Improved Clostridia to produce building-block chemicals

María Monserrat Román Lara

Doctor of Philosophy

Aston University

September 2022

©María Monserrat Román Lara 2022

María Monserrat Román Lara asserts her moral right to be identified as the author of this thesis.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright belongs to its author and that no quotation from the thesis and no information derived from it may be published without appropriate permission or acknowledgement.

Thesis summary

Improved Clostridia to produce building-block chemicals

M. Monserrat Román Lara

Doctor of Philosophy

2022

Microorganisms producing biofuels and building-block chemicals (i.e. organic acids) have the potential to be the basis of economically competitive bioprocesses. *C. saccharoperbutylacetonicum* N1-4HMT is one of the most valuable industrial organisms because of its butanol production, which produces organic acids during the first stage of fermentation, the acidogenic phase. One of the main challenges in producing organic acids is the low tolerance to the acidic environment generated by the organic acids. The project aimed to generate Clostridial strains that are more resistant to organic acid and hence generate improved bioprocesses. Several organisms can regulate their lipid membrane fluidity (saturation/unsaturation ratio) to increase their tolerance to their environment. From the different techniques available to detect unsaturation in the lipids, the iodine value was adapted to be used in biological membranes obtaining results that correlate with reported values. Another problem with organic acid production is the accumulation of protons inside the cell due to their dissociation. The cation/proton antiporters NhaA and Cpa1 were chosen candidates because they can alleviate H^+ and Na^+ accumulation. The glycerol facilitator GlpF was also chosen because its overexpression has been observed during the acidogenic phase of *C. acetobutylicum. C. saccharoperbutylacetonicum* N1-4HMT was genetically modified using the patented CLEAVE^{TM} technology to create +NhaA, +Cpa1 and ΔGlpF. Under conditions developed at Bio-Cleave Ltd, none of the mutated strains significantly improved organic acid production. 1 % (w/v) NaCl prevented the shift to the solventogenic phase in all the strains. A mutant strain ΔPssA that has the *pssA* gene knocked out, under Bio-Cleave Ltd conditions, did not produce solvents but only organic acids and was the strain that produced the highest butyric acid. The work in this thesis demonstrated that *C. saccharoperbutylacetonicum* N1-4HMT could be adapted to produce organic acids as end-products.

Key words: Organic acids, Clostridia, lipid unsaturation, iodine value, CLEAVE^{TM} technology, NhaA, Cpa1, GlpF, PssA, acidogenesis

Acknowledgements

I want to express my deepest gratitude to my supervisor Dr Alan D. Goddard; thank you so much for believing in me and creating this PhD project. For all your support in getting the CONACyT scholarship and all the other funds. For your endless support and guidance during these four years. Thank you, Professor Roslyn M. Bill, for getting on board, being my Aquamum and for all your support and guidance during my PhD. And thank you, Dr Katie J. Chong, for joining the team, giving me a different perspective on my project, and all your support and motivation.

Thank you, CONACyT and COECyTJal, for the scholarship that supported my PhD at Aston University; I couldn’t have done it without it. Thank you, COECyTJal, for supporting me in my short research stay at ITESO in Guadalajara, Jalisco. And thank you, ITESO, the Biotechnology department, and especially Dr Óscar A. Rojas-Rejón, for your help and support in the small project using the fermenters.

Thank you to the ones that were Green Biologics and to the ones that are now Bio-Cleave, for your friendship, helping me get back in England and introducing me to Alan! Thank you, Mandy, Anna and Liz, for your knowledge, advice and support during the strain engineering and characterisation.

Thank you to everybody who has been part of Goddard group and AMPL, who have been in 331, 329, 347, and the blue office. Special thanks to Sarah, John, and Carolina for the technical help, feedback, and all the journeys that science took us together.

I am grateful for my time as an APRS and ABG member and all the people and friends I met during my PhD; you made this time unforgettable. Especially during the Covid-19 lockdown, thank you for becoming my family. Salma and Juanma! Thank you for being essential persons in my life.

Thank you to the Aston Student Union, the Lounge and the Good place, especially Nat and Malc, for giving me the job and making me feel part of Aston.

Thank you for your love and support to les Roman, les Lara, and all my friends outside Birmingham. To my mum, Marilupe, my dad, Heriberto, and my sister, Fernanda, you have been my motor, inspiration and strength. You are the reason why I am here and feel capable of conquering my dreams. To Pris, for becoming part of the family. Thank you, Lucas, for all the tips and tricks using the BMG and Microsoft office and for all your love and support.

Table of contents

Contents

Thesis summary 2

Acknowledgements 3

Table of contents 4

List of abbreviations 8

List of Figures 11

List of Tables 15

1. Introduction 16

  1.1 Biotechnology 16

  1.2 White biotechnology 17

  1.3 Biorefinery 18

  1.4 Microbial fermentation 19

  1.5 Clostridia 20

    1.5.1 The acetone-butanol-ethanol fermentation 21

  1.6 Organic acids 24

  1.7 High-value chemicals 25

  1.8 Lactic acid 25

  1.9 Butyric acid 28

  1.10 Limitations in the fermentation of organic acids 30

  1.11 Lipid membrane 31

  1.12 Proton accumulation 33

  1.13 Genetic engineering in Clostridial spp. 34

    1.13.1 Type II intron 34

    1.13.2 Allele-Couple Exchange (ACE) 36

    1.13.3 CRISPR/Cas 38

    1.13.4 CLEAVE^{TM} technology by Bio-Cleave Ltd 38

  1.14 Rationale 39

  1.15 Aims of the project 40

2. Materials and Methods 42

  2.1 Results chapter I -A miniaturised iodine value assay for quantifying unsaturated fatty acids .. 42

    2.1.1 Reagents 42

    2.1.2 Lipid preparation 42

    2.1.3 *E. coli* and yeast membranes 42

    2.1.4 Iodine value assay 43

2.2 Results chapter II – Characterisation of membrane proteins for their role in organic acid tolerance 44
2.2.1 Protein modelling 44
2.2.2 Bacterial growth conditions and media 44
2.2.3 Plasmids, plasmid isolation. 44
2.2.4 PCR amplification 46
2.2.5 DNA electrophoresis and DNA purification 47
2.2.6 Double digestion 48
2.2.7 Ligation of inserts 48
2.2.8 *E. coli* transformation 48
2.2.9 Protein expression 49
2.2.10 French press cell membrane extraction 49
2.2.11 Protein purification 49
2.2.11.1 Styrene-maleic acid (SMA) 49
2.2.11.2 n-Dodecyl-β-Maltoside DDM 50
2.2.11.3. Octyl-β-Glucoside OG 50
2.2.12 SDS-PAGE 51
2.2.13 Western blot 51
2.3 Results chapter III – Biocleave *C. saccharoperbutylacetonicum* N1-4HMT characterisation 53
2.3.1 *C. saccharoperbutylacetonicum* N1-4HMT transformation 53
2.3.2 Bottle screening 53
2.3.3 HPLC quantification of sugars and solvents 53
2.3.4 Statistical Analysis 54
3. Results I – A miniaturised iodine value assay for quantifying unsaturated fatty acids 55
3.1 Introduction 55
3.2 Results 56
3.2.1 Comparisons of methods for the quantification and identification of unsaturation of lipids. 56
3.2.2 Iodine value (IV) 61
3.2.3 Optimisation of the IV method. Scale-down 64
3.2.4 Optimisation of the IV. Detection of unsaturation in lipid standards. 68
3.2.5 Miniaturised IV assay analysis of *E. coli* cell membranes 78
3.2.6 Miniaturised IV assay in total yeast lipids 80
4. Results II – Characterisation of membrane proteins for their role in organic acid tolerance 82
4.1 Introduction 82
4.2 Results 83
4.2.1 Selection of proton antiporters 83

4.2.2 Nucleotide sequencing and protein modelling 86
4.2.2.1. Proton antiporters 86
4.2.2.2 GlpF........ 93
4.2.3 Expression and characterisation of NhaA and Cpa1 in *E. coli* 98
4.2.3.1 Gene amplification and PTZ19R/U ligation 101
4.2.3.2 Ligation into expression vector pET17b 103
4.2.3.3 Re-defining strategy. Amplification of genes and ligation into pET15b and pBR322 . 104
4.2.4 Expression and characterisation of pET17b-CsGlpF in *E. coli* 107
4.2.5 Expression of pET15b-Ec-GlpF-myc 112
4.2.6 Construction of pET15b-Cs-GlpF-Histag 113
4.2.7 pET15-Cs-GlpF-Histag expression 115
5. Results III – Genetic modifications of *C. saccharoperbutylacetonicum* N1-4HMT and its characterisation for organic acid production 119
5.1 Introduction 119
5.2 Results 120
5.2.1 Generation of the genetically modified *C. saccharoperbutylacetonicum* N1-4HMT strains. 120
5.2.1.1 Overexpression of cation/proton antiporters in C. saccharoperbutylacetonicum N1-4HMT 120
5.2.1.2 CLEAVE^{TM} technology 122
5.2.2 Characterisation of the +NhaA, +Cpa1, ΔGlpF and WT strains using small bottle screening 131
5.2.3 Percentage yield and stoichiometry. 136
5.2.4 Optimisation of WT *C. saccharoperbutylacetonicum* N1-4HMT batch fermentation for lactic acid production using the response surface method 141
5.2.5 Response surface method for optimisation of WT, +NhaA, +Cpa1 and ΔGlpF for organic acid production 144
5.2.6. *C. saccharoperbutylacetonicum* N1-4HMT knock-out ΔPssA 152
5.2.6.1. Percentage yield and stoichiometry. 156
6. General discussion, alternative strategies, conclusions and future work 159
6.1 General discussion 159
6.1.1 Are any of the strains adequate for the industrial production of organic acids? 160
6.2 Alternative strategies 163
6.3 Future work 164
6.4 Conclusions 166
7. References 167
8. Appendix 180
8.1 Cpa1 sequencing 180

8.2 NhaA sequence 191

List of abbreviations

A. thiooxidans – Acidithiobacillus thiooxidans

ABE – Acetone, Butanol, Ethanol

ACE – allele couple exchange

ack – acetate kinase

adh – butanol dehydrogenase

adhE – aldehyde dehydrogenase

amp^R – ampicillin resistant

AOCS – American Oil Chemists' Society

AQPs - aquaporins

ARAi – arabinose inducer

ATP – Adenosine 5'-triphosphate

B. cereus – Bacillus cereus

bcd – butyryl-CoA dehydrogenase

BHBD – β-hydroxybutyryl-CoA dehydrogenase (oxidoreductase)

BLAST – Basic local alignment search tool

bp – base pair

BP – British Petroleum

BuAc – Butyric Acid

buk – butyrate kinase

bv – bed volume

C. – Clostridium

$C_2H_3O_2$ – acetic acid

C_2H_6O - ethanol

C_3H_6O – acetone

$C_3H_6O_3$ – lactic acid

$C_4H_8O_2$ – butyric acid

$C_4H_{10}O$ - butanol

C5/C6 – five-carbon/six-carbon

$C_6H_{12}O_6$ - Glucose

CAGR – Compound Annual Growth Rate

CaCA – Ca^{2+}: Cation Antiporter

Cam^R – chloramphenicol resistant

CAS – Chemical abstracts service

CGM – Clostridial Growth media

ClC – Chloride carrier/channel

CO_2 – Carbon dioxide

CoA – Coenzyme A

CPA – cation/proton antiporter

Cpa1 – monovalent cation:proton antiporter 1

CRISPR – Clustered, Regulated Interspaced, Short, Palindromic Repeats

crt – crotonase

ddH_2O – double-distilled water

DDM – n-Dodecyl-β-maltoside

DMPC – 1,2-dimyristoyl-sn-glycero-3-phosphocholine

DMSO – dimethyl sulfoxide

DNA – Desoxyribonucleic acid

dNTPs – deoxynucleoside triphosphates

DOE – Department of Energy

DOPC – 1,2-dioleoyl-sn-glycero-3-phosphocholine

DSB – double stranded break

E. coli – Escherichia coli

EC – Enzyme Commission

EDTA – Ethylenediaminetetraacetic acid

EMBL-EBI – European Bioinformatics Institute

EPB_NS – electroporation buffer without salt

EPB_S – electroporation buffer with salt

erm – erythromycin resistant

Ex/Em – excitation-emission

F. acidarmanus – Ferroplasma acidarmanus

F1,6DP – fructose-1,6-diphosphate

FA – Fatty acid

fab – fatty acid biosynthetic

fba – fructose biphosphate aldolase

$FeCl_3 \cdot 6H_2O$ – Iron (III) chloride hexahydrate

FD/MS – field desorption Mass spectrometry

FOA – 5-fluoroortic acid

FTIR – Fourier-transform infrared spectroscopy

g – grams

g – acceleration of gravity

g/L – grams per litre

GC – gas chromatography

gap – glyceraldehyde-3-phosphate dehydrogenase

GHG – Greenhouse gas

GlpF – Glycerol facilitator/ Glycerol uptake facilitator

G.O.I. – gene of interest

gpm – 2,3-biphosphoglycerate-independent phosphoglycerate mutase

GRAS – generally recognised as safe

GYP – Glucose-Yeast-Peptone

h – hour

HPLC – high performance liquid chromatography

HR – homologous recombination

ICl – iodine monochloride

IEA – International Energy Agency

IEP – intron encoded protein

IPTG – Isopropyl β-d-1-thiogalactopyranoside

IV – iodine value

KCl – potassium chloride

kDa – kilo Dalton

KI – potassium iodide

L – Litre

L. - Lactobacillus

LAc – Lactic acid

LB – Luria Bertani

LC – liquid chromatography

LDH – lactate dehydrogenase

M – molarity

MFS – Major facilitator superfamily

mg – milligram

min - minute

mL – millilitre

MS – mass spectrometry

MW – molecular weight

N - normality

NAD(P)H – Nicotinamide adenine dinucleotide phosphate

$Na_2S_2O_3$ – Sodium thiosulfate

NaCl – sodium chloride

NaOH – sodium hydroxide

NCBI – National Center for Biotechnology Information

NhaA - Na^+:H^+ Antiporter A

NhaC - Na^+:H^+ Antiporter C

NHEJ – non-homologous end-joining

NIR – near-infrared

nm – nanometer

NMR – nuclear magnetic resonance

nt – nucleotide

OA – organic acids

OD – optical density

OG – octyl-β-glucoside

PAM – Proto-adjacent motif

PCR – polymerase chain reaction

pdb – protein data bank

PerM – autoinducer – 2 Exporter

pfk – 6-phosphofructokinase

pfl – pyruvate-ferrodoxinoxidoreductase

pgk – phosphoglycerate kinase

pH – potential of hydrogen

PHB – poly-3-hydroxybutyrate

Pi – inorganic phosphate

pKa – acid dissociation constant

POPC – 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine

pta – phosphotransacetylase

ptb – phosphate acetyltransferase

pyk – pyruvate kinase

pyrE/pyrF – orotidine 5-phosphate decarboxylase

RAM – retrotransposition-activated selection marker

RCM – Reinforce Clostridial media

RE – restriction enzyme

RhtB – Resistance to homoserine/threonine

RNA – Ribonucelic acid

rpm – revolutions per minute

RMSD – root mean square deviation

s - second

S. cerevisiae – Saccharomyces cerevisiae

SDS – sodium dodecyl sulfate

SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEM – standard error of the mean

SFA – saturated fatty acid

SMA – styrene-maleic acid

spp. – several species

SPV – sulfo-phospho-vanillin

TAE – Tris-acetate-EDTA

TAG - triacylglyceride

TBS-T – Tris-buffered saline Tween

TC – transporter classification

tc^R – tetracycline resistant

TEMED – Tetramethylethylenadiamine

thl – thiolase

ThrE – threonine/serine exporter

TLC – Thin Layer chromatography

TM – transmembrane

UFA – unsaturated fatty acid

UV - Ultraviolet

µ - micro

v – volume

V - voltage

w – weight

WT – Wild type

List of Figures

Figure 1.1 ABE fermentation of solventogenic Clostridia 23

Figure 1.2 Lactic acid and lactate. 26

Figure 1.3 Metabolic pathway from glucose to lactate in solventogenic Clostridia 28

Figure 1.4 Butyric acid and butyrate. 29

Figure 1.5 Metabolic pathway from acetyl-CoA to butyrate in *C. saccharoperbutylacetonicum* 30

Figure 1.6 Packing of saturated, *trans* or *cis* unsaturated FA and the relation with membrane fluidity. 32

Figure 1.7 Group II intron, ClosTron. 36

Figure 1.8 Allele-coupled exchange ACE 37

Figure 1.9 Aims of the project according to each chapter 41

Figure 3.1 Chemical reactions of the Iodine Value method 62

Figure 3.2 Iodine Value of sunflower oil. 64

Figure 3.3 Comparison of the iodine values of different amounts of sunflower oil using the AOCS method and a spectrophotometer. 65

Figure 3.4 Titration of I_2 with different concentrations of 15 µL of 0.1 N $Na_2S_2O_3$ 66

Figure 3.5 Detection limit of the IV assay. 67

Figure 3.6. Performance of the IV assay under different conditions 68

Figure 3.7 Schematic representation of 18:1 (cis) DOPC and 14:0 DMPC 69

Figure 3.8 Determination of IV for lipid standards mixtures of 18:1 and 14:0. 71

Figure 3.9 Schematic representation of 22:6 (cis) PC and 22:0 PC 71

Figure 3.10 Initial determination of IV for lipid standards. 73

Figure 3.11 Miniaturised IV assay of the highest % UFA lipid standards with double the amount of ICl. 74

Figure 3.12 Miniaturised IV assay of the lowest % UFA lipid standards with double ICl, 80µL at 1M. 74

Figure 3.13 Determination of IV for the lipid standards. 76

Figure 3.14 Linear regression between the theoretical and the experimental IVs. 77

Figure 3.15 Flowchart of the miniaturised iodine Value assay and its reactions. 78

Figure 3.16 IV assay of *E. coli* lipid membrane grown at different temperatures. 79

Figure 3.17 Average double bond in *E. coli* lipid membranes 80

Figure 3.18 IV assay of yeast lipid grown at 32 ˚C and compared with commercial yeast lipids. 81

Figure 4.1 Schematic two-dimensional representation of the cation:proton antiporters 87

Figure 4.2 Sequence alignment between *E. coli* and *C. saccharoperbutylacetonicum* NhaA 88

Figure 4.3 Overall architecture of *C. saccharoperbutylacetonicum* NhaA.. 89

Figure 4.4 Ramachandran plot for the *C. saccharoperbutylacetonicum* NhaA protein 90

Figure 4.5 Comparative modelling of *E. coli* vs *C. saccharoperbutylacetonicum* NhaA 91

Figure 4.6 Sequence alignment of *P. abyssi* NhaP, *C. saccharoperbutylacetonicum* Cpa1 and *E. coli* Cpa1 92

Figure 4.7 Schematic two-dimensional representation of the Cs-Cpa1 93

Figure 4.8 Schematic representation of aquaporin.. 94

Figure 4.9 Sequence alignment between *C. saccharoperbutylacetonicum* (Cs) and *E. coli* (Ec) GlpF sequences 95

Figure 4.10 Overall architecture of *C. saccharoperbutylacetonicum* GlpF 96

Figure 4.11 Ramachandran plot for the *C. saccharoperbutylacetonicum* N1-4HMT GlpF protein 97

Figure 4.12 Comparative modelling of *E. coli* vs *C. saccharoperbutylacetonicum* GlpF 98

Figure 4.13 Cloning strategy. 99

Figure 4.14 Schematic representation of the pTZ19R/U-NhaA-His and pTZ19R/U-Cpa1-His 100

Figure 4.15 PCR products from the *cpa1* gene 101

Figure 4.16 PCR products from the *nhaA* gene.. 101

Figure 4.17 Double digestion from a random selection of colonies 102

Figure 4.18 Sequence alignment of *nhaA*.. 103

Figure 4.19 PCR products of *cpa1* and *nhaA* with their promoters 104

Figure 4.20 PCR products of NhaA and promoter-NhaA 105

Figure 4.21 PCR products of *cpa1* from different concentrations of gDNA................................... 106

Figure 4.22 Double digestion of pET15b, *nhaA* and *cpa1*... 106

Figure 4.23 Double digestion of pBR322, promoter-NhaA and promoter-Cpa1 107

Figure 4.24 pET17-CsGlpF Double digestion from random colonies.. 108

Figure 4.25 SDS-PAGE CsGlpF induction 109

Figure 4.26 Western blot of CsGlpF His-tag... 110

Figure 4.27 SDS-PAGE gel of GlpF His-tag of 1L of starting culture.. 111

Figure 4.28 Western blot of GlpF His-tag of 1 L of starting culture 111

Figure 4.29 Western blot of GlpF Hist-tag of 1 Lt of starting culture and DDM for its extraction..... 112

Figure 4.30 Ec-GlpF expression in SDS and Western blot 113

Figure 4.31 *C. saccharoperbutylacetonicum* N1-4HMT *glpF* PCR.. 113

Figure 4.32 Double digestion of pET15b and GlpF 114

Figure 4.33 Colony PCR from pET15b-Cs-GlpF transformation.. 115

Figure 4.34 Cs-GlpF solubilisation with OG or SMA 116

Figure 4.35 Cs-GlpF solubilisation with DDM 117
Figure 5.1 pSec-NhaA colony PCR 121
Figure 5.2 Colony PCR from pSec-Cpa1 121
Figure 5.3 CLEAVE knock-out procedure 123
Figure 5.4 Design of CLEAVE knock-out plasmids 124
Figure 5.5 HR- GlpF knock-out construct for genome integration 125
Figure 5.6 PAM sequences identified in GlpF gene for CLEAVE technology 126
Figure 5.7 Spacer vector construction 126
Figure 5.8 HR-GlpFKO-HR colony PCR 127
Figure 5.9 GlpF spacer (plasmid w/GlpF spacer) and S3V9 128
Figure 5.10 S3V9-GlpFspacer colony PCR 129
Figure 5.11 Primer design for ΔGlpF confirmation and sequence alignment 130
Figure 5.12 HPLC example of +NhaA strain at 48 h 131
Figure 5.13 OD600, pH and HPLC measurements of glucose, organic acids and solvents from the small bottle screening of +NhaA, +Cpa1 and WT *C. saccharoperbutylacetonicum* N1-4HMT strains 134
Figure 5.14 OD600, pH and HPLC measurements of glucose, organic acids and solvents from the small bottle screening of ΔGlpF and WT *C.saccharoperbutylacetonicum* N1-4HMT 136
Figure 5.15 Yield of organic acids, solvents and biomass of the different strains 138
Figure 5.16 Stoichiometry calculations of WT strain 139
Figure 5.17 Carbon balance of WT, +NhaA, +Cpa1 and ΔGlpF after 48 h of fermentation 140
Figure 5.18 3D response surface plot of the WT strain for lactic acid production 143
Figure 5.19 Yield of organic acids and solvents of the different strains when 1 % (w/v) of salt is added 148
Figure 5.20 Carbon distribution of WT, +NhaA, +Cpa1, ΔGlpF strains after 48 h of fermentation in salt conditions 149
Figure 5.21 Software prediction of +Cpa1 and +NhaA for organic acid production 150
Figure 5.22 Graph of the g/L of the organic acids and solvents of the different strains when salt, butyric and lactic acid are added 151
Figure 5.23 Metabolic pathway of the phospholipid headgroup synthesis in bacteria and its alteration if PssA is knocked-out. 153
Figure 5.24 HPLC example of ΔPssA strain at 24 h 154
Figure 5.25 Small bottle screening measurements for ΔPssA and WT C. saccharoperbutylacetonicum N1-4HMT 156

Figure 5.26 g/L of the organic acids and solvents after 48 h of fermentation from WT and ΔPssA strains 157

Figure 5.27 Carbon distribution of WT and ΔPssA strain after 48 h of fermentation 158

List of Tables

Table 1.1 "Top 10" target structures for biorefinery........ 25

Table 2.1 Plasmids used for Biocleave........ 45

Table 2.2 Thermal cycler parameters........ 46

Table 2.3 Primers used in the project........ 47

Table 2.4 Double digestion reaction........ 48

Table 2.5 Ligation reaction........ 48

Table 2.6 SDS-PAGE solutions........ 51

Table 2.7 Western blot solutions........ 51

Table 2.8 Clostridial transformation buffers........ 53

Table 2.9 HPLC conditions for solvents acids and sugar........ 54

Table 3.1 Advantages and disadvantages of techniques for detection and quantification of lipids and their ability to detect unsaturation........ 57

Table 3.2 Analysis of 18:1 and 14:0 lipid mixture using the miniaturised IV assay........ 70

Table 3.3 Analysis of 22:6, 22:0, 18:1 and 14:0 lipid mixtures using the miniaturised IV assay........ 72

Table 3.4 Analysis of lipid mixes using the miniaturised IV assay........ 75

Table 4.1 Proton antiporters in *C. saccharoperbutylacetonicum* obtained from TransportDB2.0 web. 84

Table 5.1 Stoichiometric reaction equations of ABE fermentation........ 137

Table 5.2 Comparison of the theoretical, actual and % yield of the different organic acids and solvents 138

Table 5.3 RSM runs given by Design-expert software with the factors of temperature, pH and NaCl on WT strains. 142

Table 5.4 RSM runs and different conditions are given by Design-expert 13 software to test the factors of lactic acid (g/L), butyric acid (g/L) and salt % (w/v) as the independent variables. 146

Table 5.5 Comparison of the theoretical, actual and % yield of the different organic acids and solvents of the different strains grown at 1 % (w/v) of NaCl........ 147

Table 5.6 Comparison of the theoretical and actual yield, % yield for butanol, acetone, ethanol, lactic acid, butyric acid and biomass between WT and ΔPssA strain........ 156

Table 6.1 Summary of the different mutated clostridial strains and the different tested conditions 161

1. Introduction

Fossil fuels provide substantial economic benefits; they have increased prosperity and lifestyle on the planet. Hydrocarbon fuels' advantages include high energy density and investment over energy returned rewards. However, fossil fuels are considered non-renewable resources. British Petroleum's (BP) annual report claimed in 2013 that Earth might have over 1688 billion barrels of oil that can last for 53 years at a present use rate. Energy experts claimed otherwise, and reserves will likely double by 2050 despite rising demands. However, there is significant societal concern about the environmental cost that fossil fuels provide. Fossil fuels' detriments are air pollution, rising costs, health issues, climate change and weather extremes on Earth. Air pollution has arisen from the long combustion of fossil fuels (Kalair *et al.*, 2020, Covert *et al.*, 2016).

Chemistry greatly impacts daily life due to its involvement in generating a wide variety of products. In 2021 the chemical industry's total worldwide revenue stood at 4.73 trillion US dollars, and it is expected to grow by 1.8 per cent in 2024 (Fernandez, 2022). The chemical industry is the largest industrial energy consumer and the third largest industrial emitter of carbon dioxide (CO_2). It is responsible for approximately 7 % of global anthropogenic greenhouse gas (GHG) emissions and 5.5 % when only counting CO_2 emissions. The international Energy Agency (IEA) estimates that to limit the global temperature rise to 2 ˚C, a 30 % reduction in direct industrial CO_2 emissions by 2050 is required. This would mean a 75 % reduction in emissions per unit production by 2050 just in the chemical sector while sustaining a 2.8-fold increase in production (Levi and Cullen, 2018). In the Net Zero Emissions by 2050 scenario, government and chemical industry efforts must address CO_2 emissions to get on track to reach this goal (IEA, 2021). Biotechnology can offer economic and environmental benefits to the chemical industry and thus has a great potential to achieve the sustainable production of existing and new products from renewable feedstocks (Yang, 2007).

1.1 Biotechnology

Biotechnology is defined as "any technological application that uses biological systems (e.g., living organisms, cells, animals, plants, tissues, enzymes or *in vitro* synthetic (enzymatic) systems) to make or modify products or processes for specific uses"; according to the convention on biological diversity (Heux *et al.*, 2015, Walther *et al.*, 2015). Biotechnology aims to benefit humankind and its natural environment (Arber, 2011). Biotechnology has many applications in medicine, pharmacy, food production and agriculture, industry and environmental protection (Kordi *et al.*, 2022).

A recent linguistic invention has adopted a colour scheme to define biotechnology according to its application area. Blue is for the exploitation of marine biodiversity for cosmetics, food and feed, and aquaculture. Green represents plant engineering for producing whole plants, or in plant cells, of metabolites or macromolecules that can be used as bulk chemicals, food and pharma ingredients, or materials. Gray is developing biotechnological tools for environmental deployment, e.g., bioremediation. The application of biotechnology in the healthcare and pharmaceutical sector is red. Black is for the use of biotechnology for military purposes, including biodefence. White describes the use of enzymes, cell extracts or whole microorganisms in industrial processes that produces a wide variety of products, such as fuels, food ingredients and chemical compounds (Heux *et al.*, 2015).

