE1* and *adhE2* have higher specific aldehyde dehydrogenase activity when acetyl-CoA is used as a substrate rather than butyryl-CoA, however they also both have higher specific alcohol dehydrogenase activity when butyraldehyde was used as a substrate rather than acetaldehyde (Inui et al., 2008).

In terms of regulation, most of the solvent forming enzymes are more active and more highly transcribed under the low pH found during solventogenesis, although there are exceptions. The acetone forming *adc* gene is not transcribed during exponential growth (Fischer, Helms and Durre, 1993) and is expressed prior to the *sol*-operon during *C. acetobutylicum* batch culture (Alsaker and Papoutsakis, 2005), however the encoded enzymes show similar responses to varying pH, with low to non-detectable specific activities at pH 6.0 and higher activities at pH 4.3 (Andersch, Bahl and Gottschalk, 1983). This is also the case for butyraldehyde and butanol dehydrogenase activities of crude *C. acetobutylicum* cell extracts (Andersch, Bahl and Gottschalk, 1983), with maximum specific activity of butanol dehydrogenase at pH 5.5 (Welch, Rudolph and Papoutsakis, 1989). *bdhA* is transcribed slightly earlier than *bdhB*, but after transcription initiation these genes have similar expression profiles (K. A. Walter, Bennett and Papoutsakis, 1992). The *adhE2* gene is only transcribed under 'alcohologenic' conditions (Fontaine et al., 2002; Yoo et al., 2016), i.e., feeding a reduced substrate (glycerol) in addition to glucose to increase intracellular NAD(P)H (Girbal, Croux, et al., 1995). This is consistent with this enzyme being the lowest contributor to alcohol formation in batch fermentation (Dai et al., 2016).

The exceptions to the trend of higher expression and specific activity at low pH are thiolase and butyryl-CoA dehydrogenase. Butyryl-CoA dehydrogenase is expressed less in the solventogenic phase of culture than in acidogenesis (Sullivan and Bennett, 2006; Wang et al., 2012). *thlA*

transcription decreases during the transition to solventogenesis, then increases again during solventogenesis (Winzer *et al.*, 2000). In *C. beijerinckii*, *thlA* is both highly expressed throughout batch fermentation (Wang *et al.*, 2012) and active from pH 5.5 to 7.0 (Wiesenborn, Rudolph and Papoutsakis, 1988). Thiolase is regulated at the post-translational level in the form of a redox switch (Kim *et al.*, 2015). The redox switch causes formation of a disulphide bond, which causes a conformational change in the catalytic cysteine loop, moving it away from the active site making the enzyme inactive (Kim *et al.*, 2015).

In summary, the acid forming enzymes are highly transcribed and active at high pH (around 6.0), where acid is the main fermentation product in continuous culture and during the exponential growth phase of batch culture. As the culture progresses and the pH drops to around 4.5, the encoding genes are transcribed less and the enzymes become less active, while the solvent forming genes are upregulated and have higher activity. Enzymes involved in both phases of culture, such as thiolase, are active and transcribed across a range of pH values.

Transcriptional Regulation and Endospore Formation

Solvent production is regulated by several DNA binding proteins, SolR, Rex, and Spo0A, the latter being responsible for both direct regulation of solvent forming genes and changing expression of other transcription factors. The *solR* gene, is located upstream of the *aad* gene in *C. acetobutylicum* and represses transcription of the genes *aad* and *adc* (Nair *et al.*, 1999). *solR* is more highly transcribed during acidogenic than solventogenic culture (Nair *et al.*, 1999), and its inactivation increased acetone and butanol titres in *C. acetobutylicum*, consistent with its proposed role as a *sol*-operon repressor (Nair *et al.*, 1999; Harris *et al.*, 2001).

The Rex protein is a transcriptional repressor and located upstream of the *bcs*-operon in *C. acetobutylicum* (Wietzke and Bahl, 2012). *rex* homologs have been identified in many *Clostridium*, but the predicted regulons vary between species (Zhang *et al.*, 2014). *rex* was found to bind promoter fragments of crotonase, thiolase, aldehyde-alcohol dehydrogenase, lactate dehydrogenase (Wietzke and Bahl, 2012; Zhang *et al.*, 2014), phosphate butyryltransferase, sulphite reductase operon and the NAD biosynthesis gene *nadA* (Zhang *et al.*, 2014). In *C. beijerinckii*, *rex* bound to *pflB* but this was not seen in *C. acetobutylicum* (Zhang *et al.*, 2014).

As in species of the genus *Bacillus* (Hoch, 1993), endospore formation in *Clostridium* is initiated by the master regulator Spo0A (Ravagnani *et al.*, 2000; Harris, Welker and Papoutsakis, 2002). Spo0A is expressed throughout fermentations (Harris, Welker and Papoutsakis, 2002; Sullivan and Bennett, 2006) but is most strongly transcribed at the end of exponential growth (Harris, Welker and Papoutsakis, 2002). This expression is dependent on the expression of multiple orphan

histidine kinases (Steiner *et al.*, 2011; Du *et al.*, 2021). Spo0A is regulated at the post-translational level by phosphorylation, which occurs under conditions such as nutrient limitation, on an aspartate residue on the 5' regulatory domain, which alters the expression of numerous genes (Dürre and Hollergschwandner, 2004). Sporulation specific sigma factors (σ^A , σ^E , σ^G , and σ^K) are transcribed prior to the formation of mature endospores (Santangelo *et al.*, 1998). Sporulation is considered detrimental from a process engineering perspective, as once cells commit to their sporulation program, they produce solvents during a short time frame before maturing into endospores and becoming inactive (Sillers *et al.*, 2008). As more cells undergo sporulation, the proportion of viable and solvent producing cells decreases, thus reducing the culture productivity (Qureshi, Lai and Blaschek, 2004).

There is a clear link between solvent formation and endospore formation (Allcock and Reid, 1982; Long, Jones and Woods, 1984b). The OA box binding motif, which is a consensus sequence bound by Spo0A, is present in the promoter sequences of several solvent forming genes (Ravagnani *et al.*, 2000) whose expression is increased when bound by Spo0A (Harris, Welker and Papoutsakis, 2002). *spo0A* knockout strains of solventogenic *Clostridium* have been developed, which had a sporulation negative phenotype but also produced virtually no solvents (Ravagnani *et al.*, 2000; Harris, Welker and Papoutsakis, 2002; Atmadjaja *et al.*, 2019; Yang *et al.*, 2020). One strain of asporogenic *C. beijerinckii* has been reported that maintained some butanol production, but titres were well below the wild type and production relied on pH control at pH 6.5 (Seo *et al.*, 2017).

Advances in RNA detection technologies have enabled researchers to gain a greater understanding of the range of genes under control of the Spo0A protein. Transcriptional analysis of the asporogenic *C. acetobutylicum* SKO1 strain, which has an inactivated *spo0A* gene, had lower expression of the *sol*-operon and the *adc* gene (Harris, Welker and Papoutsakis, 2002; Tomas *et al.*, 2003). A detailed microarray analysis of gene transcripts in this strain revealed a range of genes regulated by Spo0A, including butanol dehydrogenases *bdhAB*, many other transcriptional regulators, including *sigF*, *spollAB*, *spoVS* and *spoVAD*, sugar metabolism and motility genes (Tomas *et al.*, 2003). Reduced motility has been reported in a different *spo0A* knockout mutant of *C. acetobutylicum* (Yang *et al.*, 2020).

Aside from inactivation of the *spo0A* gene, another strain of *C. acetobutylicum*, the M5 strain, is an asporogenic strain which has lost the pSOL1 megaplasmid (Clark, Bennett and Rudolph, 1989), indicating some sporulation genes are on pSOL1 in this species (Tomas *et al.*, 2003). Like the SKO1 strain, reduced expression of *sol*-operon genes and alcohol dehydrogenases was reported in M5, along with other *spo* regulators and σ -factors (Tomas *et al.*, 2003). Both the M5 and SKO1 strains

have reduced expression of motility and chemotaxis genes during exponential growth, which increase during stationary phase (Tomas *et al.*, 2003).

Increasing the expression of Spo0A has given further insights into genes regulated by this protein. *spo0A* overexpression in *C. acetobutylicum* increased butanol titres and reduced the time to reach solventogenesis (Harris, Welker and Papoutsakis, 2002), while causing substantial changes in the transcriptome compared with the wild type strain. This included decreased expression of flagellar and glycolytic genes (Alsaker, Spitzer and Papoutsakis, 2004) and earlier transcription of the *sol*-operon (Harris, Welker and Papoutsakis, 2002). The wide range of genes regulated by Spo0A, particularly the *sol*-operon, shows the difficulties in generating asporogenic strains which are not deficient in solvent production, although with more advanced genome editing techniques this issue has been addressed. A recent study generated a strain of *C. saccharoperbutylacetonicum* which had a sporulation negative phenotype but retained solvent production, achieved by introducing a point mutation into the DNA binding domain of Spo0A (Atmadjaja *et al.*, 2019).

Aside from the regulation of sporulation and solventogenesis within cells, more recently attention has been paid to communication between cells and the potential role this has in regulating solventogenesis. Quorum sensing is a regulation mechanism which responds to changes in cell-population density and is known to be involved in sporulation in *B. subtilis* (Miller and Bassler, 2001). An RNA-Seq study in *C. beijerinckii* found that putative quorum sensing genes were upregulated in early acidogenic growth and downregulated during solventogenesis (Wang *et al.*, 2012). In *C. acetobutylicum*, eight putative quorum sensing systems were identified based on homology to known RRNPP (Rap, Rgg, NprR, PlcR and PrgsX) type regulators (Kotte *et al.*, 2020). RRNPP systems contain cytoplasmic receptors for small autoinducer peptides, which are imported through oligopeptide permeases (Neiditch *et al.*, 2017). The identified RRNPP systems consisted of regulator genes flanked by quorum sensing peptide precursors. The regulator QsrB was found to delay the onset of sporulation and reduced solvent formation (Kotte *et al.*, 2020). This regulator likely acts as a transcriptional repressor or through increasing expression of another repressor, while binding of its cognate peptide (QssB) causes inactivation, as overexpressing QssB increased solvent titres (Kotte *et al.*, 2020). Another transcriptional study investigated how butanol impacts the *C. beijerinckii* transcriptome found increased expression of an Agr-system after butanol addition (Patakova *et al.*, 2021), and the authors suggested that the cultures response to butanol shock may be cell density dependent.

While the fundamental genetics, biochemistry and regulation of ABE fermentation have been known for some time, recent advances in transcriptomic technologies have revealed a complex

regulatory network involving mechanisms such as quorum sensing which have previously been overlooked. Further understanding of the interplay between sporulation and solventogenesis will no doubt enable more efficient solvent producing strains to be developed.

1.3.3 Overview of Solventogenic Species

There are a range of solventogenic species in the genus *Clostridium*. The first and most widely used species of *Clostridium* used in industrial scale ABE fermentation was *C. acetobutylicum* (Jones and Woods, 1986), which was patented by Chaim Weizmann and used to produce acetone for cordite production (Weizmann, 1919). With ABE fermentation spreading globally, numerous strains were isolated by various commercial solvent companies, and these organisms tended to be named as 'species' based on their fermentation product pattern (Jones and Woods, 1986), e.g. certain species of solventogenic *Clostridium* reduce acetone to form isopropanol (George, Johnson and Moore, 1983), and were called *C. beijerinckii* (Jones and Woods, 1986). It was only with the advancement of 16S rRNA sequencing and DNA fingerprinting that these organisms were clearly assigned species (Keis, Shaheen and Jones, 2001) and the taxonomy of these organisms was elucidated (Keis *et al.*, 1995). The organism previously called *C. butanologenum* was found to be identical to *C. beijerinckii*, *C. kaneboi* was found to be identical to *C. acetobutylicum*, and numerous industrial strains previously identified as *C. acetobutylicum* were found to be either *C. beijerinckii* or *C. saccharobutylicum* (Keis, Shaheen and Jones, 2001). In addition, the species formerly known as *C. butylicum* was found to possess similar phenotypic characteristics and shared a high degree of DNA homology with *C. beijerinckii* (George, Johnson and Moore, 1983).

Numerous species in the genus *Clostridium* have had their genomes sequenced, which has helped to understand the relationships within this genus (**Fig. 1.4**). The *C. acetobutylicum* genome was sequenced in 2001 and was the first published genome sequence of a solventogenic species (Nölling *et al.*, 2001). This allowed identification of genes which were likely acquired through lateral gene transfer, which included various redox proteins, from archaea, bacteria, and eukaryotes (Nölling *et al.*, 2001). The draft *C. saccharoperbutylacetonicum* N1-4 genome was published in 2013 (del Cerro *et al.*, 2013), and the full, closed sequence was published a year later (Poehlein *et al.*, 2014). An unusual feature of this organism was the presence of choline in the teichoic acids of its cell walls, a feature rare in bacteria (del Cerro *et al.*, 2013). Sequencing also revealed that the pSOL1 plasmid does not harbour any genes responsible for solvent production (Gu *et al.*, 2019). *C. pasteurianum* is a solventogenic species which can utilise crude glycerol (Sandoval *et al.*, 2015) to produce butanol and 1,3-propanediol (Dabrock, Bahl and Gottschalk, 1992; Luers *et al.*, 2006). The genomes of strains ATCC 6013 (Rotta *et al.*, 2015) and DSM 525 (Poehlein *et al.*, 2015) have been published.

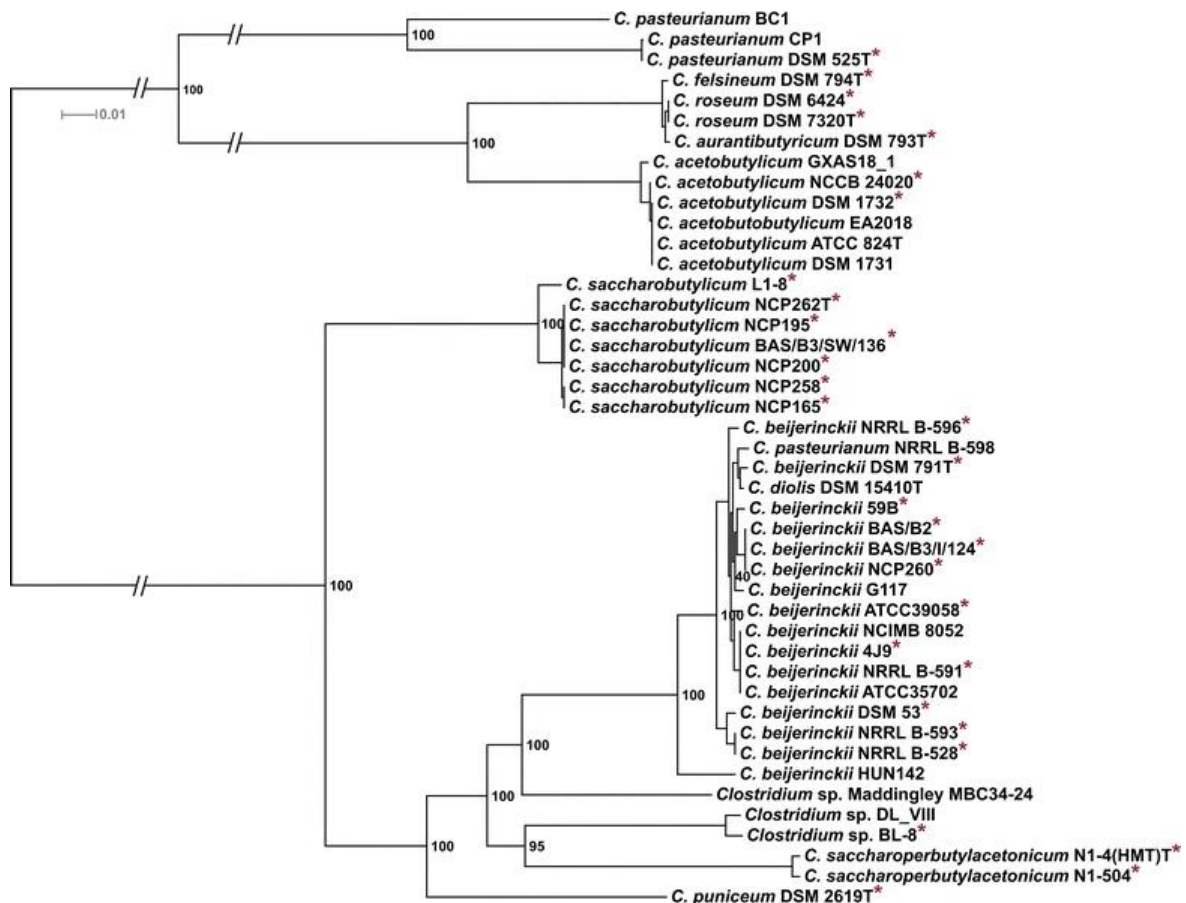


Figure 1.4 – MLSA tree of 44 sequenced solventogenic clostridia: a maximum likelihood tree of 44 solventogenic clostridial genomes was inferred with 500 bootstraps. Figure is taken without modification from (Poehlein *et al.*, 2017). Red asterisks indicate genome sequenced in the paper by Poehlein *et al.*, and ‘T’ indicates type strains. Source: <https://biotechnologyforbiofuels.biomedcentral.com/articles/10.1186/s13068-017-0742-z>. All credit goes to the original authors: Anja Poehlein, José David Montoya Solano, Stefanie K. Flitsch, Preben Krabben, Klaus Winzer, Sharon J. Reid, David T. Jones, Edward Green, Nigel P. Minton, Rolf Daniel and Peter Dürre. Reproduction of this figure was granted by the Creative commons license 4.0 (<https://creativecommons.org/licenses/by/4.0/legalcode>).

In 2008, the *C. kluyveri* genome was published (Seedorf *et al.*, 2008). An unusual species, *C. kluyveri* is a member of *Clostridium* which can ferment a range of substrates, including ethanol, acetate, mixtures of ethanol and propanol, isopropanol, butanol, ethylene glycol, and solely propanol (Kenealy and Waselefsky, 1985). The published genome sequence generated insights into the metabolism of this strain. The genes common to other ABE fermenting organisms, including alcohol dehydrogenase, aldehyde dehydrogenase, thiolase, 3-hydroxybutyryl-CoA dehydrogenase, crotonase and butyryl-CoA dehydrogenase, are found in the genome (Seedorf *et al.*, 2008). In addition to the saccharolytic *Clostridium*, some members of the genus *Clostridium* are capable of fermenting gases, termed ‘acetogenic’ metabolism, such as *C. carboxidivorans* P7^T (Paul *et al.*, 2010) and *C. autoethanogenum* (Humphreys *et al.*, 2015), with the former able to produce butanol from carbon monoxide (Fernández-Naveira *et al.*, 2016).

A wide range of solventogenic *Clostridium* exist, and given the different product patterns

and substrate ranges, for industrial production strains can be selected based on locally available materials and desired chemical production.

1.3.4 Substrate Range

A major barrier to the wide use of biofuels and production of chemicals from renewable feedstocks is the cost of the fermentation substrate, which has been estimated to contribute 26 % of the total operating costs of fermentation processes (Koutinas, Wang and Webb, 2004). Using cheap, locally available substrates is key to making these processes commercially viable (Keasling *et al.*, 2021). One potential substrate is lignocellulosic waste. This feedstock is widely available and is currently disposed of by burning (Anwar, Gulfranz and Irshad, 2014), making it an ideal substrate which does not compete with crops for human or animal feed. Currently, deployment is limited by the lack of low-cost methods of depolymerising lignocellulose into simple oligo- and monosaccharides (Brodeur *et al.*, 2011). Lignocellulose consists of cellulose, lignin, and hemicellulose, which when hydrolysed produces glucose, pentoses, and glycolytic acid (Chen, 2015). The ratio of these components varies with the type of starting material (Olsson and Hahn-Hägerdal, 1996). Thus, having an organism which can utilise a range of mono-, di- and oligosaccharides is important. Many species of *Clostridium* can grow on a range of sugars making them well-suited for lignocellulosic fermentation. For example, *C. acetobutylicum* ATCC 824 can produce butanol when grown on xylose (Mes-Hartree and Saddler, 1982; Ounine *et al.*, 1985; Hu *et al.*, 2011), arabinose (Mes-Hartree and Saddler, 1982; Servinsky *et al.*, 2012), mannose (Mes-Hartree and Saddler, 1982), galactose (Mes-Hartree and Saddler, 1982), fructose (Voigt, Bahl and Fischer, 2014) and cellobiose (Buendia-Kandia *et al.*, 2018), to name a few.

The range of sugars that *Clostridium* can metabolise makes them ideal for utilising lignocellulose and other plant waste material as a fermentation substrate, while producing solvents (**Table 1.2**). Aside from plant biomass, other waste materials that have been used as fermentation substrates include high-sucrose-content beverages (Raganati *et al.*, 2015), the biodegradable fraction of municipal waste (Farmanbordar, Karimi and Amiri, 2018) and Kraft liquor hydrolysate (Kudahettige-Nilsson *et al.*, 2015). A theoretical case study of butanol production using cheese-whey, agricultural residues, and lignocellulosic biomass (of *Miscanthus*) indicated that butanol production using these substrates could meet around 60 % of the biofuel demand in the Campania region of Italy (Procentese *et al.*, 2017). Another comparison of various fermentation parameters in a model of ABE fermentation plants showed that lignocellulosic feedstocks, combined with a low number of high working-capacity fermenters offered the lowest production costs of biobutanol (Kumar *et al.*, 2012).

Organism	Substrate	Pre-treatment	Solvent Titre	Reference
<i>C. acetobutylicum</i> DSM 1731	48 g L ⁻¹ unhydrolysed potato starch	None	8 g L ⁻¹ butanol	(Gutierrez <i>et al.</i> , 1998)
<i>C. tyrobutyricum</i> ATCC 25755	Brown alga (<i>Saccharina japonica</i>)	Acid hydrolysate, ultrasonic treatment	6.65 g L ⁻¹ butanol	(Fu <i>et al.</i> , 2021)
<i>C. saccharoperbutylacetonicum</i> N1-4	10 % algae from wastewater	Acid/base, Enzyme hydrolysis	8 g L ⁻¹ butanol	(Ellis <i>et al.</i> , 2012)
<i>C. saccharoperbutylacetonicum</i> N1-4	Cassava starch/chips	Heating, enzyme hydrolysis	21 g L ⁻¹ ABE (starch), 19.4 g L ⁻¹ ABE (chips)	(Thang, Kanda and Kobayashi, 2010)
<i>C. tyrobutyricum</i> Δ ack	Soybean hull hydrolysate with methyl viologen	Acid, enzyme hydrolysis	15.7 g L ⁻¹ butanol	(Yu <i>et al.</i> , 2015)
<i>C. beijerinckii</i> P260	Wheat straw hydrolysate	Sulphuric acid and enzymatic hydrolysis	25 g L ⁻¹ total ABE	(Qureshi, Saha and Cotta, 2007)
<i>C. beijerinckii</i> P260	Yellow top presscake	Defatting with hexane, sulphuric acid hydrolysis, enzyme hydrolysis	20.2 g L ⁻¹ total ABE, 9.44 g L ⁻¹ butanol	(Qureshi <i>et al.</i> , 2020)
<i>C. beijerinckii</i> P260	Barley straw hydrolysate	Overliming	26.64 g L ⁻¹ ABE	(Qureshi, Saha, Dien, <i>et al.</i> , 2010)
<i>C. beijerinckii</i> P260	Corn stover hydrolysate	Overliming	26.27 g L ⁻¹ ABE	(Qureshi, Saha, Hector, <i>et al.</i> , 2010)
<i>Clostridium beijerinckii</i> NCIMB 8052	Giant miscanthus	Hot water treatment	11.8 g L ⁻¹ ABE	(Zhang and Ezeji, 2014)
<i>C. beijerinckii</i> NCIMB 8052	Green seaweed (<i>Ulva lactuca</i>)	Heating	4.5 g L ⁻¹ total ABE	(van der Wal <i>et al.</i> , 2013)
<i>C. beijerinckii</i> BA101	Soy molasses	Spray dried with 2 % tri-calcium phosphate	10.7 g L ⁻¹ total ABE	(Qureshi, Lolas and Blaschek, 2001)
<i>C. beijerinckii</i> BA101	Liquified corn starch	Heat with α - amylase	18.4 g L ⁻¹ total ABE	(T. Ezeji, Qureshi and Blaschek, 2007a)

Table 1.2 – Summary of published work producing *n*-butanol in *Clostridia* using plant biomass as feedstock.

Use of lignocellulosic biomass, while desirable as a cheap fermentation substrate, requires an initial hydrolysis step to release shorter oligosaccharides and monosaccharides more easily metabolised by *Clostridium*. Upon hydrolysis, lignocellulose releases inhibitory compounds, referred to as lignocellulosic inhibitors, which include phenolic compounds, which reduce growth by depleting cellular pools of reducing power and ATP (Piotrowski *et al.*, 2014). For example, *C. beijerinckii* BA101 produced around half as much butanol when grown on corn fibre hydrolysate compared with a mixture of sugars of equivalent composition (T. Ezeji, Qureshi and Blaschek, 2007b). In *C. saccharoperbutylacetonicum*, addition of syringaldehyde and furfural to the fermentation media reduced butanol titres (Yao *et al.*, 2017). In *C. beijerinckii*, syringaldehyde concentrations greater than 0.5 g L⁻¹ reduced the peak biomass, along with higher butyric and acetic acid titres and lower solvent titres (Richmond, Ujor and Ezeji, 2012). The common phenolics vanillin, *p*-coumaric acid, syringaldehyde and ferulic acid delayed the onset of solvent production and led to lower butanol titres in *C. acetobutylicum*, even at low concentrations (0.2 g L⁻¹) (Liu *et al.*, 2020).

While these compounds are detrimental above threshold levels, low concentrations of these compounds can have positive effects. In *C. saccharoperbutylacetonicum*, the compounds ferulic acid, coumaric acid and hydroxymethylfurfural (HMF) increased butanol titres when 0.8 g L⁻¹ of the respective compounds were present (Yao *et al.*, 2017). Similarly, low concentrations (0.2 g L⁻¹) of syringaldehyde increased the growth rate of *C. beijerinckii* NCIMB 8052 (Richmond, Ujor and Ezeji, 2012). In *C. acetobutylicum* ATCC 824, addition of 1 g L⁻¹ furfuraldehyde increased titres of acetone and butanol (H. Liu *et al.*, 2019). *C. saccharoperbutylacetonicum* resisted 3 g L⁻¹ of furfural or HMF in fermentation and produced similar peak solvent titres as fermentations without inhibitor, although growth and solvent production were delayed (Yao *et al.*, 2017). Transcriptomic analyses have revealed the transcriptional response to these inhibitors. In *Clostridium acetobutylicum*, phenolic compounds caused an increase in expression of glycolytic genes and various heat shock proteins (Luo *et al.*, 2020).

Given the detrimental effects of lignocellulosic inhibitors, methods have been developed to try to overcome the toxicity of these compounds. Detoxifying various common inhibitory compounds derived from lignocellulosic hydrolysis including ferulic acid, syringaldehyde and coumaric acid with peroxidase from *Coprinus cinereus* reduced growth inhibition of *C. beijerinckii* NCIMB 8052 and restored butanol titres (Cho *et al.*, 2009). Another compound, allopurinol, reduced growth inhibition caused by furfural in the same strain by inhibiting xanthine oxidoreductase (Ujor *et al.*, 2015). Media supplementation with proline also reduced the inhibitory effects of formic-, ferulic-, *p*-coumarinic acid and syringaldehyde in *C. acetobutylicum* (Liao *et al.*, 2019). Another technique

was overexpression of heat shock proteins. Overexpression of *groESL* in *C. tyrobutyricum* ATCC 25755, improved sugar uptake from multiple hydrolysates and increased butyrate titres (Suo *et al.*, 2018; Fu *et al.*, 2021) while overexpression of *dnaK* from *Deinococcus wulumuqiensis* showed improved butanol tolerance when utilising *S. japonica* biomass as a substrate (Fu *et al.*, 2021).

Another problem arising from use of lignocellulosic feedstocks is carbon catabolite repression (CCR). This occurs when multiple sugars are present in the fermentation media, in which a preferred sugar (usually glucose) prevents simultaneous uptake of other sugars (Okeke and Obi, 1994; Moxley and Zhang, 2007). Carbon catabolite repression caused by glucose has been observed in *C. saccharoperbutylacetonicum* N1-4 (Noguchi *et al.*, 2013; Yao *et al.*, 2017), *C. acetobutylicum* (Ounine *et al.*, 1985; Bruder *et al.*, 2015), and *C. beijerinckii* (Birgen *et al.*, 2018). CCR is detrimental from a process engineering perspective, as it prolongs the run time of fermentations, increasing running costs. Carbon catabolite repression of xylose uptake by glucose in *C. acetobutylicum* was relieved by mutating a fourteen-nucleotide catabolite responsive element (CRE) upstream of the *xylB* gene (Bruder *et al.*, 2015). A strain of *C. acetobutylicum* was engineered to utilise arabinose, glucose, and xylose simultaneously through deletion of the gene encoding enzyme II of the phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS), combined with overexpression of a xylose transporter, xylose isomerase and xylulose kinase (Xiao *et al.*, 2011). Another technique to increase xylose uptake involved expressing a transaldolase from *E. coli* (Gu *et al.*, 2009). Xylose is metabolised by conversion to xylulose-5-phosphate, which enters the pentose-phosphate pathway. Transaldolase is involved in the conversion of sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate to fructose-6-phosphate and erythrose-4-phosphate and is a rate limiting step in xylose metabolism (Lu and Liao, 1997). Overexpressing the transaldolase TalA in *C. acetobutylicum* increased xylose uptake compared with the wild type, but there was still evidence of CCR (Gu *et al.*, 2009).

Despite the challenges, lignocellulosic biomass will likely be a key feedstock for making biobutanol, and any general bulk chemical, at low cost, and *Clostridium* are ideal organisms for this. With the advancement of genetic tools (discussed in section 1.4), the generation of high substrate uptake strains with reduced or eliminated CCR will improve the economics of bulk chemical production.

1.4 Genetic Manipulation Tools in *Clostridium*

1.4.1 Pre-targeted Approaches and Initial Plasmid Systems

Achieving commercial production of any target desired molecule biologically requires modification of the metabolic and regulatory pathways to achieve commercially viable titres, rates, and yields (Das, Patra and Ghosh, 2020). Prior to the advancement of cheap and reliable genome sequencing methods, random mutagenesis was, and remains, a method of adapting strains to increase

production and/or increase tolerance to inhibitory compounds (Lee *et al.*, 2011). In *Clostridium*, early efforts included using the mutagenic compound ethyl-methane sulphonate to generate asporogenic *C. acetobutylicum* strains, (Allcock and Reid, 1982), which had 17 % higher butanol titres than the wild type (Largier *et al.*, 1985). Ethidium bromide induced mutagenesis was also used to generate asporogenic strains which maintained solvent production (Long, Jones and Woods, 1984b). Another strain derived from *C. beijerinckii* NCIMB 8052 following mutagenesis induced by *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (NTG) was identified as a hyperamylolytic strain BA 101 (Annous and Blaschek, 1991), which produced over twice the butanol compared with the wild type strain (Formanek, Mackie and Blaschek, 1997).

Another method of strain development is selection on butanol to identify mutants resistant to butanol toxicity. The toxicity of end products is an issue with microbial fermentations, as it sets an upper limit on the product titres that can be achieved, reducing the yield and productivity (Ezeji *et al.*, 2010). Butanol is an antimicrobial compound, exerting stress through chaotropic (disordering macromolecular structures) effects (Cray *et al.*, 2015). In solventogenic *Clostridium*, butanol inhibits growth and substrate uptake at concentrations above 139 mM, with complete growth inhibition at concentrations above 236 mM (Bowles and Ellefson, 1985), which has led investigators to try and isolate strains which can tolerate higher concentrations of butanol. One such example in *C. acetobutylicum* selected for butanol resistant mutants exposed to NTG, which isolated a strain which produced 30 % more butanol compared with the wild type (Hermann *et al.*, 1985). Isolating such mutants will be crucial for increasing butanol titres.

While random mutagenesis and laboratory adaptation can improve strain performance, reliable methods for plasmid transformations can enable rational design of strains through gene knockouts and overexpressions (Nielsen, 2001). The development of gene knockout systems enables the study of gene loss and can infer gene function. *Clostridium* have traditionally been difficult to manipulate with plasmid transformation (Lee *et al.*, 1992) due to native restriction endonucleases, such as Cac8241 in *C. acetobutylicum* ATCC824, (Mermelstein *et al.*, 1992). Some species, such as *C. beijerinckii* do not have such a system and are amenable to transformation using shuttle vectors (Lee, Bennett and Papoutsakis, 1992). It was found that methylation of *E. coli*/*C. acetobutylicum* shuttle vectors with a methyltransferase from the *B. subtilis* phage ϕ 3T protected plasmids against digestion when transforming *C. acetobutylicum* (Mermelstein and Papoutsakis, 1993). The first knockout systems utilised non-replicative integrational plasmids, which were able to disrupt the *spo0A* and *gutD* genes in *C. beijerinckii* NCIMB 8052 (Wilkinson and Young, 1994). An issue with *Clostridium* is low frequencies of homologous recombination, with typical frequencies of between 1.7×10^{-6} to $< 4.8 \times 10^{-9}$ (Wilkinson and Young, 1994). In addition, generated mutants are not entirely

stable due to the homologous regions present on the integrational plasmid. In the absence of selection pressure, the recombination frequencies of integrants varied from 1.3×10^{-3} to 5.6×10^{-4} (Wilkinson and Young, 1994). Nevertheless, this method was successfully used to disrupt the *ack*, *butk* and *aad* genes of *C. acetobutylicum* (Green and Bennett, 1996; Green et al., 1996; Sillers et al., 2008).

1.4.2 Group II Intron Systems

To overcome the aforementioned issues using integrational plasmids, an alternative approach involving insertional activation has been developed by multiple groups. The ClosTron system utilised mobile group II introns from *Lactococcus lactis* (Heap et al., 2007). The group II intron system contains an antibiotic resistance gene which itself is interrupted by a group I intron, which is removed upon integration (**Fig. 1.5**). This is referred to as a retrotransposition activated marker (RAM). This means that only upon integration of the group II intron into the target site will the recombinant mutant be resistant to the antibiotic (Heap et al., 2007). This system was utilised to generate *spo0A* and *pyrF* knockout mutants (Heap et al., 2007). One issue with the ClosTron system is the integration of antibiotic resistance into the genome, meaning only a finite number of mutations can be produced in a single strain. This was later addressed by the same group, who flanked the antibiotic resistance gene with FLP-recognition target (FRT) sites, which allowed later removal of the erythromycin resistance gene using an FLP-recombinase expressing plasmid and was used to generate a *spo0A/agrA* double knockout (Heap et al., 2010). In addition, delivery of so-called 'cargo' sequences on the intron up to 1 kbp long were delivered into the chromosome (Heap et al., 2010). A group II intron-based system has also been used to manipulate *C. beijerinckii* by disrupting the *pta* and *buk* genes (Wang et al., 2013) and to disrupt the *solR* and *buk* genes (Shao et al., 2007).

1.4.3 Allelic Coupled Exchange

Another technique termed 'allele coupled exchange' (ACE) functions by a two-stage recombination event (Heap et al., 2012). ACE functions by combining the integration of a target sequence into the genome with the expression of a selectable marker (**Fig. 1.6**). To achieve this, a selection marker is cloned into an ACE plasmid upstream of the target sequence. Both the selection marker and target sequence are 'silent', i.e., they lack a promoter sequence and are not expressed. The plasmid contains two homologous regions, a longer 1200 bp region downstream of the insertion cassette, and a second, shorter 300 bp homologous region upstream of the insertion cassette. The longer homology arm results in a single integration event, with selection on media using an intact plasmid borne selection marker. Next, the plasmid is excised, mediated by the second, shorter homology arm. This results in integration of the insertion sequence downstream of a chromosomal promoter,

resulting in co-expression of the selection marker and the target sequence. This technique was utilised by Heap *et al.* to integrate a secondary alcohol dehydrogenase into the chromosome of *C. acetobutylicum*, enabling production of isopropyl alcohol.

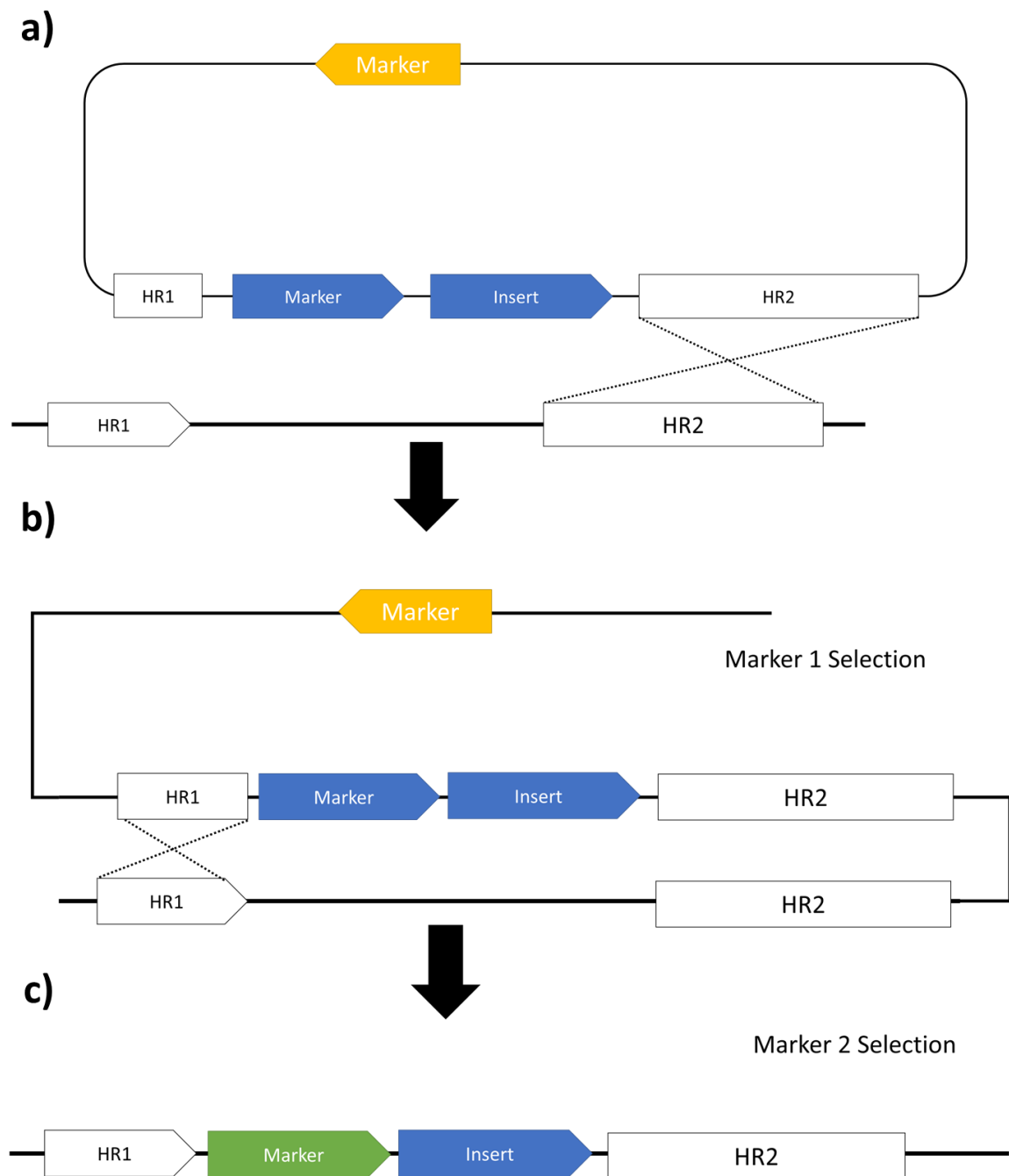


Figure 1.6 – Schematic of allelic coupled exchange (ACE). ACE utilises two homology arms of differing lengths and relies on genomic promoters to drive expression of the target sequence. a) After transformation with the ACE vector, the longer homology arm undergoes a recombination event, generating single crossover mutants. This integrates the plasmid into the genome. b) The shorter homology arm then undergoes a second recombination event, resulting in the excision of the plasmid. c) The plasmid excision results in double crossover mutants. The selection marker and insert are expressed from the genomic promoter. Adapted from (Heap et al., 2012).

A counter-selection marker using an inducible *mazF* toxin gene was developed, which could be used for targeted allelic integration events as in the previously mentioned ACE system (Al-Hinai, Fast and Papoutsakis, 2012). This was also coupled with FRT mediated excision of the antibiotic resistance marker, which was utilised by Al-Hinai *et al.* to knockout the *sigK* and *Ca_p0167* genes in *C. acetobutylicum* (Al-Hinai, Fast and Papoutsakis, 2012). Other approaches to deliver DNA into the

chromosome include the *mariner*-based transposon system, which was used to deliver the *tcdR* gene into the genome (Zhang, Grosse-Honebrink and Minton, 2015) and disrupt the *spo0A* gene (Zhang *et al.*, 2016) in *C. acetobutylicum*. In a *C. acetobutylicum* ATCC 824 strain lacking the *Cac8241I* restriction endonuclease system and an *upp* knockout, the FLP-FRT recombination system has been used to generate knockouts of the genes *Cac_3535* (encoding a potential enzyme with type II restriction endonuclease and methylase activities), *ctfAB* and *ldhA* (encoding lactate dehydrogenase) (Croux *et al.*, 2016).

1.4.4 CRISPR/Cas systems

The CRISPR/Cas system has been widely used for multigene editing in *E. coli* (Jiang *et al.*, 2015) and there have been numerous efforts to implement CRISPR/Cas systems in *Clostridium*. The first reported attempt at a CRISPR/Cas system in *Clostridium* was in *C. beijerinckii*, which used a single plasmid containing two homologous regions for recombination along with the Cas9 and guide RNA (Y. Wang *et al.*, 2015). In this case, the Cas9 protein did not directly edit the genome of the target organism, as in mammalian CRISPR/Cas systems (Cong *et al.*, 2013; Cameron *et al.*, 2019), but functioned as a selection tool, as the bacteria which underwent homologous recombination lost the gRNA recognition site (Y. Wang *et al.*, 2015). Later, the same group attempted to use a traditional CRISPR/Cas system, consisting of a DNA editing template, the Cas9 protein gene and the chimeric guide RNA, but the failure to obtain colonies using this method suggested that Cas9-induced double strand breaks were lethal in *C. beijerinckii* NCIMB 8052 (Wang *et al.*, 2016). The previously used system (i.e., using Cas as a selection tool rather than an editing tool) was again utilised to knockout both small (50 bp) and large (1.5 kbp) fragments of the *pta* gene (Wang *et al.*, 2016). In addition, Wang *et al.* integrated a 1.6 kbp fragment consisting of the *adhE* gene into the *pta* chromosomal locus. This system was later used in *C. saccharoperbutylacetonicum* to develop *pta* and *buk* knockout mutants, as well as a *pta-buk* double mutant (Wang *et al.*, 2017).

Other groups have attempted to use a more traditional CRISPR/Cas system to improve recombination efficiencies. The *Streptococcus pyogenes* Cas9 protein was used to edit genes in *C. pasteurianum*, which achieved editing efficiencies of 100 % but suffered from low transformation efficiencies of 2.6 CFU μg^{-1} plasmid DNA (Pyne *et al.*, 2016). It was found that Cas9 was toxic to the cells even in the absence of gRNA (Pyne *et al.*, 2016). As an alternative, the native type-IB CRISPR-Cas system was used to select against unmodified cells and was used to delete a region of the *cpaAIR* coding sequence (Pyne *et al.*, 2016). The lethality and low transformation efficiency (1.7 CFU μg^{-1}) was also observed in *C. acetobutylicum* expressing the *S. pyogenes* Cas9 (Li *et al.*, 2016). The authors then utilised a Cas9 nickase, which created single strand nicks rather than double strand breaks, which successfully generated *pyrE*, *adc* and *agrA* mutants. Li *et al.* also developed a

method of reducing gene expression using a 'dead Cas9' (dCas9), which bound to DNA but did not form strand breaks, which was used to reduce *spo0A* expression.

Genomic knockouts using native CRISPR/Cas systems were also reported in *C. saccharoperbutylacetonicum*, in which the use of native Cas as a selection tool was used to produce *spo0A* knockout mutants (Atmadjaja *et al.*, 2019) (overview in **Fig. 1.7**). Single nucleotide polymorphisms (SNPs) were also achieved by mutating the Cas recognition site in the homology arms simultaneously when introducing the desired SNP at the target gene (Atmadjaja *et al.*, 2019). SNPs were also achieved in *C. berjerinckii* using an alternative method. A two-step system was employed whereby first, a transformation was conducted which introduced a desired Cas recognition site near to the region to be edited, while the recognition site used for this first step was itself mutated (Wang *et al.*, 2016). In the second step, the desired SNP was introduced at the newly introduced Cas recognition site, and the mutated site used for recognition in step one was then reverted to the original sequence (Wang *et al.*, 2016).

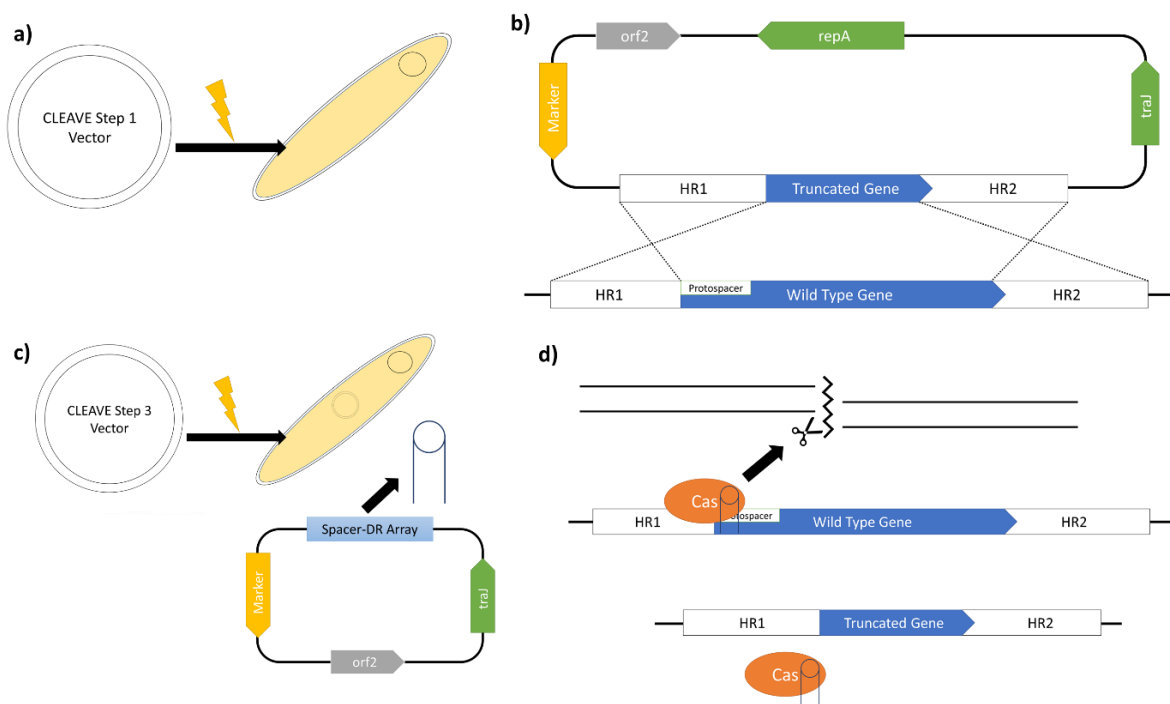


Figure 1.7 – Outline of CLEAVE system for genetic modifications, with example of a gene knockout. a) Recipient strain is transformed with a step 1 CLEAVE plasmid, which contains a truncated version of the target gene and homology arms. b) Transformed colonies are subcultured to allow homologous recombination to occur. The wild-type target gene contains a protospacer, which is removed during the homologous recombination event. c) The recipient strain is transformed with a second CLEAVE plasmid, the step 3 vector. This vector has a synthetic CRISPR array (Spacer-DR array in figure) which targets the wild type gene. This is transcribed as an RNA which guides the native Cas protein to the protospacer in the target gene. d) Where homologous recombination has not occurred (top)d, the RNA produced from the step 3 vector guides the Cas protein to the target gene and initiates a double strand break in the genome, which is lethal. Where recombination has been successful, the truncated gene loses the protospacer, preventing targeting by the Cas protein. This ensures colonies which are recovered after transformation in c) have undergone recombination.

Another approach using heterologous Cas was achieved using an inducible Cas9 from *S. pyogenes*, with one plasmid containing the *cas9* gene and another containing the gRNA (Wasels *et al.*, 2017).

This system reduced the detrimental effects of constitutive Cas9 expression mentioned above, and enabled editing of the *C. acetobutylicum upp* gene, integration of an *adh* gene into the chromosome and deletion of the pSOL1 megaplasmid (Wasels *et al.*, 2017). *C. saccharoperbutylacetonicum* has had the endogenous megaplasmid cured using CRISPR-Cas9 (Gu *et al.*, 2019). More recently, a CRIPSR/Anti-CRISPR editing approach was reported in *C. acetobutylicum* (Wasels *et al.*, 2020). This method functions as an extension of the previously reported two plasmid system (Wasels *et al.*, 2017) with an additional anti-Cas9 protein AcrIIA4 (Wasels *et al.*, 2020). The latter protein bound and repressed action of the Cas9/gRNA complex, which greatly improved transformation frequencies of butanol dehydrogenase knockouts (Wasels *et al.*, 2020).

Advances in CRISPR/Cas tools open the door to advanced metabolic engineering techniques, increasing the range of manipulations which can be made and reducing the time taken to generate mutants, which is a step toward high throughput manipulation technologies. This will increase the speed of strain generation and make *Clostridium* more attractive platforms for metabolic engineering.

1.4.5 Metabolic Engineering of *Clostridium* for *n*-Butanol Production

While *Clostridium* are native butanol producers, the generation of the side products acetone, acetate, butyrate, and ethanol cause problems by reducing the yield of butanol and necessitating the separation of compounds, which increases running costs. A common approach in metabolic engineering to improve desired product yields is the removal of competing pathways (Lee and Kim, 2015). The acetate and butyrate forming pathways have indeed been targets for knockouts. In *C. acetobutylicum*, *ack* knockouts led to a reduction in acetate titres by 22 % but the butanol titre was virtually unchanged (Green *et al.*, 1996), while other groups reported increased glucose consumption and a butanol titre increase by 23 % increase in butanol titre when *ack* was knocked out (Cooksley *et al.*, 2012). Removing the butyrate forming capabilities of *C. acetobutylicum* through *buk* knockouts increased butanol titres by 10 %, but also increased acetate titres by 30 %, while reducing acetone titres by 50 % compared to the wild type (Green *et al.*, 1996). In contrast, a *ptb* knockout reduced butanol titre in pH-controlled fed batch culture by 37 % (Lehmann, Radomski and Lütke-Eversloh, 2012a). A double knockout of both *ptb* and *buk* produced more butanol than wild type in acidogenic (pH 6.3) and solventogenic (pH 4.4) culture conditions (Yoo *et al.*, 2017). The acid forming pathways have also been knocked out in *C. saccharoperbutylacetonicum*. The *pta* knockout failed to increase the butanol titre or yield, while the *buk* knockout increased the butanol titre by only 1 % and the yield by 4 % (Wang *et al.*, 2017). The knockouts of both *pta* and *buk* increased the butanol titre by 7.3 % and the yield by 4 % (Wang

et al., 2017). In *C. beijerinckii*, knockout of the *pta* gene reduced butanol titres by 8 % but increased the yield by 12 % (Wang *et al.*, 2013b), while the *buk* knockout led to a substantial 28 % increase in butanol titre. This, however, was accompanied by a 4 % reduction in butanol yield. Increasing yield is advantageous, as more of the sugar feedstock is converted into butanol, less feedstock can be used to achieve a given amount of butanol. Titre is also an important metric, as separation costs decrease as the concentration of product increases, however as butanol is toxic, titre cannot be increased indefinitely. Above a threshold titre (which for commercialisation depends on the nature of the product (Gustavsson and Lee, 2016)), generally when butanol begins to inhibit cell growth, it becomes important to control the butanol concentration below the toxicity threshold, while generating a high yield. Concentration of butanol can be controlled through extraction technologies (such as gas stripping and pervaporation (Ezeji, Qureshi and Blaschek, 2004)), while yield is determined by the metabolism of the host strain. Provided that titres are above a certain threshold, in my view the yield is the more important metric.

Along with the acid forming pathways, the acetone forming pathway has also been a knockout target to increase butanol production. The acetoacetate decarboxylase and butyrate-acetoacetate CoA-transferase subunit A were disrupted in *C. acetobutylicum*, which in both cases effectively eliminated acetone production (Lehmann *et al.*, 2012), however butanol titres were reduced to less than half of the wild type, and butyrate and acetate titres were increased (Cooksley *et al.*, 2012; Lehmann *et al.*, 2012). Knockouts of both *ctfA* and *ctfB* genes in *C. acetobutylicum* eliminated acetone formation but reduced butanol titres by approximately 30 % and increased acetic acid titres by around six-fold, while butyric acid levels increased by around 50 % (Croux *et al.*, 2016).

As butanol production relies on NADH as a cofactor, competing NADH consuming pathways are important targets for disruption. Lactate formation is one such example. A deletion of lactate dehydrogenase in *C. acetobutylicum* ATCC 824 managed to abolish lactic acid production, while slightly increasing butanol and ethanol titres (Croux *et al.*, 2016). Another major NADH consuming knockout target for many groups has been the hydrogenase encoding gene *hydA*, with the rationale being that by removing this enzyme, the only remaining route for NADH oxidation would be the fermentation pathways. This has been unsuccessful in *C. acetobutylicum* (Cooksley *et al.*, 2012; Nguyen *et al.*, 2018), however a hydrogenase knockout strain of *C. pasteurianum* has been reported (Schwarz *et al.*, 2017). This increased the levels of butanol by 12.7 % in glycerol fermentations and by five-fold in glucose fermentations (Schwarz *et al.*, 2017). The metabolic network of *Clostridium* shows a high degree of flexibility, and from a metabolic engineering perspective care needs to be taken to ensure that reducing power is channelled towards the desired compounds.

Metabolic engineering has also been attempted by manipulating regulators of solventogenic genes. Transcriptional regulators have been identified which negatively regulate solvent producing genes. SolR, for example, is a transcriptional repressor of the *sol*-operon in *C. acetobutylicum* (Nair *et al.*, 1999), whose inactivation over doubled the butanol titre (Nair *et al.*, 1999), although later studies only showed an increase of 24 % (Harris *et al.*, 2001). Another transcriptional regulator is Rex, which binds to and represses transcription of the *bcs*-operon, *ldh* and *thlA* in *C. acetobutylicum* (Wietzke and Bahl, 2012; Zhang *et al.*, 2014). Removing the *rex* gene increased butanol titres and reduced the butyrate titres in *C. acetobutylicum* (Wietzke and Bahl, 2012).

In addition to knockouts, overexpression of genes encoding specific enzymes can divert flux toward the target compound. For example, the glycolytic enzymes 6-phosphofructokinase and pyruvate kinase were overexpressed in *C. acetobutylicum*, which increased both butanol and ethanol titres by 29.4 and 85.5 %, respectively (Ventura, Hu and Jahng, 2013). Various enzymes involved in ABE fermentation have been overexpressed in *C. saccharoperbutylacetonicum*: overexpression of the *thl* gene and a mutated *thl* increased butanol titres by 5 %, while overexpression of the *thl*, *hbd*, *crt* and *bcd* genes increased butanol titres by 13 % (Wang, Dong and Wang, 2017). The asporogenic M5 strain of *C. acetobutylicum* has also been engineered to produce solvents by transformation with overexpression constructs. Introduction of the genes for acetone (*ctfA*, *ctfB* and *adc*) and butanol (*aad*) synthesis were reintroduced on plasmids, restoring production of the respective solvents, although obtained titres were below those in the wild type (Cornillot *et al.*, 1997). Knockout of the acetate kinase (*ack*) gene in the M5 strain reduced the acetate titres compared to the parental strain, but when combined with plasmid expression of *aad* to induce butanol production the acetate titres were still > 150 mM in bioreactors (Sillers *et al.*, 2008). In other cases, overexpression was not beneficial for butanol formation. For example, the overexpression of thiolase reduced titres of butanol (Sillers *et al.*, 2008). Further efforts involved expressing *adhE1* and *ctfAB*, which produced similar butanol titres to wild type *C. acetobutylicum* ATCC824 (Jin *et al.*, 2009). In addition, acetate titres were over double those seen in the wild type (Jin *et al.*, 2009). Another engineered *C. acetobutylicum* increased the butanol titre by 60 % through deletion of *buk*, *pta* and overexpression of *adhE1* (Jang, Lee and Lee, 2012).

Certain non-solventogenic *Clostridium* have been engineered to produce butanol by introduction of solventogenic genes. The butyrate producing *C. tyrobutyricum* ATCC 25755 for example, has been engineered to produce butanol through the introduction of the *adhE2* or *aad* genes from *C. acetobutylicum* ATCC 824, producing 19 mg L⁻¹ and 67 mg L⁻¹ of butanol, respectively (Yu *et al.*, 2011). Knocking out the acetate kinase (*ack*) gene and changing the promoter of the *adhE2* gene to the native thiolase (*thl*) promoter, thus allowing stronger expression, increased the

butanol titre to 10 g L⁻¹ (Yu *et al.*, 2011). Using the FRT based system described by Croux *et al.*, *C. acetobutylicum* was engineered to produce butanol with high productivity (Nguyen *et al.*, 2018). The *ldhA*, *ctfAB*, *ptb* and *buk* genes were removed to reduce generation of side products. The *rex* regulator gene was then removed, and the *thlA* gene was replaced with *atoB* from *E. coli*, which catalysed the same reaction but was not subject to redox-switch regulation and higher catalytic efficiency (Nguyen *et al.*, 2018). Finally, the *hbd* gene was replaced with *hbd1* from *C. kluyveri*, which catalysed the same reaction but showed higher NADPH specificity (Nguyen *et al.*, 2018). This final strain had butanol yields of 0.34 g g⁻¹ glucose (Nguyen *et al.*, 2018).

To summarise, increasing the butanol titre in *Clostridium* generally requires multiple manipulations to be effective. Certain intuitive manipulations, such as acid knockouts and thiolase overexpression, can be detrimental to butanol production. This highlights the need for system scale interventions that consider the entirety of cellular metabolism.

For commercial chemical production using microbes, titres of the target compound must be as high as possible, to reduce both separation costs and the total material requiring processing (Gustavsson and Lee, 2016). This is particularly when producing low value bulk chemicals (Lee and Kim, 2015), such as butanol. As already discussed, butanol toxicity is a barrier to improved butanol production as it sets an upper limit on butanol titre. Rational engineering efforts have also been made to overcome this issue. Overexpression of various heat shock proteins has been shown to increase the butanol titre in engineered *C. tyrobutyricum*, with expression of *groESL* from *Deinococcus wulumuquiensis* R12 increasing the butanol titre by 20 % relative to the control strain expressing the genes for butanol production (Fu *et al.*, 2021). Key regulatory genes have also been targets for metabolic engineering. Multiple histidine kinase knockouts were implemented in *C. acetobutylicum*. *Cac3319* knockout increased the butanol titres by 49 %, and reduced acid titres by 19 % (Du *et al.*, 2021). Layering knockouts with an additional *cac0323* knockout further increased butanol titre by 65 % (Du *et al.*, 2021).

In the last twenty years the breakthroughs in transformation and genetic manipulation technologies have enabled a host of metabolic engineering successes, but much remains to be done before the ABE fermentation process becomes commercially viable.

1.5 Computational Modelling of Metabolism

The strategies outlined above have shown the successes of the rational metabolic engineering approach, but there are also numerous examples where increased butanol titres were not achieved. The limitation of the rational approach is a tendency to view metabolism as a set of separate, disconnected pathways, without regard to the entire metabolic network. Many

metabolites and compounds are utilised by different enzymes in different pathways, and this needs to be accounted for when designing metabolic engineering strategies. As such, cellular responses to perturbation are difficult to predict (Bailey, 1991). Various computational approaches exist which can address these issues, which I will now discuss.

1.5.1 Flux Balance Analysis

With the ever-increasing number of publicly available genomes, retrieving metabolic information from these sequences became the next challenge. The genome encapsulates the entirety of proteins present in the organism, which in turn indicates which biochemical reactions could be performed. This information can be used to generate stoichiometric models of metabolism, which consist of a tabulation of reactions (termed a stoichiometric matrix), consisting of the stoichiometric coefficients of reactants in a metabolic reaction. These stoichiometries impose constraints on the metabolic network, ensuring that the total amount of a given product does not exceed the amount of substrate. In addition, inequalities can be applied as upper and lower bounds through a given reaction, defining the minimum and maximum fluxes through a reaction (Orth, Thiele and Palsson, 2010). Reactions are added to the metabolic model if there is evidence that the reaction is present either from direct experimental evidence, from sequence homology to a known enzyme, or inferences based on the presence of other enzymes in a given metabolic pathway (Orth, Thiele and Palsson, 2010). Linear programming is used to find a set of fluxes which maximise a specified objective, usually maximisation of flux through a lumped reaction representing biomass formation (Orth, Thiele and Palsson, 2010). Early models of *E. coli* were used to simulate growth and showed that *E. coli* was adapted for maximal growth, with the model predicting acetate uptake and oxygen uptake rates that were in good agreement with experimental data (Jeremy S Edwards, Ibarra and Palsson, 2001). A genome-scale compartmentalised metabolic model of *S. cerevisiae* (Förster *et al.*, 2003) accurately predicted biomass yield and ethanol producing flux in chemostat culture (Famili *et al.*, 2003). Flux balance analysis (FBA) has been used to identify lethal genes (Gianchandani, Chavali and Papin, 2010) and to identify optimal flux distributions in the case of gene or reaction knockouts (Ibarra, Edwards and Palsson, 2002). The new flux distribution predicted by FBA may require a substantial shift in metabolic fluxes relative to the wild type, and as such FBA predictions of knockouts may be unrealistic.

Alternative methods have been developed which attempt to provide more accurate predictions of flux distributions in response to knockouts. Notable examples are minimisation of metabolic adjustment (MOMA) (Segrè, Vitkup and Church, 2002) and regulatory on-off minimisation (ROOM) (Shlomi, Berkman and Ruppin, 2005). Briefly, MOMA attempts to find a flux distribution that approximates the immediate cellular response to genetic perturbation. MOMA assumes that the

immediate response to genetic perturbation will be to minimise the number of flux changes relative to a reference 'wild type' flux distribution. In mathematical terms, MOMA finds a point in the feasible solution space (subject to the respective knockout(s)) that is a minimal distance from the reference distribution. Shlomi *et al.* developed ROOM to address the shortcomings of MOMA, mainly that MOMA's objective of minimising the Euclidian norm prevents large modifications in single fluxes. Large alterations in flux have been reported in certain mutant strains, such as a pyruvate kinase knockout strain of *E. coli* (Emmerling *et al.*, 2002). ROOM assumes that the cell minimises the genetic regulatory changes required to reach the new flux distribution, minimising the adaptation cost. ROOM also assumes that the regulatory changes in the cell can be described with Boolean on/off states, assigning a fixed cost to each regulatory change (Shlomi, Berkman and Rupp, 2005).

When compared with standard FBA, MOMA was able to identify three genes which were essential to growth on glucose, but which FBA incorrectly predicted were non-essential (Segrè, Vitkup and Church, 2002). In addition, the fluxes estimated by MOMA for a *pyk* (pyruvate kinase) knockout strain of *E. coli* were closer to ¹³C-metabolic flux analysis (MFA) estimated fluxes than FBA predictions (Segrè, Vitkup and Church, 2002). ROOM predictions were compared with the *pyk* strain flux data, along with flux maps for four additional knockout strains (*pgi*, *zwf*, *gnd* and *ppc*, encoding phosphoglucose isomerase, glucose-6-phosphate 1-dehydrogenase, and phosphoenolpyruvate carboxylase, respectively). ROOM performed either equal to or better than MOMA and FBA for eight of the nine simulations but performed worse for the *pgi* knockout under low ammonia conditions (Shlomi, Berkman and Rupp, 2005). In terms of growth rate predictions, Shlomi, Berkman and Rupp found that MOMA was better at predicting the growth rate of knockout strains immediately after gene knockouts, whereas after adaptive evolution ROOM and FBA performed better. In summary, MOMA gives better estimates of how a metabolic network will respond in the short term to genetic perturbation. After adaptation to the knockout, ROOM and FBA provide more accurate predictions.

Aside from knockouts, FBA can also be used to model the insertion of new metabolic pathways into a host organism by including these reactions in the metabolic model. With the ease of inserting new reactions and deleting them, stoichiometric models provide a rapid framework for designing metabolic engineering strategies *in silico*. Examples of this combined modelling-experimental approach include the use of a stoichiometric metabolic model which was used to guide strategies for improved lycopene production in *E. coli* (Alper *et al.*, 2005). These involved deletion of *gdhA*, *aceE* and *fdhF*, which encode an NADP-specific glutamate dehydrogenase, the E1 component of pyruvate dehydrogenase and formate dehydrogenase, respectively, increasing lycopene titres by

37 % (Alper *et al.*, 2005). L-threonine production in *E. coli* was also improved using a flux response analysis approach (Lee *et al.*, 2007). *acs* encoding acetyl-CoA synthetase was overexpressed to reduce acetate formation, which increased the L-threonine titre by 20.4 % (Lee *et al.*, 2007). Another example was the improvement of 1,4-butanediol (BDO) production in *E. coli* by knocking out competing NADH consuming reactions (alcohol dehydrogenase, pyruvate formate lyase, lactate dehydrogenase and malate dehydrogenase), which increased the BDO titre from 0.6 mM BDO to over 12 mM BDO (Yim *et al.*, 2011).

Importantly, genome scale metabolic models of *Clostridium* have been developed. Multiple *C. acetobutylicum* models have been published (Lee *et al.*, 2008; Senger and Papoutsakis, 2008; McAnulty *et al.*, 2012; Yoo *et al.*, 2015), as well as *C. beijerinckii* (Milne *et al.*, 2011), *C. sporogenes* (Kaushal *et al.*, 2018), *C. thermocellum* (Thompson *et al.*, 2015, 2016), *C. cellulolyticum* (Salimi, Zhuang and Mahadevan, 2010), *C. ljundahlii* (Nagarajan *et al.*, 2013) and *C. autoethanogenum* (Norman *et al.*, 2019). These models provide opportunity to develop *in silico* metabolic engineering strategies for improved butanol production.

1.5.2 Strain Design Algorithms

Evolution has pushed metabolic networks to respond to selection pressures, which lies in competition with chemical overproduction objectives (Burgard, Pharkya and Maranas, 2003). Achieving overproduction of a given compound relies on redirecting fluxes towards the target compound. However, this reduces the cellular growth rate, as less carbon is available for production of biomass precursors. In cases where the chemical production is coupled to biomass production, there is a trade-off between growth rate and chemical producing flux (**Fig. 1.8**).

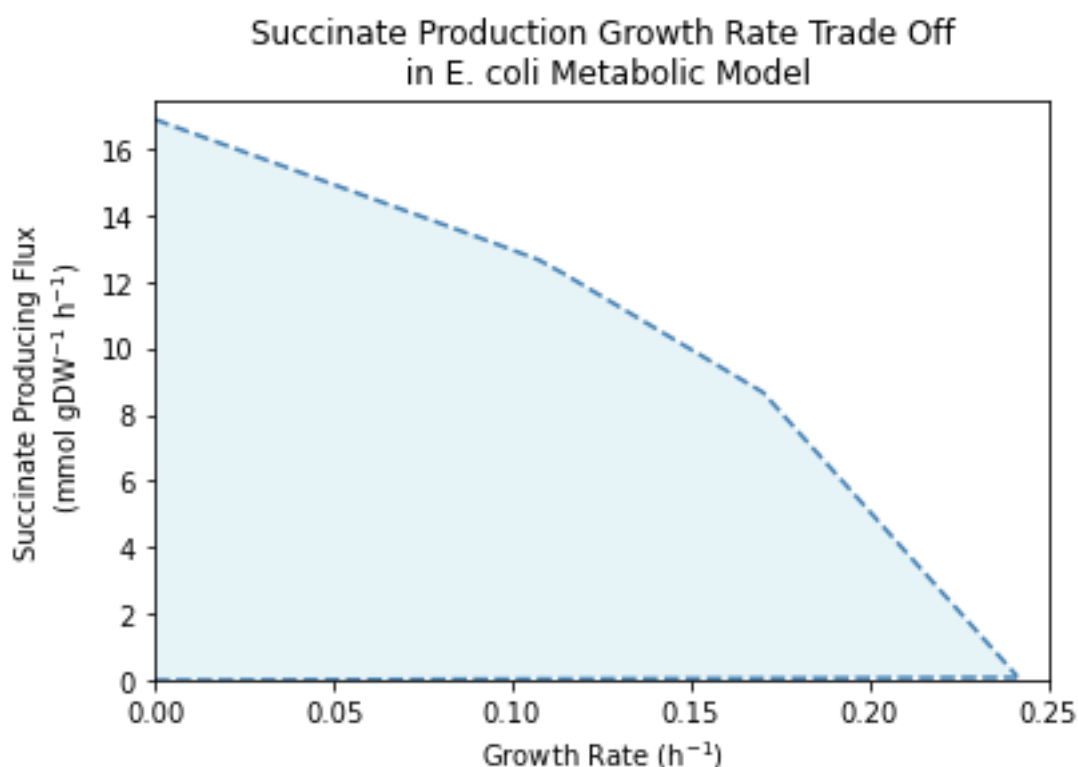


Fig. 1.8 – Pareto optimisation of succinate production versus growth rate in genome scale metabolic model of *E. coli* under anaerobic growth. Rate of glucose uptake was set to $10 \text{ mmol gDW}^{-1} \text{h}^{-1}$, with non-growth associated maintenance cost set to $3.15 \text{ gDW}^{-1} \text{h}^{-1}$. Dashed blue lines represent the upper and lower feasible bounds for succinate producing flux determined by flux variability analysis. Shaded area represents feasible succinate producing flux. As the growth rate increases, the maximum possible succinate producing flux decreases. Any solution which appears on the upper line is said to be ‘Pareto optimised’, meaning that an increase in either of the two fluxes is not possible without a decrease in the other.

FBA can be used to design a strain of an organism which can overproduce a certain compound, by removing competing side reactions or inserting heterologous pathways. A challenge when using FBA is the large number of simulations required to find the optimal network to achieve the desired overproduction target. The OptKnock framework was developed to address the issue of chemical overproduction while ensuring that growth is maintained (Burgard, Pharkya and Maranas, 2003). OptKnock maximises the flux towards a given chemical in the metabolic network through gene knockouts, subject to maximising a cellular objective such as biomass production. The aim of this framework is to ensure that target chemical production becomes a by-product of growth. A strain of *E. coli* which produced the non-natural compound 1,4-butanediol (BDO) was subject to knockout

targets (alcohol dehydrogenase, pyruvate formate lyase, lactate dehydrogenase and malate dehydrogenase) suggested by OptKnock (Yim *et al.*, 2011). This increased the BDO titre by over ten-fold, from 1.3 mM to 14 mM. OptKnock was also applied to increase lactate titres in *E. coli*. Multiple strategies were developed and implemented, accompanied with adaptive evolution to allow the metabolic network to adapt to the knockouts. Three strategies were implemented, and all showed an increase in lactate titre compared to the wild type; 1.5 g L⁻¹ lactate was produced from 2 g L⁻¹ glucose through a *pta-adhE* deletion strain after 60 days of adaptive evolution (Fong *et al.*, 2005).

It should be noted that in the aforementioned cases, regulation at the gene or protein level was not taken into account by the model. The absence of regulation meant that, in the case of BDO production, the OptKnock-suggested knockouts had to be coupled with deletion of the ArcA transcriptional repressor to increase levels of tricarboxylic acid (TCA) cycle enzymes (Yim *et al.*, 2011). A modified OptKnock procedure was applied to a genome scale metabolic model of *Pseudomonas putida* to produce polyhydroxyalkanoates (PHA) (Puchařka *et al.*, 2008). The authors searched for knockouts which could increase the availability of acetyl-CoA, a precursor in the synthesis of PHA, while also allowing the model to select different carbon sources.

A limitation of OptKnock is the problem of alternative optimal solutions. While a given set of knockouts may produce an increase in production of the target compound, there may be an alternative flux distribution whereby the same value for the inner problem (e.g., biomass production) may have less or even zero flux producing the target compound. RobustKnock was developed to address this issue (Tepper and Shlomi, 2010). Rather than maximising the production of the target compound, RobustKnock's outer problem maximises the minimum possible flux producing the target compound. RobustKnock predictions were compared to OptKnock predictions for four targets, with all RobustKnock's knockout targets reaching a higher minimum flux of target chemical production (Tepper and Shlomi, 2010).

OptGene is another framework aimed at overproducing a target chemical, in this case using a genetic algorithm (Patil *et al.*, 2005). A schematic of the OptGene algorithm is shown in **Fig. 1.9**. Mimicking directed evolution *in silico*, this algorithm begins with an initial "population" of mutants, with random removal of genes. A fitness function, such as FBA or MOMA, is applied to each individual, and the best performing individuals are selected for "mating", combining the gene knockouts in each parent to produce new individuals. After the fitness function is applied and the mating procedure has been performed, the new "offspring" are subject to an additional round of mutation, and the fitness function is applied again. This process is repeated until a desired criterion is met, such as a minimal flux towards a given chemical. OptGene was initially utilised to predict

knockout targets which overproduce vanillin, glycerol, and succinate in *S. cerevisiae* (Patil *et al.*, 2005). An advantage of OptGene is the ability to add additional criteria to the fitness function, such as the proposed biomass coupled product yield (BCPY) (Patil *et al.*, 2005). OptGene was applied to the *S. cerevisiae* network to double vanillin β -D-glucoside titres through knockout of the *pdh1* gene (Brochado *et al.*, 2010). It has also been used to increase production of the sesquiterpene cubebol in *S. cerevisiae* by over two-fold (Asadollahi *et al.*, 2009).

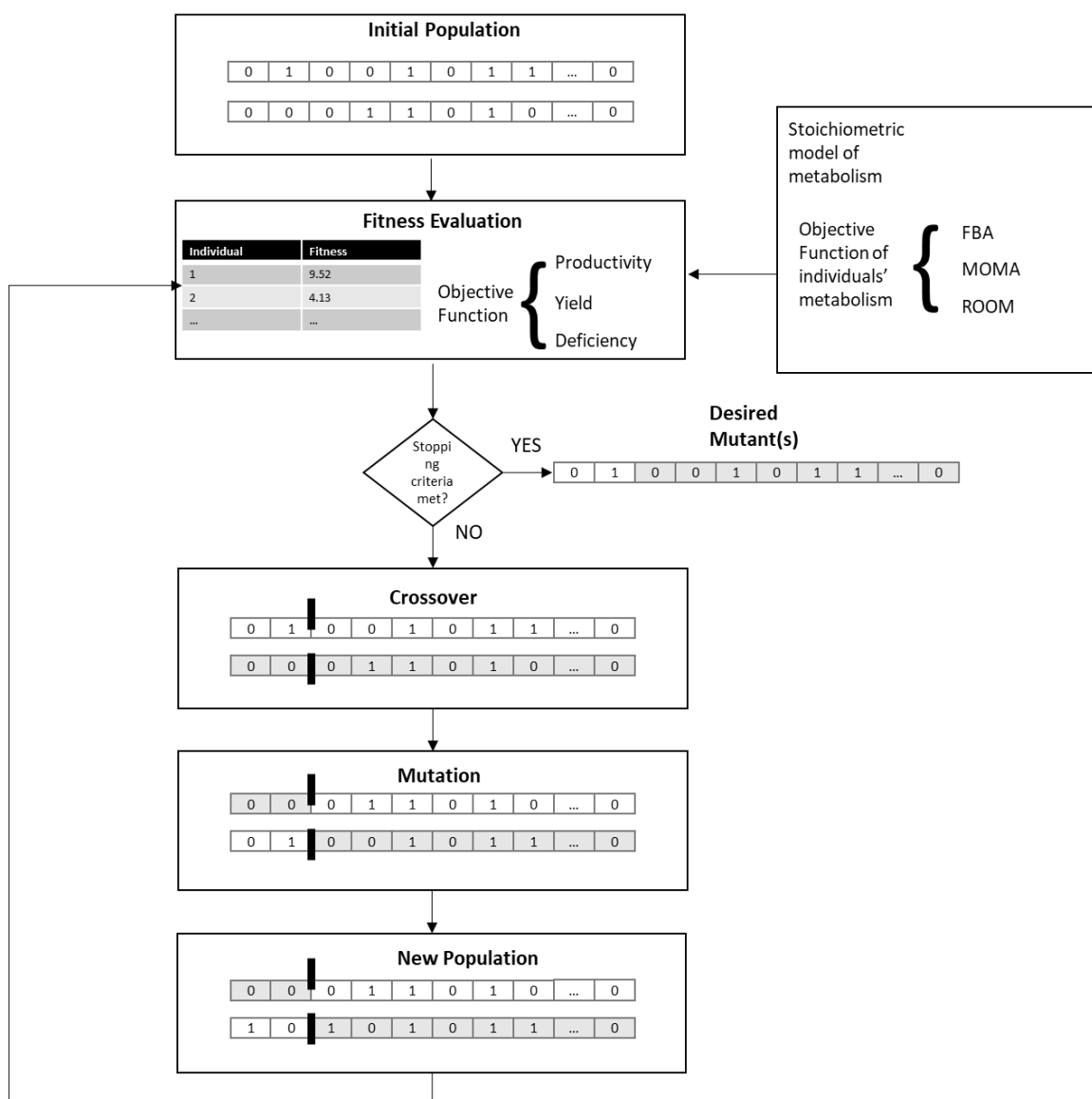


Figure 1.9 - Schematic overview of the OptGene algorithm. A population of individuals is initialised with random removal of genes or reactions from their metabolic networks. Individuals are then scored for fitness using an objective function such as FBA, MOMA etc. A set number of top individuals are selected for mating and are crossed to produce new offspring. Mutations are randomly introduced into the population at a specified mutation rate. This process is repeated until mutant(s) with the desired phenotype are obtained. Grey shaded boxes are used to represent different individuals in the cross over procedure. FBA – flux balance analysis, MOMA – minimisation of metabolic adjustment, ROOM – regulatory on/off minimisation. Adapted from (Patil *et al.*, 2005).

A similar algorithm, termed the cipher of evolutionary design (CiED), expands on OptGene by keeping beneficial ‘mutations’ by selecting the five individuals with the top fitness score without crossover, while the remaining individuals selected either undergo crossover or mutation (Fowler, Gikandi and Koffas, 2009). The CiED optimisation approach was used to improve production of the flavonoids naringenin and eriodictyol in *E. coli* by searching for single, double, triple, and quadruple gene knockout strategies (Fowler, Gikandi and Koffas, 2009). Implementing these strategies experimentally, a quadruple knockout of *sdhA*, *adhE*, *brnQ* and *citE* produced the highest titres of

both products (Fowler, Gikandi and Koffas, 2009). Using the model to investigate potential overexpression targets, the authors identified CoA synthesis genes (*acs* and *coaA*) as targets, which further increased the levels of naringenin titres (Fowler, Gikandi and Koffas, 2009). The CiED algorithm has also been applied to *E. coli* to increase NADPH production for multiple flavonoid polyphenol synthesis (Chemler *et al.*, 2010). Single, double, and triple reaction knockouts were selected which increased generation of NADPH, which were implemented *in vivo*. The best double and triple knockouts were *pgi ppc* and *pgi pldA ppc*, which increased catechin titres to over twice those in the parental strain (Chemler *et al.*, 2010).

OptForce is a further algorithm which, unlike OptKnock and OptGene, explicitly gives reactions whose flux needs to increase to meet the desired overproduction target (Ranganathan, Suthers and Maranas, 2010). The OptForce procedure is outlined in **Fig. 1.10**. OptForce aims to identify which fluxes must change to achieve an overproduction target by identifying sets of reactions whose flux must increase or decrease to achieve the overproduction target. Compared with other algorithms, OptForce has the advantage of suggesting overproduction targets, while OptKnock and OptGene may implicitly suggest overexpression targets, they require manual inspection and comparison of fluxes to a wild type to identify overexpression targets.

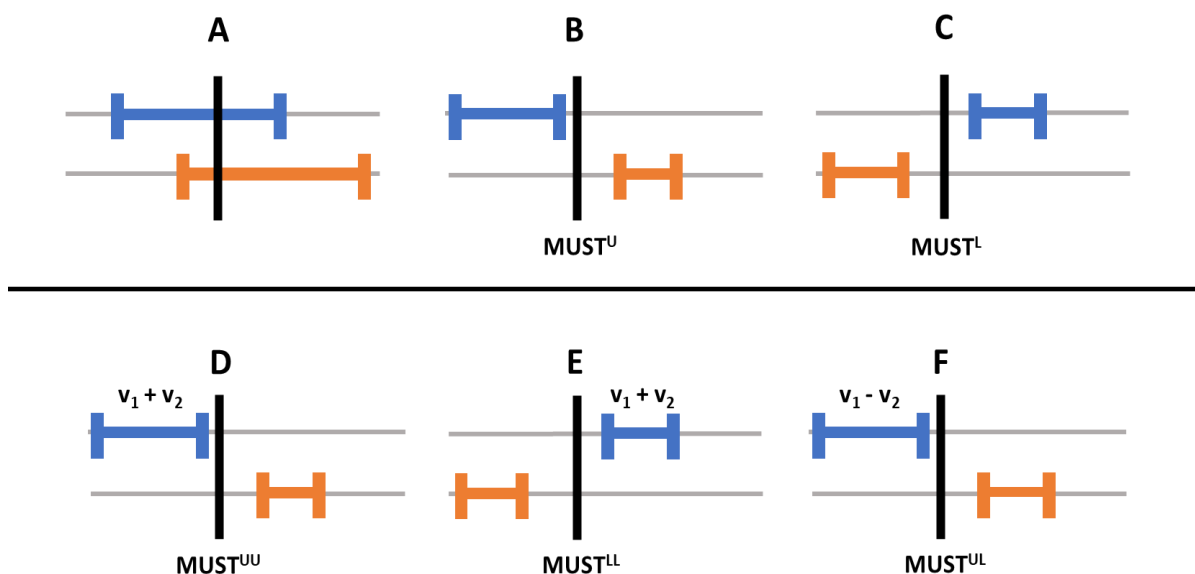


Figure 1.10 – Overview of OptForce procedure. An initial flux range is estimated for the wild type network (blue) subject to stoichiometric constraints, external rates and MFA flux data if available. Next, a set of flux ranges are estimated consistent with a desired chemical overproduction target (orange). By superimposing the flux ranges against each other, the required manipulation targets can be identified. In the cases where there is overlap between the two flux ranges (A), no manipulation might be required to reach the overproduction target. In cases where there is no overlap between the two flux ranges, the flux through the reaction will need to be increased if the overproducing flux range lies to the right of the wild type network (B) or decreased if the overproducing flux range lies to the left of the wild type network (C). Reactions whose flux must increase are placed in a set of reactions, termed $MUST^U$, while those whose flux must decrease are placed in the $MUST^L$ set. The ranges of fluxes with overlapping regions (A) may be further grouped by considering sums of two reaction fluxes and calculating their ranges in both wild type and overproducing networks. As before, the superimposed ranges of fluxes are compared. In the case of (D), either one of the two reactions fluxes must increase to reach these overproduction targets. In the case of (E), either one of the reactions fluxes must decrease. The difference between reaction fluxes can also be used to identify targets (F). Non-overlapping flux ranges imply that either the flux through the first reaction must increase or flux through the second reaction must decrease. These constitute $MUST^{UL}$ and $MUST^{LU}$ sets respectively. Adapted from (Ranganathan, Suthers and Maranas, 2010)

1.5.3 Kinetic Models of Metabolism

A limitation of stoichiometric models is their inability to model dynamic changes in metabolic networks, such as regulatory constraints, product inhibition and enzyme saturation levels. In some cases, such as continuous culture, these can be controlled by process engineering. However, in batch culture, modelling these processes requires dynamic models involving differential equations. These models tend to require more effort to produce compared to stoichiometric models but can provide more detailed models of metabolism. Early methods combined small FBA models of core metabolism with differential equations for biomass accumulation, substrate consumption and by-product secretion, with the differential equations solved by dividing the culture into small time steps then iteratively repeating the procedure (Varma and Palsson, 1994). This was used to model *E. coli* aerobic growth in batch and fed-batch culture, and anaerobic growth in batch culture, which showed good agreement with experimental data (Varma and Palsson, 1994). This approach was also used with additional constraints on the flux rate of change, which were estimated based on transcription and translation rates (Mahadevan, Edwards and Doyle, 2002). A second approach

implemented by Mahadevan, Edwards and Doyle was a 'dynamic' optimisation approach. This involved an end-point objective, in which the end point biomass was the objective function, rather than maximisation of biomass at each individual time point (Mahadevan, Edwards and Doyle, 2002). The former approach, termed the 'static' optimisation gave closer matches to the experimental data (Mahadevan, Edwards and Doyle, 2002). Another dynamic extension of FBA involved adding transcriptional regulatory constraints to a small toy metabolic network to model growth, substrate uptake and product secretion of *E. coli* with carbon catabolite repression and a shift from aerobic to anaerobic growth (Covert *et al.*, 2001).

Increasing the scale of kinetic models has been achieved with the 'ensemble modelling' (EM) approach (Tran, Rizk and Liao, 2008; Milne *et al.*, 2011). For a given stoichiometric model, the reactions are decomposed into elementary reaction steps, and kinetic parameters are determined based on reaction reversibilities and enzyme distributions (bound, unbound etc.). This process is repeated thousands of times to generate an ensemble of models allowable by thermodynamics, and this ensemble is used to determine effects of enzyme overexpression (Tran, Rizk and Liao, 2008). This approach was used to develop an *E. coli* metabolic model, by comparing ensemble model predictions with various flux datasets of *E. coli* knockout mutant strains (Khodayari *et al.*, 2014), which gave metabolite concentration predictions which were within experimental ranges for 68 % of metabolites. This model consisted of 138 reactions and 93 metabolites. A larger model was subsequently published, consisting of 457 reactions and 337 metabolites (Khodayari and Maranas, 2016). Using this model, predictive accuracy improved with 80 % of predicted metabolite concentrations within experimentally measured ranges (Khodayari and Maranas, 2016).

Clostridium fermentations have also been modelled using ordinary differential equations. Early efforts modelled batch culture of *Clostridium acetobutylicum*, utilising inhibition terms for butanol, acetate, and butyrate (Srivastava and Volesky, 1990). This model showed good agreement with experimental data for biomass, butanol, acetone, ethanol, acetate, butyrate and lactate accumulation, and glucose consumption (Srivastava and Volesky, 1990). Later efforts modelled *C. saccharoperbutylacetonicum* (Shinto *et al.*, 2007) and the impact of acetic and lactic acid on behaviour (Zhou, Liu and Yuan, 2018), an improved *C. acetobutylicum* model which incorporated butyryl-phosphate and its possible regulatory roles (Li *et al.*, 2011), a *C. acetobutylicum* model consisting of three functional modules; metabolic reactions, gene regulation by Spo0A and environmental cues such as acid uptake, whose predictions provided good agreement with wild type *C. acetobutylicum*, a *ctfA/B*, *adhE* and *pta* knockout strain and a *spo0A* knockout strain (Liao *et al.*, 2015). Continuous culture models have also been reported (Buehler and Mesbah, 2016). The aforementioned ensemble modelling method was used to generate a model of *C. thermocellum*,

which is an ethanol producer but not a butanol producer, with predicted ethanol forming fluxes which were consistent with 10 out of 19 experimental mutants (Dash *et al.*, 2017). The use of kinetic models provides more accurate predictions than FBA, particularly in batch culture, but require extensive parameterisation and/or datasets for validation.

1.5.4 ^{13}C -Metabolic Flux Analysis

FBA and kinetic models can predict an optimal flux distribution to achieve a defined objective function and can model changes in the levels of substrates and metabolites in a system, respectively. But to achieve an accurate picture of cellular metabolic fluxes in the wild type (WT) strain under a given condition, ^{13}C -metabolic flux analysis (MFA) is used. This technique uses isotopically labelled tracers to label intracellular metabolites and estimate metabolic fluxes based on the labelling pattern of said metabolites at isotopic steady state as determined by mass spectrometry or nuclear magnetic resonance (Long and Antoniewicz, 2019). ^{13}C -MFA can be used in metabolic engineering to gain a detailed quantification of central metabolic fluxes, which can be used to understand an organism's phenotype and to guide metabolic engineering strategies (Wiechert, 2001).

This technique has been extensively used in *E. coli* to understand how fluxes change under aerobic and anaerobic conditions (Chen *et al.*, 2011), compare metabolism of glucose and xylose under aerobic and anaerobic conditions (Gonzalez, Long and Antoniewicz, 2017) and to compare metabolism of wild type *E. coli* with a glucose-6-phosphate isomerase mutant (Toya *et al.*, 2010a).

Other studies investigated phenotypes of the non-model bacteria *Vibrio natriegens* and *Basfia succiniciproducens*, in the former case to identify fluxes behind the rapid growth of this species, which was found to be through high glycolytic and anaplerotic flux (Long *et al.*, 2017), and in the case of the latter to investigate the TCA cycle when growing on sucrose, which elucidated the roles of a PTS sucrose uptake system and a fructokinase (Lange *et al.*, 2017). It has also been used in *Clostridium acetobutylicum* to identify an incomplete TCA cycle and an isoleucine synthesis pathway via citramalate synthase (Au *et al.*, 2014).

In the context of metabolic engineering, ^{13}C -MFA has been used to identify an NADPH deficit in *E. coli* engineered for isopropyl alcohol (IPA) production during exponential growth which was limiting IPA production (Okahashi *et al.*, 2017). Another study of an *E. coli* strain engineered for fatty acid production used ^{13}C -MFA to compare fluxes of the engineered strain to the wild type strain to understand the metabolic response to the fatty acid overproduction. The generated flux map showed an increase in acetyl-CoA being converted to fatty acids, along with an increase in pentose-phosphate pathway, Entner-Doudoroff pathway and transhydrogenase flux converting NADH to NADPH (He *et al.*, 2014). Fatty acid production was also investigated in *S. cerevisiae* to determine

the fate of acetyl-CoA in an engineered strain of *S. cerevisiae* expressing ATP-citrate lyase (Ghosh *et al.*, 2016). This showed a near doubling of acetyl-CoA forming flux, but this was offset by an increase in malate synthase consumption of acetyl-CoA, and downregulation of this gene increased fatty acid titres by around 26 %. As these examples show, MFA is a powerful tool in metabolic engineering for identifying reactions limiting production of a desired compound and competing side reactions which may be targets for knockouts or downregulations.

1.4 Objectives of My Research

Producing butanol at scale requires a combination of high butanol titres, high conversion of substrate to butanol (yield) and high productivity (Lee and Kim, 2015), and improving these metrics in *Clostridium* will be a step toward reviving the ABE fermentation process and making it commercially viable. The first step to reach these goals would be a butanol producing strain of *Clostridium* which produces high titres of butanol, with high yield.

The work described in this thesis aims at the improvement of a previously generated asporogenic strain of *C. saccharoperbutylacetonicum* N1-4 (Atmadjaja *et al.*, 2019) to bring butanol titres up to commercially relevant levels. The work was done in partnership with Biocleave Ltd who provided access to specialised laboratory facilities for *Clostridium* culture and molecular manipulation and gave generous access to proprietary strains and molecular tools. I carried out all experiments and computational modelling experiments except where otherwise stated.

My research aimed to address the following objectives:

First, to use *in silico* guided approaches to develop novel metabolic engineering strategies for increasing butanol production in *C. saccharoperbutylacetonicum*.

Second, to quantify how the metabolism of the asporogenic strain of *C. saccharoperbutylacetonicum* compares to the wild type parental strain, and to identify what is limiting its butanol forming capabilities.

Third, to gain a detailed understanding of the intracellular fluxes of the asporogenic strain, using ¹³C-metabolic flux analysis to generate a flux map and use FBA to gain insight into the flux distribution.

Finally, to implement the devised metabolic engineering strategies and to assess the impact on butanol producing capabilities of the asporogenic strain.

Chapter 2: Materials and Methods

2.1 Materials

All chemicals were purchased from Sigma-Aldrich (Darmstadt, Germany), unless otherwise stated. For any bacteria culture, growth media were prepared using ultrapure deionised water (filtered through an Arium® Comfort, Sartorius, Goettingen, Germany). All plasticware (microcentrifuge tubes, pipette tips, etc.) were purchased from STARLAB (Milton Keynes, United Kingdom), unless otherwise stated.

2.2 Bacteria Strains

The bacteria strains used in this work are shown in **Table 2.1**.

Strain	Relevant Characteristics	Source
<i>C. saccharoperbutylacetonicum</i> N1-4	Wild type strain	Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ)
<i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T}	Asporogenic strain of <i>C. saccharoperbutylacetonicum</i> N1-4 with point mutation in Spo0A protein	(Atmadjaja <i>et al.</i> , 2019)
<i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T} S3Vb	Asporogenic strain of <i>C. saccharoperbutylacetonicum</i> N1-4 with point mutation in Spo0A protein containing S3Vb vector	(Atmadjaja <i>et al.</i> , 2019)
<i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T} Δ rex	Asporogenic strain of <i>C. saccharoperbutylacetonicum</i> N1-4 with truncated <i>rex</i> gene	This work
<i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T} pMTL82151_apy	Asporogenic strain of <i>C. saccharoperbutylacetonicum</i> N1-4 carrying pMTL82151_apy	This work
<i>C. saccharoperbutylacetonicum</i> N1-4 <i>spo0A</i> ^{I261T} pMTL82151_fdh_pfl	Asporogenic strain of <i>C. saccharoperbutylacetonicum</i> N1-4 carrying pMTL82151_fdh_pfl plasmid	This work
NEB® 10-beta Electrocompetent <i>E. coli</i>	Δ (<i>ara-leu</i>) 7697 <i>araD139 fhuA</i> Δ <i>lacX74 galK16 galE15 e14-</i> Φ 80d <i>lacZ</i> Δ M15 <i>recA1 relA1 endA1</i> <i>nupG rpsL</i> (Str ^R) <i>rph spoT1</i> Δ (<i>mrr-hsdRMS-mcrBC</i>)	New England Biolabs® Inc. (Hitchin, United Kingdom)
NEB® Turbo Competent <i>E. coli</i>	F' <i>proA</i> ⁺ <i>B</i> ⁺ <i>lacI</i> ^q Δ <i>lacZ</i> M15 / <i>fhuA2</i> Δ (<i>lac-proAB</i>) <i>glnV galK16</i> <i>galE15 R(zgb-210::Tn10)</i> Tet ^S <i>endA1</i> <i>thi-1</i> Δ (<i>hsdS-mcrB</i>)5	New England Biolabs® Inc. (Hitchin, United Kingdom)

Table 2.1 – Information of bacteria strains used in this work.

2.3 Bacteria Growth Media

Medium components were dissolved in ultrapure water. Medium pH was measured using a SevenEasy™ S20 pH meter (Mettler Toledo, Leicester, United Kingdom). pH adjustment was performed with 1 M or 2 M HCl, NaOH or H₂SO₄ (Fisher Scientific, Loughborough, United Kingdom). All media were sterilised prior to use by heating to 126 °C in a benchtop autoclave (Classic Extended model, Prestige Medical, Blackburn, United Kingdom) or by filtering through a sterile filter. For volumes ≤ 50 mL, the liquid was aspirated into an appropriate syringe and filtered through a 0.2

μm syringe filter (Corning® Inc., Amsterdam, Netherlands). For volumes > 50 mL, media was vacuum filtered into a sterile DURAN® bottle (VWR, Leicestershire, United Kingdom) through a disposable vacuum filter (Corning® Inc., Amsterdam, Netherlands).

Solid (agar) media were made by making liquid media and adding agar powder (Oxoid™, Basingstoke, United Kingdom) to a concentration of 1.5 % (w/v) prior to autoclaving. All agar plates were poured into 90 mm circular Petri dishes (Sterilin™, Thermo Scientific™, Loughborough, United Kingdom).

Antibiotics were added to maintain plasmids where appropriate, at the following concentrations: 100 $\mu\text{g mL}^{-1}$ for ampicillin, 25 $\mu\text{g mL}^{-1}$ for chloramphenicol, 40 $\mu\text{g mL}^{-1}$ for erythromycin in *C. saccharoperbutylacetonicum* N1-4 on solid media or 100 $\mu\text{g mL}^{-1}$ in liquid media or 500 $\mu\text{g mL}^{-1}$ for erythromycin in *E. coli* cultures, 50 $\mu\text{g mL}^{-1}$ for kanamycin, 75 $\mu\text{g mL}^{-1}$ for thiamphenicol. All antibiotics were purchased from Sigma-Aldrich (Darmstadt, Germany).

2.3.1 Lysogeny Broth (LB)

LB consisted of the following, per litre: 5 g yeast extract (Oxoid™, Basingstoke, United Kingdom), 10 g tryptone (Oxoid™, Basingstoke, United Kingdom), 10 g NaCl (Fisher Scientific, Loughborough, United Kingdom). pH was adjusted to between 6 and 7.

2.3.2 Super Optimal Broth with Catabolite Repression (SOC)

SOC media was prepared by preparing the following, per litre: 20 g tryptone, 5 g yeast extract, 0.58g NaCl, 0.2 g KCl and 2.46 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$. pH was adjusted to 7 with 1M NaOH, and the mixture was sterilised by autoclaving. 1M MgCl_2 and 20 % (w/v) glucose were prepared separately and sterilised by filtration. These were added to the autoclaved media at a final concentration of 10 mM MgCl_2 and 10 mM glucose.

2.3.3 *Clostridium* Growth Medium (CGM)

CGM consisted of the following, per litre: 5 g yeast extract, 0.75 g K_2HPO_4 (Fisher Scientific, Loughborough, United Kingdom), 0.75 g KH_2PO_4 (Fisher Scientific, Loughborough, United Kingdom), 0.4 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.01 g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (Fisher Scientific, Loughborough, United Kingdom), 0.01 g $\text{MnSO}_4 \cdot 4\text{H}_2\text{O}$, 1 g NaCl (Fisher Scientific, Loughborough, United Kingdom), 2 g $(\text{NH}_4)_2\text{SO}_4$ and 2 g asparagine monohydrate. All components were dissolved in ultrapure water and sterilised by autoclaving. pH was adjusted to be between 6 and 7 using 1 M HCl.

2.3.4 Reinforced *Clostridium* Medium (RCM)

RCM was purchased as a dried pre-made powder from Oxoid™ (Basingstoke, United Kingdom) and consisted of the following, per litre: 13 g yeast extract, 10 g peptone, 5 g glucose, 1 g soluble starch,

5 g sodium chloride, 3 g sodium acetate, 0.5 g cysteine hydrochloride and 0.5 g agar. RCM was prepared by dissolving 38 g pre-made powder in one litre of ultrapure water.

2.3.5 Defined *Clostridium* Growth Medium (DCGM)

DCGM was prepared following the recipe of Au et al., 2014) and consisted of the following, per litre: 0.75 g KH_2PO_4 , 0.75 g K_2HPO_4 , 1 g NaCl, 3.3 g ammonium acetate, 0.04 g CaCl_2 (Fisher Scientific, Loughborough, United Kingdom), 0.35 g MgSO_4 , 0.01 g $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.01 g $\text{FeSO}_4 \cdot 7 \text{H}_2\text{O}$, 0.004 g *p*-aminobenzoic acid (PABA) (Fisher Scientific, Loughborough, United Kingdom), 0.00001 g biotin. Components were prepared individually as stock solutions at the following concentrations: KH_2PO_4 – 100 x, K_2HPO_4 – 10 x, NaCl – 10 x, ammonium acetate – 10 x, CaCl_2 – 100 x, MgSO_4 – 100 x, $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ – 1000 x, $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ – 1000 x, PABA – 1000 x, biotin – 1000 x. All reagents except for PABA, biotin, $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ and $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ were sterilised by autoclaving. These compounds were sterilised by filtration. Reagents were combined in sterile ultrapure water to the final concentrations above.