1.2 White biotechnology

White (industrial) biotechnology is used in chemical, food, textile, packaging and sanitary industries. It was born in the 1910s at the University of Manchester by Chaim Weizmann. He isolated an anaerobic bacterium *Clostridium acetobutylicum,* that can produce acetone, butanol and ethanol (ABE) fermentation from starch and glucose (Webb, 2011, Yang, 2007). This platform's features were purified mono-culture microorganism instead of mixed cultures and large-scale anaerobic liquid fermentation. The lessons from the history of ABE fermentation are 1) a focus shift of products over time and 2) market needs driving rapid technology development. At the start of the twentieth century, the interest in alternative feedstock and routes to produce synthetic rubber was stimulated due to a natural rubber shortage produced by trees and soaring rubber prices. The chemical firm Strange and Graham Ltd decided that butadiene and isoprene could be the best precursors, which can be prepared by oxidation of *n*-butanol and isoamyl alcohol, respectively. However, the world rubber market collapsed in the same year, so making synthetic rubber from *n*-butanol was not economically appealing. In World War I, the product goal of this pioneering study was refocused on acetone due to Britain's critical need for a solvent for manufacturing smokeless explosive cordite. This fermentation scaled up rapidly to produce 30 000 tonnes of acetone per year. When acetone was produced, twice as much *n*-butanol was co-produced, second in volume as an industrial biomanufacturing product only to ethanol at that time. Surplus stock of *n*-butanol led to its use in new markets: a substitution for amyl acetate in lacquers and its use in the manufacturing of solvents, plasticisers, paints and resins (Walther *et al.*, 2015). The ABE fermentation process ceased during the early 1960s because of competition from petroleum-based solvents and the high price of sugar substrates (Yang, 2007). Now Chaim Weizmann is recognised as the father of industrial fermentation (Walther *et al.*, 2015). In the 1970s, with the first oil crisis, experts in the field realised that oil-based economic development was not sustainable (Yang, 2007). The focus of ABE fermentation was shifted back to *n*-butanol because it is an excellent drop-in biofuel that can be blended with gasoline at any ratio (Walther *et al.*, 2015).

The change from a petro-based economy to a bio-based economy can be achieved with white biotechnology (Aguiar *et al.*, 2021). It is, therefore, crucial for the future development of the economy and provides excellent growth opportunities for various industries. It can reduce waste production, energy requirements and solvent use and eliminate hazardous intermediate products, among other benefits. White biotechnology synthesises rapidly degradable products that require less energy and creates less waste during production. It is imperative to know the cost of production and application of production biomass and by-products for various applications (Kordi *et al.*, 2022). Turning starch into ethanol is neither the most environmentally nor economically efficient method, as growing plants for ethanol production involves using herbicides, pesticides, fertilisers, irrigation and machinery. Companies such as Novozymes (Denmark), Genencor (USA) and Maxygen (USA) are therefore exploring avenues to derive ethanol specifically from celluloid material in wood, grasses and, more attractively, agricultural waste (Frazzetto, 2003). The cost of raw materials in bioprocesses can be amount to around 70 – 75 % of the total production cost. One of the strengths is that it can synthesise a wide range of chemical products. Biorefinery is a branch of the white industry that addresses the circular economy proposition of processes without waste generation (Aguiar *et al.*, 2021).

1.3 Biorefinery

A biorefinery maximises the outputs of a conventional refinery by using sustainable and low-environmental impact technology, which converts biomass into high-value chemicals (Lin et al., 2013; Menon and Rao, 2012). It is the sustainable processing of biomass into a spectrum of marketable products (food, feed, materials, chemicals) and energy (fuels, power, heat) (Hingsamer and Jungmeier, 2019). Biomass consists of carbohydrates, lignin, proteins, and fats (Fernando *et al.*, 2006). During 2022-2026, the global biorefinery market is expected to grow by $651.89 mn, progressing at a CAGR (Compound Annual Growth Rate) of 11.56 % during the forecast period (ReportLinker, 2022).

Currently, biomass is needed for human food and animal feed, producing paper, fuels, power, and heat. Therefore, biomass must be used to produce these products sustainably and efficiently. All biorefineries have to be assessed over the whole value chain on their social sustainability, economic, and environmental covering the whole life cycle (construction-operation-dismantling) (Hingsamer and Jungmeier, 2019). As biomass availability is limited, it should be used efficiently, maximising biomass conversion efficiency - minimising raw material requirements and effectively producing a spectrum of marketable products and energy (de Jong and Jungmeier, 2015).

The two product groups of biorefineries are energy (bioethanol, biodiesel, synthetic biofuels, electricity, heat, and synthetic biofuels) and products (e.g., materials, chemicals, food and feed) (Hingsamer and Jungmeier, 2019). Feedstock groups can be energy crops from agriculture that include perennial grasses, starch crops (wheat and maize), sugar crops (beet and cane), lignocellulosic crops (managed forest, short rotation coppice, switchgrass), lignocellulosic residues (e.g. stover and straw), oil crops (palm and oilseed rape); aquaculture (algae, seaweeds); organic residues (industrial, commercial and post consumer waste such as cooking oils). Conversion processes are classified into four groups: biochemical (fermentation and enzymatic conversion), thermochemical (gasification and pyrolysis), chemical (synthesis and esterification) and mechanical (fractionation and pressing) (Hingsamer and Jungmeier, 2019). The use of cheap feedstocks is crucial. A large amount of research is dedicated to lignocellulosic biomass. However, few microorganisms can metabolise the pentoses derived from these raw materials, so this biomass has not been used extensively to date (Sauer *et al.*, 2008). The pretreatment of lignocellulosic biomass is necessary because it separates the constituents (lignin, cellulose, and hemicellulose) and allows the enzymes to efficiently catalyse the conversion of cellulose and hemicellulose into fermentable sugars. The most commonly used pretreatments are mechanical (milling and grinding), chemical (acid, alkali and some organic solvents), physicochemical (hydrothermal, steam explosion, ammonia, fibre explosion, wet oxidation) and biological techniques (bacterial and fungal pretreatment). The biological pretreatment methods are more suitable than physical and chemical techniques because of their cost-effectiveness and environment-friendly nature; however, they are restricted due to the long pretreatment time (Singh *et al.*, 2022).

Biorefineries aim to transform plentiful biological material (biomass feedstock) into valuable products using a combination of technologies and processes (Fernando *et al.*, 2006). Microorganisms are the basis of biorefineries and the backbone of industrial bioprocesses; they produce desired chemicals or intermediates required for the process (Goyal *et al.*, 2020). The selection of the most appropriate organism plays a vital role in the economic feasibility of the biorefinery.

1.4 Microbial fermentation

Humans have utilised spontaneous (mixed-culture) microbial processes for thousands of years to convert a food source into another form. 9 000 years ago, ancient China made wine from a mixture of rice, honey and fruits; the Sumerians and Babylonians practised the art of brewing beer before 6 000 BC; the use of yeast for baking bread by Egyptians is recorded in the Bible (Walther *et al.*, 2015). However, it was in 1857 that Pasteur proved fermentation was caused by microorganisms (Yang, 2007).

Besides being used as an alternative source to fossil fuels and reducing the impact on the environment, products from microbial fermentations are catalogued as natural products, which is currently a trend with increasing demand (Pessoa *et al.*, 2019). Consumer behaviours and lifestyles are changing. There is more knowledge and awareness of ingredients; an increase in life expectancy makes consumers more health conscious. They are also more concerned about climate change, plastic pollution, pesticides and chemicals, child labour, and working conditions. Therefore, consumers demand natural, organic, ethical products with environmentally sustainable and responsible practices (New hope network staff, 2022, Ecovia Intelligence, 2022). The use of microorganisms still needs to be economically competitive. In order to be used industrially, microorganisms must satisfy the following characteristics: grow in large-scale cultures, be genetically stable and amenable to genetic manipulation, not be harmful to humans or economically important animals and plants, to produce the desired product and be able to grow in a relatively inexpensive liquid and not be vulnerable to bacteriophages (Ditu and Gheorghe, 2017). The industrially important microorganisms include *E. coli* for Gram-negative bacteria, yeasts and some of the Gram-positive bacteria of Clostridial species (Wendisch *et al.*, 2006, Sauer *et al.*, 2008).

1.5 Clostridia

Clostridium is a rod-shaped, anaerobic, Gram-positive, spore-forming bacterium (Cho *et al.*, 2015). There are non-pathogenic Clostridial species that are relatively easy to isolate, and their vitamin and nutritional supplementation are less stringent than other organisms (Grupe and Gottschalk, 1992, Alsaker *et al.*, 2010). The simplicity of Clostridia gives it an advantage as an industrial microorganism because medium requirements and substrate selection substantially impact the production cost (Moon *et al.*, 2016). Clostridial species can convert diverse carbon sources, including C5 and C6 substrates, into organic acids and solvents that can be applied in biofuel and chemical production in various amounts and ratios (Grupe and Gottschalk, 1992, Alsaker *et al.*, 2010). These microorganisms have been considered promising hosts for producing high-value chemicals. Clostridial strains naturally produce acetate, butyric acid, ethanol, isopropanol, butanol, 1,3-propanediol, 2,3-butanediol, acetone, and hydrogen when grown on simple hexoses (Kell, 1981). Their primary products are acetone, butanol and ethanol (ABE) (Jiang *et al.*, 2009, Grimmler *et al.*, 2011, Fontaine *et al.*, 2002). The reindustrialisation of ABE fermentation is hampered by its significant production cost, linked to high product inhibition and low product yield (Gonzalez-Peñas *et al.*, 2020). The ideal microorganism should be able to produce biofuels and high-value chemicals simultaneously to maximise its value even further, optimising the microbial production process and reducing waste (da Silva *et al.*, 2014). The high-value products enhance profitability, while the high-volume fuels help to meet the global energy demand. Hence, it is essential to look at the value-added chemicals produced from the integrated biorefinery, which economically and technically support the fuel and power produced from these refineries (Fernando *et al.*, 2006). Organic acids are one of the most attractive targets in biological routes using biomass as feedstock (Li *et al.*, 2016b).

1.5.1 The acetone-butanol-ethanol fermentation

A typical acetone, butanol and ethanol (ABE) fermentation consists of two stages: the acidogenic and the solventogenic phase (Figure 1.1). During exponential growth, the acidogenesis phase commences. In the acidogenic phase, the organic acids acetate, butyrate, and lactate, as well as CO_2 and H_2, are produced inside the cell and released into the external environment. As a result, a surplus of adenosine triphosphate (ATP) is generated. However, a balance between the formation and utilisation of reducing equivalents [reduced/oxidised nicotinamide adenine dinucleotide (phosphate) ratio $NAD(P)H/NAD(P)^+$] is required. Glucose is converted to pyruvate, which undergoes a series of reactions to produce acetyl-CoA. The branch point is acetyl-CoA, which follows three separate and simultaneous pathways to produce lactic, acetic and butyric acids leading to a reduction in the pH of the medium (Li *et al.*, 2020). The energetically favourable production of acetate results in a net generation of 4 mol of ATP and 2 mol of glycolytic NADH. Only 3 mol of ATP is formed from butyrate production. In order to sustain the continued operation of glycolysis under anaerobic conditions, the regeneration of NAD^+ is necessary. For the regeneration of NAD^+, *Clostridium* can take two approaches: reducing pyruvate to lactate or forming butyryl-CoA from acetoacetyl-CoA. Therefore, even though butyrate production yields less ATP than acetate production, it is redox neutral and required by cellular metabolism. Butyrate fermentation typically results in the formation of acetate and lactate, which represent a balance between energetic efficiency NAD and a need to reoxidise NADH during the metabolism of pyruvate (Xu and Jiang, 2011). The theoretical stoichiometric equation is the following:

1 Glucose -> 0.85 butyrate + 0.1 acetate + 0.2 lactate + 1.8 CO_2 + 1.9 H_2 (Zhu and Yang, 2003)

During the transition to the solventogenic phase, growth ceases in the stationary phase. In the solventogenic phase, organic acids are re-assimilated; the acids can permeate the cell membrane and be converted into solvents acetone, butanol and ethanol in a ratio of 3:6:1, respectively. Acetone is produced via acetoacetyl-CoA, butanol via butyryl-CoA and ethanol via acetyl-CoA (Mayank *et al.*, 2013, Li *et al.*, 2020, Collas *et al.*, 2012). Endospore formation is initiated simultaneously (Cheng *et al.*, 2019), which is the dormant form of gram-positive bacteria in response to unfavourable environmental conditions stopping the production of solvents (Kubiak *et al.*, 2012). However, one is not an essential prerequisite for the other (Cheng *et al.*, 2019). The shift from the acidogenesis phase to the solventogenesis phase is due to the stress signals of the acidified environment (pH 4.5) and the higher levels of ATP and NAD(P)H/NAD(P)$^+$, which activate the synthesis of solventogenic enzymes (Cheng *et al.*, 2019, Li *et al.*, 2020, Cooksley *et al.*, 2012).

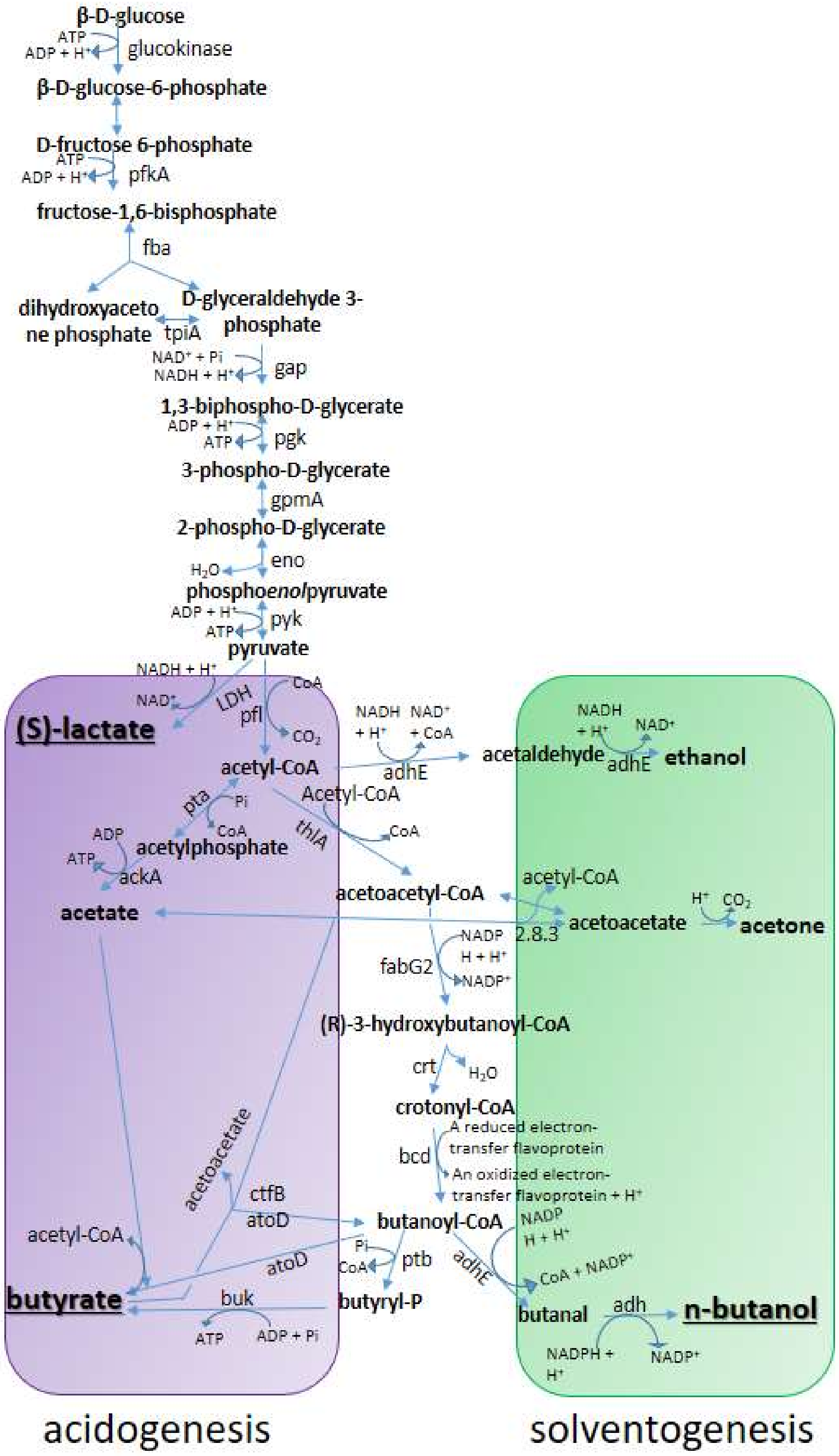

β-D-glucose
ATP
ADP + H⁺
glucokinase
β-D-glucose-6-phosphate
D-fructose 6-phosphate
ATP
ADP + H⁺
pfkA
fructose-1,6-bisphosphate
fba
dihydroxyacetone phosphate
tpiA
D-glyceraldehyde 3-phosphate
NAD⁺ + Pi
NADH + H⁺
gap
1,3-biphospho-D-glycerate
ADP + H⁺
ATP
pgk
3-phospho-D-glycerate
gpmA
2-phospho-D-glycerate
H₂O
eno
phosphoenolpyruvate
ADP + H⁺
ATP
pyk
pyruvate
NADH + H⁺
NAD⁺
LDH
pfl
CoA
CO₂
(S)-lactate
acetyl-CoA
NADH + H⁺
NAD⁺ + CoA
adhE
acetaldehyde
NADH + H⁺
NAD⁺
adhE
ethanol
pta
Pi
CoA
Acetyl-CoA
thlA
CoA
ADP
ATP
acetylphosphate
ackA
acetate
acetoacetyl-CoA
acetyl-CoA
acetoacetate
H⁺
CO₂
acetone
2.8.3
NADP
H + H⁺
NADP⁺
fabG2
(R)-3-hydroxybutanoyl-CoA
crt
H₂O
crotonyl-CoA
bcd
A reduced electron-transfer flavoprotein
An oxidized electron-transfer flavoprotein + H⁺
acetoacetate
ctfB
atoD
butanoyl-CoA
NADP
H + H⁺
CoA + NADP⁺
adhE
acetyl-CoA
atoD
Pi
CoA
ptb
butyrate
buk
butyryl-P
ATP
ADP + Pi
butanal
adh
n-butanol
NADPH + H⁺
NADP⁺
acidogenesis
solventogenesis

Figure 1.1 ABE fermentation of solventogenic Clostridia

The ABE fermentation presents a biphasic metabolism starting with acidogenesis (left panel) followed by solventogenesis (right panel). Cell growth occurs during the acidogenesis phase concurrently with acids accumulation (acetic, butyric and lactic). Acetic, butyric and lactic acid accumulation stimulates specific sets of enzymes to convert acids into solvents such as acetone butanol and ethanol. The cells enter solventogenesis, a phase revealed by a pH increase. Lactate dehydrogenase (1.1.1.27); acetate kinase (ack); butanol dehydrogenase (adh); aldehyde dehydrogenase (adhE); butyrate kinase (buk); fructose biphosphate aldolase (fba); crotonase (crt); CoA transferaseA/B (2.8.3), butyryl-CoA dehydrogenase (bcd); glyceraldehyde-3-phosphate dehydrogenase (gap); pyruvate-ferrodoxinoxidoreductase (pfl); phosphotransacetylase (pta); phosphate acetyltransferase (ptb); 6-phosphofructokinase(pfk); pyruvate kinase (pyk); phosphoglycerate kinase (pgk); 2,3-biphosphoglycerate-independent phosphoglycerate mutase (gpm); thiolase (thl)

1.6 Organic acids

Organic acids (OAs) are low-molecular-weight compounds that contain one or more carboxyl groups and are observed in almost every microorganism (Panda *et al.*, 2019). Organic acids differ in the number of carboxy groups, hydroxyl groups and carbon-carbon double bonds in their molecules. They can be classified based on the following criteria: the type of carbon chain aliphatic, alicyclic, aromatic and heterocyclic; being saturated or unsaturated; being substituted or non-substituted; the number of functional groups. Examples include acetic acid (one carboxyl group), malic acid (two carboxyl groups) and citric acid (three carboxyl groups). They exist as pure acids or in their salt forms (benzoic acid and sodium benzoate). The most common organic acids are carboxylic (-COOH) (Anyasi *et al.*, 2017). Because of their functional groups, they have numerous applications in industries related to biocomodities, such as food, feed, pharmaceuticals, cosmetics, surfactants, textile and bioplastic industries. Organic acids are generally recognised as safe (GRAS) (Panda *et al.*, 2019). Most organic acids are weak acids; in aqueous solutions, they work as a buffer, and this property is used to produce food and feed (Anyasi *et al.*, 2017). They have also been used as antioxidative agents or synergists. Production of organic acids for commercial exploitation is carried out either by chemical synthesis or fermentation.

Most organic acids are naturally obtained as intermediates of important metabolic pathways of microorganisms. The initial production practices of organic acid date back to 1823 with acetic acid and 1913 with citric acid (Panda *et al.*, 2019). The biological productions of organic acids via microbial fermentation have advantages over the chemical production processes, such as being cost-effective, feasible, reliable, environmentally friendly, sustainable and have lower carbon footprints (Coban, 2020). Their highly oxidised nature makes biosynthesis of organic acids more atom economy, which is the atoms of the starting material that end up in the product (Parent, 2003). The co-products from organic acid production in biological routes, adenosine triphosphate (ATP) and thiphophopyridine nucleotide (NADPH), can be utilised by microorganisms (Chen and Nielsen, 2016). Another critical advantage of microbial fermentation is the production of optically pure compounds compared to chemical synthesis, resulting in racemic mixtures (Singh *et al.*, 2022). Most biomolecules are chiral and consist of only two possible enantiomeric forms. As a result, molecular interactions of a racemic substrate will always be different for one enantiomer than the other, playing a critical role in many industries (Rachwalski *et al.*, 2013). Enantiomer products arise from two distinct enzymes that function in an enantiodivergent and distinct mechanistic manifold. Some single enzymes can produce both enantiomers but depend on the substrate and the stereo-specificity. Therefore, in several products, such as lactic acid, the formation of pure enantiomers through microorganisms is well known and easy to control (Finefield *et al.*, 2012). The synthesis of enantiomerically pure compounds remains a challenge in modern organic chemistry (Rachwalski *et al.*, 2013). The number of acidic compounds isolated from microorganisms exceeds 100. Only a few organic acids are produced industrially solely by microbial processes. They have attracted renewed interest as they are seen as building block chemicals (molecules with multiple functional groups) that can be transformed into new families of usable that represent high-valuable chemicals. In 2004, the US Department of Energy (DOE) released a report identifying ten chemicals that can be produced from renewable carbon sources through microbial fermentation (Papagianni, 2011).

1.7 High-value chemicals

High-value chemicals are compounds with industrial significance and range from polymer and plastic precursors to fuels. As there is a wide range of targets for biorefinery, in 2010, the US Department of Energy (DOE) released a list of the “Top 10” target structures for biorefinery (Table 1.1). This selection was based on the following criteria: the chemicals are currently commercial products, can directly substitute existing petrochemicals, have received significant attention in the literature, represent robust chemical platforms and serve as a primary building block (Bozell and Petersen, 2010). Clostridia naturally produce ethanol, glycerol derivatives, biohydrocarbons, lactic acid and succinic acid.

Table 1.1 "Top 10" target structures for biorefinery
Extracted from the US Department of Energy (2010)

Ethanol
Furans
Glycerol and derivatives
Biohydrocarbons
Lactic acid
Succinic acid
Levulinic acid
Sorbitol
Xylitol

1.8 Lactic acid

Listed as a building block for the future by the US Department of Energy (Bozell and Petersen, 2010), lactic acid or 2-hydroxypropionic acid (CAS 50-21-5) (Figure 1.2) is a chiral organic hydroxyacid. Two isomers depending on the stereochemistry of the tertiary carbon can be found, the levorotatory (L-lactic acid) and dextrorotatory (D-lactic acid) (Lebarbe, 2013). Lactic acid was first isolated from sour milk in 1798 (Papagianni, 2011). Lactic acid is commonly used in the food industry as a food acidulant, preservative and flavour-enhancing agent (Sahasrabudhe and Sankpal, 2001). It is also relevant in the cosmetic industry as a pH regulator, antimicrobial agent, moisturiser and skin lightener (Ajala *et al.*, 2020). Sodium lactate is frequently used in the pharmaceutical industry as an electrolyte in parenteral and intravenous solutions, while calcium lactate is used for calcium deficiency therapy and as an anticavities agent. Lactic acid is a precursor of acrylic acid, pyruvic acid, biosolvents and esters, which are important in the chemical industry (Vandenberghe *et al.*, 2018). Lactic acid demand has increased considerably in recent years because it can be converted to methyl lactate, lactide and it is the precursor in the synthesis of polylactic acid (PLA), a biodegradable and biocompatible plastic that is used as a fibre in textile, film catgut packaging and scaffold medical (Thongchul *et al.*, 2019). Lactic acid is a yellow to colourless, odourless liquid. It is both an alcohol and an acid. The lactic acid market is evaluated at $1.25 thousand million as of 2019, with an estimated compound annual growth rate (CAGR) of 11.5 % between 2020 and 2026 (Baptista *et al.*, 2021).

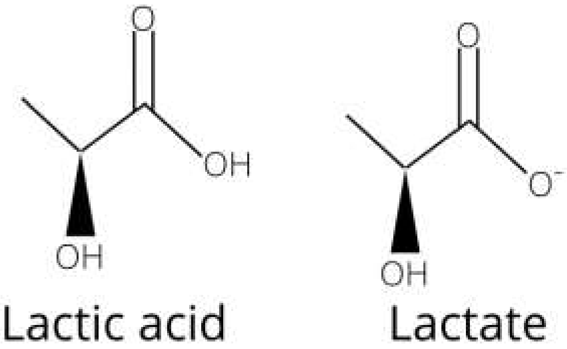


Figure 1.2 Lactic acid and lactate.
Lactic acid is a three-carbon molecule. The central carbon is bonded with a hydrogen atom, a carboxylic acid group, a hydroxyl group, and a methyl group. Lactate is lactic acid missing one proton. Created with Canva.

Lactic acid synthesis is chemically obtained with the reaction of acetaldehyde and hydrogen cyanide under high pressure in the presence of a base to produce lactonitrile. The hydrolysis of lactonitrile by strong acids (sulphuric acid) is based on chemically achieving lactic acid with a 50/50 mixture of the optically inactive L- and D-form. Ammonium salt is produced as a by-product (Li and Cui, 2010).

The main advantage of its production by a fermentation process is that pure L- or D- isomers of lactic acid can be obtained (van Maris *et al.*, 2004). The advantages of microbial fermentation include cheap renewable substrates, low production temperatures, and low energy consumption. Because of these advantages, the production process is used most often (~ 90 % of total lactic acid production) (Miller *et al.*, 2011, Komesu *et al.*, 2017). Other organic acids can be produced during lactic acid fermentation depending on the purity and quality of the inoculum, the metabolic route used, and process conditions in which carbon dioxide, and ethanol, can be produced, but for efficient lactic acid production, the formation of by-products should be avoided or kept to a minimum (Komesu *et al.*, 2017).

The biological production of lactic acid was traditionally carried out using lactic acid bacteria (Chen and Nielsen, 2016). The lactic acid formation involves the catabolism of glucose or other hexoses through the glycolytic hexose biphosphate route and subsequent NADH regeneration by pyruvate reduction. The following reaction summarises the homolactic fermentation:

$C_6H_{12}O_6 + 2NAD^+ + 2ADP + 2\ Pi \rightarrow 2C_3H_6O_3 + 2NAD^+ + 2ATP$

In this pathway, the catabolism of 1 molecule of hexose produces 2 molecules of lactic acid; therefore, the theoretical yield equals 1 g of lactic acid per gram of hexose (Papagianni, 2011).

This process is not economically viable for large industrial-scale fermentations owing to the requirements of nutritionally rich media and moderate pH conditions. The pH of the fermentation decreases according to the amount of lactic acid produced (Komesu *et al.*, 2017). There is a need for microorganisms tolerant to low pH and less stringent nutritional requirements (Chen and Nielsen, 2016).

In solventogenic Clostridia, lactate dehydrogenase (LDH [EC 1.1.1.27]) catalyses the reduction of pyruvate to lactic acid (Figure 1.3). Lactate is often not the primary fermentation product. Under stressful conditions such as iron or sulfate limitation, or CO_2 gassing at neutral pH, lactate becomes the primary fermentation product. Lactate is not produced unless the conversion of pyruvate to acetyl-CoA is partially blocked. The blockage of pyruvate increases the intracellular concentration of fructose-1,6-diphosphate (F1,6DP), which activates LDH. Clostridial LDH is a nonreversible enzyme activated by F1,6DP (Gheshlaghi *et al.*, 2009).

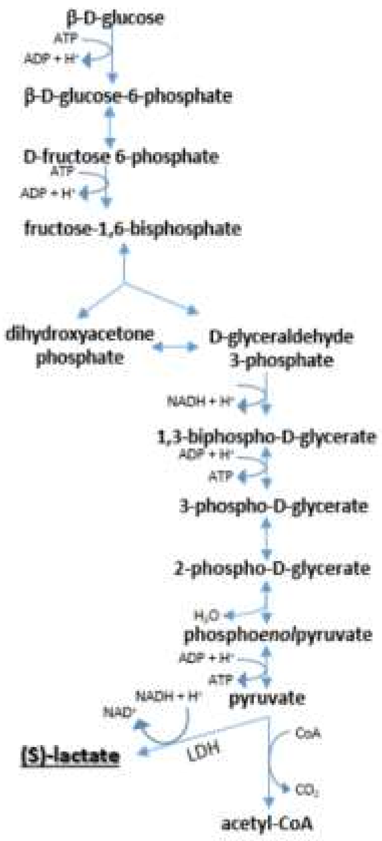

β-D-glucose
ATP
ADP + H⁺
β-D-glucose-6-phosphate
D-fructose 6-phosphate
ATP
ADP + H⁺
fructose-1,6-bisphosphate
dihydroxyacetone phosphate
D-glyceraldehyde 3-phosphate
NADH + H⁺
1,3-biphospho-D-glycerate
ADP + H⁺
ATP
3-phospho-D-glycerate
2-phospho-D-glycerate
H₂O
phosphoenolpyruvate
ADP + H⁺
ATP
NADH + H⁺
NAD⁺
pyruvate
CoA
(S)-lactate
LDH
CO₂
acetyl-CoA

Figure 1.3 Metabolic pathway from glucose to lactate in solventogenic Clostridia
Overview of the metabolic pathway for lactic acid production from glucose to lactate and acetyl-CoA in solventogenic Clostridia. LDH, Lactate dehydrogenase.

1.9 Butyric acid

Although butyric acid is not on the "top 10" list, it is the precursor for butanol production, the most valuable solvent due to its potential to substitute gasoline, diesel and kerosene (Berezina *et al.*, 2009). Butyric acid (CAS 107-92-6) (Figure 1.4) is a short-chain volatile fatty acid with an unpleasant odour (Bastidas-Oyanedel *et al.*, 2019). Pasteur, who suggested the name "anaerobic" to describe the organism since it was killed by air, discovered butyric acid-producing fermentation in 1861 (Xu and Jiang, 2011). It is used in the beverage, food and cosmetic industries as a fragrance and flavouring agent to supply butter-like notes (Bastidas-Oyanedel *et al.*, 2019). It is used to produce thermoplastics, cellulose acetate-butyrate plastics and the poly-3-hydroxybutyrate (PHB) in the chemical industry. Glycerol tributyrate and other esters also play an important role in plastic materials (Xu and Jiang, 2011). It has many applications in medicine because it can potentially treat hypercholesteraemia, cancer and hemoglobinopathies (Bastidas-Oyanedel *et al.*, 2019). It has a worldwide market of approximately 80 000 metric tons per year for $ 1.8/kg or £1.4/kg.

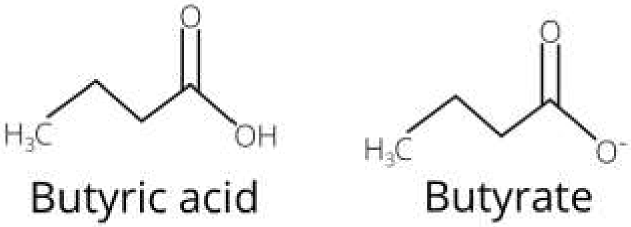


Figure 1.4 Butyric acid and butyrate.
Butyric acid is a saturated short-chain fatty acid with four carbon backbone structure. Butyrate is the conjugate base of butyric acid. Created with Canva.

In the chemical industry, butyric acid is obtained by the air oxidation of butyraldehyde. Butyraldehyde is produced from the oxo reaction of propylene, which is the addition of a hydrogen atom and a carbon monoxide to the double bond of propylene with heat and pressure (Jaros, 2014).

Biotechnological butyric acid production is not commercially competitive because of its low productivity and butyrate concentration in the fermentation broth. However, food and pharmaceutical manufacturers prefer food additives or pharmaceutical products produced biologically. Improvements in the economics and efficiency of the butyrate fermentation process are necessary (Xu and Jiang, 2011).

The theoretical glucose fermentation with complete conversion of butyrate is shown as follows:

$C_6H_{12}O_6$ -> 0.8 $C_4H_8O_2$ + 0.4 $C_2H_3O_2$ + $2CO_2$ + 2.4 H_2 (Xu and Jiang, 2011)

In this pathway, the catabolism of 1 molecule of hexose produces 0.8 molecules of butyric acid; therefore, the theoretical yield equals 0.39 g of butyric acid per gram of hexose.

In solventogenic Clostridia, butyrate and butanol formation from acetyl-CoA occurs through acetoacetyl-CoA, β-hydroxybutyryl-CoA and butyryl-CoA as intermediates (Figure 1.5). β-hydroxybutyryl-CoA dehydrogenase (oxidoreductase) (BHBD [EC 1.1.1.35]) catalyses the reduction of acetoacetyl-CoA by NAD(P)H. An essential step toward the production of butyrate and butanol. Butyryl-CoA dehydrogenase (BCD [EC 1.3.99.2]) act in a reverse direction and catalyses the conversion of crotonyl-CoA to butyryl-CoA, a key branch point in butyrate and butanol formation. Butyrate kinase (BUK [EC 2.7.2.7]) catalyses the reaction by which butyryl-P is converted to butyrate with concomitant ADP phosphorylation. This phosphorylation reaction is an important energy source during the acid formation stage. BUK plays a vital role in butyrate uptake. Meyer *et al.,* (1986) found that when the continuous butyrate-producing culture of this organism was sparged with carbon monoxide, there was a high uptake rate of butyrate with concomitant formation of butanol but no acetone (Gheshlaghi *et al.*, 2009).

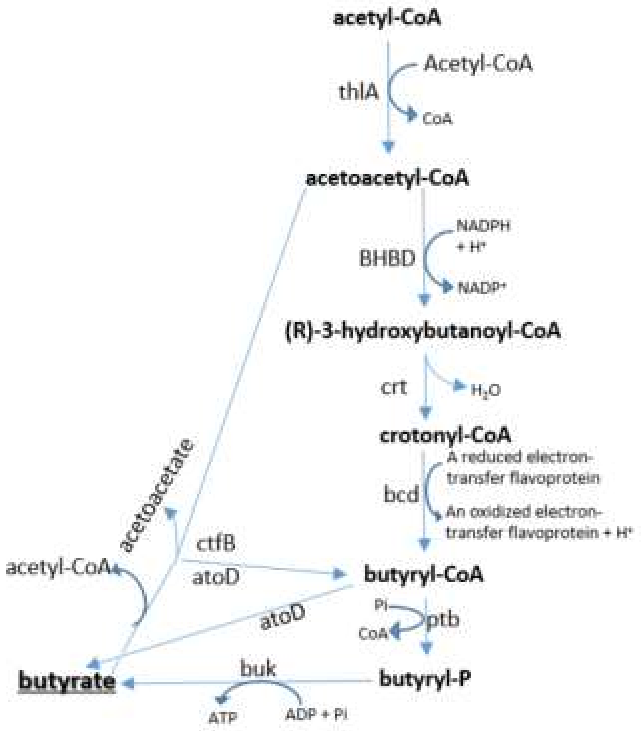

acetyl-CoA
Acetyl-CoA
thlA
CoA
acetoacetyl-CoA
NADPH + H+
BHBD
NADP+
(R)-3-hydroxybutanoyl-CoA
crt
H2O
crotonyl-CoA
A reduced electron-transfer flavoprotein
bcd
An oxidized electron-transfer flavoprotein + H+
acetoacetate
ctfB
atoD
acetyl-CoA
butyryl-CoA
atoD
Pi
CoA
ptb
butyrate
buk
butyryl-P
ATP
ADP + Pi

Figure 1.5 Metabolic pathway from acetyl-CoA to butyrate in *C. saccharoperbutylacetonicum*
The metabolic pathway involved in butyrate fermentation from acetyl-CoA. thlA, thiolase. BHBD, 3-hydroxybutyryl-CoA dehydrogenase. Ptb, phosphotransbutyrlase. Buk, butyrate kinase.

1.10 Limitations in the fermentation of organic acids

Biological reactions are slower than chemical reactions. Enzymes and microorganisms are exposed to product toxicity. Increasing the scale is one of the most challenging issues in biological processes. Setting up the large-scale production process with waste from biological processes will save time and cost (Kordi *et al.*, 2022). The economics of a fermentation process depends on the yield obtained, fermentation period and choice of raw material. An ideal fermentation process must use raw materials which are cheap, abundant and also are amenable to be converted into a valuable metabolite in high yields during a relatively short fermentation cycle (Sahasrabudhe and Sankpal, 2001). Microbial strains are required to tolerate several stresses, resulting in compounds' high yield and productivity (Goyal *et al.*, 2020).

The accumulation of organic acids acidifies the environment for microbial growth affecting it negatively (Guan and Liu, 2020). Most organic acids exist in an undissociated form (XCOOH) at low pH conditions. They can directly diffuse into the cytoplasm, where pH is more neutral, quickly releasing protons ($XCOO^-$, H^+) and reducing the intracellular pH, affecting cell growth, denaturing protein and damaging DNA (Piper, 2011) (Xu *et al.*, 2019). The charged anion cannot diffuse outside the cell, inducing protective stress responses; the proton released acidifies the cytosol (Pei *et al.*, 2019). Titres and productivities of 50 – 100 g/L and 2 – 3 g/L/h are expected for the economical manufacturing of most building block acids by fermentation. These would result in a pH reduction to around 2.0 for titres of 50 g/L (Warnecke and Gill, 2005). Thus, during the production of organic acids, bacteria would necessitate using neutralisation agents such as calcium carbonate to maintain fermentations at neutral pH, which is higher than the *pKa* values of all organic acids. Subsequently, the salt form of organic acids is obtained. An acidification step using strong acid (sulphuric acid) is thus necessary to recover the undissociated forms of organic acids. However, it leads to the formation of gypsum, rendering downstream processing costly (Tran and Zhao, 2022). Alternative technologies, such as *in situ* acid removal by extraction, could be used with the production process integration. However, choosing acid-tolerant strains as the host for development efforts is an alternative option to pH titration as it minimises the consumption of bases for pH control and thus reduces the cost of downstream processing (Chen and Nielsen, 2016). What is desired is an organism that produces high levels of organic acid chemicals and is tolerant to any associated toxicity (Warnecke and Gill, 2005). Some of the protective mechanisms the microorganisms have include: 1) protection against proton flow into the cell by changing the membrane fatty acid composition, 2) proton excretion and consumption, and 3) protection against acid-induced damage, e.g., by chaperones and nucleases (Kudo and Sasaki, 2019).

1.11 Lipid membrane

The membrane is the primary target for damage induced by environmental stresses. As the first barrier of the cell, it plays a vital role in cellular growth, metabolism, energy transduction, and maintaining a constant intracellular environment. The membrane regulates substances entering and exiting the cells and the catalysis of exchange reactions (Zhu and Yang, 2003). The membrane lipids are amphipathic: a polar end that anchors the lipid to the aqueous interface and a long hydrophobic segment that forms the oily membrane interior (Stillwell, 2016). This characteristic makes the membrane impermeable to charged ions, except by specific transport mechanisms. Protons penetrate membranes poorly, as do charged anions and, strong acids. Uncharged acid molecules diffuse rapidly through the plasma membrane if they are lipid-soluble (Gurtler and Mai, 2014). An essential lipid component of membranes is a fatty acid (FA). A FA is a monocarboxylic acid with a long (typically 14 to 24 carbons), unbranched, a hydrophobic tail that can be either saturated (S) or unsaturated (U) (double bonds; *cis* if the hydrogens are on the same side, *trans* if they are on opposite sides). FA are mainly responsible for membrane structure, function and fluidity (Stillwell, 2016). The fluidity of the membrane is commonly defined as the degree of molecular order and motion of membrane constituent (fatty acyl chain conformation, lateral diffusion, rotational diffusion, and transbilayer diffusion (“flip flop”) (Wu *et al.*, 2012), especially lipid components. Several studies have demonstrated that the degree of unsaturation of FA in phospholipids is the main factor regulating membrane fluidity (Hagve, 1988). The presence of even a single double-bond is sufficient to exert a profound influence on physical properties (Figure 1.6). Stress conditions can change the organisation, dynamic structure, and function of membrane lipids; however, the integrity and fluidity of the cytoplasmic membrane are critical factors in maintaining the viability of cells and their metabolic activities (Zhu and Yang, 2003). Maintaining membrane fluidity is essential for regular lipid-protein interaction and membrane-bound enzymes such as ATPase function.

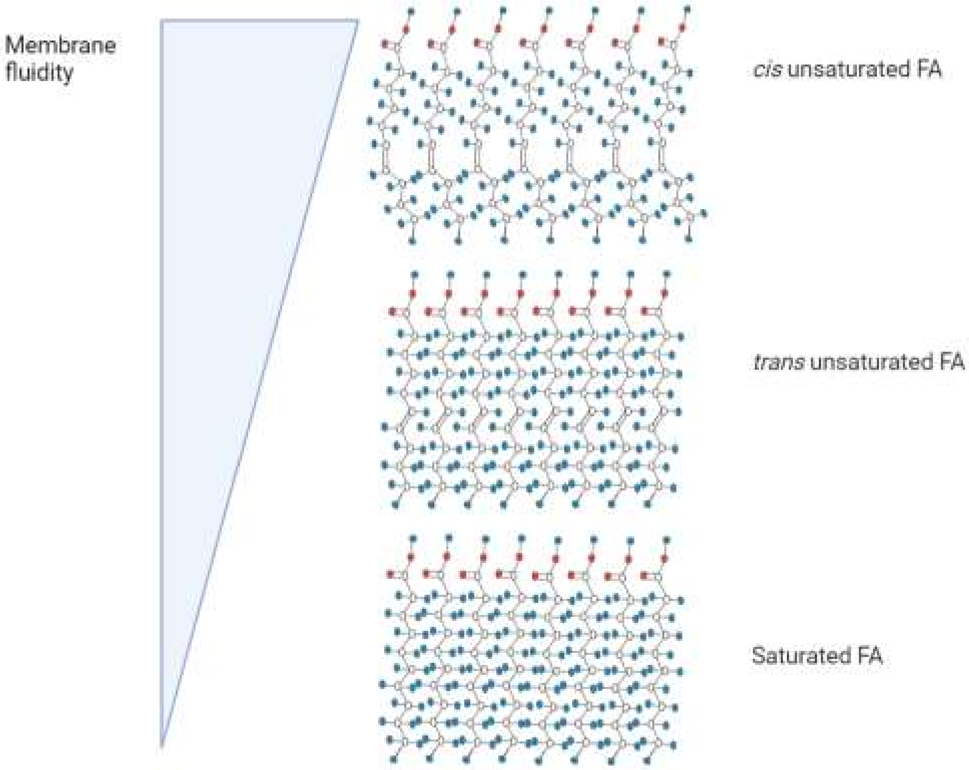


Figure 1.6 Packing of saturated, *trans* or *cis* unsaturated FA and the relation with membrane fluidity. Saturated FAs are fatty acids without double bonds. Unsaturated FA contains one or more double bonds; if the hydrogens of the double bond are on the same side are *cis;* if they are on the opposite side are *trans*. *Cis* double bond generates a kink in the FA. Saturated FA can pack more tightly; therefore, the membrane is more rigid and less fluid than unsaturated FA. Created in biorender.com

Microorganisms such as *S. cerevisiae* and *E. coli* can regulate membrane lipid composition to increase their tolerance. Increasing saturated FA composition in cell membranes would decrease the membrane fluidity and probably also change the membrane permeability (Wu *et al.*, 2012). To restrict the entry of protons into the cells, acid-tolerant microbes are generally equipped with less permeable membranes (Guan and Liu, 2020). Altering the unsaturation ratio is the most common mechanism employed by bacteria to control membrane fluidity (Sohlenkamp and Geiger, 2016). This depends on FA synthases of the anaerobic pathways and desaturase enzymes of the aerobic pathways. High unsaturation ratios have been reported to contribute to cell survival at low pH. Changes from *cis* to *trans* conformation also affect the fluidity of the bacterial membrane (Tan *et al.*, 2016). Altering the proportion or type of branching and membrane cyclopropane acyl changes are another way in which cells modulate membrane fluidity and can increase low pH sensitivity (Guan and Liu, 2020). Another important alteration to increase survival in acidic environments is the FA chain length. The increased length decreases cell membrane fluidity (Zhu and Yang, 2003, Wu *et al.*, 2012).

Serious membrane rigidification induced by lactic acid stress has been reported. A rigid membrane could reduce the fluidising effect of lactic acid so that by preventing its entry, the deleterious effect of lactic acid may be minimised. During lactic acid production, *L. bulgaricus* CFL1 reduced the ratio of both unsaturated FA:saturated FA and cyclic FAs to saturated FA in the membrane during acidification (Wu *et al.*, 2012).

In butyric acid producer cells, the toxicity of butyrate in the membrane has been observed. Butyrate affects the membrane interface, strengthening the interaction of cations and negatively charged phospholipids. Zhu and Yang (2003) reported that *C. tyrobutyricum* had more saturated FAs (C16:0 and C17:0) and less unsaturated FAs (C18:1) in its cell membrane.

Engineering or regulating membrane function can improve the robustness of industrial strains. The membrane lipid composition can be manipulated by regulating the type of phospholipid, which includes adjusting the degree of lipid unsaturation (Qi *et al.*, 2019). An integrated picture of the cellular process and properties such as transport and morphology (membrane fluidity) are needed to make rational modifications to increase strain robustness (Sauer *et al.*, 2008).

1.12 Proton accumulation

Accumulation of anions and protons induces multiple stress conditions via cytosolic acidification and modifications of membrane lipids and proteins (Palma *et al.*, 2018). Many bacteria activate enzymes when exposed to acid conditions, contributing to pH homeostasis. Cells may pump protons out of the cell, prevent protons from leaking in and counteract acidification of the cytoplasm by producing alkaline compounds (Mols and Abee, 2011). In lactic acid producers like *L. plantarum* WCFS1, the genes that have been found to play a role in pH control or the maintenance of the proton motive force are proton-translocating F_1F_0-ATPase; several sodium-proton antiporters; and amino acid decarboxylases that use an intracellular hydrogen ion for the decarboxylation of an imported amino acid (Pieterse *et al.*, 2005). In *B. cereus,* F_1F_0-ATPase has been observed to transport protons outside the cell at the expense of ATP in acidic conditions. At lethal pHs sodium - proton transporters like napA and nhaC were highly induced (Mols and Abee, 2011). *C. saccharoperbutylacetonicum* contains an RnF complex which converts reduced ferredoxin to NADH, generating an ion gradient (protons or Na^+) across the cytoplasmic membrane. This ion gradient can be used for additional ATP synthesis via the ATPase to maintain ATP concentrations during solventogenesis (Poehlein *et al.*, 2017). Exporting the protons and acid anions through H^+-ATPase and efflux pumps requires high cellular energy consumption (Baek *et al.*, 2016). Engineering strains with increased proton pumps to export protons out of the cell can improve low pH tolerance, achieving an economical and efficient fermentation process for producing organic acids.

1.13 Genetic engineering in Clostridial spp.

Strain robustness, which is the ability of the microorganism to withstand the production environment, is a crucial factor in determining whether a microbial process will be successful and industrially viable (Sauer *et al.*, 2008). Developing a robust strain with significantly improved lactic and butyric acid yield and tolerance is a prerequisite for organic acid production to be economically competitive with chemical synthesis routes (Jiang *et al.*, 2018). Genetic engineering is the required tool to alter organisms rationally to create such tolerant strains (Sauer *et al.*, 2008).

Metabolic studies of industrial Clostridia have been targeted towards: i) growth at high cell density (Xu *et al.*, 2015), ii) continued cell viability (Harris *et al.*, 2001), iii) no spore formation (Huang *et al.*, 2004), iv) high butanol selectivity (Tummala *et al.*, 2003), v) tolerance to aerobic condition (Hillmann *et al.*, 2008), vi) high tolerance to acetone, butanol and ethanol (Tomas *et al.*, 2003), vii) direct utilisation of lignocellulosic biomass (Nanda *et al.*, 2017). However, the genetic manipulation of Clostridial species is still challenging: low transformation and recombination efficiencies have hampered efforts to make stable recombinant strains. Various techniques and tools have been developed for many Clostridial species.

1.13.1 Type II intron

The group II intron (Targetron or "ClosTron" when applied in the context of the *Clostridium* genus) is a transposable genetic element based on retrohoming, a phenomenon that an intron cassette can jump into other loci in the genome through reverse splicing (Figure 1.7). After the intron is transcribed, splicing occurs by the intron-encoded RNA maturase *LtrA*, followed by the formation of a complex with *LtrA* and the spliced intron. This RNA-protein (RNP) complex can bind to other sites by base pairing between the specific region of the intron (exon-binding site; EBS) and the target, after which LtrA finishes the remaining steps for insertion: reverse splicing and reverse transcription (Lee, 2021). Among various mobile group II introns, *L1.trB* and *Tel3c/4c* have been extensively used for gene knock-out in the solventogenic *Clostridium* (Kwon *et al.*, 2020). The modified *L1.LtrB* intron includes a retrotransposition-activated selection marker (RAM) which enables antibiotic selection. The specific targeting is accomplished through protein-specific DNA binding and a programmed RNA that is complementary to a 13 nucleotide region that allows for the specification of the DNA target site by altering the intron sequence. However, the target site is limited to DNA sites compatible with the intron-encoded protein (IEP) (Joseph *et al.*, 2018). Curing and off-target manipulation are one of the main limitations. The cloning of *pyrF* (orotidine 5-phosphate decarboxylase) to the mobile intron enhanced curing efficiency. *pyrF* encodes the essential enzyme of pyrimidine biosynthesis, which can use 5-fluoroortic acid (FOA) and converts it to a toxic compound; in the ClosTron plasmid, only the cured strain could survive in the FOA added model. ClosTron accidentally affects and manipulates off-target genomes causing unexpecting genotypes and phenotypes. A highly regulated ClosTron system has been developed by inducing an L-arabinose inducer (ARAi) to reduce off-target possibility (Kwon *et al.*, 2020). This approach has many limitations, one being that it only allows in-frame insertion. On small genes, insertion might not be possible due to splicing efficiency. Suppose an intron is inserted into the sense orientation of the target gene. In that case, the LtrA protein catalyses the splicing of the original intron RNA and the mutant mRNA transcribed from the disrupted gene, resulting in the revival of the wild-type mRNA. Scarless gene knock-out is impossible. The advantages are that it is simple, efficient and convenient. It has been used in a wide range of Clostridial species (*C. acetobutylicum, C. beijerinckii, C. botulinum* as well as *C. difficile (Joseph et al., 2018)),* high insertion frequencies which can be screened by PCR or antibiotic selection and is independent of homologous recombination (Lee, 2021). Jiang *et al.* (2009), using ClosTron, eliminated the gene encoding acetoacetate decarboxylase (adc), enhancing butanol yield in *C. acetobutylicum* EA2018.

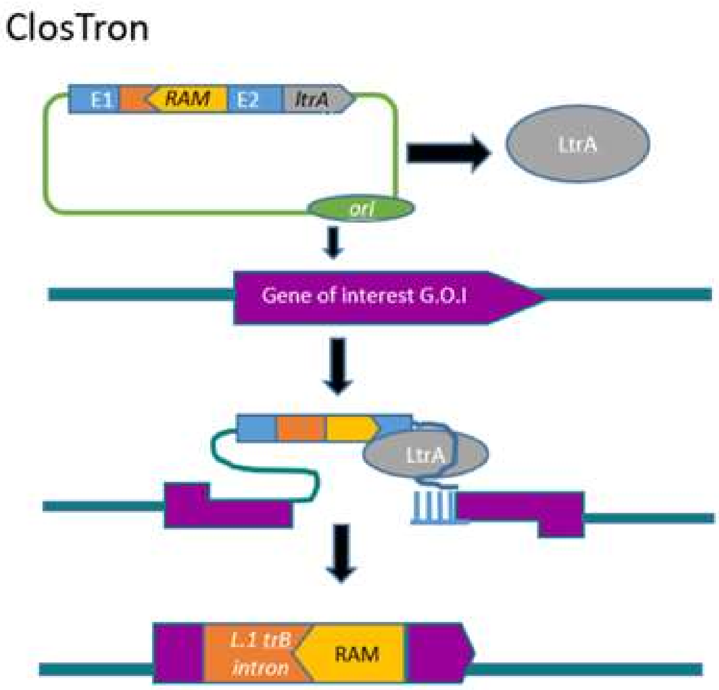


Figure 1.7 Group II intron, ClosTron.
Mobile group II intron-base genome is also known as ClosTron. Site-directed gene disruption is achieved by inserting the mobile group II intron into the target locus of the chromosome. E1 and E2 are exon-binding sites. RAM, retrostrancription-activated marker (typically kanamycin resistant marker). G.O.I. gene of interest. LtrA, encodes the RNA maturase and performs the splicing of L1. ltrB.

1.13.2 Allele-Couple Exchange (ACE)

The Allele couple exchange (ACE) method has been developed, which facilitates the insertion of complex heterologous DNA of varying sizes into the host genome (Kwon *et al.*, 2020). The strategy employs a double-crossover recombination mutation, which employs two techniques without any counter-selectable marker on the plasmid (Figure 1.8). First, the right homologous arm (~ 1 200 bp) is much longer than the left arm (~ 300 bp) to ensure the first crossover occurs dominantly through the right arm (plasmid integration). Once this is verified by colony PCR, the mutant is subjected to selection pressure to select double-crossover mutants. A promoterless selectable marker downstream of the left arm in the insertion cassette is located for this selection. This marker gene is only expressed after the second crossover under the control of the upstream promoter on the chromosome (plasmid excision) (Lee, 2021). *pyrE* or *pyrF* is the most frequently used selectable marker (Kwon *et al.*, 2020). This is because *pyrE* deficient strains have easily distinguishable phenotypes compared to wild-type, becoming auxotrophic for uracil and resistant to the toxic uracil analogue 5-fluoroorotic acid (5-FOA) (Li *et al.*, 2016a). *pyrE* is located immediately upstream of the homologous arm and directs the first crossover event in which the entire plasmid is incorporated into the genome. The excision of the plasmid backbone occurs at the second crossover event directed by the left arm, which is homologous to the internal region of the *pyrE* gene. This second recombination replaces the wild-type *pyrE* gene with a truncated form producing a mutant that can be screened based on 5-FOA resistance. Alternatively, a promoterless heterologous *pyrE* gene or antibiotic marker (usually erythromycin-resistant gene, *erm*) can be inserted in the integration vector with the regions of homology such that a successful double crossover event places the silent gene directly downstream of a constitutive promoter (Joseph *et al.*, 2018).

There are several limitations to this approach. The promoter-free marker must be expressed upstream of the promoter after integration; therefore, the site cannot be arbitrary. The first crossover must occur to the right arm. The advantage is that replicating plasmid compensates for a low transformation efficiency and/or recombination frequency. It can also integrate large gene clusters (Lee, 2021). ACE has been used in *C. acetobutylicum, C. sporogenes* and *C. difficile* (Joseph *et al.*, 2018). Willson *et al.* (2016) integrated the cellulosomal enzymes from *C. cellulolyticum* into *C. acetobutylicum* improving their growth phenotype on xylose and an increase in released reducing sugar from several substrates.