2.3.6 TYIR Medium

TYIR medium consisted of the following, per litre: 2.5 g yeast extract (Oxoid™, Basingstoke, United Kingdom), 2.5 g tryptone (Oxoid™, Basingstoke, United Kingdom), 0.025 g FeSO_4 , $(\text{NH}_4)_2\text{SO}_4$. pH was adjusted to be between 6 and 7 using 1 M HCl.

2.3.7 Sugar Solution

Sugars glucose and xylose were dissolved in pure water and sterilised by autoclaving. These were added to the media to the respective final concentration.

2.4 Glycerol Stocks

E. coli and *C. saccharoperbutylacetonicum* N1-4 strains were stored as glycerol stocks, consisting of 15 % glycerol. Exponential phase cultures with an $\text{OD}_{600} > 1.0$ were mixed with sterile 30 % (v/v) glycerol (Fisher Scientific, Loughborough, United Kingdom) at a 1:1 ratio in a 1.8 mL cryovial. Glycerol stocks were then frozen at -80°C for long term storage.

2.5 Freeze Dried Stocks

C. saccharoperbutylacetonicum N1-4 and *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{261T} were also stored as freeze-dried pellets in sucrose, using a proprietary method developed by Biocleave Ltd. (Abingdon, United Kingdom). Freeze dried stocks were stored at 4°C .

2.6 *C. saccharoperbutylacetonicum* N1-4 Growth Conditions

All *C. saccharoperbutylacetonicum* N1-4 strains were grown under anaerobic conditions at 32°C . Anaerobic conditions were maintained either by growth in sealed serum bottles or in an anaerobic

cabinet (A35 and A85 models, Don Whitley Scientific (Bingley, United Kingdom), were used in this research). A gas mixture of 85 % N₂, 10 % CO₂ and 5 % H₂ was used in the anaerobic cabinets. Any media used in the anaerobic cabinets was transferred into the cabinet and the tube cap loosened for at least 16 h prior to inoculation. This allowed the media to equilibrate. Where sealed serum bottles of medium were used, the medium was stored at 32 °C for several hours prior to inoculation.

2.6.1 Growth from Stocks

Glycerol stocks were thawed on ice for around 20 to 30 min. Between 100 and 900 µL of glycerol stock was used to inoculate RCM, which was either in a sealed serum bottle or in the anaerobic cabinet. The RCM was mixed by inversion and incubated overnight.

Freeze dried stocks were resuspended in 1 mL of warm RCM, which was then used to inoculate sealed serum bottles of RCM. After addition of cells, the RCM was mixed by inversion and kept at 32 °C.

2.6.2 Growth from Agar Plates

A single colony was picked from an agar plate using a small sterile plastic loop. The loop was transferred into 10 mL of equilibrated culture medium in a 15 mL Falcon® tube and rubbed against the side of the tube to transfer the cells into the medium. The tube was then mixed by inversion and then incubated in an anaerobic cabinet.

2.6.3 Batch Fermentations

Starter cultures were prepared from glycerol stocks as described in section 2.6.1. The OD_{600nm} of overnight cultures was measured on an Ultrospec™ 10 spectrophotometer (Amersham Biosciences, Little Chalfont, United Kingdom). Cell motility and morphology was inspected using a BA410E microscope (MoticEurope, Barcelona, Spain).

Fermentation medium consisted of the base media listed in section 2.3, along with sugar solution to a final concentration of 50 g L⁻¹. As a pH buffer, either filter sterilised 2-(N-morpholino)ethanesulfonic acid (MES) or calcium carbonate was used at a concentration of 10 mM and 10 % (w/v), respectively.

The starter culture was used to inoculate a 100 mL serum bottle (Supelco®, Sigma-Aldrich, Darmstadt, Germany) containing 54 mL of fermentation medium (10 % v/v). The culture was grown in the anaerobic cabinet for approximately 6 h. A second bottle of fermentation medium was inoculated from the day culture (again at 10 % v/v) which was grown overnight.

6 mL of the overnight culture was used to inoculate three 100 mL serum bottles containing 54 mL of fermentation medium to begin the fermentation. During fermentations, the culture was sampled at regular intervals to measure the OD_{600nm} and pH. 1 mL samples were also transferred to a sterile 1.5 mL microfuge tube. Media was centrifuged in a bench-top microfuge at 16,000 x *g* for 5 min. The supernatant was transferred to a clean microfuge tube and stored at -20 °C for HPLC analysis.

2.7 Plasmids and Primers

Primers and plasmids used in this work are listed below in **Table 2.2**.

Plasmid	Relevant Characteristics	Source
pMTL82151	<i>E. coli</i> – <i>Clostridium</i> shuttle vector, Cm ^R	(Heap <i>et al.</i> , 2009)
pMTL83251	<i>E. coli</i> – <i>Clostridium</i> shuttle vector, Ery ^R	(Heap <i>et al.</i> , 2009)
pMTL82151_apy	<i>E. coli</i> – <i>Clostridium</i> shuttle vector expressing <i>Shigella flexneri</i> apy gene codon optimised for expression in <i>C. acetobutylicum</i> under control of Cac_4351 promoter, Cm ^R	This work
pMTL82151_fdh_pflA_pflB	<i>E. coli</i> – <i>Clostridium</i> shuttle vector expressing <i>Candida boidinii</i> fdh codon optimised for expression in <i>C. acetobutylicum</i> and pflA and pflB genes amplified directly from <i>E. coli</i> . fdh and pflB genes were under control of the <i>C. saccharoperbutylacetonicum</i> thiolase promoter, while the native <i>E. coli</i> promoter was used for the pflA gene, Cm ^R	This work
pMTL83251_catP	<i>E. coli</i> – <i>Clostridium</i> shuttle vector expressing chloramphenicol acetyltransferase gene under control of <i>C. acetobutylicum</i> Cac_3451 promoter - Ery ^R	This work
S1Va_int_pthl_pflA_pflB	Step 1 CLEAVE integration vector containing the pflA and pflB genes from <i>E. coli</i> , with the latter under control of the <i>C. saccharoperbutylacetonicum</i> thl promoter into <i>C. saccharoperbutylacetonicum</i> genome at the formate acetyltransferase locus - Cm ^R	This work
S1Vb_int_flg	CLEAVE integration vector for genomic integration of genes into flagellar basal body rod protein flgB site of <i>C. saccharoperbutylacetonicum</i> genome- Cm ^R	Biocleave Ltd. (Abingdon, United Kingdom)
S1Vb_int_flg_pthl	CLEAVE integration vector for genomic integration of genes into flagellar basal body rod protein flgB site of <i>C. saccharoperbutylacetonicum</i> genome. Has thiolase promoter from <i>C. saccharoperbutylacetonicum</i> upstream of MCS - Cm ^R	This work

Table 2.2 (cont.)		
S1Vb_int_flg_pthl_fdh	CLEAVE integration vector for genomic integration of the <i>C. boidinii</i> <i>fdh</i> gene codon optimised for expression in <i>C. acetobutylicum</i> into <i>C. saccharoperbutylacetonicum</i> genome- Cm ^R	This work
pMTL82151_fdh	pMTL83251 vector expressing <i>fdh</i> gene from <i>C. boidinii</i> , codon optimised for expression in <i>C. acetobutylicum</i> under control of <i>thl</i> promoter from <i>C. saccharoperbutylacetonicum</i> - Cm ^R	This work
pMTL83251_fdh_pflA_pflB	pMTL83251_fdh vector with <i>pflA</i> and <i>pflB</i> genes from <i>E. coli</i> with latter under control of <i>thl</i> promoter from <i>C. saccharoperbutylacetonicum</i> - Cm ^R	This work
S1Va_Δrex	Step 1 CLEAVE vector with truncated <i>rex</i> gene from <i>C. saccharoperbutylacetonicum</i> N1-4 – for knockout of <i>rex</i> gene, Cm ^R	This work
S3Va_Δrex	Step 3 CLEAVE vector consisting of direct-repeats-spacers targeting <i>rex</i> gene of <i>C. saccharoperbutylacetonicum</i> N1-4, for knockout of <i>rex</i> gene, Cm ^R	This work
S3Va_ldr_spc_int	Step 3 CLEAVE vector for integration at the formate acetyltransferase locus, Ery ^R	Biocleave Ltd. (Abingdon, United Kingdom)
S3Vb_ldr_spc_int	Step 3 CLEAVE vector for integration at the flagellar basal body rod protein <i>flgB</i> site, Ery ^R	Biocleave Ltd. (Abingdon, United Kingdom)
S4Va	CLEAVE vector for curing CLEAVE plasmids, Cm ^R	Biocleave Ltd. (Abingdon, United Kingdom)
pTwist_amp_high_copy_fdh	Twist Bioscience vector containing <i>Candida boidinii</i> <i>fdh</i> gene codon optimised for expression in <i>C. acetobutylicum</i> , Amp ^R	Twist Bioscience
pTwist_amp_high_copy_synth_glu_o per	Twist bioscience vector containing synthetic glucose operon under control of thiolase promoter, Amp ^R	Twist Bioscience
21ABJSJP_pykA_design_pMK-RQ	GeneArt™ vector containing <i>E. coli</i> <i>pykA</i> gene codon optimised for <i>C. acetobutylicum</i> expression under control of thiolase promoter, Kan ^R	ThermoFisher Scientific, Loughboroug h, United Kingdom

Table 2.2 (cont.)

21ABJSKP_glk_design_pMK-RQ	GeneArt™ vector containing <i>E. coli glk</i> gene codon optimised for <i>C. acetobutylicum</i> expression under control of thiolase promoter, Kan ^R	ThermoFisher Scientific, Loughborough, United Kingdom
21ABJSLP_galP_design_pOK-RQ	GeneArt™ vector containing <i>E. coli galP</i> gene codon optimised for <i>C. acetobutylicum</i> expression under control of thiolase promoter, Kan ^R	ThermoFisher Scientific, Loughborough, United Kingdom
PMA-RQ-rex_KO_s3	GeneArt™ vector containing direct-repeat-spacer-direct repeat motif targeting the <i>C. saccharoperbutylacetonicum</i> N1-4 <i>rex</i> gene, Amp ^R	ThermoFisher Scientific, Loughborough, United Kingdom
21ABJQFP_ppx_synth_2_pMK-RQ (KanR)	GeneArt™ vector containing <i>E. coli ppx</i> gene codon optimised for <i>C. acetobutylicum</i> expression under control of <i>C. acetobutylicum</i> Cac_3451 promoter, Kan ^R	ThermoFisher Scientific, Loughborough, United Kingdom
21ABJQFP_ppk_synth_2_pMK-RQ (KanR)	GeneArt™ vector containing <i>E. coli ppx</i> gene codon optimised for <i>C. acetobutylicum</i> expression under control of <i>C. acetobutylicum</i> Cac_3451 promoter, Kan ^R	ThermoFisher Scientific, Loughborough, United Kingdom
Primer	Sequence	Source
rex_HR1 F	gcgctaaatataaattcaagaag	This work
rex_HR1 R	cacttaaatgaactggaatcttttaataaccgcatag	This work
rex_HR2 F	aagattaccagttcatttaagtgaagtttactgac	This work
rex_HR2 R	tacaattctagctaattctttgagttcc	This work
rex_s3_screen F	atggacaatagggatttaaagc	This work
rex_s3_screen R	tacctttagaatgtaggcctc	This work
pflA/B_e_coli F	aatccacttaagaaggtaggtgttac	This work
pflA/B_e_coli R	ctcaataaagttgccgctttacg	This work
fdh1 amplifying F	atgaagatagttctgttctttatg	This work
fdh1 amplifying R	tccaccatggttacttttatcatgctttccatag	This work
pThI F	gcgaacgcgtcgcgtaagaagaatagcag	This work
pThI R	aagcctcgagaatcctaagagaagactataggaac	This work
M13 F	tgtaaaacgacggccagt	Biocleave Ltd. (Abingdon,

Table 2.2 (cont)

		United Kingdom)
M13 R	caggaaacagctatgacc	Biocleave Ltd. (Abingdon, United Kingdom)
S1Va_pfl_gibs_F	gtacgatttacgaattcgaggctcggtaccaacgcgtaag	This work
S1Va_pfl_gibs_R	gaggatccccgggtaccgaggctagcataagcctgattgttg	This work
ppx_synth_2_geneart_gibs_F	accatgattacgaattcgagatgaccgcggcctacacaaa	This work
ppx_synth_2_geneart_gibs_R	gaggatccccgggtaccgagtagctcataaaaataagaagcc	This work
rex_KO_genomic_screen_F	ttggagtaatatgtcactttc	This work
rex_KO_genomic_screen_R	gaattttgctaaaaatctatttgaga	This work
galP_gibs_F_3	tatccatattgaccatgattacgaattcgaggcaaggcgattaagttgggtaacgc cagggg	This work
galP_gibs_R_3	ggcgttaccacaactcccactggaaagcgggcagtgatagaagg	This work
glk_gibs_F_3	cgctttccagtcgggaagtgggtaacgccagg	This work
glk_gibs_R_3	tatagtgagtcgtatggtttccgactggaaag	This work
pykA_gibs_F_3	tccagtcgggaaccatacgactcactataggg	This work
pykA_gibs_R_3	tccccgggtaccgaggctatgacccatgttaatg	This work
pThl F	gcgaacgcgtcgcgtaagaagaatagcag	This work
pThl R	aagcctcgagaatcctaaagagaagactataggaac	This work
S1Va_Int_pThl_Dir_Screen_F	gaatagcagatgctatattaaagtattttg	This work
S1Vb_fdh_amp_F	atgaccgcgggtacagcctaggtattcc	This work
S1Vb_fdh_amp_R	gaattcgtaaatcgactcgcaagatc	This work
S1Va_pfl_gibs_F	gtacgatttacgaattcgaggctcggtaccaacgcgtaag	This work
S1Va_pfl_gibs_R	gaggatccccgggtaccgaggctagcataagcctgattgttg	This work
MH 201	Proprietary	Biocleave Ltd. (Abingdon, United Kingdom)
MH 202	Proprietary	Biocleave Ltd. (Abingdon, United Kingdom)
MH 203	Proprietary	Biocleave Ltd. (Abingdon, United Kingdom)
Table 2.2 (cont.)		
VH 077	Proprietary	Biocleave Ltd. (Abingdon,

Table 2.2 – Information of plasmids and primers used in this work.

2.8 Molecular Biology Techniques

When working with *E. coli* all techniques were performed following aseptic technique in a laminar flow cabinet (Titan, Walker Safety Cabinets, Glossop, UK). When performing enzymatic experiments such as PCR and restriction digests, experiments were performed in 0.2 mL thin-walled PCR tubes (Axygen®, Amsterdam, Netherlands) in a T3000 thermocycler (Biometra, Gottingen, Germany). When microfuge tubes were used, the temperature was controlled using a water bath (type JB1, Grant instruments) or a block heater (Boekel BBA, Grant instruments).

2.8.1 Preparation of Chemically Competent *E. coli*

Chemically competent *E. coli* were prepared using a protocol developed by Biocleave Ltd. (Abingdon, United Kingdom). Chemically competent cells were stored at -80 °C.

2.8.2 Heat Shock Transformation of *E. coli*

Stocks of competent *E. coli* were thawed on ice for 20 – 30 min. Sterile 1.5 mL microcentrifuge tubes were cooled on ice, and 2 – 5 µL DNA was added to the tube. 50 µL of thawed *E. coli* was added to the microfuge tube and mixed by gently flicking the side of the tube. *E. coli* were incubated on ice for 30 min. Heat shock was performed by placing the tube in a water bath heated to 42 °C for 45 s, which was then returned to ice for 2 min. 250 µL of SOC medium warmed to 37 °C was added to the tube. The tube was transferred to an orbital shaking incubator (Model 4518, Forma Scientific, Cambridge, United Kingdom) at 37 °C rotating at 200 RPM for 1 h. After outgrowth, 200 µL of *E. coli* was plated onto an LB agar plate with the appropriate antibiotics, which was spread using a sterile plastic L-shaped spreader until dry. The plate was then stored in a stationary incubator overnight (compact incubator, LEEC, Nottingham, United Kingdom) at 37 °C.

2.8.3 Plasmid Purification from *E. coli*

Transformed *E. coli* colonies were picked from an LB-agar plate using a sterile P200 pipette tip. When isolating plasmids from previously made glycerol stocks, the top of the stock was scraped with a sterile P200 pipette tip. In both cases, the tip was ejected into between 2 and 10 mL (2 for high copy plasmids, 5-10 for low copy plasmids) LB in a 50 mL Falcon® tube (Corning® Inc., Amsterdam, Netherlands) with the appropriate antibiotics. The culture was incubated overnight in an orbital shaking incubator.

The following morning, the yellow tip was removed from the Falcon® tube and the *E. coli* were pelleted by centrifugation. For 2 mL culture volumes, the culture was split between two 1.5 mL microfuge tubes and cells were pelleted in a benchtop microcentrifuge (1-14 Microfuge, Sigma Centrifuges, Shropshire, United Kingdom) by spinning at 6800 x *g* for 2 min at room temperature. For culture volumes ≥ 5 mL, culture was pelleted in a 6-K15 centrifuge (Sigma Centrifuges, Shropshire, United Kingdom) at 5,000 x *g* for 5 min.

The supernatant was discarded, and plasmid was isolated using the GeneJET plasmid miniprep kit (Thermo Scientific™, Loughborough, United Kingdom) following the manufacturer's instructions. From this point, all centrifugation steps were performed in a benchtop microfuge at 12,000 x *g*. Briefly, cell pellet was resuspended in 250 µL resuspension solution and resuspended by pipetting (for pellets in microfuge tubes) or vortexing (for pellets in Falcon® tubes). Where Falcon® tubes were used for pelleting, the resuspended cells were transferred to a clean 1.5 mL microfuge tube. 250 µL lysis solution was added to the tube which was mixed by inverting 4 – 6 times. 350 µL neutralisation solution was added to the tube, which was mixed by inverting 4 – 6 times. The mixture was centrifuged for 5 min. The supernatant was transferred to a GeneJET spin column and centrifuged for 1 min. The column was washed twice, with 500 µL wash solution added followed by centrifugation for 1 min. An additional step centrifugation was performed to remove residual ethanol. The column was transferred to a clean microfuge tube and Tris-EDTA buffer (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0) was added to the membrane and incubated for 2 min. The tube was then centrifuged for 2 min to elute DNA. The purity and yield of DNA were assessed using a NanoDrop™ 1000 spectrophotometer (Thermo Scientific™, Loughborough, United Kingdom).

2.8.4 Restriction Digests

All restriction enzymes were purchased from New England Biolabs® (Inc., Hitchin, United Kingdom), unless otherwise stated. Restriction digests were prepared with the recommended buffer for the restriction enzyme at 1 X concentration, between 1 and 2 µg purified DNA, 10 – 20 U restriction enzyme, with ultrapure water added to a final volume of 50 µL. Digestions were performed in a thermal cycler (T3000 thermocycler, Biometra, Gottingen, Germany) at the recommended temperature for the enzyme for 1 – 4 h. Where enzymes could be heat inactivated, the reaction was incubated at the respective inactivation temperature for 20 min. Digestion was confirmed by comparison with uncut vector on an agarose gel.

2.8.5 Polymerase Chain Reaction

For molecular cloning, polymerase chain reaction (PCR) was performed using either Phusion Hot Start II DNA polymerase (Thermo Scientific™, Loughborough, United Kingdom) or Phusion® High-Fidelity DNA polymerase (New England Biolabs, Hitchin, United Kingdom), following manufacturer's instructions. PCRs were typically performed with 1 x Phusion HF buffer, dNTPs at a concentration of 200 µM each, forward and reverse primers at a concentration of 0.5 µM, Phusion polymerase at 0.02 U/µL, with nuclease free water added to the desired final volume (generally 50 – 100 µL). For each reaction, between 1pg – 10 ng of plasmid DNA or 50 – 250 ng genomic DNA was used as template. The general thermal cycler program is shown in **Table 2.3**. For reactions using Phusion® High-Fidelity DNA polymerase, the thermal cycler was warmed to 98 °C prior to the tubes being added, but pre-warming was not included when Hot Start II was used. Reactions were performed in a thermal cycler.

Step	Temperature (°C)	Time	Cycles
Initial Denaturation	98	30 s	1
Denaturation	98	10 s	
Annealing	X	30 s	30 – 35
Extension	72	X s	
Final extension	72	5 m	1

Table 2.3 – General thermal cycler program using Phusion polymerases. Annealing temperature was calculated using the online tool available on the Thermo Fisher website (www.thermofisher.com/tmcalculator). The extension time was calculated depending on the template: 15 s/kb for plasmid DNA and 30 s/kb for genomic DNA.

2.8.6 Colony PCR

Colony PCR was performed using GoTaq® G2 Flexi DNA polymerase (Promega, Southampton, United Kingdom). PCRs were performed with 1 x GoTaq® Green Flexi Reaction Buffer, 2 mM MgCl₂, dNTPs at a concentration of 0.2 mM each, 0.5 µM forward and reverse primers, 0.025 U/µL GoTaq® G2 Flexi DNA polymerase, with nuclease free water added to the desired final volume (generally 25 – 50 µL). The general PCR program is shown in **Table 2.4**.

Step	Temperature (°C)	Time	Cycles
Initial Denaturation	94	5 m	1
Denaturation	94	38 s	
Annealing	X	38 s	30 – 40
Extension	72	X s	
Final extension	72	5 m	1

Table 2.4 – General thermal cycler program for colony PCR. Annealing temperature was calculated using the online tool available on the Promega BioMath page (www.promega.com/biomath/). The extension time was 1 min/kb.

When screening *E. coli* transformants, colonies were picked from LB-agar plates using a small disposable loop and were transferred to the PCR tube by gently rubbing against the side of the tube. For *Clostridium* screening, colonies were first transferred to a lysis buffer using a small disposable loop, consisting of 24.75 μ L Tris-EDTA (pH 8.0) and 0.25 μ L proteinase K (New England Biolabs, Hitchin, United Kingdom). Lysis was performed at 55 °C for fifteen minutes, followed by proteinase inactivation by heating at 80 °C for fifteen minutes. 1 μ L of lysed cells was used as template for PCR.

2.8.7 Agarose Gel Electrophoresis

Restriction digests and PCR reactions were visualised on a Tris-Acetate-EDTA (TAE, 10 x concentration, Thermo Fisher Scientific, Loughborough, United Kingdom) agarose gel, containing between 0.8 and 1.5 % agarose (Sigma-Aldrich, Darmstadt, Germany) depending on the size of the expected product. GelRed® nucleic acid gel stain (Biotium, Fremont, California) was used to stain DNA. PCRs using the GoTaq® Green Flexi buffer were directly loaded on the gel. Other reactions (digests and PCRs) were mixed with 6 X purple gel loading dye (New England Biolabs, Hitchin, United Kingdom) prior to loading. Gels were run in FisherBrand™ (Fisher Scientific, Loughborough, United Kingdom) gel tanks at 60 - 90 V using either a Power Pac 300 or Power Pac Basic (Bio-Rad, Watford, United Kingdom). Gels were imaged on a BioDoc-It™ imaging system (UVP, Upland, California).

2.8.8 DNA Extraction from Agarose Gel

Bands were excised using a clean scalpel and transferred to a 1.5- or 2-mL microfuge tube. DNA was extracted using a Monarch® DNA gel extraction kit (New England Biolabs, Hitchin, United Kingdom) following manufacturer's instructions. Briefly, 4 volumes (ν/ω) of gel dissolving buffer were added to the tube with the gel slice. The tube was heated at 50 °C with periodical vortexing to dissolve the gel slice. The dissolved slice was loaded into a spin column in a collection tube, which was spun for 1 min. The column was washed twice with 200 μ L wash buffer followed by centrifugation. The column was transferred to a clean 1.5 mL microfuge tube. ≥ 6 μ L elution buffer was added to the column and allowed to stand for 1 min. DNA was eluted by spinning for 1 min. All centrifugation steps were performed at 16,000 x *g*. DNA recovery was evaluated using a spectrophotometer as in section 2.8.3.

2.8.9 Direct DNA Extraction from Enzymatic Reactions

PCRs and dephosphorylated vectors were purified from reactions using a Monarch® PCR & DNA Cleanup Kit (5 μ g) (New England Biolabs, Hitchin, United Kingdom), following manufacturer's instructions. Briefly, samples were diluted with binding buffer and mixed by pipetting. For DNA ≥ 2 kb a 2:1 ratio of binding buffer:sample was used; for DNA < 2 kb a 5:1 ratio was used. The diluted

sample was transferred to a spin column. From here, steps were identical to those used in the gel extraction protocol for the gel extraction kit (section 2.8.8).

2.8.10 Dephosphorylation of Digested Vector

Dephosphorylation of linearised plasmids digested with blunt restriction enzyme was performed using calf intestinal alkaline phosphatase (CIAP, Promega, Southampton, United Kingdom). After the restriction digest was performed, CIAP reaction buffer was added to 1 X concentration, with 0.01 U CIAP diluted in reaction buffer added for each pmole of DNA ends to be dephosphorylated. The reaction was incubated at 37 °C for fifteen minutes, then at 56 °C for fifteen minutes. A second aliquot of CIAP was added and the two incubation steps were repeated. Reactions were purified as described in section 2.8.9.

2.8.11 DNA Ligation

Joining digested plasmids with DNA inserts was performed as follows. Around 100 ng of vector was mixed with insert DNA at varying molar ratios (used ratios were 3:1, 1:1, 1:3, and 1:5 vector: insert). For standard ligation, ligase 10 X buffer (Promega, Southampton, United Kingdom) was added to 1 X, along with 3 U of T4 DNA ligase (Promega, Southampton, United Kingdom). Nuclease free water was added to a final volume of 10 µL. The reaction was incubated either at 4 °C overnight or at 22 °C for 3 h. For rapid ligation, the reaction setup was as above except the ligase 10 X buffer was replaced with 2 X rapid ligation buffer, and a 1:2 ratio of vector: insert was used. Incubations were performed for 5 minutes for cohesive ended ligations or for 15 minutes with blunt ended ligations, at 22 °C. An alternative method developed by Biocleave Ltd. (Abingdon, United Kingdom) was also used, but at the request of Biocleave Ltd. these details have been omitted. Ligation mixtures were directly used to transform *E. coli* as in section 2.8.2.

2.8.12 Gibson Assembly

Gibson assembly was performed using the NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs® Inc., Hitchin, United Kingdom) following manufacturer's instructions. Briefly, reactions were prepared as follows. Purified vector and insert DNA fragments were mixed at a vector: insert ratio of 1:2 for three or fewer DNA fragments and 1:1 when four or more fragments were used. NEBuilder® HiFi DNA assembly master mix was added to a final concentration of 1 X, and nuclease free water was added to a final volume of 20 µL. Reactions were incubated for between one and four hours at 50 °C. Reactions were then moved from the cycler to ice for use in *E. coli* transformations (section 2.8.2) or were stored at -20 °C for subsequent use.

2.8.13 DNA Sequencing

The sequences of constructed plasmids were determined by sending DNA to be sequenced by either Eurofins Genomics (Wolverhampton, United Kingdom) or Source BioScience (Cambridge,

United Kingdom). For Eurofins Scientific sequencing reactions, reactions were sent with the Mix2Seq kit. Sequencing samples consisted of 250 – 500 ng of purified plasmid, with sequencing primer at a concentration of 5 μ M. For Source BioScience sequencing reactions, the reaction consisted of 500 ng plasmid or 50 ng PCR product, and sequencing primer at a concentration of 3.2 μ M. Sequences were analysed using the MAFFT alignment algorithm (Kato *et al.*, 2002) on the online platform Benchling (Benchling, San Francisco, California).

2.8.14 *C. saccharoperbutylacetonicum* N1-4 Genomic DNA Extraction

Genomic DNA of *C. saccharoperbutylacetonicum* N1-4 was extracted as follows. A 50 mL overnight starter culture was prepared from a freeze-dried vial in CGM with 5 % glucose under the conditions described in section 2.6.1. The following day, the culture was split between 2 x 50 mL Falcon® tubes and centrifuged at 16,000 x *g* for 5 min in a 6-K15 centrifuge (Sigma Centrifuges, Shropshire, United Kingdom). The supernatant was discarded, and both pellets were resuspended in 500 μ L P1 buffer (QIAGEN, Düsserdorf, Germany) and transferred to a 2 mL microfuge tube. 50 mg/mL lysozyme solution (Thermo Scientific™, Loughborough, United Kingdom) was added to a final concentration of 1 mg mL⁻¹. The mixture was incubated at 37 °C for 20 min. 25 μ L of 10 % sodium dodecyl sulphate (SDS) was added, as well as proteinase K, to a final concentration of 100 μ g mL⁻¹. The solution was then mixed and incubated at 37 °C for a minimum of 1 h. 90 μ L of 5 M NaCl solution was added with 75 μ L cetrimonium bromide (CTAB)/NaCl solution. The mixture was mixed and then incubated at 65 °C for at least 20 min. 500 μ L of phenol: chloroform: isoamyl alcohol (IAA) was added, the solution was mixed then centrifuged at 6000 x *g* for 10 min at room temperature. 500 μ L of chloroform: isoamylalcohol 24: 1 was added to the supernatant, the solution was mixed and then centrifuged at 6,000 x *g* for 10 min. 600 μ L of isopropanol was added to the supernatant, then the solution was centrifuged at 10,000 x *g* for 10 min. The supernatant was discarded, and the pellet washed in 200 μ L pure ethanol by centrifuging at 10,000 x *g* for 10 min. The supernatant was discarded, and the pellet allowed to dry. The pellet was then resuspended in nuclease free H₂O (Invitrogen™, Loughborough, United Kingdom) and stored at 4 °C.

2.8.15 Electrotransformation of *C. saccharoperbutylacetonicum* N1-4 Strains

C. saccharoperbutylacetonicum N1-4 and derived strains were transformed with the following protocol. All steps were performed in a Whitley A85 anaerobic workstation, except for the centrifugation steps. All centrifugation steps were carried out in either a 3-16PK or 6-K15 centrifuge (Sigma Centrifuges, Shropshire, United Kingdom), chilled to 4 °C. An overnight culture was made as described in section 2.6.1. 6 mL of overnight culture was used to inoculate 54 mL CGM with 5 % glucose in a 250 mL narrow neck Erlenmeyer flask (Fisherbrand™, Fisher Scientific, Loughborough, United Kingdom), with a foam bung and foil cap. 0.5 mL samples were taken at 40 min intervals to

measured OD_{600nm}. Once OD_{600nm} reached 1.2, the culture was placed on ice for 15 min. The cells were pelleted by centrifugation at 5,000 x *g* for 10 min. Cells were washed with 20 mL ice cold electroporation buffer with salt (49 mL 300 mM sucrose, 0.5 mL phosphate solution (60 mM Na₂HPO₄, 440 mM NaH₂PO₄) and 0.5 mL 1 M MgCl₂) and centrifuged again. The supernatant was discarded, and the pellet was resuspended in 1 mL ice cold electroporation buffer without salt (same composition as with salt buffer except without MgCl₂). Cell suspension was incubated with 1 µg of the desired plasmid on ice for 5 min in a 2mm electroporation cuvette (Cell Projects Ltd., Harrietsham, United Kingdom). Electroporation was performed with at a voltage of 1.5 kV with a MicroPulser™ electroporator (Bio-Rad, Watford, United Kingdom). 1 mL of CGM warmed to 32 °C was added to the cuvette. Cells were allowed to recover for at least 16 h, then 200 µL of cell suspension was plated onto CGM agar plates with 5 % glucose and the appropriate antibiotics. Colonies were visible after 24-72 hours.

2.8.16 CLEAVE Genome Editing

For genomic manipulations of *C. saccharoperbutylacetonicum* N1-4 the CLEAVE genome editing method was used (Atmadjaja *et al.*, 2019). Briefly, CLEAVE uses homologous recombination to generate targeted modifications of the genome, followed by a selection process using native CRISPR/Cas machinery. Details of cloning are given in the respective results section.

2.8.16.1 CLEAVE Step 1

CLEAVE step 1 plasmids were constructed containing the DNA sequence to be integrated into the genome flanked by two homology arms, each 750 – 1000 bp long. This sequence was cloned into a step 1 CLEAVE vector. The parent strain was transformed with the plasmid as described in section 2.6.13.

2.8.16.2 CLEAVE Step 2

Recovered colonies from step 1 were picked from the CGM agar plate and were inoculated into RCM containing the appropriate antibiotic. Once the culture was in exponential phase, subcultures were performed twice a day as follows. 100 – 1000 µL of culture was transferred to a fresh tube of RCM with antibiotic. A minimum of seven subcultures was performed to generate genomic manipulations.

2.8.16.3 CLEAVE Step 3

Step 3 plasmids were designed to target cells which had not undergone the desired manipulation during step 2. Constructs were designed which replicated native *C. saccharoperbutylacetonicum* CRISPR clusters, consisting of a leader sequence, a direct repeat sequence, a spacer specific to the manipulation and a second direct repeat sequence. These constructs were synthesised using GeneArt Gene Synthesis (Invitrogen™, Loughborough, United Kingdom) and cloned into step 3

CLEAVE vectors. Cultures from step 2 were transformed with the respective step 3 CLEAVE plasmid following the protocol in section 2.8.13. Recovered colonies were screened for the respective manipulation using the colony PCR protocol in section 2.8.6.

2.8.16.4 CLEAVE Plasmid Curing

CLEAVE plasmids contain Cas recognition sequences which target pMTL80000 series vectors (Heap *et al.*, 2009) and destroy them. Mutant strains developed using CLEAVE require plasmid curing prior to transformation with pMTL series vectors. The parental strain used for this work contained an S3Vb derived vector. The plasmid was cured following a proprietary method developed by Biocleave Ltd. (Abingdon, United Kingdom). Plasmid curing was confirmed with colony PCR as described in section 2.8.6

2.8.17 RNA Isolation

Cultures were prepared as described in section 2.6.3 as for batch fermentations. After 24 hours of growth, 20 mL of culture was collected and transferred to a 50 mL Falcon® tube and placed on ice. Cells were pelleted by centrifugation at 5,000 x *g* for 10 min in a centrifuge cooled to 4 °C. Supernatant was discarded and pellets were stored at – 80 °C.

RNA was isolated using TRIzol™ reagent (Invitrogen™, Loughborough, United Kingdom) following manufacturer's instructions. Briefly, frozen cell pellets were resuspended in 1 mL TRIzol™ reagent and transferred to a clean 2 mL microfuge tube. The resuspended cells were incubated for 5 min at ambient room temperature. 0.2 mL chloroform (Thermo Scientific™, Loughborough, United Kingdom) was added to the sample, which was mixed by inversion then incubated for 2-3 min. Samples were then centrifuged at 12,000 x *g* at 4 °C for 15 min. The upper aqueous phase was transferred to a clean 1.5 mL centrifuge tube. 0.5 mL isopropanol was added to the aqueous phase, which was incubated for 10 min. Samples were then centrifuged at 12,000 x *g* at 4 °C for 10 min. The supernatant was discarded, and the pellet was resuspended in 1 mL 75 % ethanol (Fisher Scientific, Loughborough, United Kingdom). Samples were briefly vortexed then centrifuged for five minutes at 7,500 x *g* at 4 °C. The supernatant was then discarded, and the pellet was air dried for 10 min. The pellet was then resuspended in 20 µL RNase-free water (Fisher Scientific, Loughborough, United Kingdom), then incubated at 60 °C for 15 min. RNA content was quantified on a NanoDrop™ 1000 spectrophotometer. RNA was stored at -80 °C.

2.8.18 cDNA Synthesis

The TURBO DNA-free™ Kit (Invitrogen™, Loughborough, United Kingdom) was used to remove DNA from isolated RNA, following the manufacturer's instructions. Briefly, 2 µL 10 X TURBO DNase™ buffer was added to the RNA along with 1 µL TURBO DNase™ enzyme, which were mixed by pipetting. The reaction was incubated at 37 °C for 20-30 min. 2 µL DNase inactivation reagent was

added, and the sample was mixed by pipetting. Samples were incubated for 5 min, mixed by flicking the side of the tube. Samples were then centrifuged at 10,000 x *g* for 1.5 min. The supernatant was then transferred to a clean 1.5 mL microfuge tube.

cDNA was synthesised using SuperScript™ IV Reverse Transcriptase kit (Invitrogen™, Loughborough, United Kingdom) following manufacturer's instructions. Briefly, a sterile tube was prepared containing random hexamer primers (Thermo Scientific™, Loughborough, United Kingdom), dNTP mix, template RNA and nuclease free water to 13 µL, which was heated at 65 °C for 5 min to anneal primers to RNA. In a separate tube, SuperScript™ IV reaction buffer, DTT, and RNaseOUT™ RNase Inhibitor were combined and mixed, then was added to the annealed RNA mixture. The final concentrations of all components were as follows: 2.5 µM random hexamer primers, 0.5 mM dNTP mix (0.5 mM each of dATP, dCTP, dGTP and dTTP), 5 mM DTT, 2 U/µL RNaseOUT™, 1 X SuperScript™ reaction buffer, 2.5 µM random hexamers, and 1 µg RNA. This mixture was incubated at 23 °C for 10 min, then heated at 55 °C for 10 min. The SuperScript™ was inactivated by heating at 80 °C for 10 min. Synthesised cDNA was stored at -20 °C.

2.8.19 qPCR

Expression of heterologous genes was confirmed with qPCR. Reactions consisted of iTaq universal SYBR® Green Supermix (Bio-Rad, Watford, United Kingdom) at 1X, 500 nM forward and reverse primer, 100 ng cDNA, with nuclease free water to a final volume of 20 µL. qPCRs were performed in a 96-well plate (Bio-Rad, Watford, United Kingdom) in a C1000 Touch™ Thermal cycler with a CFX96™ Real-Time system Optics Module (both Bio-Rad, Watford, United Kingdom). The qPCR program is shown in table 2.5.

Step	Temperature (°C)	Time (s)	Cycles
Initial Denaturation	95	30	1
Denaturation	95	5	40
Annealing/Extension	60	30	

Table 2.5 – Thermal cycler program for qPCR.

The pullulanase gene (*puIA*) was used as an internal control for all samples (Sillers *et al.*, 2008).

2.9 HPLC Analysis of Culture Media

Supernatant samples from batch fermentations (section 2.6.3) were thawed at room temperature and were spun down at 16,000 x *g* for 5 min in a bench-top microfuge. The supernatant was transferred to a clean 1.5 mL microfuge tube. Samples were diluted 1:4 in ultrapure water, mixed by vortexing, then filtered through a 0.2 µm membrane into screw top HPLC vials (ThermoFisher, Loughborough, United Kingdom). Samples were run on an Dionex UltiMate™ 3000 HPLC (Thermo

Scientific™, Loughborough, United Kingdom) fitted with an Aminex® HPX-87H column (Bio-Rad, Watford, United Kingdom) with a RefractoMax521 refractive index detector. The mobile phase was 5 mM H₂SO₄, with oven temperature of 85 °C at a flow rate of 0.9 mL/min with a run time of 25 minutes. Standards were used to quantify glucose, acetate, butyrate, lactate, formate, ethanol, butanol, and acetone, using Chromeleon™ software v7.0 (Thermo Fisher Scientific, Loughborough, United Kingdom).

2.10 Enzyme Assays

All enzyme assays were performed in Fisherbrand™ (Fisher Scientific, Loughborough, United Kingdom) polystyrene semi micro cuvettes. Absorbance was measured on a 7315 spectrophotometer (Jenway, Stone, United Kingdom) at the appropriate wavelength for 2 – 5 min.

2.10.1 Preparation of Crude Cell Extracts

Cultures were prepared as described in section 2.8.15, except pellets were stored at -20 °C. To make cell lysates, the cell pellets were resuspended in 1 mL ice cold buffer (for formate dehydrogenase assays, 10 mM sodium phosphate buffer with 100 mM β-mercaptoethanol, pH 7.45; for chloramphenicol acetyltransferase activities, 100 mM Tris buffer pH 7.8) by pipetting and vortexing. Resuspended cells were transferred to a 2 mL microfuge tube. Cells were lysed on ice using a 150 model SFX sonicator (Branson Ultrasonics, Brookfield, Connecticut) on a 2-minute cycle of 5 ms on and 5 ms off at 40 % amplitude in pulse mode. The cell lysate was centrifuged at 16,000 x *g* for 1 h in a benchtop microcentrifuge (Heraeus Fresco 21 centrifuge, Thermo Scientific™, Loughborough, United Kingdom) cooled to 4 °C. The supernatant was transferred to a clean 1.5 mL microfuge tube and kept on ice.

When calculating the specific activity, the initial change in absorbance was used to calculate the change in absorbance at the indicated wavelength per minute. This was divided by the molar attenuation coefficient of NADH (6.22 mM⁻¹ cm⁻¹) or DTNB (14.15 mM⁻¹ cm⁻¹) to calculate the change in cofactor or substrate concentration, respectively. This was multiplied to correct for the dilution factor for the enzyme, to get the units per mL of enzyme. One unit was defined as the amount of enzyme that produced 1 μmol of NADH or 1 μmol of chloramphenicol acetylated per minute. This was divided by the measured protein concentration determined by the Bradford assay (section 2.10.4).

2.10.2 Formate Dehydrogenase Assay

For the formate dehydrogenase assay, the method of (Berríos-Rivera, Bennett and San, 2002a) was used. Briefly, a cuvette was prepared containing the following: 1.67 mM NAD⁺, 167 mM sodium formate, 100 mM β-mercaptoethanol in 10 mM phosphate buffer (pH 7.5). Baseline absorbance at

340 nm was recorded for approximately 15 s, then 10 μ L cell lysate added, mixed by inversion and then returned to the spectrophotometer. As a positive control, purified formate dehydrogenase from *C. boidinii* (Roche) was dissolved in deionised water and used in place of cell extract. The initial change in absorbance was divided by the molar extinction coefficient of NADH ($6.22 \text{ mM}^{-1} \text{ cm}^{-1}$) to calculate the production of NADH. One unit was defined as the amount of enzyme that produced 1 μ mol of NADH per minute.

2.10.3 Chloramphenicol Acetyltransferase Assay

For the chloramphenicol acetyltransferase assay, the method of (Shaw, 1975) was used. Briefly, a cuvette was prepared containing the following: 640 μ L 100 mM Tris, 100 μ L 5 mM (5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent) (Thermo Scientific™, Loughborough, United Kingdom), 200 μ L acetyl-Coenzyme A sodium salt (Sigma-Aldrich, Darmstadt, Germany) and 10 μ L cell lysate. The Tris had been warmed to 37 °C prior to the addition of the other components. The absorbance at 412 nm was recorded for approximately 30 seconds. The cuvette was then removed from the spectrophotometer and 10 μ L of 0.3 % chloramphenicol in deionised water was added. The cuvette was mixed by inversion and the absorbance at 412 nm was measured. When calculating the specific activity, the change in absorbance prior to the addition of chloramphenicol was subtracted from the change in absorbance after the addition of chloramphenicol.

2.10.4 Bradford Assay

Protein content of the crude lysates was determined using a Bradford assay. For each sample, 10 x and 100 x dilutions were made in deionised water. A series of standards (1 mg mL^{-1} , 750 $\mu\text{g mL}^{-1}$, 500 $\mu\text{g mL}^{-1}$, 250 $\mu\text{g mL}^{-1}$ and 125 $\mu\text{g mL}^{-1}$) were made from molecular biology grade bovine serum albumin (New England Biolabs, Hitchin, United Kingdom). 1 μ L of each sample or standard was added to 50 μ L Bradford – Coomassie protein assay reagent (Thermo Scientific™, Loughborough, United Kingdom). Samples were incubated for 10 min – 1 h at ambient room temperature. Absorbance at 595 nm was measured using a NanoDrop™ 1000 spectrophotometer. Samples were quantified using the Protein Bradford program on the NanoDrop 1000 software v3.8 by comparing sample absorbance against the standard curve.

2.11 COBRA Modelling

2.11.1 Model Setup

Flux balance analysis simulations was performed in Python 3.7.12. All modifications of the model, simulations and analysis were performed with the COBRApy package v0.19.0 (Ebrahim *et al.*, 2013), unless otherwise stated. The solver used for simulations was ILOG® CPLEX® optimiser v12.10 (IBM®, Armonk, New York). Reactions were removed from the model using the 'knock_out()' method for the respective reaction. Modifications of the model were performed as follows. Where necessary,

metabolites were defined as 'Metabolite' objects with the molecule's formula, name and compartment defined. Reactions were defined using 'Reaction' objects. Mass balances were checked to ensure that the reaction was mass and charge balanced unless the reaction in question was a source or sink reaction. Simulations were performed using the methods described in the respective sections. Simulations were analysed using JupyterLab (v3.3.2). Some simulations were performed by running a Python script, e.g., the double reaction deletion simulations. The results from this script were saved as a comma separated value (csv) file and loaded into JupyterLab for analysis and further simulations.

The starting model for this thesis was the genome scale metabolic model of *C. beijerinckii* iCM925 (Milne *et al.*, 2011), with the following modifications. The hydrogen ATP transport reaction was constrained to only allow the export of protons, as in the original model hydrogen import was allowed. This was to reflect the hypothesised role of this protein in maintaining the transmembrane pH gradient (Riebeling and Jungermann, 1976). The batch fermentations performed in chapters three and four used rich media containing yeast extract and peptone. ATP transport reactions for the amino acids valine, tryptophan, proline, threonine, histidine, isoleucine and leucine were present in the base model, but source reactions allowing amino acids into the system were not. These reactions were added to the model. No constraints were placed on hydrogen, carbon dioxide, proton or water transport into the cell. As *C. saccharoperbutylacetonicum* N1-4 is an obligate anaerobe, oxygen was not supplied to the model. A schematic of central metabolic reactions, reaction equations and metabolites are shown in Appendices A, B and C, respectively. The full model is available for download in .mat and .xml format at the following repository: <https://github.com/AJBones/PhD-Models>

2.11.2 External Rate Calculations

External rates of product secretion and substrate uptake were determined using standard methods (Milne *et al.*, 2011; Long and Antoniewicz, 2019). Briefly, the microbial growth rate was determined by plotting the natural logarithm of the OD₆₀₀ values against time. The linear portion of the graph was used to calculate the growth rate. Linear regression was used to quantify the slope, i.e., the growth rate. Biomass yield was determined by plotting biomass as gram dry weight (taken as 1 OD₆₀₀ = 0.28 gDW L⁻¹ (Milne *et al.*, 2011)) against glucose concentration. Linear regression was used as above to quantify the biomass yield as above. The product yields were determined by plotting product concentration against biomass concentration and using linear regression to calculate the product yield per amount of biomass. Linear regressions were performed in R v4.0.2 using the 'lm()' function. The glucose uptake rate was determined by dividing the growth rate by the biomass yield. The product secretion rate was determined by multiplying the growth rate by the product yield.

2.11.3 Flux Balance Analysis

Flux balance analysis uses linear programming to identify an optimal flux solution that maximises or minimises an objective function. The system is placed under various constraints, being the assumption of steady state (equation 1) and defined upper and lower flux ranges (equation 2).

$$1) \quad \mathbf{S} \cdot \mathbf{v} = 0$$

$$2) \quad v_{il} \leq v_i \leq v_{iu}$$

Where $\mathbf{S}$ is the stoichiometric matrix, $\mathbf{v}$ is the flux vector, v_i is the i^{th} element of $\mathbf{v}$, v_{il} and v_{iu} are the imposed lower and upper bounds for v_i . The objective function is defined according to equation 3

$$3) \quad Z = \mathbf{c}^T \mathbf{v}$$

The objective function, Z , is defined by a linear combination of fluxes. $\mathbf{c}$ is a vector consisting of weights, indicating how much reactions in the flux vector $\mathbf{v}$ contribute to the objective function. Model optimisation was performed using the 'optimize' method on the model object, which returns a 'solution' object.

Parsimonious enzyme usage FBA (pFBA) incorporates the minimisation of total fluxes as an additional objective when performing optimisation (Lewis *et al.*, 2010). The objective function is first maximised using FBA, then the minimisation of total flux is performed. pFBA is represented mathematically in equation 4. Note that all reversible reactions are split into two separate irreversible reactions, to prevent a reaction running in a negative direction and artificially lowering the total flux.

$$4) \quad \min \sum_{j=1}^m v_{irrev j}$$

$$\text{s.t. } \max v_{biomass} = v_{biomass, lb}$$

$$\text{s.t. } S_{irrev} \cdot v_{irrev} = 0$$

$$0 \leq v_{irrev} \leq v_{max}$$

Where m is the number of irreversible reactions in the metabolic network, $v_{biomass}$ is the growth rate, $v_{biomass, lb}$ is the lower bound for biomass production. pFBA was performed using the 'pfba()' function of the flux_analysis submodule of COBRApy.

2.11.4 Flux Variability Analysis

At any given FBA derived optimal solution, there are often multiple flux distributions which can achieve the same objective value, which can confound the understanding of metabolic networks (Mahadevan and Schilling, 2003). To find the flux ranges for a given solution, flux variability analysis

(FVA) was performed. FVA involves first solving the optimisation problem as with standard FBA. For each flux of interest (v_i), FVA solves two optimisation problems (equation 5) (Gudmundsson and Thiele, 2010)

$$\begin{aligned} 5) \quad & \max_v / \min_v v_i \\ & \text{s.t. } \mathbf{S} \cdot \mathbf{v} = 0 \\ & \mathbf{w}^T \mathbf{v} \geq \gamma Z_0 \\ & v_l \leq v \leq v_u \end{aligned}$$

Where $\mathbf{w}^T \mathbf{v}$ is an optimal solution generated by FBA, γ is a parameter controlling whether the analysis is done with respect to suboptimal network states ($0 \leq \gamma \leq 1$) or the optimal network state ($\gamma = 1$). The `flux_variability_analysis()` function from COBRApy's `flux_analysis` package was used to predict the upper and lower flux bounds of a reaction at a given fraction of optimum. The fraction of optimum is synonymous with the parameter γ in equation 5 if one were to investigate variability in the network at 90 % of the maximum achievable growth rate, the fraction of optimum parameter would be 0.9, for example.

2.11.5 Strain Design Algorithms

The strain design algorithms OptKnock (Burgard, Pharkya and Maranas, 2003) and OptForce (Ranganathan, Suthers and Maranas, 2010) were performed using the COBRA toolbox v3.0 (Heirendt *et al.*, 2019) in MATLAB version R2020b. The solver used for strain design simulations was Gurobi™ Optimiser v9.1.1 (Gurobi Optimisation, Beaverton, Oregon). The model was converted from .xml format into .mat format using the 'write_sbml_model' function of the 'io' submodule of COBRApy.

OptKnock (Burgard, Pharkya and Maranas, 2003) involves solving a bi-level optimisation problem, consisting of an inner problem which maximises the growth rate followed by an outer problem maximising production of a target chemical.

$$\begin{aligned} & \text{maximise} \quad v_{\text{chemical}} && \text{(OptKnock)} \\ & \gamma_j \\ & \text{s.t. maximise } v_{\text{biomass}} && \text{(Primal)} \\ & v_j \end{aligned}$$

subject to $\sum_{j=1}^M S_{ij}v_j = 0$

$$\left[\begin{array}{l} v_{pts} + v_{glk} = v_{glc_uptake} \\ v_{atp} \geq v_{atp_main} \\ v_{biomass} \geq v^{target_biomass} \\ v_j^{min} \cdot y_j \leq v_j \leq v_j^{max} \cdot y_j, \forall j \in M \\ y_j = \{0, 1\}, \quad \forall j \in M \end{array} \right]$$

$$\sum_{j \in M} (1 - y_j) \leq K$$

Where $v_{chemical}$ is the target reaction, v_{pts} is flux through the PTS, v_{glk} is the flux through the glucose ABC transporter, v_{atp_main} is the NGAM, v_{atp} is the total ATP forming flux, $v_{biomass}$ is the biomass reaction, $v^{target_biomass}$ is a specified lower bound of biomass formation, S is the stoichiometric matrix, M is the set of metabolic reactions, v_{glc_uptake} is the rate of glucose uptake, y_j is a binary variable of 1 if the reaction is active and 0 if the reaction is inactive. The following constraint $v_j^{min} \cdot y_j \leq v_j \leq v_j^{max} \cdot y_j, \forall j \in M$ ensures reaction flux is set to zero if variable y_j is zero. When variable y_j is one, v_j must be between the upper and lower bounds (v_j^{min} and v_j^{max} , respectively). K is the number of knockouts to search for.

OptKnock was performed as follows. A set of reactions to be removed was constructed, which included all non-essential and non-exchange reactions in the model. Essential reactions were defined as those whose removal led to a growth rate less than 10 % of a reference ‘wild type’ flux distribution which had been optimised for biomass production, with glucose uptake and NGAM as constraints. Any reaction not in this set was deemed to be non-essential. Given the intervention size K , the ‘OptKnock’ function of the COBRA toolbox was used to identify OptKnock reaction knockout sets.

OptForce (Ranganathan, Suthers and Maranas, 2010) involves performing FVA under two scenarios, a maximum growth simulation and a maximum chemical overproduction simulation. Flux ranges are calculated for all reactions in the network. These two flux ranges are compared to identify non-overlapping reaction ranges. Based on whether the flux range in the overproduction simulation is above or below the wild type, the reaction is added to the MUST^U set or the MUST^L set, respectively. Reactions which overlap with the wild type are further scrutinised. The sum of fluxes of pairs of reactions are compared to identify second order must sets. When the sum of

fluxes is below or above the wild type in the overproduction simulation, this implies that either one of the reaction fluxes in the pair (although not necessarily both) must decrease or increase, respectively. These reaction pairs are added to the MUST^{LL} or MUST^{UU} sets, respectively. This is also performed by calculating the difference in flux ranges rather than the sum and, again, comparing the flux ranges in the overproduction and wild type simulation. Here, this identifies pairs of reactions whose first reaction flux must increase, or the second reaction flux must decrease. These reactions pairs form the sets MUST^{UL} or MUST^{LU}, respectively.

The next step of OptForce is to identify how the targets identified by the MUST sets can be implemented on the wild type to reach the overproduction target. MUST sets encode Boolean choices regarding which fluxes, or flux combinations, must change. The OptForce formulation identifies the minimum number of changes to identify the minimum number of changes which must be applied so the overproducing network involves no feasible metabolic phenotypes that do not meet the production target.

OptForce was performed as follows. Optimisations were performed maximising biomass formation and butanol production. FVA was performed on both of these simulations using the 'FVAOptForce' function. Next, the MUST^U and MUST^L sets were identified using the 'findMustU' and 'findMustL' functions, respectively. Second order MUST^{UU} and MUST^{LL} sets were identified using the 'findMustUU' and 'findMustLL' functions, respectively. OptForce was then performed using the 'optForce' function, with the number of interventions and number of FORCE sets to generate as described in the main text.

2.12 ¹³C-Metabolic Flux Analysis

2.12.1 ¹³C Labelling Experiment

All tracers were purchased from Cambridge Isotope Laboratories, Tewksbury, Massachusetts. A preculture of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} was prepared as described in section 2.6.3. The culture was then subcultured in DCGM with 2 % glucose twice. The second subculture was used to inoculate six 30 mL serum bottle cultures, containing unlabelled glucose. The culture was grown for approximately six hours before a bolus of [U-¹³C]-glucose was added to 3 of the cultures and [1-¹³C]-glucose was added to the other cultures. 150 µL media samples were regularly collected for HPLC and GC-MS analysis. Cell pellets were collected for GC-MS analysis by sampling 1 mL of fermentation media, centrifuging at 16,000 x *g* for 5 min, then discarding the supernatant and storing the cell pellet at -80 °C. Growth was measured by recording the absorbance at 600 nm.

2.12.2 Preparation of Samples for Gas Chromatography-Mass Spectrometry

Isotopic labelling of proteinogenic amino acids and supernatant glucose was measured using gas chromatography-mass spectrometry (GC-MS). All steps were performed in a fume hood. All heating steps were performed using a DB-2P Dri-Block® (Techne®, Stone, United Kingdom).

Amino acids were prepared for analysis following the protocol of (Long and Antoniewicz, 2019). 500 µL 6 N HCl (Sigma-Aldrich, Darmstadt, Germany) was added to the cell pellet which was heated at 110 °C for 12-18 h. Samples were then cooled to room temperature and centrifuged at 14,000 x *g* for 5 min. 400 µL supernatant was transferred to a clean microcentrifuge tube and the supernatant was evaporated to dryness at 65 °C. Once dried, 35 µL pyridine was added to the samples, followed by 50 µL N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide (MTBSTFA) + 1 % (w/w) *tert*-butyldimethylsilyl chloride (TBDMSCIS). Samples were mixed then incubated at 60 °C for 30 min. Samples were then centrifuged at 14,000 x *g* for 1 min. 75 µL supernatant was transferred to a 0.3 mL glass GC-MS vial (Sigma-Aldrich, Darmstadt, Germany).

Supernatant glucose was prepared for analysis as follows. 20 µL supernatant samples were dried in a SpeedVac vacuum concentrator (Thermo Scientific™, Loughborough, United Kingdom). The dried sample was resuspended in 40 µL methoxyamine hydrochloride (Sigma-Aldrich, Darmstadt, Germany) in pyridine (20 mg/mL). Samples were incubated with shaking at 37 °C for 2 h. 70 µL *N*-Methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA, Sigma-Aldrich, Darmstadt, Germany) was added, and the samples were incubated with shaking at 37 °C for a further 30 min. 30 µL of sample was transferred to a GC-MS vial.

2.12.3 GC-MS of Supernatant Glucose and Amino Acid Fragments

GC-MS analysis of ¹³C-labelling was performed on an Intuvo 9000 GC system (Agilent Technologies, Santa Clara, California) with a 5977B MSD (Agilent, Santa Clara, California). For analysis of amino acids, 1 µL sample was injected with a column gas flow of 1.2 mL He per min. The oven was heated to 120 °C for 5 min, then increased to 270 °C at a rate of 4 °C per min. The temperature was then increased to 310 °C for 3 min at a rate of 20 °C per min. The run time was 40 min, with an additional 7.5 min for the initial solvent delay.

For analysis of glucose, 1 µL sample was injected with a column gas flow of 0.6 mL He per min. The oven temperature was held at 70 °C for 5 min, then was increased to 280 °C at a rate of 5 °C per min. The oven temperature was held at 280 °C for 3 min and was then decreased to 70 °C at a rate of 120 °C per min. Total run time was 39 min, with an additional 7.5 min for the initial solvent delay.

2.12.4 Analysis of GC-MS Data

GC-MS data was analysed using MassHunter qualitative workstation software (Agilent, Santa Clara, California). Chromatograms were integrated and mass spectra were extracted using the Chemstation integrator with classical baseline correction. Amino acid and glucose fragments were identified using the library search algorithm. The National Institute of Standards and Technologies (NIST, Gaithersburg, Maryland) 2011 database was used to identify compounds, with a cut-off point of 80 % match. The mass isotopomer distributions (MIDs) were extracted for the respective amino acid and exported to a csv file for input into flux estimation software. Quantitation of amino acids was performed using MassHunter quantitative workstation software (Agilent, Santa Clara, California). Quantitation was performed by generating a standard curve of the relevant amino acids.

2.12.5 Flux Estimation from ^{13}C -Dataset

A previously published metabolic network model of *C. acetobutylicum* (Au *et al.*, 2014) was used as the initial model. The following modifications were made. Formate exchange and succinate exchange reactions were added to the model, along with the glyoxylate shunt. External rates of glucose uptake and acetate, formate, lactate and butyrate secretion were calculated as for FBA (see section 2.11.2). The mass isotopomer distributions, along with the external rates were used as inputs into Isotopomer Network Compartment Analysis (INCA, MFA Suite™, Nashville, Tennessee) software (Young, 2014). Flux analysis was performed using INCA to minimise the weighted sum of squared residuals (SSR) between simulated and experimental MIDs. At least ten flux estimations were performed to attempt to identify the global optimal solution. Parameter continuation was used to identify the range of fluxes in the best fit solution. Flux precision was determined as reported by (Au *et al.*, 2014), as follows:

$$\text{Flux precision (stdev)} = [(\text{flux}_{\text{upper bound 95 \%}}) - (\text{flux}_{\text{lower bound 95 \%}})]/4$$

The model is available for download in .mat format at the following repository: <https://github.com/AJBones/PhD-Models>

2.13 Statistical Analysis

Statistical analysis was performed using scripts written in the R v4.0.2. Variance between the two samples was compared with an F-test using the 'var.test' function, and the significance of differences in sample means was assessed with a *t*-test. If a statistically significant difference in the variances of the two samples was found ($p < 0.05$) then a Welch's *t*-test was used, otherwise Student's *t*-test was used. These were performed using the 't.test' function.

Chapter 3: Exploration of Metabolic Engineering Strategies to Increase Butanol Production in *Clostridium*

3.1 Aims of Work Described in Chapter

The aim of the work described in this chapter was to explore a stoichiometric genome-scale metabolic model of solventogenic *Clostridium*. By analysing fluxes in the model with reaction deletions, pathway insertions and strain design algorithms, I developed metabolic engineering strategies to increase butanol production which I then implemented experimentally (see chapter six). At the initial stage of the project, I did not have access to the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} developed by Biocleave Ltd because of delays resolving modifications to the project agreement that would allow use of this proprietary strain. Hence, this initial phase of modelling work was based on the parental, wild type *C. saccharoperbutylacetonicum* N1-4 strain using batch-culture measurements of this strain to constrain the model.

3.2 Introduction

3.2.1 Genome-Scale Metabolic Models for Metabolic Engineering

There are many approaches to designing strategies for chemical overproduction using microorganisms, including rational design (Stephanopoulos, 2002; Nielsen and Keasling, 2016), enzyme kinetic models (Foster *et al.*, 2021), and metabolic flux analysis (Crown and Antoniewicz, 2013). Rational design is rarely successful, as the effects of network perturbation hardly ever coincide with improved production of target compound, and cellular responses to enzyme overexpression are difficult to predict (Bailey, 1991). Enzyme kinetic models have the advantage of predicting concentrations of metabolites and can be used to identify rate limiting steps in compound synthesis (Strutz *et al.*, 2019), but they require extensive parameterisation either through measuring enzyme kinetic properties or high volumes of -omics data for parameter fitting (Strutz *et al.*, 2019; Islam, Schroeder and Saha, 2021).

An alternative approach that requires minimal experimental information is FBA, which uses a model of reaction stoichiometries for the target organism, usually assembled from genome annotation (Orth, Thiele and Palsson, 2010) and values for the uptake and secretion of products, which can be obtained from experiments (O'Brien, Monk and Palsson, 2015). Genome-scale models of metabolism can be used to predict how network perturbations, such as reaction deletions (Bordbar *et al.*, 2014) and introduction of heterologous pathways (Curran, Crook and Alper, 2012) influence metabolic fluxes. The aim of this work was to use these approaches to identify targets for increased butanol production in *C. saccharoperbutylacetonicum* N1-4.

3.2.2 Metabolic Models of *Clostridium*

The first genome-scale metabolic models of the solventogenic *C. acetobutylicum* were published in 2008 (Lee *et al.*, 2008; Senger and Papoutsakis, 2008). These were followed by models of the cellulolytic *C. cellulolyticum* (Salimi, Zhuang and Mahadevan, 2010), *C. thermocellum* (Gowen and Fong, 2010; Thompson *et al.*, 2016) and more recently the acetogens *C. ljundahlii* (Nagarajan *et al.*, 2013) and *C. autoethangenum* (Norman *et al.*, 2019). The organism I used in this thesis, *C. saccharoperbutylacetonicum* N1-4, did not have a genome scale metabolic model publicly available at the time of writing. Recent analysis of solventogenic *Clostridium* genomes revealed that *C. saccharoperbutylacetonicum* N1-4 is closely related to *C. beijerinckii* NCIMB 8052 (Poehlein *et al.*, 2017). *C. beijerinckii* NCIMB 8052 has a publicly available genome scale metabolic model (Milne *et al.*, 2011). *C. beijerinckii* NCIMB 8052 and *C. saccharoperbutylacetonicum* N1-4 share similar organisation of central metabolic pathways, including glycolysis, have all ABE fermentation pathway genes and share identical organisation of the *sol*- and *bcs*-operons (Poehlein *et al.*, 2017). Given these similarities, the *C. beijerinckii* model was used to model *C. saccharoperbutylacetonicum* N1-4 metabolism.

3.3 Results

3.3.1 FBA Simulations Optimised for Growth Do Not Reflect Experimental Data

To understand the fermentative and growth characteristics of wild type *C. saccharoperbutylacetonicum* N1-4, I undertook an initial analysis of batch fermentation cultures, measuring growth rates, substrate consumption and accumulation of fermentation end-products. The concentrations of solvents, acids, glucose and the OD_{600nm} were measured across the course of the fermentation and are shown in **Fig. 3.1**. *C. saccharoperbutylacetonicum* N1-4 showed a biphasic pattern of growth, typical in batch fermentations of solventogenic *Clostridium* (Jones and Woods, 1986). During the initial six hours, the culture was in acidogenesis, as evidenced by the production of organic acids acetate, butyrate, and lactate, while the OD_{600nm} increased during this phase. The concentrations of acetone, butanol and ethanol did not increase during the six hours of the fermentation, consistent with the culture being in acidogenesis. Due to logistical issues, I was unable to take a complete growth curve, meaning there is a gap between the six- and 24-hour time points. This obviously comes with issues, mainly that the OD_{600nm} of the culture may have peaked at any time during this period. The OD_{600nm} after 24 hours decreased, which likely means that the OD_{600nm} at the 24-hour time point was already a reduction from the true peak OD_{600nm}, suggesting that the culture was in decline. The OD_{600nm} at 24 hours therefore may not be the true peak, and the growth curves presented here should not be considered accurate. What was clear despite the limitations was that by 24 hours the culture had entered solventogenesis. Acetone and butanol concentrations had increased substantially by this time point and peaked once glucose had been

exhausted from the media. Ethanol concentrations increased during solventogenesis, but the amount accumulated was not as great as for acetone or butanol. The reduction in OD_{600nm} from 24-hours onwards could be due to endospore formation, lysis of cells, or both.

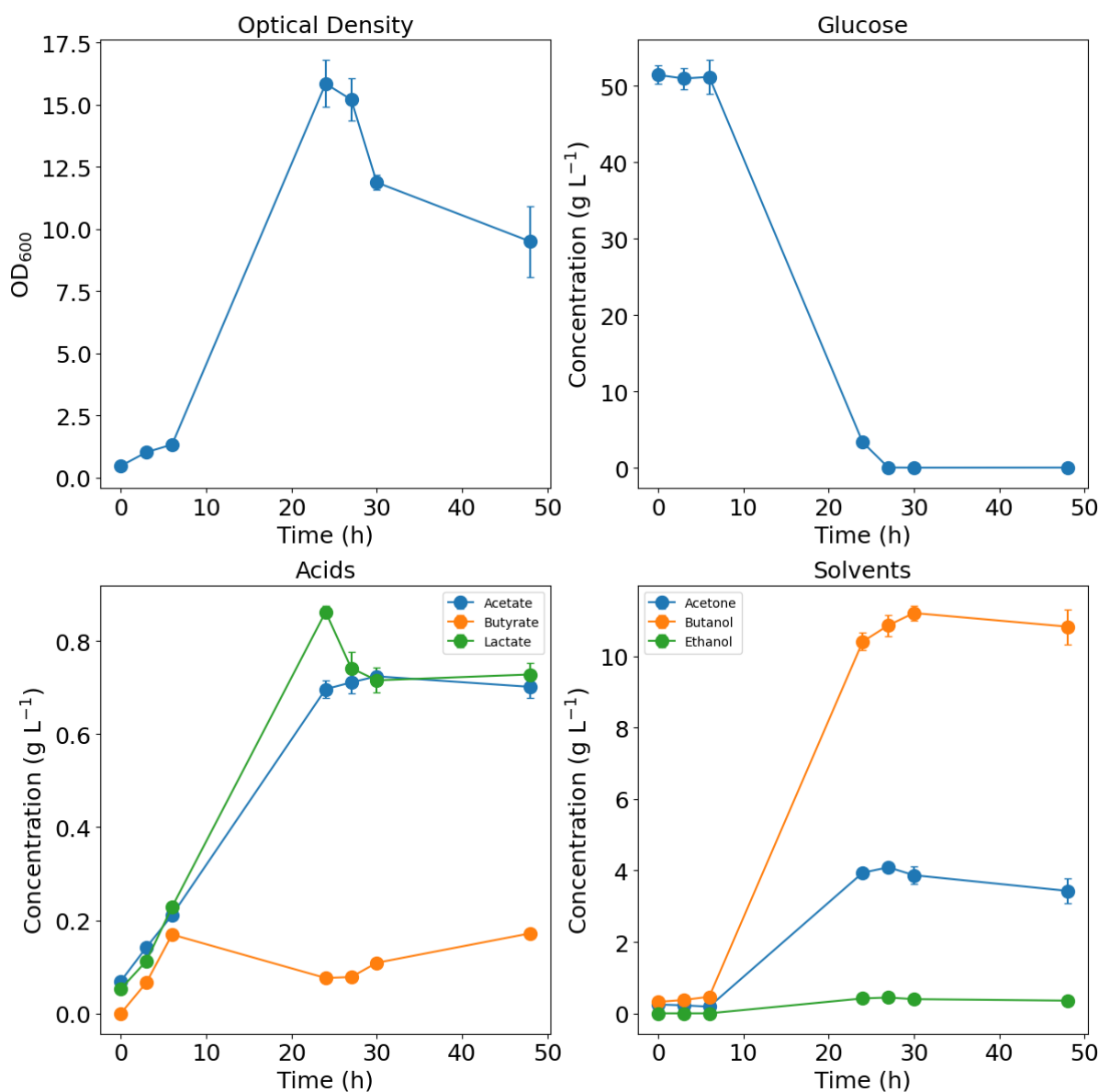


Figure 3.1 – Growth and fermentation kinetics from batch fermentations of wild type *C. saccharoperbutylacetonicum* N1-4. OD_{600nm} was determined by spectrophotometry, concentrations of glucose, acids and solvents were determined by HPLC. Batch fermentations were performed in 100 mL serum bottles with 60 mL of media. All fermentations were conducted in CGM (*Clostridia* Growth Medium) with 50 g L⁻¹ glucose and 1 % w/v CaCO₃ as a pH buffer. The starting pH of the medium was approximately 6.6. All fermentations were performed in triplicate. Error bars shown are ± 1 standard deviation.

The method of (Milne *et al.*, 2011) was followed to estimate fluxes for the rate of acetone, butanol, ethanol, acetate, butyrate and lactate production, and glucose uptake. Briefly, external fluxes were estimated as shown in equation 6:

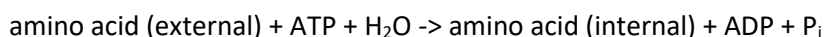
6)
$$\text{Flux} = \mu \frac{\Delta[\text{metabolite}]}{\Delta[\text{biomass}]}$$

Where μ is the specific growth rate, the [metabolite] is the concentration of the metabolite in mmol per litre and the [biomass] is the g DW of biomass per litre. A relationship of OD_{600nm} of 1 = 0.28 gDW was used to estimate the grams of biomass in the cultures (Milne *et al.*, 2011). The growth rate was determined as described by (Long and Antoniewicz, 2019). Briefly, the natural logarithm of the biomass concentration was plotted against culture time. A linear regression model was fit to the data between the six- and 24-hour time points, and the slope of the regression model was taken as the growth rate. The estimated fluxes are shown in **Table 3.1**. The lower and upper bounds for the simulations were defined as one standard deviation above or below the mean. This was performed as in the original model publication, to account for some degree of variation between the biological replicates while still replicating the observed product secretion and substrate uptake rates (Milne *et al.*, 2011).

Metabolic Process	Reaction ID	Mean Estimated Flux	Lower Bound	Upper Bound
Uptake of glucose	ex_glc_D_e	-9.02577	-9.78044	-8.2711
Secretion of acetone	ACETONEex	2.294865	2.277165	2.312565
Secretion of butanol	BUTOHex	4.767316	4.653314	4.881317
Secretion of ethanol	ETOHex	0.321893	0.293029	0.350756
Secretion of acetate	ACex	0.287374	0.276111	0.298637
Secretion of butyrate	BUTex	-0.03808	-0.04054	-0.03562
Secretion of lactate	LAC2ex	0.179115	0.163927	0.194302
Rate of biomass synthesis	biomass	0.137774	0.134361	0.141187

Table 3.1 – reaction identifiers (IDs) and constrained upper and lower reaction flux bounds from experimental estimates of nutrient consumption and product formation rates show in Fig 3.1. Upper and lower bounds of each reaction are the experimental average + or – one standard deviation, respectively. Glucose uptake was kept constant in the simulations at the mean estimated flux. Units are given as mmol gDW⁻¹ h⁻¹, except for growth rate (biomass synthesis), which is h⁻¹. All data computed were from three independent experiments. See **Appendix B** for details of reactions.

Clostridium growth medium (CGM) was used in fermentation experiments, which contained yeast extract as a source of free amino acids. To reflect this in the simulations, amino acid transport reactions were identified from the KEGG database and added to the model. It was assumed that all amino acids were imported through ATP-binding cassette transporters, the general form of import being:



I added boundary import reactions for the following amino acids: histidine, isoleucine, leucine, proline, threonine, tryptophan, and valine. Next, a non-growth associated ATP maintenance cost (NGAM) was estimated from the experimental data by adding constraints on the rate of acetate, butyrate, acetone, butanol, and ethanol secretion, and glucose uptake (constraints listed in **Table 3.1**) as upper and lower bounds through the respective exchange reactions. Then, a simulation was performed with maximisation of biomass formation as the objective value. This gave a predicted growth rate of 0.23 h^{-1} , which was higher than that seen experimentally. Many ATP consuming processes are not captured in FBA, as they do not result in conversion of chemicals. Examples include protein folding, degradation and resynthesis of DNA/RNA after turnover. To account for these processes, the lower bound through the NGAM reaction was increased sequentially. As the NGAM flux increased, the optimal growth rate decreased. The NGAM value used for the simulations was determined as that which gave a maximum growth rate which matched the experimentally determined growth rate, subject to the constraints on acid and solvent production.

This simulation had a higher growth rate than was seen experimentally, so the NGAM was increased in a stepwise manner until the flux through the biomass synthesis reaction was identical to the experimentally determined one. The estimated NGAM was $8.3 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, which was similar to that used by (Milne *et al.*, 2011) of $8.5 \text{ mmol gDW}^{-1} \text{ h}^{-1}$.

Schematics of the central metabolic reactions referred to throughout this chapter are shown in **Appendix A**, which includes the reaction identifiers and cofactors. Details of reaction names and equations are shown in **Appendix B**. Corresponding metabolite IDs and names are shown in **Appendix C**.

I first performed a simulation with all the constraints listed in **Table 3.1** added to the model as lower and upper bounds through the respective reactions. The objective in the optimisation was minimising the total sum of fluxes (termed parsimonious enzyme usage FBA, or pFBA). pFBA predictions in *E. coli* for various strains has shown that under adaptive evolution, enzyme expression data are consistent with pFBA estimations (Lewis *et al.*, 2010). From a strain design

perspective, pFBA estimates the long-term state of the cell after perturbation (Yen *et al.*, 2015).

Glucose uptake was set as the mean flux estimated from the experimental data, and the estimated NGAM value of $8.3 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ was applied. The predicted fluxes through central metabolic reactions are shown in **Fig. 3.2**. Glucose was predicted to be imported solely through the PTS, because of the minimisation of sum of fluxes constraint. The PTS imports and phosphorylates glucose in a single reaction, while the ATP-binding cassette (ABC) transporter requires a separate phosphorylation step. The imported glucose was then converted to pyruvate through the Embden-Meyerhof-Parnas (EMP) glycolysis. Around half of the pyruvate was formed from PEP dephosphorylation by pyruvate kinase, and the remainder was dephosphorylated by the PTS. There was low flux ($< 0.1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$) towards the oxidative pentose phosphate pathway (OPPP). OPPP flux was utilised to generate erythrose-4-phosphate, which entered the shikimate pathway for biosynthesis of tyrosine and phenylalanine, while tryptophan was predicted to be imported into the cell from the culture medium.

Pyruvate was predicted to be converted into acetyl-CoA mainly through pyruvate:ferredoxin oxidoreductase (PFOR), with a minor amount ($0.1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$) formed by pyruvate formate lyase (PFL). Most acetyl-CoA flux (92 %) entered the ABE fermentation pathway, synthesising acetate, acetone, ethanol, and butanol. Pyruvate also entered the tricarboxylic acid (TCA) cycle through pyruvate carboxylase. The oxaloacetate generated was used to synthesise aspartate and citrate. The citrate was converted through the TCA cycle to succinate, which was secreted from the cell. The TCA cycle intermediates α -ketoglutarate and succinyl-CoA were used to synthesise the amino acids glutamate and methionine, respectively. There was negligible predicted flux from oxaloacetate to malate, or from malate to fumarate. Fumarate was formed through the N-(L-Argininosuccinate) arginine-lyase, N6-(1,2-Dicarboxyethyl)AMP AMP-lyase and 1-(5'-Phosphoribosyl)-5-amino-4-(N-succinocarboxamide)-imidazole-lyase reactions, and was converted to succinate.

FBA solutions are often not unique, and the same objective value can be reached with different flux distributions (Mahadevan and Schilling, 2003b). Flux ranges within this solution space can give an idea of how important flux through a specific reaction is in meeting the objective function. I first wanted to examine the flux ranges and the predicted fluxes in the simulation with fermentation product constraints. I ran FVA at the experimentally determined growth rate and examined the flux ranges through central metabolic reactions (**Fig. 3.3**). Many of the reactions examined showing a high range of variability, with multiple routes possible to achieve a given flux distribution.

Glucose import could proceed through the PTS or the ATP binding cassette (ABC) transporter without impacting the growth rate. There was variation in flux ranges through glycolysis, the pentose phosphate pathway (PPP) and the TCA cycle. The ABE fermentation pathway also had a wide range of possible flux values. This is due to the stoichiometry of the ABE fermentation pathway, as one mole of glucose being converted to two moles of ethanol, or one mole of butanol, consumes the same amount of NADH; considering the two moles of NADH formed during glycolysis the production of one mole of butanol or two moles of ethanol oxidises a net two moles of NADH. This is interesting from an evolutionary perspective, as in theory ethanol formation alone would be adequate to satisfy the cells requirements for fermentation. It may be that butanol production has additional advantageous features, such as more potent antimicrobial properties or has additional

In addition, the ferredoxin: NAD(H) oxidoreductase and ferredoxin hydrogenase reactions had a wide range of acceptable fluxes. The former reaction can either oxidise reduced ferredoxin, using NAD^+ as an electron acceptor, or oxidise NADH with ferredoxin as the electron acceptor. The ferredoxin hydrogenase reaction, which oxidises reduced ferredoxin while reducing protons to form molecular hydrogen, also showed a high degree of variability.

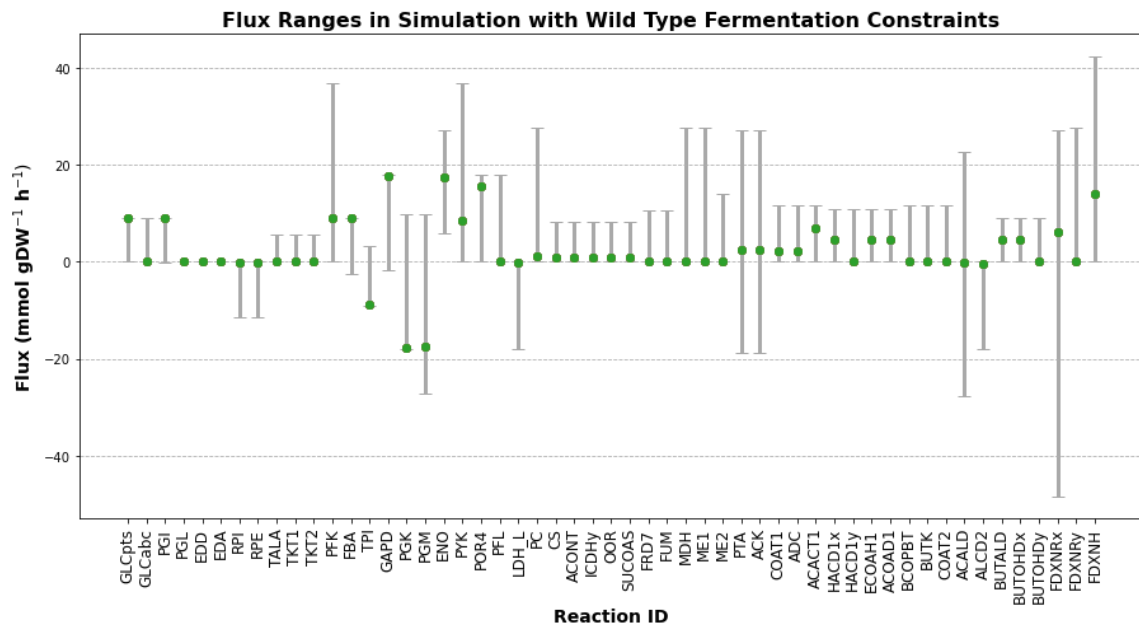


Figure 3.3 – Predicted upper and lower ranges of central metabolic fluxes in simulations of wild-type *C. saccharoperbutylacetonicum* N1-4 constrained with experimentally determined rates of glucose uptake, NGAM estimate, acetone, butanol, ethanol, lactate, acetate, and butyrate secretion rates, with minimisation of sum of fluxes as the objective. Flux ranges were calculated at the experimentally observed growth rate. Green dots indicate the predicted fluxes for the respective reactions in the simulation presented in Fig. 3.2. Reaction identifiers correspond to the names shown in Appendix B.

The usual objective function used in flux balance analysis of microbial systems is maximisation of biomass formation (Orth, Thiele and Palsson, 2010). I wanted to see how the predicted fluxes with this objective function compared to the simulation with the rates of fermentation product secretion and biomass constrained to the experimentally observed values.

I ran a pFBA simulation with the objective of maximising biomass formation while minimising the sum of fluxes, hereafter referred to as the ‘maximum growth simulation’. The only constraints applied were the NGAM cost and glucose uptake rate. This gave the flux distribution shown in Fig. 3.4. This simulation had a predicted growth rate over twice that seen in the batch fermentation experiments (0.3 h^{-1} versus 0.14 h^{-1}). As in the simulation with fermentation product constraints, glucose was converted to pyruvate through EMP glycolysis and there was low flux through the PPP. The maximum growth simulation had higher acetyl-CoA forming flux than the constrained fermentation product simulation. The only predicted ABE fermentation product in the maximum growth simulation was acetate. The TCA cycle flux was much lower in the maximum growth simulation. Most of the oxaloacetate formed through pyruvate carboxylase was used to synthesise aspartate, with less flux through citrate synthase. Another difference between the two simulations was the TCA cycle only proceeded as far as α -ketoglutarate, without appreciable flux towards succinate, as was seen in the fermentation product constrained simulation.

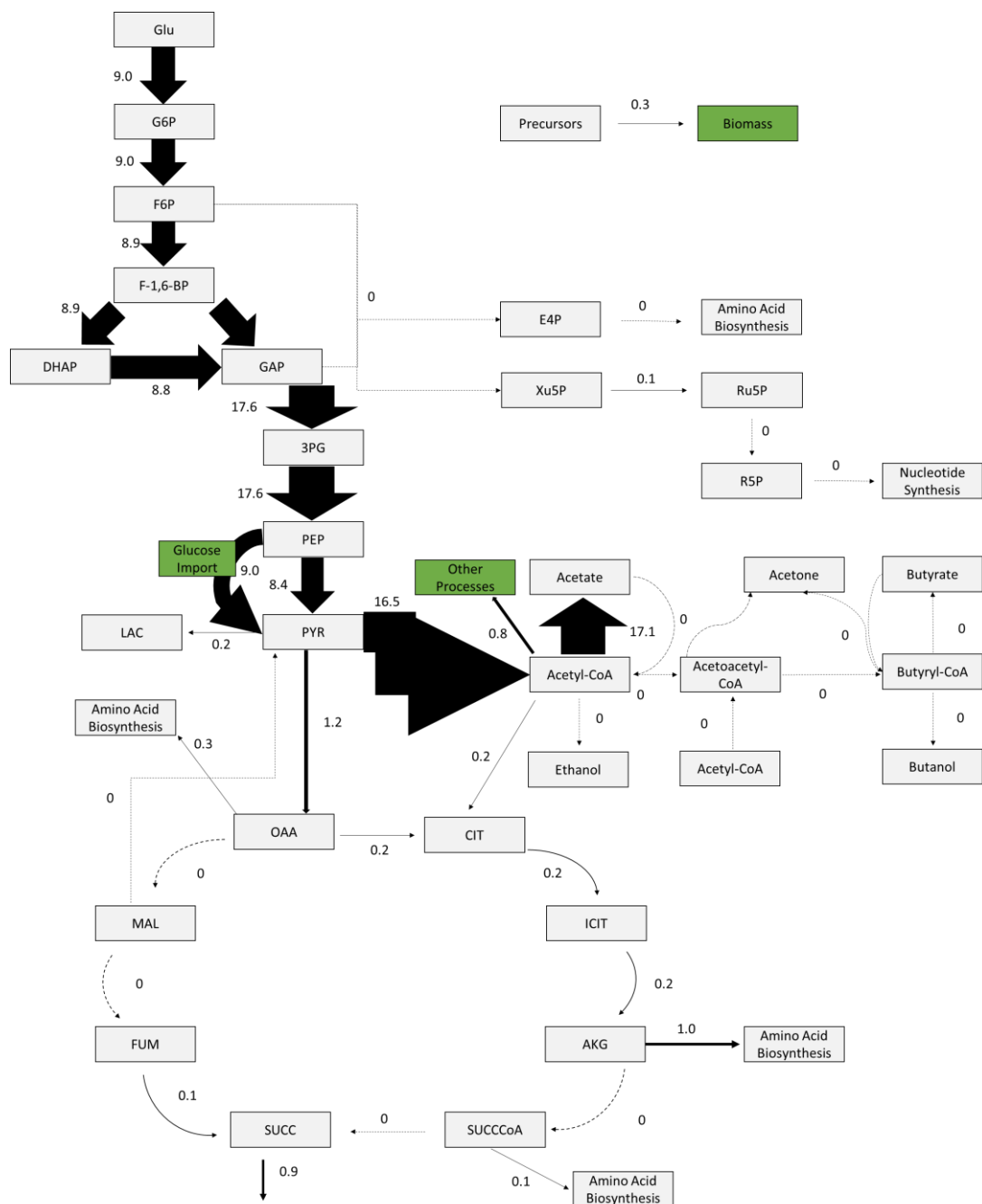


Figure 3.4 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 optimised for biomass formation with experimentally determined rates of glucose uptake and NGAM estimate. Arrow width is proportional to predicted flux through the reaction, with the flux value shown next to the respective reaction. Dashed lines indicate a reaction with a predicted flux < 0.1. The metabolites and reactions outlined in this figure are shown in detail in **Appendix A**. All fluxes are given in units of mmol gDW⁻¹ h⁻¹, except for biomass formation which is given as h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

I again investigated the variability of central metabolic reactions using FVA, this time keeping the growth rate at that achieved in the maximum growth simulation. In contrast to the flux ranges at the experimental growth rate, there was far lower variation through central metabolic reactions (Fig. 3.5). Fluxes through the PPP and TCA cycle were low, while flux through EMP glycolysis was high. The high acetate forming flux in the maximum growth simulation was necessary for the higher growth rate, as it supplied the ATP required for biomass and precursor synthesis. Production of any of the other ABE fermentation pathway end-products was not predicted to be achievable under maximum growth. In addition, high ferredoxin hydrogenase flux was also required to achieve the high growth rate. Pyruvate kinase had a degree of variation, although this was due to the reaction utilising alternative nucleoside diphosphates as phosphate acceptors (dGDP or dADP). The other reactions with high variability were PFOR and PFL. These reactions both synthesise acetyl-CoA from pyruvate, and either could be utilised without impacting the growth rate.

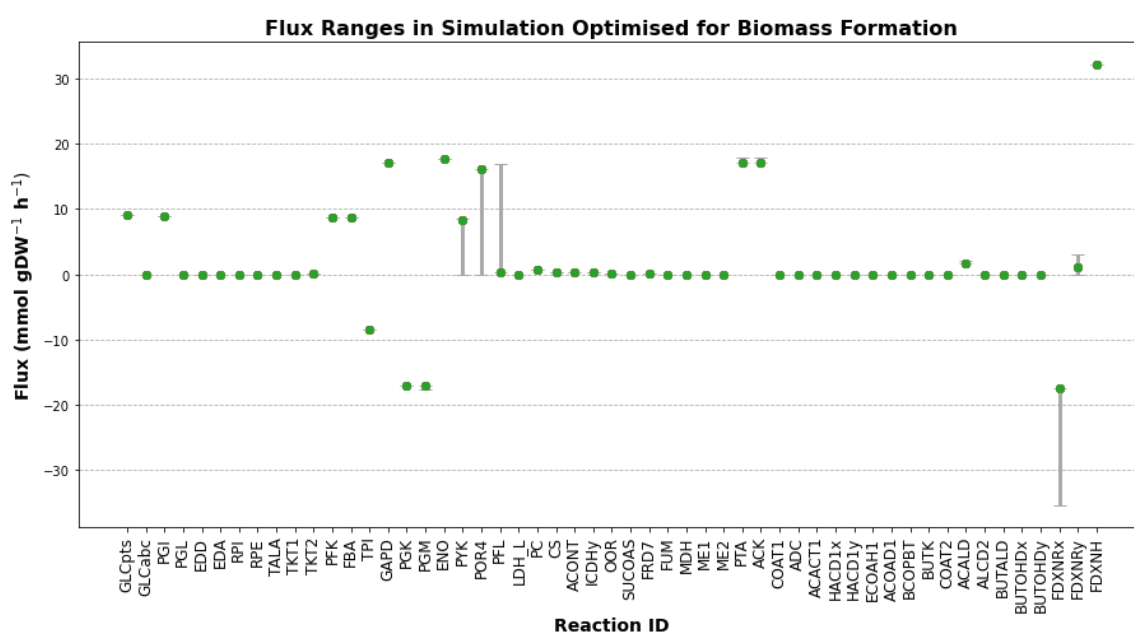


Figure 3.5 – Predicted upper and lower ranges of central metabolic fluxes in simulations of *Clostridia* metabolism with an experimentally determined glucose uptake and non-growth associated ATP maintenance cost from batch fermentation data of *C. saccharoperbutylacetonicum* N1-4 optimised for biomass formation while minimising the total sum of fluxes. Flux ranges were calculated at maximum growth. Green dots indicate the predicted fluxes for the respective reactions predicted in an FBA simulation optimised for maximising biomass formation while minimising the total sum of fluxes. Reaction identifiers correspond to the names shown in Appendix B.

The TCA cycle and PPP generate amino acids and, in the latter case, various nucleotides. Relatively small amounts of these compounds are needed to synthesise biomass, as shown by the low flux ranges. Glycolysis and acetate formation are responsible for ATP formation through substrate level phosphorylation. Given the high flux with very little flexibility in these reactions, it appeared as though the *Clostridium* growth rate is limited by ATP availability, and acetate formation resulted in the most ATP per mole glucose metabolised. To confirm this, I ran a simulation with the objective

function being the maximisation of ATP formation and minimisation of sum of fluxes. The acetate kinase flux only increased by 6 % compared to the maximum growth simulation. This increase was possible as no carbon was being used to synthesise biomass precursors in the PPP or TCA cycle. While acetate formation is clearly the most efficient method of ATP generation, I wanted to see what alternative ATP utilising reactions would be used if acetate formation was prevented. To do this, I removed the phosphate acetyltransferase reaction from the model and ran a simulation optimised for biomass formation. The predicted fluxes through central metabolism are shown in **Fig. 3.6**. The flux distribution was, for the most part, unchanged compared to the maximum growth simulation, except flux entering the ABE fermentation pathway was used to synthesise butyrate, which was the only ABE fermentation pathway product. In addition, the predicted growth rate decreased by 30 % compared to the maximum growth simulation, and the total ATP forming flux decreased by 19 %.

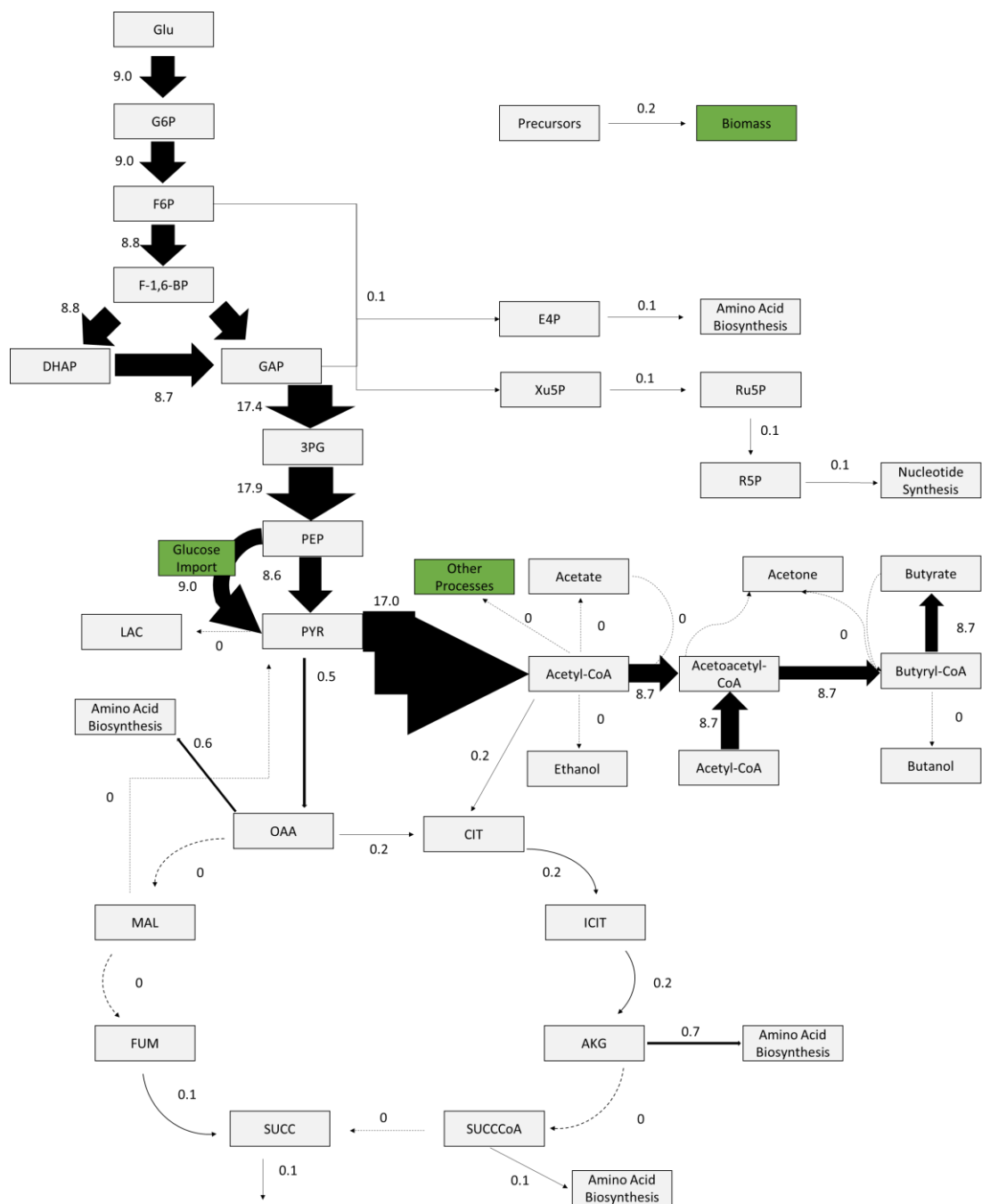


Figure 3.6 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 optimised for biomass formation, phosphate acetyltransferase reaction removed from the model with experimentally determined rates of glucose uptake and NGAM estimate. Arrow width is proportional to predicted flux through the reaction, with the flux value shown next to the respective reaction. Dashed lines indicate a reaction with a predicted flux < 0.1. The metabolites and reactions outlined in this figure are shown in detail in **Appendix A**. All fluxes are given in units of mmol gDW⁻¹ h⁻¹, except for biomass formation which is given as h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

I then checked the flux ranges through central metabolic reactions when phosphate acetyltransferase was removed from the model. As in the maximum growth simulation, EMP glycolytic reactions had high flux and PPP and TCA cycle reactions had low flux (**Fig. 3.7**). There was no variation in flux through phosphate butyryltransferase or butyrate kinase. Aside from pyruvate kinase, PFOR and PFL as already discussed, there was variation in the butyrate-acetoacetate CoA-

transferase and acetoacetate decarboxylase reactions, which catalyse the two steps of acetone formation. Flux through 3-hydroxybutyryl-CoA dehydrogenase reaction was variable because of the two possible cofactors which could be utilised in this reaction: NADH or NADPH. Unlike the maximum growth simulation, there was variation in flux through the hydrogenase reaction, although the lower bound through this reaction was above zero. This indicates that the hydrogenase is essential for oxidising ferredoxin in this simulation.

I wanted to confirm that butyrate formation is, after acetate, the most efficient fermentation product in terms of ATP generation. I removed the phosphate acetyltransferase reaction from the model and ran a simulation optimised for ATP formation (hereafter referred to as the PTA knockout

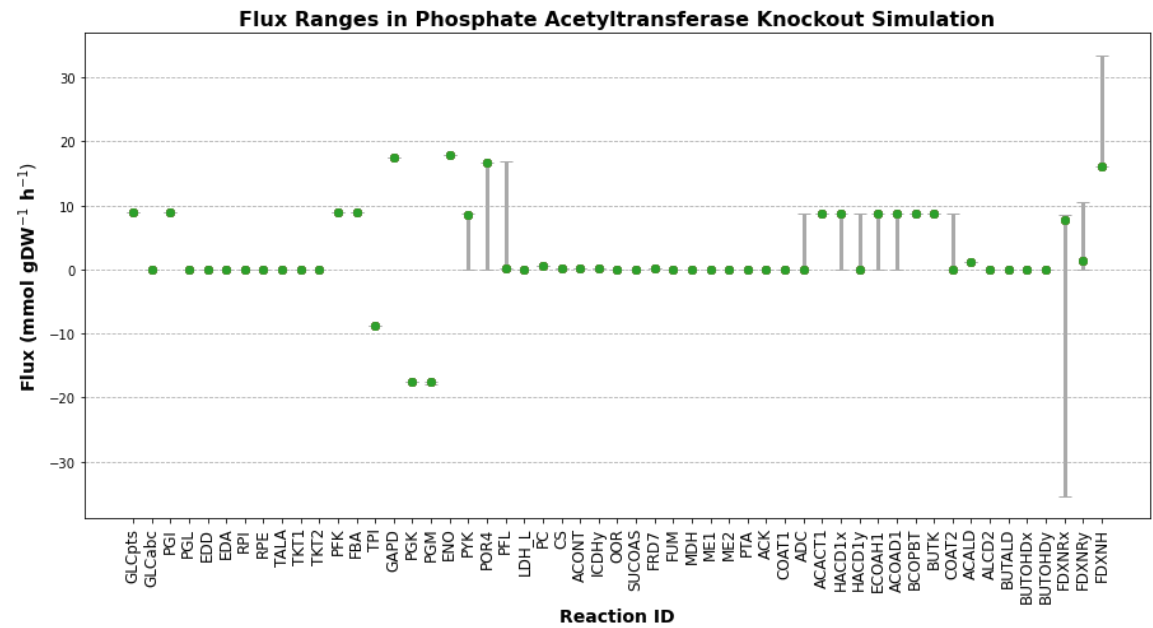


Figure 3.7 - Predicted upper and lower ranges of central metabolic fluxes in simulations of *Clostridia* metabolism with phosphate acetyltransferase removed from the model, an experimentally determined glucose uptake and non-growth associated ATP maintenance cost from batch fermentation data of *C. saccharoperbutylacetonicum* N1-4 optimised for biomass formation while minimising the total sum of fluxes. Flux ranges were calculated at maximum growth. Green dots indicate the predicted fluxes for the respective reactions predicted in an FBA simulation optimised for maximising biomass formation while minimising the total sum of fluxes. Reaction identifiers correspond to the names shown in **Appendix B**.

simulation). The butyrate kinase flux only increased by 4 % compared to the maximum growth simulation with phosphate acetyltransferase removed. This confirms that butyrate is the second most efficient ATP forming fermentation product.