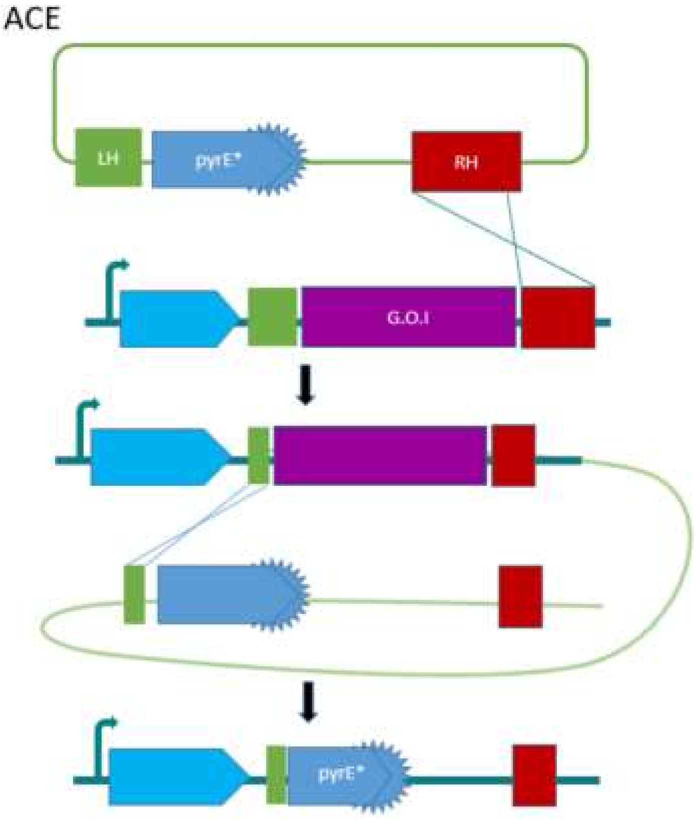


Figure 1.8 Allele-coupled exchange ACE.
Allele-coupled exchange tool for genome engineering. Double-crossover event, first by the right homology arm (RH), which is longer than the left homology arm (LH). The successful double homologous recombination inserts a promoter-less copy of the *pyrE* gene (pyrE*) directly downstream of a native constitutive promoter (light blue pentagon) in exchange for the gene of interest (G.O.I.). *pyrE* encoding orotate phosphoribosyl transferase, used

as a counter-selection marker to ensure a double crossover event, the *pyrE*-mutant (pyre*) is resistant to 5-fluoroorotic acid (FOA).

1.13.3 CRISPR/Cas

CRISPR/Cas system (CRISPR = Clustered, Regularly Interspaced, short, Palindromic Repeats) is a system that describes a "prokaryotic adaptive immune system") where the bacteria or archaea cell can protect itself from invading DNA (Nagaraju *et al.*, 2016). The CRISPR system has been engineered as a genome editing tool for eukaryotic and prokaryotic cells. If the cell encounters the invading DNA, which is directed by a fragment called a "spacer" and is flanked by direct repeats, the cell will recognise it and destroy it by cleaving it with the Cas endonuclease. The sequence that CRISPR/Cas recognises as the invading DNA is called "Protospacer", which contains the spacer copy and a short sequence called PAM (proto-adjacent motif). If the PAM is not present or mutated, the invading DNA will not be recognised and destroyed (Atmadjaja *et al.*, 2019, Kwon *et al.*, 2020). Studies in *E. coli* have shown that double-stranded breaks can enhance homologous recombination in bacteria, whereby homology-directed repair (HDR) occurs after a break has been induced. CRISPR/Cas9 represents a significant advancement in *Clostridium* gene editing as scarless edits (Joseph *et al.*, 2018). The CRISPR system has been used to implement gene deletions and insertions in *C. acetobutylicum, C. beijerinckii, C. ljundgahlii, C. autoethanogenum, C. difficile, C. pasteurianum, C. saccharoperbutylicum* and *C. cellulolyticum* (Lee, 2021, Joseph *et al.*, 2018, Kwon *et al.*, 2020). The limitations are that Cas9 expression can be toxic to the cell; therefore, its expression needs to be controlled. There is a possibility of off-target mutations; the gRNA must be carefully designed. The advantages are scarless editing; it can be used for knockdown, overexpression, and insertions (Lee, 2021, Kwon *et al.*, 2020). Wang *et al.* (2017) deleted using CRISPR/Cas *pta* and *buk* in *C. saccharoperbutylacetonicum*, obtaining a butanol production of 19 g/L.

1.13.4 CLEAVE^{TM} technology by Bio-Cleave Ltd

Clostridial spp lack or have inefficient non-homologous end-joining (NHEJ) systems, so a Cas9-mediated chromosomal DSB results in cell death (Joseph *et al.*, 2018). Jenkinson and Krabben (2015) developed a process by directing a "killing vector" to unmated targeted sequences, enabling the production of the desired mutation quickly and efficiently. The technology takes advantage of *Clostridium*, which cannot repair the cleaved DNA and is lost from the population in a dual-vector process. First, a homologous recombination step is needed, and then the use of the CRISPR-cas mechanism to target the population that did not carry the desired modification. This technology does not overcome the inherent inefficiency of homologous recombination in Clostridia, but it significantly reduces the time needed to identify the modified colonies (Atmadjaja *et al.*, 2019, Jenkinson and Krabben, 2015).

This molecular tool is ideal for modulating protein and lipids targets that can increase organic acid tolerance in Clostridia.

1.14 Rationale

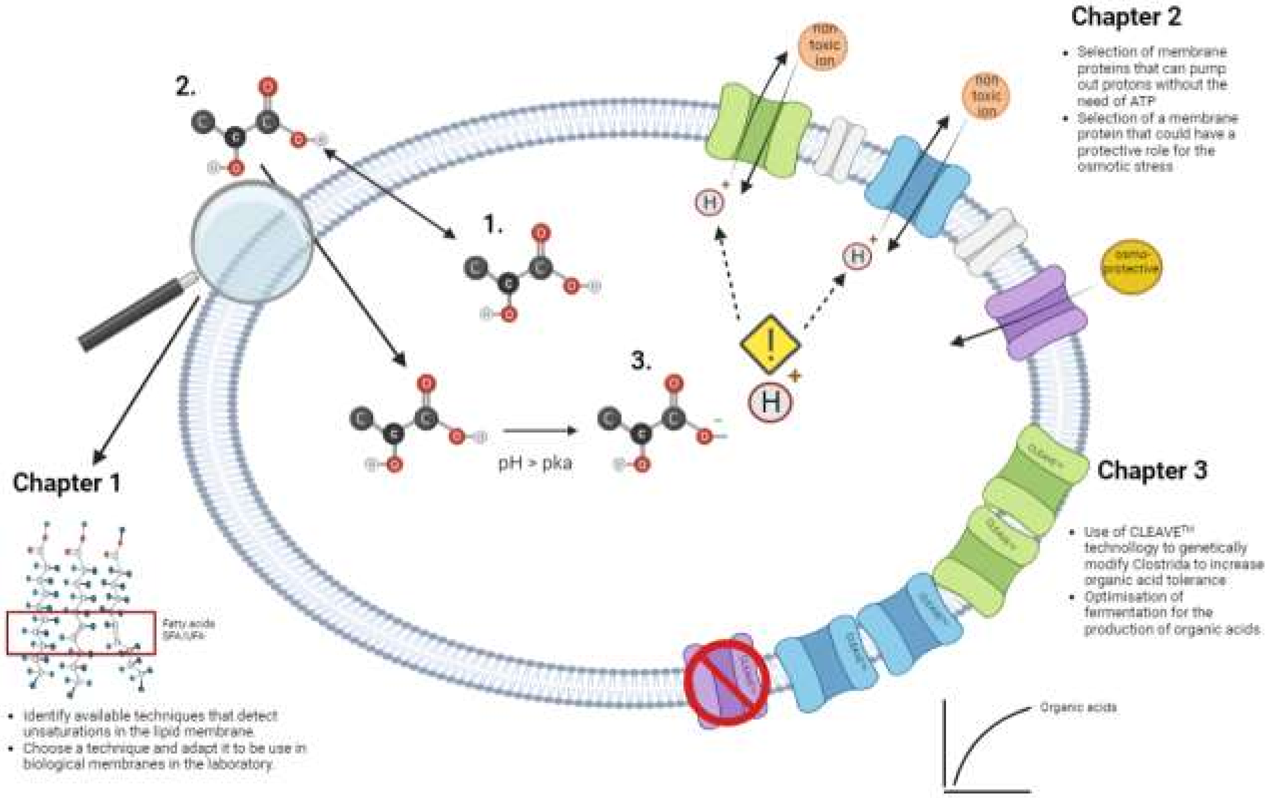
A high product titre can lower the facility and utility cost for separation (Cheng *et al.*, 2019). Tolerance engineering focuses on developing strategies to increase tolerance and increase yield. A promising method for industrial-scale fermentation is the improvement of tolerance to organic acids and acidification. Decreasing organic acid diffusion is a much more energetically efficient method than relying on the active extrusion of protons and acids (Palma *et al.*, 2018). Resistance strategies include regulating membrane lipids' composition to reduce the passive diffusion of the organic acid into the cell (Pei *et al.*, 2019). Xu *et al.* (2020) observed that *E. coli* grown at pH 4.2 increases the unsaturated FAs content in the membrane lipid. They overexpressed the genes *fabA* and *fabB,* which are essential in the biosynthesis of unsaturated FAs, and the mutants had a higher acid tolerance at 4.2. They suggested that an increase of unsaturated FAs would also change the activity of the F_0F_1-ATPase conducing to reduce membrane proton permeability and improve internal pH homeostasis, which could have a potential application in the production of organic acids. Overall, changing membrane FA composition may improve the bacterial ability to adapt to an acidic environment (Xu *et al.*, 2020). Because a higher saturation has been observed during butyric acid production in Clostridia and during lactic acid production in lactic acid bacteria, it is crucial to understand first the toxic mechanisms of the organic acids and the protective mechanisms of *C. saccharoperbutylacetonicum* N1-4HMT to then make rational membrane modifications to increase membrane robustness.

Clostridia need to be able to remove excess protons to maintain pH homeostasis (Baker-Austin and Dopson, 2007). H^+-ATPase, symporter, antiporter and secondary transporter have all been shown to promote proton efflux. However, protons are reported to be exported from cells through H^+-ATPase in bacteria, which consumes ATP (Guan and Liu, 2020). Acidophiles (*A. thiooxidans, F. acidarmanus, Sulfolobus solfataricus, Pricophilus torridus)* contain many cation transporter genes, representing an adaptation to more extremely acidic environments (Chen, 2021). Cation-proton antiporters which are secondary transporters, also play a significant role in acidification, exchanging cation (either Na^+ or K^+) for protons; its activity is driven by the concentration of Na^+ or K^+, the magnitude of membrane potential, and the pH gradient across the membrane (Booth, 1985). Secondary transporters have a potential for tolerance engineering since they can have a protective role at low pH and contribute to increasing the yield of organic acids production in *C. saccharoperbutylacetonicum* N1-4HMT.

1.15 Aims of the project

The general aim was to create Clostridial strains that are more resistant to organic acids toxicity and hence generate improved bioprocesses. The aims and chapters of this thesis are summarised in Figure 1.9.

1. Chapter 1 aimed to better understand the membrane's lipids by identifying the protective and toxic changes that occur during fermentation. Different techniques for analysing membrane fluidity (degree of unsaturation) were evaluated, and one (the Iodine Value) was adapted to be used in biological cell membranes.
2. Chapter 2 aimed to identify the membrane proteins that can pump out protons in exchange for a non-toxic ion. Two proton antiporters were chosen. The glycerol facilitator (GlpF) was also chosen to elucidate its role as a protector of osmotic stress during the transition from the acidogenic to the solventogenic phase. Their 3D structure was predicted, and their cloning, expression and purification were described to characterise their function and possible role in incrementing tolerance and organic acid production.
3. Chapter 3 built on data from previous chapters to develop a more tolerant phenotype. The chosen proton antiporters were overexpressed. Because a way to elucidate the role of a gene is to disrupt it, a GlpF knock-out strain was created using genome editing technology CLEAVE^{TM} (Bio-Cleave Ltd. The production and tolerance to organic acids in the mutant strains were evaluated, as well as strategies to prevent the solventogenic phase and optimise organic acid production were described. This will make Clostridial fermentation more attractive economically since the production of biofuels, and high-value building block chemicals will be maximised in solventogenic Clostridia.



Chapter 1

- Identify available techniques that detect unsaturations in the lipid membrane
- Choose a technique and adapt it to be use in biological membranes in the laboratory

Chapter 2

- Selection of membrane proteins that can pump out protons without the need of ATP
- Selection of a membrane protein that could have a protective role for the osmotic stress

Chapter 3

- Use of CLEAVE™ technology to genetically modify Clostrida to increase organic acid tolerance
- Optimisation of fermentation for the production of organic acids

Figure 1.9 Aims of the project according to each chapter.
The figure represents the project's aims. In a *Clostridium* cell, 1. The organic acid is produced 2. Due to its nature, it can diffuse out of the cell. 3. It can also diffuse back inside, but because the pH is higher than its pka, the organic acid will be dissociated, releasing a proton, which can be toxic to the cell. Chapter 1 identified the techniques available that detect unsaturated FAs. One was chosen and adapted to be used in biological membranes in the laboratory. Chapter 2 selected two membrane proteins that can pump out hydrogens without needing ATP and a channel that has a protective role in the acidogenic phase to express and characterise them. Chapter 3 used CLEAVE^{TM} technology on Clostridial strains to genetically modify them to increase their tolerance to organic acids. The fermentation was also optimised for producing organic acids as end-products. Created with Biorender.com

2. Materials and Methods

2.1 Results chapter I -A miniaturised iodine value assay for quantifying unsaturated fatty acids

2.1.1 Reagents

0.1 M of Iodine Monochloride solution (Acros, Organics UK) diluted in acetic acid. 15 % of potassium iodide (Acros Organics, UK); 0.1 N of Sodium Thiosulfate (Alfa Aesar, USA); 2 % Starch solution (Acros Organics, UK) diluted in water.

2.1.2 Lipid preparation

DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine), DOPC (1,2-dioleoyl-sn-glycero-3-phosphocholine), 22:0 PC (1,2-dibehenoyl-sn-glycero-3-phosphocholine), 22:6 (*Cis*) PC (1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine), *E. coli* extract total, and Yeast extract total were purchased from Avanti Polar Lipids (USA). Lipids were resuspended in chloroform. Different mixtures with those lipids were made at a final concentration of 10 mg/mL.

2.1.3 *E. coli* and yeast membranes

E. coli (DH5α) cells were grown at different temperatures. From a glycerol stock, two 50 mL conical tubes with 35 mL of Luria Bertani (LB) were inoculated with a loop (two for each temperature to be tested) and grown overnight at 180 rpm, 37 ˚C. Two overnight cultures were transferred to 300 mL of LB media and grown at 180 rpm at different temperatures (27, 30, 37, and 42 ˚C) for 48 hours. *S. cerevisiae* was grown at 32 ˚C, 200 rpm, first from a glycerol stock; two 50 mL conical tubes were inoculated with one loop in 35 mL of Glucose-Yeast-Peptone media (GYP) overnight. The next day the two overnight cultures were transferred in 300 mL of GYP and incubated at 32 ˚C, 200 rpm for 72 h. After incubation, cells were harvested by centrifugation (3 220 x *g,* 20 min). Total lipids were extracted using a modified Bligh and Dyer method (Bligh and Dyer, 1959). The pelleted cells were resuspended in 5 mL of methanol at 50 ˚C and incubated in a sonicating water bath for 15 min at 50 ˚C. Following the incubation, 5 mL of chloroform were added, and the sonication was repeated. After the second incubation, 5 mL of 0.88 % KCl were added, and the mixture was vortexed. The samples were centrifuged at 1000 x *g* for 2 min to separate the organic and aqueous phases. The lipids in the organic phase (bottom layer) were recovered and stored in glass vials at -20 ˚C. The total lipids were then quantified using the Stewart Assay (Stewart, 1980). This colourimetric assay measures the complex formed between ammonium ferrothiocyanate and phospholipids. A standard curve was required for this assay, with 0 – 50 µg of POPC (Avanti Polar, USA). 10 µL of lipids in chloroform of an unknown concentration were diluted in 2 mL of chloroform and 1 mL of ferrothiocyanate assay reagent (0.1 M $FeCl_3 \cdot 6H_2O$, 0.4 M ammonium thiocyanate). Samples were mixed, inverting the tube, and centrifuged at 14,500 x *g* for 5 minutes. The chloroform layer that became coloured was collected, and the absorbance was read at 488 nm. The results were interpolated in the standard curve to calculate the amount of mg of lipids.

2.1.4 Iodine value assay

1 mg/mL of the sample in 100 µL were placed in a glass-coated microplate reader (Thermo Fisher, USA), mixed with 40 µL of 0.1 M iodine monochloride and incubated in the dark at room temperature for 30 min. After that, 2 µL of 2 % (w/v) starch and 20 µL of 15 % potassium iodide were added. A microplate reader was used to measure the absorbance at 435 nm, dispense 5 µL of 0.1 N sodium thiosulfate, shake for 3 s and reread the absorbance. The procedure was repeated until the measurement was below 1.9, meaning the solution became colourless to the eye. The µL of sodium thiosulfate used to make the solution colourless were applied in the IV formula.

$$IV = \frac{(B - S) * N * 12.69}{W}$$

B = mL of sodium thiosulfate used for the blank
S = mL of sodium thiosulfate used for the sample
N = sodium thiosulfate normality = 0.1
W = weight of the sample used.

2.1.5 Statistical analysis

Because 1 factor was tested (IV or average double bond per FA molecule) at different temperatures, data were compared by performing a 1-way analysis of variance (ANOVA) using GraphPad Prism 8.1.0 (USA) statistical software. All values in the graphs were expressed as means +/- standard error of the mean (SEM). Differences were considered statistically significant at $P \leq 0.05$.

2.2 Results chapter II – Characterisation of membrane proteins for their role in organic acid tolerance

2.2.1 Protein modelling

Protein modelling requires at least one sequence that is sufficiently similar to the target sequence and has an experimentally defined three-dimensional structure. The experimentally defined sequence was identified through protein BLAST using the Swiss-Model tool, which compares the target sequence with its database. The template structure with the sequence closest related to the target served as the reference (at least 35% residue identity). The software superimposed the target sequence onto the three-dimension reference. The software pairs C atoms selected automatically from the highest scoring local sequence alignment. The co-ordinate file given by Swiss-Model was displayed using the molecular visualisation program PyMol (Schrodinger, 2015) to generate good-quality images and model alignment to compare the reference and the target structure.

2.2.2 Bacterial growth conditions and media

E. coli were grown in LB media at 37 ˚C. Ampicillin (amp) was used at 100 µg/mL or erythromycin at 500 µg/mL. The strains used for this project were the competent *E. coli* DH5α strain as the host for constructing *cpa1, glpF,* and *nhaA* mutant plasmids and their maintenance. *E. coli* BL21 (DE3) strain for expression of proteins and *E. coli* DH10b for construction of CLEAVE vectors.

Clostridium saccharoperbutylacetonicum N1-4HMT (kind gift from Bio-Cleave Ltd, UK) was grown anaerobically in Reinforced media (RCM) (ThermoFisher Scientific, USA); Clostridial Growth media (CGM) (5 g/L yeast extract, 0.75 g/L dipotassium phosphate, 0.75 g/L monopotassium phosphate, 0.4 g/L magnesium sulphate, 0.01 g/L iron sulphate, 0.01 g/L manganese sulphate, 1 g/L sodium chloride, 2 g/L ammonium sulphate, 2 g/L asparagine); and TYIR (2.5 g/L yeast extract, 2.5 g/L tryptone, 0.025 g/L iron sulphate, 0.5 g/L ammonium sulphate) media. CGM and TYIR were performed with 5 % (w/v) glucose, as described in Atmadjaja *et al.* (2019). Recombinant Clostridial strains were selected using 75 µg/mL of thiamphenicol or 40 µg/mL of erythromycin.

E. coli and Clostridial strains carrying plasmids were stored as glycerol stocks at -80 ˚C. Glycerol stocks were made by mixing equal parts recombinant overnight culture with sterile 50 % glycerol.

2.2.3 Plasmids, plasmid isolation.

The plasmids used in this project are listed in Table 2.1. Plasmid isolation was carried out using the Monarch® Plasmid Miniprep Kit (New England Biolabs Inc, USA).

Table 2.1 Plasmids used for Biocleave

The plasmids used in this project are described, and the reference of where they were obtained is also given. BCL = Bio-Cleave Ltd. JL = John Linney. RB = Roslyn Bill.

Plasmid	Description	Reference
pTZ19R/U	Small high copy phagemid; amp^R	ThermoFisher Scientific, USA
pET17b	Bacterial vector for expression, T7 promoter; amp^R	ThermoFisher Scientific, USA
PET15b	Bacterial vector for expression, T7*lac* promoter; amp^R	ThermoFisher Scientific, USA
pBR3222	Low copy plasmid; amp^R and tc^R	Invitrogen, USA
pMTL82151	*Clostridium-E.coli* shuttle vector, pBP1 replicon; cam^R	BCL
S3V9	CLEAVE targeting vector, ery^R	BCL
pET-GlpF	GlpF gene. pET17b vector; amp^R	JL, 2018
pMMRL01	*C. saccharoperbutylacetonicum* N1-4(HMT) Cpa1 with the insertion of *Hind*II and *Nhe*I at the 5' and C-terminal His tag and *BamH*I at the 3' end. pTZ19R/U vector; amp^R	This study
pMMRL02	*C. saccharoperbutylacetonicum* N1-4(HMT) NhaA with the insertion of *Hind*II and *Nde*I at the 5' and C-terminal His tag and *BamH*I at the 3' end. pTZ19R/U vector; amp^R	This study
pMMRL03	*C. saccharoperbutylacetonicum* N1-4(HMT) NhaA1 with the insertion of *Hind*II and *Nde*I at the 5' and *BamH*I at the 3' end. pTZ19R/U vector; amp^R	This study
pET15b-EcGlpF-myc	*E. coli* GlpF with the insertion of *Nco*I at the 5' and *BamH*I at the 3' end. pET15b vector; amp^R	RB
BC3	*C. saccharoperbutylacetonicum* N1-4(HMT) pSec-NhaA with the insertion of *Kpn*I at the 5' end and *BamH*I at the 3' end. pMTL82151 vector; cam^R	This study
BC4	*C. saccharoperbutylacetonicum* N1-4(HMT) pSec-Cpa1 with the insertion of *Kpn*I at the 5' end and *BamH*I at the 3' end. pMTL82151 vector; cam^R	This study
BC6	*C. saccharoperbutylacetonicum* N1-4(HMT) HR-GlpfKO-HR with the insertion of *Kpn*I at the 5' end and *BamH*I at the 3' end. pMTL82151 vector; cam^R	This study
BC7	*C. saccharoperbutylacetonicum* N1-4(HMT) GlpF spacer with the insertion of *Sfi*I at the 5' end and *Nco*I at the 3' end. pMA-RQ vector; amp^R	This study
BC10	*C. saccharoperbutylacetonicum* N1-4(HMT) LD-DR-GlpF spacer-DR with the insertion of *Sfi*I at the 5' end and *Nco*I at the 3' end. pS3V9 vector; ery^R	This study

2.2.4 PCR amplification

PCR amplifications were carried out using Phusion polymerase and colony PCR with OneTaq (New England Biolabs Inc., USA). The conditions are described in Table 2.2. Primers used in this project are listed in Table 2.3

Table 2.2 Thermal cycler parameters
The parameters for the PCRs for the two different enzymes are shown. Step, temperature, time, cycles and the concentration of the reagents are given.

Phusion				**Phusion reagents concentration**
Step		**Temperature**	**Time**	1x Phusion buffer
Initial denaturation		98 °C	5 min	0.2 mM dNTPs
PCR cycles (35)	Denature	98 °C	10 s	0.5 µM of each primer
	Anneal	The lowest primer Tm + 3 °C	30 s	0.02 U Phusion DNA polymerase
	Extension	72 °C	1 min/ kb	~ 50 - 250 ng DNA template
Final extension		72 °C	7 min	Final volume 50 µL with ddH_2O

OneTaq				**OneTaq reagents concentration**
Step		**Temperature**	**Time**	1x OneTaq standard reaction buffer
Initial denaturation		94 °C	10 min	0.2 mM dNTPs
PCR cycles (35)	Denature	94 °C	30 s	0.2 µM of each primer
	Anneal	The lowest primer Tm - 5 °C	30 s	0.25 U OneTaq DNA polymerase
	Extension	68 °C	1 min/ kb	~ 50 - 250 ng DNA template
Final extension		68 °C	7 min	Final volume 20 µL with ddH_2O

Table 2.3 Primers used in the project
The primers with the sequence, a brief description and the melting temperature are given.

Name	Sequence 5’ -> 3’	Description	Tm (˚C)
MRL01	GGGGAAGCTTGCTAGCATGGGAGCCTTAATTGCAATTTTTC	Forward primer of *cpa1* gene (*Hind*III and *Nhe*I restriction sites are inserted)	60.9
MRL02	CCCCGGATCCTTATCTTAAAATATCGCTTTCAATATAATTTAA	Reverse primer of *cpa1* gene (*BamH*I restriction site is inserted)	60.2
MRL03	CCCCGGATCCTTAATGATGATGATGATGATGTCTTAAAATATCGCTTTCAATATAATTTAA	Reverse primer of *cpa1* gene (*BamH*I restriction sites and His tag are inserted)	58.5
MRL04	GGGGAAGCTTCATATGAATAATAAAATCAAAAGAAAAGTGATA	Forward primer of *nhaA* gene (*Hind*III and *Nde*I restriction sites are inserted)	58.5
MRL05	CCCCGGATCCTTAGGCAATTCTTTTATTCTTAAAACTG	Reverse primer of *nhaA* gene (*BamH*I restriction site is inserted	59.9
MRL06	CCCCGGATCCTTAATGATGATGATGATGATGGGCAATTCTTTTATTCTTAAAACTG	Reverse primer of *nhaA* gene (*BamH*I restriction site and His tag are inserted)	57.6
MRL07	GGGGCCATGGAAATGAATAATAAAATCAAAAGAAAAGTG	Forward primer of *nhaA* gene (*Nco*I restriction site is inserted)	56.5
MRL08	CCCCGGATCCTTAATGATGATGATGATGATGATGATGATGATGATGATGGGCAATTCTTTTATTCTTAAAACTG	Reverse primer of *nhaA* gene (*BamH*I restriction site and 12 His tag are inserted)	57.6
MRL09	GGGGAAGCTTTGTTTGCTGTAAAAAATTGGATG	Forward primer of *nhaA* promoter gene (*Hind*III restriction site is inserted)	60.9
MRL10	GGGGCCATGGAAATGGGAGCCTTAATTGCAATTTTTC	Forward primer of *cpa1* gene (*Nco*I restriction site is inserted)	60.9
MRL11	CCCCGGATCCTTAATGATGATGATGATGATGATGATGATGATGATGATGTCTTAAAATATCGCTTTCAATATAATTTAA	Reverse primer of *cpa1* gene (*BamH*I restriction site of 12 His tag are inserted	58.5
MRL12	GGGGAAGCTTCTATTTCAATTATTAATTTCATTCATC	Forward primer of *cpa1* promoter gene (*Hind*III restriction site is inserted)	56
MRL13	GGGGCCATGGAAATGACAATTTTTTTAGCAGAATTAGTAG	Forward primer of *glpF* gene (*Nco*I restriction site is inserted)	56.6
MRL14	CCCCGGATCCTTAATGATGATGATGATGATGATGATGATGATGATGATGCATAATCATTGTATAACATACTGC	Reverse primer of *glpF* gene (*BamH*I restriction site and 12 His tag are inserted)	56.6
CH035	Property of Bio-Cleave Ltd	For detection of inserts in pMTL82151	57
CH036	Property of Bio-Cleave Ltd	For detection of inserts in pMTL82151	57
MH202	Property of Bio-Cleave Ltd	For detection of inserts (spacer) in the S3V9 plasmid	57
MH203	Property of Bio-Cleave Ltd	For detection of inserts (spacer) in the S3V9 plasmid	57

2.2.5 DNA electrophoresis and DNA purification

Agarose gels were prepared by dissolving 1 % w/v (final concentration) of agarose (Sigma-Aldrich, USA) in 1 X TAE buffer (40 mM Tris Acetate, 2 mM EDTA, pH 8.5). GelRed (Thermo Fisher Scientific, USA) was then added to a final 10 µg/mL concentration. Samples were loaded with 1 X purple loading dye (New England Biolabs Inc, USA). Gels were run in 1 X TAE buffer at 120 V for ~ 35 min, and DNA was then visualised using a UV photo documentation system. DNA fragments were purified from agarose gels using the Monarch® DNA gel extraction kit (New England Biolabs Inc, USA).

2.2.6 Double digestion

pTZ19R/U plasmids containing *cpa1* and *nhaA* with or without His-tag with a *Hind*III and *BamH*I restriction sites were double digested with the following reaction (Table 2.4), final volume of 20 µL.

Table 2.4 Double digestion reaction
The reagents for the double digestion are shown

2 µL of DNA or vector
1 x Buffer 3.1
0.5 U Restriction enzyme *Hind*III (New England Biolabs Inc. USA)
0.5 U Restriction enzyme *BamH*I (New England Biolabs Inc. USA)
Final volume 20 µL with distilled water

The vector and insert were digested at 37 ˚C for 1 h with *Hind*III and *BamH*I restriction enzyme in a reaction volume of 20 µL. Single digestion with only one enzyme was also performed. Double and single digested vectors and inserts were gel extracted and gel purified.

pET15b plasmids with 12 His-tag were double digested with *Nco*I and *BamH*I using the same reaction in Table 2.4 with the respective enzymes.

2.2.7 Ligation of inserts

The ligation reaction in Table 2.5 was incubated at room temperature for 10 min or overnight at 10 ˚C and heat-inactivated at 65 ˚C for 10 min.

Table 2.5 Ligation reaction
The reagents for the ligation reaction and their concentrations used are given.

1x T4 DNA ligase buffer
1 µL Vector DNA
3 µL Insert
0.4 U T4 DNA Ligase (NEB)
Final volume 20 µL with distilled water

2.2.8 *E. coli* transformation

50 µL of competent cells (DH5α, DH10b or BL21) stored at -80 ˚C were thawed on ice. 1 µL of plasmid was added, and cells were incubated on ice for 30 min. Heat shock was performed at 42 ˚C for 90 s, and cells were left on ice immediately for 2 min. 250 µL of LB media were added. Samples were incubated at 37 ˚C, 180 rpm for 1 hr and plated onto LB agar plates containing ampicillin at 100 µg/mL. Plates were incubated overnight at 37 ˚C.

2.2.9 Protein expression

E. coli BL21 cultures transformed with GlpF in pET17b or pET15b were inoculated into 1 L of LB/amp to a starting OD_{600} of approximately 0.05 from an overnight starter culture and incubated in an aerobic incubator at 37 ˚C, 180 rpm until an OD_{600} of ~0.5 was reached (approximately 2 h). 1 mM IPTG was added, and incubation was continued for 4 h. After the 4 h incubation, the total volume was centrifuged at 14,000 x *g* for 10 min, and the supernatant was discarded. Pellets were stored at -20 ˚C.

2.2.10 French press cell membrane extraction

Pellets from protein extraction were resuspended in 20 mL of buffer A (50 mM Hepes, pH 7.5, 500mM NaCl, 10 % glycerol, one tablet of protease inhibitor [cOmplete^{TM}(Merck KgaA, Germany)]). The sample was poured into the French pressure cell, and the cell was brought to the desired pressure of ~15 000 psi. Maintaining the pressure, the outlet flow rate was adjusted to about one drop every second. The cell lysate was collected in a 50 mL tube and kept on ice. The procedure was repeated thrice. Cell debris was removed by centrifugation at 4 ˚C 4,000 x *g* for 20 min, and then the pellet was harvested by ultracentrifugation at 4 ˚C 30,000 x *g* for 20 min using a Ti70 rotor. The pellet was resuspended in Buffer A at an 80 mg/mL wet pellet weight concentration.

2.2.11 Protein purification

In order to solubilise the membrane proteins, one polymer (styrene-maleic acid SMA) and two detergents (n-Dodecyl-β-maltoside DDM and Octyl-β-glucoside OG) were used to see which one was the most efficient.

2.2.11.1 Styrene-maleic acid (SMA)

The SMA2000 was prepared in-house according to a published protocol (Rothnie, 2016). In brief, styrene-maleic anhydride co-polymer 2000 (2:1 (Cray Valley, France)) was dissolved in a concentration of 1 g/mL in 1 M NaOH and three times autoclaved at 120 °C. Subsequently, the polymer was precipitated using concentrated HCl, washed with water and dissolved in 0.6 M NaOH for 3 h. After adjusting the pH to 8.0, the polymer was stored at -20 °C overnight before freeze-drying to obtain the powder.

The resuspended membrane preparation was diluted to a 40 mg/mL concentration of total membrane mass in buffer 1 (20 mM Tris, 150 mM NaCl, pH 8). 2.5 % (w/v) of SMA 2000 co-polymer powder was added to the membrane preparation and incubated at room temperature for 1 h. It was then centrifuged at 30 000 x *g* for 20 min at 4 ˚C. The supernatant containing the solubilised protein was harvested. Solubilisation efficiency was measured by resuspending the pellet in buffer 1 and running a Western blot (Section 2.2.13). The solubilised protein was mixed with Ni^{2+}NTA resin (pre-washed in buffer 1) at a ratio of 100 µL resin/mL solubilised protein and mixed gently overnight at 4 ˚C. The sample was poured into an empty gravity-flow column, and the resin was washed three times with 10 bed volumes (bv) of buffer 1 supplemented with 20 mM imidazole. Protein was eluted sequentially ten times with ½ bv of buffer 1 supplemented with 200 mM imidazole. Samples from each step were run on two SDS-PAGE (Section 2.2.12); one was stained with Coomassie, and the other was used for Western blot (Section 2.2.13).

2.2.11.2 n-Dodecyl-β-Maltoside DDM

The membrane preparation was resuspended to 40 mg/mL in buffer 1 plus 0.083 % (w/v) of DDM (ThermoFisher Scientifics, USA) powder and incubated at room temperature for 1 h. It was then centrifuged at 30 000 x *g* for 20 min at 4 ˚C. The supernatant containing the solubilised protein was harvested. Solubilisation efficiency was measured by resuspending the pellet in buffer 1 and running a Western blot (Section 2.2.13). The solubilised protein was mixed with Ni^{2+}NTA resin (pre-washed in buffer 1 plus 0.083 % (w/v) of DDM) at a ratio of 100 µL resin/mL of solubilised protein and mixed gently overnight at 4 ˚C. The sample was poured into an empty gravity-flow column, and the resin was washed three times with 10 bed volumes (bv) of buffer 1 supplemented with 20 mM imidazole plus 0.083 % (w/v) DDM. Protein was eluted sequentially ten times with ½ bv of buffer 1 supplemented with 200 mM imidazole plus 0.083 % (w/v) DDM. Samples from each step were run on two SDS-PAGE (Section 2.2.12); one was stained with Coomassie, and the other was used for Western blot (Section 2.2.13).

2.2.11.3. Octyl-β-Glucoside OG

This protocol was based on Borgnia and Agre (2000). The membrane preparation was resuspended to 40 mg/mL in solubilisation buffer (3 % (w/v) OG (ThermoFisher Scientific, USA) in 100 mM K_2HPO_4, 10 % glycerol, 5 mM 2-mercaptoethanol, 200 mM NaCl, pH 8) and incubated on ice for 1 h. Then it was centrifuged at 140 000 x *g* for 45 min at 4 ˚C. The solubilised protein was mixed with Ni^{2+}NTA resin (in solubilisation buffer) and incubated with agitation at 4 ˚C overnight. The sample was poured into an empty gravity-flow column, and the resin was washed with 10 bed volumes (bv) of wash buffer (solubilisation buffer with 100 mM imidazole and pH of 7) for 1 h on ice. Protein was eluted by incubation in 5 bv of elution buffer (solubilisation buffer with 1 M imidazole and pH 7) for 1 h on ice. Samples from each step were run on two SDS-PAGE (Section 2.2.12); one was stained with Coomassie, and the other was used for Western blot (Section 2.2.13).

2.2.12 SDS-PAGE

The solutions needed for SDS-PAGE are described in Table 2.6.

Table 2.6 SDS-PAGE solutions
The solutions with their components needed for SDS-PAGE gels are given.

Solution	Components
10 % SDS-PAGE	10 %, 375 mM Tris pH 8.7, 1 % SDS, 6.68 mM TEMED, 0.32 % APS
5 % stacking gel	5% 40:1 acrylamide: bis acylamide Protogel premix, 125 mM Tris pH 6.9, 1 % SDS, 6.68 mM TEMED, 0.8% APS
Coomassie Brilliant Blue R-250	40 % Methanol, 10 % glacial Acetic acid, 0.1 % Brilliant Blue R 250
6 x SDS loading dye	30 % 2-mercaptoethanol, 12 % SDS, 10 % glycerol, 0.1 % bromophenol blue, 440 mM Tris pH 6.8
1 x SDS running buffer	250 mM Tris base, 1.92 M glycine, 1 % SDS
Destain solution	40 % Methanol, 10 % glacial Acetic acid

20 µL of protein samples were mixed in 1 X (final concentration) SDS loading dye. The SDS-PAGE was done in duplicate. According to the manufacturer's instructions, two 10 % SDS-PAGE resolving gels with 5 % stacker were made using the Bio-Rad mini protean tetra apparatus (Bio-Rad, USA). 5 µL PageRuler (Invitrogen, USA) were used as a ladder, 10 µL of the sample were loaded into each well, and the gel was run in 1 X SDS running buffer at 140 V per gel for ~ 1 h. Once the dye front had reached the bottom of the gel, one gel was stained in Coomassie Brilliant Blue R-250, microwaved for 20 s and gently agitated for ~ 10 min at room temperature. Then it was destained with destain solution, microwaved for 20 s, and gently agitated overnight. The other gel was used for Western blot.

2.2.13 Western blot

The solutions used for the Western blot are described in Table 2.7

Table 2.7 Western blot solutions
The solutions with their components for Western blots are given

Solution	Components
TBS-T	20 mM Tris-Cl pH 7.5, 500 mM NaCl, 0.05 % Tween-20.
10 x Transfer Buffer	2.5 M Tris base, 19.2 M glycine
Blocking buffer	1 x Blocking reagent buffer (QIAGEN), 1 % blocking reagent (QIAGEN, Germany) diluted in TBS-T

The SDS-PAGE gel was equilibrated in a cold (~ 4 ˚C) 1 x Transfer buffer for ~ 5 min. Amersham Hybond-N+ membrane (GE Healthcare, USA) was cut into a suitable size for blotting. The membrane was wetted in deionised water for 10 s and equilibrated in cold 1 x Transfer buffer for 5 min. The blotting cassette for transfer was assembled. Transfer conditions were 1 h 100 V run at 4 ˚C. After the transfer of proteins, the remaining protein-free sites on the membrane were blocked with 1 h incubation in 20 mL blocking buffer at room temperature. The membrane was washed 3 times for 5 min each in TBS-T buffer at room temperature. Then it was incubated in Anti-His HRP conjugate solution (QIAGEN, Germany) (1/1500 dilution) or 6x-His Tag monoclonal antibody (Invitrogen, USA) (1/1000 dilution) in blocking buffer overnight at 4 ˚C or 1 h at room temperature. The membrane was washed 3 times for 5 min each in TBS-T buffer at room temperature. If the Anti-His HRP conjugate was used, the chemiluminescent detection reaction was performed. If not, it was incubated with rabbit anti-mouse IgG secondary antibody HRP (Invitrogen, USA) solution (1/2000 dilution) in blocking buffer for 1 h at room temperature. The membrane was washed 3 times for 5 min each in TSB-T buffer at room temperature. Horseradish peroxidase (HRP) catalyses the oxidation of substrates by hydrogen peroxide, resulting in a coloured or fluorescent product and the release of light (425 nm) as a by-product of the reaction, which can be captured with camera imaging devices that detect chemiluminescence (Alegria-Schaffer *et al.*, 2009). In order to reveal the membrane, SuperSignal West Atto Substrate (Thermo Fisher, USA) was used. A substrate working solution was prepared by mixing equal parts of Detection Reagents 1 and 2 (750 µL each) to incubate the blot for 1 min at room temperature. After this, the excess liquid was removed, and the membrane was placed in the G:box (Syngene, UK) to reveal the detection reagent.

2.3 Results chapter III – Biocleave *C. saccharoperbutylacetonicum* N1-4HMT characterisation

2.3.1 *C. saccharoperbutylacetonicum* N1-4HMT transformation

C. saccharoperbutylacetonicum N1-4HMT were transformed using electroporation. All steps except centrifugation were carried out in an anaerobic cabinet (Don Whitley). Buffers are listed in Table 2.8. 60 mL CGM with 5 % glucose cultures were grown until the mid-exponential phase. The media was equilibrated overnight in the anaerobic cabinet into an Erlenmeyer flask. Cells were pelleted at 4,000 x *g* for 10 min at 4 ˚C and washed in Electroporation buffer with salt (EPB_S) buffer (Table 2.8), centrifuged again and resuspended in 2 mL of Electroporation buffer without salt (EPB_NS) buffer (Table 2.8). 1000 – 1500 ng DNA were added to 200 µL cells in 2 mm electroporation cuvettes (Cell Projects, UK) and incubated on ice for 5 min before pulsing at 1.5 kV (BioRad, USA). Cells were plated on CGM agar containing 5 % glucose and relevant antibiotic.

Table 2.8 Clostridial transformation buffers

Buffers with their components used for Clostridial transformations

Buffer	**Components**
EPB_S	300 mM sucrose, 0.6 mM Na_2HPO_4, 4.4 mM NaH_2PO_4, 10 mM $MgCl_2$
EPB_NS	300 mM sucrose, 0.6 mM Na_2HPO_4, 4.4 mM NaH_2PO_4

2.3.2 Bottle screening

C. saccharoperbutylacetonicum N1-4HMT were grown anaerobically overnight in 30 mL of RCM at 32 ˚C to an $OD_{600} > 2.0$ and pH 5.0 -5.5. From this, a 15 % inoculum was used to inoculate 25 mL TYIR. The subcultures were grown anaerobically at 32 ˚C until they reached an OD_{600} of 1.5 ± 0.2 (3 – 5 h). A triplicate of 54 mL TYIR media were inoculated with 6 mL of the subculture. The bottles were incubated anaerobically at 32 ˚C for 48 h and sampled at 0, 6, 24 and 48 h. pH and OD_{600} were measured at each sample point. 4 mL were taken out for sampling, centrifuged at 4 000 x *g* for 20 min, filtered through a 0.2 µm cellulose filter and stored at – 20 ˚C. HPLC measured sugar, acids and solvent concentrations.

2.3.3 HPLC quantification of sugars and solvents

Bottle screening supernatants were removed from the – 20 ˚C freezer and allowed to thaw at room temperature. A 4 x dilution was prepared by adding 200 µL of the sample to 600 µL of HPLC grade water. Samples were analysed using the Bio-Rad Aminex HPX-87H column (Bio-Rad, USA). The running conditions are described in Table 2.9

Table 2.9 HPLC conditions for solvents acids and sugar

The conditions and settings used for the analysis by HPLC are given

Condition	Setting
Elution type	Isocratic elution
Flow rate	0.6 mL/min
Oven temperature	35 ˚C
Run time	Maximum 30 min
Injection Volume	20 µL
Detector type and detection temperature	Refractive index at 35 ˚ C
Mobile phase	5 mM H_2SO_4
Seal wash	10 % methanol
Injector wash	HPLC grade water

2.3.4 Statistical Analysis

Because 2 factors were being tested (OD_{600} or g/L, time), each with 5 groups (WT, +NhaA, +Cpa1, ΔGplF, ΔPssA) data in this factorial design were compared by performing a 2-way analysis of variance (ANOVA) using GraphPad Prism 8.1.0 (USA) statistical software. All values in bar graphs are expressed as means +/- standard error of the mean (SEM) of n = 9. Differences were considered statistically significant at $P \leq 0.05$. The * indicates a significant difference between the samples at the indicated time points. * = <0.05, **** = <0.0005.

3. Results I – A miniaturised iodine value assay for quantifying unsaturated fatty acids

3.1 Introduction

Product toxicity is one of the significant challenges in the development of bioprocesses based on the use of platform host organisms (Warnecke and Gill, 2005). Understanding the membrane stress response can improve product yields (Pei *et al.*, 2019). Zhu and Yang (2003) and Wang *et al.* (2013) observed a Clostridial mechanism to protect itself from toxic metabolites by altering the concentration and length of saturated fatty acids (SFA) in its membrane.

Membrane fluidity is mainly determined by the degree of unsaturation of fatty acids (FAs) (Hagve, 1988). It has been observed in many cells that a higher proportion of unsaturated FAs (UFA) and more short-chain FAs within the cell membrane correlates with adaptation to low temperature, where membrane viscosity would otherwise be increased. In solvent-producing cells, it has been observed that levels of saturated and cyclopropane FAs increased, whereas UFAs decreased within the cell membrane (Zhao *et al.*, 2018). The proportion of UFAs also decreases with higher pH. Interpretation of this adaptation is not clear; however, the decrease in unsaturation and the length of the chains seem to support the balance between stabilisation and flexibility of the cell membrane to prevent or enable protons or other active substances to enter the cell (Bajerski *et al.*, 2017, Wu *et al.*, 2012). Studying the mechanisms of membrane adaptation, integrity and fluidity are essential to gain new insights into maintaining cell viability and metabolic activity and providing future engineering strategies. Therefore, various complementary approaches are required to quantify fundamental membrane characteristics, such as the percentage of unsaturation in the lipid membrane.

With the general aim to increase the production of organic acids in solventogenic Clostridia and the understanding that membrane stress response can improve product yields, this first chapter aimed to implement a method to detect the unsaturation of the lipid membranes in the laboratory. The chosen method would then be applied at different time points during the fermentation of *C. saccharoperbutylacetonicum* N1-4HMT. This will provide information about the toxicity mechanisms of the organic acids and the protective changes in the cell membrane that occur during fermentation to facilitate engineering microorganisms for improved toxic tolerance.

The chapter showcases the selection and standardisation of a method that can be used to identify changes in the unsaturation of the lipids within cell membranes at different stages of fermentation. First, the existing methods for quantifying and identifying the unsaturation of lipids with project requirements and laboratory capacity were evaluated. The iodine value method was ultimately selected, but modifications were needed for laboratory-scale samples of cell membranes, namely, the ability to work with mg of lipids due to the availability of material. Therefore, this chapter describes the steps and the modifications made to the methodology. Finally, the adapted and miniaturised IV assay was tested in different cell membranes from established systems to demonstrate its effectiveness on biological samples.**Results**

3.2.1 Comparisons of methods for the quantification and identification of unsaturation of lipids.

Several methods exist for the quantification and identification of the unsaturation of lipids. Thin-layer chromatography (TLC) is a chromatographic method that can detect the degree of saturation and unsaturation in the lipid (Singh *et al.*, 2018). However, it has limitations in some applications; TLC cannot be used for precise quantification, and lipid oxidation occurs when exposed to atmospheric oxygen (Shantha and Napolitano, 1998). Gas and liquid chromatography (GC and LC, respectively) can give excellent sensitivity, molecular specificity and precision when coupled to mass spectrometry (MS). Mass spectrometry is currently the most reliable method for quantification of the level of saturation/unsaturation (Wolrab *et al.*, 2022). The previous methods are classified as “conventional” because they rely on solvent extraction and chromatographic determination (Patel *et al.*, 2019). Problems with these methods are that they can involve multiple steps, are time-consuming, and may not be suitable for high-throughput analyses. For non-conventional methods, there are high-throughput platforms available such as the spectroscopic methods: Infrared spectroscopy, nuclear magnetic resonance spectroscopy (NMR), Raman spectroscopy, fluorescence spectroscopy and dielectric spectrometry (Ginkel *et al.*, 1989, Bonzom *et al.*, 1999, Czamara *et al.*, 2015). These methods can estimate unsaturation in the lipids or, when using the proper standards, quantify the saturated and UFAs with a high correlation to LC-MS approaches. The main disadvantages of these approaches, including mass spectrometry, are the need for sophisticated equipment, specialised software and trained personnel. In addition, the volume of data can be large and complex, making data analysis challenging (Blanksby and Mitchell, 2010, Patel *et al.*, 2019). Finally, in the non-conventional category, colourimetric methods exist, but the Iodine Value assay is the only method to measure the degree of unsaturation (Pardeshi, 2020a). Table 3.1 summarises the techniques, their advantages and disadvantages and their ability to detect unsaturation.

Table 3.1 Advantages and disadvantages of techniques for detection and quantification of lipids and their ability to detect unsaturation

The methods used for detecting and quantifying lipids, their advantages, disadvantages, and if they measure unsaturations are given.

Method	Description	Advantages	Disadvantages	Measures unsaturation?
Conventional Methods				
Gravimetric method (Patel *et al.*, 2019)	Involves extraction of total lipid, centrifugation, drying and weighing	The most common way to quantify lipids can be considered a standard when developing new screening strategies. The easiest method to quantify total lipids in a short time, with no need for special equipment or training	It requires cell disruption and extraction. Not suitable for small amounts of sample. The accuracy and precision are dependent on the accuracy and precision of the weight and volume measurements. For total lipids only.	No
Chromatographic analysis				
Thin-layer chromatography (TLC) (Shantha and Napolitano, 1998)	It involves passing a mixture of the molecules to be separated across a plate containing a matrix that will selectively retarding the flow of the molecules. Based on the difference in the affinity of a component or components towards a stationary and a mobile phase.	Low cost for quick analysis of samples, easy handling and reproducibility, simple, reliable, and inexpensive.	It has limitations in preparative applications, quantification, and lipid oxidation during exposure to atmospheric oxygen.	Normal-phase TLC containing silicic acid is often used to separate TAGs with varying degrees of saturated and unsaturated lipids.
Gas and liquid chromatography coupled to mass spectrometry (Blanksby and Mitchell, 2010)	Mass spectrometry is based on the differential propensity of each lipid class to acquire either positive or negative charges under the source high voltage. In essence, lipid classes can be separated through their endogenous electric potential.	Excellent sensitivity, molecular specificity and precision, minimal sample preparation, reduced analysis time, and low sample size. Well-established procedures for lipid profiling	Sample pre-treatment makes these methods destructive, time-consuming, environmentally unfriendly and does not allow real-time lipid monitoring. Data analysis can be quite challenging.	The level of saturation/unsaturation can be quantified most reliably by GC FD/MS.
Non-conventional quantification of lipids, including high-throughput platforms				
Spectroscopic methods				
Infrared spectroscopy (Lewis and McElhaney, 2013)	It is an analytical technique that measures the interaction of infrared radiation with bonds in a chemical	A fast and non-invasive lipid quantification method that can be applied to small	FTIR band shifts are species-specific; they cannot differentiate	Lipid unsaturation can be measured with this method.

	structure. Different functional groups absorb characteristic frequencies of IR radiation.	amounts of whole cells for real-time monitoring and screening cultures It identifies lipid classes based on standards and cellular content determination without disruption of cells.	different species when a mixture of cultures is used.	
Nuclear magnetic resonance spectroscopy (NMR) (Bonzom *et al.*, 1999)	The lipid content is determined by measuring the area under a peak in an NMR chemical shift spectra that corresponds to the lipid fraction.	Requires minimal sample preparation; does not require chromatographic separation or derivation of the sample. It can be applied to liquid or solid samples. Provides detailed structure-specific information, and chemical alterations such as oxidation are kept minimal. High speed of analysis and non-destructive, enabling recovery of the sample for further analysis.	The initial investment is high. It requires a high amount of sample, incapable of differentiating between neutral lipids and polar or membrane lipids. Need for internal standards for quantification; challenging to translate to an industrial setting.	The average levels of unsaturation can be obtained from the NMR spectra
Raman spectroscopy (Czamara *et al.*, 2015)	Based on the inelastic scattering of monochromatic light produced by a laser in the visible, NIR or UV range.	A rapid and non-destructive method for quantifying the degree of lipid unsaturation. Label-free, no sample preparation, *in vivo* analysis, real-time.	The strong Raman signal of fluorescent pigments and carotenoids, which overlaps with the lipid peaks, can interfere with the analysis. The Raman signal post-processing includes cosmic noise removal and fluorescence background subtraction. Limited application in living systems. Its main limitation is the capacity to probe only one Raman band simultaneously. It is focused mainly on determining the degree of unsaturation; quantification	It can quantify saturated and UFAs with a high correlation between LC-MS.

			requires calibration based on other methods, such as a gravimetric analysis, which is difficult to translate to an industrial setting.	
Fluorescence spectroscopy (Sankaran and Wohland, 2020)	Based on the characteristic fluctuations and time it takes for a molecule that an external force has perturbated to dissipate.	Simpler and faster and requires less material than conventional methods. In addition, sensitivity is higher, allowing for minor differences in lipid content to be distinguished. High-throughput potential	Fluorescence-based detection is highly species- and strain-specific, and these methods require optimisation before they can become reliable quantification tools. Analysis of fluorescent data is a challeging task. It only detects neutral lipids.	Can measure the fluidity of the membranes, which may be correlated to the saturated/unsaturated ratio of the FAs
Nile Red (Rumin *et al.*, 2015)	A lipophilic stain. In most polar solvents, Nile red will not fluoresce; however, when in a lipid-rich environment can be intensely fluorescent, with varying colours from deep red (for polar membrane lipid) to strong yellow-gold emission (for neutral lipid in intracellular storages).	The most used in literature. Easy and cheap to use and requires a smaller amount of equipment.	Photo-bleaching of the agent; poor penetration of cells, interference with chlorophyll autofluorescence and green fluorophores. Limited specificity to lipids.	Can estimate relative polyunsaturated FA content and unsaturation index.
BODIPY 505/515 (Rumin *et al.*, 2015)	A lipophilic bright green fluorescent dye which is used to label a wide range of lipids such as FAs, phospholipids, cholesterol, cholesteryl esters, and ceramides	More lipid-specific than Nile Red, has a narrower emission spectrum, higher sensitivity, and better reproducibility; it requires a lower concentration of solvent carrier	Interference with green fluorophores	Same as Nile red: Can estimate relative polyunsaturated FA content and unsaturation index.
AC-202 (Patel *et al.*, 2019)	Lipid-specific blue fluorescent molecule.	More sensitive than BODIPY; low-background signal	It is not used for quantification yet.	Not reported
Dielectric spectrometry (Lizhi *et al.*, 2008, Atmadjaja *et al.*, 2019)	Based on the different permittivity and conductivity of cell components and measuring the dielectric properties of a medium as a function of frequency	A rapid, non-invasive, and label-free method for determining the composition of lipids in a cell. Dielectric spectrometry requires simple instrumentation and	It is used for total lipid detection with no profiling.	It can predict the main FA components and degree of unsaturation.

		minimal sample preparation as cells are centrifuged and resuspended in a fresh medium. Potential for an automated process for biomonitoring		
Colourimetry				
Sulfo-phospho-vanillin (SPV) (Anschau *et al.*, 2017)	The SPV assay produces a distinct pink colour when reacting with lipids, and the intensity of the colour can be quantified by measuring the absorbance at 530 nm using spectrophotometric methods.	The colour formation is stable for hours. Fast, low detection limit. Solvent extraction can be omitted depending on the assay. Easy method, a small amount is needed is reproducible, less labour.	It requires the presence of double bonds or free hydroxyl groups (unsaturated lipids). Its accuracy depends on the standard. The reactivity of polyunsaturated FAs decreases with increasing unsaturation.	Useful for measuring total lipid content with unsaturated lipids.
TAG kit (Patel *et al.*, 2019)	Triglycerides are converted to free FAs and glycerol. Glycerol is then oxidised to generate a product which reacts with a probe to generate colour (spectrophotometry at λ= 570 nm) and fluorescence (Ex/Em = 535/587 nm).	Fast, automated process	Detects only TAGs; requires cell disruption and extraction.	No
Copper extraction (Patel *et al.*, 2019)	Based on the hydrolysis of lipids to FAs and subsequent extraction of their copper salts into chloroform. The amount of copper in the chloroform phase is then colourimetrically determined by adding diethyldithiocarbamate to develop a yellow-coloured product (measured by optical density at 440 nm).	It requires a small amount of culture (1-2 mL culture); adaptation to microcentrifuge format allows up to 30 samples in < 2 h. Fast, high detection limit.	It requires solvent addition and cell disruption. No structural information is gained from the assay.	No
Iodine Value (IV) (Odoom and Edusei, 2015)	Measures the degree of unsaturation of oil; the amount of iodine in grams taken up by 100 grams of the oil.	Cheap, quick, and easy to perform and interpret.	A significant number of oils have the same IV, susceptible to oxidisation. 1 g is the minimum amount required.	Measures the average degree of unsaturation of a lipid

The perfect technique for determining of unsaturation during fermentation has yet to exist. There is still a need for fast and accurate quantification of lipid unsaturation method. A method that allows rapid screening, reduced sample, an inexpensive approach, sensitivity to minor differences in lipid unsaturation, without many optimisations, with accessible equipment and the resources in the laboratory, would be ideal. Many laboratories do not have specialist equipment such as gas chromatography or mass spectrometry. It was therefore decided to adapt one of the techniques that do not require specialised equipment, which could save time for data analysis and represent an alternative or an improvement on the existing methods. The aim was to make a widely applicable, simple to apply assay that required no specialist equipment. For these reasons, the IV assay was chosen.