The role of butyrate as a fermentation product is therefore unclear, as acetate formation produces more ATP, why would solventogenic *Clostridium* produce butyrate? As the formation of butyryl-CoA consumes NADH, it seemed that the role of butyrate would be formation of ATP and consumption of NADH. I compared the NADH usage and consumption in the PTA knockout simulation with the maximum biomass simulation (**Fig. 3.8**). Notably, the PTA knockout simulation produced more total NADH than the maximum growth simulation. The NADH generated through glycolysis was identical in both simulations, and the additional NADH in the PTA knockout was produced through the ferredoxin: NAD(H) oxidoreductase reaction. In the PTA knockout simulation, this reaction ran in the opposite direction to the maximum growth simulation (i.e., reducing NAD^+ rather than oxidising NADH). Almost all NADH was oxidised through the butyryl-CoA forming reactions in the PTA knockout simulation.

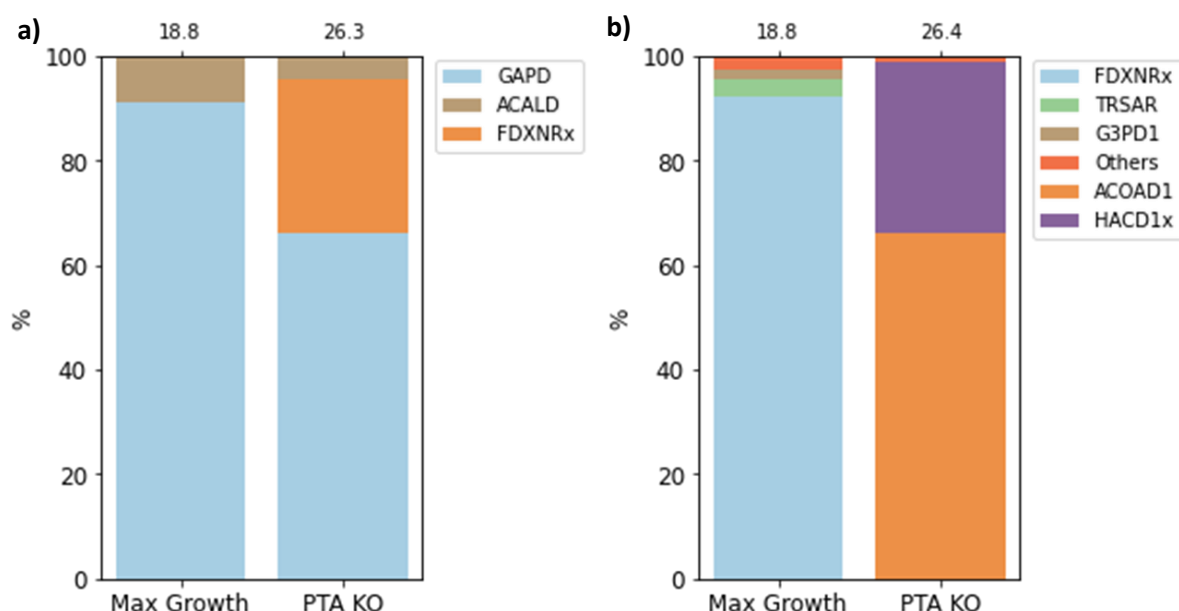


Figure 3.8 – Predicted fluxes responsible for NADH generation (a) and consumption (b) in the maximum growth simulation of *Clostridium* (labelled ‘Max Biomass’) or with the phosphate acetyltransferase reaction (PTA) removed (labelled ‘PTA KO’). The value above the bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; GAPD – glyceraldehyde-3-phosphate dehydrogenase, ACALD – acetaldehyde dehydrogenase, FDXNRx – ferredoxin: NAD(H) oxidoreductase, TRSAR – 2-hydroxy-3-oxopropionate reductase, G3PD1 – sn-Glycerol-3-phosphate: NAD⁺ 2-dehydrogenase, ACOAD1 – butyryl-CoA dehydrogenase, HACD1x – 3-Hydroxybutanoyl-CoA: NAD⁺ oxidoreductase. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/consumption flux.

This suggests that, while the most efficient method of NADH oxidation is through ferredoxin: NAD(H) oxidoreductase, the activity of this enzyme *in vivo* may not be sufficient to oxidise all the NADH produced during glycolysis. In this case, the formation of butyryl-CoA serves to oxidise NADH, which is then converted to butyrate to form ATP.

I also wanted to see what the ATP formation would be if butyrate formation was removed in addition to acetate. In this simulation, ATP was only produced through glycolysis, and lactate became the only fermentation product (**Fig. 3.9**).

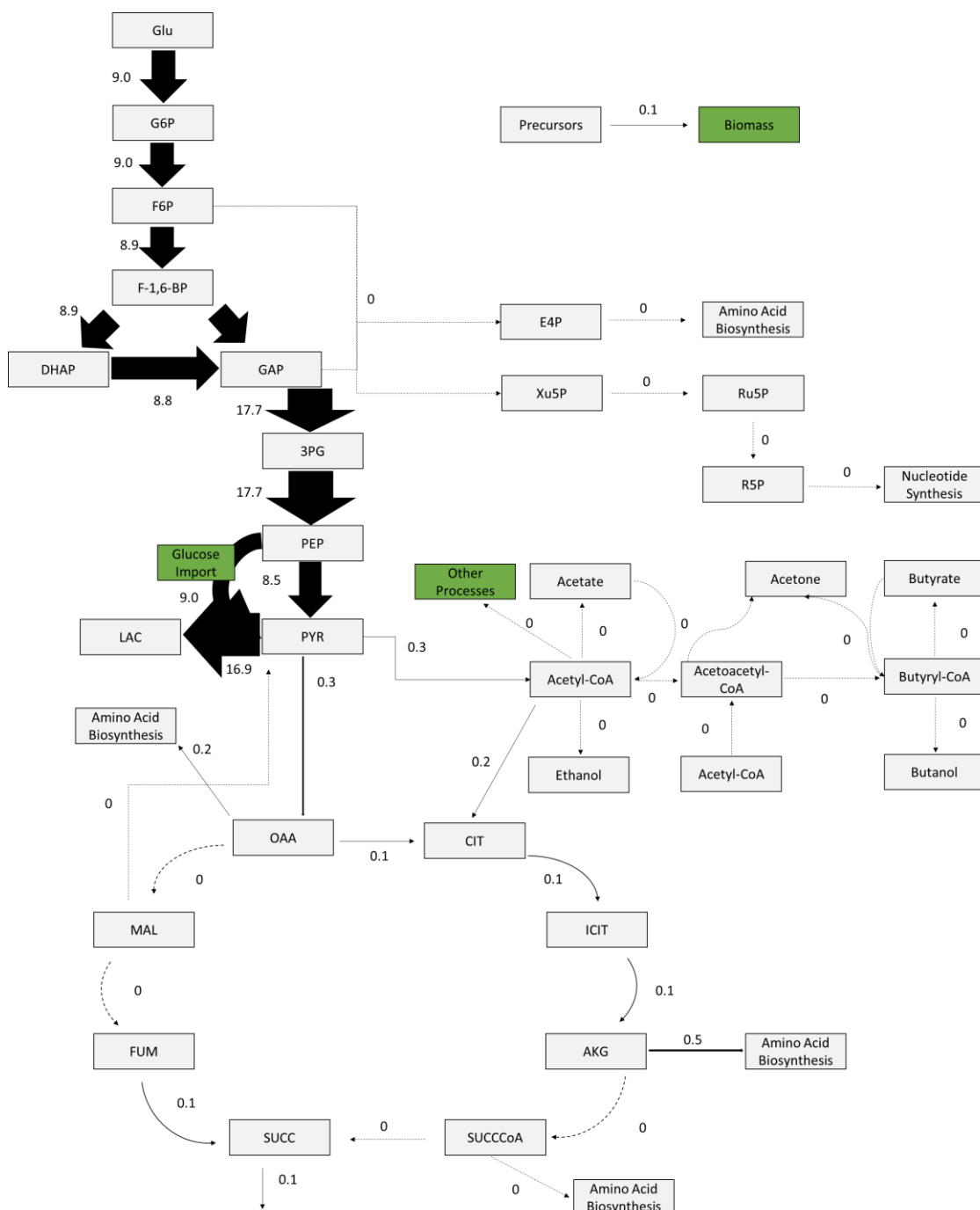


Figure 3.9 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 optimised for biomass formation, phosphate acetyltransferase and butyrate kinase reactions removed from the model with experimentally determined rates of glucose uptake and NGAM estimate. Arrow width is proportional to predicted flux through the reaction, with the flux value shown next to the respective reaction. Dashed lines indicate a reaction with a predicted flux < 0.1. The metabolites and reactions outlined in this figure are shown in detail in **Appendix A**. All fluxes are given in units of mmol gDW⁻¹ h⁻¹, except for biomass formation which is given as h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

I then checked the flux ranges when both phosphate acetyltransferase and butyrate kinase were removed from the model (**Fig. 3.10**). The pattern of flux ranges for glycolysis and the PPP were the same as in the phosphate acetyltransferase knockout. There was a large difference in the ranges through the ABE fermentation pathway and the TCA cycle. The ferredoxin: NAD(P)H oxidoreductase and ferredoxin hydrogenase reactions also had wide flux ranges.

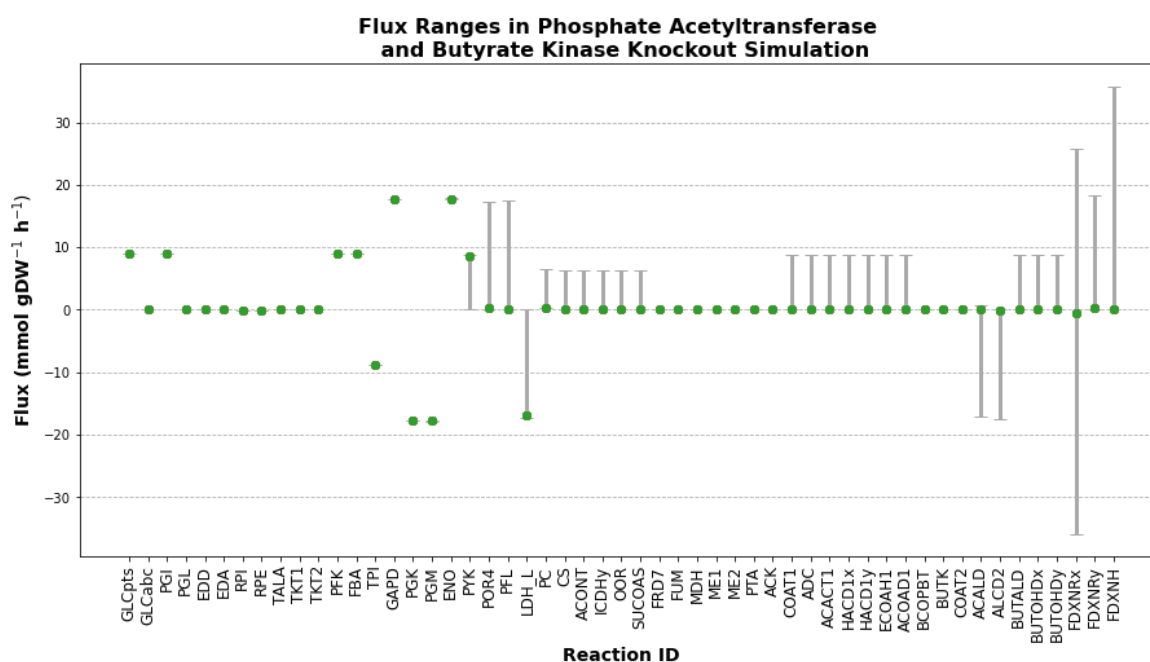


Figure 3.10 – Predicted upper and lower ranges of central metabolic fluxes in simulations of *Clostridia* metabolism with an experimentally determined glucose uptake and non-growth associated ATP maintenance cost from batch fermentation data of *C. saccharoperbutylacetonicum* N1-4, with phosphate acetyltransferase and butyrate kinase reactions removed from the model optimised for biomass formation while minimising the total sum of fluxes. Flux ranges were calculated at maximum growth. Green dots indicate the predicted fluxes for the respective reactions predicted in an FBA simulation optimised for maximising biomass formation while minimising the total sum of fluxes. Reaction identifiers correspond to the names shown in **Appendix B**.

I once again optimised for ATP formation with both acetate and butyrate pathways inactivated. Glycolysis was still the only ATP source, and the total ATP forming flux only increased by 3 % compared with the simulation optimised for biomass.

The role of various acids in the ABE fermentation pathway are as follows – acetate results in ATP formation only, butyrate results in ATP formation and NADH oxidation, and lactate results in NADH oxidation only. The wide range in fluxes through the various NADH oxidising pathways in the phosphate acetyltransferase and butyrate kinase knockout FVA show that there are numerous pathways which can oxidise NADH, but that lactate is the shortest pathway which can fully oxidise the NADH produced during glycolysis.

These simulations show that under an optimal growth scenario acetate would be the only ABE fermentation pathway product. This is because the factor limiting growth in the *C. saccharoperbutylacetonicum* N1-4 model is ATP availability, and acetate formation is the most

efficient ATP generating reaction in the ABE fermentation pathway. Each mole of glucose metabolised through glycolysis can hypothetically be converted to two moles of acetyl-CoA. Both moles of acetyl-CoA can be converted to acetate, each generating a mole of ATP, thereby generating two moles of ATP per mole of glucose metabolised. One mole of butyrate meanwhile is synthesised from two moles of acetyl-CoA and generates one mole of ATP, reducing the ATP yield per mole of glucose metabolised. While the ATP yield is reduced, each mole of butyrate formed oxidises two moles of NADH, thereby balancing the NADH formed during glycolysis. Lactate formation meanwhile does not produce any ATP, and each mole of lactate formed oxidises one mole of NADH. The physiological role of lactate is unclear, as butyrate can oxidise the same amount of NADH and produce an additional mole of ATP. Lactate may act as a method of NADH oxidation in case of capacity limits in the ABE fermentation pathway.

While the simulation optimised for biomass formation did not recreate the experimental data, it nevertheless provides a starting point with which to explore various intervention strategies.

3.3.2 Single Reaction Knockouts to Improve Butanol Production

I wanted to search for single reaction knockouts that could improve butanol production in the biomass-optimised model. The common approach to simulate how deletions impact metabolic networks is to set flux through the target reaction to zero, then predict a new optimal flux distribution. This comes with problems, particularly with large scale networks, as there are multiple flux distributions which can achieve the same objective value, termed 'alternate optima' (Lee *et al.*, 2000). My approach was to run FBA simulations for knockouts of each non-essential, non-exchange reactions in turn and to use FVA to find the range of feasible butanol secretion fluxes possible at the estimated growth rate subject to the reaction knockout. This approach ensured that any reaction knockout that could increase butanol production would be identified. Only one reaction knockout was predicted to have an upper bound of butanol secretion flux above zero, which was the ferredoxin hydrogenase reaction, with butanol secretion flux of $8.3 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. None of the other knockouts had a lower bound of butanol formation above zero. This shows that there are alternative pathways in the model which could be used to oxidise NAD(P)H, suggesting that a hydrogenase knockout alone might not be enough to force flux towards butanol formation, and that additional changes in the metabolic network may be required.

I compared the predicted fluxes in the hydrogenase knockout producing the maximum amount of butanol with the maximum growth simulation. This would identify fluxes which need to change relative to the reference flux distribution to achieve the goal of increasing butanol production. I ran a simulation where the ferredoxin hydrogenase reaction was removed from the model and butanol

formation was constrained to the upper limit predicted by FVA. This was to get an idea of what the most efficient flux distribution would be to achieve maximal butanol production.

The fluxes through central metabolism are shown in **Fig. 3.11**. The growth rate was reduced by 60 % in the hydrogenase knockout simulation, compared to the maximum growth simulation. Glycolysis was still predicted to proceed through EMP glycolysis, and there was also a slight reduction in PPP and TCA cycle flux. This was likely a result of lower biomass generation, meaning that less carbon was converted into amino and fatty acid precursors. Most carbon entered the ABE fermentation pathway, which was mostly converted to butanol, but there was also some acetate formation. No acetone, butyrate or ethanol formation was predicted.

simulation by 81 % compared to the maximum growth simulation (**Fig. 3.12**). This suggests that the amount of NADH produced from glycolysis alone would not be sufficient to provide enough reducing power for high butanol formation *in vivo*, and additional NADH forming reactions would need to be utilised.

In the maximum growth simulation, nearly all NADH formed through glycolysis was oxidised via the ferredoxin: NAD(H) oxidoreductase reaction. This demonstrates that the amount of NADH required for biomass precursors is actually very low, and that NADH formed through glycolysis is excess to cellular requirements. The reduced ferredoxin was then oxidised through the ferredoxin hydrogenase reaction. Reduced ferredoxin was also formed through the PFOR reaction, which is responsible for acetyl-CoA synthesis from pyruvate.

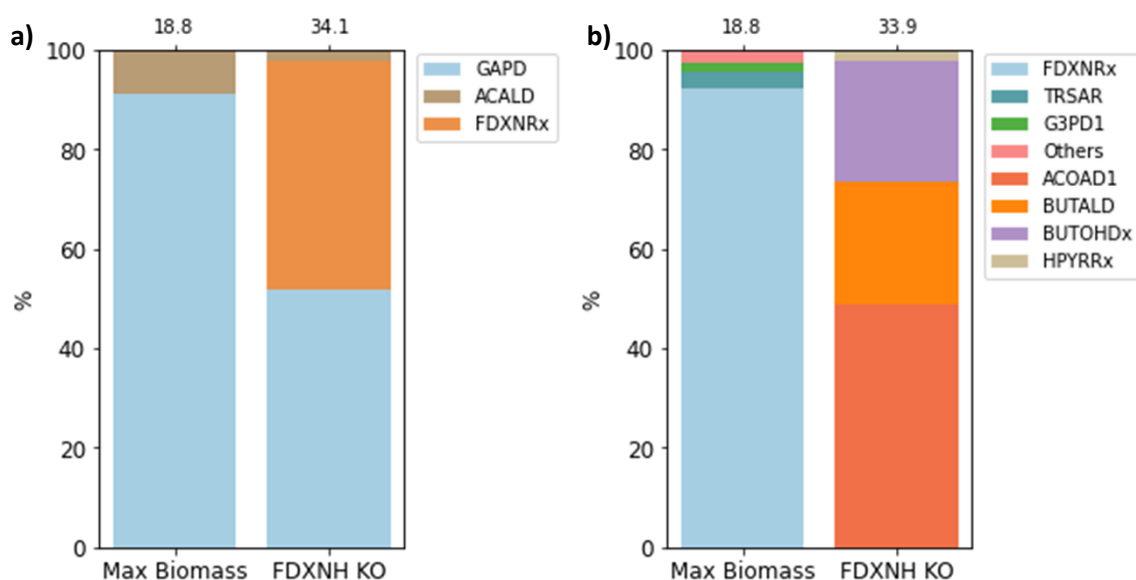


Figure 3.12 – Predicted fluxes responsible for NADH generation (a) and consumption (b) in the maximum growth simulation of *Clostridium* (labelled ‘Max Biomass’) or with the ferredoxin hydrogenase reaction removed and the butanol exchange flux constrained to the upper bound predicted by FVA (labelled ‘FDXNH KO’). The value above the bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; GAPD – glyceraldehyde-3-phosphate dehydrogenase, ACALD – acetaldehyde dehydrogenase, FDXNRx – ferredoxin: NAD(H) oxidoreductase, TRSAR – 2-hydroxy-3-oxopropionate reductase, G3PD1 – sn-Glycerol-3-phosphate: NAD⁺ 2-dehydrogenase, ACOAD1 – butyryl-CoA dehydrogenase, BUTALD – butyraldehyde dehydrogenase, BUTOHDx – NADH dependent butanol dehydrogenase, HPYRRx – glycerate dehydrogenase. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/consumption flux.

Removal of the hydrogenase reaction prevented this route of ferredoxin oxidation, and by extension oxidation of NADH. In the hydrogenase knockout, the ferredoxin: NAD(H) reaction ran in the reverse direction, oxidising ferredoxin while reducing NADH. This explained the higher NADH formation in the hydrogenase knockout simulation. NADH was oxidised almost entirely through butyryl-CoA dehydrogenase, butyraldehyde dehydrogenase and butanol dehydrogenase.

I also compared the NADPH forming fluxes in the maximum growth and hydrogenase knockout simulation (Fig. 3.13). The total NADPH forming flux increased in the ferredoxin hydrogenase knockout by nearly six-fold. In both simulations, the ferredoxin: NADP(H) oxidoreductase reaction was responsible for generating most of the NADPH. In the maximum growth simulation, NADPH was formed purely for the synthesis of amino acids and fatty acids, which were utilised for biomass formation. While there was still flux through these reactions in the hydrogenase knockout simulation, the majority of NADPH was oxidised through the 3-hydroxybutyryl-CoA dehydrogenase reaction. In these simulations flux through PFOR was similar in absolute terms, but in the hydrogenase knockout additional reduced ferredoxin was formed through the butyryl-CoA dehydrogenase reaction. The role of ferredoxin: NADP(H) oxidoreductase therefore appeared to be oxidation of this additional reduced ferredoxin.

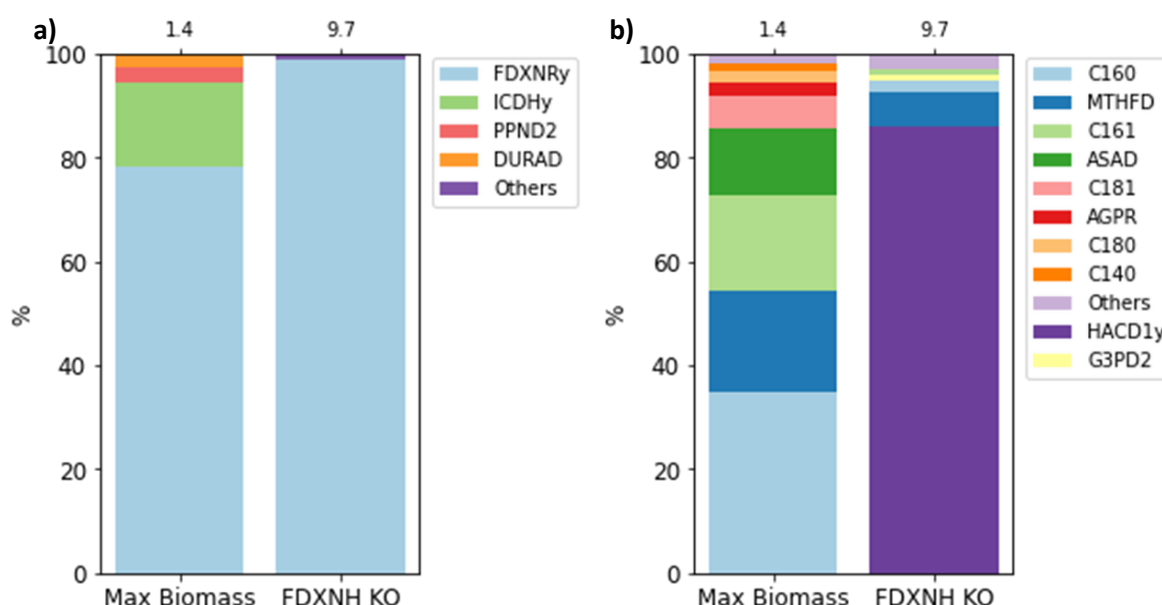


Figure 3.13 – Predicted fluxes responsible for NADPH generation (a) and consumption (b) in the maximum growth simulation of *Clostridia* metabolism or with the ferredoxin hydrogenase reaction removed and the butanol exchange flux constrained to the upper bound predicted by FVA (labelled 'FDXNH KO'). Value above bar is the total flux through all reactions shown in the bar graph, in $\text{mmol gDW}^{-1} \text{h}^{-1}$. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; FDXNRy – ferredoxin: NADP(H) oxidoreductase, ICDHy – isocitrate dehydrogenase, PPND2 – prephenate dehydrogenase, DURAD – dihydrothymine dehydrogenase, MTHFD – methylenetetrahydrofolate dehydrogenase, ASAD – aspartate semialdehyde oxidoreductase, AGPR – N-acetyl-glutamate semialdehyde oxidoreductase, HACD1y – 3-hydroxybutyryl-CoA dehydrogenase, G3PD2 – glycerol-3-phosphate dehydrogenase, C140/C160/C161/C180/C181: fatty acid synthesis. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/consumption flux.

Finally, I compared the ATP formation in these two simulations (**Fig. 3.14**). The total ATP forming flux was reduced in the hydrogenase knockout by 34 %. This was a result of lower acetate forming flux, which was 92 % lower than in the maximum growth simulation. In absolute terms the ATP forming flux from glycolysis was similar in both simulations. Unsurprisingly, biomass synthesis was the largest ATP consuming reaction in the maximum biomass simulation, followed by phosphofructokinase and the NGAM. In the hydrogenase knockout simulation, the phosphofructokinase reaction was the largest consumer of ATP, followed by NGAM and biomass synthesis.

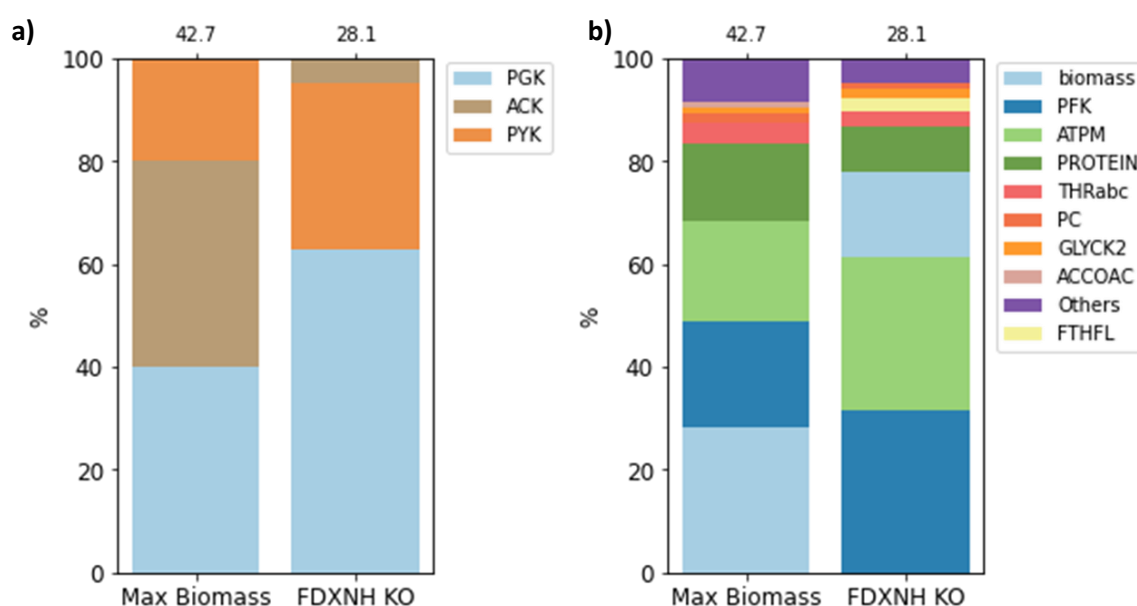


Figure 3.14 – Predicted fluxes responsible for ATP generation (a) and consumption (b) in the maximum growth simulation of *Clostridia* metabolism or with the ferredoxin hydrogenase reaction removed and the butanol exchange flux constrained to the upper bound predicted by FVA (labelled ‘FDXNH KO’). Constraints for the maximum growth simulation are shown in Table 3.3. Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Legend refers to the identifier of the reaction; PGK – phosphoglycerate kinase, ACK – acetate kinase, PYK – pyruvate kinase, biomass – lumped biomass synthesis reaction, PFK – phosphofructokinase, ATPM – NGAM, PROTEIN – lumped protein precursor for biomass synthesis reaction, THRabc – threonine-ABC transport, PC – pyruvate carboxylase, GLYCK2 – glycerate kinase, ACCOAC – Acetyl-CoA carbon dioxide ligase, FTHFL – formate: tetrahydrofolate ligase. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

To summarise, butanol formation in the hydrogenase knockout simulation was enabled by two factors; NADH generated through glycolysis was not able to be oxidised by transferring electrons to ferredoxin, and reduced ferredoxin generated through acetyl-CoA synthesis required oxidation, forming additional NADH. NADPH also played a role in butanol formation, as the final step in butyryl-CoA synthesis utilised ferredoxin as an electron acceptor. As the butanol forming flux increased in the hydrogenase knockout simulation the acetate forming flux decreased, which resulted in lower ATP formation and a lower growth rate.

As the ferredoxin: NAD(H) oxidoreductase reaction functioned in both directions in the simulation, and experimental investigations have shown this is the case *in vivo* depending on culture conditions

(Petitdemange *et al.*, 1976), I wanted to see if any other reactions could generate NADH at the levels required for high butanol production. I used FVA to investigate this. First, I identified all NADH utilising reactions in the model. I then knocked out the ferredoxin hydrogenase reaction from the model and fixed the butanol formation rate to the maximum possible. The flux ranges through all the NADH producing reactions were then compared (**Table 3.2**). Reactions with a maximum NADH forming flux $< 1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ are not shown for brevity.

Reaction ID	Minimum Flux	Maximum Flux
FDXNRx	-9.918115	24.128756
GAPD	17.713789	17.713789
FDH	0	16.922187

Table 3.2 - FVA ranges of NADH forming reactions in the ferredoxin hydrogenase knockout simulation constrained to produce butanol. Reaction IDs correspond to the following reactions: FDXNRx – ferredoxin:NAD(H) oxidoreductase, GAPD – glyceraldehyde-3-phosphate hydrogenase, FDH – formate dehydrogenase. All fluxes are in units of $\text{mmol gDW}^{-1} \text{ h}^{-1}$. The FDXNRx reaction reduces NAD^+ when running in the negative direction and oxidises NADH when running in the positive direction.

After ferredoxin:NAD(H) oxidoreductase, the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase had the second highest NADH forming flux range, of $17.7 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. The upper and lower bounds through this reaction were identical, due to its role in glycolysis. Increasing flux through this reaction would not be possible without increasing the rate of glucose uptake. The other significant NADH forming reaction was formate dehydrogenase, which had a lower bound of zero and an upper bound of $16.9 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. Aside from these reactions, none of the other NADH forming reactions in the model were predicted to have appreciable NADH forming flux (all $< 0.2 \text{ mmol gDW}^{-1} \text{ h}^{-1}$). Formate dehydrogenase utilises formate as an electron donor and generates NADH and CO_2 . 98 % of formate in the hydrogenase knockout simulation was generated through the PFL reaction, synthesising formyltetrahydrofolate, a molecule involved in one carbon metabolism. The PFL flux was low in absolute terms, being $0.7 \text{ mmol gDW}^{-1} \text{ h}^{-1}$.

As the formate forming flux was low in the hydrogenase knockout simulation, this suggested that, if formate dehydrogenase were to be utilised as an NADH generating system *in vivo*, the supply of formate would need to be increased. I utilised the same approach as above to identify all formate forming reactions in the model and see their flux ranges in the hydrogenase was knocked out simulation with forced butanol production (**Table 3.3**). The only reaction which could supply formate at the levels required was PFL, which had a predicted upper bound of 17.7 mmol gDW⁻¹ h⁻¹. Combined expression of pyruvate formate lyase and formate dehydrogenase could be used as an NADH forming system *in vivo*.

Reaction ID	Minimum Flux	Maximum Flux
PFL	0.718423	17.749303
DB4PS	0.006716	0.006716
GTPCII	0.003358	0.003358
FAPNTPDH	0	0.002177
GTPCI	0	0.002177

Table 3.3 - FVA ranges of formate forming reactions in the ferredoxin hydrogenase knockout simulation constrained to produce butanol. Reaction IDs correspond to the following reactions: PFL – pyruvate formate lyase, DB4PS - D-ribulose 5-phosphate formate-lyase, GTPCII - GTP 7,8-8,9-dihydrolase (pyrophosphate-forming), FAPNTPDH - Formamidopyrimidine nucleoside triphosphate 7,8-8,9-dihydrolase, GTPCI - GTP 7,8-8,9-dihydrolase. All fluxes are in units of mmol gDW⁻¹ h⁻¹.

3.3.3 Double Reaction Knockouts to Improve Butanol Production

As only one of the single knockouts (ferredoxin hydrogenase) was predicted to increase butanol formation, I searched for combinations of knockouts which could lead to an increase in butanol forming flux. I performed a double-reaction deletion of every non-essential reaction in the model, i.e., every possible combination of double reaction deletions. A total of 534,361 combinations were tested. As the ferredoxin hydrogenase reaction was able to achieve a high upper bound on butanol formation on its own, this was not included in the list of knockout reactions, as multiple reports have failed to generate knockouts of this gene in *C. acetobutylicum* (Cooksley *et al.*, 2012; Nguyen *et al.*, 2018). The best performing candidate reactions are shown in **Table 3.4**. The double knockout which had the highest upper bound for butanol forming flux was a double knockout of the acetate and butyrate forming pathways. Knocking out these reactions gave the highest predicted maximum butanol forming flux (8.8 mmol gDW⁻¹ h⁻¹), while the minimum butanol forming flux was zero. These knockouts also reduced the growth rate by 63 % compared to the maximum biomass simulation, due to the already discussed reduction in ATP formation. Another double knockout target was removal of NADH-dependent 3-hydroxybutyryl-CoA dehydrogenase and ferredoxin: NAD(H) oxidoreductase. The removal of these reactions reduced the growth rate by 30 % compared with the maximum biomass simulation. The minimum butanol forming flux was still zero in this knockout, and the maximum butanol forming flux was only 50 % of that in the double acid knockouts. The removal of ferredoxin: NAD(H) oxidoreductase was also a suggested knockout

combined with inactivating butyrate formation by removing phosphate butyryltransferase or butyrate kinase. These knockouts were predicted to force a minimum butanol forming flux; however, this was also the maximum achievable butanol forming flux. The butanol forming flux in these knockouts was only 40 % of the double acid knockout simulation. The final suggested knockout was ferredoxin: NAD(H) oxidoreductase combined with PFOR. This knockout again provided a guaranteed lower bound of butanol formation, but this was also the upper limit achievable.

Reaction 1	Reaction 2	Growth Rate (h^{-1})	Min Butanol (mmol $\text{gDW}^{-1} \text{h}^{-1}$)	Max Butanol (mmol $\text{gDW}^{-1} \text{h}^{-1}$)
BUTK	PTA	0.10994	0	8.75
BCOPBT	ACK	0.10994	0	8.75
PTA	BUTK	0.10994	0	8.75
BCOPBT	PTA	0.10994	0	8.75
HACD1x	FDXNRx	0.206775	0	4.41
BUTK	FDXNRx	0.226132	3.52	3.52
FDXNRx	BCOPBT	0.226132	3.52	3.52
FDXNRx	POR4	0.172474	2.29	2.29

Table 3.4 – Predicted growth rates and FVA predicted range of butanol producing fluxes from double reaction knockouts in genome scale metabolic model of *C. saccharoperbutylacetonicum*. Reaction IDs correspond to the following reactions: BUTK – butyrate kinase, PTA – phosphate acetyltransferase, ACK – acetate kinase, BCOPBT – phosphate butyryltransferase, HACD1x – 3-hydroxybutyryl-CoA dehydrogenase, FDXNRx – ferredoxin:NAD(H) oxidoreductase, POR4 – pyruvate ferredoxin oxidoreductase.

As the double acid knockout simulations were the most promising in terms of maximum achievable butanol production, I once again constrained the model to produce the maximum predicted amount of butanol with both acetate and butyrate forming pathways inactivated to enable an analysis of the overall flux distribution in this scenario (**Fig. 3.15**). Fluxes through central metabolic reactions were very similar to the ferredoxin hydrogenase knockout simulation, except that there was no acetate formation.

The ATP, NADH and NADPH generating budgets were inspected in the double acid knockout simulation and compared to those in the maximum growth simulation. The total ATP forming flux decreased by 38 % in the double acid knockout simulation than in the maximum biomass simulation as a result of removing both acid forming pathways (**Fig. 3.16**). ATP was solely formed through glycolysis when both acid forming pathways were removed. The ATP consuming fluxes in the double acid knockout simulation were similar to those in the hydrogenase knockout.

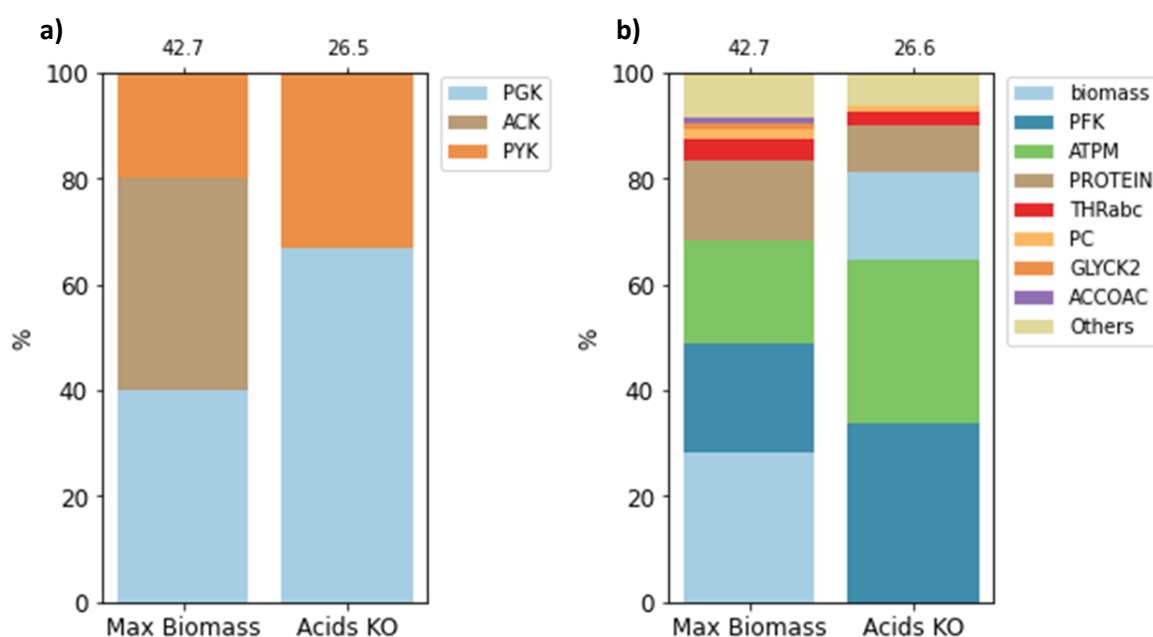


Figure 3.16 – Predicted fluxes responsible for ATP generation (a) and consumption (b) in the maximum growth simulation of *Clostridia* metabolism or with the phosphate acetyltransferase and butyrate kinase reactions removed and the butanol exchange flux constrained to the upper bound predicted by FVA (labelled 'Acids KO'). Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; PGK – phosphoglycerate kinase, ACK – acetate kinase, PYK – pyruvate kinase, biomass – lumped biomass synthesis reaction, PFK – phosphofructokinase, ATPM – NGAM, PROTEIN – lumped protein precursor for biomass synthesis reaction, THRabc – threonine-ABC transport, PC – pyruvate carboxylase, GLYCK2 – glycerate kinase, ACCOAC – Acetyl-CoA carbon dioxide ligase. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

In the double acid knockout simulation, total NADH forming flux was over double that in the maximum biomass simulation (**Fig. 3.17**). It was also a 30 % increase over that in the hydrogenase knockout simulation. Similarly to the hydrogenase knockout simulation, the ferredoxin: NAD(H) oxidoreductase reaction generated NADH in addition to that formed through glycolysis, making up the majority (58 %) of total NADH forming flux. In the double acid knockout simulation, the NADH consumption was similar to the hydrogenase knockout simulation, except that NADH was used in the 3-hydroxybutyryl-CoA dehydrogenase reaction rather than NADPH.

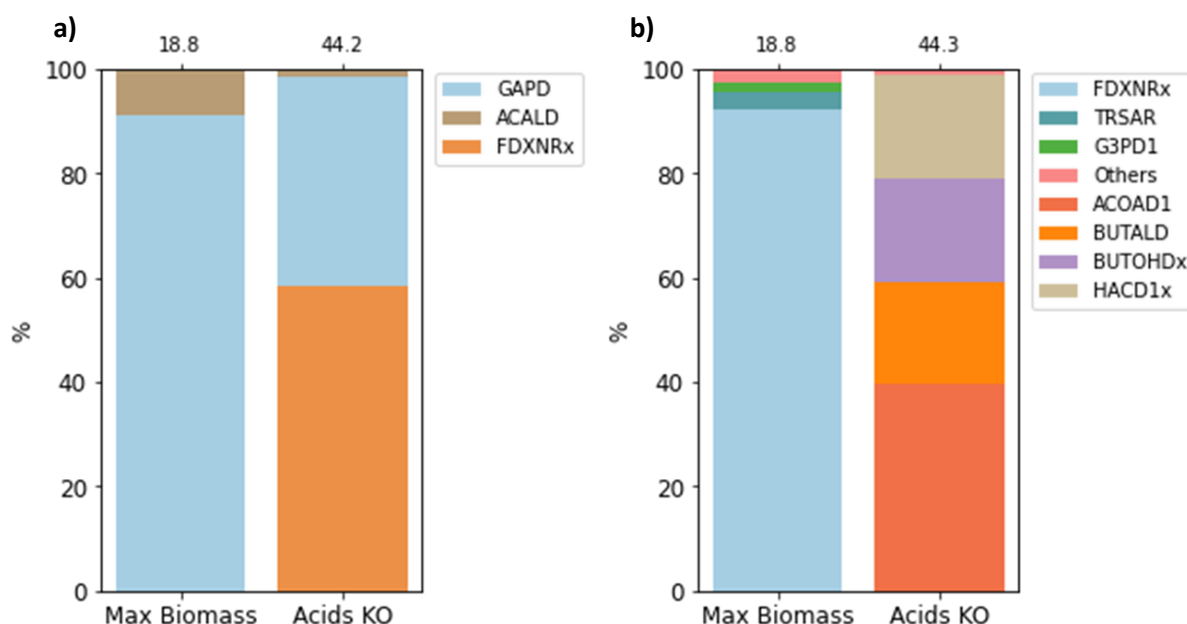


Figure 3.17 – Predicted fluxes responsible for NADH generation (a) and consumption (b) in the maximum growth simulation of *Clostridia* metabolism or with the phosphate acetyltransferase and butyrate kinase reactions removed and the butanol exchange flux constrained to the upper bound predicted by FVA (labelled ‘Acids KO’). Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; GAPD – glyceraldehyde-3-phosphate dehydrogenase, ACALD – acetaldehyde dehydrogenase, FDXNRx – ferredoxin: NAD(H) oxidoreductase, TRSAR – 2-hydroxy-3-oxopropionate reductase, G3PD1 – sn-Glycerol-3-phosphate: NAD⁺ 2-dehydrogenase, ACOAD1 – butyryl-CoA dehydrogenase, BUTALD – butyraldehyde dehydrogenase, BUTOHDx – NADH dependent butanol dehydrogenase, HACD1x – 3-Hydroxybutanoyl-CoA: NAD⁺ oxidoreductase. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/consumption flux.

In the double acid knockout simulation total NADPH forming flux decreased by 71 % compared to the maximum biomass simulation (**Fig. 3.18**). The NADPH formation was solely to synthesise fatty acids and amino acids for biomass synthesis, without any use in solvent formation. The double acid knockout simulation generated most NADPH through the ferredoxin: NAD(P)H oxidoreductase reaction, with some production from isocitrate dehydrogenase.

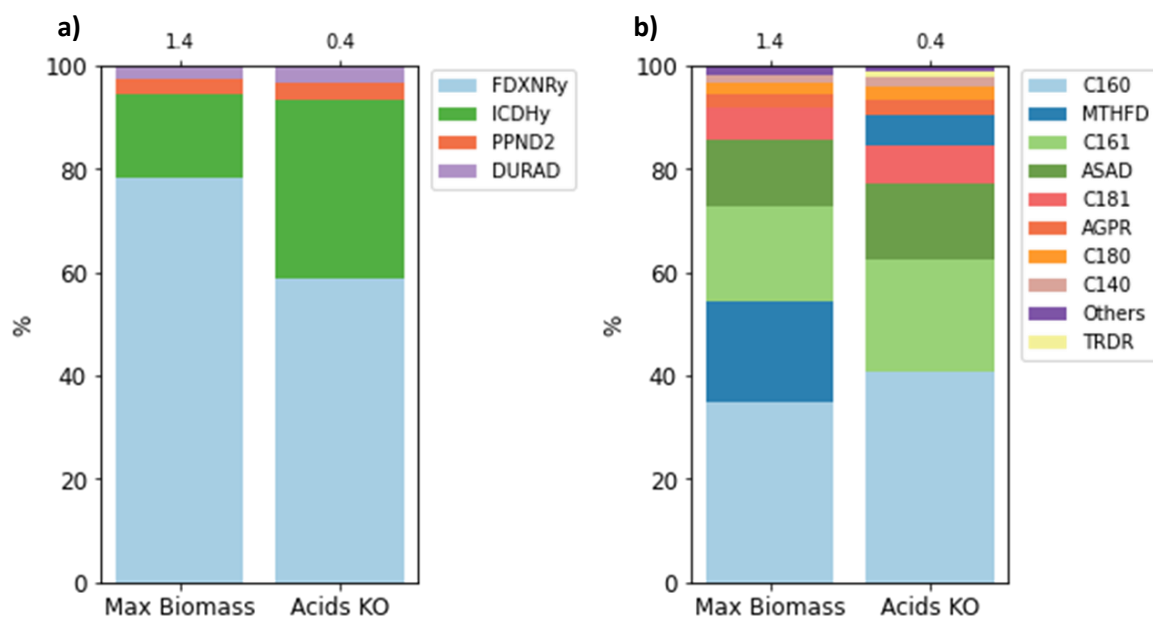


Figure 3.18 – Predicted fluxes responsible for ATP generation (a) and consumption (b) in the maximum growth simulation of *Clostridia* metabolism or with the phosphate acetyltransferase and butyrate kinase reactions removed and the butanol exchange flux constrained to the upper bound predicted by FVA (labelled ‘Acids KO’). Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Legend refers to the identifier of the reaction; FDXNRy – ferredoxin: NADP(H) oxidoreductase, ICDHy – isocitrate dehydrogenase, PPND2 – prephenate dehydrogenase, DURAD – dihydrothymine dehydrogenase, MTHFD – methylenetetrahydrofolate dehydrogenase, ASAD – aspartate semialdehyde oxidoreductase, AGPR – N-acetyl-glutamate semialdehyde oxidoreductase, TRDR – NADPH: oxidized-thioredoxin oxidoreductase, C140/C160/C161/C180/C181: fatty acid synthesis. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/consumption flux.

The hydrogenase and acid forming pathway knockout simulations both had similar features which were needed to achieve high butanol production. The biggest implication for increasing butanol production was that glycolysis alone was insufficient to generate enough NADH to drive high butanol production. In the hydrogenase knockout NADPH generation also increased, although this appeared to be interchangeable with increased NADH generation, as in the acid knockout simulation NADPH formation was lower than in the maximum growth simulation. As I found in section 3.3.1, the NADH produced from glycolysis alone can sufficiently be oxidised by lactate formation. This strongly suggests that rerouting carbon from acids to the more reduced alcohols requires an NADH driving force.

In addition, there was an increase in flux towards the four-carbon end of the ABE fermentation pathway. The maximum biomass simulation only predicted acetate formation as an ABE fermentation pathway product, while butanol formation was seen *in vivo* batch fermentations. I wanted to see what the predicted flux from acetyl-CoA to butyryl-CoA was when butanol formation was constrained to be that seen experimentally, which would roughly estimate the *in vivo* butyryl-CoA synthesis flux. With the constraints listed in **Table 3.1**, the estimated butyryl-CoA synthesis flux

was 4.6 mmol gDW⁻¹ h⁻¹. Compared to this experimentally constrained simulation, the butyryl-CoA synthesising flux increased by 79 % and 88 % in the hydrogenase and double acids knockout simulations, respectively. Given this difference in butyryl-CoA synthesising fluxes, it is likely that, in addition to the NADH driving force, increased expression of the genes responsible for butyryl-CoA synthesis will be necessary to drive increased butanol formation.

3.3.4 Strain-Design Algorithms to Improve Butanol Production

The knockout strategies I explored thus far took what is termed a ‘brute-force’ approach, which involved systematically searching every possible knockout. This is relatively simple for single knockouts, but as the number of knockouts increases, this quickly becomes computationally intensive. Strain design algorithms have been developed which are designed to overcome this problem. I decided to use the strain-design algorithms OptKnock (Burgard, Pharkya and Maranas, 2003) and OptForce (Ranganathan, Suthers and Maranas, 2010) to see what the possible knockout and overexpression targets were, and how they compared with the knockout strategies I had identified, and to see if there were additional non-intuitive manipulation targets.

I began with OptKnock, which aims to identify knockout targets which can lead to increased production of a target compound while maintaining substantial growth by solving a bilevel optimisation problem (Burgard, Pharkya and Maranas, 2003). I first ran OptKnock with a knockout size of two, which suggested inactivation of the acid forming pathways, identical to what I found using the brute-force approach (**Table 3.5**). I then increased the number of knockouts to three and excluded all of the reaction pairs identified in the knockout size of two simulations, to avoid simply getting the same results. None of the triple knockout targets were predicted to be viable.

OptKnock Set	Growth Rate after KO (h ⁻¹)	Minimum Butanol (mmol gDW ⁻¹ h ⁻¹)	Maximum Butanol (mmol gDW ⁻¹ h ⁻¹)
ACK, BCOPBT	0.11	0	8.75
ACK, BUTK	0.11	0	8.75
BCOPBT, PTA	0.11	0	8.75

Table 3.5 – Predicted growth rates and FVA predicted range of butanol producing fluxes from OptKnock determined knockouts in genome scale metabolic model of *C. saccharoperbutylacetonicum*. Reaction IDs correspond to the following reactions: ACK – acetate kinase, BCOPBT – phosphate butyryltransferase, BUTK – butyrate kinase, PTA – phosphate acetyltransferase.

I then used OptForce, which identifies overexpression and knockout targets to increase production of a target compound. A reference flux distribution is first generated representing the wild type strain, generally fulfilling an objective function such as maximisation of biomass. This is followed by a simulation optimised for production of the target compound. FVA is then performed to identify upper and lower flux bounds in each simulation. OptForce then identifies non-overlapping flux

ranges in the target and reference flux distributions, and these are placed into sets (termed must sets) based on whether the flux decreases in the target strain (requiring downregulation or knockout) or increases (requiring upregulation). Once must sets are identified, the OptForce selects a set of reactions from the must sets which maximise production of the target compound. I used OptForce to confirm the results of OptKnock and the systematic approach I had utilised, and to identify overexpression targets. I first ran OptForce with an intervention size of one (**Table 3.6**). This gave the intuitive targets of increasing butyraldehyde dehydrogenase flux and butanol export, which gave a predicted butanol forming flux of $9.02 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. The growth rate was reduced in these interventions to 0.01 h^{-1} . No other manipulations were predicted to give an appreciable butanol forming flux.

I then used an intervention size of two. This gave prevention of butyrate formation as a target (through knockout of phosphate butyryltransferase), and upregulation of butyryl-CoA forming enzymes (butyryl-CoA dehydrogenase or crotonase). These interventions had the same butanol forming flux and growth rate predicted in the butyraldehyde dehydrogenase upregulation.

I then increased the intervention size to three. Again, there was no increase in butanol forming flux in any of the interventions compared to an intervention size of two. The first OptForce set consisted of butyrate kinase knockout, upregulation of butyryl-CoA dehydrogenase and preventing hydrogen formation, effectively a hydrogenase knockout. A second set consisted of upregulating butyryl-CoA dehydrogenase and crotonase, combined with butyrate kinase knockout.

For the final round of OptForce I used an intervention size of four. These did not give any increase in butanol forming flux over the previous sets. Preventing lactate formation was an additional target in both OptForce sets, in the first set this was combined with prevention of butyrate formation, knockout of NAD-dependent butanol dehydrogenase and upregulation of butyryl-CoA dehydrogenase. Knocking out butanol dehydrogenase seems counter-intuitive, the likely reason for this knockout is forcing butanol formation to proceed by utilising NADPH as a reducing equivalent. The second OptForce set also suggested upregulation of butyryl-CoA dehydrogenase, while knocking out phosphate butyryltransferase and PFOR. The latter would force flux through PFL, which I previously identified as an overexpression target in the single deletion section.

Number of interventions	Force Set	Type of regulation	Minimum guaranteed for target (mmol gDW ⁻¹ h ⁻¹)	Maximum guaranteed for target (mmol gDW ⁻¹ h ⁻¹)	Maximum growth rate (h ⁻¹)
1	BUTALD	upregulation	9.02	9.02	0.01
	BUTOHex	upregulation	9.02	9.02	0.01
2	ACOAD1, BCOPBT	upregulation, knockout	9.02	9.02	0.01
	ECOAH1, BCOPBT	upregulation, knockout	9.01	9.01	0.01
3	ACOAD1, BUTK, H2ex	upregulation, knockout, knockout	9.02	9.02	0.01
	ACOAD1, ECOAH1, BUTK	upregulation, upregulation, knockout	9.02	9.02	0.01
	ACOAD1, BUTOHDx, BUTex, ex_lac_L_e	upregulation, knockout, knockout, knockout	9.02	9.02	0.01
4	ACOAD1, BCOPBT, POR4, ex_lac_L_e	upregulation, knockout, knockout, knockout	9.02	9.02	0.01

Table 3.6 – Predicted growth rates, minimum and maximum butanol forming fluxes in OptForce sets with intervention sizes ranging from one to four in genome scale metabolic model of *C. saccharoperbutylacetonicum*. Reaction IDs correspond to the following reactions: BUTALD – butyraldehyde dehydrogenase, BUTOHex – butanol secretion, ACOAD1 – butyryl-CoA dehydrogenase, BCOPBT – phosphate butyryltransferase, ECOAH1 – crotonase, BUTK – butyrate kinase, PTA – phosphate acetyltransferase, FDHy – NADP-dependent formate dehydrogenase, PPCDC - N-[(R)-4'-Phosphopantothienoyl]-L-cysteine carboxy-lyase, UAMAS - UDP-N-acetylmuramate:L-alanine ligase (ADP-forming), CYTD - cytidine aminohydrolase, LDH_D - (R)-Lactate:NAD⁺ oxidoreductase, MOHMT - 5,10-Methylenetetrahydrofolate:3-methyl-2-oxobutanoate.

3.3.5 Insertion of the Entner-Doudoroff Pathway

The results of the single and double knockout simulations suggested that additional NADH generation was important for increasing butanol formation. In the hydrogenase knockout, NADPH was used as a cofactor in the 3-hydroxybutyryl-CoA dehydrogenase reaction. I inspected the ABE fermentation pathway, I found that butanol formation could also use NADPH as a cofactor for certain reactions (3-hydroxybutyryl-CoA dehydrogenase and butanol dehydrogenase). Here, I wanted to identify methods of increasing NADPH in the model to drive further butanol production.

I first examined all NADPH forming reactions in the model for those which could potentially supply high NADPH, using the same FVA approach I used to identify NADH forming reactions. The highest NADPH forming fluxes were through ferredoxin: NADP(H) oxidoreductase and formate dehydrogenase (NADP-dependent). These reactions compete with ferredoxin: NAD(H) oxidoreductase and NAD-dependent formate dehydrogenase for ferredoxin and formate, respectively. Competing for common cofactors would decrease the effectiveness of NAD(P)H

generating systems, so I decided to investigate heterologous pathways which could be introduced to increase NADPH formation.

One method of generating NADPH is the use of the Entner-Doudoroff (ED) pathway. This is a glycolytic pathway that generates one mole of NADH, one mole of NADPH and one mole of ATP per mole of glucose metabolised, whereas the EMP pathway generates two moles of NADH and two moles of ATP per mole glucose metabolised (**Fig. 3.19**).

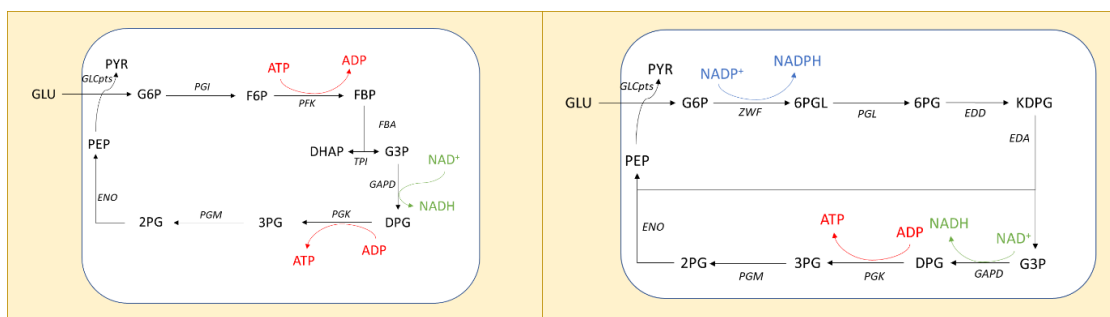


Figure 3.19 – Schematic of EMP (left) and ED (right) glycolytic pathways. Abbreviations refer to the following compounds: GLU – glucose, G6P – glucose-6-phosphate, F6P – fructose-6-phosphate, FBP – fructose-1,6-bisphosphate, DHAP – dihydroxyacetone phosphate, G3P – glyceraldehyde-3-phosphate, DPG – diphosphoglycerate, 3PG – 3-phosphoglycerate, 2PG – 2-phosphoglycerate, PEP – phosphoenolpyruvate, PYR – pyruvate, ATP – adenosine triphosphate, ADP – adenosine diphosphate, NAD⁺ – nicotinamide adenine dinucleotide, NADH – reduced nicotinamide adenine dinucleotide, NADP⁺ – nicotinamide adenine dinucleotide phosphate, NADPH – reduced nicotinamide adenine dinucleotide phosphate, 6PGL – 6-phosphogluconolactone, 6PG – 6-phosphogluconate, KDPG – 2-dehydro-3-deoxy-D-gluconate-6-phosphate. *Italics indicate respective model reaction.* Abbreviations refer to the following enzymes: *GLCpts* – PEP-dependent phosphotransferase system, *PGI* – phosphoglucose isomerase, *PFK* – phosphofructokinase, *FBA* – fructose bisphosphate aldolase, *TPI* – triose phosphate isomerase, *GAPD* – glyceraldehyde-3-phosphate dehydrogenase, *PGK* – phosphoglycerate kinase, *PGM* – phosphoglycerate mutase, *ENO* – enolase, *ZWF* – glucose-6-phosphate dehydrogenase, *PGL* – 6-phosphogluconolactonase, *EDD* – phosphogluconate dehydratase, *EDA* – KHG/KDPG aldolase.

I examined the model for ED pathway reactions, to see which enzymes would require introduction *in vivo*. All ED pathway reactions were present in the model except glucose-6-phosphate dehydrogenase, responsible for the first step in the pathway. This reaction was added to the model and a simulation was performed with maximisation of biomass and minimisation of sum of fluxes as the objective. This did not lead to changes in any fluxes from the previous maximal biomass simulation. This is likely because the ED pathway produces one less mole of ATP per mole of glucose metabolised compared with the EMP pathway. To test if it was possible to force flux through the ED pathway, I first removed the glucose-6-phosphate isomerase reaction from the model, which converts glucose-6-phosphate to fructose-6-phosphate, then ran a simulation optimised for biomass formation with minimisation of sum of fluxes as the objective. Predicted fluxes through central metabolism are shown in **Fig. 3.20**. This simulation had some flux through the ED pathway, but this was only 19 % of total glycolytic flux, with the remainder proceeding through EMP glycolysis. This also caused a reduction in predicted growth rate by 30 % compared with the maximum growth simulation. I investigated how the EMP flux was being maintained in the absence of the glucose-6-phosphate isomerase reaction. α -D-glucose ketol isomerase was found to be

To further force flux through the ED pathway, I removed the D-fructose 6-phosphotransferase

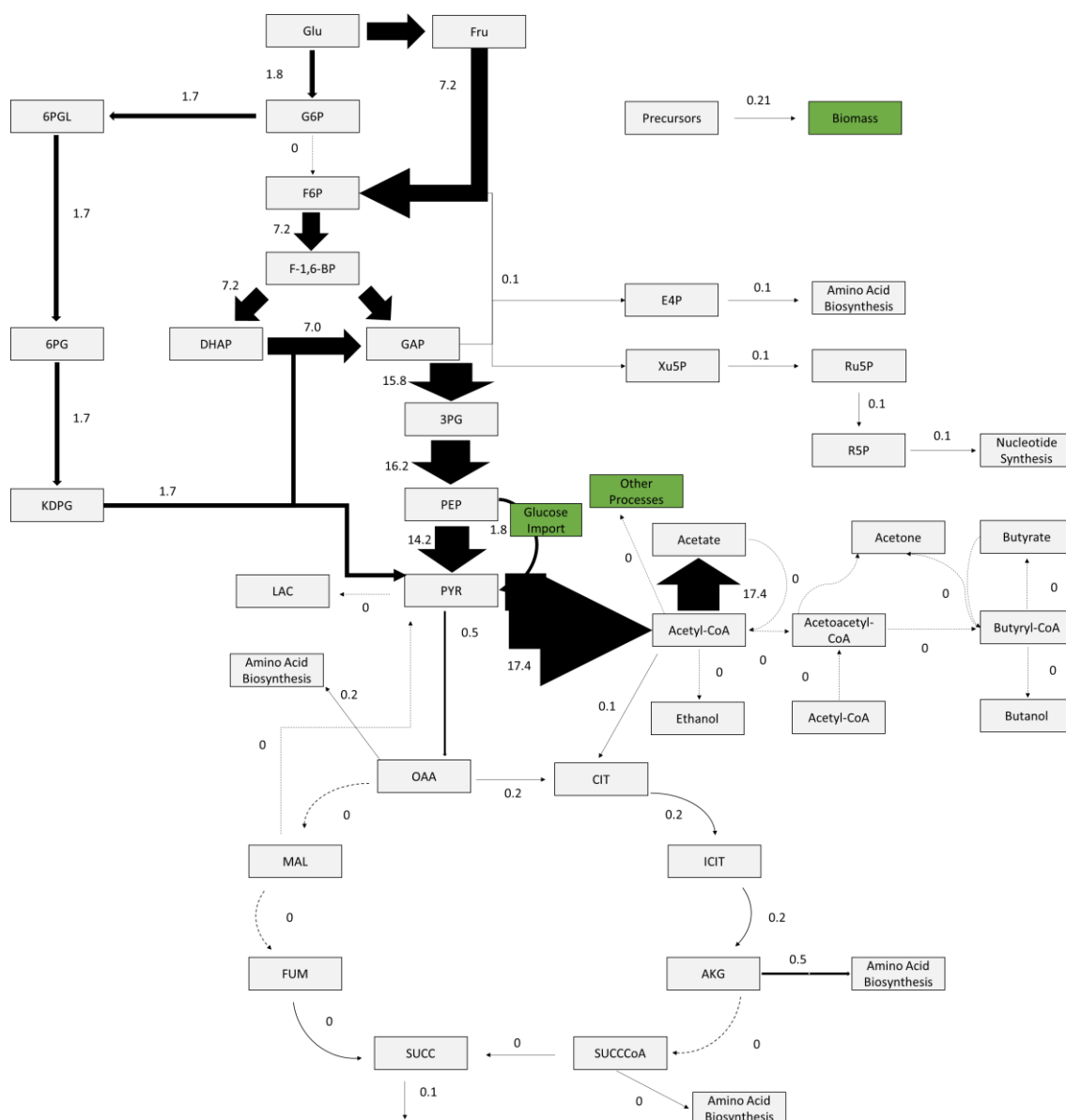


Figure 3.20 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with glucose-6-phosphate dehydrogenase added and phosphoglucose isomerase knocked out, optimised for biomass formation with experimentally determined rates of glucose uptake and NGAM estimate. Arrow width is proportional to predicted flux through the reaction, with the flux value shown next to the respective reaction. Dashed lines indicate a reaction with a predicted flux < 0.1. The metabolites and reactions outlined in this figure are shown in detail in **Appendix A**. All fluxes are given in units of mmol gDW⁻¹ h⁻¹, except for biomass formation which is given as h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

reaction from the model and ran a simulation optimised for biomass formation while minimising the total sum of fluxes. Glycolysis proceeded entirely through the ED pathway. In addition, butyrate became the major ABE fermentation product, followed by acetate (**Fig. 3.21**). The ED pathway

utilising simulation had a reduction in growth rate by 60 % compared to the maximum growth simulation.

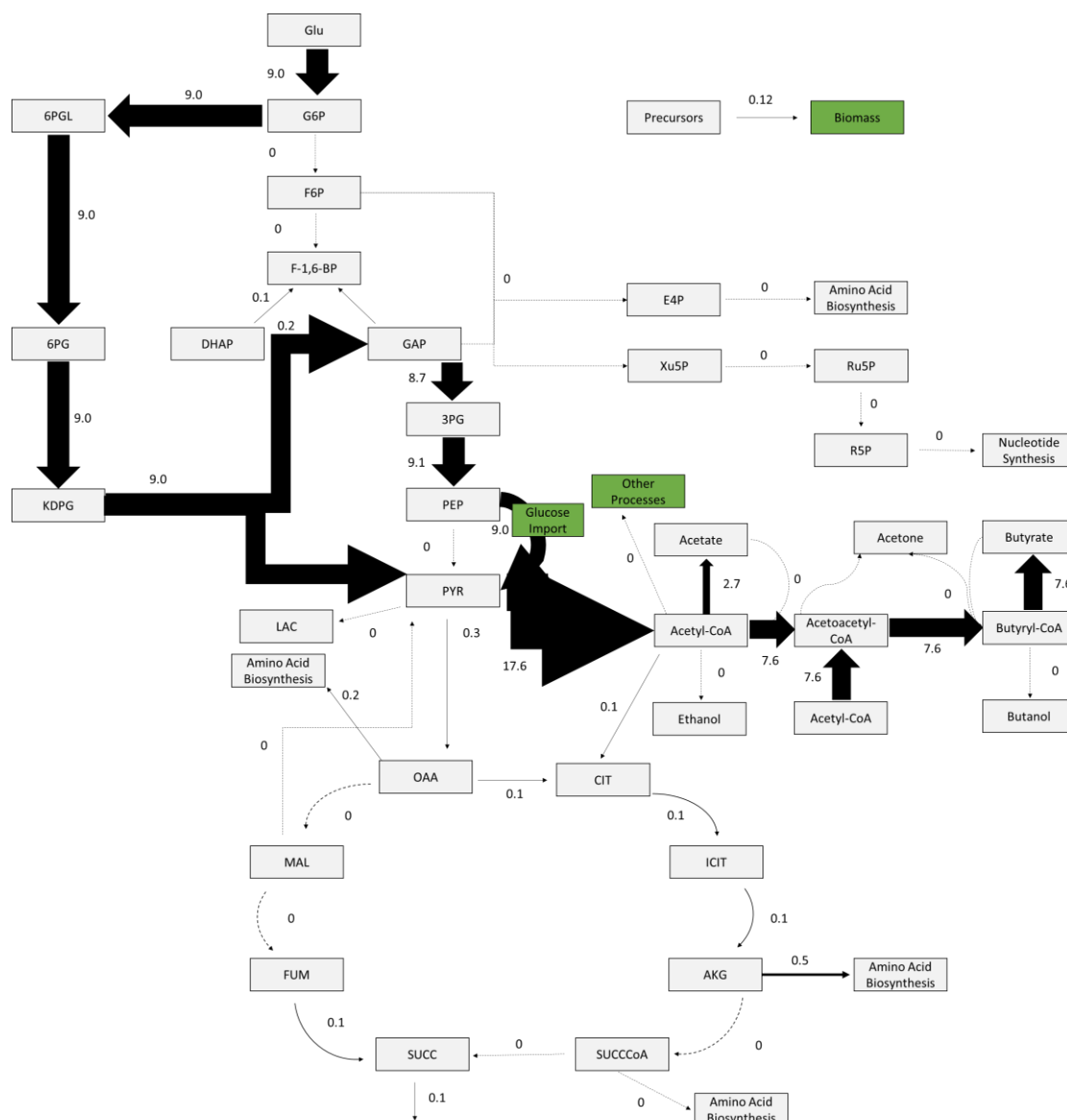


Figure 3.21 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with glucose-6-phosphate dehydrogenase added and α -D-glucose ketol isomerase and phosphoglucose isomerase knocked out, optimised for biomass formation with experimentally determined rates of glucose uptake and NGAM estimate. Arrow width is proportional to predicted flux through the reaction, with the flux value shown next to the respective reaction. Dashed lines indicate a reaction with a predicted flux < 0.1. The metabolites and reactions outlined in this figure are shown in detail in **Appendix A**. All fluxes are given in units of $\text{mmol gDW}^{-1} \text{h}^{-1}$, except for biomass formation which is given as h^{-1} . In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

In both simulations which used the ED pathway no butanol formation was predicted; however, carbon flux was diverted toward the four-carbon end of the ABE fermentation pathway. I compared the ATP, NADH and NADPH forming fluxes in the simulation which only used the ED pathway with

the maximum growth simulation. The total ATP forming flux was reduced in the ED pathway simulation by 55 % (**Fig. 3.22**). In the ED pathway simulation ATP was formed through phosphoglycerate kinase and butyrate kinase, each making up 46 % and 40 % of total ATP, respectively. Acetate kinase only contributed 14 % of total ATP forming flux. In terms of consumption, the NGAM cost was the largest drain of ATP, consuming 43 %. This was followed by biomass and protein synthesis, which consumed a combined 38 % of ATP.

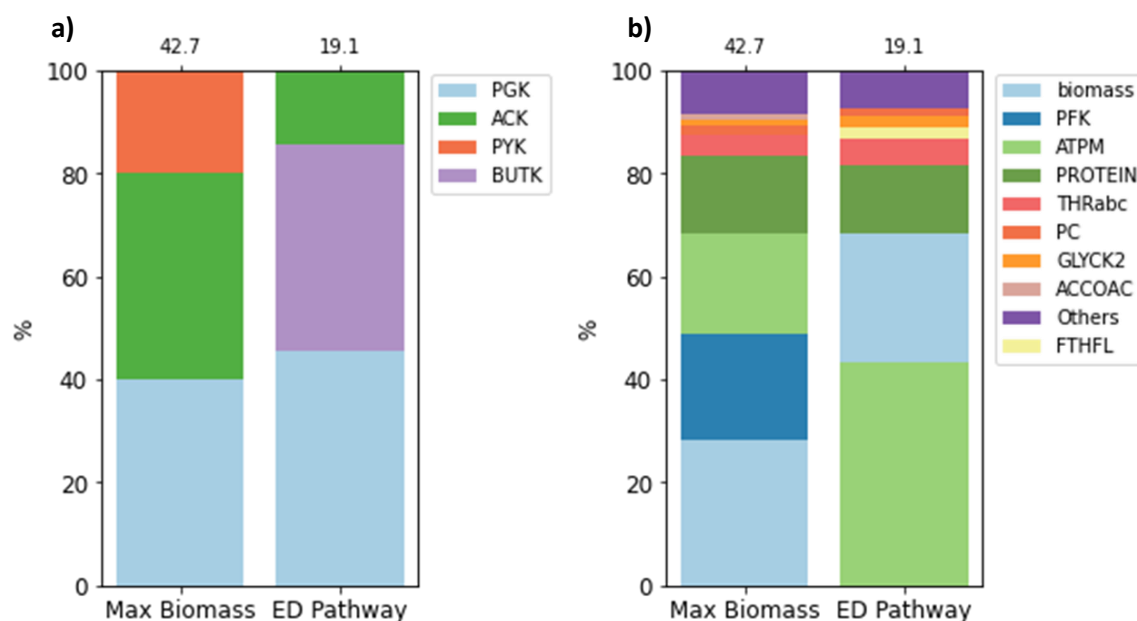


Figure 3.22 – Predicted fluxes responsible for ATP generation (a) and consumption (b) in the maximum growth simulation of *Clostridium* (labelled 'Max Biomass') or with glucose-6-phosphate dehydrogenase added to the model and α -D-glucose ketol-isomerase and phosphoglucose isomerase removed (labelled 'ED Pathway'). The value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; PGK – phosphoglycerate kinase, ACK – acetate kinase, PYK – pyruvate kinase, BUTK – butyrate kinase, biomass – lumped biomass synthesis reaction, PFK – phosphofructokinase, ATPM – NGAM, PROTEIN – lumped protein precursor for biomass synthesis reaction, THRabc – threonine-ABC transport, PC – pyruvate carboxylase, GLYCK2 – glycerate kinase, ACCOAC – Acetyl-CoA carbon dioxide ligase, FTHFL – formate: tetrahydrofolate ligase. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

The total NADPH forming flux was higher in the ED pathway simulation by over six-fold compared to the maximum biomass simulation (**Fig. 3.23**). This was almost entirely due to the glucose-6-phosphate dehydrogenase reaction, and the ferredoxin: NADP(H) oxidoreductase reaction was not active in the ED pathway simulation. The additional NADPH formed through the glucose-6-phosphate dehydrogenase reaction was oxidised through the 3-hydroxybutyryl-CoA dehydrogenase reaction.

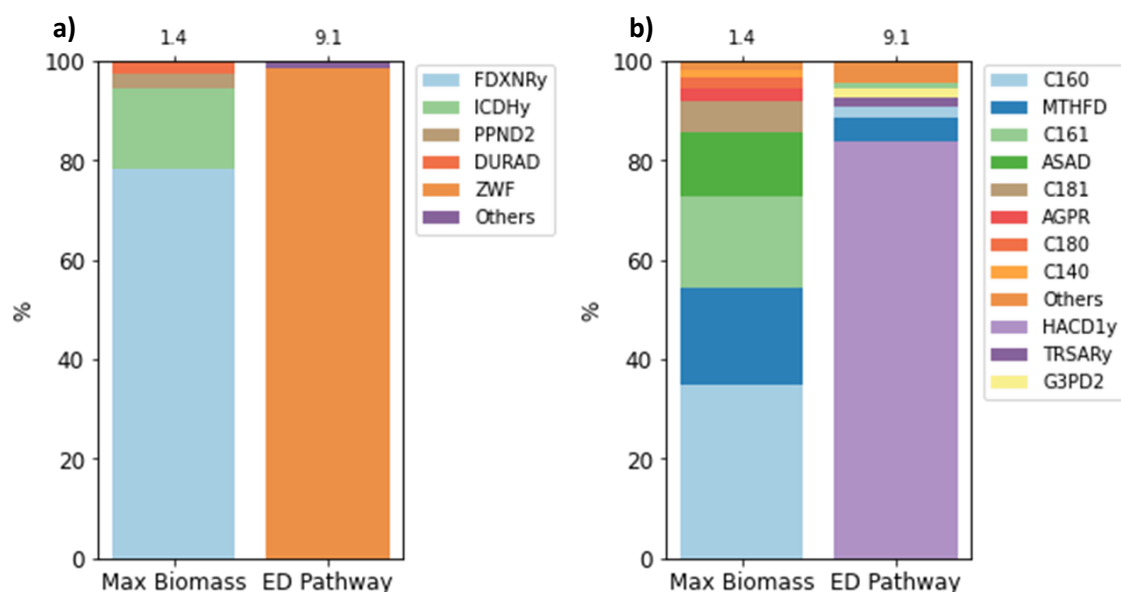


Figure 3.23 – Predicted fluxes responsible for NADPH generation (a) and consumption (b) in the maximum growth simulation of *Clostridium* (labelled ‘Max Biomass’) or with glucose-6-phosphate dehydrogenase added to the model and α -D-glucose ketol-isomerase and phosphoglucose isomerase removed (labelled ‘ED Pathway’). Value above bar is the total flux through all reactions shown in the bar graph, in $\text{mmol gDW}^{-1} \text{h}^{-1}$. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; FDXNRy – ferredoxin: NADP(H) oxidoreductase, ICDHy – isocitrate dehydrogenase, PPND2 – prephenate dehydrogenase, DURAD – dihydrothymine dehydrogenase, ZWF – glucose-6-phosphate dehydrogenase, MTHFD – methylenetetrahydrofolate dehydrogenase, ASAD – aspartate semialdehyde oxidoreductase, AGPR – N-acetyl-glutamate semialdehyde oxidoreductase, HACD1y – 3-hydroxybutyryl-CoA dehydrogenase, TRSARy – 2-hydroxy-3-oxopropionate reductase, G3PD2 – sn-Glycerol-3-phosphate:NADP+ 2-oxidoreductase, C140/C160/C161/C180/C181: fatty acid synthesis. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/consumption flux.

The NADH forming flux was lower in the ED pathway simulation than in the maximum biomass simulation (Fig. 3.24). The main reason for this being the ED pathway generates one fewer mole of NADH than EMP glycolysis, and NADH forming flux from glyceraldehyde-3-phosphate dehydrogenase was 50 % lower in the ED pathway simulation than the maximum biomass simulation. This was partially offset by the ferredoxin: NAD(H) oxidoreductase flux, which formed additional NADH in the ED pathway simulation. Almost all NADH was oxidised through the butyryl-CoA dehydrogenase reaction.

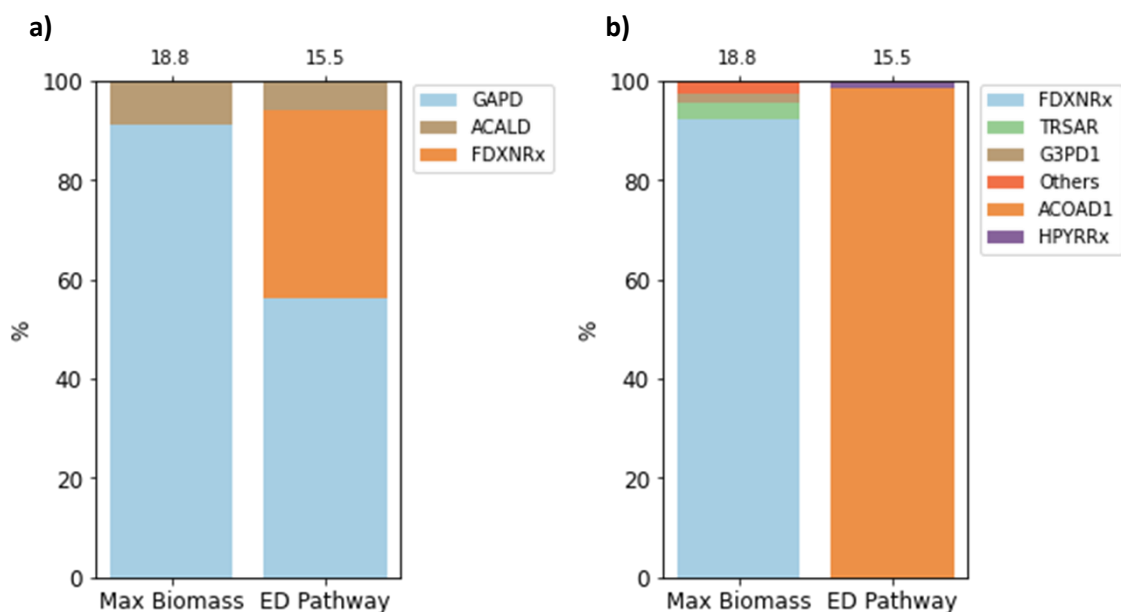


Figure 3.24 – Predicted fluxes responsible for NADH generation (a) and consumption (b) in the maximum growth simulation of *Clostridium* (labelled 'Max Biomass') or with glucose-6-phosphate dehydrogenase added to the model and α -D-glucose ketol-isomerase and phosphoglucose isomerase removed (labelled 'ED Pathway'). Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Legend refers to the identifier of the reaction; GAPD – glyceraldehyde-3-phosphate dehydrogenase, ACALD – acetaldehyde dehydrogenase, FDXNRx – ferredoxin: NAD(H) oxidoreductase, TRSAR - 2-hydroxy-3-oxopropionate reductase, G3PD1 - sn-Glycerol-3-phosphate: NAD⁺ 2-dehydrogenase, ACOAD1 – butyryl-CoA dehydrogenase, HPYRRx - glycerate dehydrogenase. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/consumption flux.

I ran FVA for the solely ED pathway-utilising simulation, i.e., inclusion of glucose-6-phosphate dehydrogenase and removal of phosphoglucose isomerase and α -D-glucose ketol isomerase, to see if the increased NADPH formed could result in butanol production. The upper predicted butanol formation flux was 3.8 mmol gDW⁻¹ h⁻¹, equivalent to 42 % of total glucose uptake. I generated a simulation with butanol production constrained to this bound then minimised the total sum of fluxes, with biomass formation as the objective (Fig. 3.25).

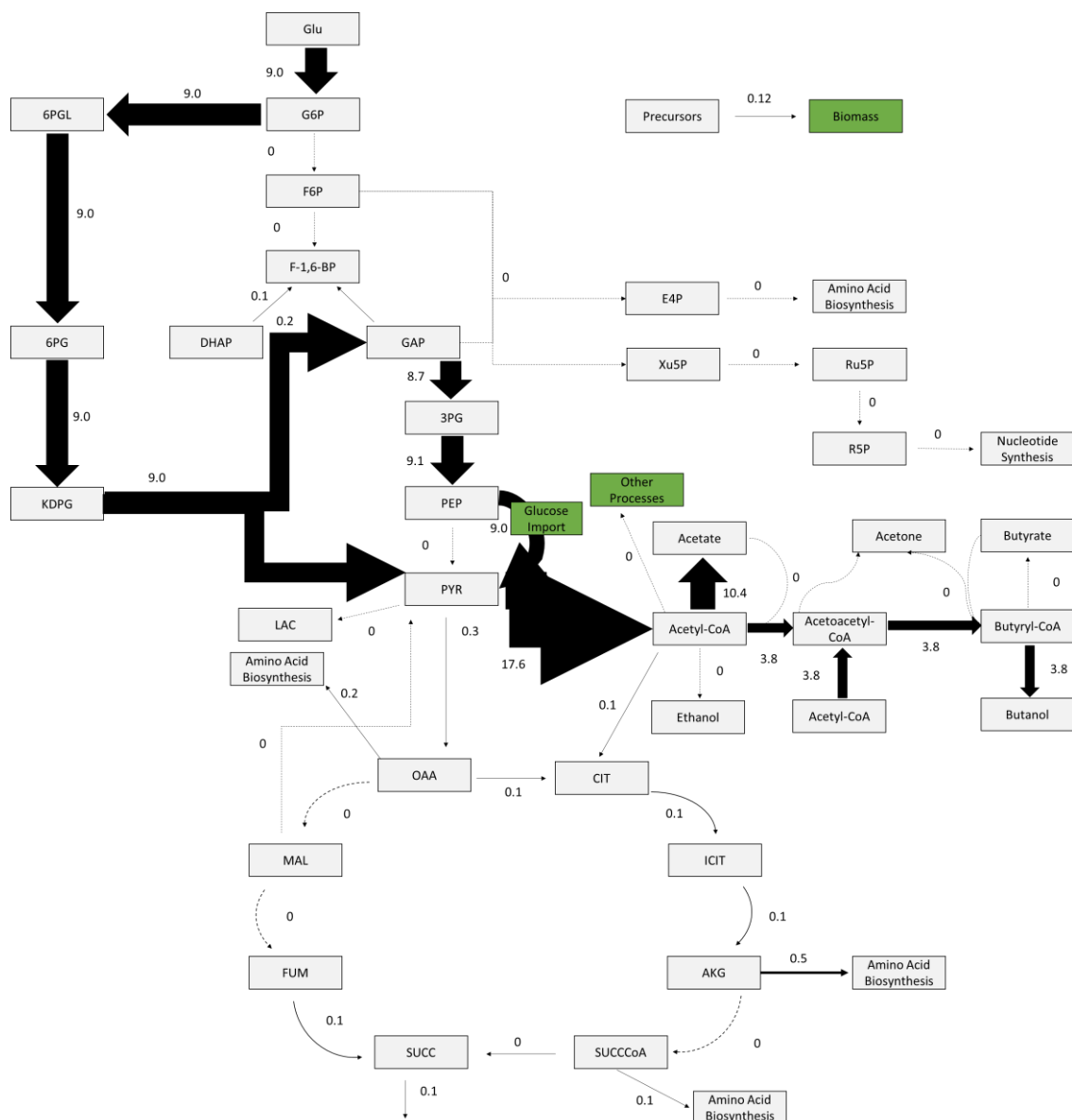


Figure 3.25 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with glucose-6-phosphate dehydrogenase added and α -D-glucose keto isomerase and phosphoglucose isomerase knocked out, optimised for biomass formation with experimentally determined rates of glucose uptake and NGAM estimate with butanol formation constrained to upper limit predicted by FVA. Arrow width is proportional to predicted flux through the reaction, with the flux value shown next to the respective reaction. Dashed lines indicate a reaction with a predicted flux < 0.1. The metabolites and reactions outlined in this figure are shown in detail in **Appendix A. All fluxes are given in units of mmol gDW⁻¹ h⁻¹, except for biomass formation which is given as h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.**

In the non-butanol forming ED pathway utilising simulation, 46 % of ATP was formed through phosphoglycerate kinase, followed by 40 % through butyrate kinase, with 14 % through acetate kinase (**Fig. 3.26a**). When butanol production was forced, 54 % of ATP was formed through acetate kinase, and 46 % through phosphoglycerate kinase. The NADPH formed through the ED pathway was oxidised mainly through 3-hydroxybutyryl-CoA dehydrogenase in the non-butanol forming simulation (**Fig. 3.26b**). In the simulation constrained to produce butanol there was half as much flux through this reaction, with the 3-hydroxybutyryl-CoA then diverted to butanol synthesis. This oxidised an equivalent amount of NADPH to the butyrate formation in the non-butanol forming simulation. The ATP consumption and NADPH forming fluxes were identical in both simulations and can be found in **Figs. 3.22b and 3.23a**.

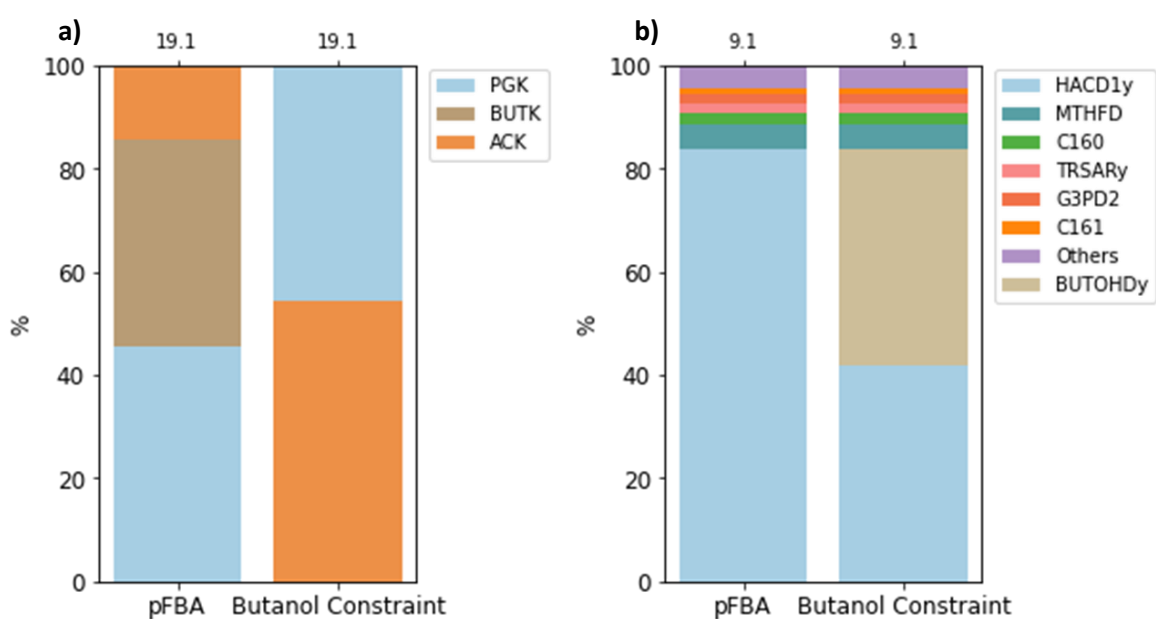


Figure 3.26 – Predicted fluxes responsible for ATP production (a) and NADPH consumption (b) in a simulation of *Clostridia* metabolism with glucose uptake and NGAM set to values estimated from batch fermentation experiments with glucose-6-phosphate dehydrogenase added to the model and α -D-glucose ketol-isomerase and phosphoglucose isomerase removed while maximising biomass formation while minimising the sum of fluxes as the objective (labelled ‘pFBA’) or constraining butanol formation to the upper bounds predicted by FVA. Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; PGK – phosphoglycerate kinase, BUTK – butyrate kinase, ACK – acetate kinase, HACD1y – 3-hydroxybutyryl-CoA dehydrogenase, MTHFD – methylenetetrahydrofolate dehydrogenase, TRSARy – 2-hydroxy-3-oxopropionate reductase, G3PD2 – sn-Glycerol-3-phosphate: NADP+ 2-oxidoreductase, BUTOHDy – butanol dehydrogenase, C140/C160/C161/C180/C181: fatty acid synthesis. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/NADPH consumption flux.

Based on the ED pathway simulations, increasing NADPH generation diverted flux towards the four-carbon end of the ABE fermentation pathway. Increasing NADPH generation may be a useful method of selecting for four carbon products in ABE fermentation.

3.3.6 4-Hydroxybutyrate Shunt

An undesirable feature of the ABE fermentation pathway is the generation of the side products acetate, butyrate, acetone, and ethanol. I wanted to implement an alternative butanol forming pathway which avoided the formation of these products. Removing the acid forming pathways as discussed forces formation of butanol, however there is the possibility that ethanol formation will increase, as the bifunctional aldehyde/alcohol dehydrogenase is responsible for catalysing both butyryl-CoA and acetyl-CoA conversion to the respective aldehyde and alcohol (R. V. Nair, Bennett and Papoutsakis, 1994). I wanted to see if there was an alternative fermentation pathway that could be designed which would produce butanol without formation of side products. Reviewing the literature, I found a potential pathway originally used to produce 1,4-butanediol in *E. coli* (Yim *et al.*, 2011) and adapted it to convert 4-hydroxybutyryl-CoA to crotonyl-CoA, which would be converted to butanol by the ABE fermentation pathway (**Fig. 3.27**). I identified pre-existing reactions of this pathway in the model and added the missing reactions, which were α -ketoglutarate decarboxylase and 4-hydroxybutyrate dehydrogenase.

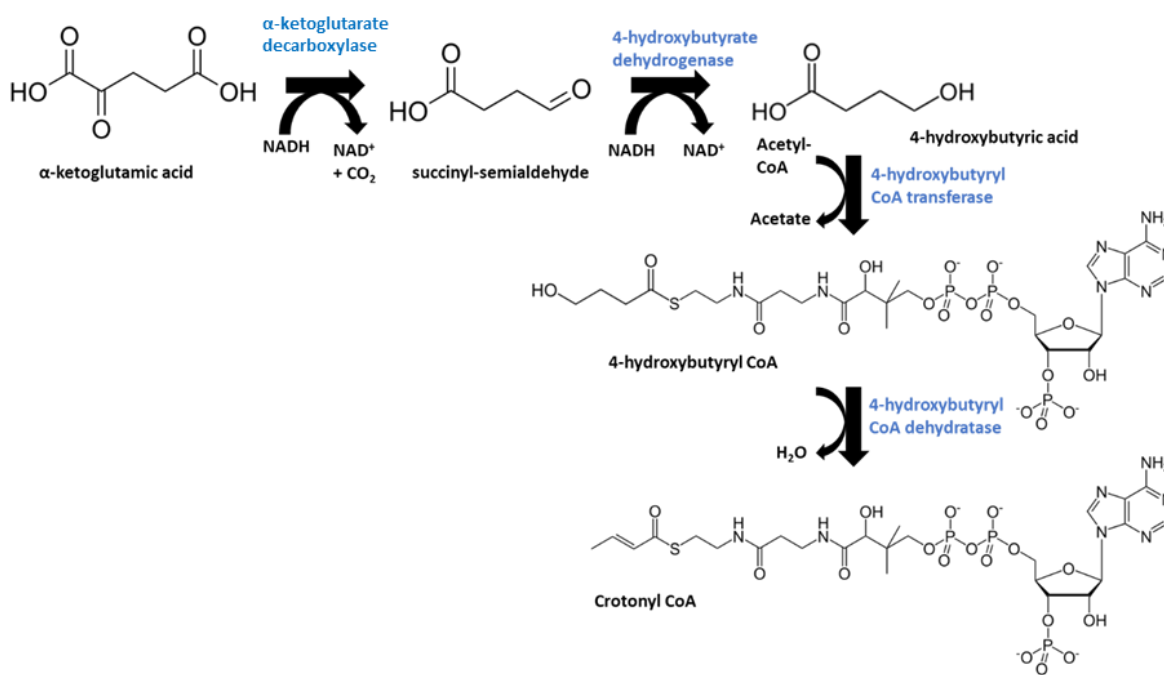


Figure 3.27 – 4-hydroxybutyrate shunt pathway. Enzyme names shown in blue, cofactors and molecule names in black. Crotonyl-CoA is used in the canonical ABE fermentation pathway.

The model was first optimised for biomass production using pFBA, which did not lead to any changes in the predicted fluxes relative to that shown in **Fig. 3.4**. I then optimised the model with butanol production as the objective, to see if the new pathway was preferred for butanol production, but the native ABE fermentation pathway was used. I then removed thiolase from the model to see if flux could be forced through the 4-hydroxybutyrate shunt. This did lead to utilisation

of the 4-hydroxybutyrate shunt (**Fig. 3.28**), but the butanol forming flux was 34 % lower when compared to a simulation optimised for butanol production using the native ABE fermentation pathway. Rather than use the introduced 4-hydroxybutyryl-CoA dehydrogenase reaction, the simulation predicted 4-hydroxybutyryl-CoA would be converted to 3-hydroxybutyryl-CoA, then subsequently converted to crotonyl-CoA. As acetyl-CoA formation was somewhat bypassed in the production of butanol here, there was less reduced ferredoxin forming flux. This meant that ferredoxin:NAD(H) oxidoreductase flux was reduced, meaning that NADH formation was lower relative to the butanol producing hydrogenase and double acid knockouts.

3.4 Discussion

3.4.1 Role of Organic Acid Production in *Clostridium* Metabolism

The results of the simulations shown in this chapter have shown that ATP availability is limiting growth in solventogenic *Clostridium*. Changing the objective from biomass formation to maximising ATP formation led to a small increase in acetate forming flux, as the maximum biomass simulation was already producing a high level of acetate. This was not consistent with experimental observations, which showed that butanol was the major fermentation product. High acetate production was also observed in unconstrained simulations of *C. cellulolyticum* (Salimi, Zhuang and Mahadevan, 2010) and *C. beijerinckii* (Milne *et al.*, 2011) optimised for maximal growth. In contrast, one of the *C. acetobutylicum* models produced more butyrate than acetate, with acetate only being produced when high ATP maintenance costs were imposed (Lee *et al.*, 2008). In these models, as in my simulations, acetate likely played a key role in ATP generation in cellular metabolism. When I removed acetate formation from the model, there was a reduction in the predicted growth rate and butyrate was the only predicted ABE fermentation product. There have been numerous reports of acid forming pathway knockouts in solventogenic *Clostridium*, with varying impacts on growth. Reports of phosphate acetyltransferase/acetate kinase knockout in the *C. acetobutylicum* WUR and ATCC 824 strains gave varied results, both with each other and compared to my simulation. The acetate kinase disruption in the WUR strain decreased acetate kinase activity to just under 2 % of the wild type and phosphate acetyltransferase activity was also reduced. Despite this, the end point acetate concentration was unchanged in the mutant and butyrate decreased (Kuit *et al.*, 2012). As the acetate kinase and phosphate acetyltransferase genes lie on an operon, it may be that altered expression led to the reported reduction in phosphate acetyltransferase activity. A similar phenotype has been reported in strain ATCC 824 with reduced acetate and butyrate titres (Cooksley *et al.*, 2012). The phenotype of these mutants contrasts with a phosphate acetyltransferase knockout which showed reduced acetate and an increase in butyrate (Green *et al.*, 1996), which is in closer agreement with the model predictions. A phosphate acetyltransferase knockout strain of *C. beijerinckii* NCIMB 8052 also showed reduced acetate and increased butyrate production, but not complete elimination of acetate (Wang *et al.*, 2013c). A phosphate acetyltransferase knockout of *C. saccharoperbutylacetonicum* N1-4 has been reported, and showed general characteristics in agreement with the simulation, with a decrease in acetate and increased butyrate (Wang *et al.*, 2017). My simulations are in agreement with a *C. thermocellum* model, which showed an increase in predicted ethanol forming flux when the phosphate acetyltransferase or acetate kinase reactions were blocked, which also reduced the predicted growth rate (Roberts *et al.*, 2010).

In my simulations, removing the butyrate forming pathway in addition to the acetate forming pathway led to lactate being the only predicted fermentation product when optimised with pFBA, which coincided with a reduction in the predicted growth rate. This is a fairly drastic shift in product pattern compared with the experimental data. Indeed, lactate production has not been reported to increase to such a dramatic extent in most published mutants, but the reduction in growth rate has been reported. There was almost no change in lactate levels and an increase in butanol titre and yield in a *C. saccharoperbutylacetonicum* N1-4 *ptabuk* double knockout mutant (Wang *et al.*, 2017). High lactate production has been reported in a *ptb* knockout strain of *C. acetobutylicum* when cultured without pH control, but an increase in butanol and ethanol titres was reported when the strain was cultured with pH control (Lehmann, Radomski and Lütke-Eversloh, 2012b). This shows that minimising the sum of fluxes in the model of *C. saccharoperbutylacetonicum* might give unrealistic predictions of behaviour. The use of FVA to compare flux ranges when modelling manipulations will be important when predicting responses to network perturbations.

My systematic knockout search, and the OptKnock and OptForce algorithms all identified the acid pathways as knockout targets. Acid knockouts have long been targeted as an intuitive approach to increase butanol production in solventogenic *Clostridium*, as they present a sink of carbon which reduces the potential butanol forming flux. OptForce predicted that knockouts of the butyrate forming arm of the ABE fermentation pathway could increase butanol production. A *ptb* knockout strain of *C. acetobutylicum* has been reported, with increased titres of ethanol and butanol, while the growth rate remained similar (Lehmann, Radomski and Lütke-Eversloh, 2012b). The already mentioned *ptabuk* double knockout mutant of *C. saccharoperbutylacetonicum* developed by Wang *et al.*, had a growth rate which was decreased by 31 % compared with the wild type, a less severe reduction than the 53 % decrease in the double acid knockout simulation here compared to the maximum biomass simulation. Also, of note, the peak butyrate increased in the *ptabuk* double knockout mutant, and the acetate levels remained similar. Multiple homologues of acetate kinase, phosphate butyryltransferase and butyrate kinase are present in the *C. saccharoperbutylacetonicum* N1-4 genome; with two, two and three isoforms, respectively, indicating that either the genes knocked out in the study were not the main contributors to the respective acid formation, or that the alternative isoforms which exist are induced when the main contributing isoforms are absent. The maintained acid production would still allow ATP generation, which explains how the growth penalty was less severe in the double knockout developed experimentally. Maintained acid production in *ptabuk* knockout strains has also been reported in a mutant strain of *C. acetobutylicum* ATCC 824 lacking both of these genes (BEKW strain), which led to a 56 % reduction in acetone levels and a 35.6 % increase in butanol levels compared to the wild

type (Jang, Lee and Lee, 2012). Interestingly, a single *pta* knockout mutant strain developed in the same paper (EKW strain) showed a 15 % reduction in growth rate compared to the parental strain, while the BEKW strain showed a 10 % increase in growth rate. This is in stark contrast with the model predictions presented here and would suggest alternative isoforms of *pta* and/or *buk* as methods of ATP generation. Importantly, this highlights the need to be careful comparing computational simulations, where flux can be prevented running through a reaction absolutely, and *in vivo* knockouts, where alternative enzyme homologues and/or spontaneous reactions can allow flux through a reaction.