3.2.2 Iodine value (IV)

The IV is a long-established analytical index related to fats and oils, which measures the total unsaturation in the sample. IV has been used beyond analytical purposes related to various physical and chemical properties, such as a quality control method in hydrogenation and quality standardisation of industrial products such as biodiesel and plant oils (Knothe, 2002). The IV has been used as an American Oil Chemists' Society (AOCS) standard method (Kyriakidis and Katsiloulis, 2000). IV is defined as the number of grams of iodine absorbed by 100 grams of sample. Because iodine (or other halogens) bind to carbon-carbon double bonds, it can quantify the number of unsaturations in fats or oils (Knothe, 2002, Odoom and Edusei, 2015). Figure 3.1 illustrates the determination of the IV, which involves the reaction of double bonds in the sample with a halogenating reagent (Hanus or Wijs solution) over 30 min; the Wijs method is the most widely used (Pardeshi, 2020b). In this step, iodine from iodine monochloride (ICl) reacts with the unsaturated bonds to produce a dihalogenated single bond; one carbon has bound to one iodine. To determine the amount of iodine that has reacted, a potassium iodide solution (KI) is added to the reaction, which causes the unreacted ICl to form molecular iodine (I_2); the solution becomes dark brown. Then a titration is performed to calculate the iodine (I_2) that did not react. Next, the excess solution is mixed with sodium thiosulfate ($Na_2S_2O_3$). The remaining free iodine (I_2) reacts with the sodium, and the solution becomes colourless as the sodium iodide (NaI) forms.

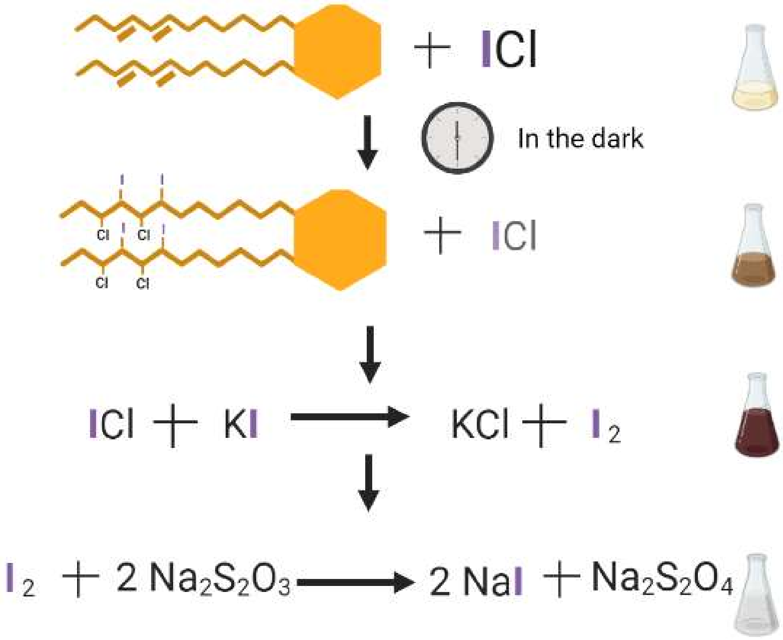


Figure 3.1 Chemical reactions of the Iodine Value method
The chemical reactions of the Iodine Value method are depicted in the image. The lipids are orange, and the reagents are written in their chemical formula. The flasks on the right indicate the approximate colour of the mixture. Created with biorender.com

The millilitres of $Na_2S_2O_3$ needed to convert a dark solution into a colourless solution were then used in the IV formula (Equation 3.1) to calculate the Iodine Value of the sample. The higher the iodine value, the greater the degree of unsaturation.

Equation 3.1 IV formula

$$IV = \frac{(B - S) * N * 126.9 * 100 * 10^{-3}}{W}$$

$$IV = \frac{(B - S) * N * 12.69}{W}$$

B = mL of $Na_2S_2O_3$ *used for the blank*
S = mL of $Na_2S_2O_3$ *used for the sample*
N = $Na_2S_2O_3$ *normality = 0.1*
126.9 = molecular weight of iodine
W = weight of the sample used.

For Equation 3.1, the blank (*B)* does not contain any lipid, meaning that no ICl will react and have the freest I_2. Therefore, it will require the largest volume of $Na_2S_2O_3$ to become colourless. For the lipid samples (*S*), the more unsaturations, the less free I_2, and the lower volume of $Na_2S_2O_3$ required. For *B-S,* the more unsaturated the lipid sample, the more significant the difference and the higher the IV. This amount is then multiplied by the normality of $Na_2S_2O_3$ (*N*). Although 2 mols of $Na_2S_2O_3$ react with

1 mol of I_2, 1 mol of thiosulfate $S_2O_3^{2-}$ reacts with 1 mol of iodide ions or iodine atoms (I^-) that have a molecular weight of 126.9. The definition of IV is for 100 grams; it needs to be multiplied by 100 and multiplied by 10^{-3} to convert the mL to L (vlab.amrita.edu, 2011).

A challenge of applying the IV method to cell membranes is the sample mass typically used. As it was established for bulk fats and oils, e.g. in the food and petrochemical industries, it is generally applied to gram-scale amounts of lipids. This presents a challenge when working with biological samples when the entire biomass available will likely be on the milligram scale, with lipids representing just a small fraction. However, the simple nature of the IV assay, lack of requirement for complex instrumentation, and linear range corresponding to biological samples make it an attractive option if the challenge of lipid mass can be solved. Therefore, this research aimed to miniaturise the IV method such that it can be applied to a range of biological samples.

The IV assay requires a titration that turns a dark solution into a colourless solution. The method approved by the AOCS requires enough sample and solution so that the eye can detect the change of colour. To start, the AOCS IV method was applied using ~0.5 g of sunflower oil to understand the assay; the exact amount weighted was 0.47 g

For the sunflower oil, 15 mL of sodium thiosulfate were needed to turn the solution colourless, whereas, for the blank, which was just chloroform, 58 mL of $Na_2S_2O_3$ were needed. Equation 3.1 was applied:

$$IV = \frac{(58 - 15)x\ 12.69\ x\ 0.1}{0.47} = 116.1$$

Sunflower oil has a reported IV between 81 and 141 (Tiefenbacher, 2017). The value of 116.1 obtained corresponds to the IV that has been reported. Once tested, the mass was reduced to 0.05 and 0.01 g to test the minimum amount that could be used. The exact amounts weighted were 0.043g and 0.013 g, and the IV was determined (Figure 3.2). Each sample was performed in duplicates. The amount of sunflower oil was ultimately reduced to 0.001 g, demonstrating highly variable results, with the determined value often being out of range (data not shown). However, these data indicated that optimisation of the method to lower the amounts of lipid was required.

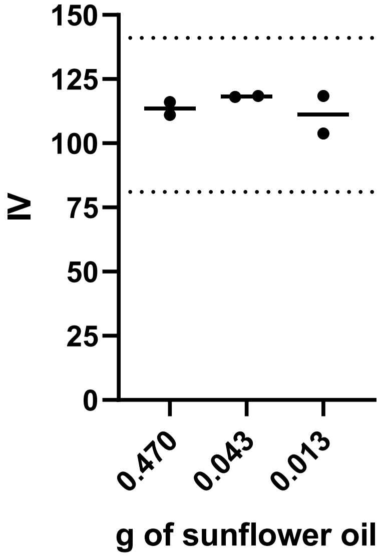


Figure 3.2 Iodine Value of sunflower oil.
The Iodine Values of 0.470, 0.043 and 0.013 g of sunflower oil were obtained following the AOCS's method and Equation 3.1 to calculate the IV. The experiment was performed in duplicate; the dot represents each experiment. The line represents the mean of the values.

3.2.3 Optimisation of the IV method. Scale-down.

First, it was determined if the loss of colour was detectable using a spectrophotometer. Because the lipids are diluted in chloroform, glass cuvettes of 2 mL were used. Since the solution has a dark blue appearance, 480 nm was chosen. The titration was performed manually by adding 100 µL of 0.1 N $Na_2S_2O_3$, shaking and reading the absorbance. In the beginning, the sample's absorbance was saturated; therefore, no accurate absorbance was measured. The loss of colour observed by the eye was obtained at an absorbance ≤ 0.3. The mL used to get an absorbance below ≤ 0.3 were used for the IV formula. In Figure 3.3, the IVs from Figure 3.1 were repeated to compare the results obtained with 0.008 g of sunflower oil and measure the absorbance to detect the loss of colour. It was possible to calculate the IV with lower concentrations of the sample and use a spectrophotometer, which means that it is possible to adapt the method and make it automatic if a spectrophotometer is used instead of human eyes.

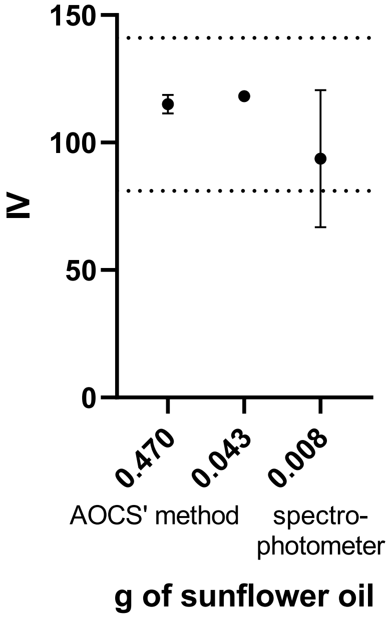


Figure 3.3 Comparison of the iodine values of different amounts of sunflower oil using the AOCS method and a spectrophotometer.
0.008 gr of sunflower oil were used for the IV assay. The loss of colour was obtained at an absorbance of 480 nm ≤0.3. (n = 3, error bars represent SEM). 0.470 g and 0.043 from Figure 3.2 are also represented there for reference.

Therefore, it was decided to miniaturise the IV assay and perform it in a multi-well plate reader to dispense liquid, shake the plate, read the absorbance and repeat, reducing the time and performing multiple samples simultaneously.

The spectrophotometer available in the lab was the BMG Labtech FluoStar multi-well plate reader, which allows running a spectrum between 350 and 650 nm. It was decided to run the spectrum with a solution of 0.1 M ICl mixed with 2 % starch and 15 % KI plus different volumes of 0.1 N $Na_2S_2O_3$ to confirm if 480 nm was the optimal wavelength and to obtain the OD that would indicate the solution has become colourless. 400, 450 and 480 nm were the wavelengths that significantly differed between the very dark solution and the colourless condition. The filters for these wavelengths are the most popular (Scientific Laboratory Supplies, 2022, Scientific, 2022), so they will be available in most laboratories. The experiment was repeated, with the plate reader, dispensing 10, 5, 3 and 1 µL of 0.1 N $Na_2S_2O_3$. Dispensing 1 or 3 µL made the process too slow to be of practical value. Some wells showed precipitation of yellow crystals, preventing obtaining an OD below 2. Iodide salts are soluble unless they contain silver, lead and mercury. Contamination with these salts would be sporadic. A possible explanation for the observed precipitation with 1 and 3 µl can be because the volume is too low. Therefore, the force that these volumes are dispensed with, and the shaking, might not be enough to properly mix all the reagents generated in some areas of the well and excess iodide concentration generating the precipitate. Dispensing 10 µL would sometimes be too much $Na_2S_2O_3$ at once; there will be an excess of the salts, turning the mixture milky white and preventing obtaining an OD below 2. From these initial experiments, it was decided to dispense 5 µL of $Na_2S_2O_3$ as it made the process fast and reduced the risk of precipitations by adding too much and reducing the assay resolution. Also, it was observed that the wavelength at 450 nm gave the most sensitivity to detect minor colour changes and gave an IV in the range of what has been reported. An OD of 1.99 or less was indicative that the solution was colourless.

Figure 3.4 shows how the mixture becomes colourless. The first well on the left has no $Na_2S_2O_3$, and the following wells represent the incremental addition of 15 µL 0.1 N sodium thiosulphate. 15 µL were chosen just for the picture to speed the reaction and fit all the colour changes in one lane. The plate reader's absorbance at 450 nm is indicated at the top of each well.

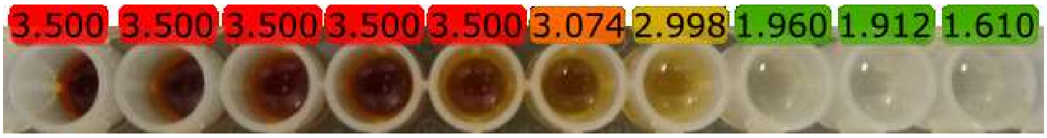


Figure 3.4 Titration of I_2 with different concentrations of 15 µL of 0.1 N $Na_2S_2O_3$
All wells contain 0.1 M ICl mixed with 2 % starch and 15 % KI. The first one on the left has no $Na_2S_2O_3$, the next one to the right has 15 µL of 0.1 N $Na_2S_2O_3$, the next one has 15 µL more and so on until the last one has a total of 135 µL. The absorbance was measured at 450 nm using the BMG plate reader and is indicated at the top of each well.

Once the optimal wavelength and OD were determined, sunflower oil was used as an established and readily available substrate to scale down the method and confirm that the values obtained were consistent with those reported in the literature. The miniaturised IV assay with the conditions mentioned above was performed on different masses of sunflower oil (from 1 to 0.1 mg) (Figure 3.5) to define the minimal mass required. Although the mean IV for all of the masses of lipids tested were in the range of the reported values, the variability of the assay increased substantially at lower masses.

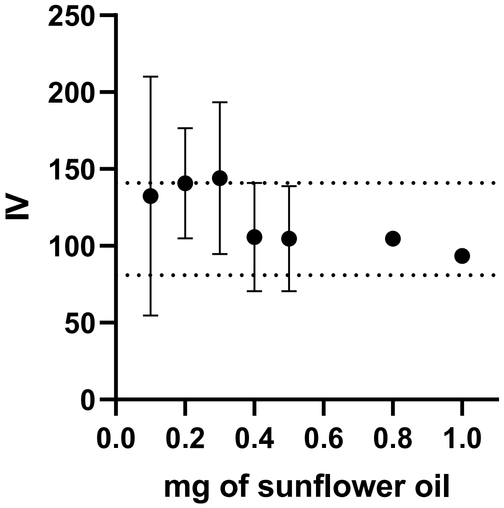


Figure 3.5 Detection limit of the IV assay.
Different amounts of sunflower oil were analysed in the miniaturised IV assay. The dotted lines represent the minimum and maximum values of the reported IV for sunflower oil (n = 3, error bars correspond to the SEM; for 0.8 and 1 mg, the SEM are too small to be seen, these are ± 4.8 and 3.0 respectively).

To understand better how the assay works and see if some modifications gave better results with less sample the same amounts of sunflower oil were tested again under different conditions (Figure 3.6):

- Eliminating the addition of starch.
- Two different concentrations of $Na_2S_2O_3$ were tested, 0.05 and 0.03 N.

Starch is an indicator; it binds to free I_2 and gives a solid dark blue colour. Without starch, the assay colour is not intense, which can reduce the number of mL of sodium thiosulfate needed, giving a lower IV that did not correlate with what has been reported (Figure 3.6, grey closed triangle). Diluting the $Na_2S_2O_3$ concentration could also turn the solution colourless. However, the IV obtained by adjusting the formula's normality did not correlate with the reported (open grey circle and triangle). Further adjustment in the formula was needed since the molecular weight of the iodine is used for 1 mol of thiosulfate. The addition of water could change or affect this relationship. However, this was not further explored. The condition that gave the more in-range value IV (between 81 and 141, dotted lines) was previously chosen for 1 and 0.8 mg (Figure 3.6, closed black circles). Therefore, adding 2 % (w/v) starch and the $Na_2S_2O_3$ solution at 0.1 N remained. Due to the small amount of starch used, it was decided to add it before the KI. The volume of KI would help to mix the starch better with the sample.

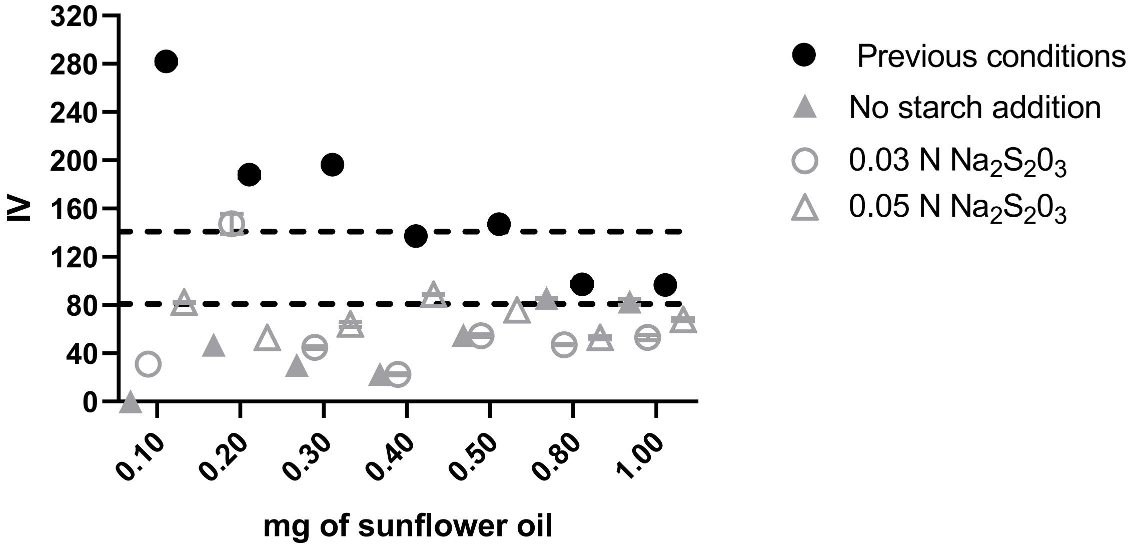


Figure 3.6. **Performance of the IV assay under different conditions**
Different amounts of sunflower oil without starch, with different normalities of $Na_2S_2O_3$ or with the previous condition were tested for IV assay. The dotted lines represent the minimum and maximum values of the reported sunflower oil IV (n=3, error bars correspond to the SEM).

Based on this and the previous results, 1 mg of lipid sample was chosen as the ideal amount required. It gave an IV that corresponds to what has been reported, had the lowest standard error and is an amount accessible for a range of biological samples.

3.2.4 Optimisation of the IV. Detection of unsaturation in lipid standards.

To mimic a cell membrane, it was decided to test the miniaturised IV assay with different concentrations of defined lipid standards. 18:1 (*cis*) DOPC (1,2-dioleoyl-sn-glycero-3-phosphocholine) and 14:0 DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine) were chosen (Figure 3.7). DOPC contains one double bond per FA chain, and DMPC is a saturated lipid. These two lipid standards were chosen because they are common FAs in animal cells, including humans (Hashimoto and Hossain, 2018), and they are present in *C. saccharoperbutylacetonicum* membrane (Dürre, 2005).

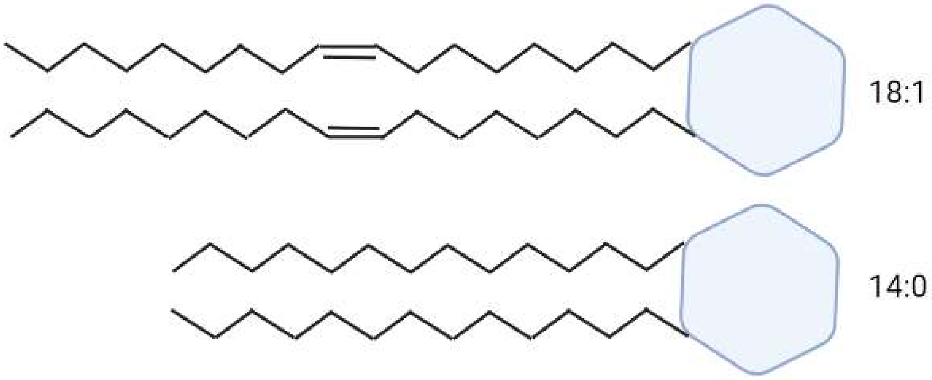


Figure 3.7 Schematic representation of 18:1 (cis) DOPC and 14:0 DMPC
The 18:1 (cis) DOPC (1,2-dioleoyl-sn-glycero-3-phosphocholine) and 14:0 DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine) are represented in the image. The head group is represented as the hexagon, and the carbon bonds are represented as black lines. Created with biorender.com.

Fifteen different mixtures were prepared at 10 mg of lipid dissolved in 1 mL of chloroform, and 100 µL of this was added to microplate wells to proceed with the miniaturised IV assay (Table 3.2, Figure 3.8). The percentage of unsaturation in each mixture was calculated as a proportion of total C-C and C=C bonds in the acyl chains based on the molar ratio. The average double bond per lipid molecule was also calculated. Equation 3.2 (Pomeranz and Meloan, 2002) calculates the theoretical IV when the molecular weight and number of double bonds are known.

Equation 3.2 Theoretical IV formula

$$IV = \frac{2 * 126.92 * no.of\ double\ bonds * 100}{molecular\ weight}$$

The equation is based on 1 mol of double bond reacting with 1 mol of I_2 (I_2 = 2 * 126.92 g/mol) * the mols of the double bond in the sample * 100 because, by definition, IV is the amount of iodine that reacts with a 100 g of sample. This is then divided by the molecular weight of the sample to get the actual mols of double bonds in the sample. The theoretical IV for each mixture was calculated (Table 3.2, Figure 3.8). Chloroform was used as a blank. After analysing these results, it was decided to use 100 % 14:0 DMPC as a blank because it is a lipid without unsaturations. Therefore, it will reduce the noise seen in the sample since it contains all the sample components apart from the unsaturations.

Table 3.2 Analysis of 18:1 and 14:0 lipid mixture using the miniaturised IV assay
The table includes the percentage of w/v of each lipid mixture, the % of UFAs in molar ratio, the average double bonds per FA, the calculated theoretical IV and the experimentally obtained IV ± SEM of each lipid mixture.

Lipid mixture (w/v)	% UFAs in molar ratio	Average double bonds per FA	Theoretical IV	Experimental IV
100 % 18:1 (*cis*) PC	5.88	1.00	32.29	30.08 ± 2.60
95.24 % 18:1 (*cis*) PC + 4.76 % 14:0 PC	5.56	0.95	30.96	20.27 ± 10.43
90.9 % 18:1 (*cis*) PC + 9.1 14:0 PC	5.27	0.91	29.72	32.16 ± 2.29
83.33 % 18:1 (*cis*) PC + 16.67 % 14:0 PC	4.77	0.83	27.54	30.45 ± 3.85
66.67 % 18:1 (*cis*) PC + 33.33 % 14:0 PC	3.72	0.67	22.56	18.55 ± 9.50
60 % 18:1 (*cis*) PC + 40 % 14:0 PC	3.32	0.60	20.50	12.16 ± 1.77
57.14 % 18:1 (*cis*) PC + 42.85 % 14:0 PC	3.15	0.57	19.61	12.16 ± 2.67
50 % 18:1 (*cis*) PC + 50 % 14:0 PC	2.72	0.50	17.34	19.86 ± 4.48
42.85 % 18:1 (*cis*) PC + 57.14 % 14:0 PC	2.31	0.43	15.02	15.64 ± 4.13
40 % 18: 1 (*cis*) PC + 60 % 14:0 PC	2.15	0.40	14.08	9.59 ± 3.00
33. 33 % 18:1 (*cis*) PC + 66.67 % 14:0 PC	1.77	0. 33	11.85	7.85 ± 2.99
16.67 % 18:1 (*cis*) PC + 83.33 % 14:0 PC	0.87	0.17	6.08	6.63 ± 3.04
9 % 18:1 (*cis*) PC + 91 % 14:0 PC	0.46	0.09	3.32	6.64 ± 2.84
4.76 % 18:1 (*cis*) PC + 95.24 % 14:0 PC	1.77	0.05	1.77	5.24 ± 3.03
100 % 14:0 PC	0	0	0	1.75 ± 1.10

These results are also represented in Figure 3.8. It can be observed that although it is difficult to see a variation in IVs between the different samples, and the majority have an extensive SEM, there is a trend where the higher the IV, the higher percentage of unsaturation in the sample. A line of best fit was drawn, obtaining an $R^2 = 0.8238$, indicating a moderate relationship between the values. The lipid mixtures were performed with two of the most common FAs in the cell membranes; the maximum percentage of unsaturation was 5.88 %, 1 double bond on average per lipid molecule. The difference in unsaturation between samples might be too small to be detected with the miniaturised IV assay. Therefore, the miniaturised IV assay still needed some adjustments to reduce errors and detect the changes better.

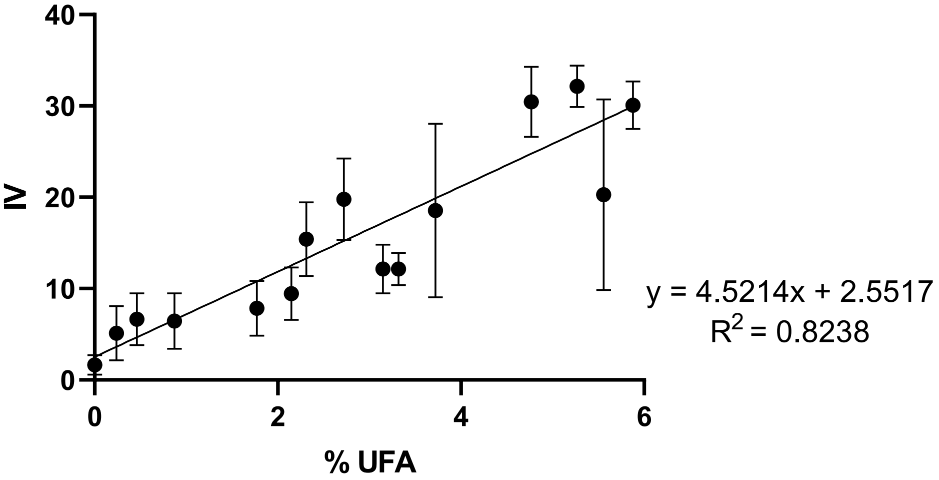


Figure 3.8 Determination of IV for lipid standards mixtures of 18:1 and 14:0.
Different mixtures of 18:1 PC and 14:0 were made, and the IV assay was performed on them (n = 5, error bars correspond to SEM). A line that best fits was drawn, and the equation and R^2 were calculated and shown on the right.

It was therefore decided to test the miniaturised IV assay with a higher percentage of unsaturation. 22:6 (*cis*) PC (1,2-didocosahexaenoyl-sn-glycero-3-phosphocoline) and 22:0 PC (1,2-dibehenoyl-sn-glycero-3-phosphocoline) were chosen. 22:6 (*cis)* PC was the most unsaturated lipid that was commercially available. 22:0 PC was chosen because it contains the same number of carbons as 22:6 (*cis*) PC, reducing any unexpected potential interference from chain length.

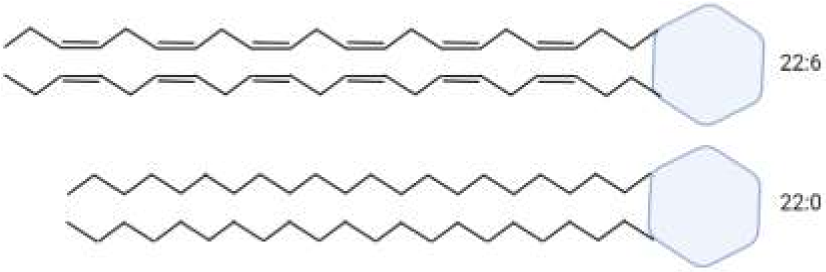


Figure 3.9 Schematic representation of 22:6 (cis) PC and 22:0 PC
22:6 (cis) PC (1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine) and 22:0 PC (1,2-dibehenoyl-sn-glycero-3-phosphocholine) are represented in the image. The head group is represented as the hexagon, and the carbon bonds are represented as black lines. Created with biorender.com

Fifteen lipid mixtures were prepared, and the percentage of unsaturation based on the molar ratio in each mixture was calculated (Table 3.3). The lipid mixtures ranged from 100 % 22:6 (*cis)* PC (28.57 % unsaturation) to 25 % 18:1 (*cis*) PC + 75 % 14:0 (1.31 % unsaturation). 100 % 14:0 DMPC (0 %

unsaturation) was used as blank. Using the plate reader, the IV for all mixtures could be obtained in less than 3 hrs.

Table 3.3 Analysis of 22:6, 22:0, 18:1 and 14:0 lipid mixtures using the miniaturised IV assay
The table includes the percentage of w/v of each lipid mixture, the % of UFAs in molar ratio, the average double bonds per FA, the calculated theoretical IV and the experimentally obtained IV ± SEM of each lipid mixture. The table is organised first by the average double bonds per FA column and in second order by the theoretical IV.