3.4.2 NAD(P)H Formation Drives Butanol Production

My exploration of the genome scale metabolic model adapted for use with *C. saccharoperbutylacetonicum* N1-4 has highlighted patterns in simulations which had high butanol forming flux. The biggest takeaway is the need for increasing NAD(P)H generation to drive synthesis of butanol, in addition to that formed through glycolysis. In both the hydrogenase knockout and phosphate acetyltransferase/butyrate kinase double knockout simulations, additional NAD(P)H was formed through oxidation of ferredoxin when constrained to produce butanol. Under a maximum growth simulation, all NADH formed through glycolysis was oxidised through ferredoxin: NAD(H) oxidoreductase, followed by ferredoxin oxidation through hydrogenase. In batch fermentations, higher hydrogen formation has been reported by investigators during acidogenesis, while hydrogen formation decreases later during solventogenesis (Jones and Woods, 1986). The role of hydrogenase in NADH oxidation has made it a target for removal by many groups, but this has rarely been successful (Cooksley *et al.*, 2012; Nguyen *et al.*, 2018). There are two likely reasons for the difficulty in generating such strains. First, formation of acetyl-CoA occurs through pyruvate ferredoxin oxidoreductase in *C. acetobutylicum* (Au *et al.*, 2014), and the removal of hydrogenase prevents the key ferredoxin oxidation route from functioning. In addition, during acidogenesis, NADH is oxidised, at least in part, through the ferredoxin: NAD(H) oxidoreductase reaction (Petitdemange *et al.*, 1976), which then uses the hydrogenase to disperse the accumulated reducing power. Removing the hydrogenase therefore prevents both ferredoxin and NADH oxidation. In my simulations, forcing butanol production in the hydrogenase knockout simulation resulted in the ferredoxin: NAD(H) oxidoreductase reaction running in the reverse, i.e., NAD⁺ reducing direction. In addition, increased flux through butyryl-CoA and butanol synthesising reactions was required. This is a substantial redistribution of fluxes, and it is unlikely that hydrogenase negative individuals when transformed can sufficiently balance NAD(H) and ferredoxin.

To further demonstrate the need for increased NADH generation, the double acid knockout

without the butanol constraint led to lactate being the only fermentation product. This is a risk given the multiple possible routes of NADH oxidation in *C. saccharoperbutylacetonicum* N1-4, being lactate, ethanol, butanol, or ferredoxin mediated hydrogen formation. In the butanol producing double acid knockout simulation, while not explicitly knocked out, there was no flux through the hydrogenase reaction, with almost all ferredoxin being oxidised through the ferredoxin: NAD(H) oxidoreductase reaction. This generated the additional NADH required for butanol formation, which would not be effectively oxidised by lactate formation. As glucose to lactate conversion generates and consumes two moles of NADH (assuming one mole of glucose is completely converted to lactate), any additional NADH formed would require alternative oxidation, such as formation of ethanol or butanol. This again shows that additional NADH must be formed (aside from that produced during glycolysis) to achieve high butanol production.

3.4.3 Consolidated Metabolic Engineering Strategy

Based on all the simulations described in this chapter, I have formulated a feasible metabolic engineering strategy for increasing butanol yield in *C. saccharoperbutylacetonicum* N1-4 that will subsequently be experimentally investigated. My double deletion search and OptKnock simulations concurred that preventing acid formation could increase butanol production. OptForce, along with my investigation of butyryl-CoA forming flux estimated from batch fermentations with the high butanol producing simulations, have shown that this alone will not be sufficient to increase butanol production. Removing acid formation must be coupled with high flux through butyryl-CoA and butanol synthesising enzymes. To ensure this occurs, it would be beneficial to overexpress these enzymes. Increasing NADH formation was achieved in the model through altering the direction of the ferredoxin: NAD(H) oxidoreductase, although this would likely be difficult to achieve *in vivo* as NADH oxidation appears to be the main role of this enzyme, at least during acidogenic growth (Petitdemange *et al.*, 1976; Jones and Woods, 1986). As an alternative, expressing pyruvate formate lyase and formate dehydrogenase could be utilised. My FVA simulations have shown that these enzymes could replace the pyruvate: ferredoxin oxidoreductase and ferredoxin: NAD(H) oxidoreductase system as an acetyl-CoA and NADH forming system. Expression of formate dehydrogenase alone has been used in other organisms for similar purposes i.e., ethanol formation (Berríos-Rivera, Bennett and San, 2002a; Berríos-Rivera *et al.*, 2004), and shows promise as an NADH forming system.

Aside from NADH formation, there was an increase in butyryl-CoA synthesising flux in the hydrogenase and acid forming knockout simulations, which was over 50 % higher than that estimated from the batch fermentation data. This suggests that overexpressing these genes might be required to maintain carbon flux towards butanol. Overexpressing the *bcs*-operon genes

individually would be a lengthy and difficult process, due to the size of the operon (around 5 kbp). An alternative method of increasing the levels of these enzymes is the removal of the Rex redox response regulator. This repressor protein is located upstream of the *bcs*-operon and in *C. acetobutylicum* and *C. beijerinckii* and binds to promoter regions of crotonase, thiolase and aldehyde-alcohol dehydrogenase genes (Wietzke and Bahl, 2012; Zhang *et al.*, 2014). A Rex homologue is present in the same position (upstream of the *bcs*-operon) in the *C. saccharoperbutylacetonicum* N1-4 genome and may have a similar role in repressing *bcs*-operon transcription. Deletion of the *rex* gene was thereby selected as another part of my manipulation strategy.

The insertion of two heterologous pathways were also simulated here. The 4-hydroxybutyrate shunt was not deemed to be worth implementing, as it required a substantial change in fluxes through the TCA cycle and did not offer an increase in butanol producing flux over the native pathway. The other insertion modelled was the ED pathway. This has been used previously to increase NADPH generation for production of neurosporene (Ng *et al.*, 2015). Here, the ED pathway was trialled as a means of NADPH formation, with the aim of increasing flux toward four carbon products in the ABE fermentation pathway. This was shown to be effective at increasing butyrate formation, but the reduction in ATP formed through this pathway rather than the EMP pathway led to a reduction in growth rate. While this was not selected as a strategy at this point, increasing NADPH formation may still prove useful in future metabolic engineering work.

Removing the hydrogenase was not considered at this stage based on difficulties generating this knockout reported by other groups (Cooksley *et al.*, 2012; Nguyen *et al.*, 2018). In addition, acid formation was relatively low in the preliminary batch fermentation, so removal of the acid forming pathways was not selected. I decided to first implement the NADH generating and *rex* knockout strains, then remove the acid forming pathways if necessary, based on the phenotype of the mutants (see chapter six).

Chapter 4: Analysis of Metabolism of an Asporogenic Mutant Strain of *C. saccharoperbutylacetonicum* N1-4 using a Genome Scale Metabolic Model

4.1 Aims of Work Described in Chapter

The work presented in this chapter aimed to understand the metabolic phenotype of a mutant, asporogenic strain of *C. saccharoperbutylacetonicum* N1-4 developed by Biocleave Ltd. The goal was to compare its performance with wild type *C. saccharoperbutylacetonicum* N1-4 (chapter 3) and to identify metabolic factors limiting butanol production. I used batch fermentations and FBA to achieve this goal.

4.2 Introduction

4.2.1 Asporogenic Strains of Solvent Producing *Clostridium*

One of the major downsides of ABE fermentation is the link between solvent formation and endospore formation (Ravagnani *et al.*, 2000). Asporogenic strains tend to show reduced ability to produce solvents (Cornillot *et al.*, 1997; Ravagnani *et al.*, 2000; Harris, Welker and Papoutsakis, 2002; Seo *et al.*, 2017; Atmadjaja *et al.*, 2019). The industrial partner for my D.Phil. project, Biocleave Ltd, have developed a proprietary mutant strain of *C. saccharoperbutylacetonicum* N1-4 which has a point mutation in the *spo0A* gene, preventing endospore formation but maintaining solvent production (Atmadjaja *et al.*, 2019). In batch fermentations, this strain (*C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}) produced a higher total solvent and acid titre than the wild type (Atmadjaja *et al.*, 2019). While the peak levels of all solvents and acids were reported, data was not provided for the levels of individual ABE fermentation products, nor were the growth characteristics quantified. To gain a better understanding of the metabolic behaviour of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, the strain was subject to more detailed analysis in batch fermentations here.

4.3 Results

4.3.1 The *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} Mutant Produced More Butanol than the Wild Type in Batch Fermentations

C. saccharoperbutylacetonicum N1-4 *spo0A*^{I261T} was subject to batch fermentations in glucose containing medium, and I compared the fermentation profile with the fermentation of wild type *C. saccharoperbutylacetonicum* N1-4 performed in chapter three (**Fig. 4.1**). As with the wild type, *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} showed a biphasic fermentation pattern, with acid production during the exponential growth phase (**Fig. 4.1f, g**), followed by appearance of solvents from 24 hours onwards (**Fig. 4.1c, d, e**). As with the wild type growth curve, logistical issues prevented me from collecting the data for a complete growth curve. Again, the peak time point at

OD_{600nm} may not have been the actual highest value of OD_{600nm}. There were notable differences between the two strains which became apparent after the shift to solvent production.

The growth of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was much slower than the wild type (**Fig. 4.1a**), with a slower pattern of product formation and slower glucose consumption (**Fig. 4.1b**). Glucose was not completely depleted from the media, with *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} consuming around the same amount of glucose as the wild type (50 g L⁻¹) but this took twice as long. There was also a slower increase in acetone, butanol, and ethanol concentration, which were produced continuously throughout the fermentation after the shift to solventogenesis (**Fig. 4.1c-e**). The concentration of lactate peaked at the same time as the wild type (**Fig. 4.1g**) but lactate was completely reassimilated by the end of the fermentation.

The peak levels of fermentation products are shown in **Fig. 4.1h**. There was a statistically significant difference in the peak levels of all metabolites between the two fermentations. The wild type had higher peak titres of acetone, butyrate, and lactate, and reached a higher peak OD_{600nm}, while *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} had higher peak titres of butanol and ethanol. These fermentations show that *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} grew more slowly

than the wild type but had a sustained pattern of solvent formation and, most importantly, produced more butanol than the wild type.

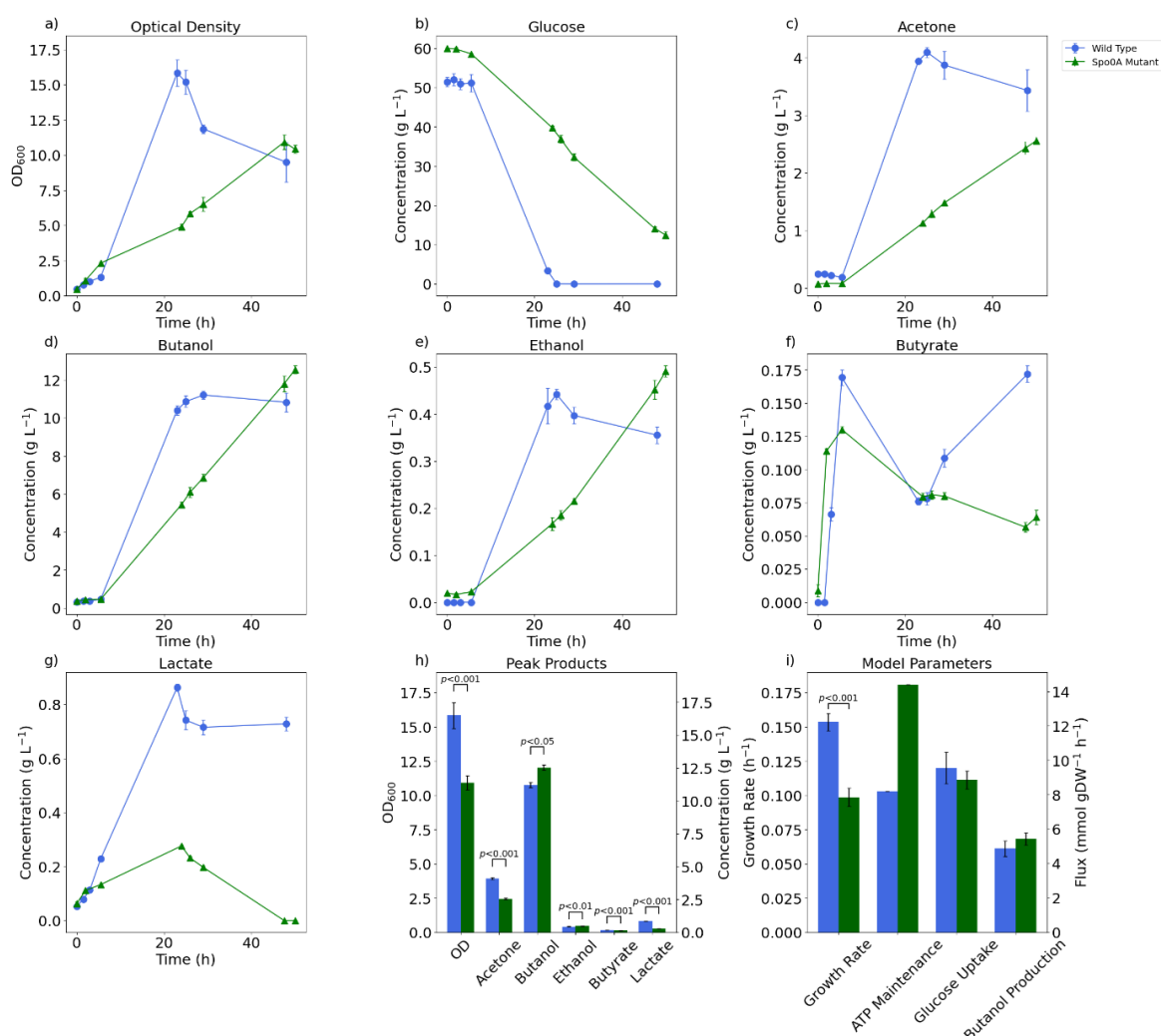


Figure 4.1 – Batch fermentation data and derived model parameters for batch fermentations of *C. saccharoperbutylacetonicum* N1-4 and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. a) Optical density as measured by absorbance at 600 nm. Concentrations of b) glucose, c) acetone, d) butanol, e) ethanol, f) butyrate, and g) lactate, in the fermentation media were measured using HPLC. Peak fermentation products are shown in h). The model parameters estimated from both sets of fermentation data are shown in i). Batch fermentations were performed in 100 mL serum bottles with 60 mL of media. Wild type fermentations were conducted in CGM with 50 g L⁻¹ glucose and 1 % w/v CaCO₃ as a pH buffer. The *spo0A* mutant strain fermentations were conducted in TYIR with 60 g L⁻¹ glucose with 200 mM MES as a pH buffer. All fermentations were performed in triplicate. Error bars shown are ± 1 standard deviation. Where a statistically significant difference between the two strains was found ($p < 0.05$) the p -value is shown above the respective parameter.

4.3.2 Comparison of Predicted Flux Distributions of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} and *C. saccharoperbutylacetonicum* N1-4

Given that *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} produced more butanol than the wild type, I also wanted to see how the pFBA predicted intracellular flux distribution differed in this strain compared with the wild type. This was particularly important with regards to the metabolic engineering strategy I proposed in chapter three, which might not be applicable to *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}.

As in chapter three, the growth rate was estimated from the log-transformed OD_{600nm} data using linear regression and the rates of glucose uptake along with butyrate, lactate, acetone, butanol, and ethanol secretion were similarly calculated (**Table 4.1**). Using the glucose uptake and product secretion rates as constraints on the model, an NGAM value was estimated, which was 14.4 mmol gDW⁻¹ h⁻¹. This was higher than that estimated for the wild type (8.3 mmol gDW⁻¹ h⁻¹). This value is not unusual however compared to other microbial fermentations. In *E. coli* simulations values of 8.4 are commonly used for aerobic growth (Feist *et al.*, 2007), but other estimates of NGAM range from 15 – 30 mmol gDW⁻¹ h⁻¹ under various stress conditions (Meadows *et al.*, 2010) and 16.4 mmol gDW⁻¹ h⁻¹ under anaerobic conditions (Chen *et al.*, 2011).

Metabolic Process		Reaction ID	Mean Estimated Flux	Lower Bound	Upper Bound
Uptake of glucose		ex_glc_D_e	-8.85127	-8.33014	-9.3724
Secretion of butyrate		BUTex	0.064314	0.068247	0.060381
Secretion of acetone		ACETONEex	1.43582	1.540194	1.331447
Secretion of butanol		BUTOHex	5.423517	5.785136	5.061898
Secretion of ethanol		ETOHex	0.253221	0.284675	0.221767
Secretion of lactate		LAC2ex	0.18975	0.197065	0.182436
Rate of biomass synthesis		biomass	0.098677	0.105442	0.091911

Table 4.1 – reaction IDs and constrained upper and lower bounds from experimental estimates of reaction fluxes for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. Upper and lower bounds of each reaction are the experimental average + or – one standard deviation, respectively. Glucose uptake was kept constant at the mean estimated flux in the simulations, unless otherwise stated. Units are given as mmol gDW⁻¹ h⁻¹, except for growth rate (biomass synthesis), which is h⁻¹.

I first ran a simulation, with growth and the rates of secretion of butyrate, lactate, acetone, butanol, and ethanol constrained to the experimental measurements for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} and an objective function of minimisation of total sum of fluxes. The upper and lower bounds listed in **Table 4.1** were imposed as the constraints. The predicted fluxes through central metabolic reactions are shown in **Fig. 4.2**. For comparison, the percent change in the respective flux in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} simulation relative to the pFBA

simulation constrained with wild type fermentation data (chapter three) is given next to the absolute flux value. As with the wild type, glucose was solely imported through the PTS, and flux through glycolysis was similar. The PPP flux was 28 % lower in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} compared with the wild type. This was a result of the lower growth rate in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, meaning less carbon flux was used in the synthesis of biomass precursors and amino acids. The glycolytic flux was similar despite the lower growth rate due to the imposed maintenance cost, which was higher than the wild type. Despite the lower ATP demand from biomass, the higher maintenance cost balanced the total ATP demand to be similar in both strains.

The production of acetyl-CoA was very similar to the wild type, i.e., formed through pyruvate: ferredoxin oxidoreductase and mostly entering the ABE fermentation pathway. In terms of ABE fermentation flux, the acetoacetyl-CoA condensation flux was lower in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. However, there was higher flux towards butyryl-CoA synthesis, which was a result of reduced acetone forming flux. There was also butyrate formation which was not predicted in the wild type simulation, and nearly all produced butyrate was reassimilated through the butyrate-acetoacetate CoA-transferase reaction. This was the only predicted route of acetoacetate formation in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, whereas in the wild-type acetone formation solely occurred through acetate reassimilation. The acetate forming flux in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} increased by 64 % compared with the wild type and was secreted from the cell.

In *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} the TCA cycle flux was reduced substantially compared to the wild type. Pyruvate carboxylase flux decreased by 74 %, and flux through the cycle from oxaloacetate to α -ketoglutarate decreased by 90 %. Unlike the wild type, flux from α -ketoglutarate to succinyl-CoA was effectively zero in the in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} simulation, and there was no flux from succinyl-CoA to succinate. Succinate was only formed via fumarate in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, resulting in a lower succinate secretion flux.

In summary, glycolytic flux in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was similar to the wild type, but there was reduced synthesis of biomass precursors, such as amino acids and nucleotides, coinciding with the reduced growth rate. Meanwhile, a higher proportion of carbon flux entering the ABE fermentation pathway was utilised in butanol formation.

As with the simulation constrained with estimated fluxes from fermentations of wild type *C. saccharoperbutylacetonicum* N1-4, I then ran FVA at the experimental growth rate of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} (Fig. 4.3). As in the wild type FVA, there was a substantial amount of variation in the fermentation pathway reactions, such as the method of glucose import, flux through the PPP, TCA cycle and the ABE fermentation pathway. In *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} however, the upper range on these reactions were lower than in the wild type FVA. Differences arose in the pyruvate: ferredoxin oxidoreductase reaction and the pyruvate formate lyase flux ranges. These were not interchangeable in the *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} FVA. This was due to the changes I made to the model. In the *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} simulation, I removed the formate dehydrogenase reaction from the model, as I had found that there was no gene in the *C. saccharoperbutylacetonicum* N1-4 genome in chapter three. Another difference was the ferredoxin hydrogenase reaction. In the wild type the lower bound on this reaction was zero, but in the *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} FVA the lower bound on this reaction was approximately 10 mmol gDW⁻¹ h⁻¹.

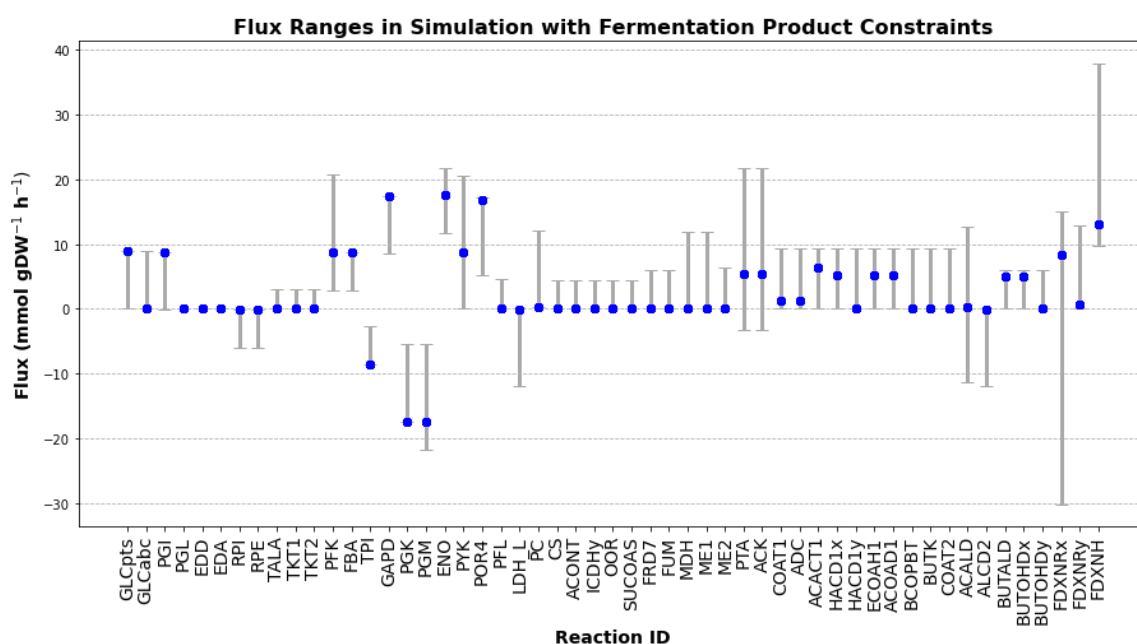


Figure 4.3 – Predicted upper and lower ranges of central metabolic fluxes in simulations of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} metabolism with an experimentally constrained glucose uptake and non-growth associated ATP maintenance cost with the rates of acetone, butanol, ethanol, butyrate, and lactate production constrained to values estimated from batch fermentation data of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T}. The NGAM cost was set to 14.4 mmol/gDW h⁻¹ and the rate of glucose uptake was set to 8.85 mmol/gDW h⁻¹. Flux ranges were calculated by taking the predicted growth rate in a simulation optimised for biomass formation where glucose uptake and NGAM were the only constraints. FVA was then used to calculate the flux ranges at a fraction of optimum of 0.43 (representing the experimental growth rate as a fraction of the theoretical maximum predicted by FBA). This was the experimentally determined growth rate divided by the maximum predicted growth rate. Blue dots indicate the predicted fluxes for the respective reactions predicted in the fermentation constrained simulation, shown in Figure 4.2. Reaction identifiers correspond to the reactions shown in Appendix B.

In summary, *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} had a higher butanol selectivity compared with the wild type. The total amount of carbon entering the ABE fermentation pathway was similar in both strains, but *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} converted more carbon to butanol, at the expense of growth.

4.3.3 *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} is Pareto Optimised for Butanol Production

The production of butanol in *Clostridium* is coupled to growth, meaning that increasing the rate of butanol production reduces the growth rate, as less ATP is generated through butanol formation than acetate or butyrate formation (see chapter three). I wanted to explore this trade off using the genome scale metabolic model of *Clostridium* metabolism. I also wanted to compare how both the wild type and *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} performed compared to the most efficient butanol producing scenario predicted by FBA.

Using the glucose uptake rate and NGAM determined from batch fermentations as constraints, I calculated the maximum possible rate of butanol production across a range of growth rates, from zero growth to the maximum predicted growth rate. This was performed twice, with NGAM and glucose uptake rates for the wild type strain and *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. The results of these simulations are shown as Pareto frontiers in **Fig. 4.4**. The dashed blue line shows the maximum achievable rate of butanol production at a given growth rate. The trade-off between growth and butanol production is clear from the dashed blue line, as at a growth rate of $\sim 0.05 \text{ h}^{-1}$ in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} and $\sim 0.012 \text{ h}^{-1}$ in the wild type the maximum achievable rate of butanol production decreases sharply. This indicates a shift to acid production, which is required to generate ATP for maintenance and biomass formation. In both strains, at the maximum predicted growth rate no butanol production was predicted. This is because all fermentation flux is directed to form acetate, which maximises ATP formation. The area below the dashed blue line, shaded blue, shows the feasible solution space of the model. In this space, butanol production is less than the predicted maximum (dashed blue line) and could be increased without reducing the growth rate. To demonstrate this concept, the predicted butanol producing flux and growth rate determined in batch fermentations for the wild type strain were plotted on the graph (**Fig. 4.4a**). Here, the wild type was sub-optimal for butanol production, with the rate of butanol production seen in batch fermentation being 61 % of the theoretical maximum. From a purely stoichiometric standpoint, at the experimentally observed glucose uptake rate and maintenance cost, there exists a more efficient butanol producing flux distribution. This contrasts with the predictions of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}, which was close to the dashed blue line, indicating that it is effectively optimised for butanol production, with the rate of butanol production in batch fermentation being 91 % of the theoretical maximum (**Fig. 4.4b**). The

stoichiometric constraints on the network mean that increasing butanol production in this strain would lead to a drastic reduction in growth rate.

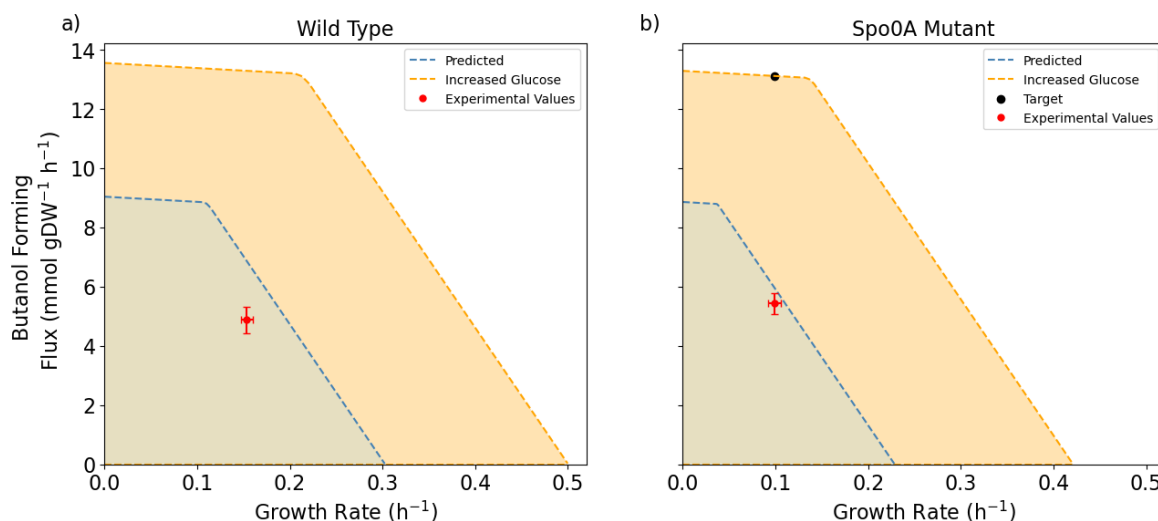


Figure 4.4 – Pareto optimisation of a) *C. saccharoperbutylacetonicum* N1-4 and b) *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} with glucose uptake and fermentation product secretion rates constrained to experimental data. The dashed blue line is the Pareto front. The dashed orange line is the Pareto front obtained when glucose uptake was increased by 50% over the experimental value. The blue shading represents the model predicted feasible solution space under the experimental data. The orange shading represents the model predicted feasible solution space when the glucose uptake rate was increased by 50 % over the experimental data. The growth rate and estimated butanol producing flux determined from batch fermentation are shown as red dots, which were performed in triplicate, with error bars ± 1 standard deviation. In b), the black dot indicates the simulated increased glucose uptake strain with optimal butanol production.

The reasons for *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} being optimal for butanol production but not the wild type is likely a result of the reduced rate of glucose uptake in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. The maximum predicted growth rate decreased in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, being 0.229 h⁻¹, whereas in the wild type the maximum predicted growth rate was 0.327 h⁻¹. To increase the rate of butanol production in the *spo0A* mutant strain, the solution space needs to be increased before attempts are made to engineer efficient conversion of glucose to butanol. To simulate this, the glucose uptake rate was increased by 50 % for both the wild type and the *spo0A* mutant, and a new Pareto frontier was generated (orange dashed line in **Fig. 4.4**). In this scenario, the experimentally observed rate of butanol accumulation in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} is sub-optimal and could be engineered to have a higher butanol producing flux. The maximum possible growth rate increased to 0.53 h⁻¹ for the wild type and 0.421 h⁻¹ for *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. In addition, the rate of butanol production could be increased to 14.07 mmol gDW⁻¹ h⁻¹ in the wild type and 13.11 mmol gDW⁻¹ h⁻¹ in the wild type and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} respectively, while achieving the same growth rate determined in batch fermentation.

4.3.4 *In silico* Investigation of Increased Glucose Uptake

Increasing the rate of glucose uptake will enable an increase in rate of butanol production by increasing carbon availability, but this would also produce more ATP as a result of increased glycolytic flux. This may need to be addressed to prevent accumulation of glycolytic intermediates caused by higher ATP levels in the cell, which may result in allosteric inhibition of glycolysis (Chandel, 2021). To see how the model balanced ATP production and consumption when provided with more glucose, two simulations were performed with a 50 % increase in glucose uptake compared to the experimental value for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. Constraints were placed on the model to simulate the 'target strain' (black dot in **Fig. 4.4b**). The growth rate was set to the value determined in batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} (0.1 h⁻¹) and the model was optimised for butanol production with and without the minimisation of sum of fluxes as an additional objective. Butanol production was selected as the objective to see how ATP was balanced in an ideal, i.e., butanol optimised strain. Minimisation of sum of fluxes prevents futile cycles from forming and will find the most efficient method of butanol production and ATP generation and consumption. In contrast, standard FBA will not avoid futile cycles forming; if these are identified in the model it could indicate that the cell may need engineering to balance its ATP consumption and formation.

I performed the two simulations with the 50 % increase in glucose uptake rate. I first compared the total ATP generation in these simulations to the simulation with *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} fermentation constraints. The increased glucose uptake simulation optimised with FBA had a 25 % increase in total ATP forming flux, while the pFBA optimised simulation had a 15 % increase in total ATP forming flux. The difference in total ATP formation between the FBA and pFBA simulations shows that the increased glycolytic flux produced more ATP than was required to meet the biomass synthesis and NGAM constraints. I wanted to investigate the fate of the additional ATP in these two simulations.

The ATP budgets in the increased glucose uptake simulations optimised with FBA and pFBA are shown in **Fig. 4.5**. The only source of ATP in both simulations was glycolysis, as butanol production was maximised no ATP was produced through acetate or butyrate synthesis. The flux through pyruvate kinase (PYK) was identical in both FBA and pFBA simulation (13 mmol gDW⁻¹ h⁻¹). The phosphoglycerate kinase (PGK) flux, however, decreased by 12 % in the pFBA simulation. Both simulations had identical flux through fructose bisphosphate aldolase and triose phosphate isomerase, so were producing the same amount of glyceraldehyde-3-phosphate. I investigated the fate of the latter in the pFBA simulation. The glyceraldehyde-3-phosphate that was not phosphorylated to 1,3-bisphosphoglycerate was directly converted to 3-phosphoglycerate through

the glyceraldehyde-3-phosphate-dehydrogenase reaction (Fig. 4.6). This avoided generating ATP through phosphoglycerate kinase. In contrast, the FBA simulation did not avoid ATP generation in this manner, with all glyceraldehyde-3-phosphate being phosphorylated and subsequently generating ATP through the phosphoglycerate kinase reaction. The excess ATP that was generated was utilised in the NGAM reaction, which had a minimum constraint but no constraint on the upper bound through this reaction. The NGAM flux in the glucose uptake simulation optimised with FBA increased by 21 % compared to the value estimated from batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{1261T}. These results show that, even accounting for the reduction in ATP formation as a result of no acid formation, the ATP generated with a 50 % increase in glucose uptake would be in excess to the cells requirements for growth and maintenance.

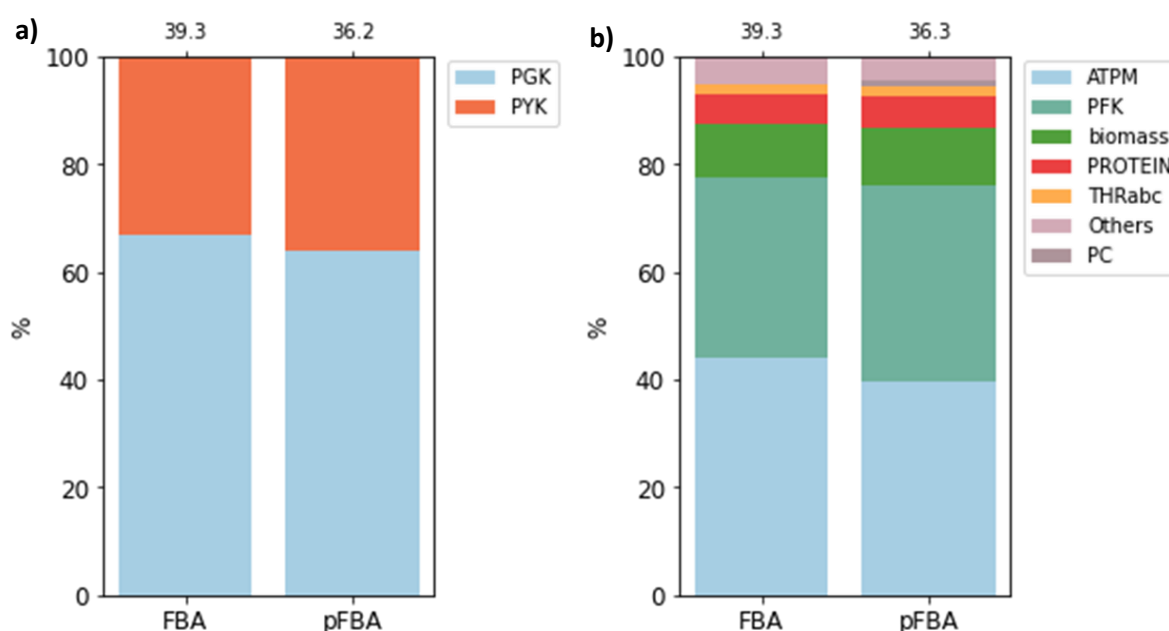


Figure 4.5 – ATP usage in simulations of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{1261T} maximising butanol production with the growth rate constrained to experimentally determined values from batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{1261T} and glucose uptake 50 % higher than values determined from batch fermentations a), ATP production fluxes, b), ATP consumption fluxes. The value above the bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Production and consumption fluxes may not be equal due to rounding. In both simulations, growth was constrained to be the same as the experimental value (~ 0.1 h⁻¹) and butanol production was the objective. Legend refers to the identifier of the reaction; PGK - ATP:3-phospho-D-glycerate 1-phosphotransferase, PYK – pyruvate kinase, ATPM – NGAM, PFK - 6-phosphofructokinase, biomass – lumped biomass synthesis reaction, PROTEIN - protein synthesis for biomass, THRabc – threonine ABC transport, PC – pyruvate carboxylase. Other represents grouped reactions which contribute less than 5 % of the total ATP forming or consuming flux.

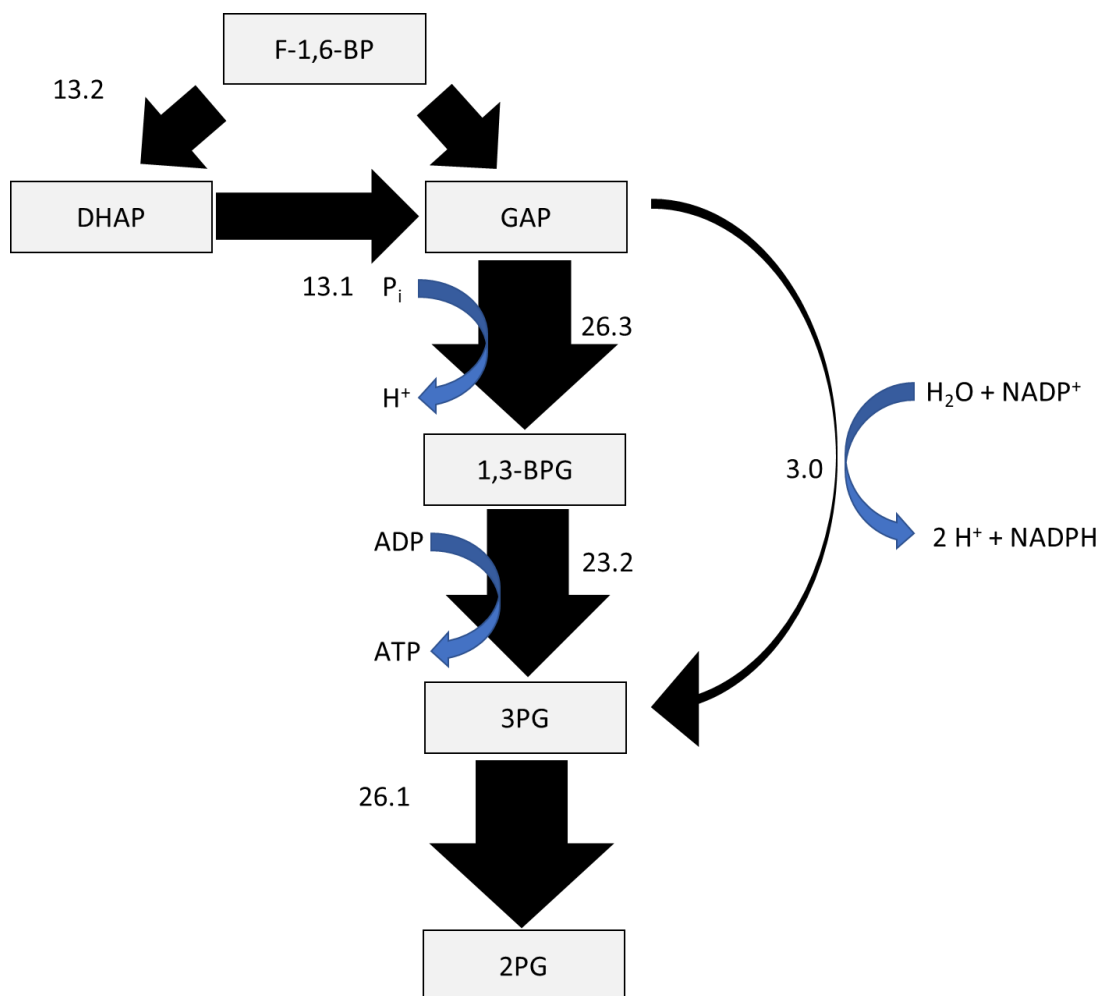


Figure 4.6 – Glycolytic bypass in pFBA simulation optimised for butanol production with glucose uptake constrained to 50 % higher than experimental values for *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, growth constrained to experimental value for *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. Blue arrows show cofactors/reactants. Numbers indicate estimated flux through the reaction. Arrow width is roughly proportional to flux through the reaction. Metabolites abbreviations are shown in **Appendix A**.

The increase in glucose uptake by 50 % was selected as an arbitrary amount for the FBA simulations. I wanted to see if ATP would be in excess across a range of increased glucose uptake values. I performed a stepwise set of simulations ranging from the experimental glucose uptake rate to a 100 % increase. At each step, the growth rate was constrained to the experimental value for *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, and the model was optimised for butanol production. The achieved butanol producing flux and the predicted NGAM flux are shown in **Fig. 4.7**. Butanol production was clearly linked to the rate of glucose uptake and increased proportionally. The glucose uptake rate could increase by 34 % over the experimental value without an increase in the NGAM value. Any further increase however resulted in an increase in NGAM, indicating that excess ATP would be generated at the given glucose uptake rate.

This finding has important implications for the metabolic engineering of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. As the butanol producing capability of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} is limited by the rate of glucose uptake, the larger the increase in glucose uptake that can be achieved the larger the potential increase in butanol production. Given this, it is likely that any strain with a substantial increased glucose uptake capability will also need to be engineered with some form of ATP consuming system.

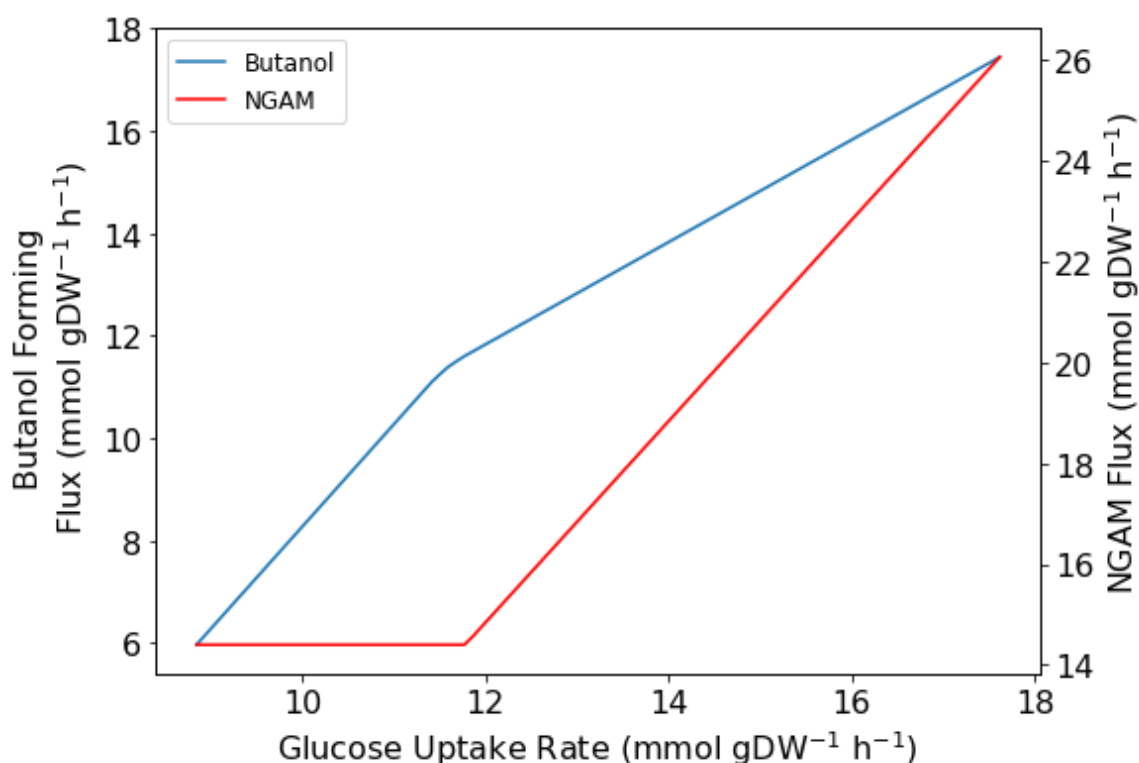


Figure 4.7 – Predicted NGAM flux and maximum butanol producing flux with increasing rate of glucose uptake rate in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. Glucose uptake rate was increased in a stepwise manner from the experimental value for *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} to twice the experimental value. Butanol production was set as the objective. The predicted flux through the NGAM at each time point is shown. The experimentally determined NGAM value for *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was set as the lower bound for the NGAM, with the upper bound set to 1000 mmol gDW⁻¹ h⁻¹. Growth rate was constrained to the experimentally determined value for *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} at all glucose uptake rates.

I next decided to investigate potential mechanisms by which increased glucose uptake could be achieved using the model. The *C. saccharoperbutylacetonicum* N1-4 model has two possible routes of glucose uptake: the PTS and the ABC transporter. In the simulations with increased glucose uptake, the glucose uptake rate was set by defining upper and lower bounds through a 'boundary' reaction. Boundary reactions are those that allow mass inequalities in the model, and are termed 'source' reactions, for those that introduce a compound into the system, and 'sink' reactions, for those that remove compounds from the system. The method of glucose import was not specified when I originally set the constraints on the model. Now I wanted to investigate how the cells imported glucose. In the simulations with a 50 % increase in glucose uptake, only the PTS was used in both the FBA and pFBA simulations. I used FVA to see if the use of the PTS was important for maintaining growth, or simply an effect of the minimisation of sum of fluxes. The lower and upper bounds through the glucose PTS transporter and the glucose ABC transporter were 0 and 13.3 mmol gDW⁻¹ h⁻¹, respectively, meaning either one could carry the increased glucose import flux.

Increasing glucose uptake experimentally could therefore be achieved by increasing flux through either of the PTS or ABC transport. The PTS system involves multiple proteins, with non-specific and specific components (Mitchell, 2016). There are many annotated PTS genes in the *C. saccharoperbutylacetonicum* N1-4 genome, but the contribution of these genes to the uptake of glucose has not been elucidated, and there is no experimental evidence for the expression of these genes. This is also the case for glucose ABC transporters. Rather than attempt to identify from the large number of candidate genes which are responsible for glucose uptake, I searched for literature examples of increased glucose uptake, with an aim of exploring their viability with FBA. An alternative to overexpressing PTS or ABC transporters would be the expression of a glucose symporter. One candidate is GalP, a H⁺: glucose symporter from *E. coli*, which has been used as an alternative to the PTS system for glucose uptake in PTS negative strains of *E. coli* (Flores *et al.*, 1996; Hernández-Montalvo *et al.*, 2003; de Anda *et al.*, 2006).

To test the viability of using this transporter to increase the glucose uptake rate experimentally, I ran FBA simulations with a H⁺: glucose symporter as a mode of glucose uptake. Utilising a H⁺: glucose symporter for glucose import might require changes in other metabolic fluxes. To investigate this, I took the following approach. A reaction for H⁺: glucose symport was added to the model. Then, I set a minimum flux through the symporter at half the flux through the PTS, modelling a situation where a 50 % increase in glucose import would be achieved by overexpressing the H⁺: glucose symporter. I set the biomass and butanol forming constraints to

those in the hypothetical ‘engineered strain’ (black dot in **Fig. 4.4b**). I then compared fluxes through central metabolic reactions when glucose uptake was increased only through the PTS system.

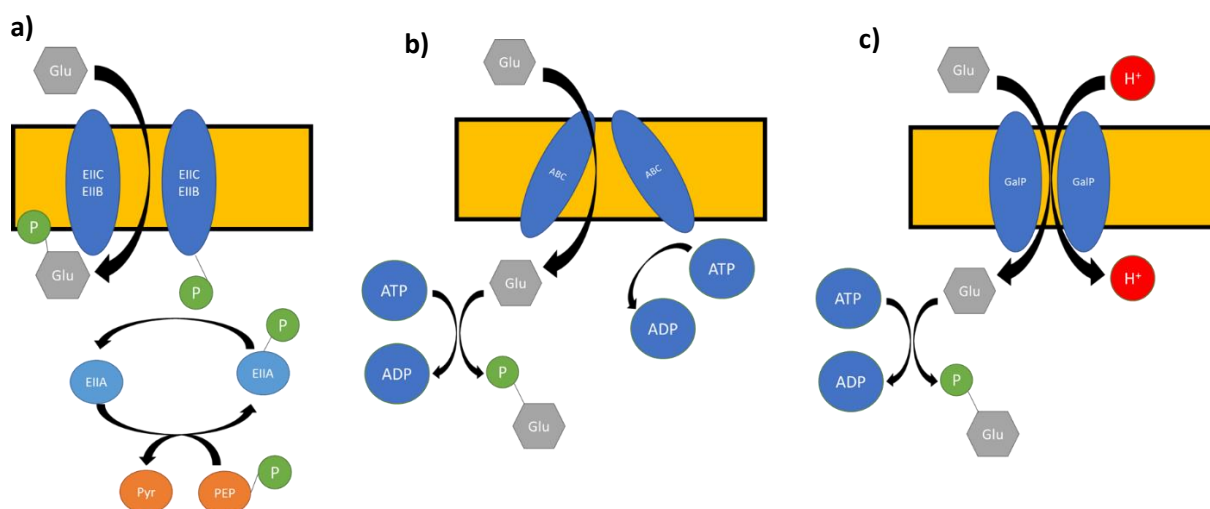


Figure 4.8 – Schematic of the three different methods of glucose import in the *C. saccharoperbutylacetonicum* N1-4 model; a) PEP-dependent phosphotransferase system (PTS), b) ATP binding cassette (ABC) transporter, c) GalP H⁺ glucose symporter. Abbreviations refer to the following compounds: ATP – adenosine triphosphate, ADP – adenosine diphosphate, EIIA/B/C – Enzyme II A/B/C, Glu – glucose, H⁺ – proton, P – phosphate group. A condensed version of phosphate transfer from PEP to EII is shown in the PTS diagram for brevity.

The first diverging flux was hexokinase, which was required to phosphorylate the imported glucose. The PTS system phosphorylates glucose as it is imported, however this is not the case using either a H⁺ glucose symporter or ABC transporter (**Fig. 4.8**). In the solely-PTS-utilising simulation, there was no predicted flux through the hexokinase reaction, while in the H⁺ glucose symporter simulation the flux was 4.4 mmol gDW⁻¹ h⁻¹. The other flux that increased was the pyruvate kinase reaction, which increased by 52 % in the H⁺ glucose symporter simulation compared to the solely PTS utilising simulation. In the latter, around half of total PEP was utilised as a phosphoryl donor during glucose import. The remaining PEP was utilised in the pyruvate kinase reaction. In contrast, as the H⁺ glucose symporter simulation did not use PEP for the additional glucose uptake, the additional PEP was utilised by pyruvate kinase. This also generated the ATP required by the hexokinase. All other fluxes in central metabolism were virtually identical. This suggests that effectively increasing glucose uptake using the H⁺ glucose symporter system will require expression of H⁺ glucose symporter, hexokinase, and pyruvate kinase simultaneously, assuming that activity of the enzymes *in vivo* is insufficient to carry increased flux.

Here, I have found that *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} was effectively optimised for butanol production. As a starting strain for metabolic engineering, the simulations have shown that the rate of glucose uptake needs to be increased before attempts are made to maximise conversion of glucose to butanol. Also, ATP consuming mechanisms will likely need to be inserted

and/or overexpressed to balance the cellular ATP budget, particularly at higher glucose uptake rates, if high butanol production is to be achieved.

4.3.5 Viability of Proposed Engineering Strategy in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}

The engineering strategy I proposed in chapter three was based on constraints and the predicted flux distribution of wild type *C. saccharoperbutylacetonicum* N1-4. Considering the finding that *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} is Pareto optimised for butanol production, I needed to see if the engineering strategy I proposed was still relevant. I ran a simulation to approximate the implementation of both the strategies I suggested in chapter three and those I developed in this chapter. I ran a simulation with the following constraints. First, the H⁺: glucose symporter was added to the model with the same constraint on glucose uptake as in the previous section. Next, the formate dehydrogenase reaction was added to the model. To model pyruvate formate lyase overexpression, the pyruvate formate lyase flux was set to be 50 % of that through the pyruvate:ferredoxin oxidoreductase reaction in the simulation with *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} fermentation constraints (**Fig. 4.2**). To model the increase in *bcs*-operon enzymes I set the lower bound on flux through the butyryl-CoA dehydrogenase reaction to 50 % higher than that in the *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} fermentation constraint simulation. Finally, the growth rate was constrained to be the same as the experimentally determined for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. This was assuming that an ATP consuming system could be expressed to balance the additional ATP produced because of the increased glucose uptake rate, preventing the formation of additional biomass. The model with these constraints was optimised with the minimisation of sum of fluxes as the objective, and the predicted flux distribution is shown in **Fig. 4.9**. Here, the butanol forming flux was 50 % higher than that seen experimentally for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. As was seen in chapter three in the double acid knockout simulation without the butanol constraint, lactate was predicted to be the major fermentation product.

As when modelling the earlier increase in glucose uptake enzymes, the 50 % increases in flux were selected as an arbitrary value to reflect overexpression, as FBA does not model enzyme abundance directly (Bordbar *et al.*, 2014). Assuming that the expression of the enzymes leads to an increase in fluxes as modelled here, implementing the metabolic engineering strategies outlined here will lead to an increase in butanol formation.

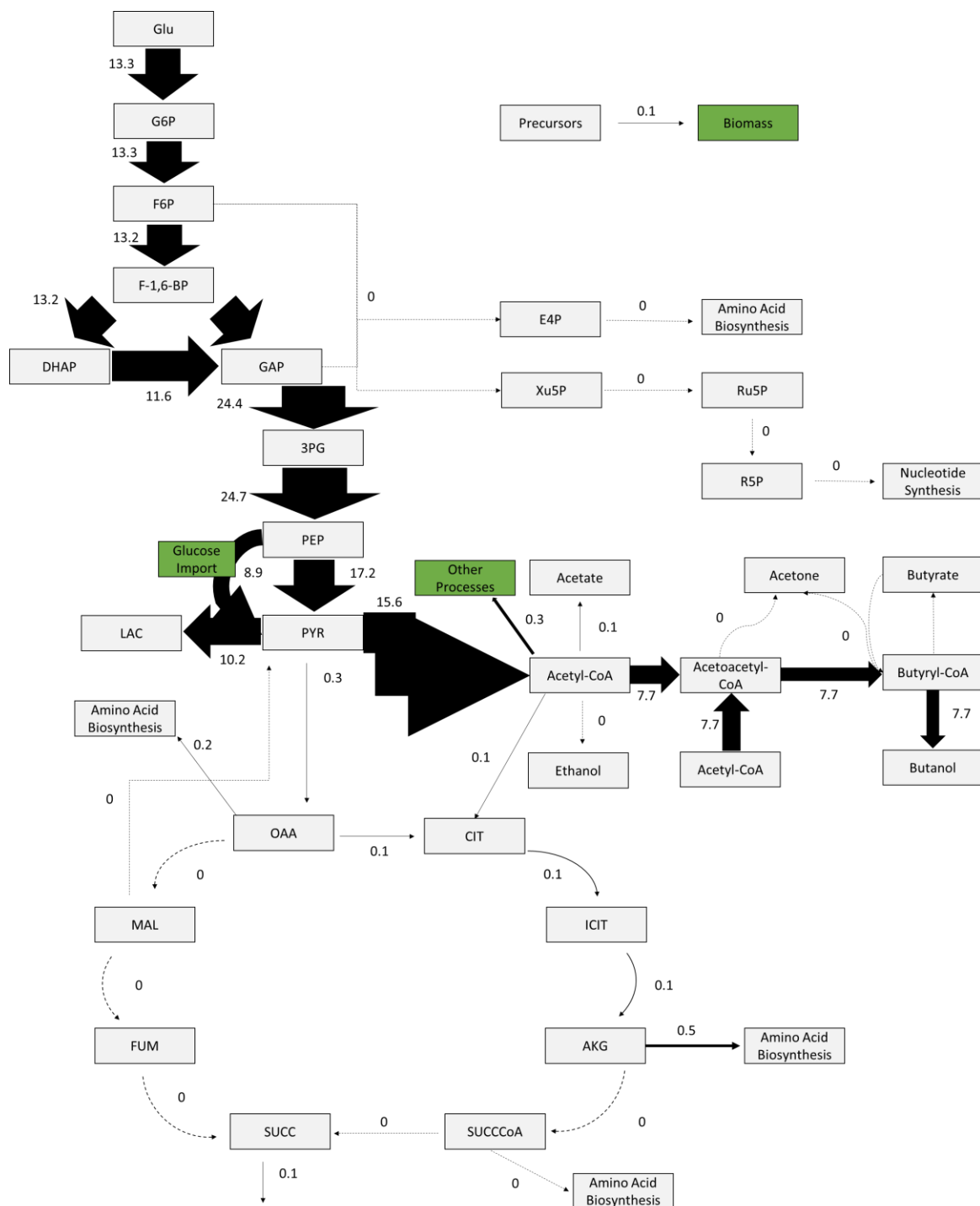


Figure 4.9 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} with implemented metabolic engineering strategy. Glucose uptake rate was set to be 50 % higher than experimental values for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T}. Formate dehydrogenase was added to the model, along with a H⁺: glucose symporter. Flux through pyruvate formate lyase were constrained to be 50 % of the pyruvate: ferredoxin oxidoreductase flux in Fig. 4.2. Butyryl-CoA dehydrogenase flux was set to be 50 % higher than the predicted flux in Fig. 4.2. These two constraints were used to simulation overexpression of the respective enzymes. Lower bound on NGAM and rate of biomass formation were set to the experimentally determined values. Arrow width is proportional to predicted flux through the reaction, with the flux value shown next to the respective reaction. Numbers next to the flux value are the percent increase or decrease in flux compared to the wild type. Red indicates a decrease in flux, blue indicates an increase. Dashed lines indicate a reaction with a predicted flux < 0.1. The metabolites and reactions outlined in this figure are shown in detail in Appendix A. All fluxes are given in units of mmol gDW⁻¹ h⁻¹, except for biomass formation which is given as h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

4.4 Discussion

4.4.1 Metabolic Engineering Approaches to Increase Glucose Uptake

The results of the batch fermentations and FBA simulations in this chapter have demonstrated that *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} is Pareto optimised for butanol production. I found that the rate of glucose uptake was limiting *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, and that increased glucose uptake would be necessary to improve butanol production further.

In solventogenic *Clostridium*, the PTS is responsible for the uptake of glucose (Mitchell, 2016). PTSs generally consist of two general proteins, EI and HPr, and two components specific to a substrate, named EIIA and EIIB. A membrane channel protein, IIC, is responsible for the import of the substrate. The *C. saccharoperbutylacetonicum* N1-4 genome contains 31 complete PTSs (Mitchell, 2016), and so far, only the PTS responsible for sucrose uptake has been identified (Zhang *et al.*, 2018). In this chapter, the PTS was the preferred system for glucose uptake with both fermentation constraints and when simulating increased glucose uptake. Increasing the native glucose PTS experimentally would first require identifying the respective PTS. While this could be achieved by bioinformatic analyses, such as sequence alignments of candidate genes to characterised glucose uptake systems from *C. acetobutylicum* (Tangney and Mitchell, 2007) and *C. beijerinckii* (Essalem and Mitchell, 2016), confirmatory experiments would still be required. Presence of a glucose specific PTS would need to be confirmed through sugar phosphorylation assays (Mitchell, Shaw and Andrews, 1991). The specificity of candidate glucose specific PTS genes could then be confirmed through heterologous expression in PTS negative non-glucose metabolising *E. coli* (Essalem and Mitchell, 2016). A disadvantage of using a native PTS is the reduction in activity after the shift to solventogenesis, reported in *C. acetobutylicum* (Hutkins and Kashket, 1986) and *C. beijerinckii* (Lee and Blaschek, 2001), which may also occur in *C. saccharoperbutylacetonicum*.

There have been efforts to improve substrate uptake through PTSs. In *C. saccharoperbutylacetonicum* N1-4, Zhang *et al.* (2018) were able to increase sucrose consumption uptake by deleting the sucrose operon repressor *scrR*. In the more well studied *E. coli*, glucose uptake rates were improved by overexpressing EI (*ptsI*) (Nishio *et al.*, 2008), which may be a limiting step in the import of glucose. Altering the expression levels of EII^{Glc} in *E. coli* revealed that while the rate of glucose import was dependent on EII^{Glc}, the rate of glucose oxidation was not impacted by EII^{Glc} level (Ruyter, Postma and van Dam, 1991). Identifying the EI and EII in *C. saccharoperbutylacetonicum* N1-4 and increasing expression would be one possible method of increasing the rate of glucose uptake, but the gene encoding this protein would still require identification.

Given the difficulties associated with engineering the PTS, I instead proposed an alternative system of heterologous expression of a H⁺: glucose symporter, which has been utilised extensively in bacteria. A PTS knockout strain of *E. coli* was developed to increase the availability of PEP for the synthesis of aromatic amino acids (Flores *et al.*, 1996). Flores *et al.* selected strains of *E. coli* with inactivated glucose PTS in continuous culture and found that the evolved strains imported glucose through the GalP H⁺: glucose symporter. Another method is expression of glucokinase with a glucose facilitator protein, which restored the ability of an *E. coli* strain with an inactive *ptsG* to utilise glucose (Snoep *et al.*, 1994). In these studies, in PTS negative strains of *E. coli*, the aim was frequently to increase the pool of PEP for amino acid synthesis. As a proof of concept, Flores *et al.* expressed 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase in a GalP utilising *E. coli* strain. This strain produced nearly three-fold as much DAHP compared to the parental (PTS-utilising) strain expressing DAHP synthase, suggesting increased PEP availability in the PTS negative strain. Plasmid borne expression of GalP and Glk (glucokinase) in a PTS negative *E. coli* strain increased the rate of glucose uptake by 40 % compared to the PTS utilising wild type (Hernández-Montalvo *et al.*, 2003).

Rather than replace the native PTS in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}, I would express GalP as a method to enhance glucose uptake. In addition, the increase in PEP seen in GalP expressing *E. coli* (Flores *et al.*, 1996) is broadly in agreement with the simulations I performed in this chapter with the H⁺: glucose symporter. The model utilised pyruvate kinase to consume the PEP in my simulations, but this required a 50 % increase in flux compared with the PTS utilising simulation. As the accumulation of PEP is undesirable for butanol production, conversion of PEP to pyruvate by expressing pyruvate kinase (Pyk) may be necessary. The system I propose to increase glucose uptake in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} therefore consists of combined expression of GalP, Glk (glucokinase) and Pyk.

4.4.2 Increased ATP Consumption for Metabolic Engineering Applications

During the simulations with increased glucose uptake, I found that over a threshold increase of 34 % above the experimental rate of glucose uptake for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}, ATP production would be in excess. Depending on the objective, the simulation predicted either bypassing the generation of ATP through phosphoglycerate kinase or utilising the NGAM reaction to simply hydrolyse the ATP. This suggested that under conditions of increased glucose uptake ATP consuming systems may need to be engineered into the cell.

There are several examples where microorganisms have been engineered through the creation of ATP deficits. A strain of *E. coli* with an inactive F₁-ATPase subunit had an increased rate of glucose import, while the activities of pyruvate kinase and phosphoglycerate kinase increased by around

50 % compared to the wild type (Yokota *et al.*, 1997). Other *E. coli* strains with deletions of the *atp* operon (encoding ATP synthase) had a 40 % increase in respiration rate compared with the wild type and an increase in acetate titres (Jensen and Michelsen, 1992). In *E. coli* the expression of the F₁-ATPase subunit increased the rate of glycolysis by 70 % compared with the wild type (Koebsmann *et al.*, 2002). The control of glycolytic flux by ATP demand is not limited to *E. coli*. A strain of *Corynebacterium glutamicum* was generated with reduced H⁺-ATPase activity had a 2.7-fold higher rate of glucose consumption per cell than the wild type strain (Sekine *et al.*, 2001). The model yeast *S. cerevisiae* has also been engineered to reduce cellular ATP, generally to divert carbon flux away from biomass towards ethanol. Expression of the ATP hydrolysing alkaline phosphatase PHO8 increased ethanol yield and productivity (Semkiv *et al.*, 2014). Another example is the expression of apyrase, which increased the ethanol titre but only when fed a mixture of 1 % glucose and 10 % galactose (Semkiv *et al.*, 2016). There was no change in ethanol titre when the *S. cerevisiae* were fed 5 % glucose and 10 % galactose (Semkiv *et al.*, 2016).

The above examples suggest that creating an ATP deficit in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} might be sufficient to increase glucose import and/or the activity of glycolytic enzymes. There are several possible methods of ATP consumption which I could exploit in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. Expressing enzymes such as the ATPase portion of a membrane proton transporter or an apyrase are methods which could reduce ATP accumulation (Semkiv *et al.*, 2016).

Futile cycles are another method of increasing substrate turnover, for example overexpressing phosphoenolpyruvate synthase in *E. coli* nearly doubled glucose consumption (Patnaik *et al.*, 1992), overexpressing phosphoenolpyruvate carboxykinase and phosphoenolpyruvate carboxylase increased glucose uptake along with acetate and pyruvate production (Chao and Liao, 1994). Semkiv *et al.* also expressed pyruvate carboxylase and PEP-carboxykinase in *S. cerevisiae*, which increased ethanol productivity (g of ethanol produced per g biomass) by nearly two-fold. Such approaches may result in trapped carbon, which would reduce carbon flux to the ABE fermentation pathway. This could be avoided by altering non-carbon metabolism, and I have developed one such approach. Polyphosphates are long chains of inorganic phosphate residues with a wide range of roles in bacteria, including motility and virulence (Rao, Gómez-García and Kornberg, 2009). ATP acts as a phosphate donor molecule, with the enzyme polyphosphate kinase extending the polyphosphate chain. Inorganic phosphate is removed from the chain through the action of exopolyphosphatase. Overexpressing both enzymes simultaneously could generate a non-carbon futile cycle. This has been reported in *E. coli*, which increased flux into the TCA cycle to increase ATP generation (Keasling, van Dien and Pramanik, 1998). As solventogenic

Clostridium rely on acid forming pathways to generate ATP rather than the TCA cycle (Amador-Noguez *et al.*, 2010), this could increase flux to the ABE fermentation pathway. Expression of these genes will require tight control to ensure that ATP is not completely drained, and an inducible promoter will also be developed. While a xylose inducible promoter exists, it is repressed by glucose (Zhang *et al.*, 2016), necessitating development of a non-glucose repressed inducible promoter. Once the substrate uptake has been improved, several different methods of ATP depletion will be trialled before the cell will be engineered to efficiently convert glucose to butanol.

4.4.3 Implications for Engineering *C. saccharoperbutylacetonicum* *spoOA*^{I261T}

Using the data from batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}, I found that this strain is Pareto optimised for butanol production. The already discussed glucose uptake system and ATP consuming reactions were designed to overcome this issue. I tested the effects of the engineering strategy I developed in chapter three in a simulation with increased glucose uptake. This resulted in an increase in butanol flux compared to the experimental value for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. It seemed that maintaining high NADH formation through the pyruvate-formate lyase/formate dehydrogenase system and increasing the abundance of the *bcs*-operon would be beneficial if the glucose uptake rate could be increased.

It is possible that the increase in glucose uptake would result in an increased rate of biomass formation due to the increase in ATP availability. The simulations I performed with increased glucose uptake showed that ATP could be hydrolysed through the NGAM to balance the additional ATP without increasing the rate of biomass formation. This trade-off between biomass formation and product yield has led many researchers to implement ATP consuming pathways or reduce ATP formation, to reduce biomass synthesis. There were efforts to replace the native *S. cerevisiae* glycolytic pathway with the Entner-Doudoroff pathway to reduce ATP formation, but the expressed 6-phosphogluconate dehydratase was not active enough to carry the desired flux (Benisch and Boles, 2014). Other, more successful approaches include expression of PEP-synthase in a lactic acid producing *E. coli* strain, which reduced the growth rate while increasing the lactate yield (Hädicke, Bettenbrock and Klamt, 2015). Many of the examples of ATP consuming reactions discussed in section 4.4.2 also reduced biomass production and increased the product yield (Yokota *et al.*, 1997; Wada, Narita and Yokota, 2007; Semkiv *et al.*, 2016). These examples suggest that the ATP utilising systems I proposed will be effective in preventing an increase in growth rate if glucose uptake is achieved in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}.

4.5 Conclusion

Here, I combined a genome scale metabolic model with batch fermentation data to gain a deeper understanding of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} than from batch fermentation

data alone. Perhaps most importantly, glucose uptake was identified as the key limitation in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{1261T}, while also showing the need for ATP consuming enzymes to be overexpressed. This chapter generated further *in silico* guided metabolic engineering approaches, which I will implement experimentally, along with the manipulations from chapter three.

Chapter 5: Metabolic Flux Analysis of an Asporogenic strain of *C. saccharoperbutylacetonicum* N1-4

5.1 Aim of Work in this Chapter

The aim of this work was to generate a detailed flux map of the asporogenic strain *C. saccharoperbutylacetonicum* *spo0A*^{I261T} using ¹³C-MFA. A secondary aim was to combine ¹³C-MFA with FBA to further analyse and explain the flux distributions found.

5.2 Introduction

5.2.1 Metabolic Flux Analysis

¹³C-MFA is the gold standard method for determining intracellular fluxes (Long and Antoniewicz, 2019). ¹³C-MFA has seen numerous uses in metabolic engineering, including to investigate intracellular fluxes in an industrial 1,3-propanediol producing *E. coli* strain (Antoniewicz *et al.*, 2007), a fatty acid overproducing strain of *E. coli* (He *et al.*, 2014), and identifying the sucrose metabolising pathways of the succinate producing bacterium *Basfia succiniciproducens* (Lange *et al.*, 2017). ¹³C-MFA has also been used in *C. acetobutylicum*. Previous MFA studies in *C. acetobutylicum* have revealed that isoleucine synthesis occurs through the citramalate synthase pathway (Au *et al.*, 2014) and was used to investigate *C. acetobutylicum* responses to various stressors in chemostat culture (Wallenius, Maaheimo and Eerikäinen, 2016). To my knowledge, no flux map of *C. saccharoperbutylacetonicum* N1-4 or derived strains has been published.

5.2.2 Synergy of Metabolic Flux Analysis and Flux Balance Analysis

¹³C-MFA metabolic network models tend to be reduced forms of networks, condensing linear sections of metabolic networks into single steps. As the network model used in MFA studies is thus a simplified representation of cellular metabolism, the balances of cofactors is usually not enforced. In contrast, the genome scale metabolic networks used for constraint-based techniques balance cofactors when performing simulations. In this manner, FBA and related techniques can be used to address this shortcoming of ¹³C-MFA and help understand the results of flux estimations. For example, a study which used ¹³C-MFA to determine fluxes in *E. coli* throughout batch culture also performed FBA to determine the cellular objective (biomass or ATP production) that best explained the ¹³C-MFA fluxes (Toya *et al.*, 2010). Throughout batch fermentation, Toya *et al.* found that during the initial few hours of culture the flux distribution estimated by ¹³C-MFA was closer to the FBA predictions optimised for biomass formation, but later in the culture ATP production as an objective gave better agreement to the ¹³C-MFA estimated fluxes. Another use of FBA was to investigate reasons for a higher maintenance cost under anaerobic versus aerobic conditions in *E. coli* (Chen *et al.*, 2011). Using FBA, Chen *et al.* identified proton secretion through a membrane bound ATPase as responsible for the additional ATP cost. ¹³C-MFA flux data has also been used to constrain FBA

models to improve the predictions of peripheral fluxes, i.e., those outside of central carbon metabolism, in an approach termed two-scale ^{13}C -MFA (García Martín *et al.*, 2015). ^{13}C -MFA has also been used to validate FBA predictions of metabolism. Non-stationary ^{13}C -MFA fluxes were compared against FBA predicted metabolic fluxes in the algae *Chlorella protothecoides* (Wu *et al.*, 2015). While most central metabolic fluxes were similar in FBA and ^{13}C -MFA, key differences in metabolic pathways (such as lower TCA cycle flux) were identified with ^{13}C -MFA, which could be important for engineering improved lipid production. In summary, FBA can provide initial estimations of metabolic fluxes based on optimal behaviour which are then verified with ^{13}C -MFA, or vice versa; the results from ^{13}C -MFA can be interrogated using a genome scale, fully mass balanced model.

The aim of this work was to generate a flux map of the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} using steady state ^{13}C -MFA. While it is possible to analyse dynamic systems using non-stationary ^{13}C -MFA approaches, these are more difficult to implement, requiring many more experimental measurements and more complex computation (Antoniewicz, 2015). Stationary ^{13}C -MFA is a mature, well-established technique that provides useful information about bacterial metabolism (Long and Antoniewicz, 2019). After generating a flux map of the asporogenic strain, I aimed to combine ^{13}C -MFA with FBA to gain an understanding of redox and energy balancing in this strain using refined flux constraints from the ^{13}C -MFA in the FBA model.

5.3 Results

5.3.1 Parallel Labelling Experiments

The first step in performing ^{13}C -MFA is the construction of the metabolic network model and the selection of isotope tracers. For both of these, I used a study of *C. acetobutylicum* (Au *et al.*, 2014) to guide my experiments. The metabolic network model of Au *et al.* was adapted for use with *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. The metabolic reactions for amino acids biosynthesis, glycolysis, the TCA cycle and the PPP were inspected, and found to be consistent with the KEGG pathways for *C. saccharoperbutylacetonicum* N1-4. In Au *et al.*'s model, glucose import proceeded through ABC transport, but it is well established that *Clostridium* use the PTS for glucose import (Mitchell, 2016). I replaced the ABC reaction with a PTS glucose import reaction. Next, I added exchange reactions for metabolites which I found to accumulate in the fermentation media not included in the base model, i.e., formate and lactate. In addition, I added a glyoxylate shunt to the model, which has been used in a previous flux analysis study of *C. acetobutylicum* (Desai, Nielsen and Papoutsakis, 1999); removal of this pathway resulted in a poor statistical fit of labelling data to the model.

Au *et al.* utilised [1-¹³C]glucose and [U-¹³C]glucose in parallel labelling experiments, and were able to estimate most central metabolic reactions with good resolution. I assumed that *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} central metabolic fluxes could also be resolved using this tracer combination, so these were selected for the tracer experiments.

The next step in ¹³C-MFA is the collection of isotopomer data for compounds labelled with the tracers. I performed isotope tracer fermentations to generate labelled amino acid data as follows. Six bottles containing defined CGM (Au *et al.*, 2014) were prepared and inoculated with asporogenic *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. Cells were grown in unlabelled glucose for the first seven hours of fermentation, at which point a bolus of 20 g L⁻¹ ¹³C isotope glucose was added to each culture. Three cultures received a bolus of [U-¹³C]-glucose and the other three culture received a bolus of [1-¹³C]-glucose. Supernatant samples were regularly collected throughout the fermentation and the concentrations of fermentation products were determined by HPLC. Cell density was measured by reading the absorbance of culture at 600 nm. The fermentation profile and growth curve are shown in **Fig. 5.1**. Technical problems arose during this experiment which meant that the OD_{600nm} readings are unreliable in this experiment, as evidenced by the drop in the density from ten to thirty-two hours, and the OD_{600nm} then increased (**Fig. 5.1a**).

The fermentation production profile was characterised by very low levels of acetone (**Fig. 5.1c**) and ethanol (**Fig. 5.1e**), with no butanol detected throughout the fermentation (**Fig. 5.1d**). In contrast to the results obtained in chapter 4, the bacteria appeared to show an acidogenic metabolism throughout the entire culture. Acetate was present in the media as a pH buffer but was also produced by the cells throughout the fermentation (**Fig. 5.1f**). There was a drop in acetate concentration seven hours into the fermentation, possibly diluted due to the addition of isotope tracer. Acetone appeared at detectable levels ten hours into the fermentation (**Fig. 5.1c**). Ethanol titres were low throughout the fermentation, never exceeding 0.01 g L⁻¹ (**Fig. 5.1e**). The organic acids butyrate (**Fig. 5.1g**) and lactate (**Fig. 5.1h**) were produced at a fairly constant rate through the fermentation. The most likely explanation for the failure to see solventogenic metabolism as before, is that a different growth medium was used. I used the defined growth medium recipe used by Au *et al.*, whose composition is detailed in materials and methods section 2.3.5. This medium only contained salts and minerals required for growth, with ammonium acetate as the pH buffer. This was a departure from the media used in the previous chapters when I performed batch fermentations. I opted for the defined media recipe as it does not contain any amino acids or additional carbon sources from the yeast extract and tryptone used in the TYIR media. It is common practice to use minimal media when performing ¹³C-labelling experiments with bacteria to prevent complications arising from the uptake of multiple unlabelled carbon sources (Long and

Antoniewicz, 2019). Amino acids and other carbon sources in yeast extract and tryptone can confound the labelling pattern and add additional complexity to the flux estimation procedure (Zamboni *et al.*, 2009).

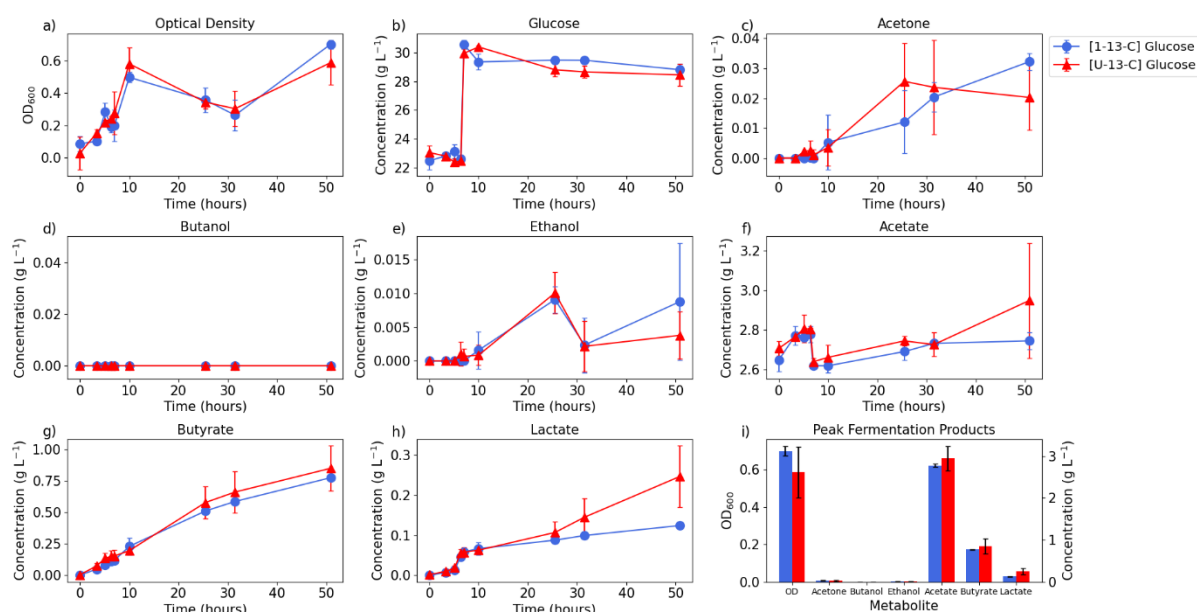


Figure 5.1 - Fermentation kinetics and peak metabolite concentration data from batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} in defined media. a) Optical density as measured by absorbance at 600 nm. Concentrations of b) glucose, c) acetone, d) butanol, e) ethanol, f) acetate, g) butyrate, and h) lactate in the fermentation media were measured using HPLC. Peak fermentation products are shown in i) Batch fermentations were performed in 30 mL serum bottles with 27 mL of media. All fermentations were conducted in defined CGM with 20 g L⁻¹ glucose. All fermentations were performed in triplicate. Error bars shown are ± 1 standard deviation. All cultures were inoculated from the same starter culture.

Exclusively acidogenic metabolism was also observed in *C. acetobutylicum* grown on the medium used here (Au *et al.*, 2014). The authors attributed this to cells being in the acidogenic phase of growth, however the presence of ammonium acetate as a pH buffer may also contribute to the lack of solvent production, as this compound has been reported to reduce solvent titres (Monot *et al.*, 1982). While any flux map estimated from this data will not capture solventogenic fluxes, a flux map of this strain will nevertheless prove useful in understanding biosynthetic fluxes during growth and will serve as a reference point for future ¹³C-MFA work on the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}.

Steady-state ¹³C-MFA relies on the metabolic network being at metabolic and isotopic steady state. To confirm metabolic steady state, I quantified the amino acids used for the ¹³C-MFA and calculated their abundance as a fraction of the dry cell pellet weight. These are shown in **Fig. 5.2**. Most amino acids were at a constant abundance throughout the experiment. Glycine had a slight increase in abundance from ten hours onwards but remained stable after 26 hours in both experiments. There was a slight increase in tyrosine from 32 to 51 hours with both tracers. It appeared that in both experiments the cultures reached a metabolic steady state by 32 hours.

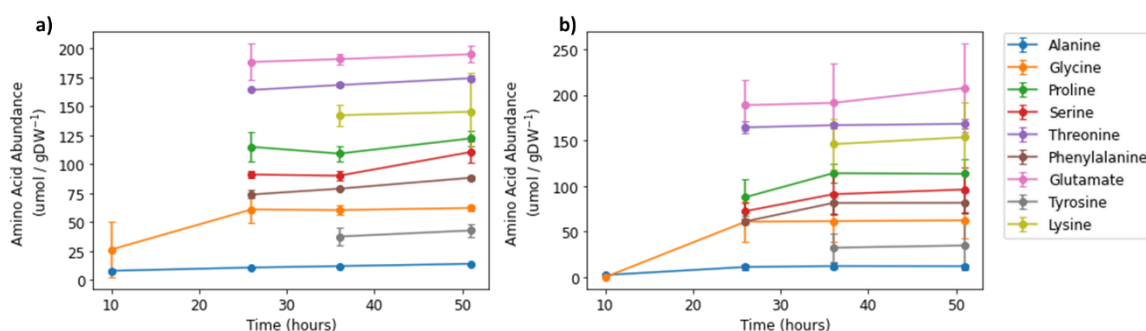


Figure 5.2 – Amino acid abundance per g of cell dry weight (DW) from cell pellet in ^{13}C -parallel labelling experiments with a) $[1\text{-}^{13}\text{C}]\text{glucose}$ as isotope tracer or b) $[\text{U-}^{13}\text{C}]\text{glucose}$ as isotope tracer. Protein hydrolysis and TBDMSCIS derivatisation were performed as described in materials and methods section 2.12.2. GC-MS analysis was performed as described in materials and methods sections 2.12.3 and 2.12.4. The microcentrifuge tube containing the dried pellet was weighed prior to derivatisation, assuming a tube weight of 1 g. Some amino acids were not detected throughout the whole experiment. Lysine and tyrosine were only detected from 32 hours onwards. Most other amino acids were detected from 26-hours onwards. This could be due to poor separation or insufficient cell mass, particularly at the ten-hour time point.

To confirm that, in addition to being at metabolic steady state, that the cultures were at isotopic steady state, I compared the mass isotopomer distributions of amino acids throughout the fermentation. The results from two representative amino acid fragments (Ala260 and Asp418, (Leighty & Antoniewicz, 2012)) are shown in **Fig. 5.3** and **Fig. 5.4** for both tracers. The time course profiles for the remaining amino acids are shown in **Appendix D**. Very few amino acid fragments were detected at the ten-hour time point, which was likely due to low biomass concentration. Most of the amino acids not detected at the ten-hour time point were later detected at the 26-hour time point, except for lysine and tyrosine, which were only detected from the 32-hour time point onwards. The mass isotopomer distributions for all measured amino acid fragments were unchanged throughout the experiments, indicating isotopic steady state.

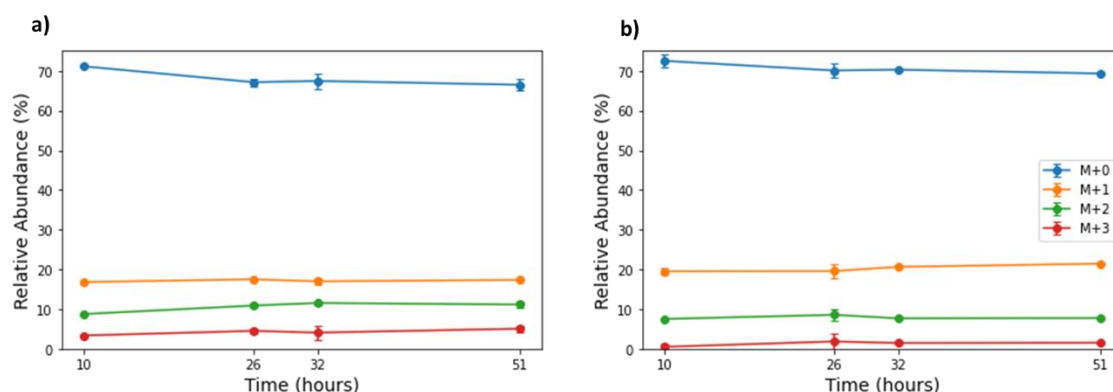


Figure 5.3 – Mass isotopomer distributions over time for Ala260 fragment in ^{13}C -MFA parallel labelling experiments with a) $[\text{U-}^{13}\text{C}]\text{glucose}$ as isotope tracer or b) $[1\text{-}^{13}\text{C}]\text{glucose}$ as isotope tracer. Protein hydrolysis and TBDMSCIS derivatisation were performed as described in materials and methods section 2.12.2. GC-MS analysis was performed as described in materials and methods sections 2.12.3 and 2.12.4.

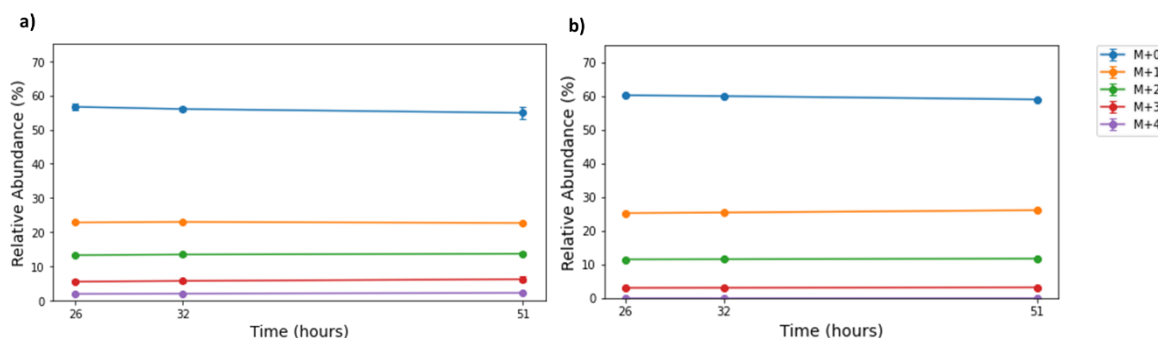


Figure 5.4 – Mass isotopomer distributions over time for Asp418 fragment in ^{13}C -MFA parallel labelling experiments with a) $[\text{U-}^{13}\text{C}]\text{glucose}$ as isotope tracer or b) $[1\text{-}^{13}\text{C}]\text{glucose}$ as isotope tracer. Protein hydrolysis and TBDMSCIS derivatisation were performed as described in materials and methods section 2.12.2. GC-MS analysis was performed as described in materials and methods sections 2.12.3 and 2.12.4. Asparagine was not detected at the ten-hour timepoint.

To confirm labelling of amino acid fragments, the extracted mass isotopomer distributions were compared to those predicted by the INCA software (Young, 2014) at metabolic steady state. These are listed in **Appendix E** for each amino acid fragment and each tracer. For most fragments, the measured data was in good agreement with the estimation. Proline had a higher-than-expected M+0 abundance, which could have been due to low turnover of this residue, or low abundance in cellular protein.

5.3.2 Metabolic Flux Map of Asporogenic *C. saccharoperbutylacetonicum* *spoOA*^{1264T}

A flux map using labelled data from the 32-hour timepoint (25 hours after the addition of tracer, data given in **Appendix F**) is shown in **Fig. 5.5**. Parallel fitting of the three $[\text{U-}^{13}\text{C}]\text{-glucose}$ datasets and the three $[1\text{-}^{13}\text{C}]\text{-glucose}$ datasets gave an acceptable fit, with an SSR of 353.6 with 326 degrees of freedom, which was within the acceptable threshold (277.9 – 377.9). The 32-hour time point was chosen for presentation as at the ten-hour time point very few amino acid fragments were detected, likely a result of low cell mass in the media. In addition, there was no significant

difference in the levels of fermentation products as measured by HPLC, whereas the final time point had significant differences in the concentrations of lactate, indicating different flux distributions in the two sets of parallel cultures. Reliable labelling experiments require consistent behaviour between cultures (Long and Antoniewicz, 2019).

Glucose oxidation occurred solely through the EMP pathway, with no predicted flux through the ED pathway. Very low flux was predicted through the non-oxidative steps of the oxidative pentose-phosphate pathway, which was used to generate ribose-5-phosphate for biomass synthesis, and erythrose-4-phosphate for synthesis of aromatic amino acids in the Shikimate pathway.

In *C. saccharoperbutylacetonicum* N1-4 there are two routes for acetyl-CoA formation, pyruvate: ferredoxin oxidoreductase and pyruvate formate lyase. In the flux map 96 % of acetyl-CoA was formed through pyruvate: ferredoxin oxidoreductase, with the remaining 4 % through pyruvate formate lyase. The formate produced through this reaction was secreted from the cell, satisfying the external constraint on formate secretion. Most acetyl-CoA (56 %) was utilised in the ABE fermentation pathway to form acetate and butyrate. Approximately 24 % of acetyl-CoA entered the TCA cycle and was converted to citrate and isocitrate via citrate synthase and aconitase, respectively. The remaining acetyl-CoA was used in leucine biosynthesis. Almost all the isocitrate generated in the TCA cycle (98 %) was converted into glyoxylate and succinate by isocitrate lyase. The remaining isocitrate was converted to α -ketoglutarate. A low portion of α -ketoglutarate was utilised in biomass formation, and the remainder in glutamate synthesis. The glyoxylate and succinate produced through the isocitrate lyase reaction were both converted to malate through malate synthase (glyoxylate) and through succinate dehydrogenase and fumarase (succinate). There was flux into the TCA cycle via pyruvate carboxylase. Approximately 25 % of oxaloacetate was converted to malate, and 1 % was utilised to synthesise aspartate. The remaining 74 % was converted to malate by malate dehydrogenase. All malate which was produced by malate dehydrogenase, malate synthase and fumarase was converted back to pyruvate through the NADP-dependent malic enzyme. This appeared to be a futile cycle, shown in detail in **Fig. 5.6**. The net effect of the futile cycle was transhydrogenation of NADH and NADP⁺ to NAD⁺ and NADPH, with consumption of one mole of ATP for each mole of NADPH formed. There are important caveats here. The cofactor dependencies of these enzymes were unchanged from the original model, which

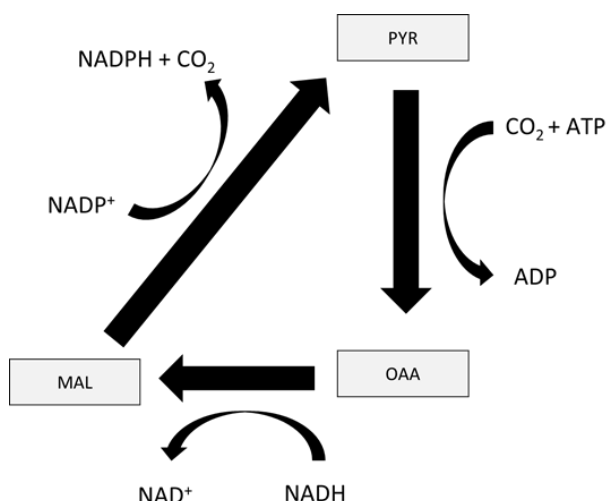


Figure 5.6 – Detail of futile cycle from MFA estimations of intracellular fluxes. Pyruvate was converted to oxaloacetate via pyruvate carboxylase. Oxaloacetate was then converted to malate by malate dehydrogenase. Finally, malate was converted back to pyruvate by NADP-dependent malic enzyme. Metabolite abbreviations are as follows: PYR: pyruvate, OAA: oxaloacetate, MAL: malate.

was based on previous models and metabolic pathway databases (Au *et al.*, 2014). To my knowledge, the exact cofactor dependencies of the above futile cycle enzymes has not been demonstrated *in vivo*, which has implications for both the flux estimations performed here and interpreting the physiological role of this cycle. In addition, determining the metabolic fluxes around the anaplerotic node is particularly tricky in many prokaryotic organisms (Kappelmann, Wiechert and Noack, 2016). Given these uncertainties, these fluxes should be viewed as, at best, preliminary estimates.

The futile cycle warranted further investigation. I examined the upper and lower bounds predicted by parameter continuation analysis through the ATP consuming pyruvate carboxylase reaction, which had a lower bound of $3.1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ and an upper bound of $23.2 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. If the purpose of this reaction was solely to dispose of excess ATP, I would expect the lower bound on the flux through the futile cycle to be zero. This is because the MFA model contains an ATP hydrolysing reaction, which dephosphorylates ATP to ADP. The lower bound through the ATP hydrolysis reaction was $0 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ and the upper bound was $14.7 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. The non-zero upper bound through this ATP hydrolysis reaction suggests that surplus ATP was being generated, i.e., greater than required for biomass synthesis.

The fluxes through glycolysis and acetate and butyrate formation produced more ATP than was required for biomass synthesis. As the medium used in these experiments included sodium acetate, as already discussed, solvent formation was prevented. Solvent formation would be a route of preventing excess ATP formation (by reducing the ATP-forming acidogenic fluxes), as discussed in chapter three. It may be that the growth medium prevented the formation of solvents, which led to acid fluxes above what would be expected. An alternative reason for the ATP generation above that required for biomass synthesis is an ATP demand from additional ATP consuming reactions. One such candidate is the membrane bound H^+ : ATPase. Solventogenic *Clostridium* maintain a ΔpH across their membranes (Huang, Gibbins and Forsberg, 1985), which requires membrane bound ATPases to pump protons across the membrane (Bowles and Ellefson, 1985). This could cause a significant ATP demand, but this would not be captured by the MFA model as protons were not included in the reaction balances.

The net effect of the futile cycle was also a transhydrogenase reaction, oxidising NADH in the malate dehydrogenase reaction, then reduction of NADP^+ through the malate: NADP dehydrogenase reaction. The futile cycle had an NADPH generating flux of $22.9 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. As is common when performing MFA, the cofactors NAD^+ , NADP^+ , NADH and NADPH were present but allowed to be unbalanced, as the metabolic network model does not account for all cofactor

utilising enzymes in the cell. To investigate the role of the futile cycle in a cofactor balanced model, further investigation of this was performed using the FBA model and is described in section 5.3.3 and 5.3.4.

5.3.3 pFBA Simulation with Fermentation Product Constraints

I wanted to compare the flux map obtained by MFA with various FBA simulations, to see if the fluxes predicted by FBA were consistent with the MFA estimations. To simulate growth on the minimal medium used in the MFA experiments, no import of amino acids was allowed into the genome scale metabolic model. The glucose uptake rate and non-growth associated maintenance (NGAM) cost were estimated from the batch fermentation data prior to the addition of isotopic tracers, i.e., from the fermentation start to the seven-hour time point (**Fig. 5.1**). The same method was used as in chapters three and four, i.e., placing glucose uptake and produce secretion constraints on the model, optimising for growth and increasing the maintenance cost in a stepwise manner, until the estimated growth rate matched that seen experimentally. The NGAM was slightly below that estimated in chapter four ($14.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$), of $11.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. A potential explanation for this is reduced cellular stress in minimal media due to lack of butanol formation. Butanol is cytotoxic to *Clostridium* by increasing the fluidity of the membrane (Vollherbst Schneck, Sands and Montenecourt, 1984). This disrupts the transmembrane proton gradient (Bowles and Ellefson, 1985), and it is likely that the cells would have to expend additional ATP to maintain the same ΔpH , which would reflect the higher NGAM in the fermentations in chapter four (i.e., when in butanol producing culture).

The secretion rates of the fermentation products acetate, butyrate, lactate, and formate were also determined based on the ^{13}C -MFA fermentation data (**Fig. 5.1, Table 5.1**) and applied as constraints to the FBA model. In addition, the FBA model was extended to include the glyoxylate shunt (reactions shown in **Table 5.2**). As the ED pathway was not predicted to carry flux in the MFA flux map, and there is a lack of annotated glucose-6-phosphate dehydrogenase in the *C. saccharoperbutylacetonicum* N1-4 genome, this was not added to the FBA model.

Metabolic Process	Reaction ID	Constrained Flux
Uptake of glucose	ex_glc_D_e	-8.8182
Secretion of acetate	ACex	4.9068
Secretion of butyrate	BUTex	3.6779
Secretion of acetone	ACETONEex	0
Secretion of formate	FORex	0.8876
Secretion of lactate	LAC2ex	0.4793
Rate of biomass synthesis	biomass	0.1485

Table 5.1 – Constraints used for the ‘externally constrained’ simulation. Reaction IDs correspond to the following reactions: ex_glc_D_e - glucose uptake, ACex - acetate secretion, BUTex - butyrate secretion, ACETONEex - acetone secretion, FORex - formate secretion, LAC2ex - L-lactate secretion, biomass - growth rate. Units are given as $\text{mmol gDW}^{-1} \text{h}^{-1}$, except for growth rate (biomass synthesis), which is h^{-1} .

Reaction ID	Reaction Name	Reaction Equation
ICL	Isocitrate lyase	icit_c --> glx_c + succ_c
MALS	Malate synthase	accoa_c + glx_c + h2o_c --> coa_c + h_c + mal_L_c

Table 5.2 – Glyoxylate shunt reactions added to the FBA model used in this chapter. The metabolite identifiers in the reaction equation correspond to the following compounds: icit_c – isocitrate, glx_c – glyoxylate, succ_c – succinate, accoa_c – acetyl-CoA, h2o_c – water, h_c – proton, mal_L_c – L-malate.

The first simulation I conducted was a pFBA simulation constrained with the glucose uptake rate, along with the experimentally determined biomass synthesis and fermentation product constraints, along with the NGAM. The predicted flux map with the constraints listed in **Table 5.1** is shown in **Fig. 5.7**. Fluxes through glycolysis and the PPP were in good agreement with the ^{13}C -MFA estimations, however the acetyl-CoA forming flux was lower in the pFBA simulation compared with the ^{13}C -MFA estimates. Pyruvate: ferredoxin oxidoreductase was predicted as the major route of acetyl-CoA formation, with minor flux through pyruvate formate lyase. This was forced by the constraint on formate secretion. The flux entering the ABE fermentation pathway, unsurprisingly, showed good agreement as the fluxes were constrained to estimated secretion fluxes based on HPLC data in both simulations.

The TCA cycle fluxes, however, differed from the ^{13}C -MFA estimation. The pyruvate carboxylase flux was under a tenth of that estimated by ^{13}C -MFA. Oxaloacetate was converted via citrate to isocitrate, and there was flux through the glyoxylate shunt. Malate was converted to oxaloacetate by malate dehydrogenase. There was also flux from isocitrate to α -ketoglutarate, which participated in glutamate synthesis, but there was also flux through oxaloacetate oxidoreductase. Succinyl-CoA was partially used in biosynthesis, but was mostly converted to succinate, which was secreted from the cell. There was no predicted flux through malic-enzyme, and the futile cycle found in the ^{13}C -MFA estimation was inactive.