Lipid mixture (w/v)	% UFAs in molar ratio	Average double bonds per FA	Theoretical IV	Experimental IV
100 % 22:6 (*Cis)* PC	28.57	6.00	173.43	135.36 ± 0.00
75 % 22:6 (*Cis*) PC + 25 % 18:1 (*Cis)* PC	22.42	4.75	141.00	131.13 ± 4.23
75 % 22:6 (*Cis*) PC + 25 % 14:0 PC	19.96	4.50	137.94	135.00 ± 0.00
75 % 22:6 (*Cis*) PC + 25 % 22:0 PC	21.57	4.50	129.19	135.00 ± 0.00
50 % 22:6 (*Cis*) PC + 50 % 18:1 (*Cis)* PC	16.60	4.00	106.77	135.6 ± 0.00
50 % 22:6 (*Cis*) PC + 50 % 14:0 PC	12.45	3.00	97.88	118.00 ± 4.23
25 % 22:6 (*Cis*) PC + 75 % 18:1 (*Cis)* PC	11.10	2.25	70.59	97.3 ± 7.33
25 % 22:6 (*Cis*) PC + 75 % 14:0 PC	5.85	1.50	52.30	42.3 ± 25.7
25 % 22:6 (*Cis*) PC + 75 % 22:0 PC	7.29	1.50	42.48	8.46 ± 12.70
100 % 18:1 (*Cis)* PC	5.88	1.00	32.29	50.8 ± 29.60
75 % 18:1 (*Cis*) PC + 25 % 14:0 PC	4.24	0.75	25.08	4.2 ± 8.5
50 % 18:1 (*Cis)* PC + 50 % 22:0 PC	3.14	0.50	15.03	4.23 ± 4.23
50 % 18:1 (*Cis)* PC + 50 % 14:0 PC	2.72	0.50	17.34	0.00 ± 0.00
25 % 18:1 (*Cis*) PC + 75 % 14:0 PC	1.31	0.25	9.00	0.00 ± 0.00
100 % 14:0 PC	0	0	Blank	Blank

Figure 3.10 shows the data obtained with the miniaturised IV assay. This indicates a defined linear range, with samples above 16.00% or below 1.31% lying outside of this.

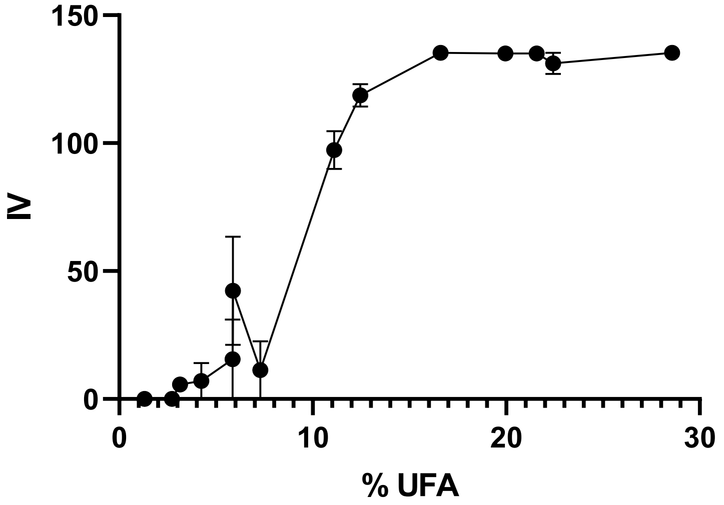


Figure 3.10 Initial determination of IV for lipid standards.
15 mixtures of 22:6 (cis) PC, 22:0 PC, 18:1 (cis) PC and 14:0 PC at different ratios (Table 3.3) were analysed with IV assay (n=3), error bars correspond to SEM.

To further optimise the assay, it was decided to make two changes. First, 1 mg of the highest percentage of unsaturation (from 28.57 to 16.6 %) was used with double the amount of ICl. ICl needs to be in excess since the titration is performed on the I_2 that did not react with the double bond. The more unsaturations, the less free I_2; if the amount of ICl is doubled, more free I_2 can be titrated, which might give an optimised performance at the highest percentage of unsaturation. Second, 0.5 mg of the same lipid mixtures were used to see if increasing the ICl and having a higher excess could decrease the sample amount and still detect unsaturations (Figure 3.11). The theoretical IV of the samples was calculated to determine the most accurate value (black squares). Doubling the amount of ICl with 1 mg of sample (full red circles) did not improve the assay since the experimental IVs were all the same regards the percentage of unsaturation. 0.5 mg of sample (open red circles) gave better results as it could detect different IVs between the samples that followed the expected trend: a higher IV for a higher percentage of unsaturation. However, the SEM was more significant than 1 mg.

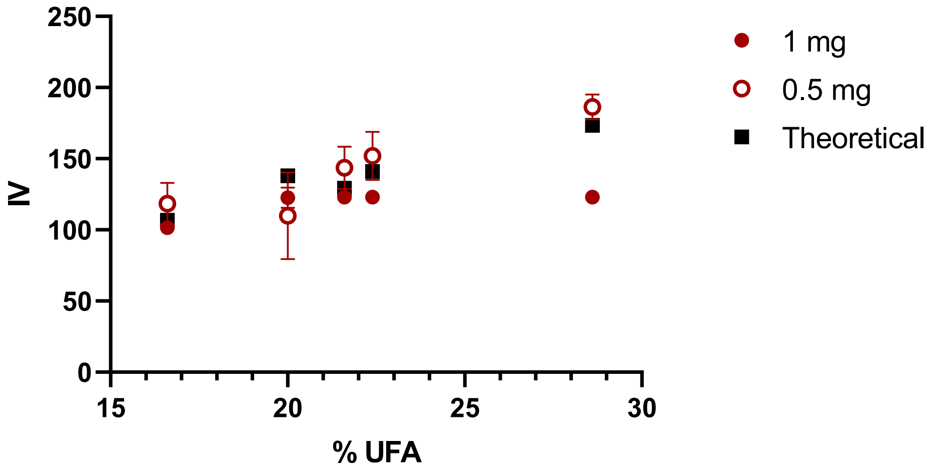


Figure 3.11 Miniaturised IV assay of the highest % UFA lipid standards with double the amount of ICl. Black squares represent the theoretical IV of the chosen lipids. The experimental IVs of 1 mg sample are represented with full red circles. Red open circles represent the experimental IV of 0.5 mg of sample. The assay was performed in triplicates; error bars correspond to SEM.

Since 0.5 mg of sample and doubling the amount of ICl gave better results than 1 mg, it was decided to try with a lower percentage of unsaturations (5.85 % to 1.31%) and just 0.5 mg of sample and double the amount of ICl (Figure 3.12). For the lowest percentage of unsaturation, using 0.5 mg of sample and doubling the amount of ICl did not improve the assay (open red circle), and the SEM was much more prominent. The obtained values were outside the theoretical values (black squares), and the IV was undetected in some replicates.

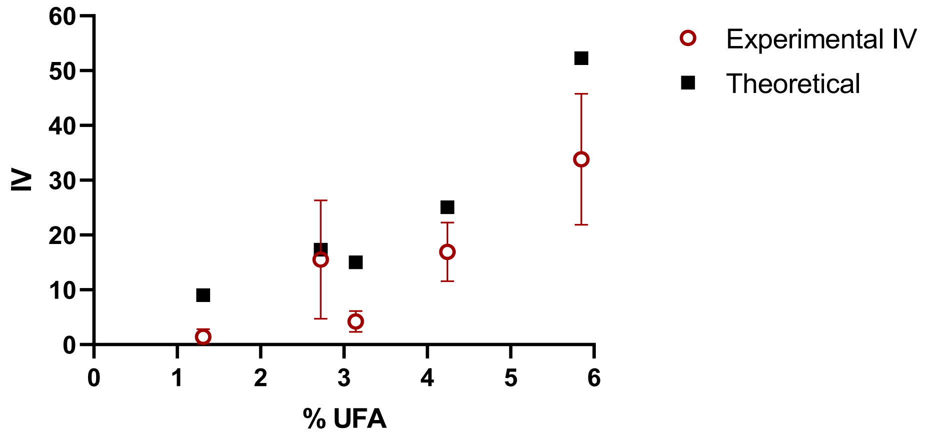


Figure 3.12 Miniaturised IV assay of the lowest % UFA lipid standards with double ICl, 80µL at 1M. Black squares represent the theoretical IV of the chosen lipids. Red open circles represent the experimental IV of 0.5 mg of sample. (n = 2, error bars correspond to SEM).

Because the unsaturation in cell membranes is most commonly low (≤ 6 %), it was decided to stay with 1 mg of sample and 0.1 M ICl.

The miniaturised IV assay was repeated to obtain replicates with the 15 different lipid mixtures. The theoretical IV for each mixture was calculated (Equation 3.2) and plotted alongside the experimental IV of the mixtures (Table 3.4 and Figure 3.13). The iodine value was determined for each mixture in triplicate, illustrated as “Experimental IV” in Table 3.4. A trend can be observed, the higher the IV, the higher the double bonds per FA. The Experimental IV is very close to the Theoretical IV. This looked very optimistic and was confirmed with statistical analysis.

Table 3.4 Analysis of lipid mixes using the miniaturised IV assay.
The table includes the percentage of w/v of each lipid mixture, the % of UFAs in molar ratio, the average double bonds per FA, the calculated theoretical IV and the experimentally obtained IV ± SEM of each lipid mixture. The table is organised first by the average double bonds per FA column and in second order by the theoretical IV.

Lipid mixture (w/w)	**% UFAs in molar ratio**	**Average double bonds per FA**	**Theoretical IV**	**Experimental IV**	**Calculated double bonds per FA**
100 % 22:6 (*Cis*) PC	28.57	6.00	173.43	152.35 ± 7.71	5.27
75 % 22:6 (*Cis*) PC + 25 % 18:1 (*Cis*) PC	22.42	4.75	141.00	141.77 ± 5.95	4.78
75 % 22:6 (*Cis*) PC + 25 % 14:0 PC	19.96	4.50	137.94	130.08 ± 8.81	4.24
75 % 22:6 (*Cis*) PC + 25 % 22:0 PC	21.57	4.50	129.19	143.67 ± 14.72	5.01
50 % 22:6 (*Cis*) PC + 50 % 18:1 (*Cis*) PC	16.60	4.00	106.77	131.18 ± 5.56	4.3
50 % 22:6 (*Cis*) PC + 50 % 14:0 PC	12.45	3.00	97.88	81.51 ± 11.43	2.5
25 % 22:6 (*Cis*) PC + 75 % 18:1 (*Cis*) PC	11.10	2.25	70.59	83.63 ± 10.16	2.67
25 % 22:6 (*Cis*) PC + 75 % 14:0 PC	5.85	1.50	52.30	29.62 ± 10.6	0.85
25 % 22:6 (*Cis*) PC + 75 % 22:0 PC	7.29	1.50	42.48	24.9 ± 5.05	0.88
100 % 18:1 (*Cis*) PC	5.88	1.00	32.29	32.31 ± 8.68	1
75 % 18:1 (*Cis*) PC + 25 % 14:0 PC	4.24	0.75	25.08	40.89 ± 10.08	1.22
50 % 18:1 (*Cis*) PC + 50 % 14:0 PC	2.72	0.50	17.34	16.15 ± 8.94	0.47
50 % 18:1 (*Cis*) PC + 50 % 22:0 PC	3.14	0.50	15.03	12.60 ± 4.42	0.42
25 % 18:1 (*Cis*) PC + 75 % 14:0 PC	1.31	0.25	9.00	8.44 ± 6.81	0.23
100 % 14:0 PC	0	0	0	Blank	Blank

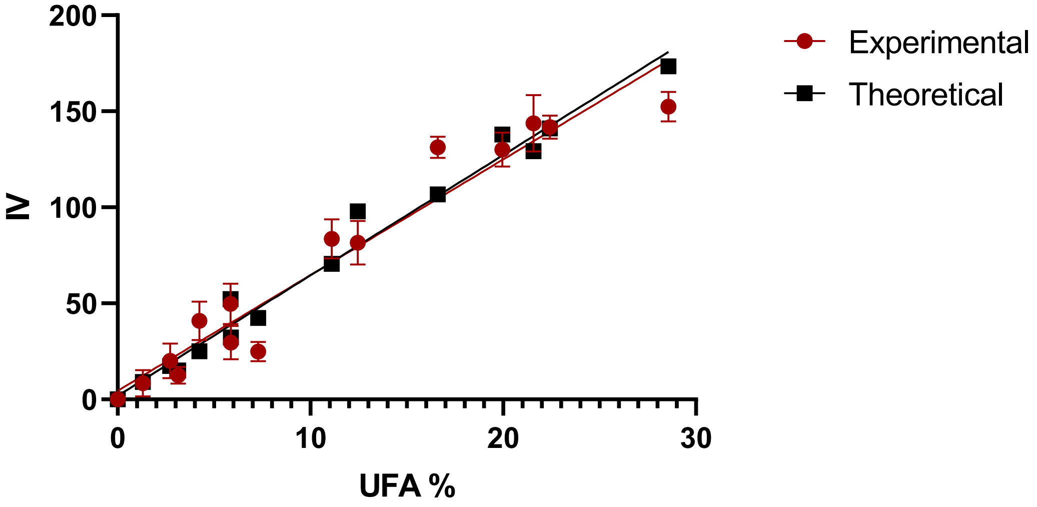


Figure 3.13 Determination of IV for the lipid standards.
Black squares represent the theoretical IV of the standards, and red circles the experimental values (n = 4, error bars correspond to SEM). A line that best fits both groups was drawn.

If the molecular weight of the substances is known, it can calculate the average number of double bonds by rearranging Equation 3.2; Equation 3.3 (Odoom and Edusei, 2015). Both equations are based on the rationale that 1 mol of double bond reacts with 1 mol of I_2 (I_2 MW = I MW * 2 = 126.9 * 2) and that IV is the grams of iodine absorbed by a 100 gr of sample. The average number of double bonds per fatty acid of the mixtures was calculated using Equation 3.3 and plotted in Table 3.4 as calculated double bonds per lipid molecule.

Equation 3.3 Calculation of the average number of double bonds per FA

$$The\ average\ number\ of\ double\ bonds\ per\ FA = \frac{IV\ x\ MW\ of\ substance}{253.8\ x\ 100}$$

In order to prove that the theoretical and experimental IVs were equal, a linear regression model was used. To determine if there was a significant difference between the theoretical and the experimental a line that best fits both values was calculated (Figure 3.14). With linear regression, an $R^2 = 0.9559$ was obtained, indicating with a high confidence level that the experimental IV strongly relates to the theoretical IV. It can also be observed that the assay's sensitivity is relatively high since the IV of a mixture that contains 1.31% and an average double bond per lipid molecule of 0.25 could be detected.

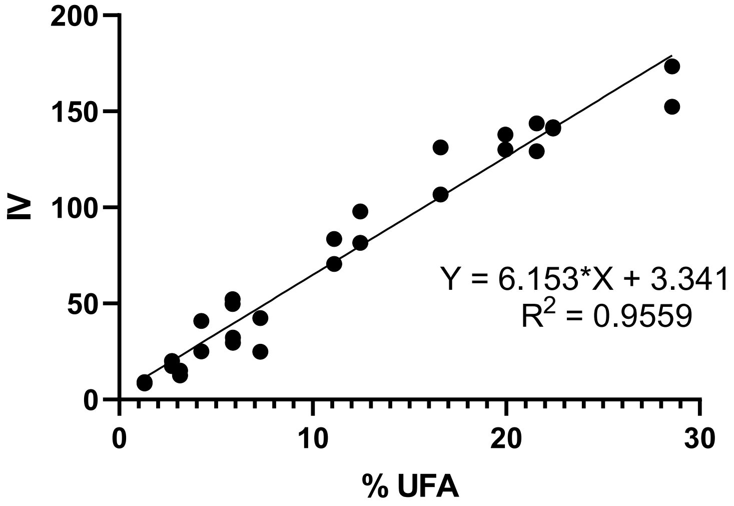


Figure 3.14 Linear regression between the theoretical and the experimental IVs.
The values obtained from the theoretical and experimental IV were put together as one set of data and a line that best fits was calculated. The equation and R^2 were written in the graph.

The results gave enough confidence to believe that the miniaturisation of the IV assay was achieved, and it was ready to be used in cell membranes. The optimised miniaturised IV assay protocol is represented in Figure 3.15

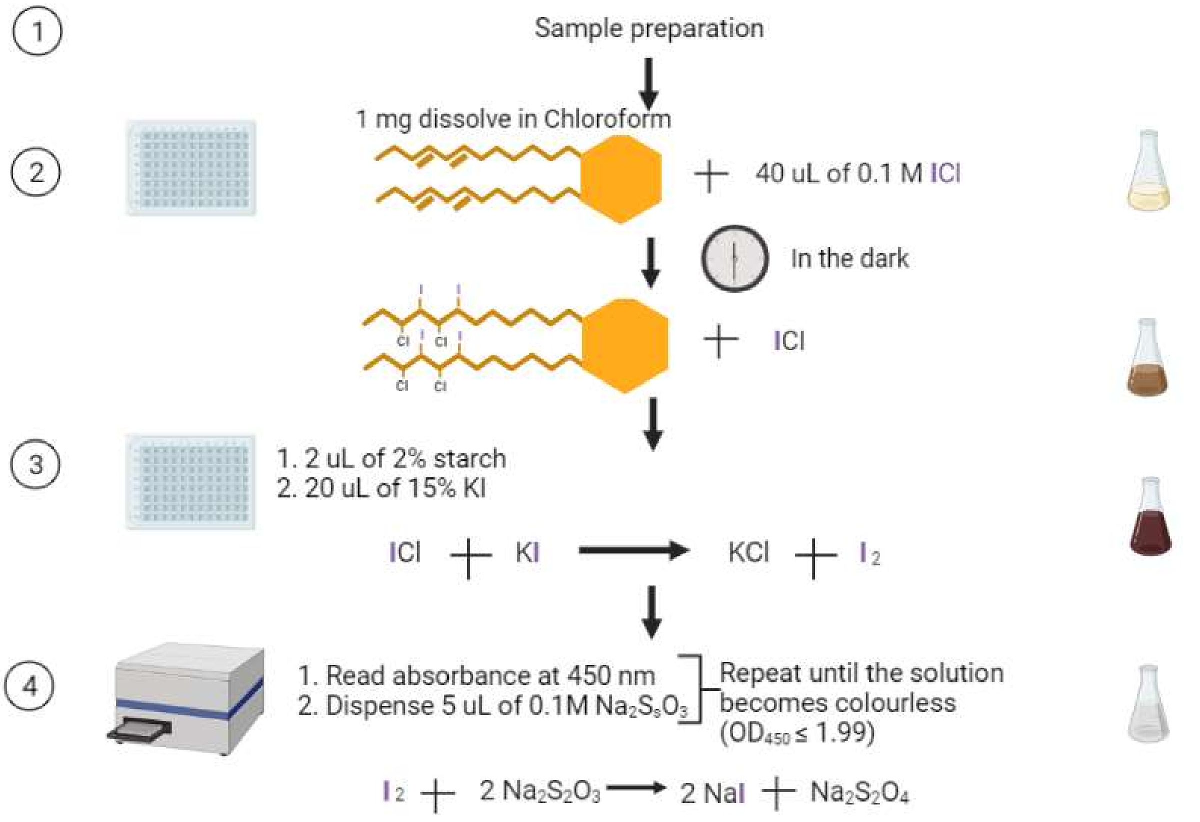

1
Sample preparation
2
1 mg dissolve in Chloroform
+ 40 uL of 0.1 M ICl
In the dark
Cl
Cl
I
I
+ ICl
3
1. 2 uL of 2% starch
2. 20 uL of 15% KI
ICl + KI → KCl + I2
4
1. Read absorbance at 450 nm
2. Dispense 5 uL of 0.1M Na2S2O3
Repeat until the solution becomes colourless (OD450 ≤ 1.99)
I2 + 2 Na2S2O3 → 2 NaI + Na2S2O4

Figure 3.15 Flowchart of the miniaturised iodine Value assay and its reactions.
(1) Sample preparation. (2) 1 mg of sample dissolved in chloroform is added to 40 µL of 0.1 M ICl and incubated for 30 min in the dark to allow the formation of dihalogenated single bonds. (3) 2 µL of 2 % starch plus 20 µL of 15 % KI is added to allow the formation of I_2 with the ICl that did not react. (4) The plate is read at 450 nm, dispensed 5 µL of 0.1 M $Na_2S_2O_3$, shaken and repeated until the OD is ≤ 1.99 due to the formation of NaI. The flasks depicted to the right indicate the colour of the solution in each step. Created with biorender.com

Two options can be taken depending on how much information is known about the sample. 1) If the percentage of unsaturation in the sample is known and it is 20 % or above or has an average double bond per lipid molecule of 4.5, 0.5 mg of sample and double the amount of ICl (80 µL) will give a better estimation. It must be kept in mind that using 0.5 mg has a more considerable standard error of the mean. Therefore, some optimisation might still be needed. 2) If there is no knowledge about the percentage of unsaturation, 1 mg and 40 µL of 0.1 M ICl will be a good place to start.

3.2.5 Miniaturised IV assay analysis of *E. coli* cell membranes

Because the primary goal of the miniaturised IV assay is to be applied to biological samples, it was decided to test it first with *E. coli* cell membranes. In order to increase the fluidity at low temperatures, the cell membrane needs to have a reduced melting temperature. To achieve that, bacteria have adapted a mechanism that increases the proportion of UFAs (Bajerski *et al.*, 2017). Therefore, cultures of *E. coli* DH5α were grown at different temperatures (30, 37, and 42 ˚C), as this is expected to cause a difference in the proportion of unsaturated bonds to control membrane fluidity. The lipids of the *E.*

coli membrane were extracted using the Bligh and Dyer method (Bligh and Dyer, 1959). 1 mg of lipid was used for the IV assay performed in triplicate. 14:0 DMPC was used as blank. *E. coli* total lipid extract from Avanti (USA) was also run and used as standard (*E. coli* std) (Figure 3.16). Figure 3.16 shows that compared to 37 ˚ C, on average, the IV at 30 ˚C is slightly higher and lower at 42 ˚C, but there is no significant difference. The lower the temperature, the higher the IV, which correlates with what has been reported. The lower the temperature, the higher the unsaturations.

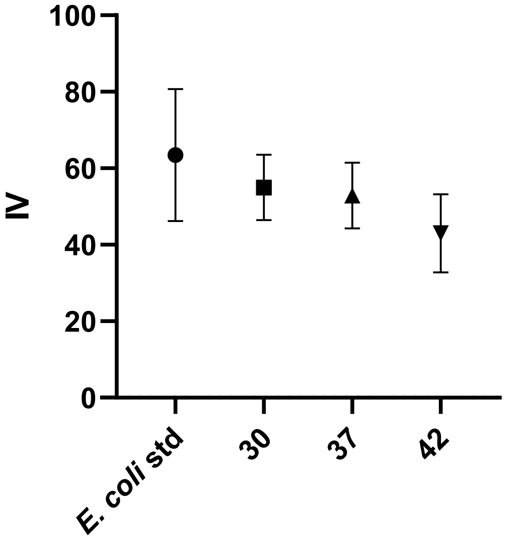


Figure 3.16 IV assay of *E. coli* lipid membrane grown at different temperatures.
The miniaturised IV assay was performed on *E. coli* lipid membranes grown at 30, 37 and 42 ˚C (n = 3, error bars correspond to SEM). A 1-way ANOVA was performed to calculate if there was a significant difference between the samples. There was no significant difference.

Several publications have reported for *E. coli* strains grown at 37 ˚C a percentage unsaturation from 1.44 to 3.77 % and an average double bond number from 0.24 to 0.64 (Bright-Gaertner and Proulx, 1971, Mavis and Vagelos, 1972, Pramanik and Keasling, 1997). The *E. coli* std from Avanti has known concentrations of *E. coli* FAs, corresponding to 0.85 average double bonds per FA (2.16 % of unsaturation). Based on the direct proportion between the IV and the number of double bonds in the sample, the average number of double bonds per FA at three different temperatures was determined using the rule of three (Figure 3.17). The values obtained were 0.79 ± 0.9, 0.75 ± 0.9, and 0.61 ± 0.1 at 30, 37 and 42 ˚C, respectively.

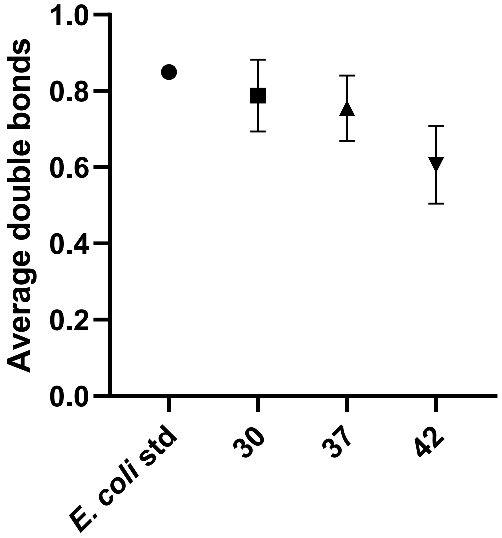


Figure 3.17 Average double bond in *E. coli* lipid membranes
Based on the *E. coli* standard and the obtained experimental IV, the average double bond per lipid molecule was calculated on the different *E. coli* lipid membranes grown at different temperatures (n = 3, error bars represent SEM). A 1-way ANOVA was performed to calculate if there was a significant difference between samples. There was no significant difference.

No significant difference was obtained between 30 and 37 ˚ C. Morein *et al.* (1996) determined the acyl composition of *E. coli* grown at 37 and 27 ˚C by gas chromatography, obtaining an average double bond of 0.47 and 0.54, respectively. Demonstrating that the average double bond increases as the temperature decreases, but the difference is too small to be significant. The results from the miniaturised IV assay correlate well with what has been reported. The miniaturised IV assay can detect the small changes of unsaturation between samples under conditions where unsaturation would be expected to change.

3.2.6 Miniaturised IV assay in total yeast lipids

It was then decided to test the IV with yeast to demonstrate the broader applicability of the assay. A 300 mL culture of *S. cerevisiae* was grown at 30 ˚C in GPY media for 72 hrs. The lipids were extracted using the Bligh and Dyer method (Bligh and Dyer, 1959). 1 mg of lipid was used for the miniaturised IV assay performed in triplicate. 1 mg of 14:0 DMPC lipid standard was used as blank. 1 mg of Total Yeast Lipid Extract total from Avanti (USA) was used as standard (yeast std) (Figure 3.18).

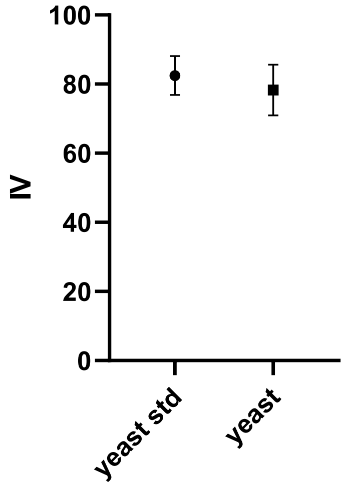


Figure 3.18 IV assay of yeast lipid grown at 32 ˚C and compared with commercial yeast lipids.
The miniaturised IV assay was performed in yeast lipid membranes grown at 30 ˚C; the samples were run in triplicate. Error bars correspond to SEM. A 1-way ANOVA was performed to calculate if there was a significant difference between samples. There was no significant difference.

The total yeast lipid extract has a calculated 4.25 % unsaturation and an average double bond of 0.75. Using the same procedure as with *E. coli* (1.2.5), the average double bond of the yeast sample was calculated, and a 0.71 average double bond was obtained. Michalik *et al.* (2014) reported 4.97 % of unsaturation and 0.80 average double bonds per lipid molecule in *S. cerevisiae* when grown on industrial glycerol at 28 ˚C for 48 h. These results correlate well with the standard and with what has been reported.

To summarise, IV, a colourimetric method that measures the total unsaturation in the sample and is used as an American Oil Chemists' Society (AOCS) method, was miniaturised to apply to a range of biological samples. In this chapter, the IV method was performed in microplates using 1 mg sample. The miniaturised IV assay is valuable since it can be used in biological samples, multiple samples can be analysed simultaneously, and it is reproducible and easy to implement in any laboratory. Furthermore, it does not require sophisticated equipment, and data analysis involves simple mathematical calculations. The miniaturised IV assay was used to determine the average of double bonds per lipid molecule in cell membranes of *E. coli* and yeast obtaining results that correlate well with what has been reported in the literature. The main reason for the miniaturisation of the IV method was to use it at different time points during the fermentation of different strains of *C. saccharoperbutylacetonicum* N1-4HMT. To observe the bacteria's protective mechanisms at the membrane level during the acidogenic phase and use the information to modify the membrane of the bacteria and increase its tolerance to organic acids. Unfortunately, this was not performed due to the delay caused by the closure of the laboratories because of the Covid-19 pandemic.