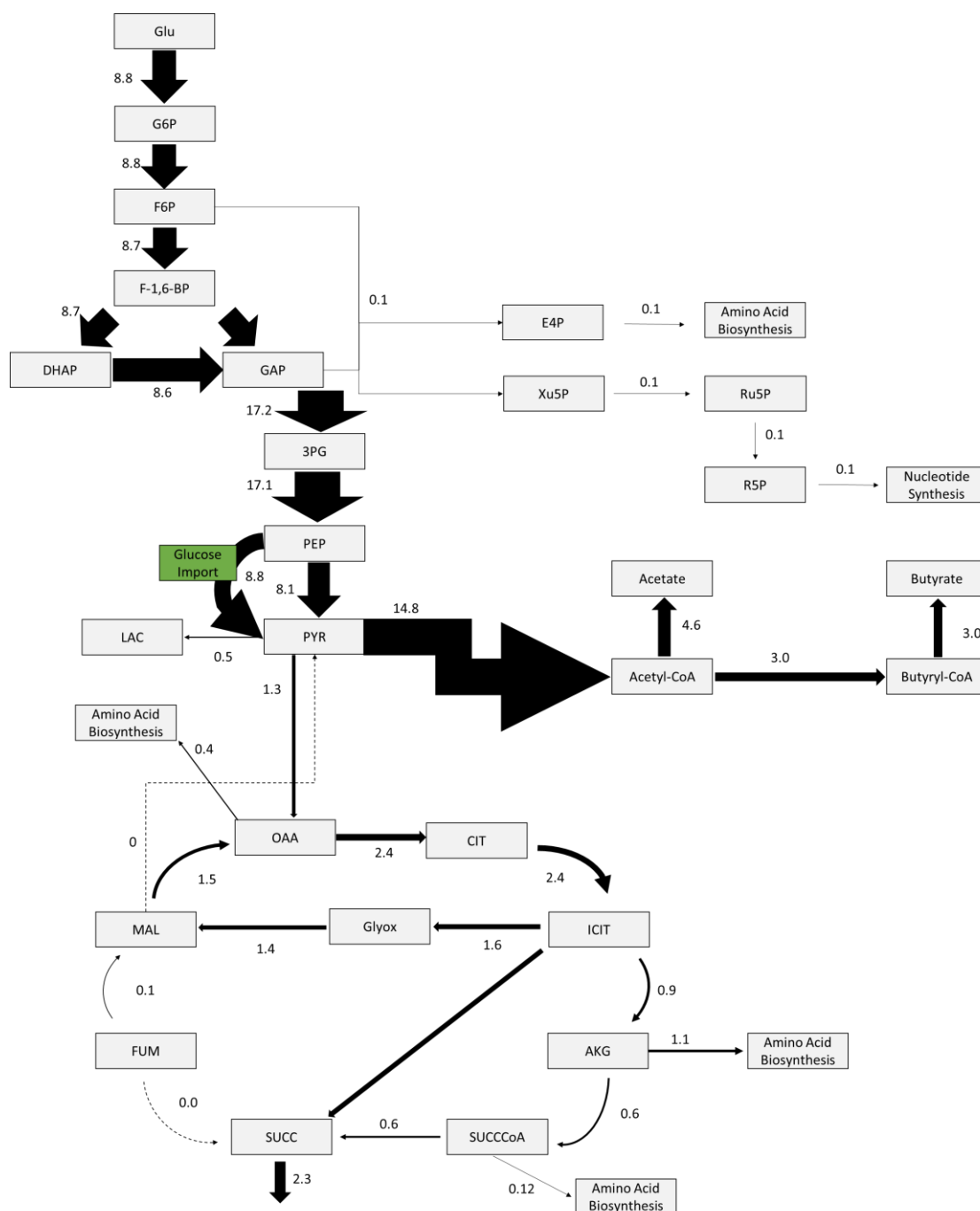


Figure 5.7 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constraints on the rate of biomass formation, glucose uptake and secretion rates of acetate, butyrate, lactate, and formate. Constraints are listed in Table 5.1. Abbreviations of metabolites are given in legend of Figure 5.4. Arrow width corresponds to the flux through the reaction. Dashed lines indicate a flux < 0.1 mmol gDW⁻¹ h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

To investigate whether ATP is in excess in the pFBA simulation and, if so whether an alternative route to ATP disposal was being used, the ATP budget of this simulation was calculated (**Fig. 5.8a**). These fluxes were compared to an ATP budget in the ^{13}C -MFA flux estimation (**Fig. 5.8b**). The ATP producing fluxes were similar in the two flux maps, both in terms of absolute ATP forming flux and the respective reaction contributions. Glycolysis totalled around 78 % of ATP production in both flux maps, with the remainder being produced through acetate and butyrate synthesis. The ATP consumption fluxes differed. In the ^{13}C -MFA flux estimation, the majority of ATP (57 %) was consumed by pyruvate carboxylase. This was not the case in the pFBA simulation, where the NGAM was the largest ATP consuming flux, accounting for 34 % of total ATP consumption. This was followed by phosphofructokinase, which consumed 26 % of total ATP, which was similar to the ^{13}C -MFA estimation (27 % of total ATP). Synthesis of biomass and protein precursors in the pFBA simulation consumed around 27 % of total ATP. In the ^{13}C -MFA flux estimation, biomass synthesis consumed 12 % of total ATP, with amino acid synthesis consuming most of the remaining ATP. In the pFBA simulation, pyruvate carboxylase flux was much lower than that predicted in the ^{13}C -MFA flux estimation. The role of pyruvate carboxylase in the pFBA simulation appeared to be ATP consumption, as most of the produced oxaloacetate was converted to succinate and secreted from the cell. Another ATP consuming reaction in the pFBA simulation was the glutamine synthase reaction. This appeared to serve no purpose aside from consuming ATP, as the produced glutamine was converted back to glutamate. This suggests that ATP generation was in excess to that required for biomass synthesis and maintenance, with the constraints on glucose uptake and organic acid production. Given the technical issues presented here, it may be that the growth rate and/or maintenance costs are being underestimated, or that the external fluxes estimated were unreliable. This would lead the pFBA simulation to use such ATP dissipating reactions as the glutamine synthase reaction.

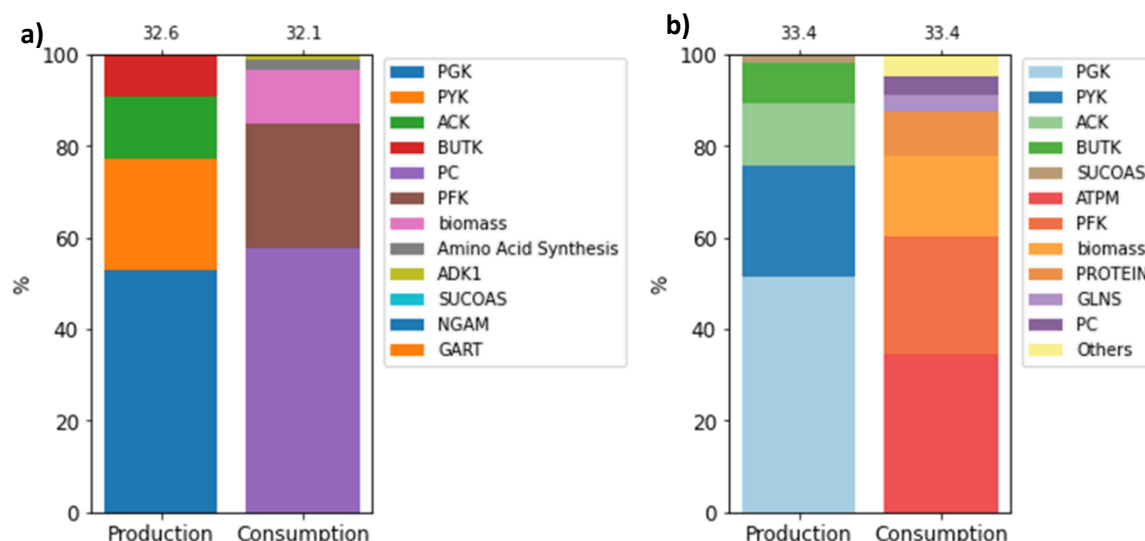


Figure 5.8 – Predicted fluxes responsible for ATP generation and consumption in the (a) ¹³C-MFA flux estimation and (b) the pFBA simulation with constraints on glucose uptake, ABE fermentation products and NGAM. The value above the bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Production and consumption fluxes may not be equal due to rounding. Legend refers to the identifier of the reaction; PGK – phosphoglycerate kinase, PYK – pyruvate kinase, ACK – acetate kinase, BUTK – butyrate kinase, PC – pyruvate carboxylase, PFK – phosphofructokinase, biomass – lumped biomass synthesis equations, Amino Acid Synthesis – sum of ATP requirement for synthesis of all amino acids, ADK1 – ATP: AMP phosphotransferase, SUCOAS – succinyl-CoA synthetase, NGAM /ATPM– ATP hydrolysis reaction (non-growth associated maintenance), GART - glycineamide ribonucleotide transformylase, PROTEIN – lumped protein synthesis reaction, GLNS - L-Glutamate: ammonia ligase. Others represent remaining reactions whose individual fluxes contribute less than 1 % of the total ATP generation/consumption flux.

In both flux maps, ATP was in excess to the biomass synthesis requirements. The ¹³C-MFA estimation consumed the additional ATP through the pyruvate carboxylase reaction, while in the pFBA simulation the ATP was disposed of through both pyruvate carboxylase and glutamine synthase.

Disposal of ATP was one role of the futile cycle in the ¹³C-MFA flux map. Additional potential roles were the synthesis of NADPH and/or the oxidation of NADH. These cofactors were not balanced in the MFA model, so I wanted to compare budgets of these compounds in the pFBA simulation with the equivalent reactions in the ¹³C-MFA model.

The NADH budget in the pFBA simulation with fermentation product constraints is shown in **Fig. 5.9**. The glyceraldehyde-3-phosphate dehydrogenase flux in the pFBA simulation was identical to that in the ^{13}C -MFA flux estimation. The malate dehydrogenase flux was vastly different in the two flux estimation procedures. The pFBA simulation predicted a flux of only 1.5 mmol gDW $^{-1}$ h $^{-1}$, while the ^{13}C -MFA estimation had a flux of -13.8 mmol gDW $^{-1}$ h $^{-1}$, i.e., running in the oxaloacetate reducing direction rather than the malate oxidising direction. The only NADH producing reaction in the ^{13}C -MFA estimation was glyceraldehyde-3-phosphate dehydrogenase. This was also the major NADH producing reaction in the pFBA simulation, but malate dehydrogenase produced additional NADH. In both simulations, the ferredoxin: NAD(H) oxidoreductase reaction oxidised NADH, although in the pFBA simulation this reaction was a greater contributor to NADH oxidation (oxidising 48 % of total NADH versus 38 % of total NADH). In the pFBA simulation, the synthesis of butyryl-CoA consumed a total of 47 % of total NADH. In the ^{13}C -MFA estimation, butyryl-CoA synthesis consumed 34 % of total NADH.

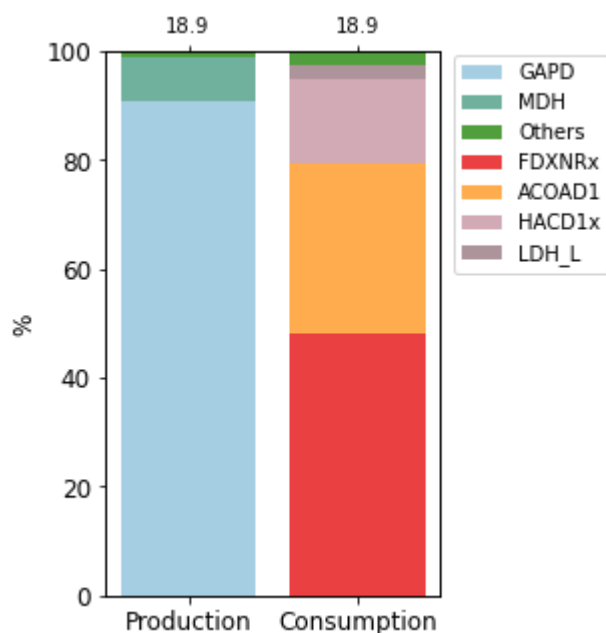


Figure 5.9 – Predicted fluxes responsible for NADH generation and consumption in pFBA simulations of a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constraints on the rates of biomass formation, glucose uptake and secretion rates of acetate, butyrate, lactate, and formate, and NGAM cost. Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW $^{-1}$ h $^{-1}$. Reaction identifiers in figure legend correspond to the following reactions: GAPD – glyceraldehyde-3-phosphate dehydrogenase, MDH – malate dehydrogenase, FDXNRx – ferredoxin: NAD(H) oxidoreductase, ACOAD1 – butyryl-CoA dehydrogenase, HACD1x – 3-hydroxybutyrate dehydrogenase, LDH_L – lactate dehydrogenase. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

I next compared the NADPH budgets in the pFBA simulation (**Fig. 5.10**). In the pFBA simulation the majority of NADPH was produced through the ferredoxin: NADP(H) oxidoreductase reaction, with minor production through prephenate dehydrogenase and dihydrothymine dehydrogenase. NADPH was consumed through the synthesis of various biomass precursors, both amino acids and fatty acids. In the ^{13}C -MFA flux estimation, ferredoxin: NADP(H) forming flux was negligible, and the NADP-dependent malic enzyme effectively produced the entirety of cellular NADPH. The total NADPH forming flux was nearly 26-fold higher in the ^{13}C -MFA flux estimation. In terms of NADPH consumption, the ^{13}C -MFA estimated that the major NADPH consuming flux was 3-hydroxybutyryl-CoA dehydrogenase (NADPH dependent), followed by amino acid biosynthesis reactions.

The individual reactions in the pFBA reaction correspond to lumped reactions in the ^{13}C -

MFA model. The aspartate semialdehyde oxidoreductase (reaction in L-aspartate synthesis pathway) flux was similar in the pFBA simulation and the ^{13}C -MFA simulation ($0.2 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ and $0.17 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, respectively), as was the acetolactate dehydrogenase (reaction in valine and leucine biosynthesis pathways) flux ($0.11 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ and $0.09 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, respectively). The ketol-acid reductoisomerase (reaction in isoleucine biosynthesis pathway) flux was double that in the ^{13}C -MFA estimation. There was low ($0.04 \text{ mmol gDW}^{-1} \text{ h}^{-1}$) flux predicted through 5,10-methylenetetrahydrofolate: NADP⁺ dehydrogenase, which was not active in the ^{13}C -MFA estimation. Instead, methylenetetrahydrofolate was synthesised using NADH as a cofactor. Glutamate-semialdehyde oxidoreductase (reaction in proline synthesis pathway) flux was similar in both the pFBA simulation and ^{13}C -MFA estimation ($0.02 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ and $0.03 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, respectively). The flux through N-Acetyl-L-glutamate-5-semialdehyde: NADP⁺ 5-oxidoreductase (reaction in arginine synthesis pathway) was also similar in both pFBA and ^{13}C -MFA ($0.02 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ and $0.01 \text{ mmol gDW}^{-1} \text{ h}^{-1}$). For the majority of these reactions, the fluxes were similar in both ^{13}C -MFA and pFBA, and the flux through these reactions tended to be low. The total NADPH oxidising flux required for amino acid biosynthesis was $1.5 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, somewhat higher than that estimated for the pFBA simulation. This was due to small differences in the biomass composition between the two models, which were not updated from their respective original publications.

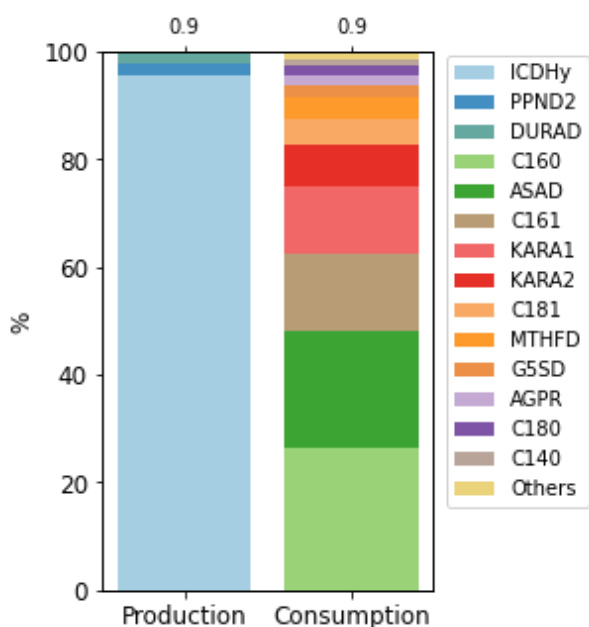


Figure 5.10 – Predicted fluxes responsible for NADPH generation and consumption in pFBA simulations of a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constraints on the rates of biomass formation, glucose uptake and secretion rates of acetate, butyrate, lactate, and formate, and NGAM cost. Value above bar is the total flux through all reactions shown in the bar graph, in $\text{mmol gDW}^{-1} \text{ h}^{-1}$. Reaction identifiers in figure legend correspond to the following reactions: ICDHy - isocitrate dehydrogenase, PPND2 - prephenate dehydrogenase, DURAD - dihydrothymine dehydrogenase, ASAD - aspartate semialdehyde oxidoreductase, KARA1 - acetolactate dehydrogenase, KARA2 - ketol-acid reductoisomerase, MTHFD - methylenetetrahydrofolate dehydrogenase, G5SD - glutamate-semialdehyde oxidoreductase, AGPR - N-Acetyl-L-glutamate-5-semialdehyde:NADP⁺ 5-oxidoreductase, C160/C161/C181 - fatty acid synthesis. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

The slightly higher NADPH requirement for biomass synthesis in the ^{13}C -MFA model was not nearly high enough to explain the high malic enzyme flux. The consumption of NADH was another possible reason for the futile cycle. In the pFBA simulation, glycolysis provided sufficient NADH to form

butyryl-CoA, which was oxidised mainly through ferredoxin. The remaining NADH was used to synthesise the fermentation products butyrate and lactate.

These simulations have shown that in both the ^{13}C -MFA flux map and the pFBA flux map constrained with fermentation product secretion rates, the ATP produced through glycolysis and acids was in excess of biomass synthesis requirements. This could result in activation of a futile cycle to consume the additional ATP. The oxidation of NADH was another potential role of the futile cycle. This was achieved in the pFBA simulation mainly through ferredoxin: NAD(H) oxidoreductase. There may be capacity limitations *in vivo* which prevent ferredoxin: NAD(H) oxidoreductase from carrying the flux predicted by pFBA, leading to the use of the futile cycle. The final potential role was the production of NADPH. This did not seem to be the main role, as the NADPH requirements for biomass synthesis were met through the isocitrate dehydrogenase reaction in the pFBA simulation. To further investigate the futile cycle, I performed additional pFBA simulations with the futile cycle imposed as a constraint.

5.3.4 pFBA Simulations with Futile Cycle Constraints

I imposed the constraints on the futile cycle reactions shown in **Table 5.3** and ran a pFBA simulation, in order to deduce the role of the futile cycle. No NGAM was imposed on this simulation, as the futile cycle was potentially contributing to this metric.

Reaction ID	Lower Bound	Upper Bound
ex_glc_D_e	-8.8182	-8.8182
MDH	0.447	16.4339
PC	4.5998	21.3832
ME2	8.6322	25.5441

Table 5.3 – Constraints used for the ‘futile cycle’ simulation. Reaction IDs correspond to the following reactions: ex_glc_D_e: glucose uptake, MDH: malate dehydrogenase, PC: pyruvate carboxylase, ME2: NADP-dependent malic enzyme. Units are given as mmol gDW⁻¹ h⁻¹.

The central metabolic fluxes predicted using pFBA, with biomass formation and futile cycle constraints are shown in **Fig. 5.11**. This simulation is referred to as the ‘futile cycle’ simulation for the remainder of the chapter. In this simulation, flux through the glycolysis was similar to that seen in the simulation constrained with rates of fermentation product formation. The PPP flux was also similar. Pyruvate was predicted to be formed mostly through PFL, with a roughly 3:1 ratio of pyruvate formate lyase to pyruvate ferredoxin oxidoreductase flux. Lactate was predicted to be the major fermentation product, followed by acetate and butanol.

The constraint on pyruvate carboxylase forced high flux from pyruvate to oxaloacetate. The oxaloacetate formed was partially used in amino acid biosynthesis, but the majority proceeded

through malate dehydrogenase. There was no predicted flux from oxaloacetate to citrate or citrate to isocitrate. Isocitrate was instead formed through the reverse TCA cycle. Succinate generated from isocitrate lyase was converted via succinyl-CoA synthetase, oxoglutarate oxidoreductase and isocitrate dehydrogenase back to isocitrate. There was predicted flux through the glyoxylate shunt, which generated malate. Malate was cycled back to pyruvate through NADP-dependent malic enzyme. In terms of secreted products, formate had the highest secretion flux ($8.7 \text{ mmol gDW}^{-1} \text{ h}^{-1}$) as pyruvate formate lyase formed the majority of acetyl-CoA. This was followed by lactate, with acetate secretion slightly below this, and butanol formation was predicted, which was not seen in the ^{13}C -MFA batch fermentations. This suggested that excess NAD(P)H was being formed.

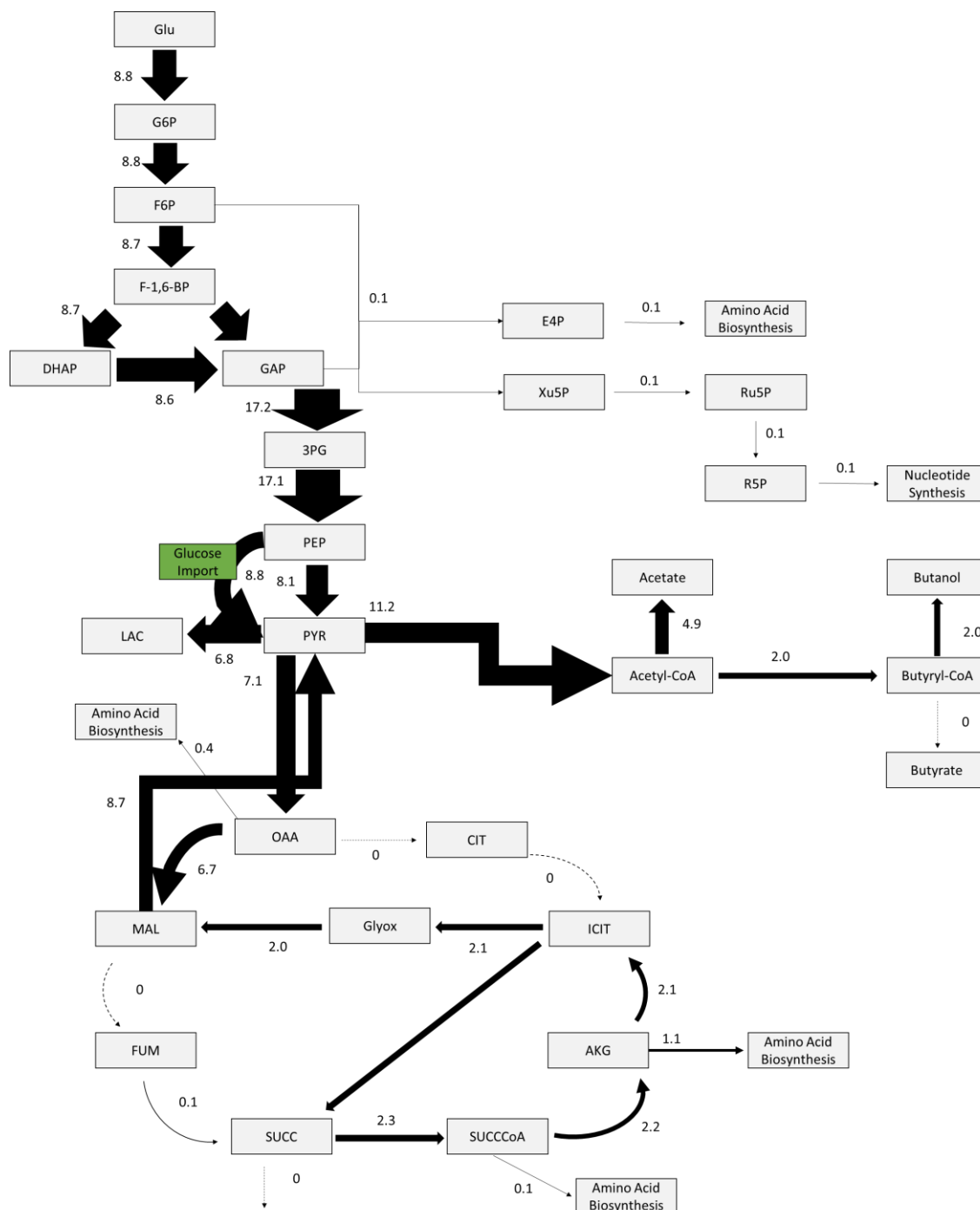


Figure 5.11 – Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constraints on the rate of biomass formation, glucose uptake and flux through MFA estimated futile cycle of pyruvate carboxylase, malate dehydrogenase, and NADP-dependent malic enzyme. Constraints are listed in Table 5.3. Abbreviations are given in legend of Figure 5.2. Arrow width corresponds to the flux through the reaction. Dashed lines indicate a flux < 0.1 mmol gDW⁻¹ h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

The ATP budget of the futile cycle simulation is shown in **Fig. 5.12**. As in the ^{13}C -MFA flux estimation glycolysis produced the majority of ATP, with acetate kinase producing the remainder. There was no predicted butyrate synthesis in the futile cycle simulation, however. The ATP consuming fluxes were effectively identical to the simulation constrained with fermentation products. Pyruvate carboxylase consumed less ATP than the NGAM, but unlike the simulation with constraints on the rate of fermentation product secretion, the succinyl-CoA synthetase reaction ran in the reverse (ATP consuming) direction. Compared with the ^{13}C -MFA estimations, the pFBA simulation did not consume as much ATP through pyruvate carboxylase (only 38 % flux compared to the ^{13}C -MFA estimation). Meanwhile, synthesis of biomass and protein precursors consumed a combined 30 % of total ATP in the pFBA simulation, while in the ^{13}C -MFA estimation biomass synthesis only consumed 12 % of total ATP. In the futile cycle pFBA simulation, ATP was also consumed through succinyl-CoA synthetase and glutamine synthase. The succinyl-CoA synthetase flux was relatively much higher in the pFBA simulation than in the ^{13}C -MFA estimation ($2.3 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ compared with $0.1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$). As with the pFBA simulation with fermentation product constraints, the glutamine synthase reaction only appeared to hydrolyse ATP, as the majority (87 %) of produced glutamine was converted back to glutamate.

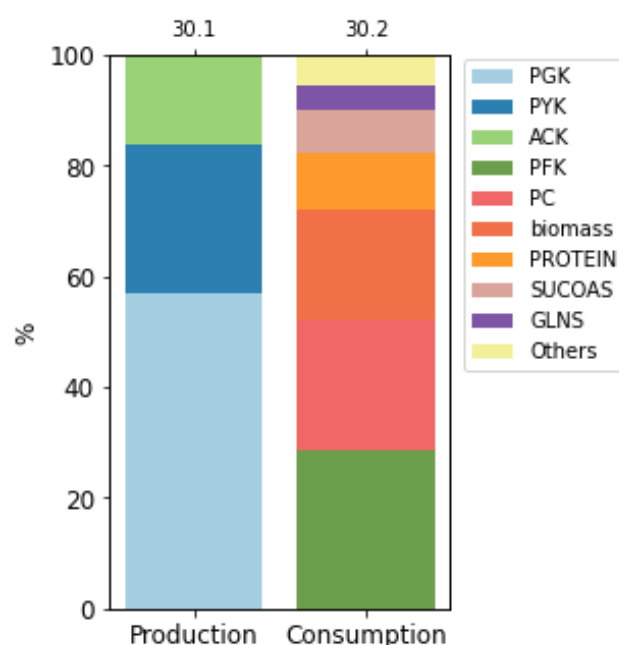


Figure 5.12 – Predicted fluxes responsible for ATP generation and consumption in pFBA simulations of a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constraints on the rate of biomass formation, glucose uptake and flux through MFA estimated futile cycle of pyruvate carboxylase, malate dehydrogenase, and NADP-dependent malic enzyme. Value above bar is the total flux through all reactions shown in the bar graph, in $\text{mmol gDW}^{-1} \text{ h}^{-1}$. Reaction identifiers in figure legend correspond to the following reactions: PGK - phosphoglycerate kinase, PYK - pyruvate kinase, ACK - acetate kinase, PFK - phosphofructokinase, PC - pyruvate carboxylase, biomass - lumped biomass synthesis reaction, PROTEIN - lumped protein synthesis reaction, SUCOAS - succinyl-CoA synthetase, GLNS - glutamine synthase. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

I then examined the NADPH forming budget (**Fig. 5.13**). NADP-dependent malic enzyme was the only source of NADPH, as was the case in the ^{13}C -MFA estimation. The high rate of NADPH formation resulted in the production of butanol. During the production of both 3-hydroxybutyryl-CoA dehydrogenase and butanol, NADPH was used as a cofactor in the pFBA simulation. Isocitrate dehydrogenase also consumed approximately 24 % of total NADPH. Flux through this reaction was much higher than in the ^{13}C -MFA estimation ($2.1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ versus $0.1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$).

Another reaction with substantial NADPH consumption was glutamate synthase, which consumed around 13 % of total NADPH. This flux was slightly higher than in the ^{13}C -MFA flux estimation (pFBA flux was $1.1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, ^{13}C -MFA flux estimation $0.9 \text{ mmol gDW}^{-1} \text{ h}^{-1}$). The flux through these two reactions appeared to be to balance the additional NADPH formed as a result of the imposed futile cycle. As in the ^{13}C -MFA estimation, the remaining NADPH was used to synthesise amino and fatty acids.

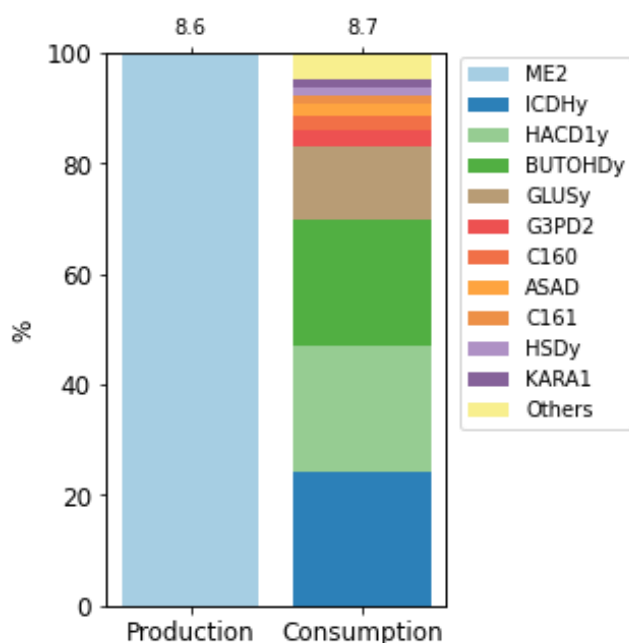


Figure 5.13 – Predicted fluxes responsible for NADPH generation and consumption in pFBA simulations of a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constraints on the rate of biomass formation, glucose uptake and flux through MFA estimated futile cycle of pyruvate carboxylase, malate dehydrogenase, and NADP-dependent malic enzyme. Value above bar is the total flux through all reactions shown in the bar graph, in $\text{mmol gDW}^{-1} \text{ h}^{-1}$. Values may not be equal due to rounding. Reaction identifiers in figure legend correspond to the following reactions: ME2 - NADP-dependent malic enzyme, ICDH_y - isocitrate dehydrogenase, HACD1_y - 3-hydroxybutyryl-CoA dehydrogenase, BUTOHD_y: butanol dehydrogenase, GLUS_y - glutamate synthase, G3PD2: glycerol-phosphate dehydrogenase, ASAD - aspartate semialdehyde oxidoreductase, HSD_y: homoserine oxidoreductase, KARA1 - acetolactate dehydrogenase, C160/C161 - fatty acid synthesis. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

The NADH budget in the futile cycle simulation is shown in **Fig. 5.14**. Glyceraldehyde-3-phosphate dehydrogenase generated the majority of NADH, with some formation through ferredoxin:NAD(H). Lactate dehydrogenase was the largest consumer of NADH, followed by malate dehydrogenase. Combined, these two reactions consumed around 70 % of total NADH. The remaining 30 % was consumed through butyryl-CoA dehydrogenase and butyraldehyde dehydrogenase. The lactate flux was much higher than in the ^{13}C -MFA estimation, while the butyryl-CoA forming flux was nearly identical. In addition, the ferredoxin: NAD(H) oxidoreductase reaction ran in the opposite direction in the ^{13}C -MFA estimation. The reason for the additional NADH formation was likely to produce butyraldehyde, as this step of butanol synthesis cannot use NADPH as a cofactor. This allowed the model to balance NADPH through the formation of butanol and 3-hydroxybutyryl-CoA, as discussed when I examined the NADPH budget.

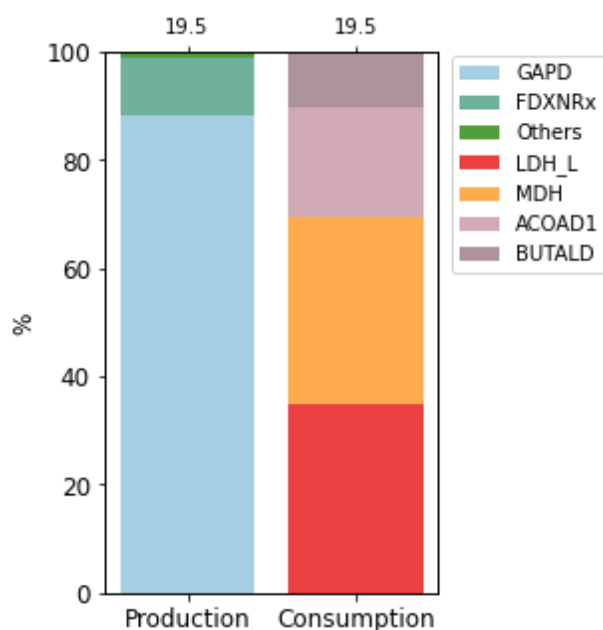


Figure 5.14 – Predicted fluxes responsible for NADH generation and consumption in pFBA simulations of a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constraints on the rate of biomass formation, glucose uptake and flux through MFA estimated futile cycle of pyruvate carboxylase, malate dehydrogenase, and NADP-dependent malic enzyme. Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Reaction identifiers in figure legend correspond to the following reactions: GAPD - glyceraldehyde-3-phosphate dehydrogenase, FDXNRx - ferredoxin: NAD(H) oxidoreductase, LDH_L - lactate dehydrogenase, MDH - malate dehydrogenase, ACOAD1 - butyryl-CoA dehydrogenase, BUTALD - butyraldehyde dehydrogenase. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

This simulation has further confirmed that ATP produced through glycolysis and acid formation was in excess to biomass synthesis requirements. In addition, the NADPH was excess to the requirements for biomass synthesis and was mostly used in the synthesis of butanol. The NADH produced in the simulation with the futile cycle constraint resulted in the production of both lactate and butanol, while the malate dehydrogenase constraint consumed a substantial amount of total NADH.

Given that the higher NADPH formation because of the futile cycle constraint resulted in formation of butanol, I wanted to see what would happen if I removed the NADP-dependent malic enzyme and only allowed flux through the NAD-dependent malic enzyme. I removed the NADP-dependent malic enzyme (ME) from the model and replacing the futile constraint through it with the same constraint through NAD-dependent malic enzyme. This gave the predicted flux distribution shown in **Fig. 5.15**. The central metabolic fluxes in this simulation were nearly identical to those with the NADP-utilising malic enzyme (ME) futile cycle simulation, except the butanol forming flux in the NADP-dependent ME knockout simulation was reduced by 25 %, and ethanol formation was predicted. In contrast to the futile cycle simulation with NADP-dependent malic enzyme active, there was very low flux through the glyoxylate shunt or the TCA cycle, aside from the constraints on the futile cycle. The pyruvate carboxylase flux was also higher when NADP-dependent malic enzyme was removed.

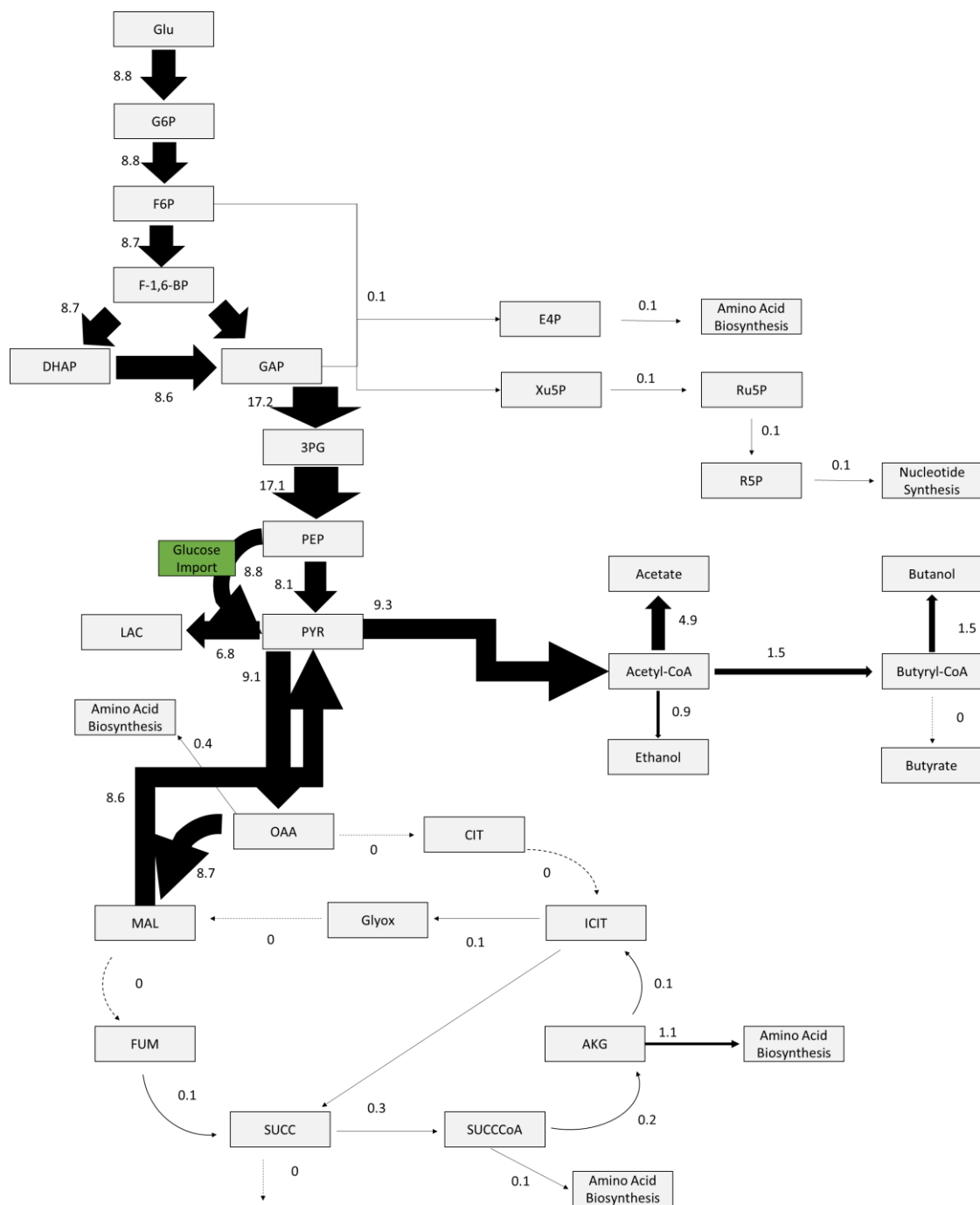


Figure 5.15 - Central metabolic fluxes predicted by pFBA using a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constrained flux through MFA estimated futile cycle of pyruvate carboxylase, malate dehydrogenase, and NAD-dependent malic enzyme. Constraints are listed in table 5.3, except the ME2 reaction was removed from the model and the constraint was applied to the ME1 reaction. Abbreviations are given in legend of Figure 5.4. Arrow width corresponds to the flux through the reaction. Dashed lines indicate a flux < 0.1 mmol gDW⁻¹ h⁻¹. In the interest of brevity, not all reactions involving a given metabolite are presented, as such, there may be small discrepancies between flux balances.

I examined the ATP, NADPH and NADH usage in the NADP-dependent ME knockout simulation. The total ATP forming flux, and the formation fluxes were identical to the NADP-dependent ME futile cycle simulation (**Fig. 5.16**). The ATP consumption in the NADP-dependent ME knockout had a slight shift, with the increased pyruvate carboxylase flux accounting for a higher proportion of ATP consumption (30 % when NAD-dependent ME was used vs 24 % when NADP-dependent ME was used). This was still far below that estimated by the ^{13}C -MFA, however (57 %). The other difference between the two pFBA simulations was the reduced succinyl-CoA synthetase flux in the NADP-dependent malic enzyme knockout, which consumed a very small (< 1 %) of total ATP consuming flux. All the other major ATP consuming fluxes were identical in both simulations.

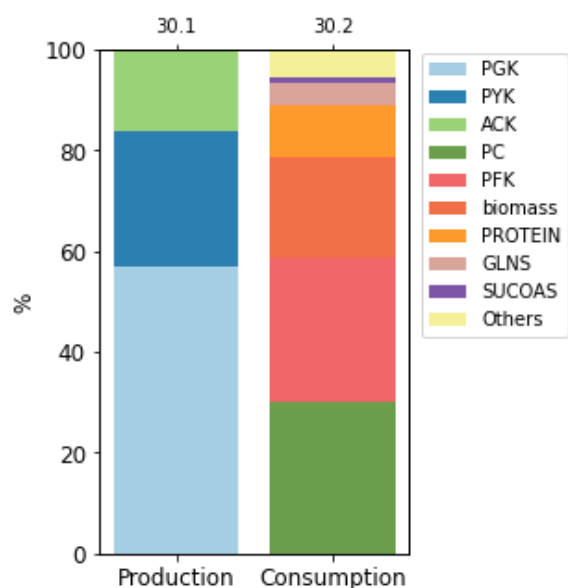


Figure 5.16 – ATP usage summary in pFBA simulations of a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constrained flux through MFA estimated futile cycle of pyruvate carboxylase, malate dehydrogenase, and NAD-dependent malic enzyme. Value above bar is the total flux through all reactions shown in the bar graph, in $\text{mmol gDW}^{-1} \text{h}^{-1}$. Reaction identifiers in figure legend correspond to the following reactions: PGK - phosphoglycerate kinase, PYK - pyruvate kinase, ACK - acetate kinase, PC - pyruvate carboxylase, PFK - phosphofructokinase, biomass - lumped biomass synthesis reaction, PROTEIN - lumped protein synthesis reaction, GLNS - glutamine synthase. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

The total NADPH forming flux in the NADP-dependent ME knockout simulation was much lower than the NADP-dependent ME utilising futile cycle simulation and was effectively identical to the simulation with fermentation product constraints (**Fig. 5.17**). Again, the majority of NADPH was formed through ferredoxin: NADP(H) oxidoreductase, and this was utilised in the synthesis of fatty acids and amino acids.

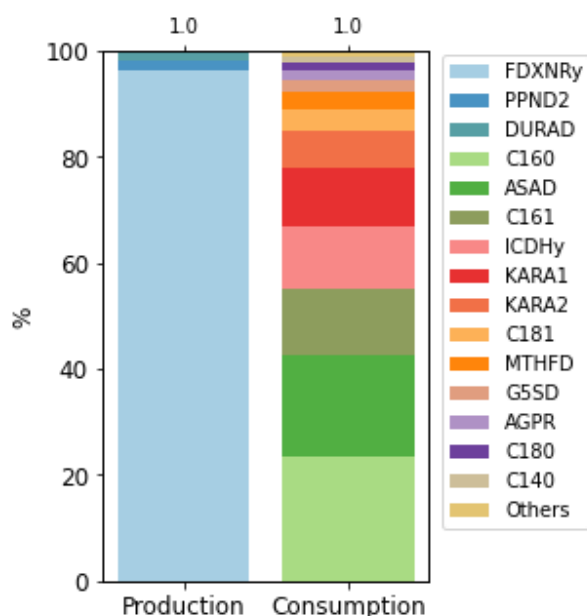


Figure 5.17 – NADPH usage summary in pFBA simulations of a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constrained flux through MFA estimated futile cycle of pyruvate carboxylase, malate dehydrogenase, and NAD-dependent malic enzyme. Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Reaction identifiers in figure legend correspond to the following reactions: FDXNRy - ferredoxin:NADP oxidoreductase, PPND2 - prephenate dehydrogenase, DURAD - dihydrothymine dehydrogenase, ASAD - aspartate semialdehyde oxidoreductase, KARA1 - acetolactate dehydrogenase, ICDHy - isocitrate dehydrogenase, KARA2 - ketol-acid reductoisomerase, MTHFD - methylenetetrahydrofolate dehydrogenase, G5SD - glutamate-semialdehyde oxidoreductase, AGPR - N-acetyl-glutamate semialdehyde oxidoreductase, C140/C160/C161/C180/C181 - fatty acid synthesis. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

The NADH forming flux was 33 % higher in the NADP-dependent ME knockout simulation compared to the futile cycle simulation using NAD-dependent ME (**Fig. 5.18**). In absolute terms, the flux through glyceraldehyde-3-phosphate dehydrogenase was identical with NADP-dependent ME active and knocked out simulations, but in the latter, there was also flux through the NAD-dependent malic enzyme. Malate dehydrogenase was the largest consumer of NADH in the NADP-dependent ME knockout, consuming 33 % of total NADH, followed by lactate dehydrogenase (26 %). The combined formation of ethanol and butanol consumed 48 % of total NADH. The latter was synthesised solely using NADH, including ABE fermentation pathway intermediates.

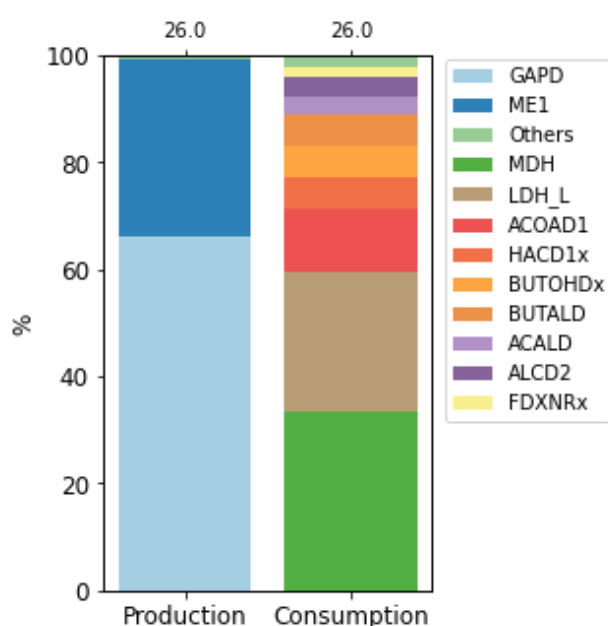


Figure 5.18 – NADH usage summary in pFBA simulations of a genome scale metabolic model of *C. saccharoperbutylacetonicum* N1-4 with constrained flux through MFA estimated futile cycle of pyruvate carboxylase, malate dehydrogenase, and NAD-dependent malic enzyme. Value above bar is the total flux through all reactions shown in the bar graph, in mmol gDW⁻¹ h⁻¹. Reaction identifiers in figure legend correspond to the following reactions: GAPD - glyceraldehyde-3-phosphate dehydrogenase, ME1 - NAD-dependent malic enzyme, MDH - malate dehydrogenase, LDH_L - lactate dehydrogenase, ACOAD1 - butyryl-CoA dehydrogenase, HACD1x - 3-hydroxybutyryl-CoA dehydrogenase, BUTOHDx - butanol dehydrogenase, BUTALD - butyraldehyde dehydrogenase, ACALD - acetaldehyde dehydrogenase, ALCD2 - alcohol dehydrogenase, FDXNRx - ferredoxin:NAD(H) oxidoreductase. Others represent remaining reactions which individually contribute less than 1 % to total ATP generation/consumption.

From these simulations, the following has become clear. The ATP produced by glycolysis and ABE fermentation (acetate and butyrate synthesis) is in excess to the requirements for biomass synthesis. The NADPH requirement for biomass synthesis was low in both the ^{13}C -MFA model and the pFBA model, which was met through ferredoxin: NADP(H) oxidoreductase in the latter. The flux through the futile cycle in the ^{13}C -MFA flux map produced far more NADPH than required for biomass synthesis. The pFBA simulation disposed of this excess NADPH through butanol production, glutamine synthase and isocitrate dehydrogenase. Finally, the NADH oxidation role of the futile cycle could be achieved through the synthesis of lactate, solvents, or ferredoxin: NAD(H) oxidoreductase. The cofactor use for the synthesis of butyryl-CoA and butanol was flexible in the pFBA model, so depending on the cofactor preferences *in vivo*, the production of additional NADPH might be beneficial when producing solvents. The primary role appeared to be the dissipation of ATP, with potential cofactor generating/consuming roles needing further investigation.

5.4 Discussion

5.4.1 Metabolic Flux Map of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}

In this chapter I generated a flux map of the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} from steady-state ^{13}C isotope labelling during acidogenic culture. I found that glucose was metabolised through Embden-Meyerhof-Parnas (EMP) glycolysis, without Entner-Doudoroff pathway flux. This was also found to be the case in *C. acetobutylicum* (Au *et al.*, 2014). I identified a futile cycle that was not reported by Au *et al.*, and the TCA cycle fluxes in my flux map are in general disagreement with their work. Au *et al.* found no flux from fumarate to succinate, α -ketoglutarate to succinyl-CoA or from malate to oxaloacetate. I also found no flux from α -ketoglutarate to succinyl-CoA, however my ^{13}C -MFA flux map showed high flux from oxaloacetate to malate through the futile cycle and from succinate to fumarate. I also identified an active glyoxylate shunt, which was not reported by (Au *et al.*, 2014).

The high pyruvate carboxylase flux I identified was not in agreement with Au *et al.*, who found low (around 6 %) of glucose uptake flux entering the TCA cycle through pyruvate carboxylase. Higher pyruvate carboxylase flux was reported by (Amador-Noguez *et al.*, 2011), who used kinetic flux profiling during the shift from acidogenesis to solventogenesis in *C. acetobutylicum*. They found that pyruvate carboxylase flux reduced by $\text{Log}_2(-4)$ fold in solventogenesis, however the oxaloacetate formed was mainly converted to aspartic acid, without the futile cycle I found in my flux map (Amador-Noguez *et al.*, 2011). In addition, the absolute flux through pyruvate carboxylase was much higher than reported in their study. How much of this is due to differences between the species *C. acetobutylicum* and *C. saccharoperbutylacetonicum* N1-4, and how much is due to the difference in the asporogenic strain I used, is unclear. Future efforts using ^{13}C -MFA to understand

C. saccharoperbutylacetonicum N1-4 could first generate a map of the wild type strain to elucidate this. For the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, it will be important to determine the physiological cofactor dependence of the TCA cycle enzymes, and update the model based on these findings. Labelling with glutamine tracers has been reported to offer good characterisation of net TCA cycle fluxes, specifically with [1,2]Gln, [3,4]Gln, and [U]Gln (Metallo, Walther and Stephanopoulos, 2009). It may be worth using these tracers in combination with glucose tracers, however feeding glutamine would lead to a change in cellular metabolism, so fluxes determined with these tracers may not reflect fluxes in standard fermentation conditions. In addition, while the model provided a statistically acceptable fit to the data, it might be worth experimentally determining the cellular composition of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, so the flux estimations are more representative of cellular conditions *in vivo*.

5.4.2 Role of the Futile Cycle

The futile cycle predicted by MFA had three potential roles, which I have tried to elucidate using FBA and various constraints from the ¹³C-MFA flux map. Transhydrogenation of NADH and NADP⁺ to NAD⁺ and NADPH does not appear to be the main function. In the simulation constrained to force flux through the futile cycle, the NADPH generated was far above that required to synthesise biomass components and was predicted to lead to butanol production. This was not seen experimentally; however, butyrate formation was predicted, which can utilise NADPH as a cofactor during 3-hydroxybutyryl-CoA synthesis. While the model can freely use NADH or NADPH as a cofactor in this reaction, this has not been shown to be the case *in vivo*. If NADPH is a preferred cofactor, then it might make sense for the bacteria to increase NADPH formation. The *C. saccharoperbutylacetonicum* N1-4 3-hydroxybutyryl-CoA dehydrogenase would need to be purified and its substrate preferences evaluated before any firm judgement can be made on this role of the futile cycle.

The other potential role of the futile cycle was disposing excess ATP. As I found in chapter three that ATP availability limits biomass synthesis, it seems unusual that ATP would be in excess of cellular requirements. If the cells were not actively growing, then it may be possible that ATP would be in excess to cellular requirements. Due to technical issues, it could not be determined whether the cells were actively growing or not, and the growth curve needs repeating to confirm the results I found, and to shed light on the futile cycle role.

The initial MFA model which the model used here was based on was for *C. acetobutylicum* (Au *et al.*, 2014). In this work, the initial model had a similar range through the futile cycle as I found. The upper bound through the pyruvate carboxylase and malic enzyme were 217 mmol gDW⁻¹ h⁻¹, when normalised to a glucose uptake rate of 100 mmol gDW⁻¹ h⁻¹. However, Au *et al.* removed this

reaction from their minimal model. In their minimal model, the ATP hydrolysis reaction had a flux of 205 mmol gDW⁻¹ h⁻¹. This shows that during acidogenesis, the ATP produced is in excess to biomass synthesis requirements. There are numerous other roles of ATP, such as maintaining the transmembrane proton gradient. At pH 6.5 the gradient has been reported at 0.2, while at pH 4.5 the gradient increases to 1.5 (Huang, Gibbins and Forsberg, 1985). Throughout acidogenesis, as the pH drops the ATP cost of maintaining the increased pH gradient increases. This could explain the reason for the differences in ATP requirements for biomass synthesis and the ATP produced through acid synthesis and glycolysis.

The final potential role of the futile cycle was dissipation of NADH. The pFBA simulations with futile cycles predicted the formation of alcohols and high fluxes toward lactate, the formation of all these compounds oxidised NADH. This was not seen experimentally, suggesting that another mechanism of NADH oxidation was being utilised, which could be the futile cycle. This could also be in combination with the ATP hydrolysing role, but, as above, this experiment would need to be repeated first. Further flux analysis work with different tracers could improve the resolution of the TCA cycle, which could elucidate the role of this futile cycle.

This futile cycle has been reported to function in other organisms too, and mainly appears to be a method of NADPH formation. The bacterium *Pseudomonas fluorescens*, for example, utilises the same futile cycle predicted by MFA to generate NADPH used for anti-oxidative defence mechanisms e.g., catalase and glutathione peroxidase, under oxidative stress conditions (Singh *et al.*, 2008). The fungus *Aspergillus nidulans* is another example, which utilises malic enzyme to produce NADPH for lipid synthesis (Wynn and Ratledge, 1997). What could be an interesting application of this futile cycle is utilising it to increase NADPH generation. This was discussed in chapter three using the Entner-Doudoroff pathway, but this futile cycle could provide an alternative route of NADPH formation. Indeed, this system has been used in other organisms for metabolic engineering applications. Overexpressing pyruvate carboxylase in *E. coli* increased the succinate yield while decreasing formate and acetate yields (Gokarn, Eiteman and Altman, 2000). In the yeast, *S. cerevisiae*, this futile cycle has been utilised to increase NADPH generation (Moreira Dos Santos *et al.*, 2004), and has been used to increase isobutanol production in *S. cerevisiae* without increasing ethanol titres (Morita *et al.*, 2017). This futile cycle would be interesting to implement, especially as it may channel flux towards the C4 end of the ABE fermentation pathway by exploiting the fact that butyryl-CoA forming enzymes utilise NADPH as a cofactor. Given the ATP consuming role of the futile cycle, ideally expression would be driven from inducible promoters, to prevent complete consumption of ATP. Another drawback of a futile cycle such as this is a risk of trapping carbon, preventing flux towards the ABE fermentation pathway.

5.4.3 Impact of ^{13}C -MFA Flux Map on the Proposed Metabolic Engineering Strategy

The flux map that I generated in this chapter for *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T}, despite being of acidogenic cells, nevertheless provides additional support for some aspects of the metabolic engineering strategy that I proposed in chapters three and four. The NADH generating strategy developed in chapter three assumed that pyruvate: ferredoxin oxidoreductase was the major acetyl-CoA forming reaction, while pyruvate formate lyase was only a minor contributor. This agrees with the FBA predictions in chapters three and four when constraints were set on the rate of fermentation end products secretion. The role of pyruvate formate lyase in the aforementioned simulations was the generation of one carbon units for the formation of one carbon units for serine biosynthesis, which was also the case in the ^{13}C -MFA flux map. This is also consistent with the hypothesised anabolic role of pyruvate formate lyase in *Clostridium* (Thauer, Kirchniawy and Jungermann, 1972). The low flux through PFL in the ^{13}C -MFA flux map confirms that overexpression of this enzyme will be necessary to drive high formate forming flux, which would then be oxidised by formate dehydrogenase.

Another finding of the ^{13}C -MFA flux map was that ATP was produced in excess of the cellular requirements for biomass synthesis. This led to the use of a futile cycle to balance excess ATP. The formation of ATP in excess to the biomass synthesis requirements is also the case with *C. acetobutylicum* (Au et al., 2014). The excess ATP production shown by the ^{13}C -MFA is in line with my proposed engineering strategy developed in chapter 4 of introducing additional ATP-sink reactions.

My strategy also assumed as increased uptake of glucose could be achieved, as this appeared to be limiting solvent production, which would generate additional ATP. However, it may also be the case that increasing the ATP demand could be sufficient to drive an increase in glucose uptake. Given that there is an ATP surplus in the cell, it may be that certain central metabolic such as phosphofructokinase and pyruvate kinase have reduced activity due to allosteric inhibition by ATP (Chandel, 2021). Relieving such inhibition by depleting ATP may therefore drive an increase in glucose uptake. As discussed in chapter four, this has been implemented in several bacteria (Yokota et al., 1997; Sekine et al., 2001; Koebmann et al., 2002; Hädicke, Bettenbrock and Klamt, 2015) and yeast (Ebert et al., 2011; Semkiv et al., 2014, 2016; Luo et al., 2019). The presence of an existing futile cycle also suggests that low expression, or inducible control of any introduced ATP consuming reactions will be required. As some portion of cellular ATP is already being disposed of, consuming too much ATP could lead to growth cessation (Chao and Liao, 1994; Koebmann et al., 2002).

The final aspect confirmed by the ^{13}C -MFA was the importance, for high butanol yield, of NADH production and its subsequent oxidation through reactions relevant to butanol production. When

using I constrained the FBA model with experimentally measured fermentation product secretion rates, NADH was produced from glycolysis, but was oxidised completely without solvent formation through butyryl-CoA synthesis and ferredoxin: NAD(H) oxidoreductase. When the futile cycle identified by ^{13}C -MFA was imposed on the model, NADH was oxidised through malate dehydrogenase, lactate formation and butanol formation. As acid production and transfer of electrons to ferredoxin may be sufficient to oxidise NADH produced from glycolysis, further production of NADH will likely be key to ensuring flux is diverted to fermentation products. The alternative routes of NADH oxidation which were active in the ^{13}C -MFA flux map also highlight the importance of ensuring NADH is channelled towards the four-carbon end of the ABE fermentation pathway. Both of these findings confirm the results of chapter three, that the introduction of an NADH forming pathway combined with increased expression of butyryl-CoA synthesis genes will be important for driving butanol formation.

5.5 Conclusions

In this chapter, I have generated a flux map of the asporogenic *Clostridium* strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}. I identified major central metabolic fluxes, and a futile cycle, which I investigated using FBA simulations. I tentatively identified the role of this pathway to be oxidation of NADH, but this needs further study to be confirmed. The flux map will serve as a starting point to gain an understanding of metabolic fluxes of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} during acidogenic growth.

Chapter 6: Experimental Implementation of Metabolic Engineering Interventions Suggested by Computational Modelling

6.1 Introduction

6.1.1 Aim of Work Described in this Chapter

In chapter four it was established that the strain *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} was nearly optimal for butanol production and was potentially being limited by the rate of glucose uptake. Possible genetic manipulations to increase glucose uptake and improve the efficiency of butanol production were investigated using flux balance analysis of a genome scale metabolic model of solventogenic *Clostridium* in chapters three and four. The aim of the work presented in this chapter was to undertake an initial implementation of the desired manipulations into *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} and assess their performance in batch fermentations.

6.1.2 Engineering Solventogenic *Clostridium* for *n*-Butanol Production

Targeted manipulations of metabolic pathways are important for increasing the titre and yield of the desired target compound. This can be achieved by removing production of competing side products and increase the levels of enzymes catalysing conversion of precursors or increasing the availability of cofactors (Lee and Kim, 2015). *Clostridium* have traditionally been difficult to manipulate (Jones and Woods, 1986; Lee, Bennett and Papoutsakis, 1992), which has hampered efforts to engineer strains with improved butanol producing capacity. In recent years, several technologies have been developed to engineer these bacteria. These include non-replicative integrational plasmids (Wilkinson and Young, 1994; Green and Bennett, 1996; Green and et al., 1996), group II intron systems (Heap *et al.*, 2007, 2010) and, more recently, the development of CRISPR/Cas techniques (Y. Wang *et al.*, 2015; Li *et al.*, 2016; Pyne *et al.*, 2016; Wang *et al.*, 2016, 2017; Atmadjaja *et al.*, 2019). These approaches have increased the understanding of the molecular-genetics of ABE fermentation (Cooksley *et al.*, 2012; Dai *et al.*, 2016) and have enabled targeted metabolic engineering approaches to be implemented (Jiang *et al.*, 2009; Lehmann and Lütke-Eversloh, 2011; Jang, Lee and Lee, 2012; Lehmann *et al.*, 2012; Wang *et al.*, 2017; Nguyen *et al.*, 2018). Despite these advances, substantial increases in butanol production have not been achieved for many strains. Many of these approaches have relied on intuitive manipulation of the target strain, rather than system-scale guided engineering strategies. While these intuitive approaches have achieved some success; they often fail to achieve the desired outcome. An integrated approach combining *in silico* analysis with rational target selection may be able to produce better butanol producing strains.

6.1.3 Engineering Approaches Implemented in This Chapter

The aim of this chapter was to implement the metabolic engineering strategies based on my FBA simulations in chapters three and four, which I will briefly review here. The engineering approach designed from simulations in chapter three involved increasing butyryl-CoA forming flux and increasing NADH generation.

The increased NADH formation served to provide the cofactor requirements to increase flux through the butyryl-CoA forming and butanol forming pathways. This coincided with increased flux through the butyryl-CoA synthesis pathway to provide sufficient butyryl-CoA to be reduced to form butanol. I took the following approaches to engineer these flux changes into *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}.

To increase the butyryl-CoA forming flux, direct expression of all enzymes responsible for butyryl-CoA synthesis (the *bcs*-operon) was considered, as has previously been reported in wild type *C. saccharoperbutylacetonicum* N1-4 (Wang, Dong and Wang, 2017). This approach comes with limitations, however. The size of the *bcs*-operon in *C. saccharoperbutylacetonicum* N1-4 is approximately 6 kbp, which would result in a large (> 10 kbp) transformation vector. The size of this vector may have proven difficult to transform into *C. saccharoperbutylacetonicum* N1-4 (M. Nicolle, personal communication). An alternative approach involved removing the Rex redox response regulator, which is a repressor of the *bcs*-operon in *C. acetobutylicum* (Wietzke and Bahl, 2012; Zhang et al., 2014) and *C. pasteurianum* (Schwarz et al., 2017). A *rex* gene is also present in the *C. saccharoperbutylacetonicum* N1-4 genome, located upstream of the *bcs*-operon, as is the case in *C. acetobutylicum*. To investigate whether the *C. saccharoperbutylacetonicum* *rex* protein has a similar role to the *C. acetobutylicum* and *C. pasteurianum*, I performed a sequence alignment of these three protein sequences (Fig. 6.1).

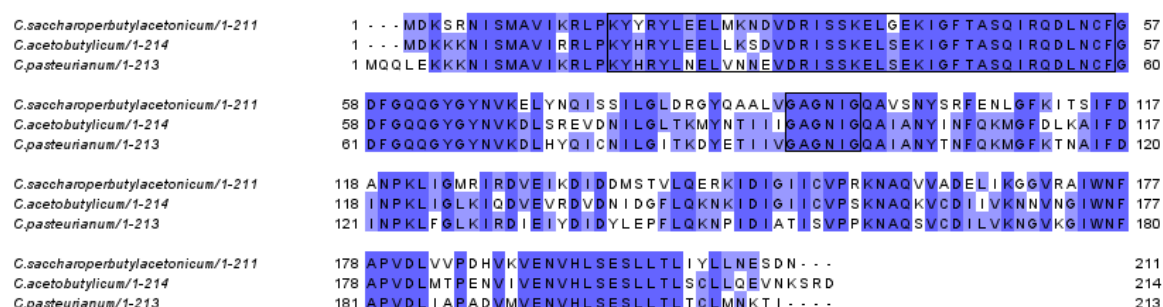


Figure 6.1 – Sequence alignment of Rex protein sequences from *C. saccharoperbutylacetonicum* N1-4 (UniProtKB accession number: M1LN94), *C. acetobutylicum* ATCC 824 (UniProtKB accession number: M1LN94) and *C. pasteurianum* DSM 525 (UniProtKB accession number: C. saccharoperbutylacetonicum N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl). The alignment was performed with the ClustalΩ algorithm (Sievers et al., 2011). Visualisation was performed using Jalview v2.11.2 (Waterhouse et al., 2009). Residues are coloured according to their conservation. Dark blue indicates that the residue is conserved in all three species. Light blue indicates residue is conserved in two of the three sequences. Black box around residues 20 – 56 is the DNA binding domain, and the black box around residues 92 – 97 is the nucleotide (NAD(H)) binding domain.

The three sequences had a 59 % identity and showed strong conservation of the NAD(H) binding domain and relatively strong conservation of the DNA binding domain. It seemed likely that the *rex* protein serves a similar physiological role in *C. saccharoperbutylacetonicum* as in *C. acetobutylicum* and *C. pasteurianum* and was targeted for knockout to increase the butyryl-CoA forming flux.

I explored the genome scale metabolic model of *C. saccharoperbutylacetonicum* for reactions which could increase NADH formation in chapter three. I identified two reactions which would be able to carry the required flux to meet the butanol synthesis target: ferredoxin: NAD(H) oxidoreductase and formate dehydrogenase (FDH). As discussed in chapter three, the fact that the former enzyme can function in both the NAD⁺ reducing and NADH oxidising direction made it unsuitable for overexpression. In contrast, the FDH reaction is irreversible (Tishkov and Popov, 2006). The *Candida boidinii* gene was selected for overexpression, as this enzyme has previously been expressed in an active form in *C. acetobutylicum* (S. Wang *et al.*, 2011), and has been widely used as an NADH generating system in both cellular and cell-free systems (Shaked and Whitesides, 1980; Neuhauser *et al.*, 1998; Berríos-Rivera, Bennett and San, 2002a; Mäde *et al.*, 2012; Barin *et al.*, 2018).

To use FDH as an NADH synthesising system, a high supply of formate also needs to be engineered into the cell. In chapter three, the only enzyme capable of this which was already in the model was pyruvate formate lyase (PFL). *C. saccharoperbutylacetonicum* has two predicted PFL genes, but these enzymes have not been functionally characterised. PFL has been proposed to have an anabolic role of providing C1 units in *Clostridium*, rather than its catabolic function in *E. coli* (Thauer, Kirchner and Jungermann, 1972). This role has been confirmed with ¹³C-MFA in *C. acetobutylicum*, where PFL functioned to supply C1 units (Au *et al.*, 2014). I also found that PFL flux was minor in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} (chapter five), with only a minor contribution to acetyl-CoA formation.

To ensure that the formate forming flux was at the levels required in the high butanol producing simulations in chapter three, a heterologous protein was chosen. The *E. coli* PFL is induced and active during anaerobiosis (Knappe and Sawers, 1990), replacing the pyruvate dehydrogenase complex as the pyruvate oxidising enzyme in anaerobic conditions (Knappe *et al.*, 1974). As PFL carries the entirety of acetyl-CoA forming flux under anaerobic conditions in *E. coli*, it seemed likely that the *E. coli* enzyme could carry out the desired catabolic role of acetyl-CoA synthesis if expressed in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. Thus, *E. coli* PFL was selected for co-expression with FDH. The catalytic activity of PFL relies on a glycyl radical, which is introduced by the PFL-activating enzyme (PFL-AE). The *E. coli* PFL-AE was therefore also included for coexpression.

In chapter four I found that *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} was Pareto optimised for butanol production, and to further increase the butanol forming flux in this strain increasing that rate of glucose uptake would be necessary. In chapter four, I performed simulations using three different pathways of glucose uptake: PTS, an ABC transporter, and a H⁺: glucose symporter. The PTS is an efficient system of glucose uptake and is utilised in many *Clostridium* (Mitchell, Shaw and Andrews, 1991; Lee et al., 2005; Tangney and Mitchell, 2007); however, it involves a complex phosphorelay system (Gosset, 2005). The PTS genes involved in glucose uptake have not been identified in *C. saccharoperbutylacetonicum* N1-4. To effectively increase glucose uptake through this system, the many candidate genes in the *C. saccharoperbutylacetonicum* genome would need to be characterised in order to identify the glucose specific components.

The ABC system, as with the PTS, consists of multiple distinct subunits in many bacteria (Verdon *et al.*, 2003), would require identification of the correct genes from the *C. saccharoperbutylacetonicum* genome. Heterologous expression of the PTS and ABC systems could have been utilised. These both involve large protein complexes (Robillard and Broos, 1999; Cui and Davidson, 2011) which would have to be targeted to the cell membrane. In contrast, the H⁺: glucose symporter only requires a single protein to be targeted to the membrane (Jahreis *et al.*, 2008).

The simulations I performed in chapter four using a H⁺: glucose symporter indicated that additional expression of hexokinase and pyruvate kinase would be necessary. As the H⁺: glucose symporter does not phosphorylate glucose during import (unlike the PTS), separate expression of a hexokinase is necessary for glucose to enter glycolysis. When I analysed the glycolytic fluxes in the H⁺: glucose symporter simulation in chapter four, I found that the pyruvate kinase flux also had to increase to maintain glycolytic flux. This was because the PTS is responsible for oxidising a significant portion of PEP, however the use of the H⁺: glucose symporter and glucokinase requires ATP instead of PEP to phosphorylate glucose. Thus, expression of three proteins was deemed necessary; H⁺: glucose symporter, a hexokinase, and a pyruvate kinase.

E. coli was chosen as the donor organism for these genes. This was based on a review of the literature, where expressing the *galP* (H⁺: glucose symporter) and *glk* (glucokinase) genes in an *E. coli* strain with an inactive PTS restored glucose uptake to wild type levels (Flores *et al.*, 1996; Hernández-Montalvo *et al.*, 2003). The aforementioned reports of Flores *et al.* and Hernández-Montalvo *et al.* suggest that combined *galP* and *glk* expression import glucose at a similar rate to the PTS and can be utilised in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} to augment the glucose uptake capacity of this strain. As mentioned, the modelling work suggested that increased pyruvate kinase flux would be necessary in the *galP* and *glk* expressing cell. *pykA* was selected from *E. coli* as the pyruvate kinase, for reasons discussed later.

The other feature of the increased glucose uptake simulations in chapter four was that the additional ATP formed from increased glycolytic flux would be excess to the cells requirements for biomass formation and non-growth associated maintenance. When increased glucose uptake was forced on the model, the model either avoided ATP generation by bypassing phosphoglycerate kinase or by increasing ATP hydrolysis through the NGAM reaction. Given that these fluxes deviated substantially from the predicted wild-type flux distribution in chapter four, it seemed likely that I would need to directly manipulate the ATP consumption of *C. saccharoperbutylacetonicum spo0A^{I261T}*. Further confirmation of this approach was identified in chapter five, which found that ATP was already in excess to cellular requirements during acidogenic growth. This suggests that additional ATP consuming reactions could be engineered into the cell to increase the rate of glucose uptake.

I developed two interventions which would increase ATP demand. The first was expression of an apyrase (Semkiv *et al.*, 2016), which cleaves terminal phosphate groups from ATP, reducing the ATP/ADP ratio. The apyrase also cleaves phosphate groups from ADP to form AMP (Babu *et al.*, 2002). This feature was the reason for selecting the *E. coli pykA* gene as the pyruvate kinase. The *E. coli pykA* gene product is allosterically activated by AMP (Kotlarz, Garreau and Buc, 1975), and the eventual aim was to express both apyrase and the designed glucose uptake system simultaneously. This would ensure that the pyruvate kinase was active under the high AMP conditions induced by the apyrase. The second intervention to increase ATP demand was introduction of a futile cycle as in the chapter four simulations. The problem with both the futile cycle suggested by the model and commonly implemented approaches e.g., overexpression of PEP synthase (Hädicke, Bettenbrock and Klamt, 2015) or PEP carboxylase and PEP carboxykinase (Liao, Chao and Patnaik, 1994), is that carbon becomes trapped in the cycles. This is clearly detrimental when trying to move carbon towards butanol. In chapter five I identified an existing futile cycle was present in the cell, suggesting that increasing the ATP demand on the cell might free additional carbon for butanol production.

I identified a potential non-carbon futile cycle centred around polyphosphate. Combining expression of polyphosphate kinase and exopolyphosphatase would results in simultaneous extension and cleaving of phosphate groups from a chain of inorganic phosphate (Keasling, van Dien and Pramanik, 1998). This would result in a net consumption of ATP, which could be used to balance out the additional ATP formed in a strain with an increased rate of glucose uptake. Given that this system had successfully been used in *E. coli*, the polyphosphate kinase and exopolyphosphatase genes were selected from this organism for expression.

The aim of this chapter was to implement these strategies in the asporogenic *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} using both genomic manipulations and plasmid based overexpressions to increase the butanol titre of this strain.

6.2 Results

6.2.1 Development and Characterisation of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} Δ *rex*

The *rex* gene of *C. saccharoperbutylacetonicum* was targeted for knockout using the CLEAVE system (Atmadjaja *et al.*, 2019). A step 1 CLEAVE vector was generated containing a truncated *rex* gene flanked by two DNA sequences homologous to the area surrounding the *C. saccharoperbutylacetonicum* N1-4 chromosomal *rex* locus. 774 bp upstream of the *rex* gene was amplified using primers *rex*_HR1 F/R, and the 842 bp downstream of the *rex* gene was amplified using primers *rex*_HR2 F/R. Purified genomic DNA of wild type *C. saccharoperbutylacetonicum* N1-4 was used as the template. The homology region PCR products contained complementary tails and were used to generate a truncated in-frame *rex* gene (Fig. 6.2) through overlapping PCR (Heckman and Pease, 2007). The wild type *rex* gene is 636 bp and the truncated gene is 108 bp.

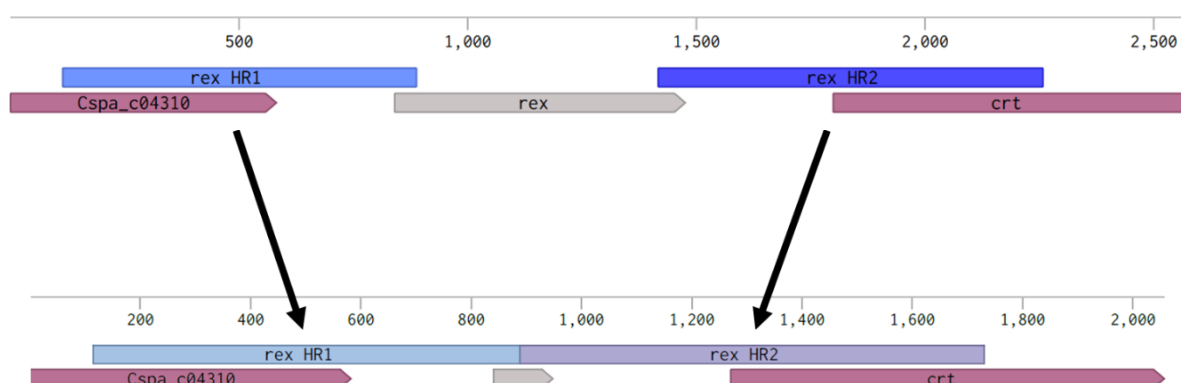


Figure 6.2 – Outline of the generation of the truncated *rex* gene. The top portion of the figure shows the genomic region of the wild type *rex* gene. The bottom portion shows the truncated *rex* gene (centre, grey). The homology regions (HR1 and HR2) either side of the *rex* gene were amplified by PCR, with primers containing complimentary tails. Following purification, a second PCR was performed. Here, the complimentary tails anneal, allowing extension by DNA polymerase. Finally, multiple copies of the full-length DNA construct, consisting of a truncated *rex* gene flanked by both homology arms, were generated. This construct was used to generate the Δ *rex* strain as described in the main text.

The full-length DNA construct consisting of the homology arms and truncated *rex* gene was cloned into the S1Va plasmid digested with *Pme*I, generating S1Va_ Δ *rex*. A synthetic CRISPR array designed to target the wild type was synthesised by Invitrogen™ (Darmstadt, Germany). This was cloned into the S3Va plasmid digested with *Zra*I and *Xho*I, generating plasmid S3Va_ Δ *rex*. The CLEAVE steps were performed as described in materials and methods section 2.8.16 with S1Va_ Δ *rex* as the step 1 vector followed by S3Va_ Δ *rex* as the step 3 vector. After the final transformation step with

S3Va_Δ*rex*, nineteen colonies were recovered. These were screened by colony PCR at the *rex* locus for the truncated *rex* gene with primers *rex_s3_screen* F/R (**Fig. 6.3**). While most of the colonies were parental cells which had not been successfully killed by S3Va_Δ*rex*, two were found to have the correct size sequence. One of these colonies was sent for confirmatory Sanger sequencing with the primer *rex_KO_genomic_screen_F*. This primer binds outside of the homology arms, ensuring that the sequence returned is of the genomic DNA rather than S1Va_Δ*rex*. The sequencing chromatogram alignment with the truncated *rex* gene is shown in **Appendix G**. This confirmed that the bacteria had undergone homologous recombination and that the *rex* gene had been truncated. The recovered strain is hereafter referred to as *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} Δ*rex*.

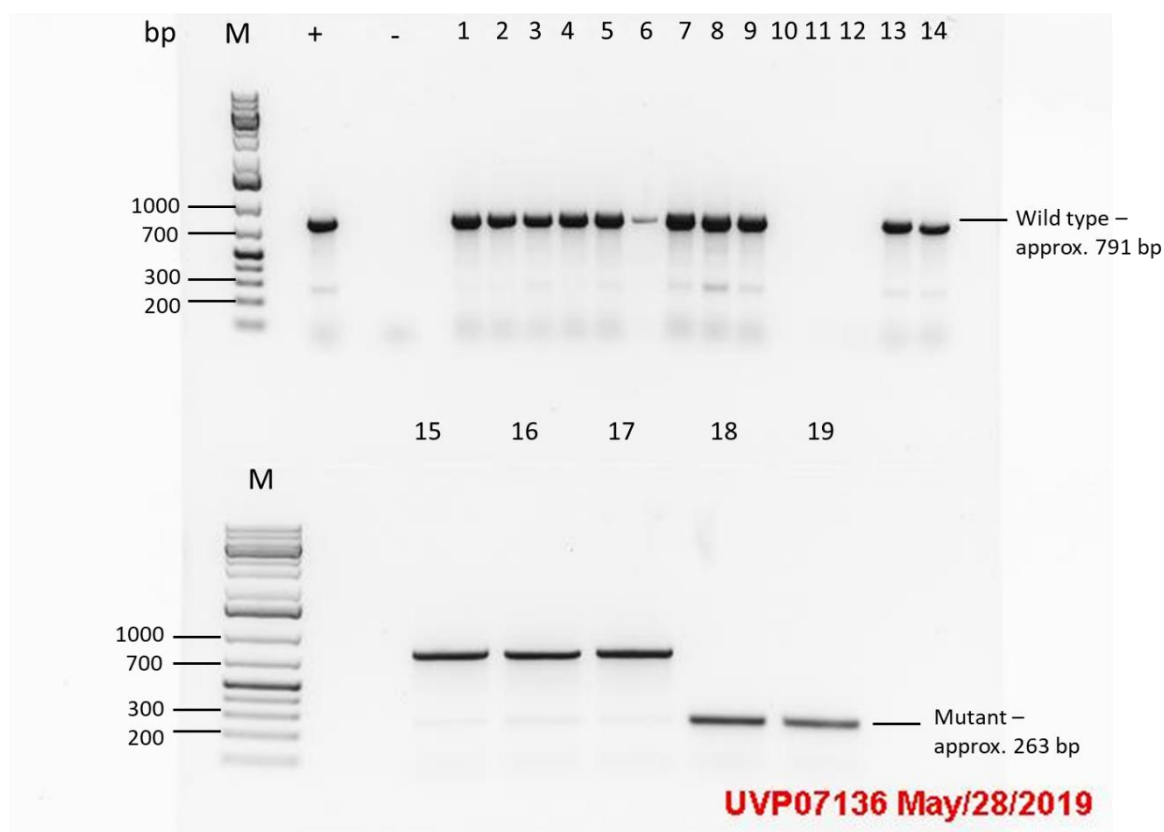


Figure 6.3 –Colony PCR analysis of genomic *rex* locus in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} transformed with S3Va_Δ*rex*. Colony PCR was performed as described in materials and methods section 2.8.6 with primers *rex_s3_screen* F/R. The sizes of the expected PCR products for wild type and *rex* knockout are indicated. No band in lanes 10-12 was likely due to PCR failure. This was most likely due to insufficient transfer of material into the tube when performing cell lysis. Lower lanes are larger than top lanes as only different sized combs were available when running the gel. Positive control consisted of wild type *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} transformed with S3Va. Negative control was water.

C. saccharoperbutylacetonicum N1-4 *spo0A*^{I261T} Δ*rex* was then subject to batch fermentations with *C. saccharoperbutylacetonicum* N1-4 as the control strain. Fermentations were performed as described in materials and methods section 2.6.3 in TYIR with 5 % glucose as the carbon source.

The fermentation kinetics are shown in **Fig. 6.4**. The Δrex strain showed a slight reduction in OD_{600nm} compared to the parental strain throughout the fermentation (**Fig. 6.4a**), however there was no statistically significant difference between the peak OD_{600nm} values (**Fig. 6.4i**). The production profile of acetone (**Fig. 6.4c**) and butanol (**Fig. 6.4d**) showed similar patterns in both strains. Unexpectedly, peak butanol and acetone titres were slightly reduced in the Δrex strain, while peak ethanol titres more than doubled (**Fig. 6.4i**). During the initial seventeen hours of growth the ethanol titres in both strains were similar, but after this the Δrex strain showed a large increase in ethanol titre (**Fig. 6.4e**). Acetate levels were also slightly higher in the Δrex strain at seventeen hours (**Fig. 6.4f, i**), and showed a clear reduction after seventeen hours indicating reutilisation. Butyrate titres were fairly similar in both strains (**Fig. 6.4g**), with a slight decrease in peak titres in the Δrex strain. The peak lactic acid levels were initially below the control in the Δrex strain, but then increased from 24 hours onwards (**Fig. 6.4h**). Key fermentation parameters are shown in **Table 6.1**. On a per biomass unit basis, all compounds were similar throughout the fermentation between the two strains, except for ethanol, which was higher in the Δrex strain (end point ratio: 0.03 vs 0.13 g L⁻¹ butanol OD_{600nm}⁻¹, $p < 0.001$, two-tailed Students t -test). The peak levels of fermentation products are shown in **Fig. 6.5**. The total amount of fermentation products was nearly identical in the two strains. There was no significant difference in the total solvent concentration, as the reduction in butanol titre in strain *C. saccharoperbutylacetonicum* N1-4 $spo0A^{1261T} \Delta rex$ was accompanied by an increase in ethanol titre. There was a significantly higher peak concentration of all acids ($p = 8.136e-05$) in *C. saccharoperbutylacetonicum* N1-4 $spo0A^{1261T} \Delta rex$, caused by the increase in lactate and acetate concentrations.

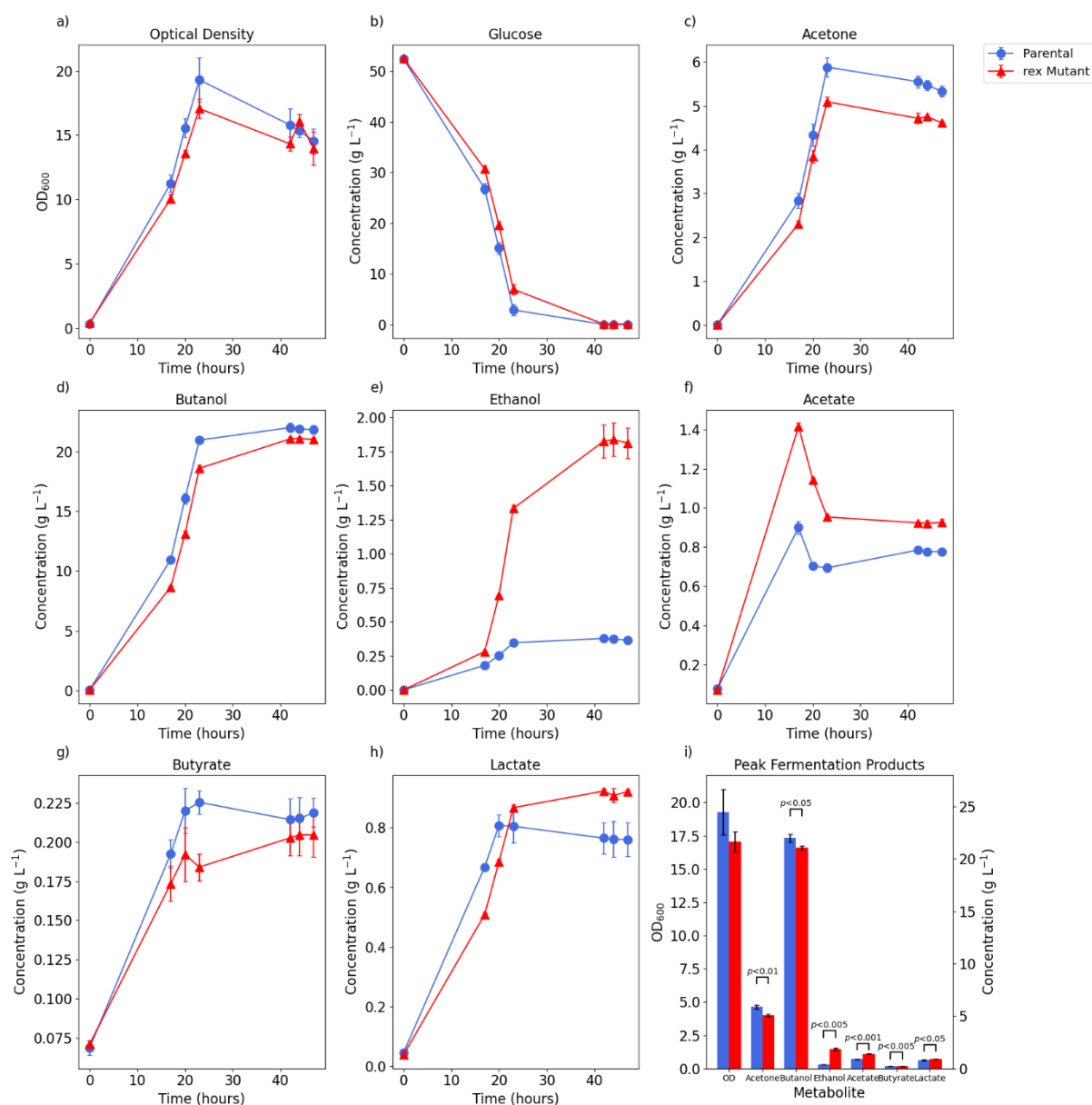


Figure 6.4 - Fermentation kinetics and peak metabolite concentration data from batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{261T} S3Va and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{261T} *Δrex* S3Va_Δ*rex*. a) Optical density as measured by absorbance at 600 nm. Concentrations of b) glucose, c) acetone, d) butanol, e) ethanol, f) acetate, g) butyrate, and h) lactate in the fermentation media were measured using HPLC. Peak fermentation products are shown in i). Batch fermentations were performed in 100 mL serum bottles with 60 mL of media. All fermentations were conducted in TYIR with 50 g L⁻¹ glucose and 1 % w/v CaCO₃ as a pH buffer. All fermentations were performed in triplicate. Error bars shown are ± 1 standard deviation. Where a statistically significant difference between the two strains was found ($p < 0.05$) the p -value is shown above the respective compound in i).

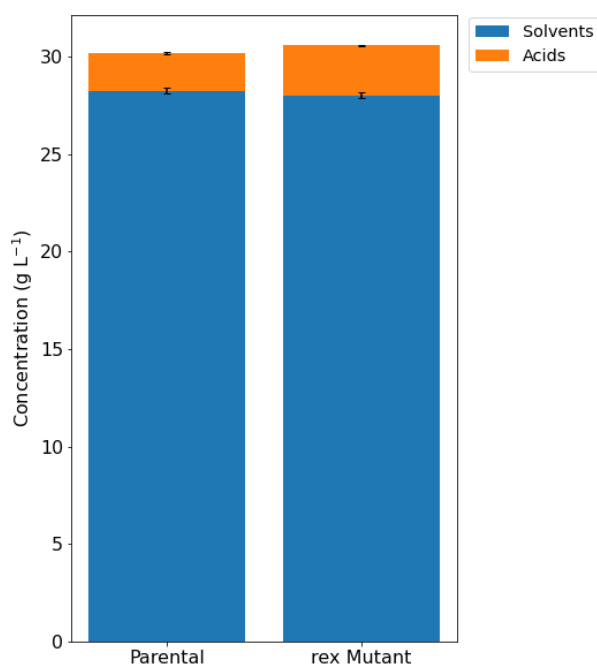


Figure 6.5 – Peak total solvents and acids in batch fermentation of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} S3Va and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} Δ rex S3Va_ Δ rex. Fermentation conditions were as described in legend of **Figure 6.4**. All fermentations were performed in triplicate. Error bars are $\pm$ 1 standard deviation.

6.2.2 *pflA*, *pflB* & *fdh* Overexpressing Strain

The NADH generating system consisting of FDH, PFL and PFL-AE was constructed as follows. The *fdh* gene from *Candida boidinii* was codon optimised for expression in *C. acetobutylicum* and synthesised by Twist bioscience (San Francisco, California). After subcloning into integration vectors (described in section 6.2.5) the *fdh* gene with the upstream thiolase promoter was amplified from S1Vb_int_flg_pThl_fdh using primers S1Vb_fdh_amp_F/R and cloned into pMTL82151 digested with *SacII* and *EcoRI*, generating plasmid pMTL82151_fdh. The PFL and PFL-AE genes (*pflB* and *pflA*) from *E. coli* were also subcloned into integration vectors as a single PCR product (see section 6.2.5). Both genes along with the thiolase promoter were amplified from S1Va_int_pThl_pflA_pflB with primers S1Va_pfl_gibs_F/R. The PCR product was cloned into pMTL82151_fdh digested with *Eco53kI* using Gibson assembly. This generated plasmid pMTL82151_fdh_pflA_pflB. The presence of all the desired genes on the single plasmid was confirmed by Sanger sequencing (see **Supplementary file 1** for complete sequence alignment read). Strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was cured of the S3Vb plasmid, which targets pMTL vectors using a synthetic CRISPR/Cas array, using a method developed by Biocleave Ltd. *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was then transformed with pMTL82151_fdh_pflA_pflB, referred to hereafter as strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl.

The expression of the introduced genes was confirmed through two methods. The expression of FDH was measured in crude cell extracts by measuring the production of NADH (described in

materials and methods section 2.10). The specific FDH activity of strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl was 0.11 U/mg. This was similar to previously reported values in *E. coli* expressing the *C. boidinii* *fdh* gene (0.231 U mg⁻¹, (Berríos-Rivera, Bennett and San, 2002b) and 0.18 U mg⁻¹ (Kratzer *et al.*, 2008)). No FDH activity was detected in the control strain carrying empty pMTL82151 vector.

PFL is a highly oxygen-sensitive enzyme (Crain and Broderick, 2014), and I did not have access to facilities to perform enzyme assays under anoxic conditions. Instead, I used RT-qPCR to confirm the expression of both the *pflB* gene and the activating enzyme *pflA* as described in materials and methods sections 2.8.17 – 2.8.19. The pullulanase gene (*pulA*) was used as an internal control for each of the samples. This gene has previously been used as a control in *C. acetobutylicum* (Sillers *et al.*, 2008). Here, it was found that there was no significant difference ($p > 0.05$) in the C_q of the *pul* between *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151 (empty vector control) (**Fig. 6.6**). This confirmed that the RT-qPCR was functioning as expected, and that differences in C_q values of the other genes investigated were not due to qPCR failure. Additional negative controls were included, which were reactions with no template, and RNA only controls. The latter consisted of a cDNA synthesis reaction (performed as described in materials and methods section 2.8.18) without reverse transcriptase added. None of these samples produced a C_q value above the threshold, indicating that reagents were free from contamination, and that any carry-over of DNA during RNA isolation was effectively removed by the DNase treatment.

Expression of *pflB* was only detected in strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl (**Fig. 6.6**). *pflA* on the other hand was detected in both *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151. This is not entirely unexpected, as *C. saccharoperbutylacetonicum* N1-4 contains a predicted PFL-AE gene in its genome. There was a significant difference in the C_q values between the two strains. The much higher C_q value in strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151 indicates that expression was far lower than in strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl. These results, along with those of the FDH assay, confirmed that all of the introduced genes were being expressed.

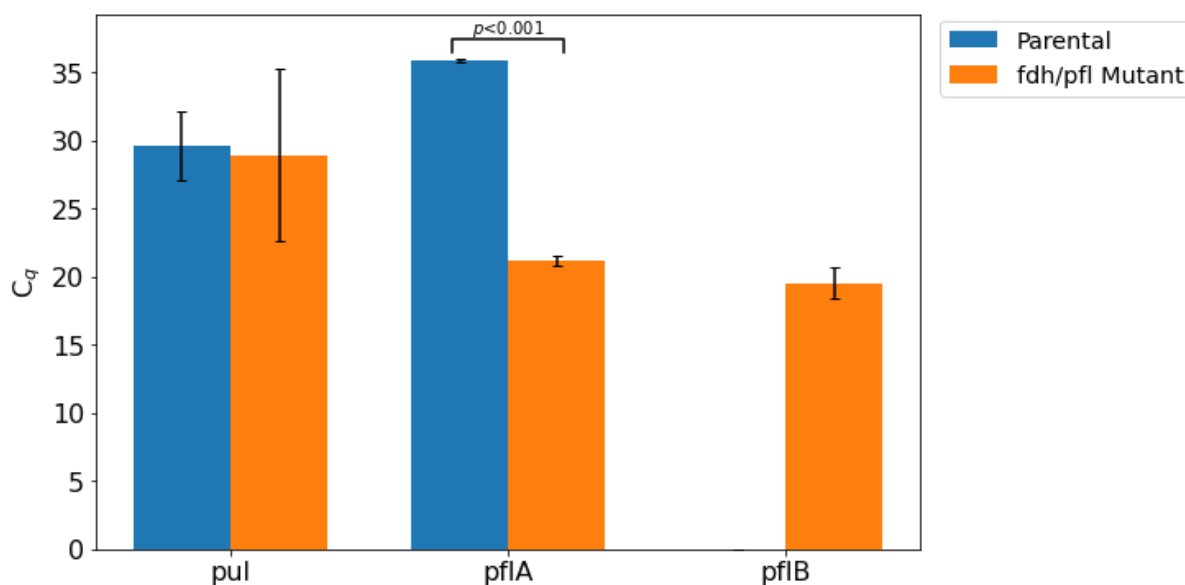


Figure 6.6 – RT-qPCR analysis of *pflA* and *pflB* expression in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151 and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151_fdh_pfl. RT-qPCR was performed as described in materials and methods sections 2.8.15 – 2.8.17. The pullulanase gene (*pulA*) was used as a positive control for these experiments. Data are presented as quantification cycle (C_q). A lower C_q indicates more abundant target. A C_q value of zero indicates that the amplified target did not reach the threshold and was below the detection level. There was no significant difference in the levels of the control gene (*pulA*) between the two strains. *pflB* was not detected in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151. *pflA* transcript was detected in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151, but the abundance was much lower than in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151_fdh_pfl. Significance testing was performed using a two-tailed Student's *t*-test. *n*=3 biological samples. For each biological replicate, the qPCR was run twice and the C_q values were averaged.

C. saccharoperbutylacetonicum N1-4 *spo0A*^{Δ261T} pMTL82151_fdh_pfl and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151 were then compared in batch fermentations. The fermentation profiles are shown in **Fig. 6.7** and summarised in **Table 6.1**. Growth and substrate consumption were similar in both strains (**Fig. 6.7a, b**), with no significant difference in peak OD_{600nm} (**Fig. 6.7i**). In terms of solvent production, both strains showed a similar production pattern of all three solvents, being first detected at around 24 hours and continuing throughout the fermentation. *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151_fdh_pfl had higher concentrations of butanol than the control between 24 and 30 hours (**Fig. 6.7d**), but the end point concentrations of butanol were not significantly different. Acetone levels were similar in both strains (**Fig. 6.7c**). During the initial portion of the fermentation, the ethanol titres were similar in both strains, but from 24 hours onward *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151_fdh_pfl had higher ethanol titres (**Fig. 6.7e**), but the end point concentrations were not significantly different (**Fig. 6.7i**). After the initial exponential growth phase, the acetate and butyrate levels were slightly lower than the control strain from 24 to 30 hours of fermentation (**Fig. 6.7f, g**), however the lactic acid levels were higher than the control during this time. The lactic acid was produced from the start of the fermentation in both strains (**Fig. 6.7h**) and showed a similar pattern of reutilisation. Higher peak acetate was detected in *C. saccharoperbutylacetonicum* N1-4

spo0A^{I261T} pMTL82151_fdh_pfl (**Fig. 6.7i**). Reutilisation of butyrate was also apparent after six hours of fermentation in both strains (**Fig. 6.7g**). When compared on a per biomass unit basis, end point acetate and lactate levels were not significantly different ($p=0.17$, $p=0.85$, respectively, two-tailed Students *t*-test). Peak levels of solvents and acids are shown in **Fig. 6.8**. There was no significant difference in either peak acid levels or peak solvent levels, as the increase lactate in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl was offset by the reduced acetate titre.

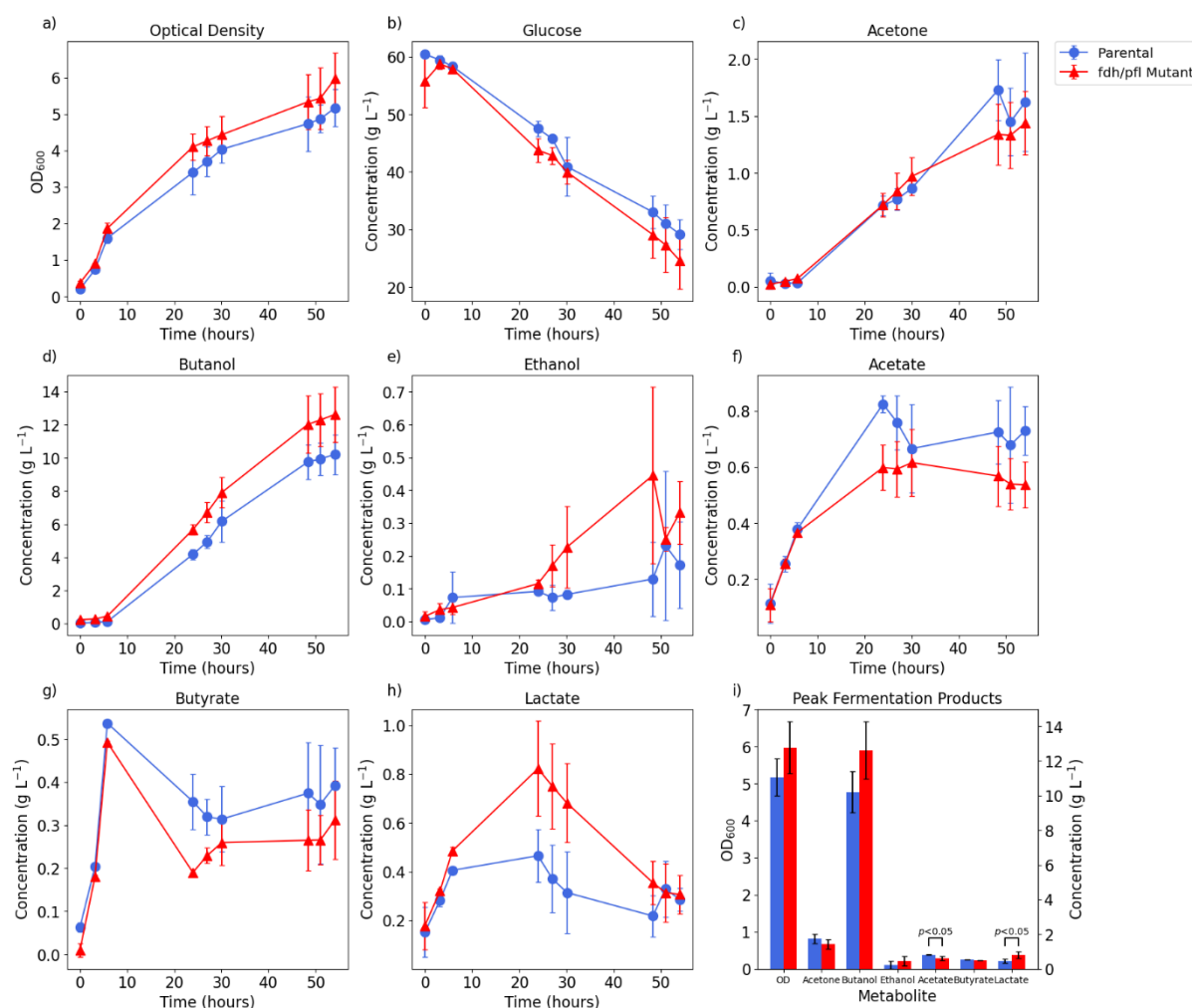


Figure 6.7 - Fermentation kinetics and peak metabolite concentration data from batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{1261T} pMTL82151 and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{1261T} pMTL82151_fdh_pfl. a) Optical density as measured by absorbance at 600 nm. Concentrations of b) glucose, c) acetone, d) butanol, e) ethanol, f) acetate, g) butyrate, and h) lactate in the fermentation media were measured using HPLC. Peak fermentation products are shown in i). Batch fermentations were performed in 100 mL serum bottles with 60 mL of media. All fermentations were conducted in TYIR with 50 g L⁻¹ glucose and 1 % w/v CaCO₃ as a pH buffer. All fermentations were performed in triplicate. Error bars shown are ± 1 standard deviation. Where a statistically significant difference between the two strains was found ($p < 0.05$) the p -value is shown above the respective compound in i).

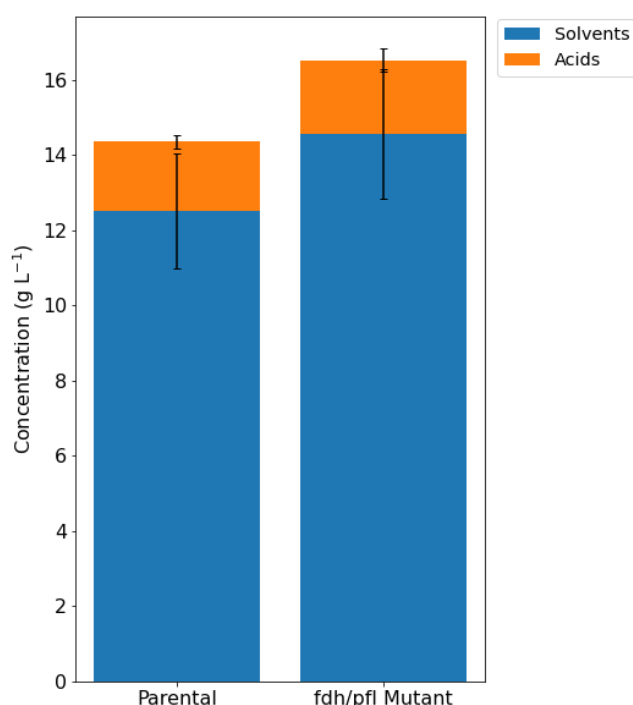


Figure 6.8 – Peak total solvents and acids in batch fermentation of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151 and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151 *fdh_pfl*. Fermentation conditions were as described in legend of Figure 6.7. All fermentations were performed in triplicate. Error bars are ± 1 standard deviation. There was no statistically significant difference in the levels of either total acids or solvents, as determined by a two-tailed Students *t*-test.

6.2.3 Characterisation of Cac_3451 Promoter

For the expression of the *apy*, *ppx* and *ppk* genes, which might be toxic to the cell, I attempted to develop an inducible promoter for expression experiments. Expression studies in *C. acetobutylicum* indicated that the Cac_3451 gene, encoding a putative sugar/Na⁺(H⁺) symporter (Grimmler *et al.*, 2010), would provide a good candidate: the gene was expressed when xylose was present in the media but did not show carbon catabolite repression when grown on glucose and xylose. The expression was first evaluated using a chloramphenicol acetyltransferase (CAT) reporter construct. A sequence was designed containing the chloramphenicol resistance gene (*catP*) under control of the Cac_3451 promoter, with the *C. pasteurianum* *fdx* terminator placed downstream of the *catP* gene. This construct was synthesised by synthesised by Invitrogen™ (Darmstadt, Germany). The reporter construct was amplified from the GeneArt vector using the M13 primer pair and was then blunt cloned into pMTL83251 digested with *Eco53kl*, generating plasmid pMTL83251_ *catP*. The correct insertion sequence was validated by Sanger sequencing (**Appendix H**). *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was transformed with pMTL83251_ *catP*, making strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL83251_ *catP*. To compare the differences in expression on glucose and xylose, cultures of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL83251_ *catP* were grown on xylose- and glucose-containing medium. Pellets were then harvested, and crude cell extracts were made. The CAT assay was performed to assess the strength of the Cac_3451 promoter in xylose and glucose relative to the *catP* promoter on the pMTL82151 plasmid. The specific activity of strains grown under both conditions are shown in **Fig. 6.9**. Activity was detected under both conditions, even in the absence of xylose. There was no

significant difference in CAT specific activity between the two conditions ($p=0.47$, two-tailed Students t -test), but both conditions showed significantly lower activity than the pMTL82151 positive control (xylose versus positive control $p=0.03$, Welch's t -test, glucose versus positive control $p=0.007$, two-tailed Student's t -test). On this basis, it seemed that, although not inducible, the promoter might be suitable for the constitutive, but weak expression of the ATP consuming enzymes.

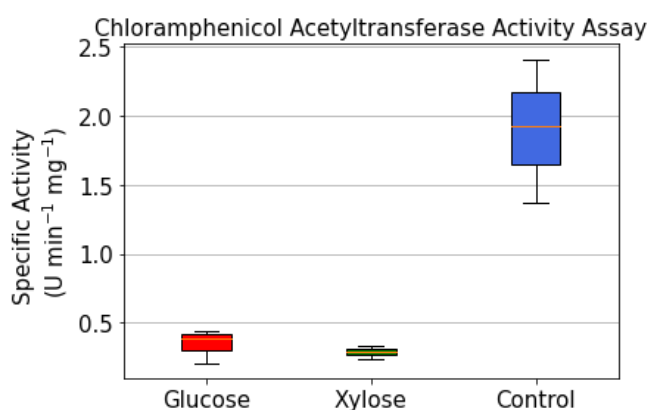


Figure 6.9 - Specific activities of crude extracts of asporogenic *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151 (control) and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL83251_ catP. The rate of chloramphenicol acetylation was measured using the spectrophotometric assay described in materials and methods section 2.10.3.

6.2.4 *apy* Expressing Strain

The apyrase (APY) expressing construct was produced as follows. The *apy* gene from *Shigella flexneri* strain FDAARGOS was codon optimised for expression in *C. acetobutylicum*. The codon optimised gene was designed with the Cac_3451 promoter and the *C. pasteurianum* *fdx* terminator. This construct was synthesised by Invitrogen™ (Darmstadt, Germany). The designed construct was extracted using a double digest of *Mlu*I-HF and *Sac*II. The construct was then cloned into pMTL82151 digested with *Mlu*I-HF and *Sac*II, forming plasmid pMTL82151_apy. The sequence of the insert was confirmed with Sanger sequencing (**Appendix I**). *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was then transformed with pMTL82151_apy, making strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy.

Due to various restrictions caused by the COVID-19 pandemic there was insufficient time to develop and troubleshoot an APY activity assay. As the RT-qPCR assay had already been developed for gene expression in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl and would be a faster way of confirming gene expression. This was used to confirm the expression of apyrase in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy. Again, the *pulA* gene was used as an internal control for all samples. RNA was isolated from *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy grown under glucose containing medium and xylose containing medium. The C_q of *pulA* under both conditions was not statistically different between *C.*

saccharoperbutylacetonicum N1-4 *spoOA*^{I261T} pMTL82151_apy and *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151 (empty vector control), again confirming the RT-qPCR assay was functional. The *apy* gene was detected in strain *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151_apy only, during growth on either glucose or xylose, confirming expression under the Cac_3451 promoter.

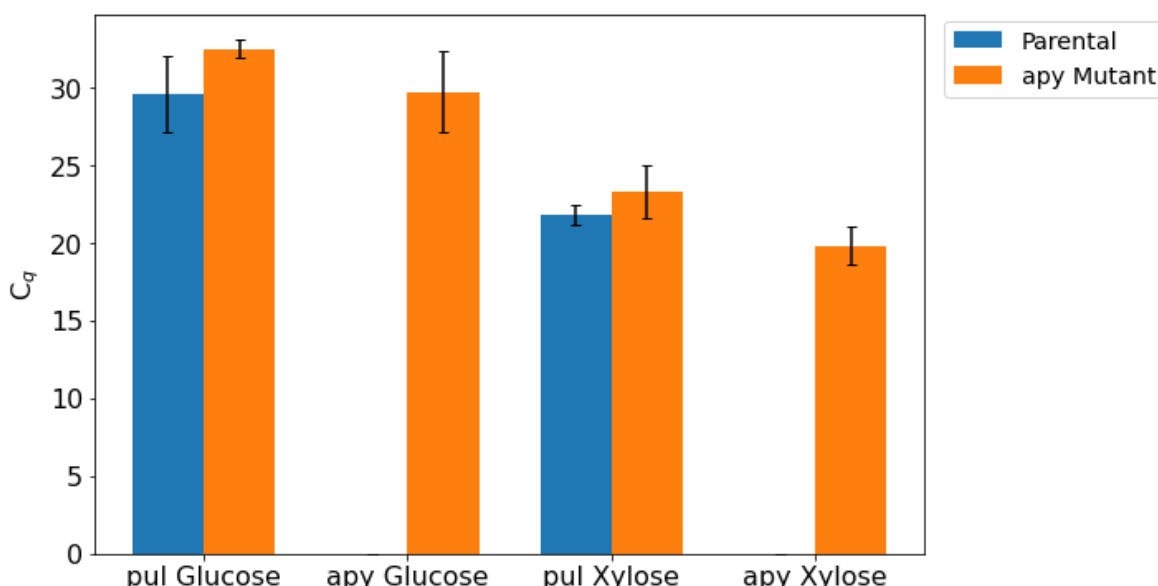


Figure 6.10 – RT-qPCR analysis of *apy* expression in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151 and *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151_apy. RT-qPCR was performed as described in materials and methods sections 2.8.15 – 2.8.17. The pullulanase gene (*pul*) was used as a positive control for these experiments. Data are presented as quantification cycle (C_q). A lower C_q indicates more abundant target. A C_q value of zero indicates that the amplified target did not reach the threshold and is below the detection level. There was no significant difference in the levels of the control gene (*pulA*) between the two strains on either glucose or xylose. *apy* was not detected in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151 in either glucose or xylose containing medium. Significance testing was performed using a two-tailed Students t-test. $n=3$ biological samples. For each biological replicate, the qPCR was run twice and the C_q values were averaged.

C. saccharoperbutylacetonicum N1-4 *spoOA*^{I261T} pMTL82151_apy was then subject to batch fermentations in TYIR media with xylose **Fig. 6.11** and glucose **Fig. 6.13** as carbon sources. In glucose containing medium, *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151_apy showed similar growth to the control during the initial six hours of fermentation, but quickly reached a plateau of around 1 OD_{600nm}. The control continued to grow until around 30 hours suggesting that the APY protein is inhibitory for growth (**Fig. 6.11a**). *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151_apy produced no detectable levels of acetone or butanol (**Fig 6.11c, d**). However, it did produce much more ethanol during the initial fermentation (**Fig. 6.11e**). The acid profile of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151_apy showed an even more dramatic alteration compared to the control. There was a minor decrease in butyrate concentration (**Fig. 6.11g**) in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} pMTL82151_apy, but a substantial increase in both acetate (**Fig. 6.11f**) and lactate (**Fig. 6.11h**) titres. The total acid and

solvent titres in the fermentation are shown in **Fig. 6.12**. Here, there was a major increase in peak acid titre in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} pMTL82151_apy, as the increase in the concentrations of these two acids more than offset the decrease in butyrate. There was also a substantial reduction in total solvent titre in this strain versus the control.

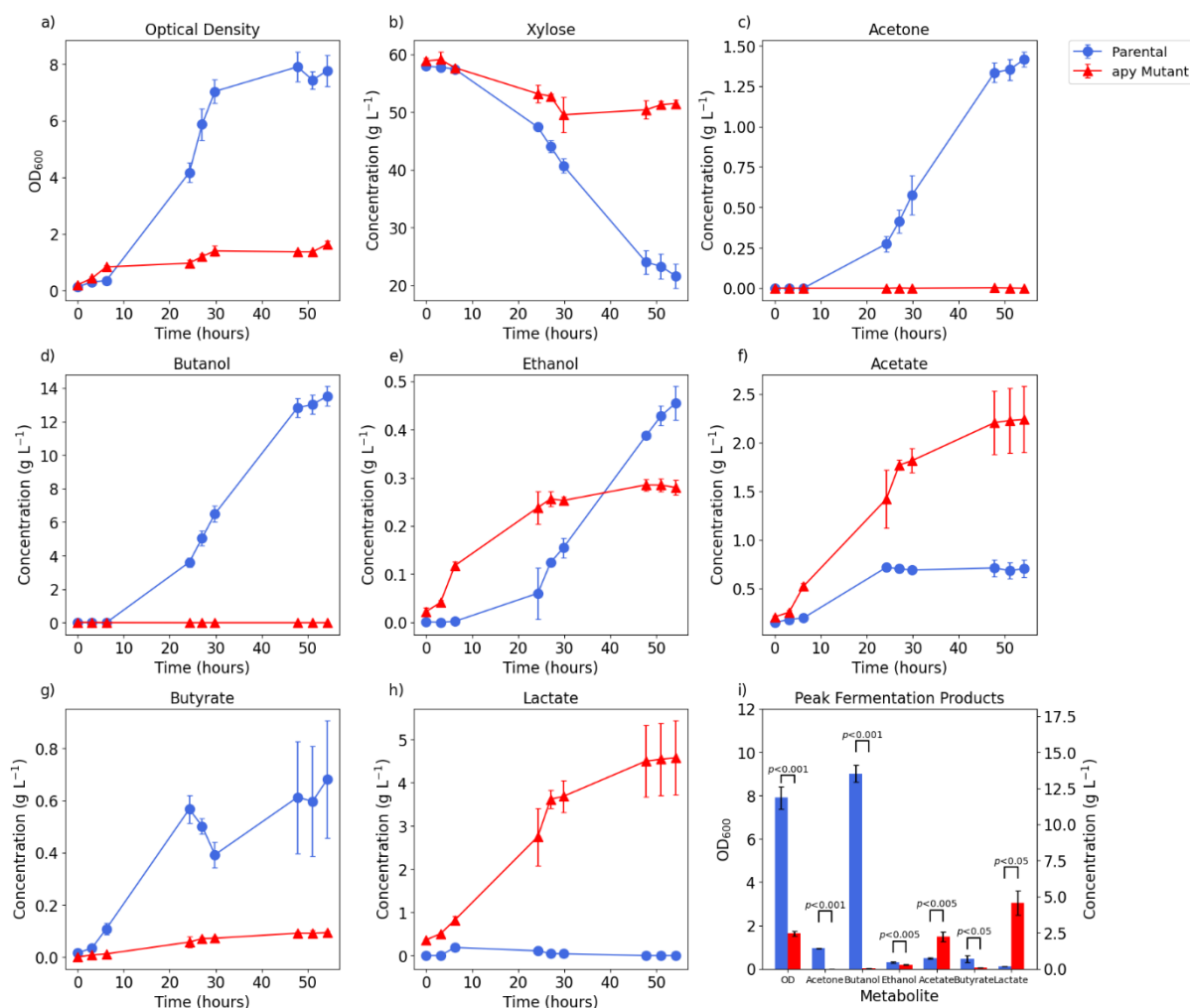


Figure 6.11 - Fermentation kinetics and peak metabolite concentration data from batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} pMTL82151 and *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T} pMTL82151_apy in xylose containing medium. a) Optical density as measured by absorbance at 600 nm. Concentrations of b) xylose, c) acetone, d) butanol, e) ethanol, f) acetate, g) butyrate, and h) lactate in the fermentation media were measured using HPLC. Batch fermentations were performed in 100 mL serum bottles with 60 mL of media. All fermentations were conducted in TYIR with 50 g L⁻¹ xylose and 1 % w/v CaCO₃ as a pH buffer. All fermentations were performed in triplicate. Error bars shown are ± 1 standard deviation. Where a statistically significant difference between the two strains was found ($p < 0.05$) the p -value is shown above the respective compound in i).

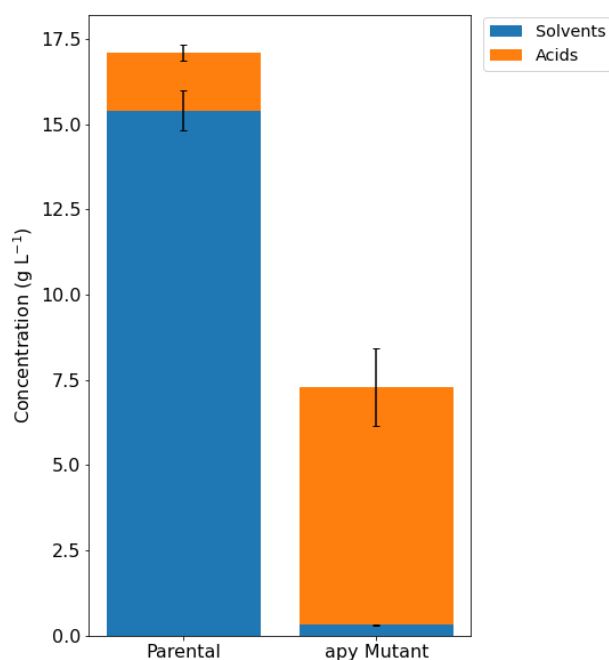


Figure 6.12 – Peak total solvents and acids in batch fermentation of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151 and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy in xylose containing medium. Fermentation conditions were as described in legend of **Figure 6.11**. All fermentations were performed in triplicate. Error bars are ± 1 standard deviation. There was a statistically significant difference in both total solvent ($p=0.0005$) and acid ($p=0.0014$) titres. Significance was determined with two-tailed Student's *t*-test (for acids) or two-tailed Welch's *t*-test (for solvents).

When grown on glucose, a similar cessation of growth after six hours was observed in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy (**Fig. 6.13a**). While initially, this strain showed a similar growth rate to the control strain, the OD_{600nm} quickly plateaued after six hours (**Fig. 6.13a**). Less glucose was consumed compared to the control (**Fig. 6.13b**), and neither acetone nor butanol was detected during the fermentation (**Fig. 6.13c, d**). Ethanol levels were also barely detectable in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy (**Fig. 6.13e**). Unlike the fermentation in xylose, acetate levels were much lower than the control strain (**Fig. 6.13f**), and butyrate levels were virtually zero throughout the fermentation (**Fig. 6.13g**). Lactate was the major fermentation product (**Fig. 6.13h**). The biomass specific rate of xylose uptake in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy was not significantly different from the control strain (10.16 mmol gDW⁻¹ h⁻¹ and 14.45 mmol gDW⁻¹ h⁻¹, respectively, $p=0.34$, two-tailed Student's *t*-test), but the glucose uptake rate was significantly higher in the *apy* expressing mutant compared to the parental control strain (10.27 mmol gDW⁻¹ h⁻¹ and 4.94 mmol gDW⁻¹ h⁻¹, respectively, $p=0.01$, two-tailed Student's *t*-test). In terms of total acids and solvents, *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy produced even more total acids on glucose than on xylose, with a similar low solvent titre (**Fig. 6.14**). Here though, lactate was effectively the only acid produced.

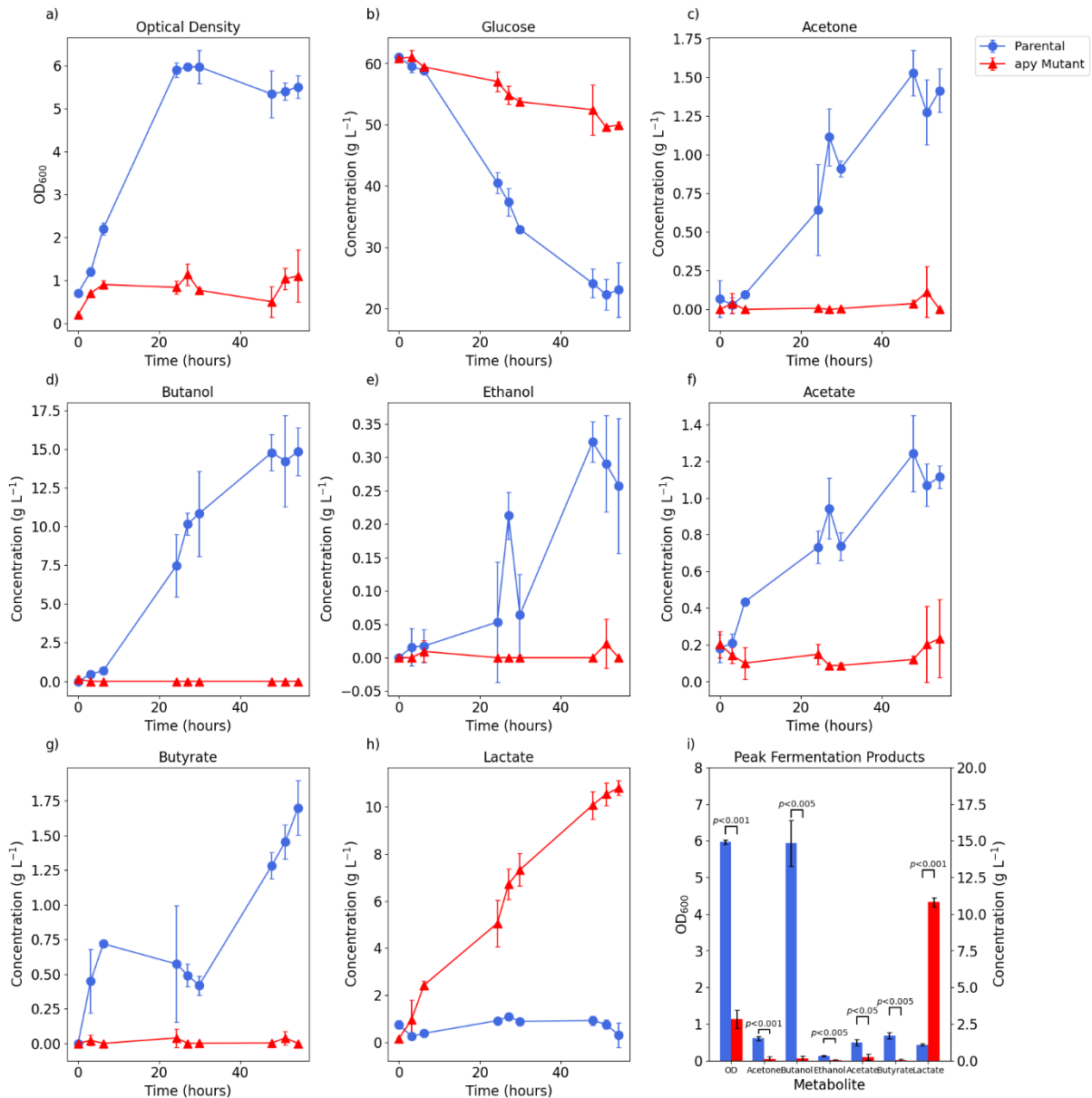


Figure 6.13 - Fermentation kinetics and peak metabolite concentration data from batch fermentations of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{261T} pMTL82151 and *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{261T} pMTL82151_apy in glucose containing medium. a) Optical density as measured by absorbance at 600 nm. Concentrations of b) glucose, c) acetone, d) butanol, e) ethanol, f) acetate, g) butyrate, and h) lactate in the fermentation media were measured using HPLC. Batch fermentations were performed in 100 mL serum bottles with 60 mL of media. All fermentations were conducted in TYIR with 50 g L⁻¹ glucose and 1 % w/v CaCO₃ as a pH buffer. All fermentations were performed in triplicate. Error bars shown are ± 1 standard deviation. Where a statistically significant difference between the two strains was found ($p < 0.05$) the p -value is shown above the respective compound in i). Note that the spike in concentration of ethanol, acetate and acetone was likely due to errors in loading samples into the tray.

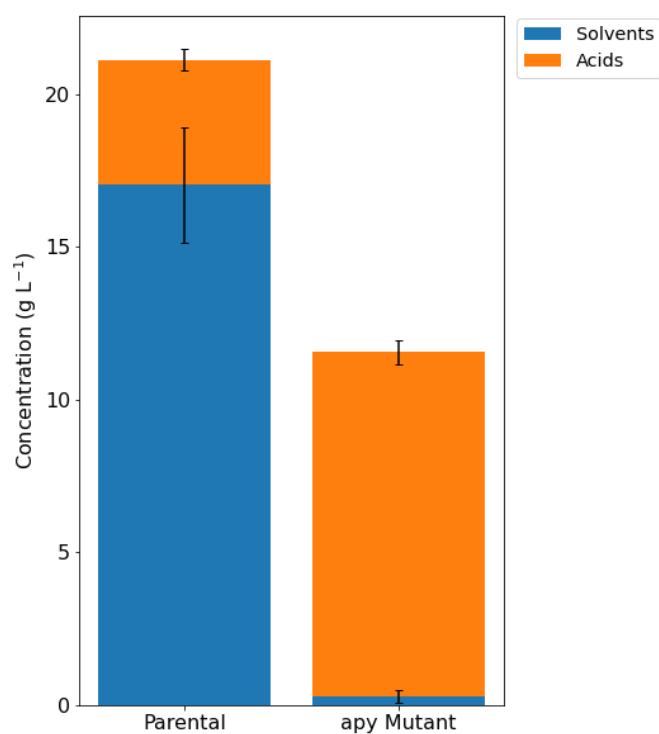


Figure 6.14 – Peak total solvents and acids in batch fermentation of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151 and *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{Δ261T} pMTL82151_apy in glucose containing medium. Fermentation conditions were as described in legend of **Figure 6.13**. All fermentations were performed in triplicate. Error bars are ± 1 standard deviation. There was a statistically significant difference in both total solvent ($p=0.004$) and acid ($p=1.962e-05$) titres. Significance was determined with two-tailed Student's *t*-test (for acids) or two-tailed Welch's *t*-test (for solvents).

Parameter	Experiment							
	rex KO		fdh/pfl		apy		apy	
			overexpression		overexpression (Glc)		overexpression (Xyl)	
	control	mutant	control	mutant	control	mutant	control	mutant
Growth rate	0.21	0.19	0.35	0.28	0.2	0.26	0.17	0.23
Substrate consumed	52.36	52.18	31.26	34.78	42.23	28.26	36.44	11.16
Acetate (peak)	0.9	1.41	0.86	0.64	1.29	0.39	0.76	2.25
Acetate (end point)	0.78	0.93	0.73	0.54	1.11	0.23	0.7	2.25
Butyrate (peak)	0.23	0.21	0.5	0.49	1.7	0.08	0.76	0.09
Butyrate (end point)	0.22	0.2	0.39	0.31	1.7	0	0.68	0.09
Lactate (peak)	0.82	0.93	0.47	0.82	1.1	10.81	0.19	4.63
Lactate (end point)	0.76	0.92	0.29	0.31	0.3	10.81	0	4.63
Total acids	1.97	2.55	1.83	1.96	4.09	11.28	1.7	6.98
Acetone	5.88	5.09	1.85	1.44	1.53	0.16	1.42	0
Butanol	22	21.1	10.34	12.61	15.19	0.09	13.53	0.02
Ethanol	0.38	1.84	0.33	0.51	0.33	0.03	0.45	0.29
ABE	28.26	28.03	12.52	14.56	17.05	0.28	15.4	0.31
ABE ratio	15:58:01	03:11:01	06:31:01	03:25:01	05:46:01	05:03:01	03:30:01	00:01:01
Total fermentation products	30.23	30.58	14.35	16.52	21.14	11.56	17.1	7.29
Butanol yield (g g ⁻¹)	0.42	0.4	0.33	0.36	0.36	0	0.37	0

Table 6.1 – Summary of fermentation parameters in the mutant strains developed in this chapter. Control strain for all experiments was asporogenic *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{261T} carrying empty pMTL82151 vector, with the exception of the *rex* KO experiment, which was the S3Va plasmid. All concentrations were determined by HPLC as described in section 2.9. Growth rate was determined as described in materials and methods section 2.11.2. All concentrations are given in g L⁻¹.

6.2.5 Unsuccessful Genetic Manipulation Attempts

In order to express the glucose uptake system consisting of *galP*, *glk* and *pykA*, a synthetic operon was designed. The operon was designed for expression in *C. autoethanogenum* using the Operon Calculator tool developed by De Novo DNA in design mode (Ng *et al.*, 2015). The selected promoter sequence was the ferredoxin promoter of *C. saccharoperbutylacetonicum*. The designed operon was synthesised by Twist Bioscience (San Francisco, California). I first attempted to use PCR to amplify the designed construct and insert it into the pMTL83251 shuttle vector, however no PCR product could be generated after multiple attempts. Inspection of the vector showed multiple *Hae*III cut sites on the vector backbone which were not present in the construct. I digested the plasmid with this enzyme, then purified the product via gel extraction. Blunt ligations were

attempted with pMTL83251 digested with *Eco53kl*. *E. coli* was transformed with the ligation mixture, and recovered colonies were screened by colony PCR. Positive colonies were identified by screening with the M13 primer pair, however when these were sent for confirmatory Sanger sequencing the *pykA* gene was truncated. I then repeated the cloning, and while the *pykA* gene was not as truncated there were still several missing and incorrect nucleotides at the end of the *pykA* gene. I repeated the cloning again with similar results. An alternative strategy was implemented by synthesising each gene individually under control of the thiolase promoter and ferredoxin terminator separately. Gene synthesis was performed by Invitrogen™ (Darmstadt, Germany). To generate a single construct containing all the genes, Gibson assembly was performed, with pMTL83251 digested with *Eco53kl* as the vector. The mixture was used to transform *E. coli*, and the recovered colonies were screened with colony PCR, as described in materials and methods section 2.8.6. Positive colonies were identified, which were sent for confirmatory Sanger sequencing. This revealed that all genes were inserted but each gene had been inserted into a separate plasmid, due to homologous regions between the GeneArt vectors and the pMTL83251 vector. Blunt cloning was attempted with these plasmids as recipients to generate dual expression vectors, but this was unsuccessful. Due to time constraints, this was not taken further.

Aside from the *apy* gene, simultaneous expression of the polyphosphate kinase and exopolyphosphatase genes was expected to generate the same ATP consuming effect. I attempted to construct pMTL82151_*ppk_ppx*, which would be a dual expression plasmid of both genes under control of the *Cac_3451* promoter. The *E. coli* *ppk* and *ppx* genes were codon optimised for expression in *C. acetobutylicum*. The genes were both designed as constructs under control of the *Cac_3451* promoter with the *C. pasteurianum* *fdx* terminator, and were synthesised by Invitrogen™ (Darmstadt, Germany). Initially, I attempted to generate plasmid pMTL82151_*ppx* by inserting the *ppx* construct into the *Eco53kl* construct of pMTL82151. The first attempt was unsuccessful, so this was attempted again. A single colony was recovered and screened with colony PCR, which gave a product at the expected size. The genes were sent for confirmatory Sanger sequencing. The sequence was correct except for a point mutation at position 868, causing amino acid 290 to become a stop codon. A third attempt was made to generate the target plasmid, but no colonies were recovered, therefore further attempts to develop the polyphosphate futile cycle system were abandoned.

In the main body of this thesis the NADH generating system of FDH, PFL and PFL-AE was expressed on a plasmid. Initially, I attempted to integrate the NADH generating system genes into *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} Δ *rex*. While the Δ *rex* strain alone failed to produce more butanol, it is more likely to work better in combination with the other interventions, as

suggested by the modelling in chapter three. Genomic integration would be desirable for this as it offers greater stability than plasmid born gene expression, which requires selection pressure to be maintained. I initially attempted to generate a *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} Δ *rex fdh_pfl* int strain, using the CLEAVE system. The plasmid S1Vb_int_flg_pThl was constructed by subcloning the thiolase promoter from S1Va_int_pThl into S1Vb_int_flg vector digested with *Xho*I and *Mlu*I, making plasmid S1Vb_int_flg_pThl. The *fdh* construct (see section 6.2.2) was then cloned into S1Vb_int_flg_pThl digested with *Stu*I, generating plasmid S1Vb_int_flg_pThl_fdh. An integration plasmid for PFL and PFL-AE was also constructed. S1Va_int_pThl was digested with *Stu*I. The *pflB* and *pflA* genes were amplified from *E. coli* as a single PCR product. This was cloned into S1Va_int_pThl digested with *Stu*I, making plasmid S1Va_int_pThl_pflA_pflB.

Initially S1Vb_int_flg_pThl_fdh was transformed into *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} Δ *rex*, and recovered colonies were picked and cultured in RCM media in the presence of thiamphenicol. This underwent subculture eighteen times before being transformed with the step 3 plasmid. A putative positive clone was identified through colony PCR, and this strain was transformed with S1Va_int_pThl_pflA_pflB. However, upon sending the PCR product for Sanger sequencing, the *fdh* integration was not present. At this stage, as the *pflB* integration plasmid had been transformed in, recovered colonies were subcultured in RCM as described above. If a *pflB* integrant strain could be developed, I could then repeat the transformation with the *fdh* integration plasmid to generate the double integration strain. After 43 subcultures with *pflB* integration plasmid, the strain was transformed with the step 3 integration plasmid. Colonies were recovered, however when screened with colony PCR no integrants were identified. A further five subcultures were performed, and the transformation was attempted again, but again no integrants were identified. Further attempts to generate genomic integrants were abandoned. The plasmids generated for integration here were used to generate the pMTL vectors in section 6.2.2.

6.3 Discussion

In this work, I generated several strains of solvent producing *Clostridium* from the parental asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}, which already produced increased levels of solvents compared to the wild type (Atmadjaja *et al.*, 2019). Here, FBA guided strategies were developed focused on alleviating potential substrate uptake limits, increased expression of butanol biosynthesis pathway enzymes via manipulation of the *rex* regulator gene, introduction of an NADH-generating system to encourage high rates of fermentation for NAD⁺ balancing and an ATP consumption system to deal with the potential excess ATP as a result of increased glucose uptake. All of these were guided by the simulations I performed in chapters three and four. Unfortunately, due to a combination of technical difficulties and time constraints, not all

of these strategies could be tested, and none were able to be tested in combination as would have been ideal.

Issues arose when cloning the genes for the polyphosphate futile cycle and the glucose uptake system, and due to time constraints, these were not taken further. Even in the current era of rapid and easily available gene synthesis, the low GC-content of *Clostridium* genomes leads to difficulties when synthesising genes optimised for expression in *Clostridium*. In addition, the gene products that were synthesised appeared to be toxic to the *E. coli* cloning host, as the polyphosphate kinase was only detected in the *E. coli*/*Clostridium* shuttle vector when a stop codon was introduced into the middle of the gene. Also, it was not possible to successfully clone the large synthetic operon consisting of the *galP*, *glk* and *pykA* genes into the *Clostridium* shuttle vector. This construct may have been toxic due to its large size. In addition, the failed integration of FDH and PFL and PFL-AE meant that these enzymes had to be expressed from a shuttle vector. The CLEAVE system has previously been used for integrations of up to 5 kb of λ -phage DNA, although efficiencies decreased as the size of the insert increased (Atmadjaja *et al.*, 2019). The size of the *fdh* insert was approximately 1.5 kbp and the *pfl* gene insert was approximately 3.5 kbp. Perhaps these integrations could have been achieved with prolonged subculturing or attempting to use other transformation systems such as flippase recombinases (Croux *et al.*, 2016) or Cas9 nickase (Li *et al.*, 2016).

Of the strains developed, none showed the desired increase in butanol, although they did show some expected characteristics and evidence of metabolic perturbations. The *rex* knockout has previously been described in *C. pasteurianum* and *C. acetobutylicum*, which in the former increased butanol titres (Schwarz *et al.*, 2017), while in the latter butanol production occurred earlier in the fermentation but the wild type reached similar titres by the end of the fermentation (Wietzke and Bahl, 2012). In this work, the *rex* knockout showed a decrease in peak butanol titre and an increase in ethanol titres. This might have been expected, as in *C. acetobutylicum* the Rex protein bound to *adhE2*, *ldh*, *crt* and *thlA* genes (Wietzke and Bahl, 2012; Zhang *et al.*, 2014). The *adhE2* gene, encoding bifunctional aldehyde/alcohol dehydrogenase, is responsible for formation of both butanol and ethanol from butyryl-CoA and acetyl-CoA, respectively (Inui *et al.*, 2008). Increasing the levels of this enzyme might increase the flux through either of these reactions. The increase in ethanol seen in the Δrex strain suggests that there was increased aldehyde/alcohol dehydrogenase activity. A similar reduction in butanol titre was reported when *adhE2* was overexpressed in *C. saccharoperbutylacetonicum* (Wang, Dong and Wang, 2017) and *C. beijerinckii* (Lu *et al.*, 2017). In addition, the increased peak lactate levels indicated that there was increased lactate dehydrogenase activity in the Δrex mutant, suggesting that transcription of at least one *ldh* is

repressed by Rex in *C. saccharoperbutylacetonicum*. The rationale for the *rex* knockout was that increasing the *bcs*-operon enzymes would increase flux towards butanol (Wang, Dong and Wang, 2017), but this was not seen here. This could indicate that the Rex protein in *C. saccharoperbutylacetonicum* N1-4 was not repressing transcription of the *bcs*-operon in the same manner as in *C. acetobutylicum*. An alternative reason could be that the mutation in the Spo0A protein of the parental strain (Atmadjaja *et al.*, 2019) was already increasing expression levels of the *bcs*-operon genes, or that other regulatory pathways were influencing expression/activity of these enzymes. A better characterisation of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} in terms of protein abundance is needed, mainly to confirm or refute the idea that the Rex actively represses the *bcs*-operon in this strain. If not, then future efforts to increase butyryl-CoA synthesis should focus on direct overexpression of butyryl-CoA synthesising enzymes.

The other system I incorporated was the NADH generating system that was used in high butanol producing simulations in chapter three. The PFL/FDH expression system suggested by the model has previously been used in *Klebsiella pneumoniae* and *E. coli* to drive production of reduced compounds (Berríos-Rivera, Bennett and San, 2002b, 2002a; Zhang *et al.*, 2009). In order to confirm the system I introduced was functional, I measured the FDH activity of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl. Due to technical limitations however, the PFL activity of this strain was not measured, only expression was confirmed by RT-qPCR. This does not necessarily mean that the PFL is active in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl. Aside from performing an assay of PFL activity following published protocols e.g., measuring [¹⁴C]-formate formation from [1-¹⁴C]-pyruvate (Thauer, Kirchniawy and Jungermann, 1972), the NADH/NAD⁺ ratio in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl should be compared to the control strain, to confirm that expression of both genes increases the NADH/NAD⁺ ratio in the cell. This was not performed here due to time constraints, but there are multiple ways to achieve this. One would be to directly measure the levels of NAD⁺ and NADH using fluorimetry, as has been reported in *C. acetobutylicum* (Vasconcelos, Girbal and Soucaille, 1994). An alternative method would be to use redox biosensing proteins. A recently published example in *E. coli* used a mutated *Bacillus subtilis* Rex (B-Rex) protein (Liu, Landick and Raman, 2019), which could be adapted for use in *C. saccharoperbutylacetonicum* N1-4. This would require development of a suitable promoter which appropriately responds to the NADH/NAD⁺ ratio (Liu, Landick and Raman, 2019).

If the NADH/NAD⁺ ratio was increased in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_fdh_pfl, it may be that additional NADH cannot drive increased fermentation flux as a result of capacity limitations in the butanol synthesis pathway. As the ethanol and lactate titres

increased in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_ *fdh_pfl* compared to the control, this suggests that increased NADH needed to be oxidised through alternative routes. The likely capacity constraint is thiolase. In *C. acetobutylicum*, this protein is regulated by the formation of a disulphide bond between the catalytic residues Cys88 and Cys378 (Kim *et al.*, 2015). Replacing the native thiolase with the *E. coli* AtoB, which does not form disulphide bonds, resulted in increased butanol forming flux in *C. acetobutylicum* (Nguyen *et al.*, 2018). While it is not known if the *C. saccharoperbutylacetonicum* N1-4 thiolase also forms disulphide bonds, the catalytic cysteine residues are conserved across all four predicted thiolase genes in *C. saccharoperbutylacetonicum* N1-4. Given that overexpressing thiolase only led to minor increases in butanol titre in this species (Wang, Dong and Wang, 2017), it might be worth combining the *rex* knockout with expression of a thiolase lacking the ability to form disulphide bonds. This might lead to increased capacity for butanol biosynthesis.

The apyrase expressing strain showed the most drastic change in fermentation profile compared with the control. The strains initially grew at a similar rate to the wild type, but by 24 hours the growth plateaued when grown on either glucose or xylose as carbon substrates. Solvent production was effectively eliminated in this strain, and lactate became the major fermentation product. When grown on xylose containing medium, acetate levels were higher than the control throughout the fermentation. The increased acetate could be explained as a method to increase the amount of ATP produced per mole of glucose metabolised, as each mole of acetate produced generates one mole of ATP through substrate level phosphorylation. In contrast, this was not seen when the strain was grown on glucose, where lactate was the major fermentation product, and acetate levels were below those seen in the wild type. The reason for this is not immediately clear; while the lactate production is sufficient to balance NADH, the lack of ATP clearly inhibited growth. It may be that regulation of the lactate dehydrogenase and/or aldehyde/alcohol dehydrogenase enzymes in response to the reduced cellular ATP/ADP/AMP levels. The enzymes phosphate butyryltransferase and butyrate kinase, which are responsible for butyrate production, have reduced specific activities when cellular ATP increases (Girbal and Soucaille, 1994), so higher butyrate levels when cells are operating under low ATP levels might be expected. Low solvents have been reported in continuous culture under low intracellular ATP levels, although this is generally accompanied by higher acetate titres (Monot *et al.*, 1982; Meyer, Roos and Papoutsakis, 1986; Meyer and Papoutsakis, 1989; Girbal, Croux, *et al.*, 1995). The high levels of lactate have previously been reported in *C. acetobutylicum* under iron limitation and when pH was maintained above pH 5.0 (Bahl *et al.*, 1986; Diez-Gonzalez, Russell and Hunter, 1995). Under glucose, the lactate per biomass yield was 35 g g⁻¹, and under xylose this was 10 g g⁻¹. For comparison, the wild type produced butanol at a yield of

8.8 g g⁻¹ on glucose and 6.2 g g⁻¹ on xylose. If the flux toward lactate could be redirected towards butanol through knockouts of lactate dehydrogenase genes, this would be advantageous as the strain could produce more butanol from less biomass.

One of the most important findings was that the apyrase expressing strain had an increased rate of glucose uptake compared to the control. This shows that the native PTS is in fact not limiting the rate of glucose uptake, which was suggested by the chapter four simulations. Assuming that cellular ATP was depleted by the expression of apyrase in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy, this indicates that the cell can increase glucose uptake when required. This kind of regulation, termed 'control-by-demand' has been reported in *E. coli*, where expressing the F₁-ATPase portion of F₁F₀ H⁺-ATP synthase increased glucose uptake in proportion to the expression level of the ATPase (Koebmann *et al.*, 2002). This ATP demand was able to increase the rate of glucose uptake by 70 % (Koebmann *et al.*, 2002). Another example of control-by-demand regulation has been reported in *Pseudomonas*. In the bacterium *Pseudomonas putida*, adding the compound 2,4-dinitrophenol (DNP), which uncouples oxidative phosphorylation (thereby increasing the cells ATP demand), to the media increased the glucose uptake rate (Ebert *et al.*, 2011). The preliminary work presented here suggests that an increased ATP demand is sufficient to drive increased glucose uptake in *C. saccharoperbutylacetonicum* N1-4. This needs further work to be confirmed, mainly calculating the cellular ATP/ADP ratio in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} pMTL82151_apy and the parental strain.

If the apyrase was indeed active, it might be possible to increase glucose uptake by expressing apyrase under a properly inducible promoter. Ideally, I would follow the approach taken by Koebmann *et al.* of expressing the apyrase under control of varying strengths of promoter to find the level of expression resulting in the biggest increase in glucose uptake rate. While the selection of promoters is somewhat limited in *Clostridium*, efforts are underway to expand the genetic toolkit of these bacteria. A recently developed library of synthetic promoters with varying expression strengths in *Clostridium* was shown to function similarly across distantly related species (*C. acetobutylicum* and *C. sporogenes*) (Mordaka and Heap, 2018). This library could be used in *C. saccharoperbutylacetonicum* to gradually increase expression of the apyrase to find the optimal expression level of apyrase.

Metabolic engineering is often a cyclical process, building from the previous round of experimentation to guide further strategies (Lee and Kim, 2015). Here, I found that, in contrast to the FBA simulations, the glucose uptake rate in the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} was not limited by the glucose uptake system.

Instead, placing an ATP demand on the cell was sufficient to drive the increase in glucose uptake. This could be combined with the *rex* knockout and the NADH generating system could potentially be enough to drive increased glucose uptake.

The increased glucose uptake in the apyrase expressing strain highlights the shortcomings of stoichiometric models, mainly the inability to incorporate regulatory information. This would have been useful to model the impact of increasing the ATP demand in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{1261T}. Such an approach has been used in the bacterium *Pseudomonas putida* to model increased ATP demand, which reflected experimental data where increasing ATP demand increased the glucose uptake flux (Tokic, Hatzimanikatis and Miskovic, 2020). Future engineering efforts would no doubt benefit from combined analysis of stoichiometric models of metabolism and kinetic models of metabolism, where available.

6.4 Conclusion

Model guided approaches can reveal non-intuitive approaches to improving butanol production, however the interventions here have not had the desired impact. Metabolic engineering is often a sequential process, requiring new strategies to be developed based on the results of the previous round(s) of manipulations. Here, multiple strains were developed using the modelling work described in previous chapters. The work here has highlighted some of the shortcomings of stoichiometric modelling to design metabolic engineering strategies, such as the alternative optima that may exist, and the inability of the model to replicate dynamic and kinetic effects. Further work would likely involve expanding the toolkit for manipulating *Clostridium* genetics, such as more effective tools for genomic integrations and a wider selection of promoters for altering enzyme expression levels.

Chapter 7: General Discussion

7.1 Summary of Results

This thesis aimed to gain an improved understanding of the fermentative metabolism of the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} and to devise and test engineering strategies to increase butanol production in the strain.

In chapter three, I performed preliminary batch fermentations of wild type *C. saccharoperbutylacetonicum* N1-4. By quantifying the concentrations of fermentation end products with HPLC data, I estimated experimental constraints on substrate consumption and product secretion rates, which were applied to a genome scale metabolic model. Systematic explorations in the model of single and double reaction knockouts were used, in conjunction with the strain design algorithms OptKnock and OptForce, to identify targets which could improve butanol production. I also modelled the insertion of the Entner-Doudoroff pathway and a synthetic 4-hydroxybutyrate shunt. These latter strategies were not chosen for implementation based on a reduced growth rate and a reduction in butanol yield compared with the canonical ABE fermentation pathway, respectively.

In chapter four, I performed batch fermentations on the asporogenic strain *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} and found that this strain not only produced more butanol than the wild type but was Pareto optimised for butanol production. This suggested that the rate of glucose uptake was limiting the butanol forming capability of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}. I performed further FBA simulations where the rate of glucose uptake was increased and explored in-silico the metabolic consequences of introducing a membrane permease for glucose uptake, a hexokinase, and a pyruvate kinase. I also found that increasing the rate of glucose uptake would form ATP surplus to cellular requirements for maintenance and growth. This suggested that increased glucose uptake would also require introduction of ATP consuming pathways.

In chapter five, I used ¹³C-MFA to resolve the intracellular fluxes of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} in detail. I adapted a previously published metabolic network model to estimate fluxes using parallel labelling experiments with [1-¹³C]- and [U-¹³C]-glucose. I generated a flux map with a statistically acceptable fit and revealed that glycolysis proceeded through the EMP pathway rather than the ED pathway. An unexpected finding was that *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261} utilised a futile cycle, which was investigated using constraint-based modelling. The role of this cycle appeared to be the oxidation of NADH.

In chapter six, I made initial attempts to implement a metabolic engineering strategy to increase butanol yield in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} based on the work in the preceding chapters. This involved expression of an NADH generating system consisting of pyruvate-formate lyase and formate dehydrogenase, knocking out the *rex* redox response regulator gene and expressing an apyrase. Other strategies of expressing the sugar symporter GalP, glucokinase and pyruvate kinase as a single synthetic operon, and expressing polyphosphate kinase and exopolyphosphatase were attempted but I was unable to clone the genes into the appropriate shuttle vectors. Knocking out the *rex* gene led to a reduction in butanol titre and an increase in ethanol titre. Expressing pyruvate formate lyase and formate dehydrogenase did not lead to a significant increase in butanol titre. Expressing the apyrase led to drastic changes in fermentation profile in both glucose and xylose fermentations. In the former, lactate was the sole fermentation product, while in the latter there was some acetate formation.

7.2 Evaluating the Metabolic Engineering Strategy in this Thesis

7.2.1 *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} as a Chassis for Butanol Production

This thesis set out to explore the phenotype of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} compared with *C. saccharoperbutylacetonicum* N1-4. As in the original work describing this strain, I found an increase in butanol titre compared with the wild type (Atmadjaja *et al.*, 2019), but also found that this strain was Pareto optimised for butanol production (chapter four). This fact, coupled with the asporogenic phenotype, makes *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} ideal as a starting strain for butanol production.

The ¹³C-MFA flux map that I developed for this strain resolved the EMP glycolysis and PPP fluxes, which were similar to *C. acetobutylicum* fluxes (Au *et al.*, 2014). However, the finding of a futile cycle in chapter five was unexpected. Based on FBA simulations performed in chapter five, it seemed that the role of the futile cycle was the oxidation of NADH. While this work did shed light on the internal fluxes of *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}, which is important for guiding metabolic engineering strategies (Long and Antoniewicz, 2019), the experiment was performed in the acidogenic phase of the culture. The fluxes in solventogenic phase should also be determined. Previous isotopic tracer experiments in *C. acetobutylicum* have explored the shifts in the metabolome during the transition from acidogenesis to solventogenesis (Amador-Noguez *et al.*, 2011). Substantial changes in the metabolome were found, such as a reduction in glycolytic intermediates and certain amino acids during solventogenesis, accompanied by a reduction in TCA cycle flux (Amador-Noguez *et al.*, 2011). Given the biphasic fermentation profile in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} it seems likely that there will be similar changes in the

flux distribution. Identifying the fluxes in solventogenesis will be needed to find the most effective targets for metabolic engineering.

Determining solventogenic fluxes could be achieved through several methods. One would be to repeat the fermentation as in chapter five but replace ammonium acetate with an alternative pH buffer, such as calcium carbonate or MES, to allow the culture to naturally progress from acidogenesis to solventogenesis. This would require extending MFA from steady state to model this nonstationary system, utilising the framework reported by (Antoniewicz *et al.*, 2007) for determining fluxes throughout a fed-batch fermentation of *E. coli*. Here, the MFA model has additional parameters in the flux model accounting for fractions of labelled product and the contribution of feed to the labelling of products (Antoniewicz *et al.*, 2007). An alternative approach would be performing chemostat fermentations under conditions which maintain solventogenesis, i.e., phosphate limited conditions (Girbal, Croux, et al., 1995). Running fermentations for at least five residence times prior to addition of isotopic tracer would ensure the culture is in metabolic steady state (Long and Antoniewicz, 2019). Regardless of the method, determining the flux distribution in solventogenic culture will enable targeted metabolic engineering interventions to be designed.

Determining a more complete picture of metabolic fluxes in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} would be beneficial from a metabolic engineering perspective, as it would determine the fluxes which need to be altered to increase butanol production. This would not, however, show what the reasons are for the improved butanol production in terms of gene expression. The SpoOA protein is implicated in the expression of a huge range of genes (Alsaker et al., 2004; Sullivan and Bennett, 2006), both through binding directly to the respective genes (Ravagnani *et al.*, 2000) and through regulating downstream transcription regulators (Harris, Welker and Papoutsakis, 2002). Given the point mutation in the SpoOA DNA binding domain in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} (Atmadjaja *et al.*, 2019), it is likely that this strain has a unique transcriptional profile that deserves investigation. While there is not always a link between transcript abundance and protein levels (Muers, 2011), the identification of genetic regulators which control expression of genes of interest could enable more efficient metabolic engineering strategies to be developed, as this can increase the expression of multiple genes simultaneously. For example, deleting the ArcA transcriptional repressor in *E. coli* increased flux through the TCA cycle by increasing the expression of citrate synthase, aconitase and isocitrate dehydrogenase simultaneously (Yim *et al.*, 2011). In *Clostridium*, the Rex protein binds to genes in the butyryl-CoA synthesis operon (Wietzke and Bahl, 2012; Zhang et al., 2014), which led me to inactivate the gene in *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} in chapter six. This inactivation did not increase butanol titre, which contrasts with

the inactivation in *C. acetobutylicum* (Wietzke and Bahl, 2012) and *C. pasteurianum* (Schwarz *et al.*, 2017). This could be due to alterations in Rex expression because of the mutated Spo0A protein in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}.

While the work performed in this thesis provided some insight into the metabolism of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, further studies examining the alterations in transcript and protein abundance would reveal the reasons for the altered behaviour of this strain compared to the wild type, as has been performed with *C. acetobutylicum* Spo0A mutants (Sullivan and Bennett, 2006) and *C. beijerinckii* (Máté de Gérando *et al.*, 2018).

7.2.2 Evaluating the Metabolic Engineering Attempts in this Thesis

Using the superior strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} as a starting point, the aim of this thesis was to use *in silico* guided approaches to guide metabolic engineering approaches (chapters three and four), which I then implemented in the laboratory (chapter six). The first approaches I developed were based on data from wild type *C. saccharoperbutylacetonicum* N1-4. Increasing flux towards butyryl-CoA from acetyl-CoA was one strategy which the modelling suggested would increase butanol formation. Increasing butyryl-CoA forming flux was successful in *C. saccharoperbutylacetonicum* N1-4 by expressing additional copies of the *bcs*-operon genes increased butanol titres by 14 % (Wang, Dong and Wang, 2017). As discussed earlier, I deleted the *rex* gene to increase flux toward butyryl-CoA synthesis rather than directly expressing the butyryl-CoA synthesis genes. I only found an increase in ethanol and not butanol in chapter six with strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} Δ *rex*, which was also reported in *C. acetobutylicum* (Zhang *et al.*, 2014). It would have been interesting to directly overexpress the *bcs*-operon in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}, but this was not attempted due to the relative ease of generating a plasmid with a single insert i.e., the truncated *rex* gene with the homology arms, rather than generate a plasmid with four genes. Confirmatory proteomics of strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} Δ *rex* would be necessary to confirm the *rex* knockout led to increased levels of the *bcs*-operon.

The other strategy identified in chapter three was the increased generation of NADH to force flux towards more reduced fermentation products. My strategy involved the combined expression of pyruvate formate lyase and formate dehydrogenase. While expression of formate dehydrogenase in *C. acetobutylicum* has been reported previously (S. Wang *et al.*, 2011), this was not utilised to form additional NADH. Expression of formate dehydrogenase in *Klebsiella oxytoca* increased titres of ethanol, 1,3-propanediol and lactate (Zhang *et al.*, 2009). In *E. coli* similar results were obtained, as the ethanol/acetate ratio increased when formate dehydrogenase was overexpressed (Sánchez, Bennett and San, 2005). Clearly increasing NADH levels increases flux towards more reduced

compounds. While I did not find an increase in butanol titre when I expressed pyruvate formate lyase and formate dehydrogenase, I did see an increase in lactate titre (chapter six). While I confirmed that the formate dehydrogenase was active, I was only able to confirm expression of pyruvate formate lyase through RT-qPCR, not activity, due to the oxygen sensitivity of this enzyme (Sawers and Watson, 1998). To confirm expression of formate dehydrogenase and pyruvate formate lyase led to the increase in NADH, the NADH/NAD⁺ ratio would need to be evaluated through, for example, the Rex biosensor reporter method (Y. Liu et al., 2019).

The other strategies which were developed based on FBA of *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} were increasing the rate of glucose uptake by expressing a H⁺: glucose symporter, hexokinase, and pyruvate kinase. The FBA work also suggested that this would need to be coupled with an ATP consuming enzyme to prevent ATP accumulation. These strategies were much more difficult to implement, and I was only able to generate an expression plasmid for the apyrase gene. Much like the pyruvate formate lyase expressing strain, I was unable to confirm enzyme activity of the apyrase, and only confirmed expression via RT-qPCR (chapter six). To confirm the apyrase activity, I would implement a fluorescent reporter to analyse the ATP/ADP ratio in the apyrase expressing strain compared to wild type, such as the Perceval system (Berg, Hung and Yellen, 2009). Expressing apyrase in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} effectively abolished solvent production, with lactate becoming the sole fermentation product. This was unexpected, although the possibility of lactate production was identified in the FVA simulations in chapters three and four. Futile ATP consuming cycles have been used in *S. cerevisiae* to increase the ethanol yield through reducing biomass formation (Semkiv *et al.*, 2016). While expressing the apyrase did not lead to an increase in butanol yield, the lactate yield in apyrase expressing *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} increased dramatically, suggesting that reducing ATP availability diverts flux away from biomass formation and towards fermentation products in *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T}.

I was unable to generate the polyphosphate kinase and exopolyphosphatase plasmid or the membrane permease, glucokinase and pyruvate kinase plasmid. Given longer, it would have been interesting to generate these plasmids and see the effects of expressing these genes. While to my knowledge, no previous work has expressed the polyphosphate kinase or exopolyphosphatase genes in *Clostridium*, these genes have both been expressed in *E. coli* (Keasling, van Dien and Pramanik, 1998). FBA of *E. coli* overexpressing both polyphosphate kinase and exopolyphosphatase predicted an increase in TCA cycle flux under aerobic conditions to meet the increased ATP demand (Keasling, van Dien and Pramanik, 1998). Given that *Clostridium* generate ATP through acetate kinase and butyrate kinase, increased acetate and/or butyrate would be expected to counteract

reduced cellular ATP levels. I found this to be the case when apyrase expressing *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} was grown on xylose, but not glucose. The reasons for this difference are unclear.

Part of the ATP reducing strategy involved generating a novel promoter to control the expression of apyrase. Excessive ATP consumption could halt growth entirely, so tight control of gene expression is necessary when expressing ATP consuming enzymes. In chapter four I attempted to design a xylose inducible promoter for the apyrase gene, but this promoter did not show a difference in the reporter assay when grown on xylose or glucose. This highlights an additional problem working with *Clostridium*, in that the genetic toolkit is not as developed as for model organisms (Mordaka and Heap, 2018).

Due to time constraints, I was unable to generate strains which had multiple manipulations. In published metabolic engineering efforts, particularly with solventogenic *Clostridium*, achieving high butanol production required multiple manipulations, such as removal of acid pathways and replacing enzymes in the *bcs*-operon (Nguyen *et al.*, 2018), or overexpressing *adhE* while removing acid forming pathways (Jang, Lee and Lee, 2012). It is possible that combined manipulations, such as deleting *rex* or overexpressing the *bcs*-operon while expressing pyruvate formate lyase and formate dehydrogenase, would lead to an increase in butanol titre. This could be combined with apyrase expression to drive increased glucose uptake, as found in chapter six.

7.2.3 Limitations of Constraint Based Modelling of *Clostridium*

One of the core strengths of FBA are the small amount of data required to build a metabolic model and to generate flux predictions (Orth, Thiele and Palsson, 2010). While base models may, as has been seen in this thesis, not accurately reproduce experimental data, they can provide a starting point for analysing behaviour satisfying various objective functions. The disadvantages of FBA are the inability to predict absolute concentrations of metabolites, the lack of regulatory effects and the prediction of fluxes only at metabolic steady state (Orth, Thiele and Palsson, 2010). As FBA gives a flux distribution that is optimised for a specific objective, the choice of objective function is of key importance when attempting to replicate experimental behaviour *in silico* (García Sánchez and Torres Sáez, 2014). The biomass objective function is commonly used, and for *E. coli* metabolic networks, has been shown to predict accurate values for substrate uptake and product secretion rates (Varma and Palsson, 1994; Jeremy S. Edwards, Ibarra and Palsson, 2001). As I found in chapter three, the objective function of maximising biomass formation did not reproduce the extracellular fluxes estimated from batch fermentation data of *C. saccharoperbutylacetonicum* N1-4 (chapter three) or *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T} (chapter four). This was because biomass formation was ATP limited, and acetate production was the most efficient method of generating

ATP. This has been found by other reconstructed models of *Clostridium*, including *C. cellulolyticum* (Salimi, Zhuang and Mahadevan, 2010), *C. beijerinckii* (Milne *et al.*, 2011), and *C. acetobutylicum* (Lee *et al.*, 2008).

The major problem with using FBA to model processes with distinct phases of growth, substrate uptake and/or product secretion is that FBA only provides a single flux distribution. As ABE fermentation in batch culture is a dynamic process (chapters three and four), steady-state FBA cannot completely replicate experimental behaviour. One attempt to model ABE fermentation involved splitting simulations into acidogenic and solventogenic phases (Lee *et al.*, 2008). The maximisation of biomass formation was used as the objective in the former phase, while in the latter a non-linear objective function of minimising the difference in fluxes between acidogenesis and solventogenesis, except for acid and solvent forming fluxes, was used (Lee *et al.*, 2008). This objective assumed that *C. acetobutylicum* minimises the change in metabolite concentrations between the two phases. The validity of this assumption could be called into question, as studies profiling the metabolome and the intracellular fluxes of *C. acetobutylicum* have shown a substantial shifting of both metabolite concentrations and metabolic fluxes, which were not limited to the ABE fermentation pathway (Amador-Noguez *et al.*, 2011; Crown *et al.*, 2011).

The obvious way to model the complete course of batch ABE fermentation would be to use a kinetic model. These models require more data to construct than stoichiometric models but give a more informative set of predictions. Kinetic models of *C. saccharoperbutylacetonicum* N1-4 have been published (Shinto *et al.*, 2007, 2008; Zhou, Liu and Yuan, 2018, 2020). These models have been used to explore how initial glucose concentration impacts butanol production (Shinto *et al.*, 2007), the impacts of exogenous lactate, acetate (Zhou, Liu and Yuan, 2018) and butyrate (Zhou, Liu and Yuan, 2020) on butanol titre, and have been used to suggest metabolic engineering targets.

In chapter four, I found that increasing the rate of glucose uptake in strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} would be necessary to increase butanol production. In one kinetic model of *C. saccharoperbutylacetonicum*, sensitivity analysis predicted that increasing the glucose uptake rate would increase butanol titre during the early stage of fermentation, but that the end-point butanol titre would be lower (Shinto *et al.*, 2007). This was also reported for a separate model of *C. saccharoperbutylacetonicum* (Zhou, Liu and Yuan, 2018). This again highlights the shortcomings of FBA, which does not take enzyme kinetics into account.

While the glucose uptake target in my FBA simulations disagreed with the kinetic models, my other targets were in better agreement. It was predicted that increasing flux from acetoacetyl-CoA to butyryl-CoA, reducing flux from butyryl-CoA to butyrate and increasing flux from butyryl-

CoA to butanol would increase the end butanol titre (Shinto *et al.*, 2007). Zhou *et al.* also performed a sensitivity analysis, with their model predicting that increasing conversion of glyceraldehyde-3-phosphate (G3P) to pyruvate and conversion of butyryl-CoA to butanol would increase butanol titre. In general, these predictions are in agreement with my simulations in chapter three, where increasing the flux from acetyl-CoA to butyryl-CoA was predicted to be key to increasing butanol formation, as well as blocking the formation of the acid forming pathways. My simulations captured the need for increased flux from phosphoenolpyruvate to pyruvate in chapter four when glucose uptake was increased (chapter four). Zhou *et al.*'s suggestion that increasing glyceraldehyde-3-phosphate to pyruvate flux would increase butanol titre mean that increased expression of pyruvate kinase alone might be sufficient to increase butanol titre.

Kinetic models do provide important insights into metabolism and are useful for metabolic engineering, but the shortcoming of the aforementioned models is their limited scope, only consisting of the ABE fermentation pathway and glycolysis (Shinto *et al.*, 2007; Zhou, Liu and Yuan, 2018, 2020). In contrast, FBA considers the entire metabolic network of the cell. The classical 'bottom-up' approach of constructing kinetic models involves characterising enzymes or pathways *in vitro*, which can be a technically challenging process (Saa and Nielsen, 2017). In contrast, the 'top-down' approach uses larger omics datasets to infer parameters (Saa and Nielsen, 2017), and this technique has recently been used to generate large scale kinetic models. Using the ensemble modelling (EM) method (Tran, Rizk and Liao, 2008), a kinetic model of *E. coli* was produced consisting of 457 reactions, which outperformed FBA predictions of various product yields in numerous mutant strains (Khodayari and Maranas, 2016). The same group used the EM method to construct a kinetic model of *C. thermocellum*, consisting of 118 reactions (Dash *et al.*, 2017). This model was used to model several scenarios, including nitrogen limitation and ethanol stress (Dash *et al.*, 2017). Both the *E. coli* and *C. thermocellum* model required a substantial amount of data to determine the parameters for the kinetic model. The *C. thermocellum* model relied on fermentation data of fermentation products for nineteen knockout strains for parameterisation (Dash *et al.*, 2017). The *E. coli* model was parameterised with steady state flux maps for nineteen knockout strains under varying growth conditions (Khodayari and Maranas, 2016). Unfortunately, datasets of this scale are not publicly available for *C. saccharoperbutylacetonicum* N1-4, which prevented me from using this kind of approach in this thesis.

Despite the shortcomings of FBA, it still provides a rapid method of analysing genome scale metabolic models for optimal behaviour. The many successes of FBA in metabolic engineering mean it will still be useful for developing metabolic engineering strategies *in silico*, particularly for less well characterised organisms. Where applicable, FBA simulations should be complemented

with kinetic models to gain a more detailed understanding of metabolic responses to changing enzyme levels.

7.3 Future Directions for Metabolic Engineering *Clostridium* for Butanol Production

C. saccharoperbutylacetonicum N1-4 *spoOA*^{I261T} is a strain that is superior to wild type *C. saccharoperbutylacetonicum* N1-4 as a chassis for industrial butanol production. While this thesis failed to generate a mutant strain which increased butanol production, this was likely because I was only able to implement single manipulations. The *in silico* guided strategies required multiple manipulation to be fully realised but were not fully implemented due to technical issues constructing plasmids and time constraints. The major issues with metabolic engineering *C. saccharoperbutylacetonicum* N1-4 *spoOA*^{I261T}, and *Clostridium* in general, is the lack of tools for rapid mutant generation (Pyne *et al.*, 2014). First and foremost, there is a need to develop a robust toolkit for the manipulation of *Clostridium*. A wider library of promoters, ribosome binding sites and terminators would enable more fine-tuned control of gene expression. Progress in this area is already underway, as a library of synthetic promoters was constructed and characterised in *C. acetobutylicum* through randomisation of the thiolase and glucuronidase promoters (Mordaka and Heap, 2018). A key area which would be a boon to scaling up industrial processes would be phase-dependent promoters, which do not require the addition of inducer molecules. This would keep fermentation costs low while ensuring coordinated gene expression at certain times of the culture (Nielsen and Keasling, 2016). This could be used to ensure that a certain biomass is reached before shifting cells from growth to a production phenotype.

Generating knockout strains of *Clostridium* is now more achievable than ever, with technologies such as ClosTron (Heap *et al.*, 2007, 2010) and CRISPR/Cas systems (Li *et al.*, 2016; Wang *et al.*, 2016; Wasels *et al.*, 2017, 2020; Atmadjaja *et al.*, 2019) being used in several diverse species. Other developments include gene knockdown technologies, such as the recently described small-regulatory RNA method (Cho and Lee, 2017) and CRISPR interference (Woolston *et al.*, 2018). Knockdown methods would prove useful for those genes whose removal is lethal, such as the hydrogenase gene in *C. acetobutylicum* (Cooksley *et al.*, 2012; Nguyen *et al.*, 2018).

Using these emerging manipulation technologies, a deeper understanding of genetic regulation in *Clostridium* would be invaluable in designing effective metabolic engineering strategies. There is some knowledge of various regulons in solventogenic *Clostridium* (Dürre *et al.*, 2002), but regulators of central metabolic pathways such as glycolysis, the pentose phosphate pathway and TCA cycle are poorly understood. In addition, many *Clostridium* have multiple isozymes of important proteins such as aldehyde/alcohol dehydrogenase (Dai *et al.*, 2016) and butanol dehydrogenase (Walter, Bennett and Papoutsakis, 1992). With high throughput knockout

technologies, it should be possible to identify key regulatory genes and to identify isozymes contributing to major enzymatic activities.

Generating a broad range of mutants through these technologies would be a first step in creating a public database of *Clostridium* mutants, in a similar vein to the ASAP database of *E. coli* mutants (Glasner *et al.*, 2003). Merely having a database of growth phenotypes and fermentation product yields would enable researchers to identify productive mutants and avoid generating mutants which have a detrimental impact on target product titre. Further information could be even more useful to the research community. Collecting various -omics datasets for mutant strains, while labour intensive, would be a huge boon to the *Clostridium* research community. Both transcriptomics and proteomics could be used to identify gene functions (Aslam *et al.*, 2017; Lowe *et al.*, 2017), the former has already been used in *C. beijerinckii* to confirm the role of various uncharacterised genes (Máté de Gérando *et al.*, 2018). Further insight could be gained through incorporating metabolomics and fluxomics datasets. Metabolomics data can be used to add additional constraints to stoichiometric models by estimating the net direction of a given reaction (Toya and Shimizu, 2013). Metabolic bottlenecks can also be identified based on accumulation of pathway intermediates (Toya and Shimizu, 2013). Fluxomics data would not only provide a detailed understanding of mutant phenotypes (Long and Antoniewicz, 2019) but could be used as training data for kinetic models, as performed by (Khodayari *et al.*, 2014; Khodayari and Maranas, 2016).

With a detailed understanding of enzyme kinetics and pathways, metabolic engineering of *Clostridium* can move towards both enzyme engineering and fine-tuned control of target enzymes. As an example of what can be achieved with large quantities of biological data, a recently published example described acetone and isopropanol production using recombinant *C. autoethanogenum* in a carbon negative production process (Liew *et al.*, 2022). Using a recently assembled collection of ABE producing *Clostridium* genome sequences to identify acetone synthesising genes, the authors generated 247 strains of *C. autoethanogenum* with unique acetone synthesis pathway designs. After initial studies with plasmid-based expression, the top performing acetone pathways were integrated into the genome to increase stability. Improvements in acetone titre were investigated using a genome scale metabolic model and an evolutionary algorithm to generate knockout targets. Overexpression targets were identified using both proteomic analysis of pathway enzymes, combined with kinetic modelling, identifying butyrate-acetoacetate CoA-transferase enzyme was limiting acetone formation. This was confirmed through cell free expression using the iPROBE system, before being implemented by genomic integration of an additional copy of the butyrate-acetoacetate CoA-transferase genes. The final strains were scaled up to field pilot reactors with high productivities. This study highlights the importance of combined approaches to

identifying knockout and overexpression targets, in combination with a large volume of data for pathway design and kinetic model construction.

Given the broad substrate range of *Clostridium*, these organisms can serve as important chassis organisms for future biotechnological applications. With the advances in genome editing, these organisms will soon be amenable to routine manipulation, and will likely be able to rival established model organisms as cellular factories. Given the current climate crisis, the importance of developing carbon neutral or, better yet, carbon negative processes to replace petrochemical production routes is vital. I hope that the findings from this thesis provide useful information for the metabolic engineering community in future, and a starting point for engineering *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} and solventogenic *Clostridium* in general.

References

- Adams, M.W.W. and Mortenson, L.E. (1984) 'The physical and catalytic properties of hydrogenase II of *Clostridium pasteurianum*. A comparison with hydrogenase I.', *Journal of Biological Chemistry*, 259(11), pp. 7045–7055. Available at: [https://doi.org/10.1016/S0021-9258\(17\)39835-6](https://doi.org/10.1016/S0021-9258(17)39835-6).
- Adams, M.W.W., Mortenson, L.E. and Chen, J.S. (1980) 'Hydrogenase', *Biochimica et Biophysica Acta (BBA) - Reviews on Bioenergetics*, 594(2–3), pp. 105–176. Available at: [https://doi.org/10.1016/0304-4173\(80\)90007-5](https://doi.org/10.1016/0304-4173(80)90007-5).
- Ahmad, A., Banat, F. and Taher, H. (2020) 'A review on the lactic acid fermentation from low-cost renewable materials: Recent developments and challenges', *Environmental Technology and Innovation*. Elsevier, p. 101138. Available at: <https://doi.org/10.1016/j.eti.2020.101138>.
- Al-Hinai, M.A., Fast, A.G. and Papoutsakis, E.T. (2012) 'Novel system for efficient isolation of clostridium double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration', *Applied and Environmental Microbiology*, 78(22), pp. 8112–8121. Available at: <https://doi.org/10.1128/AEM.02214-12>.
- Allcock, E.R. and Reid, S.J. (1982) 'Solvent Production and Morphological Changes in *Clostridium acetobutylicum* Solvent Production and Morphological Changes in *Clostridium acetobutylicum*', *Applied and Environmental Microbiology*, 43(6), pp. 1434–1439.
- Alper, H. *et al.* (2005) 'Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*', *Metabolic Engineering*, 7(3), pp. 155–164. Available at: <https://doi.org/10.1016/j.ymben.2004.12.003>.
- Alsaker, K. v. and Papoutsakis, E.T. (2005) 'Transcriptional program of early sporulation and stationary-phase events in *Clostridium acetobutylicum*', *Journal of Bacteriology*, 187(20), pp. 7103–7118. Available at: <https://doi.org/10.1128/JB.187.20.7103-7118.2005>.
- Alsaker, K. v, Spitzer, T.R. and Papoutsakis, E.T. (2004) 'Transcriptional Analysis of *spo0A* Overexpression in *Clostridium acetobutylicum* and Its Effect on the Cell ' s Response to Butanol Stress Transcriptional Analysis of *spo0A* Overexpression in *Clostridium acetobutylicum* and Its Effect on the Cell ' s Response', *Journal of Bacteriology*, 186(7), pp. 1959–1971. Available at: <https://doi.org/10.1128/JB.186.7.1959>.
- Al-Shorgani, N.K.N. *et al.* (2012) 'Bioconversion of butyric acid to butanol by clostridium saccharoperbutylacetonicum N1-4 (ATCC 13564) in a limited nutrient medium', *Bioenergy Research*, 5(2), pp. 287–293. Available at: <https://doi.org/10.1007/s12155-011-9126-6>.
- Amador-Noguez, D. *et al.* (2010) 'Systems-level metabolic flux profiling elucidates a complete, bifurcated tricarboxylic acid cycle in *Clostridium acetobutylicum*', *Journal of Bacteriology*, 192(17), pp. 4452–4461. Available at: <https://doi.org/10.1128/JB.00490-10>.
- Amador-Noguez, D. *et al.* (2011) 'Metabolome Remodeling during the Acidogenic-Solventogenic Transition in *Clostridium acetobutylicum*', *Applied and Environmental Microbiology*, 77(22), pp. 7984–7997. Available at: <https://doi.org/10.1128/AEM.05374-11>.
- de Anda, R. *et al.* (2006) 'Replacement of the glucose phosphotransferase transport system by galactose permease reduces acetate accumulation and improves process performance of *Escherichia coli* for recombinant protein production without impairment of growth rate',

Metabolic Engineering, 8(3), pp. 281–290. Available at:
<https://doi.org/10.1016/j.ymben.2006.01.002>.

Andersch, W., Bahl, H. and Gottschalk, G. (1982) 'Acetone-butanol production by *Clostridium acetobutylicum* in an ammonium-limited chemostat at low pH values', *Biotechnology Letters*, 4(1), pp. 29–32. Available at: <https://doi.org/10.1007/BF00139278>.

Andersch, W., Bahl, H. and Gottschalk, G. (1983) 'Level of enzymes involved in acetate, butyrate, acetone and butanol formation by *Clostridium acetobutylicum*', *European Journal of Applied Microbiology and Biotechnology*, 18(6), pp. 327–332. Available at:
<https://doi.org/10.1007/BF00504740>.

Annous, B.A. and Blaschek, H.P. (1991) 'Isolation and characterization of *Clostridium acetobutylicum* mutants with enhanced amylolytic activity.', *Applied and Environmental Microbiology*, 57(9).

Antoniewicz, M.R. *et al.* (2007) 'Metabolic flux analysis in a nonstationary system: Fed-batch fermentation of a high yielding strain of *E. coli* producing 1,3-propanediol', *Metabolic Engineering*, 9(3), pp. 277–292. Available at: <https://doi.org/10.1016/j.ymben.2007.01.003>.

Antoniewicz, M.R. (2015) 'Methods and advances in metabolic flux analysis: a mini-review', *Journal of Industrial Microbiology and Biotechnology*, 42(3), pp. 317–325. Available at:
<https://doi.org/10.1007/s10295-015-1585-x>.

Anwar, Z., Gulfraz, M. and Irshad, M. (2014) 'Agro-industrial lignocellulosic biomass a key to unlock the future bio-energy: A brief review', *Journal of Radiation Research and Applied Sciences*, 7(2), pp. 163–173. Available at: <https://doi.org/10.1016/j.jrras.2014.02.003>.

Asadollahi, M.A. *et al.* (2009) 'Enhancing sesquiterpene production in *Saccharomyces cerevisiae* through in silico driven metabolic engineering', *Metabolic Engineering*, 11(6), pp. 328–334. Available at: <https://doi.org/10.1016/J.YMBEN.2009.07.001>.

Aslam, B. *et al.* (2017) 'Proteomics: Technologies and Their Applications', *Journal of Chromatographic Science*, 55(2), pp. 182–196. Available at:
<https://doi.org/10.1093/CHROMSCI/BMW167>.

Atmadjaja, A.N. *et al.* (2019) 'CRISPR-Cas, a highly effective tool for genome editing in *Clostridium saccharoperbutylacetonicum* N1-4(HMT)', *FEMS Microbiology Letters*, 366(6), pp. 1–10. Available at: <https://doi.org/10.1093/femsle/fnz059>.

Atsumi, S. *et al.* (2008) 'Metabolic engineering of *Escherichia coli* for 1-butanol production', *Metabolic Engineering*, 10(6), pp. 305–311. Available at:
<https://doi.org/10.1016/j.ymben.2007.08.003>.

Au, J. *et al.* (2014) 'Parallel labeling experiments validate *Clostridium acetobutylicum* metabolic network model for ¹³C metabolic flux analysis', *Metabolic Engineering*, 26, pp. 23–33. Available at: <https://doi.org/10.1016/j.ymben.2014.08.002>.

Babu, M. *et al.* (2002) 'Shigella apyrase – a novel variant of bacterial acid phosphatases?', *FEBS Letters*, 512(1–3), pp. 8–12. Available at: [https://doi.org/10.1016/S0014-5793\(02\)02287-1](https://doi.org/10.1016/S0014-5793(02)02287-1).

Bahl, H. *et al.* (1982) 'Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium acetobutylicum* grown in continuous culture', *European Journal of Applied*

Microbiology and Biotechnology, 14(1), pp. 17–20. Available at:
<https://doi.org/10.1007/BF00507998>.

Bahl, H. *et al.* (1986) 'Nutritional Factors Affecting the Ratio of Solvents Produced By *Clostridium-Acetobutylicum*', *Applied and Environmental Microbiology*, 52(1), pp. 169–172.

Bahl, H., Andersch, W. and Gottschalk, G. (1982) 'Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a two-stage phosphate limited chemostat', *European Journal of Applied Microbiology and Biotechnology*, 15(4), pp. 201–205. Available at:
<https://doi.org/10.1007/BF00499955>.

Bailey, J.E. (1991) 'Toward a science of metabolic engineering', *Science*, 252(5013), pp. 1668–1675. Available at: <https://doi.org/10.1126/science.2047876>.

Barbeau, J.Y., Marchal, R. and Vandecasteele, J.P. (1988) 'Conditions promoting stability of solventogenesis or culture degeneration in continuous fermentations of *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology* 1988 29:5, 29(5), pp. 447–455. Available at: <https://doi.org/10.1007/BF00269067>.

Barin, R. *et al.* (2018) 'Enzymatic CO₂ reduction to formate by formate dehydrogenase from *Candida boidinii* coupling with direct electrochemical regeneration of NADH', *Journal of CO₂ Utilization*, 28, pp. 117–125. Available at: <https://doi.org/10.1016/J.JCOU.2018.09.020>.

Benisch, F. and Boles, E. (2014) 'The bacterial Entner–Doudoroff pathway does not replace glycolysis in *Saccharomyces cerevisiae* due to the lack of activity of iron–sulfur cluster enzyme 6-phosphogluconate dehydratase', *Journal of Biotechnology*, 171(1), pp. 45–55. Available at: <https://doi.org/10.1016/J.JBIOTEC.2013.11.025>.

Berg, J., Hung, Y.P. and Yellen, G. (2009) 'A genetically encoded fluorescent reporter of ATP:ADP ratio', *Nature Methods* 2009 6:2, 6(2), pp. 161–166. Available at:
<https://doi.org/10.1038/nmeth.1288>.

Berrios-Rivera, S.J. *et al.* (2004) 'Effect of different levels of NADH availability on metabolite distribution in *Escherichia coli* fermentation in minimal and complex media', *Applied Microbiology and Biotechnology*, 65(4), pp. 426–432. Available at: <https://doi.org/10.1007/s00253-004-1609-3>.

Berrios-Rivera, S.J., Bennett, G.N. and San, K.Y. (2002a) 'Metabolic engineering of *Escherichia coli*: Increase of NADH availability by overexpressing an NAD⁺-dependent formate dehydrogenase', *Metabolic Engineering*, 4(3), pp. 217–229. Available at: <https://doi.org/10.1006/mben.2002.0227>.

Berrios-Rivera, S.J., Bennett, G.N. and San, K.Y. (2002b) 'The effect of increasing NADH availability on the redistribution of metabolic fluxes in *Escherichia coli* chemostat cultures', *Metabolic Engineering*, 4(3), pp. 230–237. Available at: <https://doi.org/10.1006/mben.2002.0228>.

Birgen, C. *et al.* (2018) 'The Effect of Feeding Strategy on Butanol Production by *Clostridium beijerinckii* NCIMB 8052 using Glucose and Xylose', *Chemical Engineering Transactions*, 65, pp. 283–288. Available at: <https://doi.org/10.3303/CET1865048>.

Bordbar, A. *et al.* (2014) 'Constraint-based models predict metabolic and associated cellular functions', *Nature Reviews Genetics*, 15(2), pp. 107–120. Available at:
<https://doi.org/10.1038/nrg3643>.

Bowles, L.K. (1985) 'Effects of Butanol', *Cultures*, 50(5), pp. 1165–1170.

- Boynton, Z.L., Bennett, G.N. and Rudolph, F.B. (1996) 'Cloning, sequencing, and expression of genes encoding phosphotransacetylase and acetate kinase from *Clostridium acetobutylicum* ATCC 824', *Applied and Environmental Microbiology*, 62(8), pp. 2758–2766. Available at: <https://doi.org/10.1128/aem.62.8.2758-2766.1996>.
- Brochado, A.R. *et al.* (2010) 'Improved vanillin production in baker's yeast through in silico design', *Microbial Cell Factories* 2010 9:1, 9(1), pp. 1–15. Available at: <https://doi.org/10.1186/1475-2859-9-84>.
- Brodeur, G. *et al.* (2011) 'Chemical and physicochemical pretreatment of lignocellulosic biomass: A review', *Enzyme Research*, 2011(1). Available at: <https://doi.org/10.4061/2011/787532>.
- Brosseau, J.D., Yan, J.-Y. and Lo, K.V. (1986) 'The Relationship Between Hydrogen Gas and Butanol Production', *Biotechnology and Bioengineering*, 28, pp. 305–310.
- Bruder, M. *et al.* (2015) 'Elimination of carbon catabolite repression in *Clostridium acetobutylicum*—a journey toward simultaneous use of xylose and glucose', *Applied Microbiology and Biotechnology*, 99(18), pp. 7579–7588. Available at: <https://doi.org/10.1007/S00253-015-6611-4/FIGURES/3>.
- Buehler, E.A. and Mesbah, A. (2016) 'Kinetic study of acetone-butanol-ethanol fermentation in continuous culture', *PLoS ONE*, 11(8), pp. 1–21. Available at: <https://doi.org/10.1371/journal.pone.0158243>.
- Buendia-Kandia, F. *et al.* (2018) 'Diauxic growth of *Clostridium acetobutylicum* ATCC 824 when grown on mixtures of glucose and cellobiose', *AMB Express*, 8(1), pp. 1–9. Available at: <https://doi.org/10.1186/S13568-018-0615-2/FIGURES/2>.
- Burgard, A.P., Pharkya, P. and Maranas, C.D. (2003) 'OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization', *Biotechnology and Bioengineering*, 84(6), pp. 647–657. Available at: <https://doi.org/10.1002/bit.10803>.
- Cameron, P. *et al.* (2019) 'Harnessing type I CRISPR–Cas systems for genome engineering in human cells', *Nature Biotechnology* 2019 37:12, 37(12), pp. 1471–1477. Available at: <https://doi.org/10.1038/s41587-019-0310-0>.
- del Cerro, C. *et al.* (2013) 'Genome Sequence of the Butanol Hyperproducer *Clostridium saccharoperbutylacetonicum* N1-4.', *Genome announcements*, 1(2), p. e0007013. Available at: <https://doi.org/10.1128/genomeA.00070-13>.
- Chandel, N.S. (2021) 'Glycolysis', *Cold Spring Harbor Perspectives in Biology*, 13(5), p. a040535. Available at: <https://doi.org/10.1101/CSHPERSPECT.A040535>.
- Chao, Y.P. and Liao, J.C. (1994) 'Metabolic responses to substrate futile cycling in *Escherichia coli*.', *Journal of Biological Chemistry*, 269(7), pp. 5122–5126. Available at: [https://doi.org/10.1016/S0021-9258\(17\)37663-9](https://doi.org/10.1016/S0021-9258(17)37663-9).
- Chemler, J.A. *et al.* (2010) 'Improving NADPH availability for natural product biosynthesis in *Escherichia coli* by metabolic engineering', *Metabolic Engineering*, 12(2), pp. 96–104. Available at: <https://doi.org/10.1016/J.YMBEN.2009.07.003>.

- Chen, C. K. and Blaschek, H.P. (1999) 'Acetate enhances solvent production and prevents degeneration in *Clostridium beijerinckii* BA101', *Applied Microbiology and Biotechnology*, 52(2), pp. 170–173. Available at: <https://doi.org/10.1007/s002530051504>.
- Chen, Chih Kuang and Blaschek, H.P. (1999a) 'Effect of acetate on molecular and physiological aspects of *Clostridium beijerinckii* NCIMB 8052 solvent production and strain degeneration', *Applied and Environmental Microbiology*, 65(2), pp. 499–505. Available at: <https://doi.org/10.1128/aem.65.2.499-505.1999>.
- Chen, Chih Kuang and Blaschek, H.P. (1999b) 'Examination of physiological and molecular factors involved in enhanced solvent production by *Clostridium beijerinckii* BA101', *Applied and Environmental Microbiology*, 65(5), pp. 2269–2271.
- Chen, H. (2015) 'Lignocellulose biorefinery feedstock engineering', *Lignocellulose Biorefinery Engineering*, pp. 37–86. Available at: <https://doi.org/10.1016/B978-0-08-100135-6.00003-X>.
- Chen, X. *et al.* (2011) 'Synergy between ¹³C-metabolic flux analysis and flux balance analysis for understanding metabolic adaption to anaerobiosis in *E. coli*', *Metabolic Engineering*, 13(1), pp. 38–48. Available at: <https://doi.org/10.1016/j.ymben.2010.11.004>.
- Cho, C. and Lee, S.Y. (2017) 'Efficient gene knockdown in *Clostridium acetobutylicum* by synthetic small regulatory RNAs', *Biotechnology and Bioengineering*, 114(2), pp. 374–383. Available at: <https://doi.org/10.1002/bit.26077>.
- Cho, D.H. *et al.* (2009) 'Detoxification of model phenolic compounds in lignocellulosic hydrolysates with peroxidase for butanol production from *Clostridium beijerinckii*', *Applied Microbiology and Biotechnology*, 83(6), pp. 1035–1043. Available at: <https://doi.org/10.1007/s00253-009-1925-8>.
- Cho, D.H., Shin, S.J. and Kim, Y.H. (2012) 'Effects of acetic and formic acid on ABE production by *Clostridium acetobutylicum* and *Clostridium beijerinckii*', *Biotechnology and Bioengineering*, 17(2), pp. 270–275. Available at: <https://doi.org/10.1007/s12257-011-0498-4>.
- Choi, J. and Lee, S.Y. (1999) 'Factors affecting the economics of polyhydroxyalkanoate production by bacterial fermentation', *Applied Microbiology and Biotechnology* 51:1, 51(1), pp. 13–21. Available at: <https://doi.org/10.1007/S002530051357>.
- Choi, K.R. *et al.* (2016) 'Systems Metabolic Engineering of *Escherichia coli*', *EcoSal Plus*, 7(1). Available at: <https://doi.org/10.1128/ECOSALPLUS.ESP-0010-2015>
- Clark, S.W., Bennett, G.N. and Rudolph, F.B. (1989) 'Isolation and Characterization of Mutants of *Clostridium acetobutylicum* ATCC 824 Deficient in Acetoacetyl-Coenzyme A:Acetate/Butyrate:Coenzyme A-Transferase (EC 2.8.3.9) and in Other Solvent Pathway Enzymes', *Applied and Environmental Microbiology*, 55(4).
- Colby, G.D. and Chen, J.S. (1992) 'Purification and properties of 3-hydroxybutyryl-coenzyme A dehydrogenase from *Clostridium beijerinckii* ('*Clostridium butylicum*') NRRL B593', *Applied and Environmental Microbiology*, 58(10), pp. 3297–3302.
- Cong, L. *et al.* (2013) 'Multiplex genome engineering using CRISPR/Cas systems', *Science*, 339(6121), pp. 819–823. Available at: https://doi.org/10.1126/SCIENCE.1231143/SUPPL_FILE/PAPV2.PDF.

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

Cornillot, E. *et al.* (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), pp. 5442–5447. Available at: <https://doi.org/10.1128/jb.179.17.5442-5447.1997>.

Covert, M.W. *et al.* (2001) 'Regulation of gene expression in flux balance models of metabolism', *Journal of Theoretical Biology*, 213(1), pp. 73–88. Available at: <https://doi.org/10.1006/jtbi.2001.2405>.

Crain, A. v. and Broderick, J.B. (2014) 'Pyruvate formate-lyase and its activation by pyruvate formate-lyase activating enzyme', *Journal of Biological Chemistry*, 289(9), pp. 5723–5729. Available at: <https://doi.org/10.1074/jbc.M113.496877>.

Cray, J.A. *et al.* (2015) 'Chaotropicity: a key factor in product tolerance of biofuel-producing microorganisms', *Current Opinion in Biotechnology*, 33, pp. 228–259. Available at: <https://doi.org/10.1016/J.COPBIO.2015.02.010>.

Croux, C. *et al.* (2016) 'Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*', *Biotechnology for Biofuels*, 9(1), pp. 1–13. Available at: <https://doi.org/10.1186/s13068-016-0432-2>.

Crown, S.B. *et al.* (2011) 'Resolving the TCA cycle and pentose-phosphate pathway of *Clostridium acetobutylicum* ATCC 824: Isotopomer analysis, in vitro activities and expression analysis', *Biotechnology Journal*, 6(3), pp. 300–305. Available at: <https://doi.org/10.1002/biot.201000282>.

Crown, S.B. and Antoniewicz, M.R. (2013) 'Publishing ¹³C metabolic flux analysis studies: A review and future perspectives', *Metabolic Engineering*, 20, pp. 42–48. Available at: <https://doi.org/10.1016/j.ymben.2013.08.005>.

Cui, J. and Davidson, A.L. (2011) 'ABC solute importers in bacteria', *Essays in Biochemistry*, 50(1), pp. 85–99. Available at: <https://doi.org/10.1042/BSE0500085/78319>.

Curran, K.A., Crook, N.C. and Alper, H.S. (2012) 'Using Flux Balance Analysis to Guide Microbial Metabolic Engineering', *Methods in Molecular Biology*, 834, pp. 197–216. Available at: https://doi.org/10.1007/978-1-61779-483-4_13.

Dabrock, B., Bahl, H. and Gottschalk, G. (1992) 'Parameters Affecting Solvent Production by *Clostridium pasteurianum*', *Applied and Environmental Microbiology*, 58(4), pp. 1233–1239.

Dai, Z. *et al.* (2016) 'Elucidating the contributions of multiple aldehyde/alcohol dehydrogenases to butanol and ethanol production in *Clostridium acetobutylicum*', *Scientific Reports* 2016 6:1, 6(1), pp. 1–9. Available at: <https://doi.org/10.1038/srep28189>.

Das, M., Patra, P. and Ghosh, A. (2020) 'Metabolic engineering for enhancing microbial biosynthesis of advanced biofuels', *Renewable and Sustainable Energy Reviews*, 119, p. 109562. Available at: <https://doi.org/10.1016/J.RSER.2019.109562>.

- Dash, S. *et al.* (2017) 'Development of a core *Clostridium thermocellum* kinetic metabolic model consistent with multiple genetic perturbations', *Biotechnology for Biofuels*, 10(1), pp. 1–16. Available at: <https://doi.org/10.1186/s13068-017-0792-2>.
- Datta, R. and Zeikus, J.G. (1985) 'Modulation of acetone-butanol-ethanol fermentation by carbon monoxide and organic acids', *Applied and Environmental Microbiology*, 49(3), pp. 522–529. Available at: <https://doi.org/10.1128/AEM.49.3.522-529.1985>.
- Demmer, J.K. *et al.* (2017) 'The semiquinone swing in the bifurcating electron transferring flavoprotein/butyryl-CoA dehydrogenase complex from *Clostridium difficile*', *Nature Communications*, 8(1), pp. 1–10. Available at: <https://doi.org/10.1038/s41467-017-01746-3>.
- Desai, R.P., Harris, L.M. and Welker, N.E. (1999) 'Metabolic Flux Analysis Elucidates the Importance of the Acid-Formation Pathways in Regulating Solvent.pdf', *Metabolic Engineering*, 213, pp. 206–213.
- Desai, R.P., Nielsen, L.K. and Papoutsakis, E.T. (1999) 'Stoichiometric modeling of *Clostridium acetobutylicum* fermentations with non-linear constraints', *Journal of Biotechnology*, 71(1–3), pp. 191–205. Available at: [https://doi.org/10.1016/S0168-1656\(99\)00022-X](https://doi.org/10.1016/S0168-1656(99)00022-X).
- Desai, S.H. and Atsumi, S. (2013) 'Photosynthetic approaches to chemical biotechnology', *Current Opinion in Biotechnology*, pp. 1031–1036. Available at: <https://doi.org/10.1016/j.copbio.2013.03.015>.
- Diallo, M., Kengen, S.W.M. and López-Contreras, A.M. (2021) 'Sporulation in solventogenic and acetogenic clostridia', *Applied Microbiology and Biotechnology* 2021 105:9, 105(9), pp. 3533–3557. Available at: <https://doi.org/10.1007/S00253-021-11289-9>.
- Diez-Gonzalez, F., Russell, J.B. and Hunter, J.B. (1995) 'The role of an NAD-independent lactate dehydrogenase and acetate in the utilization of lactate by *Clostridium acetobutylicum* strain P262', *Archives of Microbiology*, 164(1), pp. 36–42. Available at: <https://doi.org/10.1007/BF02568732>.
- Dong, H. *et al.* (2017) 'A systematically chromosomally engineered *Escherichia coli* efficiently produces butanol', *Metabolic Engineering*, 44(1), pp. 284–292. Available at: <https://doi.org/10.1016/j.ymben.2017.10.014>.
- Doremus, M.G., Linden, J.C. and Moreira, A.R. (1985) 'Agitation and pressure effects on acetone-butanol fermentation', *Biotechnology and Bioengineering*, 27(6), pp. 852–860. Available at: <https://doi.org/10.1002/BIT.260270615>.
- Du, G. *et al.* (2021) 'Energy-efficient butanol production by *Clostridium acetobutylicum* with histidine kinase knockouts to improve strain tolerance and process robustness', *Green Chemistry*, 23(5), pp. 2155–2168. Available at: <https://doi.org/10.1039/D0GC03993D>.
- Dürre, P. *et al.* (2002) 'Transcriptional regulation of solventogenesis in *Clostridium acetobutylicum*', in *Journal of Molecular Microbiology and Biotechnology*, pp. 295–300. Available at: <https://pubmed.ncbi.nlm.nih.gov/11931561/>
- Dürre, P. and Hollergschwandner, C. (2004) 'Initiation of endospore formation in *Clostridium acetobutylicum*', *Anaerobe*, 10(2), pp. 69–74. Available at: <https://doi.org/10.1016/j.anaerobe.2003.11.001>.

- Ebert, B.E. *et al.* (2011) 'Response of *Pseudomonas putida* KT2440 to increased NADH and ATP demand', *Applied and Environmental Microbiology*, 77(18), pp. 6597–6605. Available at: <https://doi.org/10.1128/AEM.05588-11>.
- Ebrahim, A. *et al.* (2013) 'COBRApy: CONstraints-Based Reconstruction and Analysis for Python', *BMC Systems Biology*, 7. Available at: <https://doi.org/10.1186/1752-0509-7-74>.
- Edwards, J. S., Ibarra, R.U. and Palsson, B.O. (2001) 'In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data', *Nature Biotechnology*, 19(2), pp. 125–130. Available at: <https://doi.org/10.1038/84379>.
- Egbumokei, C. (2021) 'UK Environmental Accounts - Office for National Statistics', *UK Environmental Accounts: 2021* [Preprint]. Available at: <https://www.ons.gov.uk/economy/environmentalaccounts/bulletins/ukenvironmentalaccounts/2021> (Accessed: 26 April 2022).
- Ellis, J.T. *et al.* (2012) 'Acetone, butanol, and ethanol production from wastewater algae', *Bioresource Technology*, 111, pp. 491–495. Available at: <https://doi.org/10.1016/j.biortech.2012.02.002>.
- Emmerling, M. *et al.* (2002) 'Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*', *Journal of Bacteriology*, 184(1), pp. 152–164. Available at: <https://doi.org/10.1128/JB.184.1.152-164.2002>.
- Essalem, M.E.E. and Mitchell, W.J. (2016) 'Identification of a glucose–mannose phosphotransferase system in *Clostridium beijerinckii*', *FEMS Microbiology Letters*, 363(8), p. 53. Available at: <https://doi.org/10.1093/FEMSLE/FNW053>.
- Ezeji, T. *et al.* (2010) 'Achievements and perspectives to overcome the poor solvent resistance in acetone and butanol-producing microorganisms', *Applied Microbiology and Biotechnology*, 85(6), pp. 1697–1712. Available at: <https://doi.org/10.1007/s00253-009-2390-0>.
- Ezeji, T., Qureshi, N. and Blaschek, H.P. (2007) 'Butanol production from agricultural residues: Impact of degradation products on *Clostridium beijerinckii* growth and butanol fermentation', *Biotechnology and Bioengineering*, 97(6), pp. 1460–1469. Available at: <https://doi.org/10.1002/bit.21373>.
- Ezeji, T.C., Qureshi, N. and Blaschek, H.P. (2004) 'Butanol fermentation research: upstream and downstream manipulations', *The Chemical Record*, 4(5), pp. 305–314. Available at: <https://doi.org/10.1002/TCR.20023>.
- Ezeji, T.C., Qureshi, N. and Blaschek, H.P. (2007) 'Production of acetone butanol (AB) from liquefied corn starch, a commercial substrate, using *Clostridium beijerinckii* coupled with product recovery by gas stripping', *Journal of Industrial Microbiology and Biotechnology*, 34(12), pp. 771–777. Available at: <https://doi.org/10.1007/s10295-007-0253-1>.
- Famili, I. *et al.* (2003) 'Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network', *Proceedings of the National Academy of Sciences*, 100(23), pp. 13134–13139. Available at: <https://doi.org/10.1073/pnas.2235812100>.

Farmanbordar, S., Karimi, K. and Amiri, H. (2018) 'Municipal solid waste as a suitable substrate for butanol production as an advanced biofuel', *Energy Conversion and Management*, 157(December 2017), pp. 396–408. Available at: <https://doi.org/10.1016/j.enconman.2017.12.020>.

Feist, A.M. *et al.* (2007) 'A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information', *Molecular Systems Biology*, 3(1), p. 121. Available at: <https://doi.org/10.1038/MSB4100155>.

Fernández-Naveira, Á. *et al.* (2016) 'Efficient butanol-ethanol (B-E) production from carbon monoxide fermentation by Clostridium carboxidivorans', *Applied Microbiology and Biotechnology*, 100(7), pp. 3361–3370. Available at: <https://doi.org/10.1007/s00253-015-7238-1>.

Finn, R.K. and Nowrey, J.E. (1959) 'A Note on the Stability of Clostridia when Held in Continuous Culture', *Applied Microbiology*, 7(1), pp. 29–32. Available at: <https://doi.org/10.1128/am.7.1.29-32.1959>.

Fischer, R.J., Helms, J. and Durre, P. (1993) 'Cloning, sequencing, and molecular analysis of the sol operon of Clostridium acetobutylicum, a chromosomal locus involved in solventogenesis', *Journal of Bacteriology*, 175(21), pp. 6959–6969. Available at: <https://doi.org/10.1128/JB.175.21.6959-6969.1993>.

Flores, N. *et al.* (1996) 'Pathway engineering for the production of aromatic compounds in Escherichia coli', *Nature Biotechnology*, 14(5), pp. 620–623. Available at: <https://doi.org/10.1038/nbt0596-620>.

Foley, P.M., Beach, E.S. and Zimmerman, J.B. (2011) 'Algae as a source of renewable chemicals: opportunities and challenges', *Green Chemistry*, 13(6), pp. 1399–1405. Available at: <https://doi.org/10.1039/C1GC00015B>.

Fong, S.S. *et al.* (2005) 'In silico design and adaptive evolution of Escherichia coli for production of lactic acid', *Biotechnology and Bioengineering*, 91(5), pp. 643–648. Available at: <https://doi.org/10.1002/bit.20542>.

Fontaine, L. *et al.* (2002) 'Molecular Characterization and Transcriptional Analysis of adhE2, the Gene Encoding the NADH-Dependent Aldehyde/Alcohol Dehydrogenase Responsible for Butanol Production in Alcohologenic Cultures of Clostridium acetobutylicum ATCC 824', *Journal of Bacteriology*, 184(3), pp. 821–830. Available at: <https://doi.org/10.1128/JB.184.3.821>.

Formanek, J., Mackie, R. and Blaschek, H.P. (1997) 'Enhanced butanol production by Clostridium beijerinckii BA101 grown in semidefined P2 medium containing 6 percent maltodextrin or glucose', *Applied and Environmental Microbiology*, 63(6), pp. 2306–2310. Available at: <https://doi.org/10.1128/AEM.63.6.2306-10>.

Förster, J. *et al.* (2003) 'Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network', *Genome Research*, 13(2), pp. 244–253. Available at: <https://doi.org/10.1101/gr.234503>.

Foster, C.J. *et al.* (2021) 'Building kinetic models for metabolic engineering', *Current Opinion in Biotechnology*, 67, pp. 35–41. Available at: <https://doi.org/10.1016/j.copbio.2020.11.010>.

Fowler, Z.L., Gikandi, W.W. and Koffas, M.A.G. (2009) 'Increased malonyl coenzyme A biosynthesis by tuning the Escherichia coli metabolic network and its application to flavanone production', *Applied and Environmental Microbiology*, 75(18), pp. 5831–5839. Available at: <https://doi.org/10.1128/AEM.00270-09>.

- Fu, H. *et al.* (2021) 'Butanol production from *Saccharina japonica* hydrolysate by engineered *Clostridium tyrobutyricum*: The effects of pretreatment method and heat shock protein overexpression', *Bioresource Technology*, 335, p. 125290. Available at: <https://doi.org/10.1016/j.biortech.2021.125290>.
- Gangl, D. *et al.* (2015) 'Biotechnological exploitation of microalgae', *Journal of Experimental Botany*, 66(22), pp. 6975–6990. Available at: <https://doi.org/10.1093/JXB/ERV426>.
- García Martín, H. *et al.* (2015) 'A Method to Constrain Genome-Scale Models with ¹³C Labeling Data', *PLOS Computational Biology*, 11(9), p. e1004363. Available at: <https://doi.org/10.1371/JOURNAL.PCBI.1004363>.
- García Sánchez, C.E. and Torres Sáez, R.G. (2014) 'Comparison and analysis of objective functions in flux balance analysis', *Biotechnology Progress*, 30(5), pp. 985–991. Available at: <https://doi.org/10.1002/btpr.1949>.
- George, H.A. and Chen, J.-S. (1983) 'Acidic Conditions Are Not Obligatory for Onset of Butanol Formation by *Clostridium beijerinckii* (Synonym, *C. butylicum*)', *Applied and Environmental Microbiology*, 46(2), pp. 321–327. Available at: <https://doi.org/10.1128/AEM.46.2.321-327.1983>.
- George, H.A., Johnson, J.L. and Moore, W.E.C. (1983) 'Acetone, isopropanol, and butanol production by *Clostridium beijerinckii* (syn. *Clostridium butylicum*) and *Clostridium aurantibutyricum*', *Applied and Environmental Microbiology*, 45(3), pp. 1160–1163. Available at: <https://doi.org/10.1128/aem.45.3.1160-1163.1983>.
- Ghosh, A. *et al.* (2016) '¹³C metabolic flux analysis for systematic metabolic engineering of *S. cerevisiae* for overproduction of fatty acids', *Frontiers in Bioengineering and Biotechnology*, 4(8), p. 76. Available at: <https://doi.org/10.3389/fbioe.2016.00076>.
- Gianchandani, E.P., Chavali, A.K. and Papin, J.A. (2010) 'The application of flux balance analysis in systems biology', *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2(3), pp. 372–382. Available at: <https://doi.org/10.1002/wsbm.60>.
- Girbal, L., Vasconcelos, I., *et al.* (1995) 'How neutral red modified carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH', *FEMS Microbiology Reviews*, 16(2–3), pp. 151–162. Available at: <https://doi.org/10.1111/J.1574-6976.1995.TB00163.X>.
- Girbal, L., Croux, C., *et al.* (1995) 'Regulation of metabolic shifts in *Clostridium acetobutylicum* ATCC 824', *FEMS Microbiology Reviews*, 17(3), pp. 287–297. Available at: <https://doi.org/10.1111/j.1574-6976.1995.tb00212.x>.
- Girbal, L. and Soucaille, P. (1994a) 'Regulation of *Clostridium acetobutylicum* metabolism as revealed by mixed- substrate steady-state continuous cultures: Role of NADH/NAD ratio and ATP pool', *Journal of Bacteriology*, 176(21), pp. 6433–6438. Available at: <https://doi.org/10.1128/jb.176.21.6433-6438.1994>.
- Girbal, L. and Soucaille, P. (1994b) 'Regulation of *Clostridium acetobutylicum* metabolism as revealed by mixed- substrate steady-state continuous cultures: Role of NADH/NAD ratio and ATP pool', *Journal of Bacteriology*, 176(21), pp. 6433–6438. Available at: <https://doi.org/10.1128/jb.176.21.6433-6438.1994>.

- Girbal, L., Vasconcelos, I. and Soucaille, P. (1994) 'Transmembrane pH of *Clostridium acetobutylicum* is inverted (more acidic inside) when the in vivo activity of hydrogenase is decreased', *Journal of Bacteriology*, 176(19), pp. 6146–6147. Available at: <https://doi.org/10.1128/JB.176.19.6146-6147.1994>.
- Glasner, J.D. *et al.* (2003) 'ASAP, a systematic annotation package for community analysis of genomes', *Nucleic Acids Research*, 31(1), p. 147. Available at: <https://doi.org/10.1093/NAR/GKG125>.
- Gokarn, R.R., Eiteman, M.A. and Altman, E. (2000) 'Metabolic analysis of *Escherichia coli* in the presence and absence of the carboxylating enzymes phosphoenolpyruvate carboxylase and pyruvate carboxylase', *Applied and Environmental Microbiology*, 66(5), pp. 1844–1850. Available at: <https://doi.org/10.1128/AEM.66.5.1844-1850.2000>.
- Gonzalez, J.E., Long, C.P. and Antoniewicz, M.R. (2017) 'Comprehensive analysis of glucose and xylose metabolism in *Escherichia coli* under aerobic and anaerobic conditions by ¹³C metabolic flux analysis', *Metabolic Engineering*, 39, pp. 9–18. Available at: <https://doi.org/10.1016/j.ymben.2016.11.003>.
- Gosset, G. (2005) 'Improvement of *Escherichia coli* production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system', *Microbial Cell Factories*, 4(1), pp. 1–11. Available at: <https://doi.org/10.1186/1475-2859-4-14>.
- Gottschal, J.C. and Morris, J.G. (1981) 'Non-production of acetone and butanol by *Clostridium acetobutylicum* during glucose- and ammonium- limitation in continuous culture', *Biotechnology Letters* 1981 3:9, 3(9), pp. 525–530. Available at: <https://doi.org/10.1007/BF00147566>.
- Gowen, C.M. and Fong, S.S. (2010) 'Genome-scale metabolic model integrated with RNAseq data to identify metabolic states of *Clostridium thermocellum*', *Biotechnology Journal*, 5(7), pp. 759–767. Available at: <https://doi.org/10.1002/biot.201000084>.
- Green, E.M. *et al.* (1996) 'Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824', *Microbiol.* 142:2079-2086, (May), pp. 2079–2086.
- Green, E.M. and Bennett, G.N. (1996) 'Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824', *Applied Biochemistry and Biotechnology*, 57–58(1), pp. 213–221. Available at: <https://doi.org/10.1007/BF02941702>.
- Grimmler, C. *et al.* (2010) 'Transcriptional analysis of catabolite repression in *Clostridium acetobutylicum* growing on mixtures of d-glucose and d-xylose', *Journal of Biotechnology*, 150(3), pp. 315–323. Available at: <https://doi.org/10.1016/j.jbiotec.2010.09.938>.
- Grupe, H. and Gottschalk, G. (1992) 'Physiological events in *Clostridium acetobutylicum* during the shift from acidogenesis to solventogenesis in continuous culture and presentation of a model for shift induction', *Applied and Environmental Microbiology*, 58(12), pp. 3896–3902.
- Gu, Y. *et al.* (2009) 'Improvement of xylose utilization in *Clostridium acetobutylicum* via expression of the talA gene encoding transaldolase from *Escherichia coli*', *Journal of Biotechnology*, 143(4), pp. 284–287. Available at: <https://doi.org/10.1016/J.JBIOTEC.2009.08.009>.
- Gu, Y. *et al.* (2019) 'Curing the endogenous megaplasmid in *Clostridium saccharoperbutylacetonicum* N1-4 (HMT) using CRISPR-Cas9 and preliminary investigation of the

role of the plasmid for the strain metabolism', *Fuel*, 236(January 2018), pp. 1559–1566. Available at: <https://doi.org/10.1016/j.fuel.2018.09.030>.

Guðmundsson, S. and Thiele, I. (2010) 'Computationally efficient flux variability analysis', *BMC Bioinformatics*, 11(1), pp. 1–3. Available at: <https://doi.org/10.1186/1471-2105-11-489/TABLES/2>.

Guerrero, K. *et al.* (2021) 'Continuous biohydrogen production by a degenerated strain of *Clostridium acetobutylicum* ATCC 824', *International Journal of Hydrogen Energy*, 46(7), pp. 5100–5111. Available at: <https://doi.org/10.1016/J.IJHYDENE.2020.11.104>.

Gustavsson, M. and Lee, S.Y. (2016) 'Prospects of microbial cell factories developed through systems metabolic engineering', *Microbial Biotechnology*, 9(5), pp. 610–617. Available at: <https://doi.org/10.1111/1751-7915.12385>.

Gutierrez, N.A. *et al.* (1998) 'Strain comparison and medium preparation for the acetone-butanol-ethanol (ABE) fermentation process using a substrate of potato', *Bioresource Technology*, 66(3), pp. 263–265. Available at: [https://doi.org/10.1016/S0960-8524\(98\)00051-0](https://doi.org/10.1016/S0960-8524(98)00051-0).

Hädicke, O., Bettenbrock, K. and Klamt, S. (2015) 'Enforced ATP futile cycling increases specific productivity and yield of anaerobic lactate production in *Escherichia coli*', *Biotechnology and Bioengineering*, 112(10), pp. 2195–2199. Available at: <https://doi.org/10.1002/BIT.25623>.

Harris, L. *et al.* (2001) 'Fermentation characterization and flux analysis of recombinant strains of *Clostridium acetobutylicum* with an inactivated *solR* gene', *Journal of Industrial Microbiology and Biotechnology*, 27(5), pp. 322–328. Available at: <https://doi.org/10.1038/sj.jim.7000191>.

Harris, L.M., Welker, N.E. and Papoutsakis, E.T. (2002) 'Northern, morphological, and fermentation analysis of *[i]spo0A[i]* inactivation and overexpression in *[i]Clostridium acetobutylicum[i]* ATCC 824', *J Bacteriol*, 184(13), pp. 3586–3597. Available at: <https://doi.org/10.1128/JB.184.13.3586-3597.2002>.

Hartmanis, M.G.N. (1987) 'Butyrate kinase from *Clostridium acetobutylicum*.', *Journal of Biological Chemistry*, 262(2), pp. 617–621. Available at: [https://doi.org/10.1016/S0021-9258\(19\)75828-1](https://doi.org/10.1016/S0021-9258(19)75828-1).

Hartmanis, M.G.N., Klason, T. and Gatenbeck, S. (1984) 'Uptake and activation of acetate and butyrate in *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 20(1), pp. 66–71. Available at: <https://doi.org/10.1007/BF00254648>.

He, L. *et al.* (2014) 'Central metabolic responses to the overproduction of fatty acids in *Escherichia coli* based on ¹³C-metabolic flux analysis', *Biotechnology and Bioengineering*, 111(3), pp. 575–585. Available at: <https://doi.org/10.1002/bit.25124>.

Heap, J.T. *et al.* (2007) 'The ClosTron: A universal gene knock-out system for the genus *Clostridium*', *Journal of Microbiological Methods*, 70(3), pp. 452–464. Available at: <https://doi.org/10.1016/j.mimet.2007.05.021>.

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

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

- Heap, J.T. *et al.* (2012a) 'Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker', *Nucleic Acids Research*, 40(8), pp. e59–e59. Available at: <https://doi.org/10.1093/NAR/GKR1321>.
- Heap, J.T. *et al.* (2012b) 'Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker', *Nucleic Acids Research*, 40(8), pp. e59–e59. Available at: <https://academic.oup.com/nar/article/40/8/e59/2411801>.
- Heckman, K.L. and Pease, L.R. (2007) 'Gene splicing and mutagenesis by PCR-driven overlap extension', *Nature Protocols*, 2(4), pp. 924–932. Available at: <https://doi.org/10.1038/nprot.2007.132>.
- Heirendt, L. *et al.* (2019) 'Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0', *Nature Protocols* 2019 14:3, 14(3), pp. 639–702. Available at: <https://doi.org/10.1038/s41596-018-0098-2>.
- Hermann, M. *et al.* (1985) 'Isolation and characterization of butanol-resistant mutants of *Clostridium acetobutylicum*.' , *Applied and Environmental Microbiology*, 50(5)., pp. 1238-1243.
- Hernández-Montalvo, V. *et al.* (2003) 'Expression of galP and glk in a *Escherichia coli* PTS mutant restores glucose transport and increases glycolytic flux to fermentation products', *Biotechnology and Bioengineering*, 83(6), pp. 687–694. Available at: <https://doi.org/10.1002/bit.10702>.
- Higo, A. and Ehira, S. (2019) 'Anaerobic butanol production driven by oxygen-evolving photosynthesis using the heterocyst-forming multicellular cyanobacterium *Anabaena* sp. PCC 7120', *Applied Microbiology and Biotechnology*, 103(5), pp. 2441–2447. Available at: <https://doi.org/10.1007/S00253-019-09635-Z>.
- Hoch, J.A. (1993) 'Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*', *Annual Review of Microbiology*, 47(1), pp. 441–465. Available at: <https://doi.org/10.1146/annurev.mi.47.100193.002301>.
- Holt, R.A, Stephens, G.M. and Morris, J.G. (1984) 'Production of Solvents by *Clostridium acetobutylicum* Cultures Maintained at Neutral pH', *Applied and environmental microbiology*, 48(6), pp. 1166–1170.
- Hu, S. *et al.* (2011) 'Comparative genomic and transcriptomic analysis revealed genetic characteristics related to solvent formation and xylose utilization in *Clostridium acetobutylicum* EA 2018', *BMC Genomics*, 12. Available at: <https://doi.org/10.1186/1471-2164-12-93>.
- Huang, H.J., Ramaswamy, S. and Liu, Y. (2014) 'Separation and purification of biobutanol during bioconversion of biomass', *Separation and Purification Technology*, 132, pp. 513–540. Available at: <https://doi.org/10.1016/J.SEPPUR.2014.06.013>.
- Huang, L., Gibbins, L.N. and Forsberg, C.W. (1985) 'Transmembrane pH gradient and membrane potential in *Clostridium acetobutylicum* during growth under acetogenic and solventogenic conditions', *Applied and Environmental Microbiology*, 50(4), pp. 1043–1047.
- Humphreys, C.M. *et al.* (2015) 'Whole genome sequence and manual annotation of *Clostridium autoethanogenum*, an industrially relevant bacterium', *BMC Genomics*, 16(1), pp. 1–10. Available at: <https://doi.org/10.1186/S12864-015-2287-5/FIGURES/3>.

- Hüsemann, M.H.W. and Papoutsakis, E.T. (1988) 'Solventogenesis in *Clostridium acetobutylicum* fermentations related to carboxylic acid and proton concentrations', *Biotechnology and Bioengineering*, 32(7), pp. 843–852. Available at: <https://doi.org/10.1002/BIT.260320702>.
- Hutkins, R.W. and Kashket, E.R. (1986) 'Phosphotransferase Activity in *Clostridium acetobutylicum* from Acidogenic and Solventogenic Phases of Growth.', *Applied and environmental microbiology*, 51(5), pp. 1121–3. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16347058> (Accessed: 8 March 2020).
- Ibarra, R.U., Edwards, J.S. and Palsson, B.O. (2002) 'Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth', *Nature*, 420(6912), pp. 186–189. Available at: <https://doi.org/10.1038/nature01149>.
- Inui, M. *et al.* (2008) 'Expression of *Clostridium acetobutylicum* butanol synthetic genes in *Escherichia coli*', *Applied Microbiology and Biotechnology* 77:6, 77(6), pp. 1305–1316. Available at: <https://doi.org/10.1007/S00253-007-1257-5>.
- Islam, M.M., Schroeder, W.L. and Saha, R. (2021) 'Kinetic modeling of metabolism: Present and future', *Current Opinion in Systems Biology*, 26, pp. 72–78. Available at: <https://doi.org/10.1016/J.COISB.2021.04.003>.
- Jahreis, K. *et al.* (2008) 'Ins and outs of glucose transport systems in eubacteria', *FEMS Microbiology Reviews*, 32(6), pp. 891–907. Available at: <https://doi.org/10.1111/J.1574-6976.2008.00125.X>.
- Jang, M.O. and Choi, G. (2018) 'Techno-economic analysis of butanol production from lignocellulosic biomass by concentrated acid pretreatment and hydrolysis plus continuous fermentation', *Biochemical Engineering Journal*, 134, pp. 30–43. Available at: <https://doi.org/10.1016/J.BEJ.2018.03.002>.
- Jang, Y., Lee, J.Y. and Lee, J. (2012) 'Enhanced Butanol Production Obtained by Reinforcing the Direct Butanol-Forming Route in *Clostridium acetobutylicum*', 3(5), pp. 1–9. Available at: <https://doi.org/10.1128/mBio.00314-12>. Updated.
- Janssen, H. *et al.* (2012) 'A transcriptional study of acidogenic chemostat cells of *Clostridium acetobutylicum*—Solvent stress caused by a transient n-butanol pulse', *Journal of Biotechnology*, 161(3), pp. 354–365. Available at: <https://doi.org/10.1016/J.JBIOTEC.2012.03.027>.
- Jawed, K. *et al.* (2016) 'Engineered Production of Short Chain Fatty Acid in *Escherichia coli* Using Fatty Acid Synthesis Pathway', *PLOS ONE*, 11(7), p. e0160035. Available at: <https://doi.org/10.1371/JOURNAL.PONE.0160035>.
- Jawed, K. *et al.* (2020) 'Improved Butanol Production Using FASII Pathway in *E. coli*', *ACS Synthetic Biology*, 9(9), pp. 2390–2398. Available at: <https://doi.org/10.1021/acssynbio.0c00154>.
- Jensen, P.R. and Michelsen, O. (1992) 'Carbon and energy metabolism of atp mutants of *Escherichia coli*', *Journal of Bacteriology*, 174(23), pp. 7635–7641. Available at: <https://doi.org/10.1128/JB.174.23.7635-7641.1992>.
- Jiang, Y. *et al.* (2009) 'Disruption of the acetoacetate decarboxylase gene in solvent-producing *Clostridium acetobutylicum* increases the butanol ratio', *Metabolic Engineering*, 11(4–5), pp. 284–291. Available at: <https://doi.org/10.1016/j.ymben.2009.06.002>.

- Jiang, Y. *et al.* (2015) 'Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system', *Applied and Environmental Microbiology*, 81(7), pp. 2506–2514. Available at: <https://doi.org/10.1128/AEM.04023-14>.
- Jin, C. *et al.* (2011) 'Progress in the production and application of n-butanol as a biofuel', *Renewable and Sustainable Energy Reviews*, 15(8), pp. 4080–4106. Available at: <https://doi.org/10.1016/J.RSER.2011.06.001>.
- Jin, Y.L. *et al.* (2009) 'Metabolic engineering of Clostridium acetobutylicum M5 for highly selective butanol production', *Biotechnology Journal*, 4(10), pp. 1432–1440. Available at: <https://doi.org/10.1002/biot.200900142>.
- Jones, D.T. and Woods, D.R. (1986) 'Acetone-butanol fermentation revisited.', *Microbiological reviews*, 50(4), pp. 484–524. Available at: <https://doi.org/3540574>.
- Kappelmann, J., Wiechert, W. and Noack, S. (2016) 'Cutting the Gordian Knot: Identifiability of anaplerotic reactions in Corynebacterium glutamicum by means of ¹³C-metabolic flux analysis', *Biotechnology and Bioengineering*, 113(3), pp. 661–674. Available at: <https://doi.org/10.1002/BIT.25833>.
- Kashket, E.R. and Cao, Z.Y. (1993) 'Isolation of a Degeneration-Resistant Mutant of Clostridium acetobutylicum NCIMB 8052', *Applied and Environmental Microbiology*, 59(12), pp. 4198–4202. Available at: <https://doi.org/10.1128/AEM.59.12.4198-4202.1993>.
- Katoh, K. *et al.* (2002) 'MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform', *Nucleic Acids Research*, 30(14), pp. 3059–3066. Available at: <https://doi.org/10.1093/NAR/GKF436>.
- Kaushal, M. *et al.* (2018) 'Understanding regulation in substrate dependent modulation of growth and production of alcohols in Clostridium sporogenes NCIM 2918 through metabolic network reconstruction and flux balance analysis', *Bioresource Technology*, 249(October 2017), pp. 767–776. Available at: <https://doi.org/10.1016/j.biortech.2017.10.080>.
- Keasling, J. *et al.* (2021) 'Microbial production of advanced biofuels', *Nature Reviews Microbiology*, pp. 701–715. Available at: <https://doi.org/10.1038/s41579-021-00577-w>.
- Keasling, J.D., van Dien, S.J. and Pramanik, J. (1998) 'Engineering polyphosphate metabolism in Escherichia coli: Implications for bioremediation of inorganic contaminants', *Biotechnology and Bioengineering*, 58(2–3), pp. 231–239. Available at: [https://doi.org/10.1002/\(SICI\)1097-0290\(19980420\)58:2/3<231::AID-BIT16>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-0290(19980420)58:2/3<231::AID-BIT16>3.0.CO;2-F).
- Keis, S. *et al.* (1995) 'Taxonomy and phylogeny of industrial solvent-producing clostridia', *International Journal of Systematic Bacteriology*, 45(4), pp. 693–705. Available at: <https://doi.org/10.1099/00207713-45-4-693>.
- Keis, S., Shaheen, R. and Jones, D.T. (2001) 'Emended descriptions of Clostridium acetobutylicum and Clostridium beijerinckii, and descriptions of Clostridium saccharoperbutylacetonicum sp. nov. and Clostridium saccharobutylicum sp. nov.', *International journal of systematic and evolutionary microbiology*, 51(Pt 6), pp. 2095–103. Available at: <https://doi.org/10.1099/00207713-51-6-2095>.
- Kenealy, W.R. and Waselefsky, D.M. (1985) 'Studies on the substrate range of Clostridium kluyveri; the use of propanol and succinate', *Archives of Microbiology*, 141(3), pp. 187–194. Available at: <https://doi.org/10.1007/BF00408056>.

- Khodayari, A. *et al.* (2014) 'A kinetic model of Escherichia coli core metabolism satisfying multiple sets of mutant flux data', *Metabolic Engineering*, 25, pp. 50–62. Available at: <https://doi.org/10.1016/j.ymben.2014.05.014>.
- Khodayari, A. and Maranas, C.D. (2016) 'A genome-scale Escherichia coli kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains.', *Nature communications*, 7, p. 13806. Available at: <https://doi.org/10.1038/ncomms13806>.
- Kim, B. *et al.* (1984) 'Control of Carbon and Electron Flow in Clostridium acetobutylicum Fermentations: Utilization of', *Applied and Environmental Microbiology*, 48(4), pp. 764–770. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=13241823564203067459.
- Kim, S. *et al.* (2015) 'Redox-switch regulatory mechanism of thiolase from Clostridium acetobutylicum', *Nature Communications* 2015 6:1, 6(1), pp. 1–11. Available at: <https://doi.org/10.1038/ncomms9410>.
- Knappe, J. *et al.* (1974) 'Pyruvate Formate-Lyase of Escherichia coli: the Acetyl-Enzyme Intermediate', *European Journal of Biochemistry*, 50(1), pp. 253–263. Available at: <https://doi.org/10.1111/j.1432-1033.1974.tb03894.x>.
- Knappe, J. and Sawers, G. (1990) 'A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of Escherichia coli', *FEMS Microbiol. Rev.*, 75, pp. 383–398. Available at: <https://doi.org/10.1111/j.1574-6968.1990.tb04108.x>.
- Koebmann, B.J. *et al.* (2002) 'The glycolytic flux in Escherichia coli is controlled by the demand for ATP', *Journal of Bacteriology*, 184(14), pp. 3909–3916. Available at: <https://doi.org/10.1128/JB.184.14.3909-3916.2002>.
- Kosaka, T. *et al.* (2007) 'Characterization of the sol Operon in Butanol-Hyperproducing Clostridium saccharoperbutylacetonicum Strain N1-4 and Its Degeneration Mechanism', *Bioscience, Biotechnology, and Biochemistry*, 71(1), pp. 58–68. Available at: <https://doi.org/10.1271/bbb.60370>.
- Kotlarz, D., Garreau, H. and Buc, H. (1975) 'Regulation of the amount and of the activity of phosphofructokinases and pyruvate kinases in Escherichia coli', *Biochimica et Biophysica Acta (BBA) - General Subjects*, 381(2), pp. 257–268. Available at: [https://doi.org/10.1016/0304-4165\(75\)90232-9](https://doi.org/10.1016/0304-4165(75)90232-9).
- Kotte, A.K. *et al.* (2020) 'RRNPP-type quorum sensing affects solvent formation and sporulation in clostridium acetobutylicum', *Microbiology (United Kingdom)*, 166(6), pp. 579–592. Available at: <https://doi.org/10.1099/mic.0.000916>.
- Koutinas, A.A., Wang, R. and Webb, C. (2004) 'Restructuring Upstream Bioprocessing: Technological and Economical Aspects for Production of a Generic Microbial Feedstock from Wheat', *Biotechnology and Bioengineering*, 85(5), pp. 524–538. Available at: <https://doi.org/10.1002/bit.10888>.
- Kratzer, R. *et al.* (2008) 'Whole-cell bioreduction of aromatic alpha-keto esters using Candida tenuis xylose reductase and Candida boidinii formate dehydrogenase co-expressed in Escherichia coli', *Microbial cell factories*, 7. Available at: <https://doi.org/10.1186/1475-2859-7-37>.

- Kudahettige-Nilsson, R.L. *et al.* (2015) 'Biobutanol production by *Clostridium acetobutylicum* using xylose recovered from birch Kraft black liquor', *Bioresource Technology*, 176, pp. 71–79. Available at: <https://doi.org/10.1016/J.BIORTECH.2014.11.012>.
- Kuit, W. *et al.* (2012) 'Disruption of the acetate kinase (ack) gene of *Clostridium acetobutylicum* results in delayed acetate production', *Applied Microbiology and Biotechnology*, 94(3), pp. 729–741. Available at: <https://doi.org/10.1007/s00253-011-3848-4>.
- Kumar, M. *et al.* (2012) 'Comparative economic assessment of ABE fermentation based on cellulosic and non-cellulosic feedstocks', *Applied Energy*, 93, pp. 193–204. Available at: <https://doi.org/10.1016/j.apenergy.2011.12.079>.
- Kutzenok, A. and Aschner, M. (1952) 'Degenerative processes in a strain of *Clostridium butylicum*.', *Journal of bacteriology*, 64(6), pp. 829–836. Available at: <https://doi.org/10.1128/jb.64.6.829-836.1952>.
- Lan, E.I. and Liao, J.C. (2012) 'ATP drives direct photosynthetic production of 1-butanol in cyanobacteria', *Proceedings of the National Academy of Sciences*, 109(16), pp. 6018–6023. Available at: <https://doi.org/10.1073/pnas.1200074109>.
- Lan, E.I., Ro, S.Y. and Liao, J.C. (2013) 'Oxygen-tolerant coenzyme A-acylating aldehyde dehydrogenase facilitates efficient photosynthetic n-butanol biosynthesis in cyanobacteria', *Energy and Environmental Science*, 6(9), pp. 2672–2681. Available at: <https://doi.org/10.1039/c3ee41405a>.
- Lange, A. *et al.* (2017) 'Bio-based succinate from sucrose: High-resolution ¹³C metabolic flux analysis and metabolic engineering of the rumen bacterium *Basfia succiniciproducens*', *Metabolic Engineering*, 44, pp. 198–212. Available at: <https://doi.org/10.1016/j.ymben.2017.10.003>.
- Largier, S.T. *et al.* (1985) 'Immobilized *Clostridium acetobutylicum* P262 mutants for solvent production', *Applied and Environmental Microbiology*, 50(2), pp. 477–481. Available at: <https://doi.org/10.1128/AEM.50.2.477-481.1985>.
- Lee, J. *et al.* (2005) 'Evidence for the presence of an alternative glucose transport system in *Clostridium beijerinckii* NCIMB 8052 and the solvent-hyperproducing mutant BA101', *Applied and Environmental Microbiology*, 71(6), pp. 3384–3387. Available at: <https://doi.org/10.1128/AEM.71.6.3384-3387.2005>.
- Lee, J. *et al.* (2008) 'Genome-scale reconstruction and in silico analysis of the *Clostridium acetobutylicum* ATCC 824 metabolic network', *Applied Microbiology and Biotechnology*, 80(5), pp. 849–862. Available at: <https://doi.org/10.1007/s00253-008-1654-4>.
- Lee, J. and Blaschek, H.P. (2001) 'Glucose Uptake in *Clostridium beijerinckii* NCIMB 8052 and the Solvent-Hyperproducing Mutant BA101', *Applied and Environmental Microbiology*, 67(3–12), pp. 5025–5031. Available at: <https://doi.org/10.1128/aem.67.11.5025-5031.2001>.
- Lee, J.W. *et al.* (2011) 'Systems metabolic engineering for chemicals and materials', *Trends in Biotechnology*, 29(8), pp. 370–378. Available at: <https://doi.org/10.1016/J.TIBTECH.2011.04.001>.
- Lee, K.H. *et al.* (2007) 'Systems metabolic engineering of *Escherichia coli* for L-threonine production', *Molecular Systems Biology*, 3(1), p.149. Available at: <https://doi.org/10.1038/msb4100196>.

- Lee, S. *et al.* (2000) 'Recursive MILP model for finding all the alternate optima in LP models for metabolic networks', *Computers & Chemical Engineering*, 24(2–7), pp. 711–716. Available at: [https://doi.org/10.1016/S0098-1354\(00\)00323-9](https://doi.org/10.1016/S0098-1354(00)00323-9).
- Lee, S.Y. *et al.* (1992) 'Vector Construction, Transformation, and Gene Amplification in *Clostridium acetobutylicum* ATCC 824', *Annals of the New York Academy of Sciences*, 665(1), pp. 39–51. Available at: <https://doi.org/10.1111/j.1749-6632.1992.tb42572.x>.
- Lee, S.Y., Bennett, G.N. and Papoutsakis, E.T. (1992) 'Construction of *Escherichia coli*-*Clostridium acetobutylicum* shuttle vectors and transformation of *Clostridium acetobutylicum* strains', *Biotechnology Letters*, 14(5), pp. 427–432. Available at: <https://doi.org/10.1007/BF01021259>.
- Lee, S.Y. and Kim, H.U. (2015) 'Systems strategies for developing industrial microbial strains', *Nature Biotechnology*, 33(10), pp. 1061–1072. Available at: <https://doi.org/10.1038/nbt.3365>.
- Lehmann, D. *et al.* (2012) 'Modifying the product pattern of *Clostridium acetobutylicum*: Physiological effects of disrupting the acetate and acetone formation pathways', *Applied Microbiology and Biotechnology*, 94(3), pp. 743–754. Available at: <https://doi.org/10.1007/s00253-011-3852-8>.
- Lehmann, D. and Lütke-Eversloh, T. (2011) 'Switching *Clostridium acetobutylicum* to an ethanol producer by disruption of the butyrate/butanol fermentative pathway', *Metabolic Engineering*, 13(5), pp. 464–473. Available at: <https://doi.org/10.1016/j.ymben.2011.04.006>.
- Lehmann, D., Radomski, N. and Lütke-Eversloh, T. (2012a) 'New insights into the butyric acid metabolism of *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 96(5), pp. 1325–1339. Available at: <https://doi.org/10.1007/s00253-012-4109-x>.
- Lehmann, D., Radomski, N. and Lütke-Eversloh, T. (2012b) 'New insights into the butyric acid metabolism of *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 96(5), pp. 1325–1339. Available at: <https://doi.org/10.1007/s00253-012-4109-x>.
- Lewis, N.E. *et al.* (2010) 'Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models', *Molecular Systems Biology*, 6(390). Available at: <https://doi.org/10.1038/msb.2010.47>.
- Li, Q. *et al.* (2016) 'CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*', *Biotechnology Journal*, 11(7), pp. 961–972. Available at: <https://doi.org/10.1002/biot.201600053>.
- Liao, J.C., Chao, Y.P. and Patnaik, R. (1994) 'Alteration of the biochemical valves in the central metabolism of *Escherichia coli*.' , *Annals of the New York Academy of Sciences*, 745, pp. 21–34. Available at: <https://doi.org/10.1111/J.1749-6632.1994.TB44361.X>.
- Liao, Z. *et al.* (2019) 'The significance of proline on lignocellulose-derived inhibitors tolerance in *Clostridium acetobutylicum* ATCC 824', *Bioresource Technology*, 272, pp. 561–569. Available at: <https://doi.org/10.1016/J.BIORTECH.2018.10.038>.
- Liew, F.E. *et al.* (2022) 'Carbon-negative production of acetone and isopropanol by gas fermentation at industrial pilot scale', *Nature Biotechnology*, 40(3), pp. 335–344. Available at: <https://doi.org/10.1038/s41587-021-01195-w>.
- Liu, H. *et al.* (2019) 'Omics-based analyses revealed metabolic responses of *Clostridium acetobutylicum* to lignocellulose-derived inhibitors furfural, formic acid and phenol stress for

butanol fermentation', *Biotechnology for Biofuels*, 12(1), pp. 1–20. Available at: <https://doi.org/10.1186/S13068-019-1440-9/FIGURES/7>.

Liu, H. *et al.* (2020) 'Gene coexpression network analysis reveals a novel metabolic mechanism of *Clostridium acetobutylicum* responding to phenolic inhibitors from lignocellulosic hydrolysates', *Biotechnology for Biofuels*, 13(1), pp. 1–15. Available at: <https://doi.org/10.1186/s13068-020-01802-z>.

Liu, X. *et al.* (2019) 'Modular engineering for efficient photosynthetic biosynthesis of 1-butanol from CO₂ in cyanobacteria', *Energy & Environmental Science*, 12(9), pp. 2765–2777. Available at: <https://doi.org/10.1039/C9EE01214A>.

Liu, Y., Landick, R. and Raman, S. (2019) 'A Regulatory NADH/NAD⁺ Redox Biosensor for Bacteria', *ACS Synthetic Biology*, 8(2), pp. 264–273. Available at: https://doi.org/10.1021/ACSSYNBIO.8B00485/SUPPL_FILE/SB8B00485_SI_001.PDF.

Long, C.P. *et al.* (2017) 'Metabolism of the fast-growing bacterium *Vibrio natriegens* elucidated by ¹³C metabolic flux analysis', *Metabolic Engineering*, 44, pp. 191–197. Available at: <https://doi.org/10.1016/j.ymben.2017.10.008>.

Long, C.P. and Antoniewicz, M.R. (2019) 'High-resolution ¹³C metabolic flux analysis', *Nature Protocols*, 14(10), pp. 2856–2877. Available at: <https://doi.org/10.1038/s41596-019-0204-0>.

Long, S., Jones, D.T. and Woods, D.R. (1984a) 'Initiation of solvent production, clostridial stage and endospore formation in *Clostridium acetobutylicum* P262', *Applied Microbiology and Biotechnology*, 20(4), pp. 256–261. Available at: <https://doi.org/10.1007/BF00250635>.

Long, S., Jones, D.T. and Woods, D.R. (1984b) 'The relationship between sporulation and solvent production in *Clostridium acetobutylicum* P262', *Biotechnology Letters*, 6(8), pp. 529–534. Available at: <https://doi.org/10.1007/BF00139997>.

Lowe, R. *et al.* (2017) 'Transcriptomics technologies', *PLOS Computational Biology*, 13(5), p. e1005457. Available at: <https://doi.org/10.1371/JOURNAL.PCBI.1005457>.

Lu, C. *et al.* (2017) 'Enhanced robustness in acetone-butanol-ethanol fermentation with engineered *Clostridium beijerinckii* overexpressing adhE2 and ctfAB', *Bioresource Technology*, 243, pp. 1000–1008. Available at: <https://doi.org/10.1016/j.biortech.2017.07.043>.

Lu, J.-L. and Liao, J.C. (1997) 'Metabolic Engineering and Control Analysis for Production of Aromatics: Role of Transaldolase', *Biotechnology and Bioengineering*, 53(2), pp.132–138. Available at: [https://doi.org/10.1002/\(SICI\)1097-0290\(19970120\)53:2](https://doi.org/10.1002/(SICI)1097-0290(19970120)53:2).

Luers, F. *et al.* (2006) 'Glycerol conversion to 1,3-propanediol by *Clostridium pasteurianum*: cloning and expression of the gene encoding 1,3-propanediol dehydrogenase', *FEMS Microbiology Letters*, 154(2), pp. 337–345. Available at: <https://doi.org/10.1111/j.1574-6968.1997.tb12665.x>.

Luo, H. *et al.* (2020) 'Efficient bio-butanol production from lignocellulosic waste by elucidating the mechanisms of *Clostridium acetobutylicum* response to phenolic inhibitors', *Science of The Total Environment*, 710, p. 136399. Available at: <https://doi.org/10.1016/J.SCITOTENV.2019.136399>.

Luo, Z. *et al.* (2019) 'Enhanced Pyruvate Production in *Candida glabrata* by Engineering ATP Futile Cycle System', *ACS Synthetic Biology*, 8(4), pp. 787–795. Available at: https://doi.org/10.1021/ACSSYNBIO.8B00479/SUPPL_FILE/SB8B00479_SI_001.PDF.

- Maddox, I.S. *et al.* (2000) 'The cause of "acid crash" and "acidogenic fermentations" during the batch acetone-butanol-ethanol (ABE-) fermentation process', *Journal of Molecular Microbiology and Biotechnology*, 2(1), pp. 95–100.
- Mädje, K. *et al.* (2012) 'Host cell and expression engineering for development of an E. coli ketoreductase catalyst: Enhancement of formate dehydrogenase activity for regeneration of NADH', *Microbial Cell Factories*, 11(1), pp. 1–8. Available at: <https://doi.org/10.1186/1475-2859-11-7/TABLES/3>.
- Mahadevan, R., Edwards, J.S. and Doyle, F.J. (2002) 'Dynamic Flux Balance Analysis of diauxic growth in Escherichia coli', *Biophysical Journal*, 83(3), pp. 1331–1340. Available at: [https://doi.org/10.1016/S0006-3495\(02\)73903-9](https://doi.org/10.1016/S0006-3495(02)73903-9).
- Mahadevan, R. and Schilling, C.H. (2003) 'The effects of alternate optimal solutions in constraint-based genome-scale metabolic models', *Metabolic Engineering*, 5(4), pp. 264–276. Available at: <https://doi.org/10.1016/j.ymben.2003.09.002>.
- Mancini, E. *et al.* (2022) 'Economic and environmental analysis of bio-succinic acid production: From established processes to a new continuous fermentation approach with in-situ electrolytic extraction', *Chemical Engineering Research and Design*, 179, pp. 401–414. Available at: <https://doi.org/10.1016/J.CHERD.2022.01.040>.
- Martin, J.R. *et al.* (1983) 'Effects of acetic and butyric acids on solvents production by Clostridium acetobutylicum', *Biotechnology Letters* 1983 5:2, 5(2), pp. 89–94. Available at: <https://doi.org/10.1007/BF00132165>.
- Máté de Gérand, H. *et al.* (2018) 'Genome and transcriptome of the natural isopropanol producer Clostridium beijerinckii DSM6423', *BMC Genomics*, 19(1), pp. 1–12. Available at: <https://doi.org/10.1186/s12864-018-4636-7>.
- Maxon, W.D. (1955) 'Continuous fermentation; a discussion of its principles and applications.', *Applied microbiology*, 3(2), pp. 110–120. Available at: <https://doi.org/10.1128/aem.3.2.110-122.1955>.
- McAnulty, M.J. *et al.* (2012) 'Genome-scale modeling using flux ratio constraints to enable metabolic engineering of clostridial metabolism in silico', *BMC Systems Biology*, 6. Available at: <https://doi.org/10.1186/1752-0509-6-42>.
- Meadows, A.L. *et al.* (2010) 'Application of dynamic flux balance analysis to an industrial Escherichia coli fermentation', *Metabolic Engineering*, 12(2), pp. 150–160. Available at: <https://doi.org/10.1016/J.YMBEN.2009.07.006>.
- Mermelstein, L.D. *et al.* (1992) 'Expression of cloned homologous fermentative genes in Clostridium acetobutylicum ATCC 824', *Nature Biotechnology*, 10, pp. 190–195.
- Mermelstein, L.D. and Papoutsakis, E.T. (1993) 'In vivo methylation in [i]Escherichia coli by the Bacillus subtilis phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of Clostridium acetobutylicum ATCC 824', *Appl Environ Microbiol*, 59(4), pp. 1077–1081. Available at: <https://doi.org/10.1016/j.bbrc.2011.01.095>.
- Mes-Hartree, M. and Saddler, J.N. (1982) 'Butanol production of Clostridium acetobutylicum grown on sugars found in hemicellulose hydrolysates', *Biotechnology Letters*, 4(4), pp. 247–252. Available at: <https://doi.org/10.1007/BF00132395>.

- Metallo, C.M., Walther, J.L. and Stephanopoulos, G. (2009) 'Evaluation of ^{13}C isotopic tracers for metabolic flux analysis in mammalian cells', *Journal of Biotechnology*, 144(3), pp. 167–174. Available at: <https://doi.org/10.1016/J.JBIOTEC.2009.07.010>.
- Meyer, C.L., McLaughlin, J.K. and Papoutsakis, E.T. (1985) 'The effect of CO on growth and product formation in batch cultures of *Clostridium acetobutylicum*', *Biotechnology Letters*, 7(1), pp. 37–42. Available at: <https://doi.org/10.1007/BF01032417>.
- Meyer, C. L. and Papoutsakis, E.T. (1989) 'Continuous and biomass recycle fermentations of *Clostridium acetobutylicum*', *Bioprocess Engineering* 1989 4:1, 4(1), pp. 1–10. Available at: <https://doi.org/10.1007/BF00612664>.
- Meyer, C.L. and Papoutsakis, E.T. (1989) 'Increased levels of ATP and NADH are associated with increased solvent production in continuous cultures of *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 30(5), pp. 450–459. Available at: <https://doi.org/10.1007/BF00263849>.
- Meyer, C.L., Roos, J.W. and Papoutsakis, E.T. (1986) 'Carbon monoxide gasing leads to alcohol production and butyrate uptake without acetone formation in continuous cultures of *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 24(2), pp. 159–167. Available at: <https://doi.org/10.1007/BF01982561>.
- Miller, M.B. and Bassler, B.L. (2001) 'Quorum Sensing in Bacteria', *Annual Reviews in Microbiology*, 55(1), pp.165-199. Available at: <https://doi.org/10.1146/ANNUREV.MICRO.55.1.165>.
- Milne, C.B. *et al.* (2011) 'Metabolic network reconstruction and genome-scale model of butanol-producing strain *Clostridium beijerinckii* NCIMB 8052', *BMC Systems Biology*, 5. Available at: <https://doi.org/10.1186/1752-0509-5-130>.
- Mitchell, W.J. (2016) 'Sugar uptake by the solventogenic clostridia', *World Journal of Microbiology and Biotechnology*, 32(2), pp. 1–10. Available at: <https://doi.org/10.1007/s11274-015-1981-4>.
- Mitchell, W.J., Shaw, J.E. and Andrews, L. (1991) 'Properties of the glucose phosphotransferase system of *Clostridium acetobutylicum* NCIB 8052', *Applied and Environmental Microbiology*, 57(9), pp. 2534–2539. Available at: <https://doi.org/10.1128/aem.57.9.2534-2539.1991>.
- Mohr, A. and Raman, S. (2015) 'Lessons from first generation biofuels and implications for the sustainability appraisal of second generation biofuels', *Efficiency and Sustainability in Biofuel Production: Environmental and Land-Use Research*, 63, pp. 281–310. Available at: <https://doi.org/10.1016/j.enpol.2013.08.033>.
- Monot, F. *et al.* (1982) 'Acetone and butanol production by *Clostridium acetobutylicum* in a synthetic medium', *Applied and Environmental Microbiology*, 44(6), pp. 1318–1324. Available at: <https://doi.org/10.1128/aem.44.6.1318-1324.1982>.
- Monot, F., Engasser, J.M. and Petitdemange, H. (1984) 'Influence of pH and undissociated butyric acid on the production of acetone and butanol in batch cultures of *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 19(6), pp. 422–426. Available at: <https://doi.org/10.1007/BF00454381>.

- Mordaka, P.M. and Heap, J.T. (2018) 'Stringency of Synthetic Promoter Sequences in *Clostridium* Revealed and Circumvented by Tuning Promoter Library Mutation Rates', *ACS Synthetic Biology*, 7(2), pp. 672–681. Available at: <https://doi.org/10.1021/acssynbio.7b00398>.
- Moreira Dos Santos, M. *et al.* (2004) 'Manipulation of malic enzyme in *Saccharomyces cerevisiae* for increasing NADPH production capacity aerobically in different cellular compartments', *Metabolic Engineering*, 6(4), pp. 352–363. Available at: <https://doi.org/10.1016/J.YMBEN.2004.06.002>.
- Morita, K. *et al.* (2017) 'Heterologous expression of bacterial phosphoenol pyruvate carboxylase and Entner–Doudoroff pathway in *Saccharomyces cerevisiae* for improvement of isobutanol production', *Journal of Bioscience and Bioengineering*, 124(3), pp. 263–270. Available at: <https://doi.org/10.1016/J.JBIOSEC.2017.04.005>.
- Moxley, G. and Zhang, Y.H.P. (2007) 'More accurate determination of acid-labile carbohydrates in lignocellulose by modified quantitative saccharification', *Energy and Fuels*, 21(6), pp. 3684–3688. Available at: https://doi.org/10.1021/EF7003893/SUPPL_FILE/EF7003893-FILE003.PDF.
- Muers, M. (2011) 'Gene expression: Transcriptome to proteome and back to genome', *Nature Reviews Genetics*, p. 518. Available at: <https://doi.org/10.1038/nrg3037>.
- Müller, V. (2008) 'Bacterial Fermentation', in *eLS*. John Wiley & Sons, Ltd. Available at: <https://doi.org/10.1002/9780470015902.a0001415.pub2>.
- Nagarajan, H. *et al.* (2013) 'Characterizing acetogenic metabolism using a genome-scale metabolic reconstruction of *Clostridium ljungdahlii*', *Microbial Cell Factories*, 12(1), pp. 1–13. Available at: <https://doi.org/10.1186/1475-2859-12-118>.
- Nair, R. v. *et al.* (1999) 'Regulation of the sol locus genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 by a putative transcriptional repressor', *Journal of Bacteriology*, 181(1), pp. 319–330.
- Nair, R. v., Bennett, G.N. and Papoutsakis, E.T. (1994) 'Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824', *Journal of Bacteriology*, 176(3), pp. 871–885. Available at: <https://doi.org/10.1128/jb.176.3.871-885.1994>.
- Nair, R. V., Bennett, G.N. and Papoutsakis, E.T. (1994) 'Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824', *Journal of Bacteriology*, 176(3), pp. 871–885. Available at: <https://doi.org/10.1128/jb.176.3.871-885.1994>.
- Neiditch, M.B. *et al.* (2017) 'Genetic and Structural Analyses of RRNPP Intercellular Peptide Signaling of Gram-Positive Bacteria', *Annual review of genetics*, 51, p. 311. Available at: <https://doi.org/10.1146/ANNUREV-GENET-120116-023507>.
- Neuhauser, W. *et al.* (1998) 'A pH-Controlled Fed-Batch Process Can Overcome Inhibition by Formate in NADH-Dependent Enzymatic Reductions Using Formate Dehydrogenase-Catalyzed Coenzyme Regeneration', *Biotechnol Bioeng*, 60, pp. 277–282. Available at: [https://doi.org/10.1002/\(SICI\)1097-0290\(19981105\)60:3](https://doi.org/10.1002/(SICI)1097-0290(19981105)60:3).
- Ng, C.Y. *et al.* (2015) 'Rational design of a synthetic Entner–Doudoroff pathway for improved and controllable NADPH regeneration', *Metabolic Engineering*, 29, pp. 86–96. Available at: <https://doi.org/10.1016/j.ymben.2015.03.001>.

- Nguyen, N.-P.-T. *et al.* (2018) 'Reviving the Weizmann process for commercial n-butanol production', *Nature Communications*, 9(1), p. 3682. Available at: <https://doi.org/10.1038/s41467-018-05661-z>.
- Nielsen, J. (2001) 'Metabolic engineering', *Applied Microbiology and Biotechnology*. Springer, pp. 263–283. Available at: <https://doi.org/10.1007/s002530000511>.
- Nielsen, J. and Keasling, J.D. (2016) 'Engineering Cellular Metabolism', *Cell*, 164(6), pp. 1185–1197. Available at: <https://doi.org/10.1016/j.cell.2016.02.004>.
- Nigam, P. and Singh, D. (1995) 'Enzyme and microbial systems involved in starch processing', *Enzyme and Microbial Technology*, 17(9), pp. 770–778. Available at: [https://doi.org/10.1016/0141-0229\(94\)00003-A](https://doi.org/10.1016/0141-0229(94)00003-A).
- Nishio, Y. *et al.* (2008) 'Computer-aided rational design of the phosphotransferase system for enhanced glucose uptake in Escherichia coli', *Molecular Systems Biology*, 4(1), p. 160. Available at: <https://doi.org/10.1038/msb4100201>.
- Noguchi, T. *et al.* (2013) 'Efficient butanol production without carbon catabolite repression from mixed sugars with Clostridium saccharoperbutylacetonicum N1-4', *Journal of Bioscience and Bioengineering*, 116(6), pp. 716–721. Available at: <https://doi.org/10.1016/j.jbiosc.2013.05.030>.
- Nölling, J. *et al.* (2001) 'Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum', *Journal of Bacteriology*, 183(16), pp. a823–4838. Available at: <https://doi.org/10.1128/JB.183.16.4823>.
- Norman, R.O.J. *et al.* (2019) 'Genome-scale model of C. autoethanogenum reveals optimal bioprocess conditions for high-value chemical production from carbon monoxide', *Engineering Biology*, 3(2), pp. 32–40. Available at: <https://doi.org/10.1049/enb.2018.5003>.
- O'Brien, E.J., Monk, J.M. and Palsson, B.O. (2015) 'Using genome-scale models to predict biological capabilities', *Cell*, pp. 971–987. Available at: <https://doi.org/10.1016/j.cell.2015.05.019>.
- Ohtake, T. *et al.* (2017) 'Metabolomics-driven approach to solving a CoA imbalance for improved 1-butanol production in Escherichia coli', *Metabolic Engineering*, 41(April), pp. 135–143. Available at: <https://doi.org/10.1016/j.ymben.2017.04.003>.
- Okahashi, N. *et al.* (2017) 'Metabolic engineering of isopropyl alcohol-producing Escherichia coli strains with 13C-metabolic flux analysis', *Biotechnology and Bioengineering*, 114(12), pp. 2782–2793. Available at: <https://doi.org/10.1002/bit.26390>.
- Okeke, B.C. and Obi, S.K.C. (1994) 'Lignocellulose and sugar compositions of some agro-waste materials', *Bioresource Technology*, 47(3), pp. 283–284. Available at: [https://doi.org/10.1016/0960-8524\(94\)90192-9](https://doi.org/10.1016/0960-8524(94)90192-9).
- Olsson, L. and Hahn-Hägerdal, B. (1996) 'Fermentation of lignocellulosic hydrolysates for ethanol production', *Enzyme and Microbial Technology*, 18(5), pp. 312–331. Available at: [https://doi.org/10.1016/0141-0229\(95\)00157-3](https://doi.org/10.1016/0141-0229(95)00157-3).
- Orth, J.D., Thiele, I. and Palsson, B.O. (2010) 'What is flux balance analysis?', *Nature Biotechnology*, 28(3), pp. 245–248. Available at: <https://doi.org/10.1038/nbt.1614>.

- Ounine, K. *et al.* (1985) 'Regulation and butanol inhibition of D-xylose and D-glucose uptake in *Clostridium acetobutylicum*', *Applied and Environmental Microbiology*, 49(4), pp. 874–878.
- Papoutsakis, E.T. and Alsaker, K. v. (2005) 'The transcriptional program of early sporulation and stationary phase events in *Clostridium acetobutylicum*', *AIChE Annual Meeting, Conference Proceedings*, 187(20), p. 11608. Available at: <https://doi.org/10.1128/JB.187.20.7103>.
- Patakova, P. *et al.* (2021) 'Deeper below the surface—transcriptional changes in selected genes of *Clostridium beijerinckii* in response to butanol shock', *MicrobiologyOpen*, 10(1). Available at: <https://doi.org/10.1002/mbo3.1146>.
- Patil, K.R. *et al.* (2005) 'Evolutionary programming as a platform for in silico metabolic engineering', *BMC Bioinformatics*, 6, pp. 1–12. Available at: <https://doi.org/10.1186/1471-2105-6-308>.
- Patnaik, R. *et al.* (1992) 'Stimulation of glucose catabolism in *Escherichia coli* by a potential futile cycle', *Journal of Bacteriology*, 174(23), pp. 7527–7532. Available at: <https://doi.org/10.1128/jb.174.23.7527-7532.1992>.
- Paul, D. *et al.* (2010) 'Genome Sequence of the Solvent-Producing Bacterium *Clostridium carboxidivorans* Strain P7T', *Journal of Bacteriology*, 192(20), pp. 5554–5555. Available at: <https://doi.org/10.1128/JB.00877-10>.
- Peguin, S. *et al.* (1994) 'Metabolic flexibility of *Clostridium acetobutylicum* in response to methyl viologen addition', *Applied Microbiology and Biotechnology* 1994 42:4, 42(4), pp. 611–616. Available at: <https://doi.org/10.1007/BF00173928>.
- Petersen, D.J. *et al.* (1991) 'Molecular cloning of an alcohol (butanol) dehydrogenase gene cluster from *Clostridium acetobutylicum* ATCC 824', *Journal of Bacteriology*, 173(5), pp. 1831–1834. Available at: <https://doi.org/10.1128/JB.173.5.1831-1834.1991>.
- Petitdemange, H. *et al.* (1976) 'Regulation of the NADH and NADPH-ferredoxin oxidoreductases in Clostridia of the butyric group', *BBA - General Subjects*, 421(2), pp. 334–347. Available at: [https://doi.org/10.1016/0304-4165\(76\)90300-7](https://doi.org/10.1016/0304-4165(76)90300-7).
- Piotrowski, J.S. *et al.* (2014) 'Death by a thousand cuts: The challenges and diverse landscape of lignocellulosic hydrolysate inhibitors', *Frontiers in Microbiology*, 5(MAR), p. 90. Available at: <https://doi.org/10.3389/FMICB.2014.00090/BIBTEX>.
- Poehlein, A. *et al.* (2014) 'Complete genome sequence of the solvent producer *Clostridium saccharoperbutylacetonicum* strain DSM 14923', *Genome Announcements*, 2(5), pp. 1056–1070. Available at: <https://doi.org/10.1128/genomeA.01056-14>.
- Poehlein, A. *et al.* (2015) 'Complete genome sequence of the nitrogen-fixing and solvent-producing *Clostridium pasteurianum* DSM 525', *Genome Announcements*, 3(1), pp. 1591–1605. Available at: <https://doi.org/10.1128/genomeA.01591-14>.
- Poehlein, A. *et al.* (2017) 'Microbial solvent formation revisited by comparative genome analysis', *Biotechnology for Biofuels*, 10(1), pp. 1–15. Available at: <https://doi.org/10.1186/s13068-017-0742-z>.
- Pontrelli, S. *et al.* (2018) '*Escherichia coli* as a host for metabolic engineering', *Metabolic Engineering*, 50, pp. 16–46. Available at: <https://doi.org/10.1016/J.YMBEN.2018.04.008>.

- Procentese, A. *et al.* (2017) 'Renewable feedstocks for biobutanol production by fermentation', *New Biotechnology*, 39, pp. 135–140. Available at: <https://doi.org/10.1016/j.nbt.2016.10.010>.
- Puchalka, J. *et al.* (2008) 'Genome-Scale Reconstruction and Analysis of the *Pseudomonas putida* KT2440 Metabolic Network Facilitates Applications in Biotechnology', *PLOS Computational Biology*, 4(10), p. e1000210. Available at: <https://doi.org/10.1371/JOURNAL.PCBI.1000210>.
- Pugazhendhi, A. *et al.* (2019) 'Biobutanol as a promising liquid fuel for the future - recent updates and perspectives', *Fuel*, 253, pp. 637–646. Available at: <https://doi.org/10.1016/J.FUEL.2019.04.139>.
- Pyne, M.E. *et al.* (2014) 'Technical guide for genetic advancement of underdeveloped and intractable *Clostridium*', *Biotechnology Advances*. Elsevier, pp. 623–641. Available at: <https://doi.org/10.1016/j.biotechadv.2014.04.003>.
- Pyne, M.E. *et al.* (2016) 'Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*', *Scientific Reports*, 6(December 2015), pp. 1–15. Available at: <https://doi.org/10.1038/srep25666>.
- Qureshi, N., Saha, B.C., Dien, B., *et al.* (2010) 'Production of butanol (a biofuel) from agricultural residues: Part I - Use of barley straw hydrolysate', *Biomass and Bioenergy*, 34(4), pp. 559–565. Available at: <https://doi.org/10.1016/j.biombioe.2009.12.024>.
- Qureshi, N., Saha, B.C., Hector, R.E., *et al.* (2010) 'Production of butanol (a biofuel) from agricultural residues: Part II - Use of corn stover and switchgrass hydrolysates', *Biomass and Bioenergy*, 34(4), pp. 566–571. Available at: <https://doi.org/10.1016/j.biombioe.2009.12.023>.
- Qureshi, N. *et al.* (2020) 'Production of acetone–butanol–ethanol (ABE) from concentrated yellow top presscake using *Clostridium beijerinckii* P260', *Journal of Chemical Technology and Biotechnology*, 95(3), pp. 614–620. Available at: <https://doi.org/10.1002/jctb.6242>.
- Qureshi, N., Lai, L.L. and Blaschek, H.P. (2004) 'Scale-Up of a High Productivity Continuous Biofilm Reactor to Produce Butanol by Adsorbed Cells of *Clostridium Beijerinckii*', *Food and Bioprocess Processing*, 82(2), pp. 164–173. Available at: <https://doi.org/10.1205/0960308041614891>.
- Qureshi, N., Lolas, A. and Blaschek, H.P. (2001) 'Soy molasses as fermentation substrate for production of butanol using *Clostridium beijerinckii* BA101', *Journal of Industrial Microbiology and Biotechnology*, 26(5), pp. 290–295. Available at: <https://doi.org/10.1038/sj.jim.7000131>.
- Qureshi, N., Saha, B.C. and Cotta, M.A. (2007) 'Butanol production from wheat straw hydrolysate using *Clostridium beijerinckii*', *Bioprocess and Biosystems Engineering*, 30(6), pp. 419–427. Available at: <https://doi.org/10.1007/s00449-007-0137-9>.
- Raganati, F. *et al.* (2015) 'Butanol Production from Leftover Beverages and Sport Drinks', *Bioenergy Research*, 8(1), pp. 369–379. Available at: <https://doi.org/10.1007/s12155-014-9531-8>.
- Rahimi, A. *et al.* (2019) 'Enhancing the efficiency of recombinant hepatitis B surface antigen production in *Pichia pastoris* by employing continuous fermentation', *Biochemical Engineering Journal*, 141, pp. 112–119. Available at: <https://doi.org/10.1016/J.BEJ.2018.10.019>.
- Ranganathan, S., Suthers, P.F. and Maranas, C.D. (2010) 'OptForce: An optimization procedure for identifying all genetic manipulations leading to targeted overproductions', *PLoS Computational Biology*, 6(4). Available at: <https://doi.org/10.1371/journal.pcbi.1000744>.

- Rao, G. and Mutharasan, R. (1987) 'Altered electron flow in continuous cultures of *Clostridium acetobutylicum* induced by viologen dyes', *Applied and Environmental Microbiology*, 53(6), pp. 1232–1235. Available at: <https://doi.org/papers2://publication/uuid/660100BA-2E74-4E86-BE9D-CC1FDFB7EA0C>.
- Rao, N.N., Gómez-García, M.R. and Kornberg, A. (2009) 'Inorganic Polyphosphate: Essential for Growth and Survival', *Annual Review of Biochemistry*, 78(1), pp. 605–647. Available at: <https://doi.org/10.1146/annurev.biochem.77.083007.093039>.
- Rass-Hansen, J. *et al.* (2007) 'Bioethanol: fuel or feedstock?', *Journal of Chemical Technology & Biotechnology*, 82(4), pp. 329–333. Available at: <https://doi.org/10.1002/JCTB.1665>.
- Ravagnani, A. *et al.* (2000) 'Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia', *Molecular Microbiology*, 37(5), pp. 1172–1185. Available at: <https://doi.org/10.1046/j.1365-2958.2000.02071.x>.
- Richmond, C., Ujor, V. and Ezeji, T.C. (2012) 'Impact of syringaldehyde on the growth of *Clostridium beijerinckii* NCIMB 8052 and butanol production', *3 Biotech*, 2(2), pp. 159–167. Available at: <https://doi.org/10.1007/s13205-011-0042-4>.
- Riebeling, V. and Jungermann, K. (1976) 'Properties and function of clostridial membrane ATPase', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 430(3), pp. 434–444. Available at: [https://doi.org/10.1016/0005-2728\(76\)90019-0](https://doi.org/10.1016/0005-2728(76)90019-0).
- Roberts, S.B. *et al.* (2010) 'Genome-scale metabolic analysis of *Clostridium thermocellum* for bioethanol production', *BMC Systems Biology*, 4. Available at: <https://doi.org/10.1186/1752-0509-4-31>.
- Robillard, G. and Acta, J. (1999) 'Structure/function studies on the bacterial carbohydrate transporters, enzymes II, of the phosphoenolpyruvate-dependent phosphotransferase system', *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 1422(2), pp.73-104, Available at: [https://doi.org/10.1016/S0304-4157\(99\)00002-7](https://doi.org/10.1016/S0304-4157(99)00002-7).
- Robson, R.L., Robson, R.M. and Morris, J.G. (1974) 'The biosynthesis of granulose by *Clostridium pasteurianum*', *Biochemical Journal*, 144(3), pp. 503–511. Available at: <https://doi.org/10.1042/BJ1440503>.
- Rotta, C. *et al.* (2015) 'Closed genome sequence of *Clostridium pasteurianum* ATCC 6013', *Genome Announcements*, 3(1), pp. 1596–1610. Available at: <https://doi.org/10.1128/genomeA.01596-14>.
- Li, R.D., Li, Y.Y., Lu, L.Y., Ren, C., Li, Y.X. and Liu, L., (2011) 'An improved kinetic model for the acetone-butanol-ethanol pathway of *Clostridium acetobutylicum* and model-based perturbation analysis' *BMC systems biology*, 5, pp.1-11. Available at: <https://doi.org/10.1186/1752-0509-5-S1-S12>
- Ruyter, G.J.G., Postma, P.W. and van Dam, K. (1991) 'Control of glucose metabolism by enzyme II_{Glc} of the phosphoenolpyruvate-dependent phosphotransferase system in *Escherichia coli*', *Journal of Bacteriology*, 173(19), pp. 6184–6191. Available at: <https://doi.org/10.1128/JB.173.19.6184-6191.1991>.

- Saa, P.A. and Nielsen, L.K. (2017) 'Formulation, construction and analysis of kinetic models of metabolism: A review of modelling frameworks', *Biotechnology Advances*, 35(8), pp. 981–1003. Available at: <https://doi.org/10.1016/j.biotechadv.2017.09.005>.
- Salimi, F., Zhuang, K. and Mahadevan, R. (2010) 'Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing', *Biotechnology Journal*, 5(7), pp. 726–738. Available at: <https://doi.org/10.1002/biot.201000159>.
- Sánchez, A.M., Bennett, G.N. and San, K.Y. (2005) 'Effect of different levels of NADH availability on metabolic fluxes of Escherichia coli chemostat cultures in defined medium', *Journal of Biotechnology*, 117(4), pp. 395–405. Available at: <https://doi.org/10.1016/j.jbiotec.2005.02.006>.
- Sandoval, N.R. *et al.* (2015) 'Whole-genome sequence of an evolved Clostridium pasteurianum strain reveals Spo0A deficiency responsible for increased butanol production and superior growth', *Biotechnology for Biofuels*, 8(1), pp. 1–18. Available at: <https://doi.org/10.1186/s13068-015-0408-7>.
- Santangelo IV, J.D. *et al.* (1998) 'Sporulation and time course expression of sigma-factor homologous genes in Clostridium acetobutylicum', *FEMS Microbiology Letters*, 161(1), pp. 157–164. Available at: <https://doi.org/10.1111/J.1574-6968.1998.TB12943.X>.
- Sawers, G. and Watson, G. (1998) 'A glycyl radical solution: Oxygen-dependent interconversion of pyruvate formate-lyase', *Molecular Microbiology*, 29(4), pp. 945–954. Available at: <https://doi.org/10.1046/j.1365-2958.1998.00941.x>.
- Schwarz, K.M. *et al.* (2017) 'Towards improved butanol production through targeted genetic modification of Clostridium pasteurianum', *Metabolic Engineering*, 40(January), pp. 124–137. Available at: <https://doi.org/10.1016/j.ymben.2017.01.009>.
- Seedorf, H. *et al.* (2008) 'The genome of Clostridium kluyveri, a strict anaerobe with unique metabolic features', *Proceedings of the National Academy of Sciences*, 105(6), pp. 2128–2133. Available at: <https://doi.org/10.1073/pnas.0711093105>.
- Segrè, D., Vitkup, D. and Church, G.M. (2002) 'Analysis of optimality in natural and perturbed metabolic networks', *Proceedings of the National Academy of Sciences of the United States of America*, 99(23), pp. 15112–15117. Available at: <https://doi.org/10.1073/pnas.232349399>.
- Sekine, H. *et al.* (2001) 'H⁺-ATPase defect in Corynebacterium glutamicum abolishes glutamic acid production with enhancement of glucose consumption rate', *Applied Microbiology and Biotechnology* 2001 57:4, 57(4), pp. 534–540. Available at: <https://doi.org/10.1007/S002530100778>.
- Semkiv, M. v. *et al.* (2014) 'Increased ethanol accumulation from glucose via reduction of ATP level in a recombinant strain of Saccharomyces cerevisiae overexpressing alkaline phosphatase', *BMC Biotechnology*, 14(1), pp. 1–9. Available at: <https://doi.org/10.1186/1472-6750-14-42/FIGURES/4>.
- Semkiv, M. v. *et al.* (2016) 'Activation of futile cycles as an approach to increase ethanol yield during glucose fermentation in Saccharomyces cerevisiae', *Bioengineered*, 7(2), pp. 106–111. Available at: <https://doi.org/10.1080/21655979.2016.1148223>.

- Senger, R.S. and Papoutsakis, E.T. (2008) 'Genome-scale model for *Clostridium acetobutylicum*: Part I. Metabolic network resolution and analysis', *Biotechnology and Bioengineering*, 101(5), pp. 1036–1052. Available at: <https://doi.org/10.1002/bit.22010>.
- Seo, S.O. *et al.* (2017) 'Characterization of a *Clostridium beijerinckii* spo0A mutant and its application for butyl butyrate production', *Biotechnology and Bioengineering*, 114(1), pp. 106–112. Available at: <https://doi.org/10.1002/bit.26057>.
- Servinsky, M.D. *et al.* (2012) 'Arabinose is metabolized via a phosphoketolase pathway in *Clostridium acetobutylicum* ATCC 824', *Journal of Industrial Microbiology and Biotechnology*, 39(12), pp. 1859–1867. Available at: <https://doi.org/10.1007/S10295-012-1186-X>.
- Shaked, Z. and Whitesides, G.M. (1980) 'Enzyme-Catalyzed Organic Synthesis: NADH Regeneration by Using Formate Dehydrogenase', *Journal of the American Chemical Society*, 102(23), pp. 7104–7105. Available at: <https://doi.org/10.1021/ja00543a038>.
- Shao, L. *et al.* (2007) 'Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*', *Cell Research*, 17(11), pp. 963–965. Available at: <https://doi.org/10.1038/cr.2007.91>.
- Shaw, W. v. (1975) '[57] Chloramphenicol Acetyltransferase from Chloramphenicol-Resistant Bacteria', *Methods in Enzymology*, 43(C), pp. 737–755. Available at: [https://doi.org/10.1016/0076-6879\(75\)43141-X](https://doi.org/10.1016/0076-6879(75)43141-X).
- Shen, C.R. *et al.* (2011) 'Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*', *Applied and Environmental Microbiology*, 77(9), pp. 2905–2915. Available at: <https://doi.org/10.1128/AEM.03034-10>.
- Shinto, H. *et al.* (2007) 'Kinetic modeling and sensitivity analysis of acetone-butanol-ethanol production', *Journal of Biotechnology*, 131(1), pp. 45–56. Available at: <https://doi.org/10.1016/j.jbiotec.2007.05.005>.
- Shinto, H. *et al.* (2008) 'Kinetic study of substrate dependency for higher butanol production in acetone-butanol-ethanol fermentation', *Process Biochemistry*, 43(12), pp. 1452–1461. Available at: <https://doi.org/10.1016/j.procbio.2008.06.003>.
- Shlomi, T., Berkman, O. and Ruppin, E. (2005) 'Regulatory on/off minimization of metabolic flux changes after genetic perturbations', *Proceedings of the National Academy of Sciences of the United States of America*, 102(21), pp. 7695–7700. Available at: <https://doi.org/10.1073/pnas.0406346102>.
- Sievers, F. *et al.* (2011) 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Molecular Systems Biology*, 7(1), p. 539. Available at: <https://doi.org/10.1038/MSB.2011.75>.
- Sillers, R. *et al.* (2008) 'Metabolic engineering of the non-sporulating, non-solventogenic *Clostridium acetobutylicum* strain M5 to produce butanol without acetone demonstrate the robustness of the acid-formation pathways and the importance of the electron balance', *Metabolic Engineering*, 10(6), pp. 321–332. Available at: <https://doi.org/10.1016/j.ymben.2008.07.005>.

- Singh, R. *et al.* (2008) 'A novel strategy involved anti-oxidative defense: The conversion of NADH into NADPH by a metabolic network', *PLoS ONE*, 3(7), p. e2682. Available at: <https://doi.org/10.1371/journal.pone.0002682>.
- Snoep, J.L. *et al.* (1994) 'Reconstruction of glucose uptake and phosphorylation in a glucose-negative mutant of *Escherichia coli* by using *Zymomonas mobilis* genes encoding the glucose facilitator protein and glucokinase', *Journal of Bacteriology*, 176(7), pp. 2133–2135. Available at: <https://doi.org/10.1128/JB.176.7.2133-2135.1994>.
- Soni, B.K., Soucaille, P. and Goma, G. (1987) 'Continuous acetone-butanol fermentation: a global approach for the improvement in the solvent productivity in synthetic medium', *Applied Microbiology and Biotechnology*, 25(4), pp. 317–321. Available at: <https://doi.org/10.1007/BF00252540>.
- Srivastava, A.K. and Volesky, B. (1990) 'Updated model of the batch acetone-butanol fermentation', *Biotechnology Letters*, 12(9), pp. 693–698. Available at: <https://doi.org/10.1007/BF01088196>.
- Steiner, E. *et al.* (2011) 'Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*', *Molecular Microbiology*, 80(3), pp. 641–654. Available at: <https://doi.org/10.1111/J.1365-2958.2011.07608.X>.
- Stephanopoulos, G. (2002) 'Metabolic engineering: Perspective of a chemical engineer', *AIChE Journal*, pp. 920–926. Available at: <https://doi.org/10.1002/aic.690480502>.
- Stim-Herndon, K.P., Petersen, D.J. and Bennett, G.N. (1995) 'Characterization of an acetyl-CoA C-acetyltransferase (thiolase) gene from *Clostridium acetobutylicum* ATCC 824', *Gene*, 154(1), pp. 81–85. Available at: [https://doi.org/10.1016/0378-1119\(94\)00838-J](https://doi.org/10.1016/0378-1119(94)00838-J).
- Strutz, J. *et al.* (2019) 'Metabolic kinetic modeling provides insight into complex biological questions, but hurdles remain', *Current Opinion in Biotechnology*, 59, pp. 24–30. Available at: <https://doi.org/10.1016/J.COPBIO.2019.02.005>.
- Sullivan, L. and Bennett, G.N. (2006) 'Proteome analysis and comparison of *Clostridium acetobutylicum* ATCC 824 and Spo0A strain variants', *Journal of Industrial Microbiology and Biotechnology*, 33(4), pp. 298–308. Available at: <https://doi.org/10.1007/s10295-005-0050-7>.
- Suo, Y. *et al.* (2018) 'Butyric acid production from lignocellulosic biomass hydrolysates by engineered *Clostridium tyrobutyricum* overexpressing Class I heat shock protein GroESL', *Bioresource Technology*, 250(October 2017), pp. 691–698. Available at: <https://doi.org/10.1016/j.biortech.2017.11.059>.
- Tangney, M. and Mitchell, W.J. (2007) 'Characterisation of a glucose phosphotransferase system in *Clostridium acetobutylicum* ATCC 824', *Applied Microbiology and Biotechnology*, 74(2), pp. 398–405. Available at: <https://doi.org/10.1007/s00253-006-0679-9>.
- Tashiro, Y. *et al.* (2004) 'High Butanol Production by *Clostridium saccharoperbutylacetonicum* N1-4 in Fed-Batch Culture with pH-Stat Continuous Butyric Acid and Glucose Feeding Method', *Journal of bioscience and bioengineering*, 98(4), pp. 263–268. Available at: [https://doi.org/10.1016/S1389-1723\(04\)00279-8](https://doi.org/10.1016/S1389-1723(04)00279-8).

- Tepper, N. and Shlomi, T. (2010) 'Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways', *Bioinformatics*, 26(4), pp. 536–543. Available at: <https://doi.org/10.1093/bioinformatics/btp704>.
- Terracciano, J.S. and Kashket, E.R. (1986) 'Intracellular Conditions Required for Initiation of Solvent Production by *Clostridium acetobutylicum* Intracellular Conditions Required for Initiation of Solvent Production by *Clostridium acetobutylicum*', *Applied and environmental microbiology*, 52(1), pp. 86–91.
- Thang, V.H., Kanda, K. and Kobayashi, G. (2010) 'Production of Acetone-Butanol-Ethanol (ABE) in direct fermentation of cassava by *Clostridium saccharoperbutylacetonicum* N1-4', *Applied Biochemistry and Biotechnology*, 161(1–8), pp. 157–170. Available at: <https://doi.org/10.1007/s12010-009-8770-1>.
- Thauer, R.K., Kirchniawy, F.H. and Jungermann, K.A. (1972) 'Properties and Function of the Pyruvate-Formate-Lyase Reaction in Clostridia', *European Journal of Biochemistry*, 27(2), pp. 282–290. Available at: <https://doi.org/10.1111/J.1432-1033.1972.TB01837.X>.
- Thompson, R.A. *et al.* (2015) 'Elucidating central metabolic redox obstacles hindering ethanol production in *Clostridium thermocellum*', *Metabolic Engineering*, 32, pp. 207–219. Available at: <https://doi.org/10.1016/j.ymben.2015.10.004>.
- Thompson, R.A. *et al.* (2016) 'Exploring complex cellular phenotypes and model-guided strain design with a novel genome-scale metabolic model of *Clostridium thermocellum* DSM 1313 implementing an adjustable cellulosome', *Biotechnology for Biofuels*. BioMed Central Ltd., p. 194. Available at: <https://doi.org/10.1186/s13068-016-0607-x>.
- Tishkov, V.I. and Popov, V.O. (2006) 'Protein engineering of formate dehydrogenase', *Biomolecular Engineering*, 23(2–3), pp. 89–110. Available at: <https://doi.org/10.1016/J.BIOENG.2006.02.003>.
- Tokic, M., Hatzimanikatis, V. and Miskovic, L. (2020) 'Large-scale kinetic metabolic models of *Pseudomonas putida* KT2440 for consistent design of metabolic engineering strategies', *Biotechnology for Biofuels*, 13(1), pp. 1–19. Available at: <https://doi.org/10.1186/S13068-020-1665-7/FIGURES/3>.
- Tomas, C.A. *et al.* (2003) 'DNA array-based transcriptional analysis of asporogenous, nonsolventogenic *Clostridium acetobutylicum* strains SKO1 and M5', *Journal of Bacteriology*, 185(15), pp. 4539–4547. Available at: <https://doi.org/10.1128/JB.185.15.4539-4547.2003>.
- Toya, Y. *et al.* (2010) '13C-metabolic flux analysis for batch culture of *Escherichia coli* and its pyk and pgi gene knockout mutants based on mass isotopomer distribution of intracellular metabolites', *Biotechnology Progress*, 26(4), pp. 975–992. Available at: <https://doi.org/10.1002/BTPR.420>.
- Toya, Y. and Shimizu, H. (2013) 'Flux analysis and metabolomics for systematic metabolic engineering of microorganisms', *Biotechnology Advances*. Elsevier, pp. 818–826. Available at: <https://doi.org/10.1016/j.biotechadv.2013.05.002>.
- Tran, L.M., Rizk, M.L. and Liao, J.C. (2008) 'Ensemble modeling of metabolic networks', *Biophysical Journal*, 95(12), pp. 5606–5617. Available at: <https://doi.org/10.1529/biophysj.108.135442>.

- Twarog, R. and Wolfe, R.S. (1962) 'Enzymatic phosphorylation of butyrate.', *The Journal of biological chemistry*, 237, pp. 2474–2477. Available at: [https://doi.org/10.1016/S0021-9258\(19\)73776-4](https://doi.org/10.1016/S0021-9258(19)73776-4).
- Ujor, V. *et al.* (2015) 'Allopurinol-mediated lignocellulose-derived microbial inhibitor tolerance by *Clostridium beijerinckii* during acetone–butanol–ethanol (ABE) fermentation', *Applied Microbiology and Biotechnology*, 99(8), pp. 3729–3740. Available at: <https://doi.org/10.1007/s00253-015-6450-3>.
- Valentine, R.C. and Wolfe, R.S. (1960) 'Purification and Role of Phosphotransbutyrylase', *Journal of Biological Chemistry*, 235(7), pp. 1948–1952. Available at: [https://doi.org/10.1016/S0021-9258\(18\)69341-X](https://doi.org/10.1016/S0021-9258(18)69341-X).
- Vallino, J.J., Hopkinson, C.S. and Hobbie, J.E. (1996) 'Modeling bacterial utilization of dissolved organic matter: Optimization replaces Monod growth kinetics', *Limnology and Oceanography*, 41(8), pp. 1591–1609. Available at: <https://doi.org/10.4319/LO.1996.41.8.1591>.
- Varma, A. and Palsson, B.O. (1994) 'Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By product secretion in wild type *Escherichia coli* W3110', *Applied and Environmental Microbiology*, 60(10), pp. 3724–3731. Available at: <https://doi.org/PMC201879>.
- Vasconcelos, I., Girbal, L. and Soucaille, P. (1994) 'Regulation of carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol', *Journal of Bacteriology*, 176(5), pp. 1443–1450. Available at: <https://doi.org/10.1128/jb.176.5.1443-1450.1994>.
- Ventura, J.R.S.S., Hu, H. and Jahng, D. (2013) 'Enhanced butanol production in *Clostridium acetobutylicum* ATCC 824 by double overexpression of 6-phosphofructokinase and pyruvate kinase genes', *Applied Microbiology and Biotechnology*, 97(16), pp. 7505–7516. Available at: <https://doi.org/10.1007/s00253-013-5075-7>.
- Verdon, G. *et al.* (2003) 'Crystal Structures of the ATPase Subunit of the Glucose ABC Transporter from *Sulfolobus solfataricus*: Nucleotide-free and Nucleotide-bound Conformations', *Journal of Molecular Biology*, 330(2), pp. 343–358. Available at: [https://doi.org/10.1016/S0022-2836\(03\)00575-8](https://doi.org/10.1016/S0022-2836(03)00575-8).
- Veza, I., Muhamad Said, M.F. and Latiff, Z.A. (2021) 'Recent advances in butanol production by acetone-butanol-ethanol (ABE) fermentation', *Biomass and Bioenergy*, 144, p. 105919. Available at: <https://doi.org/10.1016/J.BIOMBIOE.2020.105919>.
- Voigt, C., Bahl, H. and Fischer, R.J. (2014) 'Identification of PTSFru as the major fructose uptake system of *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 98(16), pp. 7161–7172. Available at: <https://doi.org/10.1007/S00253-014-5809-1/FIGURES/6>.
- Vollherbst Schneck, K., Sands, J.A. and Montenecourt, B.S. (1984) 'Effect of butanol on lipid composition and fluidity of *Clostridium acetobutylicum* ATCC 824', *Applied and Environmental Microbiology*, 47(1), pp. 193–194.
- Wada, M., Narita, K. and Yokota, A. (2007) 'Alanine production in an H⁺-ATPase- and lactate dehydrogenase-defective mutant of *Escherichia coli* expressing alanine dehydrogenase', *Applied Microbiology and Biotechnology*, 76(4), pp. 819–825. Available at: <https://doi.org/10.1007/s00253-007-1065-y>.

- van der Wal, H. *et al.* (2013) 'Production of acetone, butanol, and ethanol from biomass of the green seaweed *Ulva lactuca*', *Bioresource Technology*, 128, pp. 431–437. Available at: <https://doi.org/10.1016/j.biortech.2012.10.094>.
- Wallenius, J., Maaheimo, H. and Eerikäinen, T. (2016) 'Carbon 13-Metabolic Flux Analysis derived constraint-based metabolic modelling of *Clostridium acetobutylicum* in stressed chemostat conditions', *Bioresource Technology*, 219, pp. 378–386. Available at: <https://doi.org/10.1016/J.BIORTECH.2016.07.137>.
- Walter, K.A. *et al.* (1993) 'Sequence and arrangement of two genes of the butyrate-synthesis pathway of *Clostridium acetobutylicum* ATCC 824', *Gene*, 134(1), pp. 107–111. Available at: [https://doi.org/10.1016/0378-1119\(93\)90182-3](https://doi.org/10.1016/0378-1119(93)90182-3).
- Walter, K.A., Bennett, G.N. and Papoutsakis, E.T. (1992) 'Molecular characterization of two *Clostridium acetobutylicum* ATCC 824 butanol dehydrogenase isozyme genes', *Journal of Bacteriology*, 174(22), pp. 7149–7158. Available at: <https://doi.org/10.1128/jb.174.22.7149-7158.1992>.
- Wang, M. *et al.* (2015) 'Enhanced 1-butanol production in engineered *Klebsiella pneumoniae* by NADH regeneration', *Energy and Fuels*, 29(3), pp. 1823–1829. Available at: <https://doi.org/10.1021/acs.energyfuels.5b00009>.
- Wang, M., Fan, L. and Tan, T. (2014) '1-Butanol production from glycerol by engineered *Klebsiella pneumoniae*', *RSC Advances*, 4(101), pp. 57791–57798. Available at: <https://doi.org/10.1039/C4RA09016K>.
- Wang, S. *et al.* (2011) 'Formic Acid Triggers the "Acid Crash" of Acetone-Butanol-Ethanol Fermentation by *Clostridium acetobutylicum*', *Applied and Environmental Microbiology*, 77(5), pp. 1674–1680. Available at: <https://doi.org/10.1128/AEM.01835-10>.
- Wang, S. *et al.* (2017) 'Genome Editing in *Clostridium saccharoperbutylacetonicum* N1-4 with the CRISPR-Cas9 System', *Applied and Environmental Microbiology*, 83(10), pp. 1–16.
- Wang, S., Dong, S. and Wang, Y. (2017) 'Enhancement of solvent production by overexpressing key genes of the acetone-butanol-ethanol fermentation pathway in *Clostridium saccharoperbutylacetonicum* N1-4', *Bioresource Technology*, 245(September), pp. 426–433. Available at: <https://doi.org/10.1016/j.biortech.2017.09.024>.
- Wang, Y. *et al.* (2011) 'Single-nucleotide resolution analysis of the transcriptome structure of *Clostridium beijerinckii* NCIMB 8052 using RNA-Seq', *BMC Genomics*, 12(1), pp. 1–10. Available at: <https://doi.org/10.1186/1471-2164-12-479>.
- Wang, Y. *et al.* (2012) 'Genome-wide dynamic transcriptional profiling in *Clostridium beijerinckii* NCIMB 8052 using single-nucleotide resolution RNA-Seq', *BMC Genomics*, 13(1), pp. 1–15. Available at: <https://doi.org/10.1186/1471-2164-13-102/FIGURES/5>.
- Wang, Y. *et al.* (2013) 'Development of a gene knockout system using mobile group II introns (Targetron) and genetic disruption of acid production pathways in *Clostridium beijerinckii*', *Applied and Environmental Microbiology*, 79(19), pp. 5853–5863. Available at: <https://doi.org/10.1128/AEM.00971-13>.

- Wang, Y. *et al.* (2015) 'Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9 system', *Journal of Biotechnology*, 200, pp. 1–5. Available at: <https://doi.org/10.1016/j.jbiotec.2015.02.005>.
- Wang, Y. *et al.* (2016) 'Bacterial Genome Editing with CRISPR-Cas9: Deletion, Integration, Single Nucleotide Modification, and Desirable "clean" Mutant Selection in *Clostridium beijerinckii* as an Example', *ACS Synthetic Biology*, 5(7), pp. 721–732. Available at: <https://doi.org/10.1021/acssynbio.6b00060>.
- Wasels, F. *et al.* (2017) 'A two-plasmid inducible CRISPR/Cas9 genome editing tool for *Clostridium acetobutylicum*', *Journal of Microbiological Methods*, 140(April), pp. 5–11. Available at: <https://doi.org/10.1016/j.mimet.2017.06.010>.
- Wasels, F. *et al.* (2020) 'A CRISPR/Anti-CRISPR genome editing approach underlines the synergy of butanol dehydrogenases in *Clostridium acetobutylicum* DSM 792', *Applied and Environmental Microbiology*, 86(13). Available at: <https://doi.org/10.1128/AEM.00408-20>.
- Waterhouse, A.M. *et al.* (2009) 'Jalview Version 2—a multiple sequence alignment editor and analysis workbench', *Bioinformatics*, 25(9), pp. 1189–1191. Available at: <https://doi.org/10.1093/BIOINFORMATICS/BTP033>.
- Weizmann, C. (1919) 'Production of acetone and alcohol by bacteriological processes', *US patent*, 1(315), p.585.
- Welch, R.W., Rudolph, F.B. and Papoutsakis, E.T. (1989) 'Purification and characterization of the NADH-dependent butanol dehydrogenase from *Clostridium acetobutylicum* (ATCC 824)', *Archives of Biochemistry and Biophysics*, 273(2), pp. 309–318. Available at: [https://doi.org/10.1016/0003-9861\(89\)90489-X](https://doi.org/10.1016/0003-9861(89)90489-X).
- Wiechert, W. (2001) '13C metabolic flux analysis', *Metabolic Engineering*, 3(3), pp. 195–206. Available at: <https://doi.org/10.1006/mben.2001.0187>.
- Wiesenborn, D.P., Rudolph, F.B. and Papoutsakis, E.T. (1988) 'Thiolase from *Clostridium acetobutylicum* ATCC 824 and Its Role in the Synthesis of Acids and Solvents.', *Applied and environmental microbiology*, 54(11), pp. 2717–22.
- Wietzke, M. and Bahl, H. (2012) 'The redox-sensing protein Rex, a transcriptional regulator of solventogenesis in *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 96(3), pp. 749–761. Available at: <https://doi.org/10.1007/s00253-012-4112-2>.
- Wilkinson, S.R. and Young, M. (1994) 'Targeted integration of genes into the *Clostridium acetobutylicum* chromosome', *Microbiology*, 140(1), pp. 89–95. Available at: <https://doi.org/10.1099/13500872-140-1-89>.
- Winzer, K. *et al.* (2000) 'Differential Regulation of Two Thiolase Genes from *Clostridium acetobutylicum* DSM 792', *J. Mol. Microbiol. Biotechnol*, 2(4), pp. 531–541. Available at: www.caister.com/bacteria-plant (Accessed: 3 January 2022).
- Winzer, K., Lorenz, K. and Dürre, P. (1997) 'Acetate kinase from *Clostridium acetobutylicum*: A highly specific enzyme that is actively transcribed during acidogenesis and solventogenesis', *Microbiology*, 143(10), pp. 3279–3286. Available at: <https://doi.org/10.1099/00221287-143-10-3279>.

- Woolley, R.C. and Morris, J.G. (1990) 'Stability of solvent production by *Clostridium acetobutylicum* in continuous culture: strain differences', *Journal of Applied Bacteriology*, 69(5), pp. 718–728. Available at: <https://doi.org/10.1111/j.1365-2672.1990.tb01569.x>.
- Woolston, B.M. *et al.* (2018) 'Rediverting carbon flux in *Clostridium ljungdahlii* using CRISPR interference (CRISPRi)', *Metabolic Engineering*, 48, pp. 243–253. Available at: <https://doi.org/10.1016/J.YMBEN.2018.06.006>.
- Wu, C. *et al.* (2015) 'Genome-Based Metabolic Mapping and ¹³C Flux Analysis Reveal Systematic Properties of an Oleaginous Microalga *Chlorella protothecoides*', *Plant Physiology*, 167(2), pp. 586–599. Available at: <https://doi.org/10.1104/PP.114.250688>.
- Wu, H. *et al.* (2017) 'Effects of pH and ferrous iron on the coproduction of butanol and hydrogen by *Clostridium beijerinckii* IB4', *International Journal of Hydrogen Energy*, 42(10), pp. 6547–6555. Available at: <https://doi.org/10.1016/J.IJHYDENE.2017.02.094>.
- Wynn, J.P. and Ratledge, C. (1997) 'Malic enzyme is a major source of NADPH for lipid accumulation by *Aspergillus nidulans*', *Microbiology*, 143(1), pp. 253–257. Available at: <https://doi.org/10.1099/00221287-143-1-253>.
- Xiao, H. *et al.* (2011) 'Confirmation and elimination of xylose metabolism bottlenecks in glucose phosphoenolpyruvate-dependent phosphotransferase system-deficient *Clostridium acetobutylicum* for simultaneous utilization of glucose, xylose, and arabinose', *Applied and Environmental Microbiology*, 77(22), pp. 7886–7895. Available at: <https://doi.org/10.1128/AEM.00644-11>.
- Yang, Z. *et al.* (2020) 'Effects of Spo0A on *Clostridium acetobutylicum* with an emphasis on biofilm formation', *World Journal of Microbiology and Biotechnology*, 36(3), p. 80. Available at: <https://doi.org/10.1007/s11274-020-02859-6>.
- Yao, D. *et al.* (2017) 'Robustness of *Clostridium saccharoperbutylacetonicum* for acetone-butanol-ethanol production: Effects of lignocellulosic sugars and inhibitors', *Fuel*, 208, pp. 549–557. Available at: <https://doi.org/10.1016/j.fuel.2017.07.004>.
- Yen, J.Y. *et al.* (2015) 'Designing metabolic engineering strategies with genome-scale metabolic flux modeling', *Advances in Genomics and Genetics*, 5, pp. 93–105. Available at: <https://doi.org/10.2147/AGG.S58494>.
- Yerushalmi, L., Volesky, B. and Szczesny, T. (1985) 'Effect of increased hydrogen partial pressure on the acetone-butanol fermentation by *Clostridium acetobutylicum*', *Applied Microbiology and Biotechnology*, 22(2), pp. 103–107. Available at: <https://doi.org/10.1007/BF00250028>.
- Yim, H. *et al.* (2011) 'Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol', *Nature Chemical Biology*, 7(7), pp. 445–452. Available at: <https://doi.org/10.1038/nchembio.580>.
- Yokota, A. *et al.* (1997) 'Enhancement of glucose metabolism in a pyruvic acid-hyperproducing *Escherichia coli* mutant defective in F1-ATPase activity', *Journal of Fermentation and Bioengineering*, 83(2), pp. 132–138. Available at: [https://doi.org/10.1016/S0922-338X\(97\)83571-4](https://doi.org/10.1016/S0922-338X(97)83571-4).

- Yoo, M. *et al.* (2015) 'A quantitative system-scale characterization of the metabolism of *Clostridium acetobutylicum*', *mBio*, 6(6), pp. 1–12. Available at: <https://doi.org/10.1128/mBio.01808-15>.
- Yoo, M. *et al.* (2016) 'Elucidation of the roles of adhE1 and adhE2 in the primary metabolism of *Clostridium acetobutylicum* by combining in-frame gene deletion and a quantitative system-scale approach', *Biotechnology for Biofuels*, 9(1), pp. 1–15. Available at: <https://doi.org/10.1186/s13068-016-0507-0>.
- Yoo, M. *et al.* (2017) 'Metabolic flexibility of a butyrate pathway mutant of *Clostridium acetobutylicum*', *Metabolic Engineering*, 40(December 2016), pp. 138–147. Available at: <https://doi.org/10.1016/j.ymben.2017.01.011>.
- Yoo, M., Nguyen, N.P.T. and Soucaille, P. (2020) 'Trends in Systems Biology for the Analysis and Engineering of *Clostridium acetobutylicum* Metabolism', *Trends in Microbiology*, 28(2), pp. 118–140. Available at: <https://doi.org/10.1016/J.TIM.2019.09.003>.
- Young, J.D. (2014) 'INCA: a computational platform for isotopically non-stationary metabolic flux analysis', *Bioinformatics*, 30(9), pp. 1333–1335. Available at: <https://doi.org/10.1093/BIOINFORMATICS/BTU015>.
- Yu, L. *et al.* (2015) 'Metabolic engineering of *Clostridium tyrobutyricum* for n-butanol production through co-utilization of glucose and xylose', *Biotechnology and Bioengineering*, 112(10), pp. 2134–2141. Available at: <https://doi.org/10.1002/bit.25613>.
- Yu, M. *et al.* (2011) 'Metabolic engineering of *Clostridium tyrobutyricum* for n-butanol production', *Metabolic Engineering*, 13(4), pp. 373–382. Available at: <https://doi.org/10.1016/j.ymben.2011.04.002>.
- Zamboni, N. *et al.* (2009) '13C-based metabolic flux analysis', *Nature Protocols*, 4(6), pp. 878–892. Available at: <https://doi.org/10.1038/nprot.2009.58>.
- Zeikus, J.G. (1980) 'Chemical and fuel production by anaerobic bacteria', *Ann. Rev. Microbiol*, 34, pp. 423–64. Available at: www.annualreviews.org (Accessed: 26 April 2022).
- Zhang, J. *et al.* (2018) 'Enhancement of sucrose metabolism in *Clostridium saccharoperbutylacetonicum* N1-4 through metabolic engineering for improved acetone–butanol–ethanol (ABE) fermentation', *Bioresource Technology*, 270, pp. 430–438. Available at: <https://doi.org/10.1016/j.biortech.2018.09.059>.
- Zhang, L. *et al.* (2014) 'Redox-responsive repressor rex modulates alcohol production and oxidative stress tolerance in *Clostridium acetobutylicum*', *Journal of Bacteriology*, 196(22), pp. 3949–3963. Available at: <https://doi.org/10.1128/JB.02037-14>.
- Zhang, Y. *et al.* (2009) 'Introduction of an NADH regeneration system into *Klebsiella oxytoca* leads to an enhanced oxidative and reductive metabolism of glycerol', *Metabolic Engineering*, 11(2), pp. 101–106. Available at: <https://doi.org/10.1016/j.ymben.2008.11.001>.
- Zhang, Y. *et al.* (2016) 'Development of an inducible transposon system for efficient random mutagenesis in *Clostridium acetobutylicum*', *FEMS Microbiology Letters*, 363(8), pp. 1–7. Available at: <https://doi.org/10.1093/femsle/fnw065>.
- Zhang, Y. and Ezeji, T.C. (2014) 'Elucidating and alleviating impacts of lignocellulose-derived microbial inhibitors on *Clostridium beijerinckii* during fermentation of *Miscanthus giganteus* to

butanol', *Journal of Industrial Microbiology and Biotechnology*, 41(10), pp. 1505–1516. Available at: <https://doi.org/10.1007/s10295-014-1493-5>.

Zhang, Y., Grosse-Honebrink, A. and Minton, N.P. (2015) 'A universal mariner transposon system for forward genetic studies in the genus *Clostridium*', *PLoS ONE*. Edited by D. Paredes-Sabja, 10(4), p. e0122411. Available at: <https://doi.org/10.1371/journal.pone.0122411>.

Zhao, Y. *et al.* (2005) 'Intracellular butyryl phosphate and acetyl phosphate concentrations in *Clostridium acetobutylicum* and their implications for solvent formation', *Applied and Environmental Microbiology*, 71(1), pp. 530–537. Available at: <https://doi.org/10.1128/AEM.71.1.530-537.2005>.

Zhen, X., Wang, Y. and Liu, D. (2020) 'Bio-butanol as a new generation of clean alternative fuel for SI (spark ignition) and CI (compression ignition) engines', *Renewable Energy* 147, pp. 2494–2521. Available at: <https://doi.org/10.1016/j.renene.2019.10.119>.

Zhou, Q., Liu, Y. and Yuan, W. (2018) 'Kinetic modeling of lactic acid and acetic acid effects on butanol fermentation by *Clostridium saccharoperbutylacetonicum*', *Fuel*, 226(October 2017), pp. 181–189. Available at: <https://doi.org/10.1016/j.fuel.2018.04.019>.

Zhou, Q., Liu, Y. and Yuan, W. (2020) 'Kinetic modeling of butyric acid effects on butanol fermentation by *Clostridium saccharoperbutylacetonicum*', *New Biotechnology*, 55, pp. 118–126.

Appendix A

The diagram illustrates the central carbon metabolism of *E. coli*, showing the flow of carbon compounds and the associated enzymatic reactions. Key pathways include:

- Glycolysis:** Glucose (Glu) is converted to Glucose-6-phosphate (G6P) by PEP-dependent Glucose-6-phosphatase (GLCpts). G6P is then converted to Fructose-6-phosphate (F6P) by Phosphoglucose isomerase (PGI). F6P is converted to Fructose-1,6-bisphosphate (F-1,6-BP) by Phosphofructokinase (PFK), which is regulated by ATP/ADP. F-1,6-BP is cleaved by FBPase into Dihydroxyacetone phosphate (DHAP) and Glyceraldehyde-3-phosphate (GAP). DHAP is converted to GAP by Triose phosphate isomerase (TPI).
- Gluconeogenesis:** GAP is converted to 1,3-bisphosphoglycerate (1,3-DPG) by GAPDH, which is regulated by NAD⁺/NADH. 1,3-DPG is converted to 3-phosphoglycerate (3PG) by Phosphoglycerate kinase (PGK), which is regulated by ADP/ATP. 3PG is converted to 2-phosphoglycerate (2PG) by Phosphoglycerate mutase (PGM). 2PG is converted to Phosphoenolpyruvate (PEP) by Enolase (ENO). PEP is converted to Pyruvate (PYR) by Pyruvate kinase (PYK), which is regulated by ADP/ATP.
- Biosynthetic Pathways:**
 - Amino Acid Biosynthesis:** PEP is converted to Lactate (LAC) by Lactate dehydrogenase (LDH), which is regulated by NAD⁺/NADH. PYR is converted to Amino Acid Biosynthesis by Amino Acid Biosynthesis (AAB). PYR is also converted to Acetyl-CoA (AcCoA) by Pyruvate dehydrogenase (PDH), which is regulated by NAD⁺/NADH. AcCoA is converted to Citrate (CIT) by Citrate synthase (CS). CIT is converted to Isocitrate (ICIT) by Aconitase (ACONT). ICIT is converted to α-Ketoglutarate (αKG) by Isocitrate dehydrogenase (ICDH), which is regulated by NAD⁺/NADPH. αKG is converted to Succinyl-CoA (SUCCoA) by α-Ketoglutarate dehydrogenase (αKGDH), which is regulated by NAD⁺/NADPH. SUCCoA is converted to Succinate (SUCC) by Succinyl-CoA lyase (SCL), which is regulated by ATP/ADP. SUCC is converted to Succinyl-CoA by Succinate dehydrogenase (SDH), which is regulated by Ubiquinol/Ubiquinone.
 - Other Biosynthetic Pathways:** F6P is converted to Glucose-6-phosphate (G6P) by Glucose-6-phosphatase (GLCpts), which is regulated by PEP/Glcpts and ATP/Glcpts. F6P is converted to Fructose-6-phosphate (F6P) by Phosphoglucose isomerase (PGI). F6P is converted to Fructose-1,6-bisphosphate (F-1,6-BP) by Phosphofructokinase (PFK), which is regulated by ATP/ADP. F-1,6-BP is cleaved by FBPase into DHAP and GAP. DHAP is converted to GAP by TPI. GAP is converted to 1,3-DPG by GAPDH, which is regulated by NAD⁺/NADH. 1,3-DPG is converted to 3PG by PGK, which is regulated by ADP/ATP. 3PG is converted to 2PG by PGM. 2PG is converted to PEP by ENO. PEP is converted to PYR by PYK, which is regulated by ADP/ATP. PYR is converted to LAC by LDH, which is regulated by NAD⁺/NADH. PYR is converted to AcCoA by PDH, which is regulated by NAD⁺/NADH. AcCoA is converted to CIT by CS. CIT is converted to ICIT by ACONT. ICIT is converted to αKG by ICDH, which is regulated by NAD⁺/NADPH. αKG is converted to SUCCoA by αKGDH, which is regulated by NAD⁺/NADPH. SUCCoA is converted to SUCC by SCL, which is regulated by ATP/ADP. SUCC is converted to SUCCoA by SDH, which is regulated by Ubiquinol/Ubiquinone.

Appendix B

Table of reactions in central metabolism shown in metabolic network in Appendix A. Arrows in the reaction equation show the reversibility of the reaction; arrows shown as ' $\rightarrow$ ' are only allowed to proceed in the forward direction, while arrows shown as ' $\rightleftharpoons$ ' are reversible. The ' $_c$ ' and ' $_e$ ' designate the location of the metabolite, being intracellular and extracellular, respectively. Lumped equations for synthesis of biomass and pseudometabolite precursors are included.

Reaction ID	Reaction Name	Reaction Equation
biomass	biomass synthesis	$40.0 \text{ atp}_c + 0.0432 \text{ carbo_met}_c + 0.026 \text{ dna_met}_c + 40.0 \text{ h}_2\text{o}_c + 0.1008 \text{ peptido_met}_c + 0.076 \text{ plipid_met}_c + 0.5284 \text{ protein_met}_c + 0.0655 \text{ rna_met}_c + 0.1106 \text{ teich_met}_c + 0.0495 \text{ trace_met}_c \rightarrow 40.0 \text{ adp}_c + 40.0 \text{ pi}_c$
GLCpts	Transport of glucose	$\text{glc_D}_e + \text{pep}_c \rightarrow \text{g6p}_c + \text{pyr}_c$
GLCabc	Transport of glucose	$\text{atp}_c + \text{glc_D}_e + \text{h}_2\text{o}_c \rightarrow \text{adp}_c + \text{glc_D}_c + \text{h}_c + \text{pi}_c$
PGI	D-glucose-6-phosphate aldose-ketose-isomerase	$\text{g6p}_c \rightleftharpoons \text{f6p}_c$
PGL	6-Phospho-D-glucono-1,5-lactone lactonohydrolase	$6\text{pgl}_c + \text{h}_2\text{o}_c \rightarrow 6\text{pgc}_c + \text{h}_c$
EDD	phosphogluconate dehydratase	$6\text{pgc}_c \rightarrow 2\text{ddg6p}_c + \text{h}_2\text{o}_c$
EDA	2-dehydro-3-deoxy-D-gluconate-6-phosphate	$2\text{ddg6p}_c \rightarrow \text{g3p}_c + \text{pyr}_c$
RPI	D-Ribose-5-phosphate ketol-isomerase	$\text{r5p}_c \rightleftharpoons \text{ru5p_D}_c$
RPE	D-Ribulose-5-phosphate 3-epimerase	$\text{ru5p_D}_c \rightleftharpoons \text{xu5p_D}_c$
TALA	sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate	$\text{e4p}_c + \text{f6p}_c \rightarrow \text{g3p}_c + \text{s7p}_c$
TKT1	Sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate	$\text{g3p}_c + \text{s7p}_c \rightleftharpoons \text{r5p}_c + \text{xu5p_D}_c$
TKT2	D-Fructose 6-phosphate:D-glyceraldehyde-3-phosphate glycolaldehyde	$\text{f6p}_c + \text{g3p}_c \rightarrow \text{e4p}_c + \text{xu5p_D}_c$
PFK	6-phosphofructokinase	$\text{atp}_c + \text{f6p}_c \rightarrow \text{adp}_c + \text{f6p}_c + \text{h}_c$
FBA	D-Fructose-1,6-bisphosphate D-glyceraldehyde-3-phosphate-lyase	$\text{f6p}_c \rightleftharpoons \text{dhap}_c + \text{g3p}_c$
TPI	D-Glyceraldehyde-3-phosphate ketol-isomerase	$\text{g3p}_c \rightleftharpoons \text{dhap}_c$
GAPD	D-Glyceraldehyde-3-phosphate:NAD ⁺ oxidoreductase(phosphorylating)	$\text{g3p}_c + \text{nad}_c + \text{pi}_c \rightleftharpoons 1,3\text{dpg}_c + \text{h}_c + \text{nadh}_c$
PGK	ATP:3-phospho-D-glycerate 1-phosphotransferase	$3\text{pg}_c + \text{atp}_c \rightleftharpoons 1,3\text{dpg}_c + \text{adp}_c$
PGM	2-Phospho-D-glycerate 2,3-phosphomutase	$2\text{pg}_c \rightleftharpoons 3\text{pg}_c$
ENO	2-Phospho-D-glucerate hydro-lyase	$2\text{pg}_c \rightleftharpoons \text{h}_2\text{o}_c + \text{pep}_c$
PYK	ATP:pyruvate O ₂ -phosphotransferase	$\text{adp}_c + \text{h}_c + \text{pep}_c \rightarrow \text{atp}_c + \text{pyr}_c$
POR4	pyruvate synthase	$\text{coa}_c + \text{fdxox}_c + \text{pyr}_c \rightarrow \text{accoa}_c + \text{co}_2_c + \text{fdxrd}_c + \text{h}_c$

PFL	Acetyl-CoA:formate C- acetyltransferase	coa_c + pyr_c --> accoa_c + for_c
FDH	Formate:NAD+ oxidoreductase	for_c + nad_c --> co2_c + nadh_c
FDHy	formate dehydrogenase (NADP+)	for_c + nadp_c --> co2_c + nadph_c
LDH_L	(S)-Lactate:NAD+ oxidoreductase	lac_L_c + nad_c <=> h_c + nadh_c + pyr_c
PC	Pyruvate:carbon-dioxide ligase (ADP-forming)	atp_c + hco3_c + pyr_c --> adp_c + h_c + oaa_c + pi_c
CS	Citrate Synthase	accoa_c + h2o_c + oaa_c --> cit_c + coa_c + h_c
ACONT	citrate hydro-lyase	cit_c <=> icit_c
ICDHy	Isocitrate:NADP+ oxidoreductase (decarboxylating)	icit_c + nadp_c <=> akg_c + co2_c + nadph_c
OOR	Pyruvate ferredoxin/flavodoxin oxidoreductase	akg_c + coa_c + fdxox_c <=> co2_c + fdxrd_c + h_c + succoa_c
SUCOAS	succinyl-CoA synthetase (ADP-forming)	adp_c + pi_c + succoa_c --> atp_c + coa_c + succ_c
FRD7	succinate dehydrogenase (ubiquinone)	fum_c + q8h2_c --> q8_c + succ_c
FUM	(S)-Malate hydro-lyase	mal_L_c --> fum_c + h2o_c
MDH	malate dehydrogenase	h_c + nadh_c + oaa_c --> mal_L_c + nad_c
ME1	(S)-Malate:NADP+ oxidoreductase(oxaloaceta te-decarboxylating)	mal_L_c + nad_c --> co2_c + nadh_c + pyr_c
ME2	(S)-Malate:NADP+ oxidoreductase(oxaloaceta te-decarboxylating)	mal_L_c + nadp_c --> co2_c + nadph_c + pyr_c
PTA	phosphate acetyltransferase	accoa_c + pi_c <=> actp_c + coa_c
ACK	ATP:acetate phosphotransferase	actp_c + adp_c <=> ac_c + atp_c
COAT1	butyryl-CoA-acetoacetate CoA-transferase	aacoa_c + ac_c --> acac_c + accoa_c
ADC	Acetoacetate carboxy- lyase	acac_c + h_c --> acetone_c + co2_c
ACACT1	Acetyl-CoA:acetyl-CoA C- acetyltransferase	2.0 accoa_c <=> aacoa_c + coa_c
HACD1x	(S)-3-Hydroxybutanoyl- CoA:NAD+ oxidoreductase	aacoa_c + h_c + nadh_c --> 3hbcoa_c + nad_c
HACD1y	(S)-3-Hydroxybutanoyl- CoA:NADP+ oxidoreductase	aacoa_c + h_c + nadph_c --> 3hbcoa_c + nadp_c
ECOA1	(S)-3-Hydroxybutanoyl- CoA hydro-lyase	3hbcoa_c <=> b2coa_c + h2o_c
ACOAD1	butyryl-CoA dehydrogenase	b2coa_c + fdxox_c + 2.0 nadh_c --> btcoa_c + fdxrd_c + 2.0 nad_c
BCOPBT	butanoyl-CoA:phosphate butanoyltransferase	btcoa_c + pi_c <=> butylphos_c + coa_c
BUTK	ATP:butyrate 1- phosphotransferase	adp_c + butylphos_c --> atp_c + but_c
COAT2	butyryl-CoA-acetoacetate CoA-transferase	aacoa_c + but_c --> acac_c + btcoa_c
ACALD	acetaldehyde:NAD+ oxidoreductase (CoA- acetylating)	acald_c + coa_c + nad_c <=> accoa_c + h_c + nadh_c
ALCD2	Ethanol:NAD+ oxidoreductase	etoh_c + nad_c <=> acald_c + h_c + nadh_c

BUTALD	butyraldehyde:NAD+ oxidoreductase (CoA-acetylating)	btcoa_c + h_c + nadh_c --> btal_c + coa_c + nad_c
BUTOHDx	butanol dehydrogenase	btal_c + h_c + nadh_c --> butoh_c + nad_c
BUTOHDy	butanol dehydrogenase	btal_c + h_c + nadph_c --> butoh_c + nadp_c
FDXNRx	ferredoxin-NAD+ reductase	fdxrd_c + h_c + nad_c <=> fdxox_c + nadh_c
FDXNRy	ferredoxin-NADP+ reductase	fdxrd_c + h_c + nadp_c --> fdxox_c + nadph_c
FDXNH	ferredoxin hydrogenase	fdxrd_c + 2.0 h_c --> fdxox_c + h2_c
ACex	Transport of acetate	ac_c + h_c <=> ac_e + h_e
BUTex	Transport of butyrate	but_c + h_c <=> but_e + h_e
LAC2ex	Transport of lactate	h_c + lac_L_c --> h_e + lac_L_e
ACETONEex	Transport of acetone	acetone_c --> acetone_e
BUTOHex	Transport of butanol	butoh_c --> butoh_e
ETOHex	Transport of ethanol	etoh_c --> etoh_e
SUCCex	N/A	h_e + succ_e <=> h_c + succ_c

Appendix C

Metabolite IDs and names of metabolites present in the reactions listed in Appendix B.

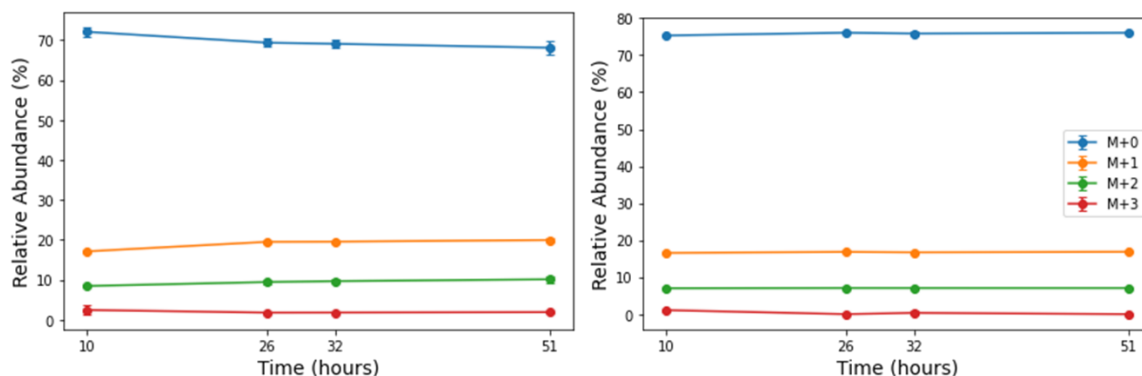
Metabolite ID	Metabolite Name
atp	ATP
carbo_met	Carbohydrate for biomass synthesis
dna_met	DNA for biomass synthesis
h2o	H ₂ O
peptido_met	Peptidoglycan for biomass synthesis
plipid_met	Phospholipid for biomass synthesis
protein_met	Protein for biomass synthesis
rna_met	RNA for biomass synthesis
teich_met	Teichoic acid for biomass synthesis
trace_met	Trace compounds for biomass synthesis
glc_D	D-Glucose
pep	Phosphoenolpyruvate
g6p	D-Glucose 6-phosphate
6pgl	6-phospho-D-glucono-1,5-lactone
6pg	6-Phospho-D-gluconate
2ddg6p	2-Dehydro-3-deoxy-6-phospho-D-gluconate
r5p	alpha-D-Ribose 5-phosphate
ru5p_D	D-Ribulose 5-phosphate
e4p	D-Erythrose 4-phosphate
f6p	D-Fructose 6-phosphate
g3p	Glyceraldehyde 3-phosphate
s7p	Sedoheptulose 7-phosphate
fdp	D-Fructose 1,6-bisphosphate
nad	Nicotinamide adenine dinucleotide
pi	Phosphate
3pg	3-Phospho-D-glycerate
2pg	D-Glycerate 2-phosphate
adp	ADP
h	H ⁺
coa	Coenzyme A
fdxox	Oxidized ferredoxin
pyr	Pyruvate
for	Formate
nadp	Nicotinamide adenine dinucleotide phosphate
lac_L	L-Lactate
hco3	Bicarbonate
accoa	Acetyl-CoA
oaa	Oxaloacetate
cit	Citrate
icit	Isocitrate
akg	2-Oxoglutarate

succoa	Succinyl-CoA
fum	Fumarate
q8h2	Ubiquinol
mal_L	L-Malate
nadh	Nicotinamide adenine dinucleotide - reduced
actp	Acetyl phosphate
aacoa	Acetoacetyl-CoA
ac	Acetate
acac	Acetoacetate
nadph	Nicotinamide adenine dinucleotide phosphate - reduced
3hbcoa	(S)-3-Hydroxybutanoyl-CoA
b2coa	Crotonoyl-CoA
btcoa	Butanoyl-CoA
butylphos	Butanoylphosphate
but	Butanoic acid
acald	Acetaldehyde
etoh	Ethanol
btal	Butanal
fdxrd	Reduced ferredoxin
acetone	Acetone
butoh	1-Butanol
h	H ⁺
succoa	Succinate

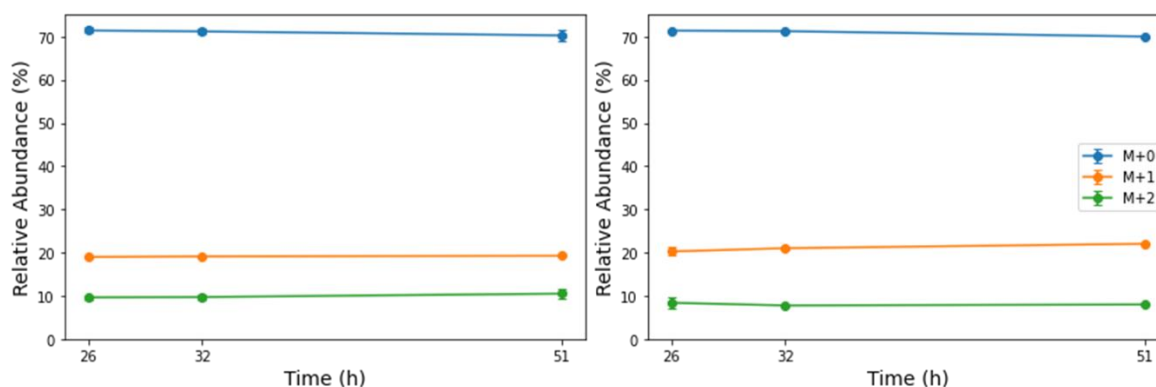
Appendix D

Time course of labelling for amino acids used in this chapter. In the interest of brevity, only a single fragment for each amino acid is presented here. Time points are only shown where the respective fragment was detected. All graphs are presented with the [U-¹³C]glucose experiment on the left and [1-¹³C]glucose experiment on the right. Error bars indicate ± 1 standard deviation from three replicates. Protein hydrolysis and TBDMSCIS derivatisation were performed as described in materials and methods section 2.12.2. GC-MS analysis was performed as described in materials and methods sections 2.12.3 and 2.12.4.

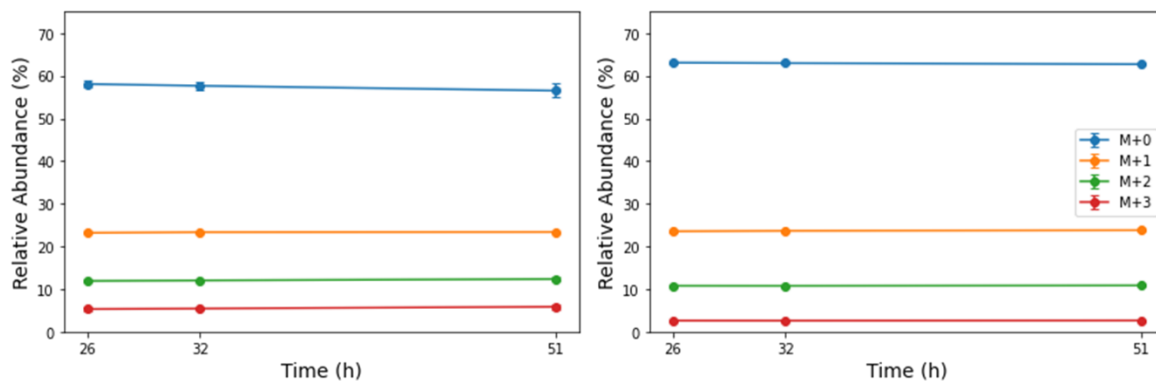
Gly246



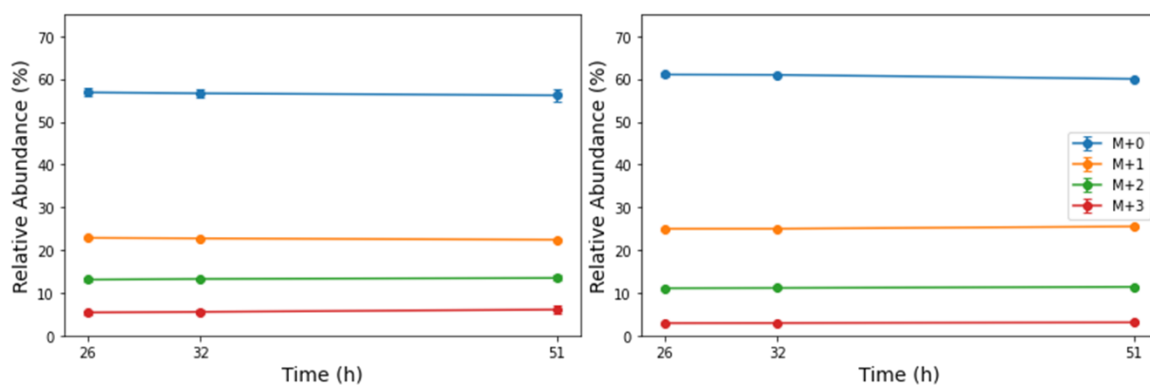
Pro258



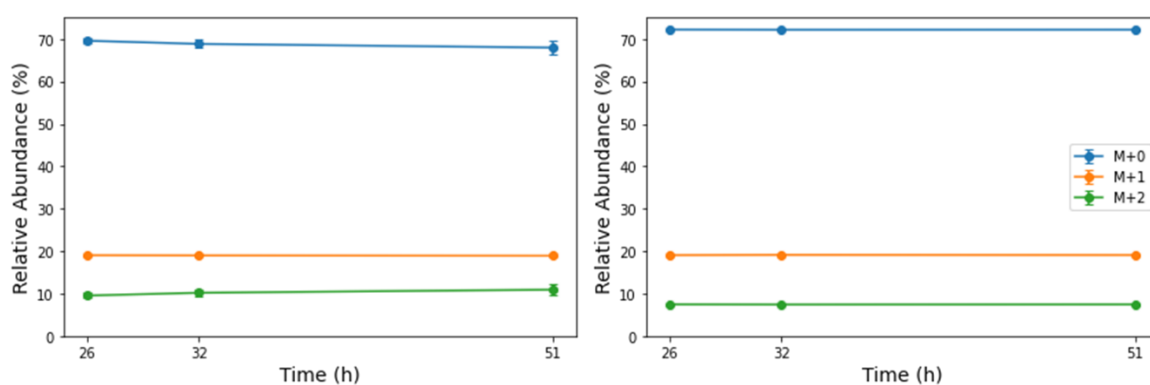
Ser362



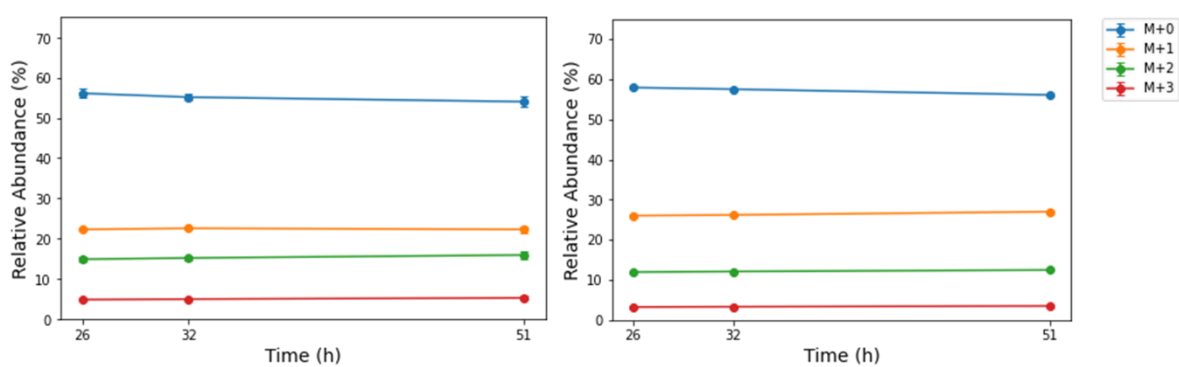
Thr404



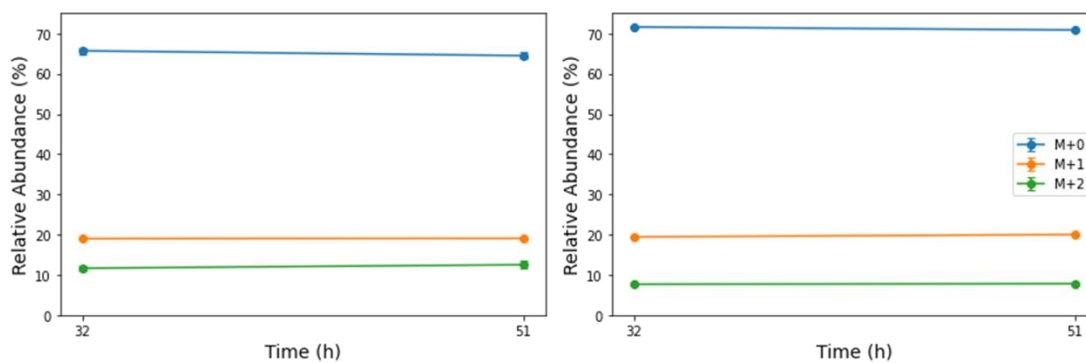
Phe302



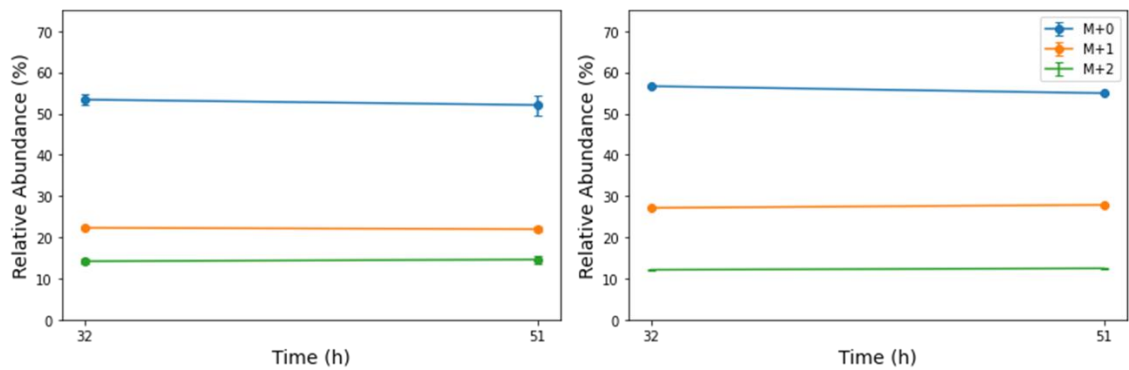
Glu432



Tyr302



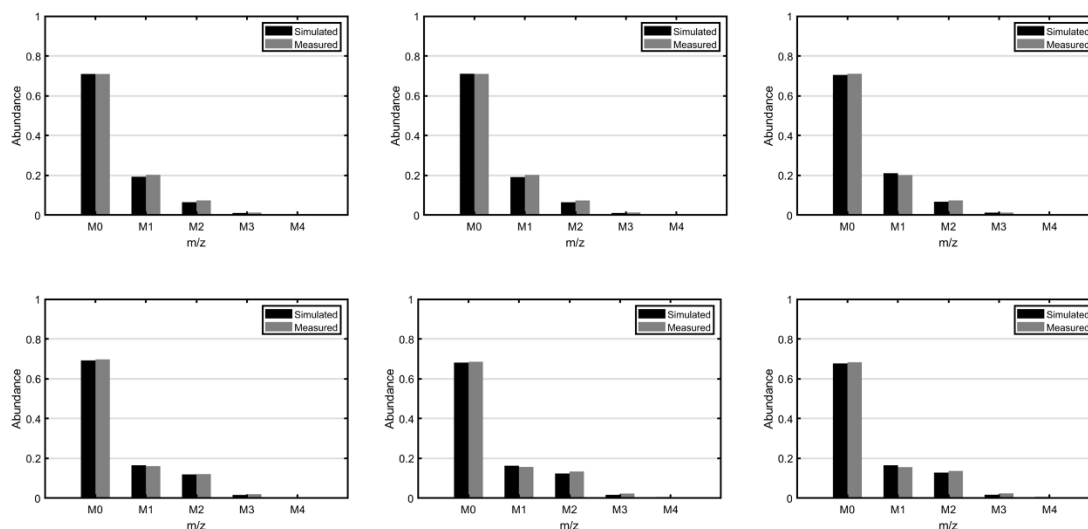
Lys431



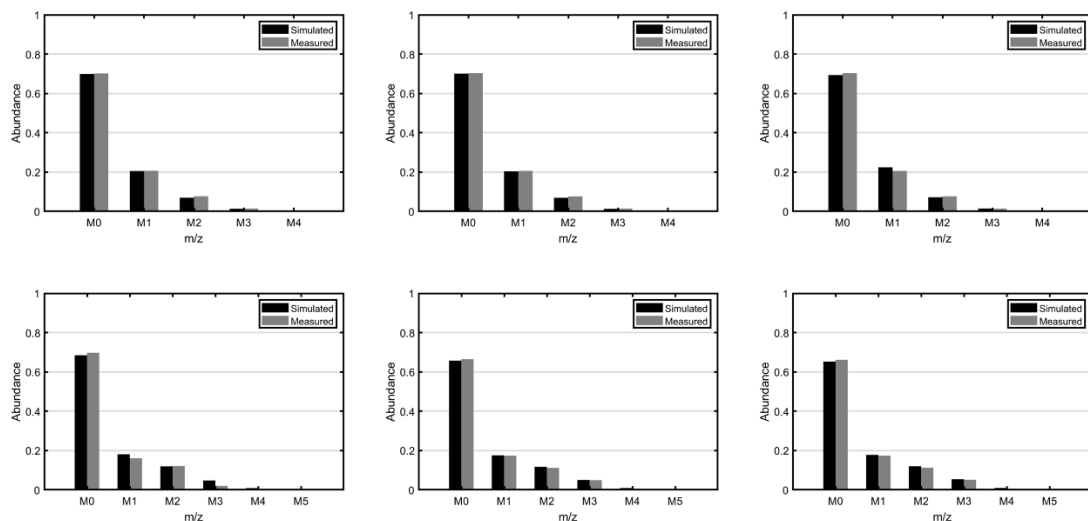
Appendix E

Expected and measured mass isotopomer data for amino acid fragments used in this chapter. All data were from the 32-hour timepoint. The graph layout is as follows: top row is [1-¹³C]glucose, bottom row is [U-¹³C]glucose. Replicates are arranged left to right as 1, 2, 3. Expected measurement was calculated by INCA at steady state with the fraction of glucose measured in the culture supernatant. Protein hydrolysis and TBDMSCIS derivatisation were performed as described in materials and methods section 2.12.2. GC-MS analysis was performed as described in materials and methods sections 2.12.3 and 2.12.4.

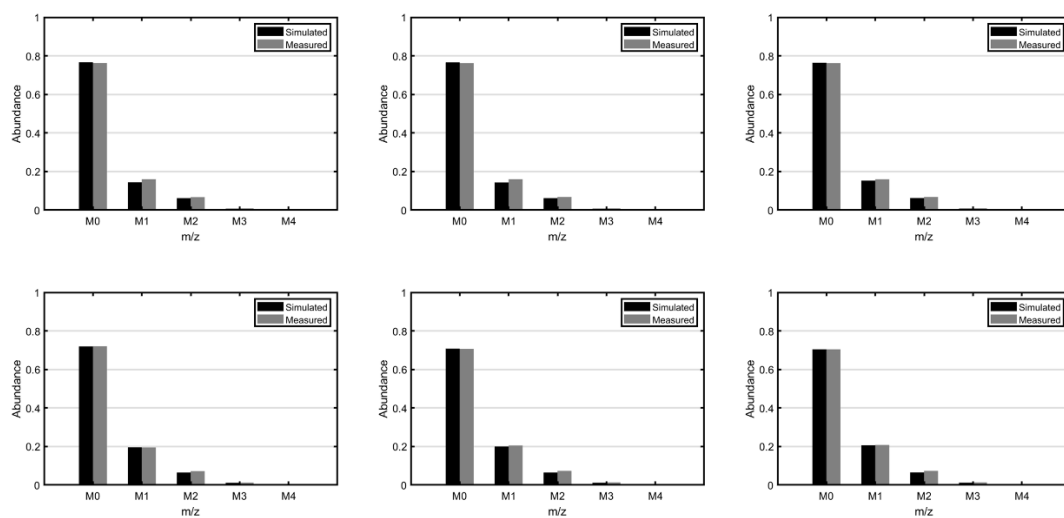
Ala232



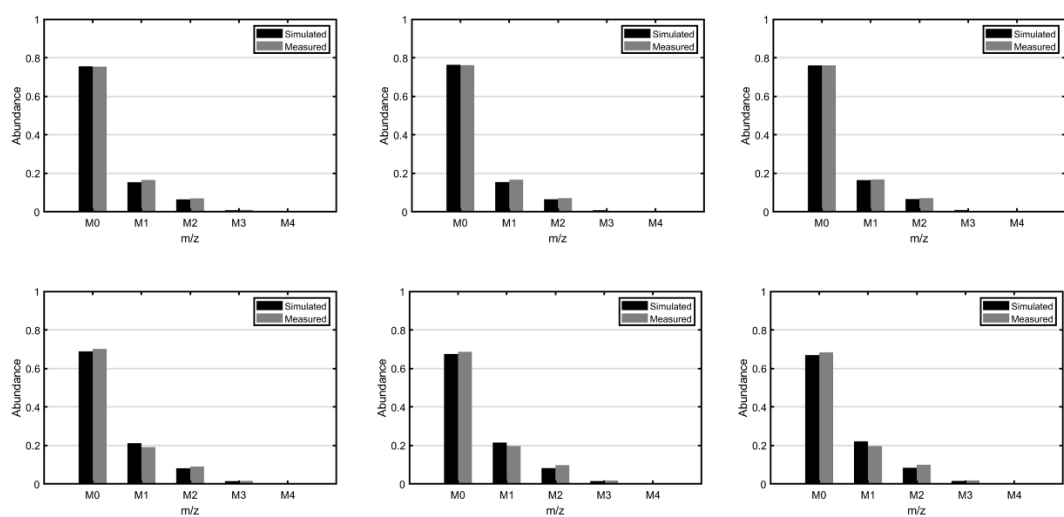
Ala260



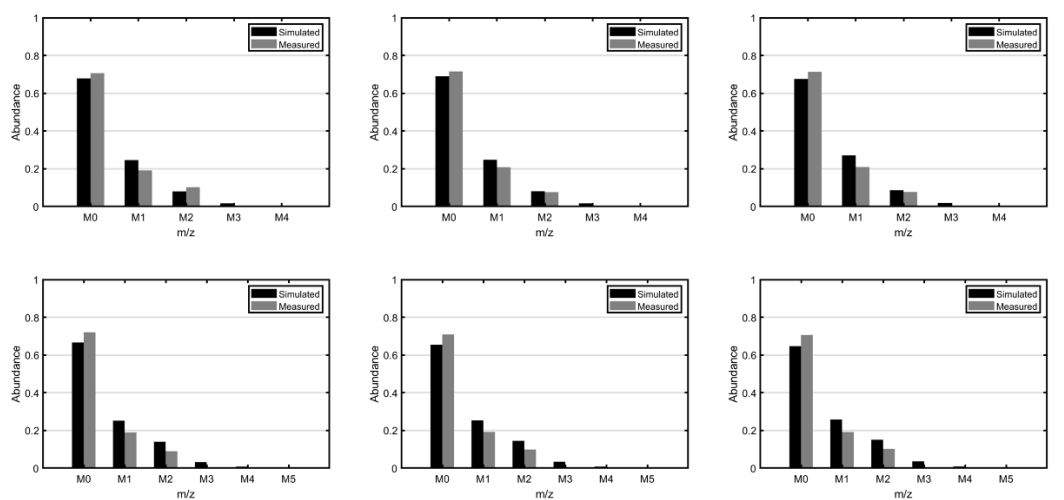
Gly218



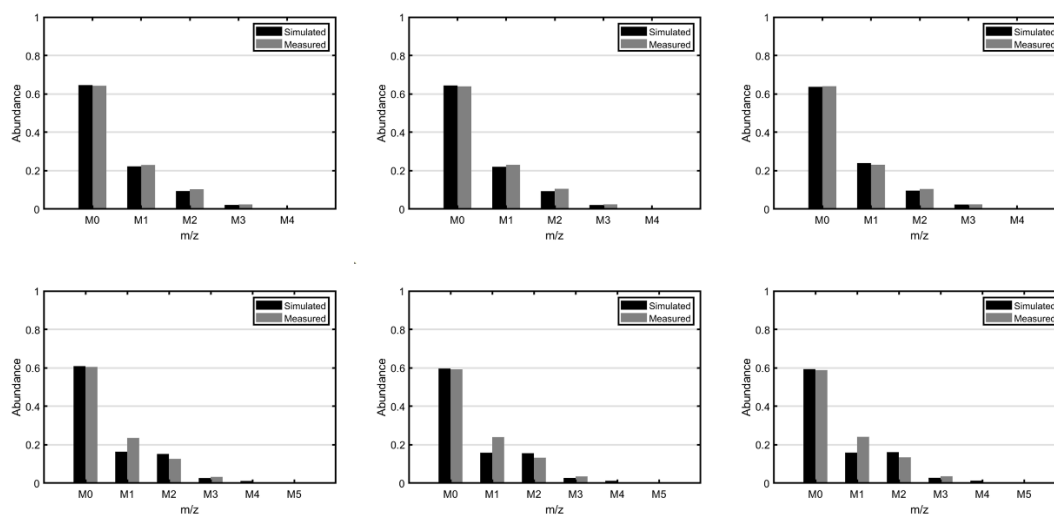
Gly246



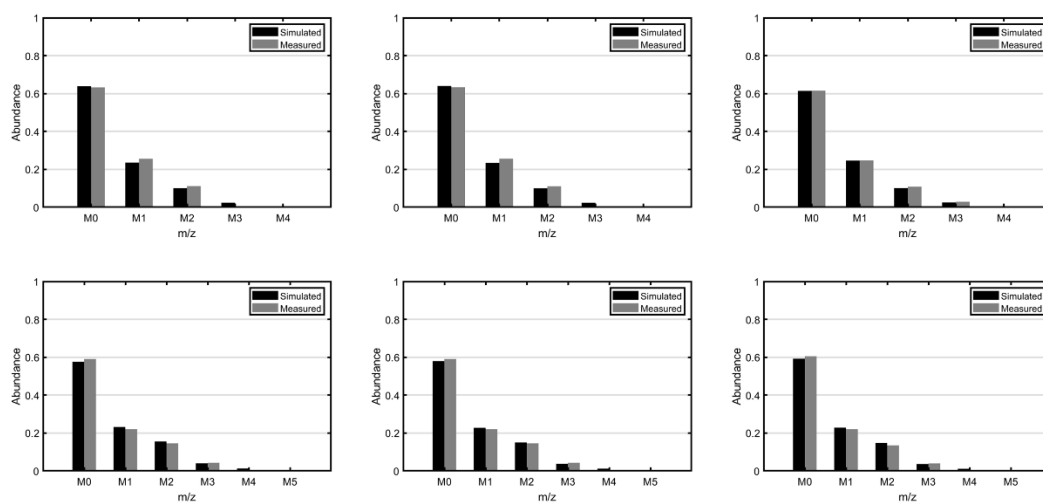
Pro258



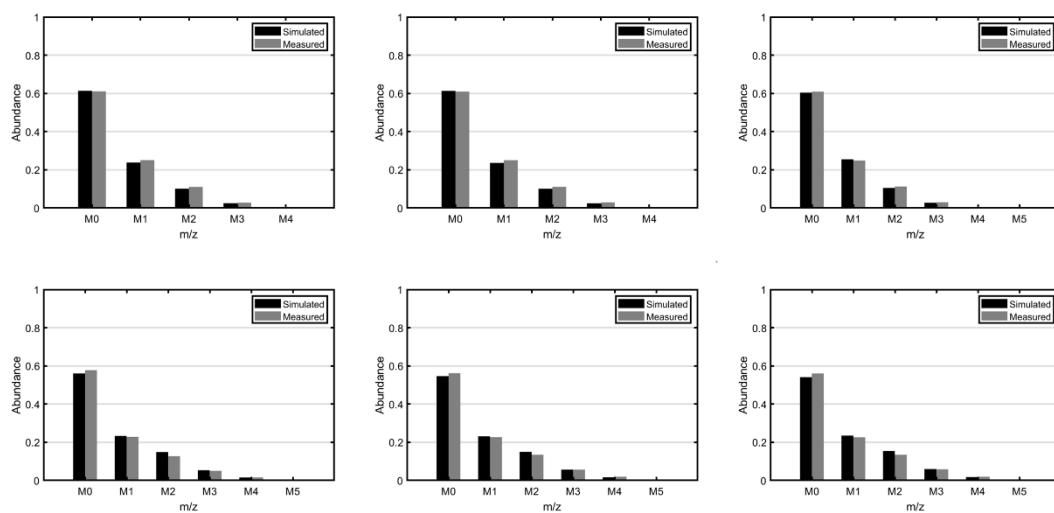
Ser362



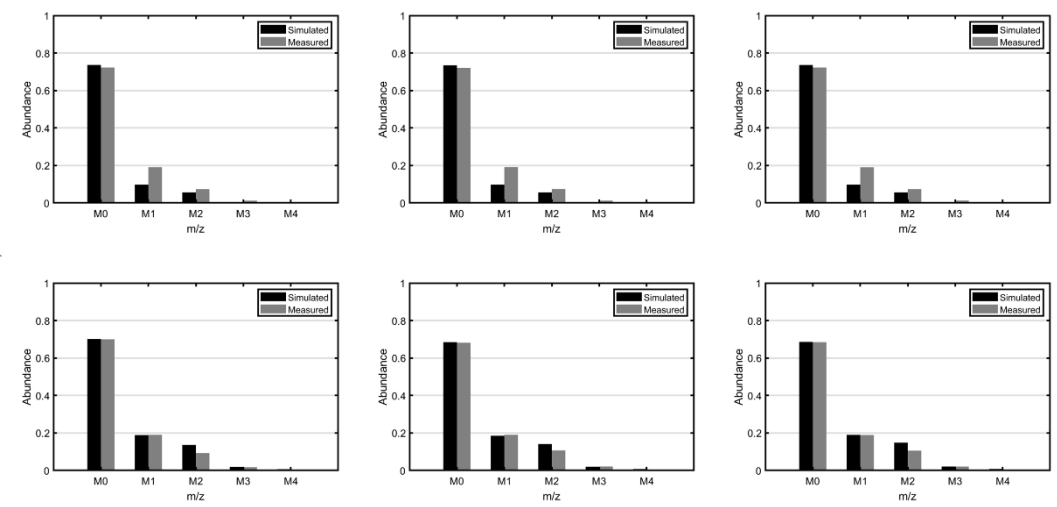
Thr376



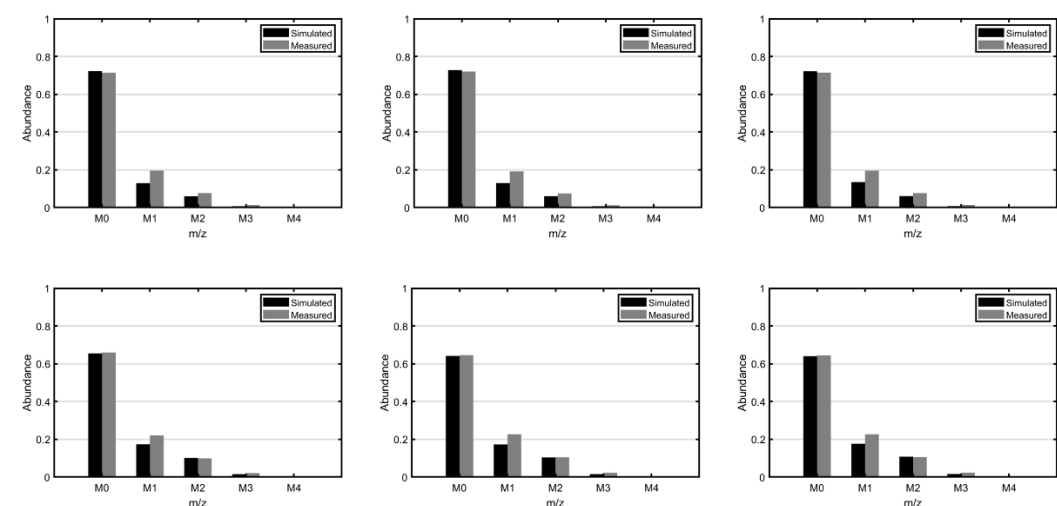
Thr404



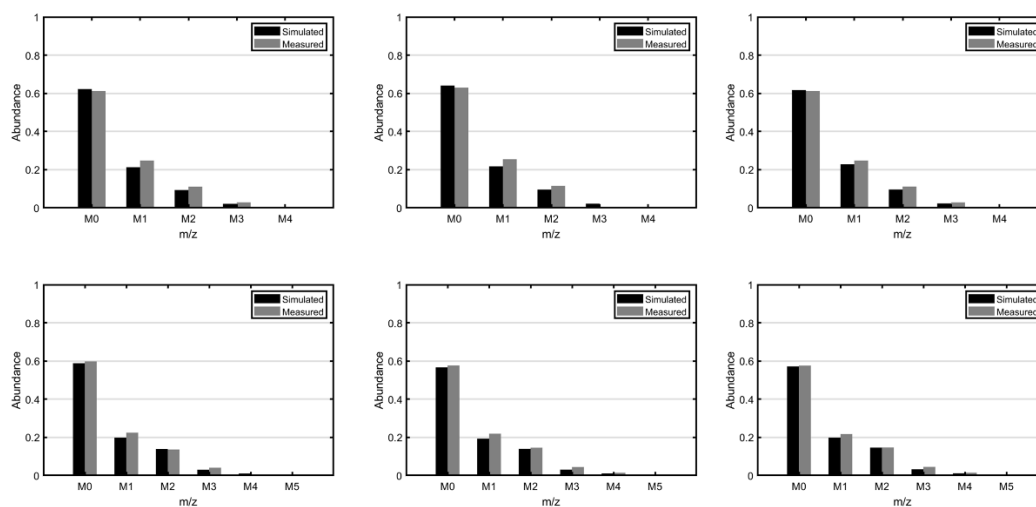
Phe302



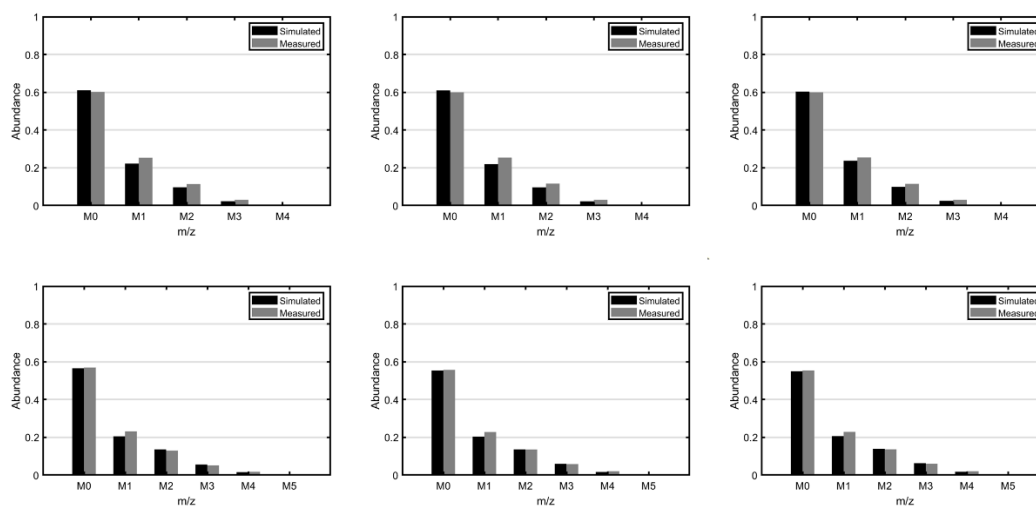
Asp302



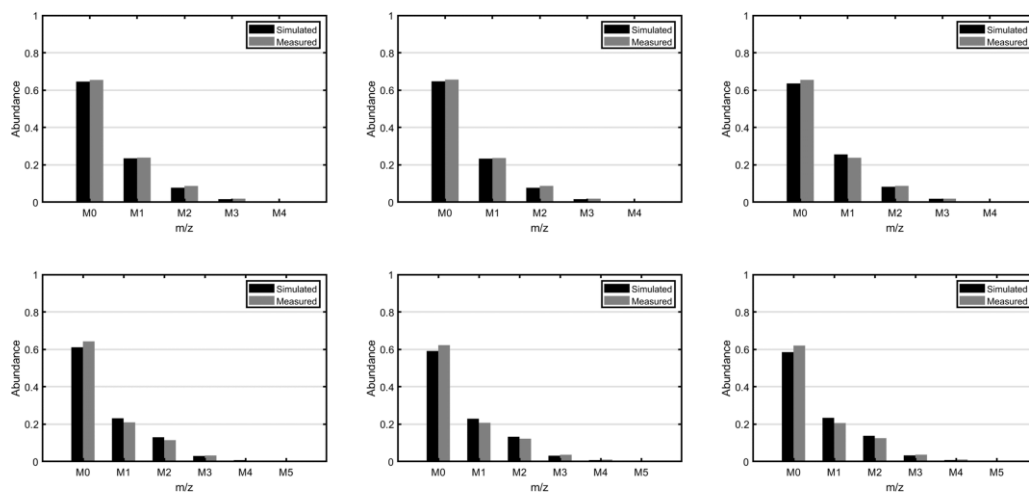
Asp390



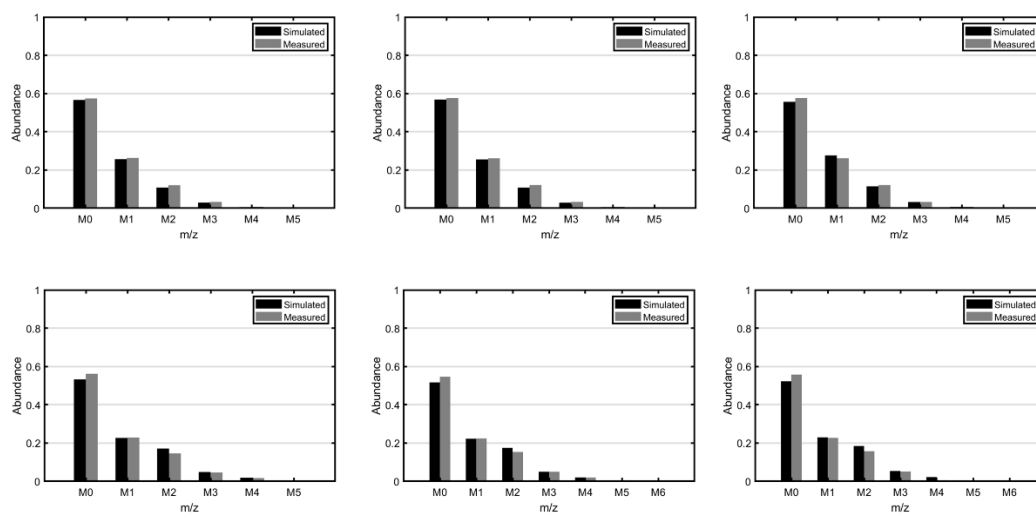
Asp418



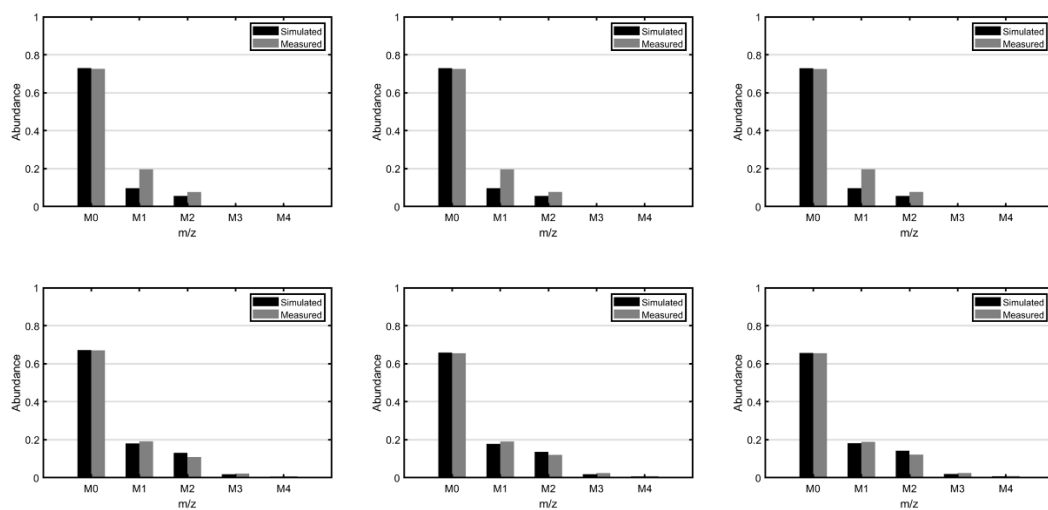
Glu330



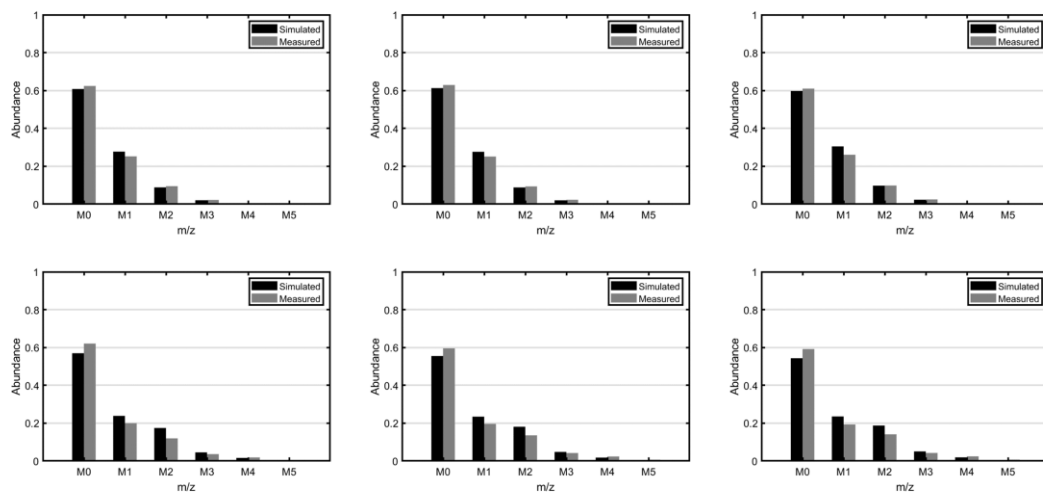
Glu432



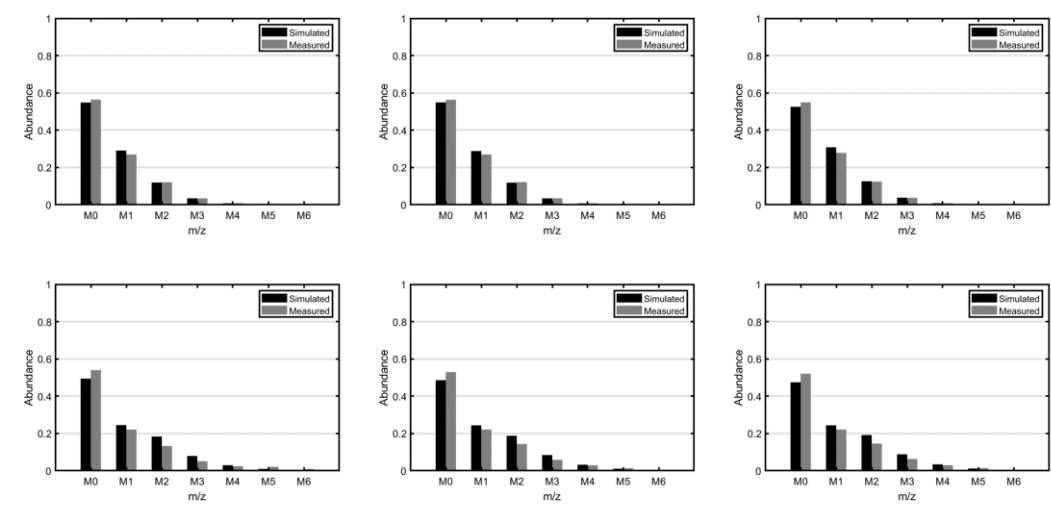
Tyr302



Lys329



Lys431



Appendix F

Table of mass isotopomer data used in ^{13}C -MFA estimation procedure.

	[U- ^{13}C]Glucose	[U- ^{13}C]Glucose	[U- ^{13}C]Glucose	[1- ^{13}C]Glucose	[1- ^{13}C]Glucose	[1- ^{13}C]Glucose
	1	2	3	1	2	3
Ala232 (M0)	69.74923	68.52665	68.25679	70.96197	70.9877	71.03953
Ala232 (M1)	16.08554	15.74055	15.6265	20.27482	20.26167	20.19421
Ala232 (M2)	12.11409	13.40484	13.70434	7.427518	7.398298	7.43361
Ala232 (M3)	2.051139	2.327958	2.412371	1.335699	1.35234	1.332652
Ala232 (M4)	0	0	0	0	0	0
Ala260 (M0)	68.1573	66.50122	66.14098	70.17925	70.34851	70.3505
Ala260 (M1)	17.41692	17.39137	17.40413	20.65057	20.63228	20.56606
Ala260 (M2)	10.37388	11.1787	11.27574	7.710907	7.603022	7.669224
Ala260 (M3)	4.051909	4.928707	5.179147	1.459278	1.416187	1.414221
Gly218 (M0)	72.03053	70.69443	70.4481	76.28994	76.24802	76.26044
Gly218 (M1)	19.50025	20.55267	20.80093	16.03516	16.01558	15.99609
Gly218 (M2)	7.21236	7.37751	7.356297	6.73508	6.776913	6.776527
Gly218 (M3)	1.256868	1.375385	1.394678	0.939823	0.959488	0.966935
Gly246 (M0)	70.12652	68.63259	68.39155	75.38144	76.15357	76.06317
Gly246 (M1)	19.16053	19.72774	19.69498	16.58686	16.72103	16.81523
Gly246 (M2)	9.076705	9.814758	10.01369	6.978029	7.125402	7.121602
Gly246 (M3)	1.636247	1.824908	1.899775	1.053672	0	0
Pro258 (M0)	72.03899	70.90639	70.63627	70.89224	71.53263	71.33028
Pro258 (M1)	18.96263	19.20695	19.15757	21.24775	20.82425	20.92483
Pro258 (M2)	8.998378	9.886657	10.20616	7.860015	7.643122	7.744895
Ser362 (M0)	59.92751	58.64909	58.24979	64.27953	63.94092	63.98991
Ser362 (M1)	23.34317	23.74807	23.87377	22.98642	23.06491	23.10281
Ser362 (M2)	12.56922	13.11151	13.34748	10.33127	10.54508	10.4634
Ser362 (M3)	3.24916	3.474457	3.525345	2.402782	2.44909	2.443883
Thr376 (M0)	60.5159	59.05648	58.35407	63.25575	63.31014	61.55049
Thr376 (M1)	22.03231	22.05251	22.29796	25.56715	25.61612	24.69673
Thr376 (M2)	13.48803	14.57279	14.92685	11.1771	11.07373	10.85222
Thr376 (M3)	3.963764	4.318224	4.421123	0	0	2.900558
Thr404 (M0)	57.73986	56.1879	56.02621	61.0561	60.94349	60.90415
Thr404 (M1)	22.86481	22.67587	22.6173	25.07698	25.01517	24.84422
Thr404 (M2)	12.74856	13.47139	13.47811	11.04501	11.08294	11.22994
Thr404 (M3)	5.047346	5.674818	5.874406	2.82191	2.958394	3.021691
Thr404 (M4)	1.599433	1.990023	2.003973	0	0	0

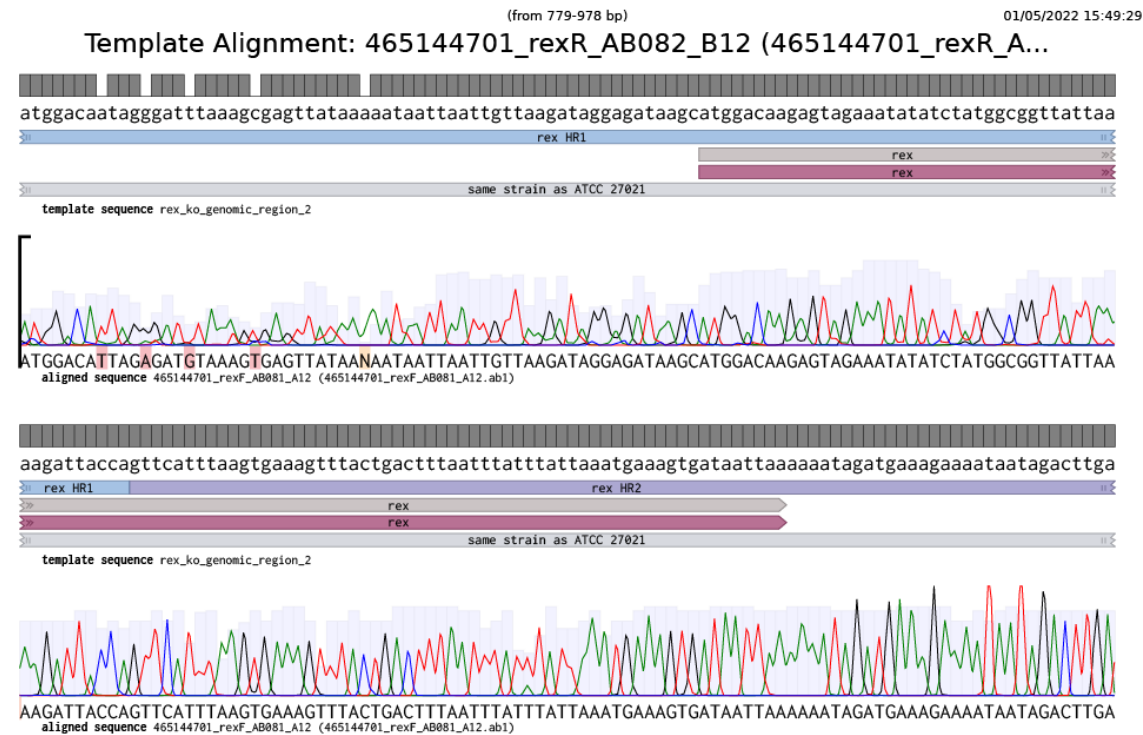
Phe302 (M0)	69.95946	68.14001	68.43752	72.26139	72.04658	72.24929
Phe302 (M1)	19.0107	19.0114	18.90688	19.10754	19.20235	19.05669
Phe302 (M2)	9.281976	10.71206	10.54815	7.357088	7.485016	7.415363
Phe302 (M3)	1.747872	2.136537	2.107455	1.273978	1.266053	1.27865
Asp418 (M0)	56.92624	55.67234	55.43228	60.13747	59.93027	59.93041
Asp418 (M1)	23.15313	22.82201	22.87851	25.32521	25.40476	25.48436
Asp418 (M2)	12.96961	13.59828	13.65841	11.46297	11.60097	11.53146
Asp418 (M3)	5.177143	5.853532	5.99071	3.074351	3.063997	3.05377
Asp418 (M4)	1.773873	2.053838	2.040092	0	0	0
Asp302 (M0)	65.93682	64.53522	64.44568	71.40458	71.50562	71.44569
Asp302 (M1)	22.00578	22.6472	22.64934	19.57246	19.51349	19.54258
Asp302 (M2)	9.932542	10.51021	10.56616	7.684992	7.614512	7.663295
Asp302 (M3)	2.124852	2.307375	2.338814	1.33796	1.366379	1.348441
Asp390 (M0)	59.76638	57.63114	57.58682	61.17969	61.20388	61.18885
Asp390 (M1)	22.44343	21.90131	21.72572	24.7995	24.75749	24.78198
Asp390 (M2)	13.67421	14.56464	14.69209	11.14995	11.16517	11.19133
Asp390 (M3)	4.115981	4.445215	4.515535	2.87086	2.87346	2.837839
Asp390 (M4)	0	1.457692	1.479834	0	0	0
Asp418 (M0)	56.92624	55.67234	55.43228	60.13747	59.93027	59.93041
Asp418 (M1)	23.15313	22.82201	22.87851	25.32521	25.40476	25.48436
Asp418 (M2)	12.96961	13.59828	13.65841	11.46297	11.60097	11.53146
Asp418 (M3)	5.177143	5.853532	5.99071	3.074351	3.063997	3.05377
Asp418 (M4)	1.773873	2.053838	2.040092	0	0	0
Glu330 (M0)	64.26366	62.26011	62.01695	65.46188	65.63558	65.49079
Glu330 (M1)	20.99127	20.75189	20.62658	23.86837	23.69111	23.80025
Glu330 (M2)	11.47168	12.26646	12.5488	8.739132	8.748379	8.761904
Glu330 (M3)	3.273388	3.673344	3.692953	1.930621	1.924936	1.947055
Glu330 (M4)	0	1.048194	1.11472	0	0	0
Glu432 (M0)	56.17815	54.62651	54.64369	57.43638	57.61798	57.61115
Glu432 (M1)	22.8624	22.42394	22.26196	26.33781	26.11924	26.152
Glu432 (M2)	14.59263	15.35273	15.43148	12.06836	12.1256	12.12832
Glu432 (M3)	4.633648	5.019494	5.02752	3.335157	3.317129	3.310649
Glu432 (M4)	1.733182	2.011516	2.024518	0.822306	0.820051	0.797881
Glu432 (M5)	0	0.56581	0.610833	0	0	0
Tyr302 (M0)	66.53682	65.03037	65.50361	71.64849	71.62757	71.58335
Tyr302 (M1)	19.05794	18.99709	18.91144	19.41199	19.42015	19.42342
Tyr302 (M2)	10.78843	11.96382	12.15808	7.60597	7.639118	7.655333
Tyr302 (M3)	2.157068	2.470723	2.512437	1.333552	1.313156	1.337905
Tyr302 (M4)	0.78395	0.882207	0.914435	0	0	0

Lys329 (M0)	62.06985	59.55927	59.16768	62.41215	62.90776	60.98637
Lys329 (M1)	19.97323	19.6257	19.34401	25.23161	25.12691	26.07735
Lys329 (M2)	11.95901	13.54889	14.10684	9.51995	9.366452	9.863293
Lys329 (M3)	3.654643	4.195913	4.242194	2.238313	2.295468	2.499093
Lys329 (M4)	1.856821	2.394188	2.433134	0.516269	0.225518	0.519922
Lys329 (M5)	0.486443	0.676037	0.706151	0.081711	0.077889	0.053967
Lys431 (M0)	54.93044	52.93583	52.30418	56.42495	56.4174	57.09109
Lys431 (M1)	22.49384	22.11888	22.24144	26.97516	26.99274	27.42537
Lys431 (M2)	13.53539	14.34051	14.64019	12.08938	12.17032	12.05366
Lys431 (M3)	5.219781	5.893594	6.367372	3.479186	3.472715	3.429876
Lys431 (M4)	2.455153	2.918133	2.988456	0.844695	0.834838	0
Lys431 (M5)	1.014811	1.342381	0.596518	0.165301	0.086626	0
Lys431 (M6)	0.350583	0.450667	0.861846	0.021334	0.025354	0

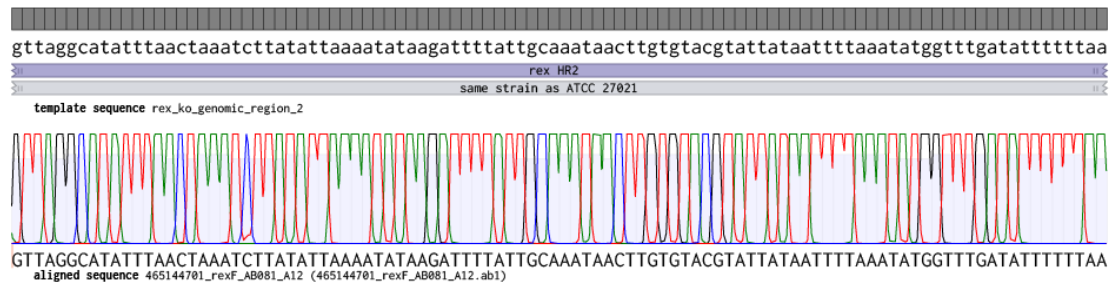
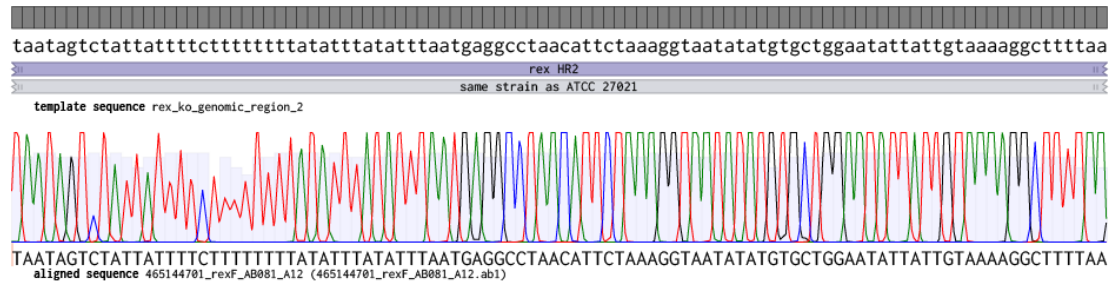
Appendix G

In all the chromatograms presented in these Appendices and Supplementary file 1 any yellow bases are bases which could not be identified from the chromatogram. Red bases are those which are mutated from the template sequence.

Sequence alignment of the *rex* genomic region in strain *C. saccharoperbutylacetonicum* N1-4 *spo0A*^{I261T} Δ *rex*. Template sequence for alignment was the truncated *rex* gene generated through overlapping PCR. Sequence was correct for homology arm 1 and truncated *rex* gene. Near the end of the read there were a few mismatches, but this was likely due to the drop in signal which commonly occurs near the end of sequencing reads.



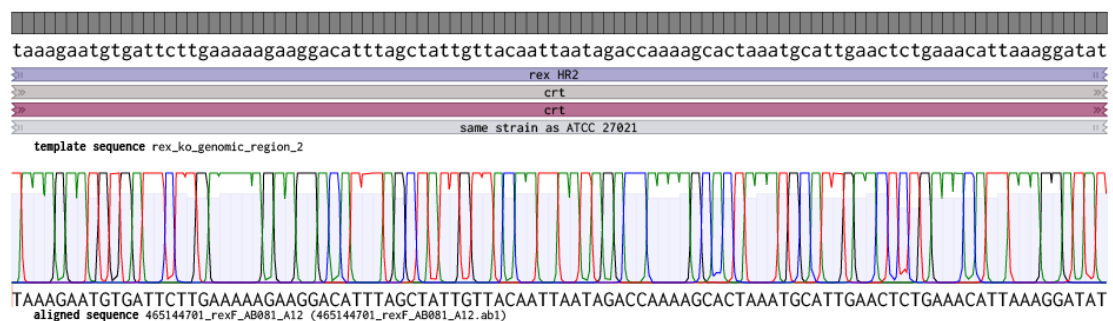
Template Alignment: 465144701_rexR_AB082_B12 (465144701_rexR_AB082_B12.ab1), 465144701_rexF_AB081_A12 (465144701... 01/05/2022 15:49:29



https://benchling.com/alexander-bones/f/lib_w11nsTka-rex-knockout/seq_Y5cUfmyT-rex_ko_genomic_region_2/edit

2/5

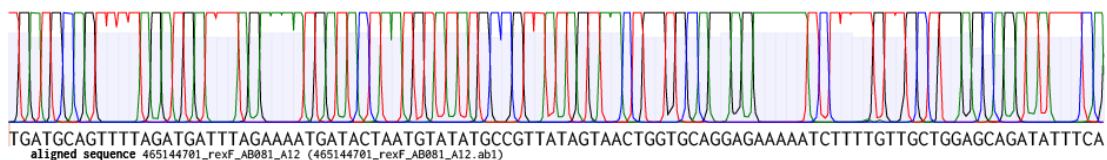
Template Alignment: 465144701_rexR_AB082_B12 (465144701_rexR_AB082_B12.ab1), 465144701_rexF_AB081_A12 (465144701... 01/05/2022 15:49:29



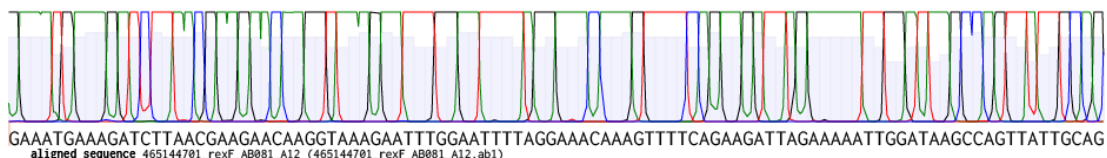
https://benchling.com/alexander-bones/f/lib_w11nsTka-rex-knockout/seq_Y5cUfmyT-rex_ko_genomic_region_2/edit

3/5

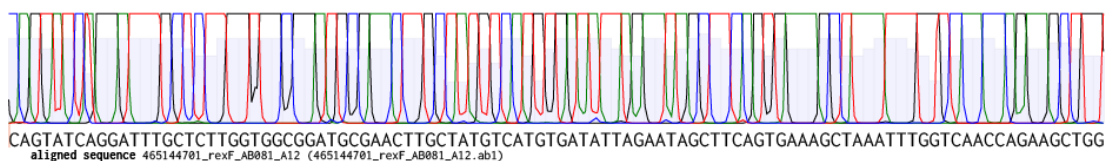
tgatgcagtttttagatgatttagaaaatgataactaatgtatatgccggtatagtaactgggtgcaggagaaaaatcttttggctgctggagcagatatattca																																																																																																			
rex HR2																																																																																																			
crt																																																																																																			
crt																																																																																																			
same strain as ATCC 27021																																																																																																			
template sequence rex ko genomic region 2																																																																																																			



gaaatgaagatcttaacgaagaacaaggtaaagaatttggaatttaggaaacaaagttttcagaagattagaaaaattggataagccagttattgcag
 rex HR2
 crt
 crt
 same strain as ATCC 27021
 template sequence rex ko genomic region 2



4/5



tcttggaaataccaggatttgggtggaactcaaagattagctagaattgtagggcctggaaaggctaagaattaatatatacttgtga

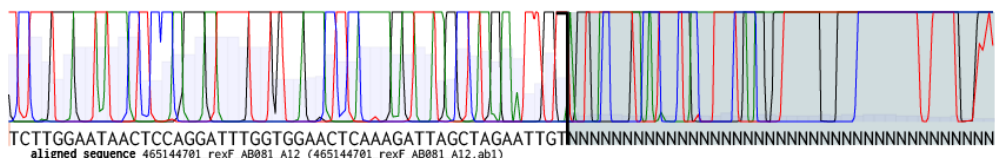
rex HR2

crt

crt

same strain as ATCC 27021

template sequence rex_ko_genomic_region_2



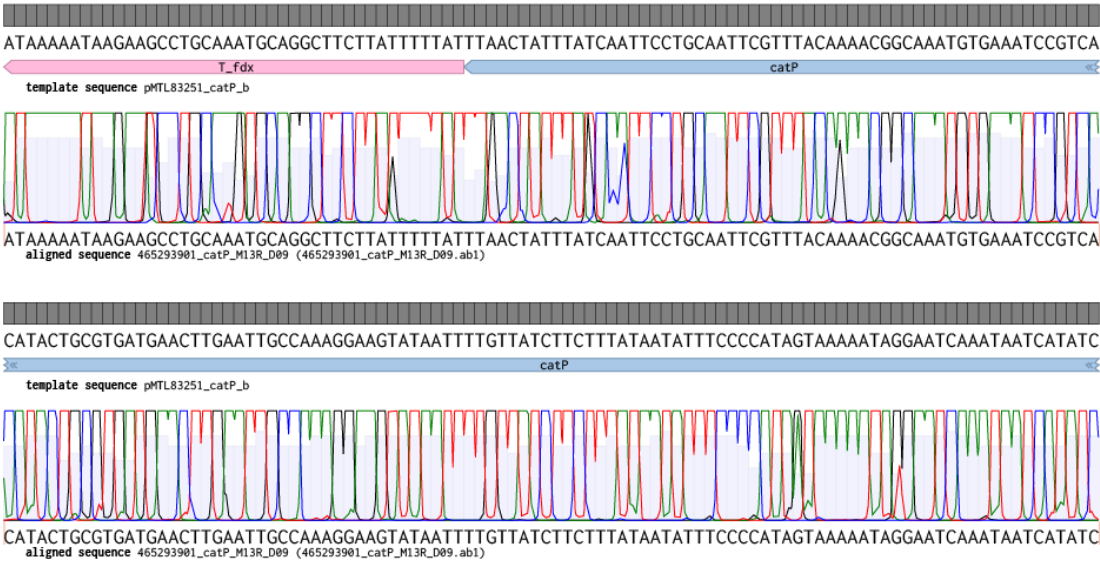
5/5

Appendix H

Sequence alignment of plasmid pMTL83251_catP to template. No mutations were identified in the sequenced plasmid.

(from 212-411 bp)02/05/2022 14:18:51

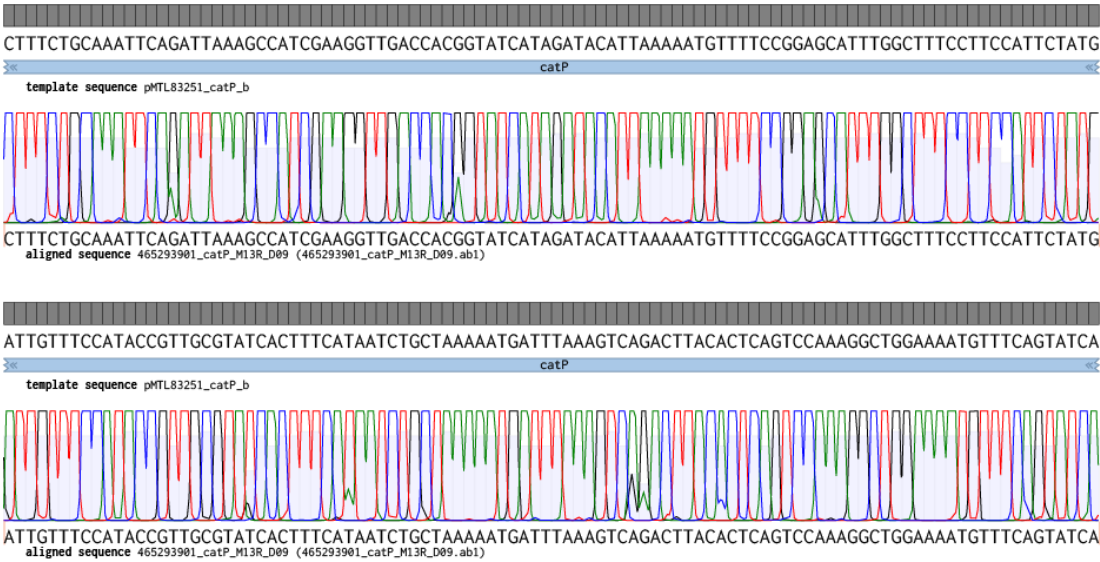
Template Alignment: 465293901_catP_M13R_D09 (465293901_catP_M1...



https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_Bi2ialMk-pmtl83251_catp_b/edit

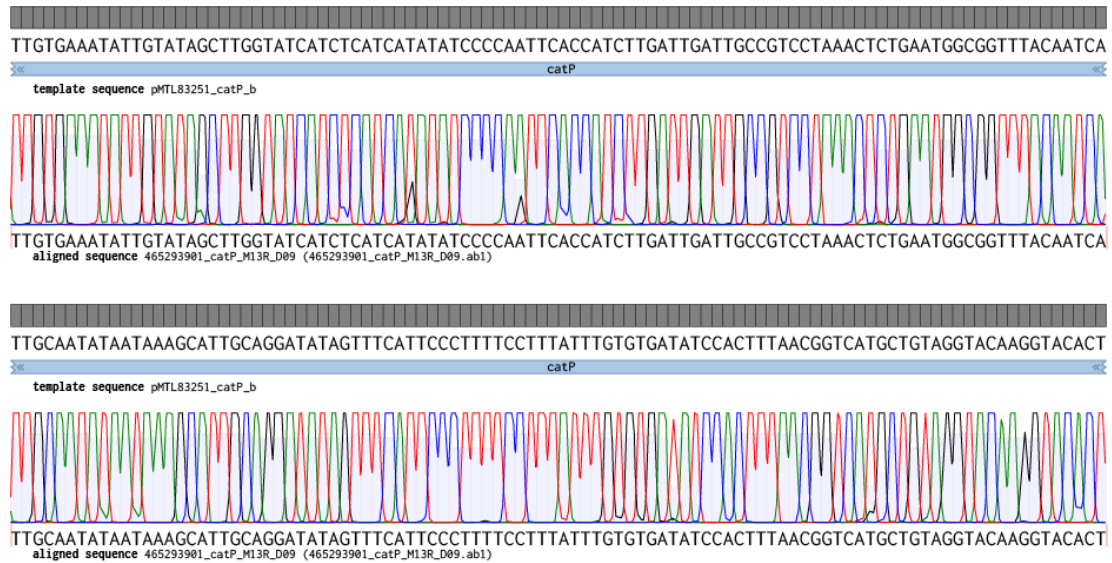
1/5

Template Alignment: 465293901_catP_M13R_D09 (465293901_catP_M13R_D09.ab1) (from 412-611 bp)02/05/2022 14:18:51



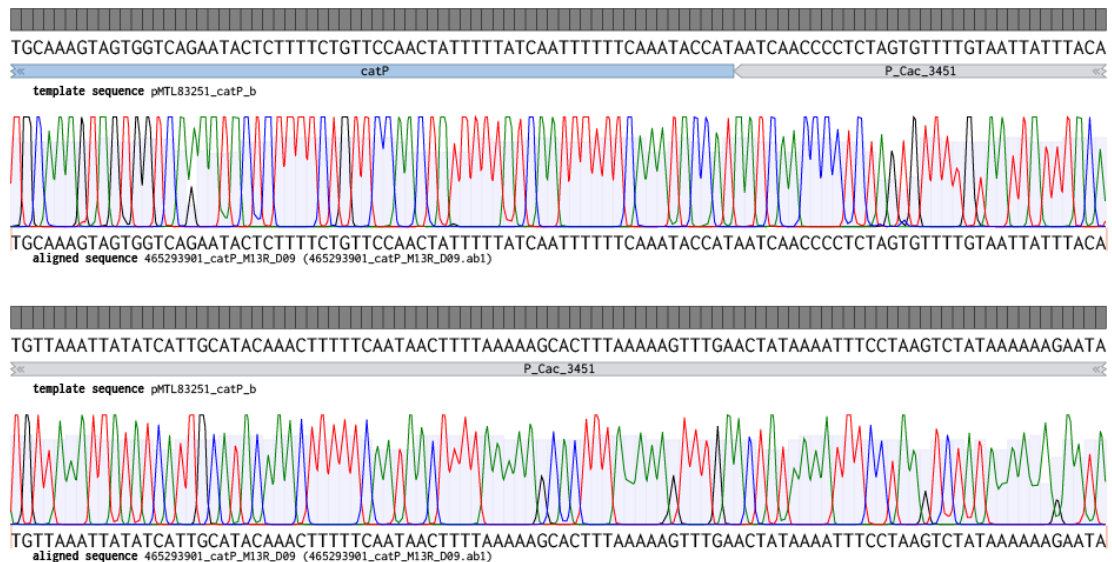
https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_Bi2ialMk-pmtl83251_catp_b/edit

2/5



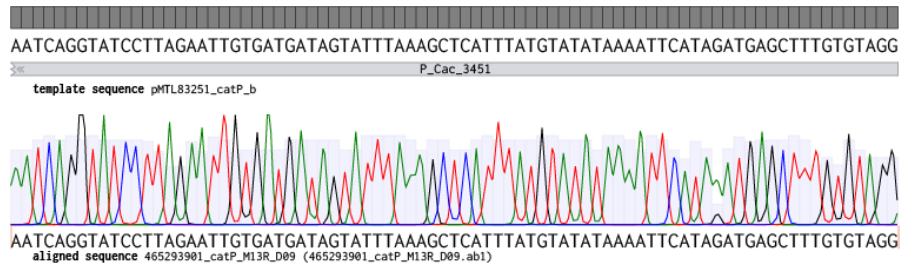
https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_Bi2ialMk-pmtl83251_catp_b/edit

3/5



https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_Bi2ialMk-pmtl83251_catp_b/edit

4/5



https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_Bi2ialMk-pmtl83251_catp_b/edit

5/5

Appendix I

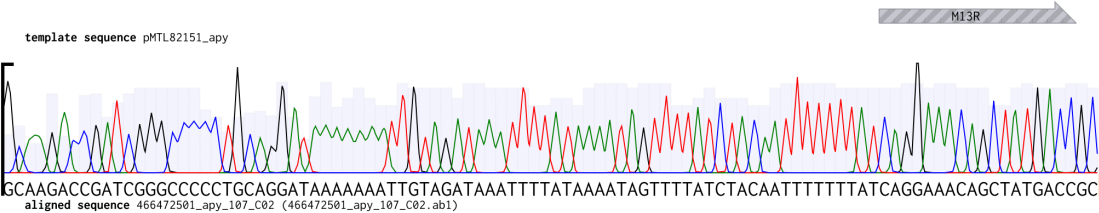
Sequence alignment of pMTL82151_apv with template. No mutations were identified in the plasmid.

(from 2470-2669 bp)

30/04/2022 15:59:09

Template Alignment: 466472501_apy_107_C02 (466472501_apy_107_C...

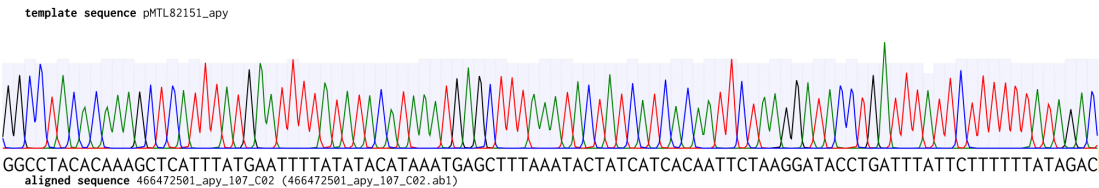
gcaagaccgatcgggcccccctgcaggataaaaaattgtagataaattttataaaatagttttatctacaattttttatcaggaaacagctatgaccGC



GGCCTACACAAAGCTCATTTATGAATTTTATATACATAAATGAGCTTTAAATACTATCATCACAAATTCTAAGGATACCTGATTTATTCTTTTTATAGAC

Cac3451 Promoter

P_Cac3451_apy_T_fdx



Template Alignment: 466472501_apy_107_C02 (466472501_apy_107_C02.ab1) (from 2670-2869 bp)

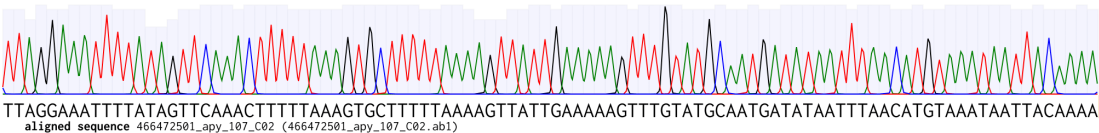
30/04/2022 15:59:09

TTAGGAAATTTTATAGTTCAAACCTTTTAAAGTGCTTTTAAAAGTTATTGAAAAAGTTTGTATGCAATGATATAATTTAACATGTAATAATTACAAAA

Cac3451 Promoter

P_Cac3451_apy_T_fdx

template sequence pMTL82151_apy

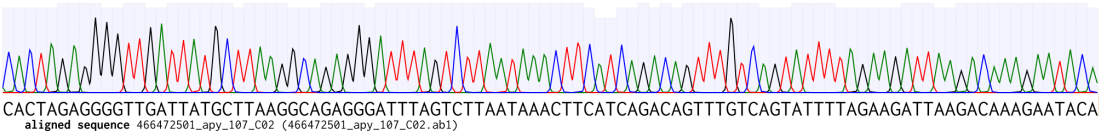


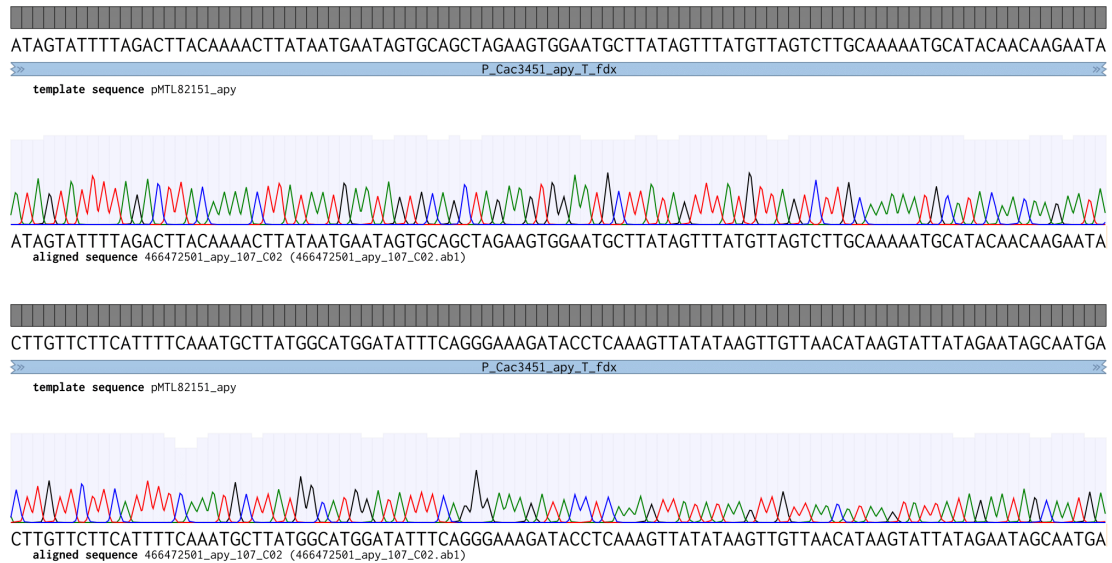
CACTAGAGGGTTGATTATGCTTAAGGCAGAGGGATTTAGTCTTAATAAACTTCATCAGACAGTTTGTCAAGTATTTAGAAGATTAAGACAAAGAATACA

Cac3451 Promoter

P_Cac3451_apy_T_fdx

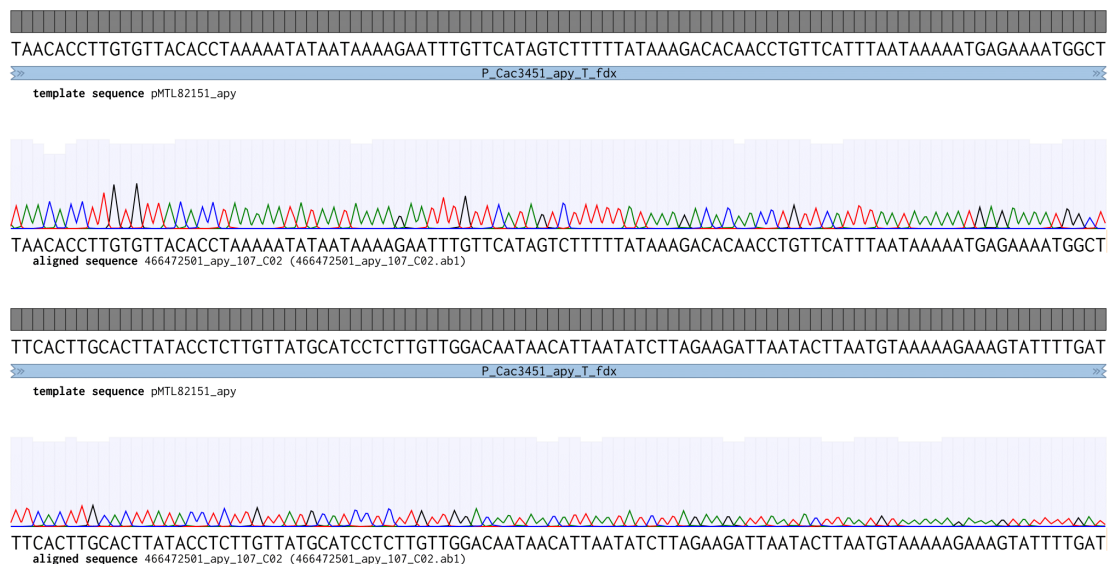
template sequence pMTL82151_apy





https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_TNP44tR5-pmtl82151_apy/edit

3/6

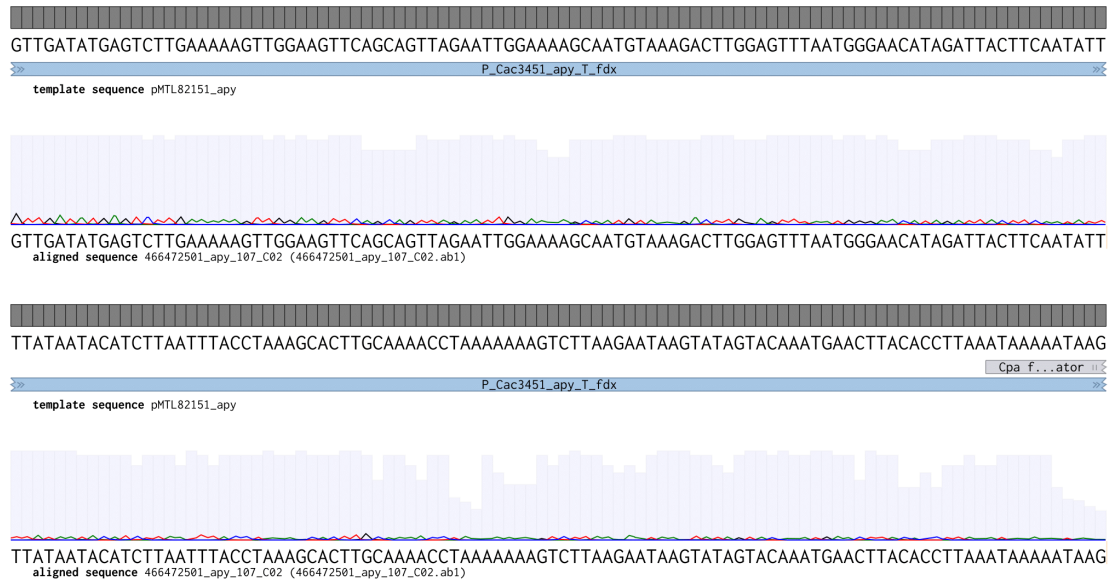


https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_TNP44tR5-pmtl82151_apy/edit

4/6

Template Alignment: 466472501_apy_107_C02 (466472501_apy_107_C02.ab1) (from 3270-3469 bp)

30/04/2022 15:59:09

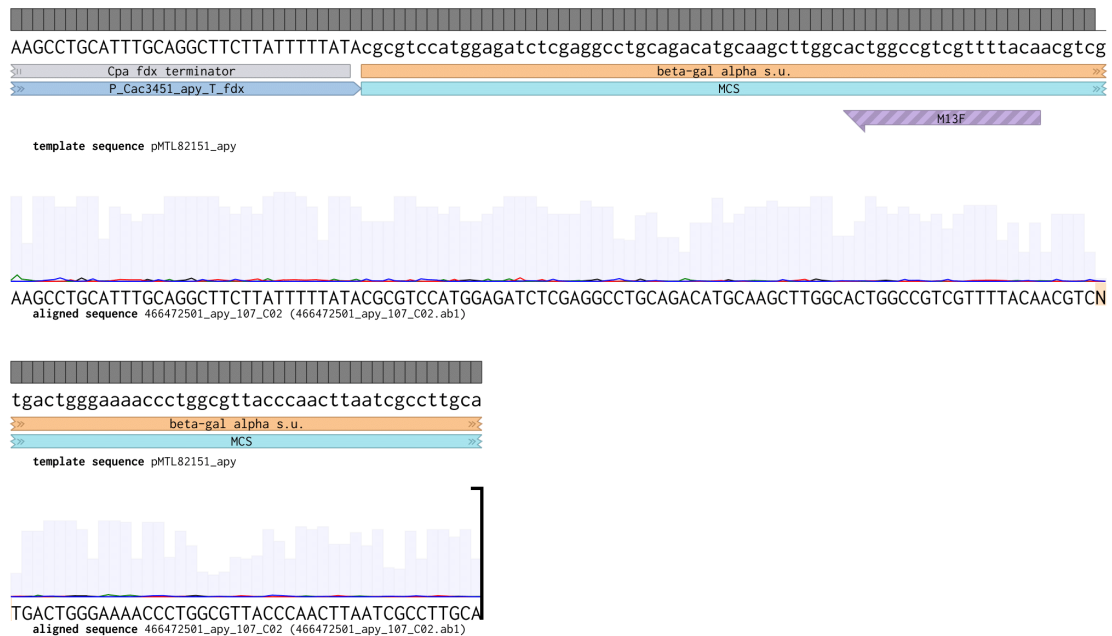


https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_TNP44tR5-pmtl82151_apy/edit

5/6

Template Alignment: 466472501_apy_107_C02 (466472501_apy_107_C02.ab1) (from 3470-3612 bp)

30/04/2022 15:59:09



https://benchling.com/alexander-bones/f/lib_eRt8CLfn-restriction-cloning/seq_TNP44tR5-pmtl82151_apy/edit

6/6