4. Results II – Characterisation of membrane proteins for their role in organic acid tolerance

4.1 Introduction

Many bioprocesses are hindered by product toxicity, which may occur at the cell membrane. Understanding the cellular membrane stress response to chemicals will facilitate strain improvement to increase yields. Undissociated organic acids can permeate the cell membrane. Once inside, they dissociate, releasing the anion and the proton and disrupting the transmembrane potential (Wu *et al.*, 2005). Therefore, a common problem in producing organic acids is cytoplasm acidification.

In bacteria, the transmembrane electrochemical gradient of protons, also known as the proton motive force, is generated by extruding protons from the cytoplasm through the membrane-bound proton pumps. Regulation of cytoplasmic pH requires control of the permeability of the cell membrane to protons. Cation/proton antiporters play a significant role in its acidification, exchanging cations (Na^+ or K^+) for protons (Booth, 1985). Monovalent cation/proton antiporters (CPA) are found in all domains of life. They regulate cell homeostasis, including pH, Na^+/Li^+ tolerance, cell cycle and proliferation (Furrer *et al.*, 2007). Most CPA proteins from prokaryotes and eukaryotes fall into two major families: CPA1 (2.A.36) and CPA2 (2.A.37). The CPA1 family includes the Na^+/H^+ exchanger (NHX) family and the CPA2 superfamily consists of three families NHA, KEA and CHX (Chanroj *et al.*, 2012). Cation/proton antiporters have been suggested to play a dominant role in pH control during acidification (Pieterse *et al.*, 2005). The recognition that protons can be extruded in response to cation uptake, causing an elevation of internal pH, has led to a discussion of the role of proton antiporters in preventing toxicity. No (published) attempts have been made to exploit proton antiporters to improve tolerance in acidogenic Clostridia.

If there is an overexpression of cation/proton antiporters during acidogenesis, where protons will be pumped out in exchange for cations, cells will then have a cation accumulation which may result in osmotic stress and ion toxicity. Osmotic stress is a change in the solute concentration around a cell. Cells respond to osmotic stress by accumulating or releasing solutes, attenuating water fluxes. Those solutes include inorganic ions (often K^+) and organic molecules denoted “osmolytes” or “osmoprotective compounds”, which are selected to perturb cellular functions minimally, even after accumulating to high concentrations. If the osmotic pressure rises, these solutes accumulate via uptake transporters or synthesis and are released via mechanosensitive channels if the osmotic pressure falls (Wood, 2015).

Glycerol is an osmolyte also necessary for cellular nutrition (Borgnia and Agre, 2000). The glycerol facilitator (GlpF), a protein of ~28 kDa, facilitates the flux of glycerol and linear polyalcohols across the cellular membrane (Atmadjaja *et al.*, 2019, Stroud *et al.*, 2003). GlpF is a channel that belongs to the aquaporin family. It is a highly conserved structure that consists of 6 transmembrane helices and two half-spanning helices that can form tetramers (Stroud *et al.*, 2003, Trefz *et al.*, 2018). GlpF has a crucial role since it regulates the water flow in response to osmotic stress without disturbing the electrochemical properties of the membrane (Chakrabarti *et al.*, 2004). An overexpression of GlpF has been observed during the acidogenic phase of *Clostridium acetobutylicum* (Dong *et al.*, 2016, Mao *et al.*, 2010). Therefore, GlpF is another candidate for tolerance engineering to evaluate its role as a protective mechanism during the Clostridial acidogenic phase and/or osmotic stress.

This chapter will explore the characteristics of the proton antiporters in *C. saccharoperbutylacetonicum* and their relevance in alleviating proton accumulation and cytoplasm acidification. The chapter aimed to identify two proton antiporters that could improve organic acid resistance and increase their production and express them in *E. coli* to get a deeper characterisation and better understanding of the mechanism of the protein and its role in increasing tolerance. To fully characterise and understand the mechanism of glycerol uptake, its role in the acidogenic phase and protection of osmotic stress, this chapter also aimed to express and purify the *C. saccharoperbutylacetonicum* GlpF. The *in silico* model was obtained for the three membrane proteins. The different strategies for the amplification, digestion, ligation, transformation in *E. coli,* induction, expression and purification are described in this chapter.

4.2 Results

4.2.1 Selection of proton antiporters

A search in TransportDB2.0 web, a MySQL database which provides annotations for predicted transporters (Elbourne *et al.,* 2017), resulted in the identification of *Clostridium saccharoperbutylacetonicum* N1-4HMT proton antiporters. A summary of the identified proteins is given in Table 4.1:

Table 4.1 Proton antiporters in *C. saccharoperbutylacetonicum* obtained from TransportDB2.0 web. The abbreviation, full name (Transporter classification code) and the catalysed reaction are given.

	Abbreviation	Full name (TC)	Catalysed reaction
1	PerM	Autoinducer-2 Exporter (9.B.22)	AI-2 (in) ⇌ AI-2 (out) Na^+ or Li^+ (out) + H^+ (in) → Na^+ or Li^+ (in) + H^+ (out)
2	CaCA	Ca^{2+}: Cation Antiporter (2.A.19)	Ca^{2+} (in) + [nH^+ or nNa^+ (out)] ⇌ Ca^{2+} (out) + [nH^+ or nNa^+] (in)
3	ClC	Chloride Carrier/Channel (2.A.49)	Carriers: 2 Anions (in) + H^+ (out) ⇌ 2 Anions (out) + H^+ (in).
4	Cpa1	Monovalent Cation:Proton Antiporter-1 (2.A.36)	Na^+ (out) + H^+ (in) ⇌ Na^+ (in) + H^+ (out)
5	MFS	Major Facilitator Superfamily (2.A.1)	(1) Uniport: S (out) ⇌ S (in) (2) Symport: S (out) + [H+ (or Na+)] (out) ⇌ S (in) + [H+ (or Na+)] (in) (3) Antiport: S1 (out) + S2 (in) ⇌ S1 (in) + S2 (out) (S1 may be H+ or a solute)
6	NhaA	Na^+:H^+ Antiporter A (2.A.33)	Na^+ (in) + $2H^+$ (out) ⇌ Na^+ (out) + $2H^+$ (in)
7	NhaC	Na^+:H^+ Antiporter C (2.A.35)	Na^+ (in) + nH^+ (out) ⇌ Na^+ (out) + nH^+ (in). ($n > 1$)
8	RhtB	Resistance to Homoserine/Threonine (2.A.76)	amino acid (in) + nH^+ (out) ⇌ amino acid (out) + nH^+ (in)
9	ThrE	Threonine/Serine Exporter (2.A.79)	threonine or serine (in) + H^+ (out) ⇌ threonine or serine (out) + H^+ (in)

The ideal proton antiporter for reducing weak acid toxicity is the one that exports multiple protons in exchange for a non-toxic and highly available molecule. From Table 4.1, ClC (3), MFS (5), RhtB (8) and ThrE (9) were discarded first because they do not export protons in exchange for a non-toxic, highly available molecule. (3) Chloride Carrier/Channel (ClC) has a clear role in proton expulsion in extreme acid response (Chen, 2005, Li *et al.*, 2002). However, two molecules of Cl^- are needed to extrude one H^+, making it less desirable for this application. (5) The major facilitator superfamily (MFS) is a superfamily whose main function is to expel harmful substances (Du *et al.*, 2015); sugars, drugs, metabolites, oligosaccharides, amino acids, and oxyanions are being transported by MFS and therefore, is not ideal to be used as proton extruders. (8) Resistance to Homoserine/Threonine (RhtB) and (9) Threonine/Serine Exporter (ThrE) regulate the efflux of homoserine/threonine/serine, which are involved in cell signalling; their overexpression activates the expression of the σ^S subunit of RNA polymerase, which acts during starvation and the stationary growth phase. The overexpression of RhtB and ThrE play a role in tolerance and improved metabolite production; nevertheless, the metabolites are not organic acids, nor is its mechanism of tolerance related to proton extrusion, but metabolite extrusion (Jones *et al.*, 2015, Zakataeva *et al.*, 1999). Based on these, the selection was reduced to Autoinducer-2 Exporter (PerM) (1), the Ca^{2+}/Cation antiporter (CaCA) (2), Monovalent cation:proton antiporter 1 (Cpa1) (4), and the Na^+:H^+ antiporter A and C (NhaA (6), and NhaC (7) respectively).

(1) The Autoinducer-2 Exporter (AI-2 [PerM, previously known as putative efflux pumps]) is involved in the interspecies communication “quorum-sensing”, biofilm formation and cell motility. However, only two members of the AI-2 family have been fully characterised, suggesting that some members, such as TqsA of *E. coli* or all, can catalyse the pmf-dependent efflux of intracellular and extracellular signalling molecules (Rettner and Saier, 2010). Several researchers have observed an increase in AI-2 production in acidic conditions and a decrease under alkaline conditions (Zhao *et al.*, 2018).

(2) The Ca^{2+}/cation antiporter (CaCA) members transport Ca^{2+} and/or other cations in exchange for another anion, such as H^+ or Na^+. Animal proteins use Na^+ gradients as the driving force, while plants and bacteria use H^+ (Emery *et al.*, 2012). CaCA proteins are essential for maintaining Ca^{2+} homeostasis (Wu *et al.*, 2013a), mineral nutrition, ion stress tolerance and signal transduction (Emery *et al.*, 2012).

(4) Na^+/H^+ antiporters play crucial roles in regulating intracellular pH, sodium homeostasis and cell volume. Monovalent cation:proton antiporter 1 (Cpa1) members are electroneutral and exchange protons and Na^+ ions with a 1:1 stoichiometry. This group strongly depend on extracellular Na^+ to extrude protons from the cell (Paulino and Kuhlbrandt, 2014).

(6) The Na^+:H^+ antiporter A (NhaA) is an electrogenic antiporter which exchanges Na^+ ions and protons with a 1:2 stoichiometry. The family’s best-known member is NhaA from *E. coli,* which enables it to survive at high salinity or alkaline pH (Paulino and Kuhlbrandt, 2014, Wohlert *et al.*, 2014). Wu *et al.* (2013b) observed that overexpression of NhaA in *E. coli* led to a 25% increase in lactate production.

(7) The Na^+:H^+ antiporter C (NhaC) plays a role in pH homeostasis and lowers the cytoplasmic Na^+ concentration. However, NhaA is the dominant antiporter. It is believed that NhaC V_{max} is insufficiently high and/or the H^+/Na^+ ratio is inefficient; therefore, it has been suggested that NhaC plays an auxiliary role (Ito *et al.*, 1997).

Based on the proton antiporters found in *C. saccharoperbutylacetonicum* and the aim of identifying two that export protons in exchange for a non-toxic and highly available molecule. The choice was between PerM (1), CaCA (2), Cpa1 (4), NhaA (6) and NhaC (7). The principal role of PerM, AI-2 (1) is quorum sensing, which significantly regulates the solvent formation, cell motility and sporulation (Feng *et al.*, 2020). It is still uncertain whether the Clostridial CaCA (2) uses H^+ or Na^+_+ as their cation-gradient driving forces. Moreover, NhaC (7) is believed to have an auxiliary role. The Na_+:H_+ antiporter A (NhaA) (6) and the monovalent cation:proton antiporter 1 (Cpa1) (4) were chosen for further analysis because they contribute to endogenous pH, cation and osmotic homeostasis, and there is enough evidence of their role as pH-sensors and in cellular adaptation (Bartsch 2015).

Once the two proton antiporters were chosen, their nucleotide and protein sequences were obtained to generate an *in silico* model of NhaA and Cpa1 and also to proceed with their overexpression in *E. coli*.

4.2.2 Nucleotide sequencing and protein modelling

There currently exists no structures produced for the Clostridial proteins. The atoms in the three-dimensional structure dictate the protein's function. Protein modelling can provide a level of understanding of how a protein works and can help create a hypothesis about how to affect it, control it, or modify it. Because Clostridial proteins are relatively uncharacterised, structural models might help understand their function. Protein structure prediction relies on homology modelling, where homologue proteins descend from a common ancestor and are likely to present the same structure and function. The 3D structure of a target protein can be modelled based on one or a combination of several template molecules if it has a sequence identity of at least 30 % with another protein with a known structure.

4.2.2.1. Proton antiporters

Cation/proton antiporters (CPA) are a family of proteins in charge of pH homeostasis and ion concentration. They are classified based on their electrogenicity in two groups: CPA1 and CPA2; CPA1s are electroneutral (1 Na^+ is exchanged by 1 H^+), while CPA2s are electrogenic (1 Na^+ is exchanged by 2 H^+). From the protein sequence, it was possible to confirm that NhaA was electrogenic (CPA2) and Cpa1 was electroneutral (CPA1). CPA1 can be further divided into six main clades, whereas CPA2 comprises nine main clades (Masrati *et al.*, 2018). Masrati *et al.* (2018) looked for the amino acids that differentiated these phenotypes. Using ConSurf to identify structurally and functionally essential residues that tend to be evolutionary conserved among homologous proteins, they identified a eight highly conserved positions that capture the different CPA clades and may lay the basis for the electrogenicity and ion selectivity. At the protein core, all eight positions clustered and included the substrate binding site. The residues are $X_1X_2X_3X_4[E/-]_5X_6D_7[R/K_8]$. Positions 1 to 4 feature various residues that discriminate between different CPA groups within the two main clades. E_5 and R_8 for CPA1 members and any residue in the fifth position and K_8 for CPA2 members. Figure 4.1 shows a schematic representation of the cation:proton antiporter and the eight highly conserved residues in their transmembrane location with one example for each family.

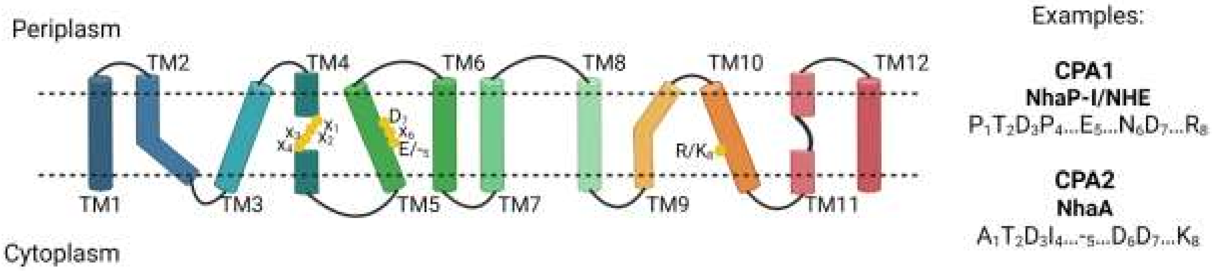
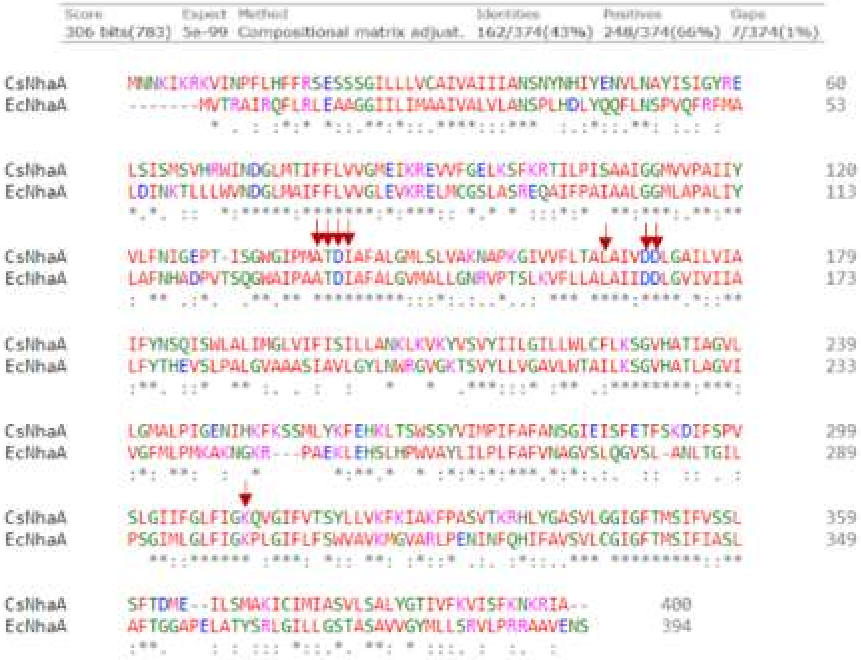


Figure 4.1 Schematic two-dimensional representation of the cation:proton antiporters
Located in the core domain is the CPA motif. The dashed lines show the membrane boundaries; the helices are numbered TM-1 through- TM-12, and the yellow spheres are the motif residues. Residues X1-X4 distinguish between different CPA groups within each of the 2 main clades. On the 5th position, CPA1 accommodates only glutamate while CPA2 any residue. 6th and 7th comprise the ND/DD motif, and the 8th position features arginine or lysine. Examples of CPA1 and CPA2 motif residues are written on the right. Created with biorender.com

4.2.2.1.1. NhaA

For CPA2 members, such as *E. coli* (Ec) NhaA, the residues are $X_1X_2X_3X_4X_5X_6D_7K_8$. Specifically, EcNhaA residues are TM4- $A131_1$, $T132_2$, $D133_3$ and $I134_4$. $L159_5$ is located on TM-5. The sixth and seventh positions, D163 and D164, also on TM-5, comprise the well-known DD motif in the binding site, which gives the capacity to exchange 2 H^+. The eighth position features K300 on TM-10 (Masrati *et al.*, 2018). The residues Asp163, Asp164, Asp 133, Thr 132 and Lys 300 have been suggested to form the NhaA active site.

The sequence for *C. saccharoperbutylacetonicum* N1-4HMT NhaA (CsNhaA) was obtained from Uniprot (https://www.uniprot.org/uniprot/M1MPG3) and NCBI Reference sequence: WP_015392945.1. The sequence for Ec NhaA was obtained by performing a protein Blast using the NCBI blast tool and the NCBI reference sequence: NEY29192.1. Using Clustal Omega from EMBL-EBI (McWilliam *et al.*, 2013), an alignment with related structures and *E. coli* proteins was performed. The eight residues highly conserved in the CPA2 family can be found in the NhaA for *C. saccharoperbutylacetonicum* (Asp170, Asp171, Asp140, Thr139 and Lys311 [Figure 4.2, red arrows]), strongly suggesting it is electrogenic. The first step to obtaining each protein's three-dimensional (3D) model is to perform an alignment (Figure 4.2). *C. saccharoperbutylacetonicum* NhaA has a 42.15 % sequence similarity to *E. coli* NhaA, EcNhaA structure has been experimentally obtained, and because it is 42.15 % sequence similar to the CsNhaA, the 3D model would be reliable.



```
Score           Expect  Method                          Identities      Positives       Gaps
306 bits(783)   5e-99   Compositional matrix adjust.    162/374(43%)    248/374(66%)    7/374(1%)

CsNhaA    MNNKIKRKVINPFLHFFRSESSSGILLLVCAIVAIIIANSNYNHIYENVLNAYISIGYRE    60
EcNhaA    -------MVTRAIRQFLRLEAAGGIILIMAAIVALVLANSPLHDLYQQFLNSPVQFRFMA    53

CsNhaA    LSISMSVHRWINDGLMTIFFLVVGMEIKREVVFGELKSFKRTILPISAAIGGMVVPAIIY    120
EcNhaA    LDINKTLLLWVNDGLMAIFFLVVGLEVKRELMCGSLASREQAIFPAIAALGGMLAPALIY    113

CsNhaA    VLFNIGEPT-ISGWGIPMATDIAFALGMLSLVAKNAPKGIVVFLTALAIVDDLGAILVIA    179
EcNhaA    LAFNHADPVTSQGWAIPAATDIAFALGVMALLGNRVPTSLKVFLLALAIIDDLGVIVIIA    173

CsNhaA    IFYNSQISWLALIMGLVIFISILLANKLKVKYVSVYIILGILLWLCFLKSGVHATIAGVL    239
EcNhaA    LFYTHEVSLPALGVAAASIAVLGYLNWRGVGKTSVYLLVGAVLWTAILKSGVHATLAGVI    233

CsNhaA    LGMALPIGENIHKFKSSMLYKFEHKLTSWSSYVIMPIFAFANSGIEISFETFSKDIFSPV    299
EcNhaA    VGFMLPMKAKNGKR---PAEKLEHSLHPWVAYLILPLFAFVNAGVSLQGVSL-ANLTGIL    289

CsNhaA    SLGIIFGLFIGKQVGIFVTSYLLVKFKIAKFPASVTKRHLYGASVLGGIGFTMSIFVSSL    359
EcNhaA    PSGIMLGLFIGKPLGIFLFSWVAVKMGVARLPENINFQHIFAVSVLCGIGFTMSIFIASL    349

CsNhaA    SFTDME--ILSMAKICIMIASVLSALYGTIVFKVISFKNKRIA--      400
EcNhaA    AFTGGAPELATYSRLGILLGSTASAVVGYMLLSRVLPRRAAVENS      394
```

Figure 4.2 Sequence alignment between *E. coli* and *C. saccharoperbutylacetonicum* NhaA
Sequence alignment of *C. saccharoperbutylacetonicum* NhaA (CsNhaA) vs *E. coli* NhaA (EcNhaA). The numbers on the right refer to protein residue numbers. (*) Indicates a single, fully conserved residue. (:) Indicates conservation between groups of strongly similar properties. (.) Indicates conservation between groups of weakly similar properties. (-) Indicates gaps in the sequences. Alignment obtained by Clustal Omega from EMBL-EBI. The conserved residues are indicated with a red arrow at the top.

The predicted structure of *C. saccharoperbutylacetonicum* NhaA was obtained using SWISS-MODEL from Expasy software. The predicted conformation was based on sequence identity (42 %) from the experimentally determined protein structure of Ec-NhaA (pdb: 1zcd). The co-ordinate file given by SWISS-MODEL was displayed using the molecular visualisation program of PyMol to generate good-quality images. NhaA (Figure 4.3) consists of 400 amino acid residues with the amino and carboxy terminus exposed to the cytoplasm. The model is arranged in 12 transmembrane segments, a characteristic shared by MFS transporters, which might be critical for antiporter activities (Padan, 2014, Lee *et al.*, 2016).

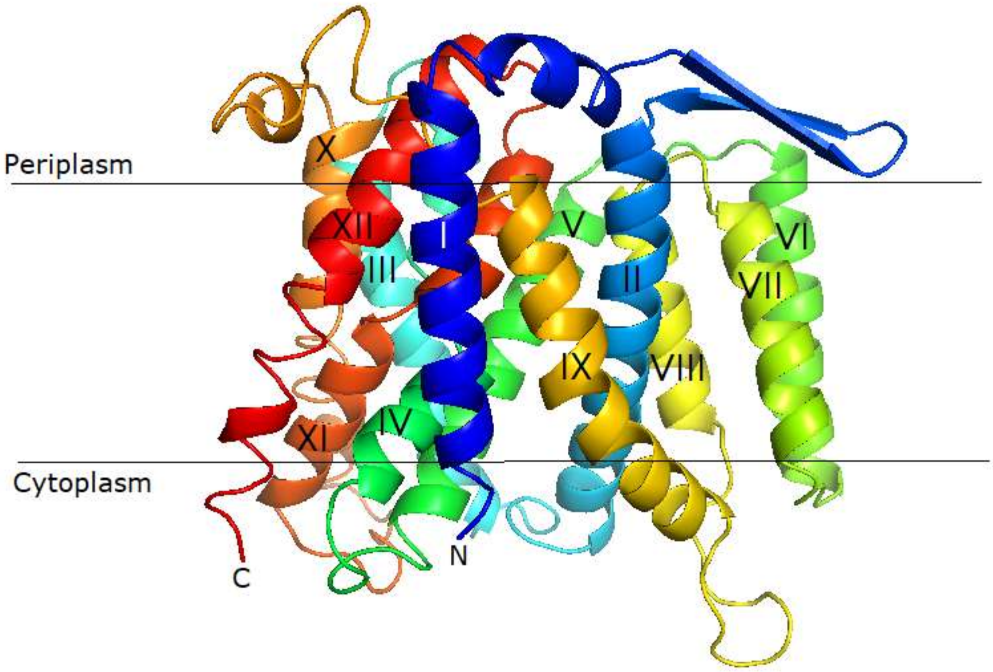


Figure 4.3 Overall architecture of *C. saccharoperbutylacetonicum* NhaA
Stereo view of a ribbon representation viewed parallel to the membrane (grey lines). The transmembrane domains are labelled with Roman numerals. Amino and carboxy termini are indicated by N and C. Protein was obtained using Swiss-model and PyMol.

The Ramachandran plot that is used to evaluate the quality of the 3D prediction of the structure gave an 85.52 % of favoured residues (Figure 4.4). Ramachandran plots visualise energetically favoured regions for backbone dihedral angles against amino acid residues in protein structure. Each point represents a protein residue; the coloured regions represent favoured and allowed regions. Ideally, no residue should be an outlier except for glycine and proline. Because it is unique and lacks a side chain, glycine can adopt phi and psi angles in all four quadrants. Because the pyrrolidine is very restricted, a specific proline Ramachandran plot is needed (Ho and Brasseur, 2005). A range between 85-90% of residues falling in the allowed areas is acceptable (Elsliger and Wilson, 2012, Rao *et al.*, 2020).

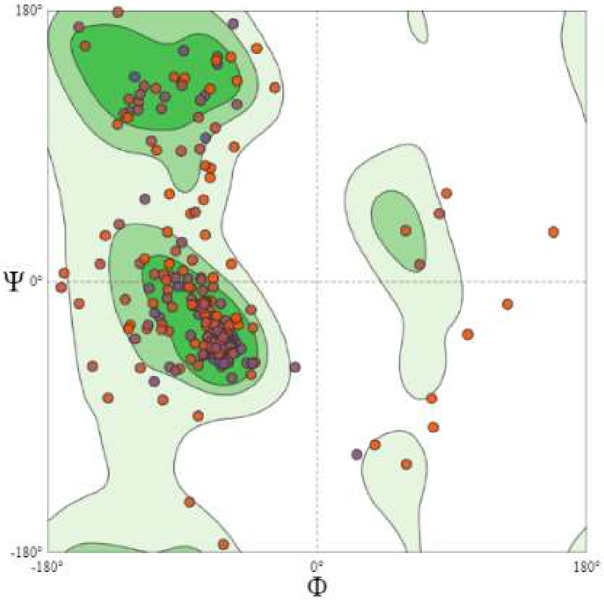


Figure 4.4 Ramachandran plot for the *C. saccharoperbutylacetonicum* NhaA protein
Plot prepared with SWISS-MODEL that shows the φ/ψ torsion angles. The red circles are the residues of the protein. The dark green regions are the favoured, and the light green are the allowed regions. For Cs-NhaA, most residues are found in the favoured and allowed regions.

Figure 4.5 shows the superimposition of the two 3D models of *C. saccharoperbutylacetonicum* and *E. coli* NhaA. The root mean square deviation (RMSD), the extent to which the main chains of the template and the target protein overlap, is the most popular method for protein structure. However, it is the least representative since errors dominate it because it depends on the accuracy of the experimentally determined protein structure (Kufareva and Abagyan, 2012). An RMSD value of 0 means perfect fit, and less than 0.5 means that the target and the template have essentially identical 3D structures. NhaA of *C. saccharoperbutylacetonicum* and *E. coli* has an RSMD of 0.169, meaning they have an identical 3D structure.

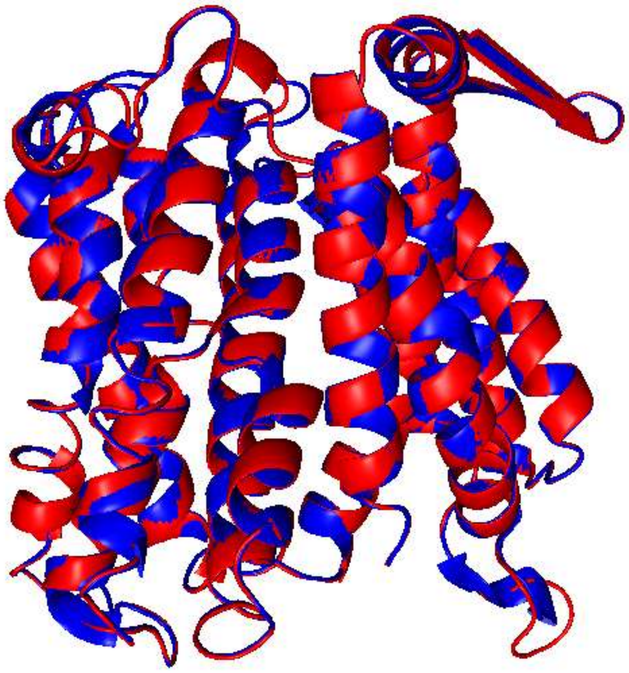


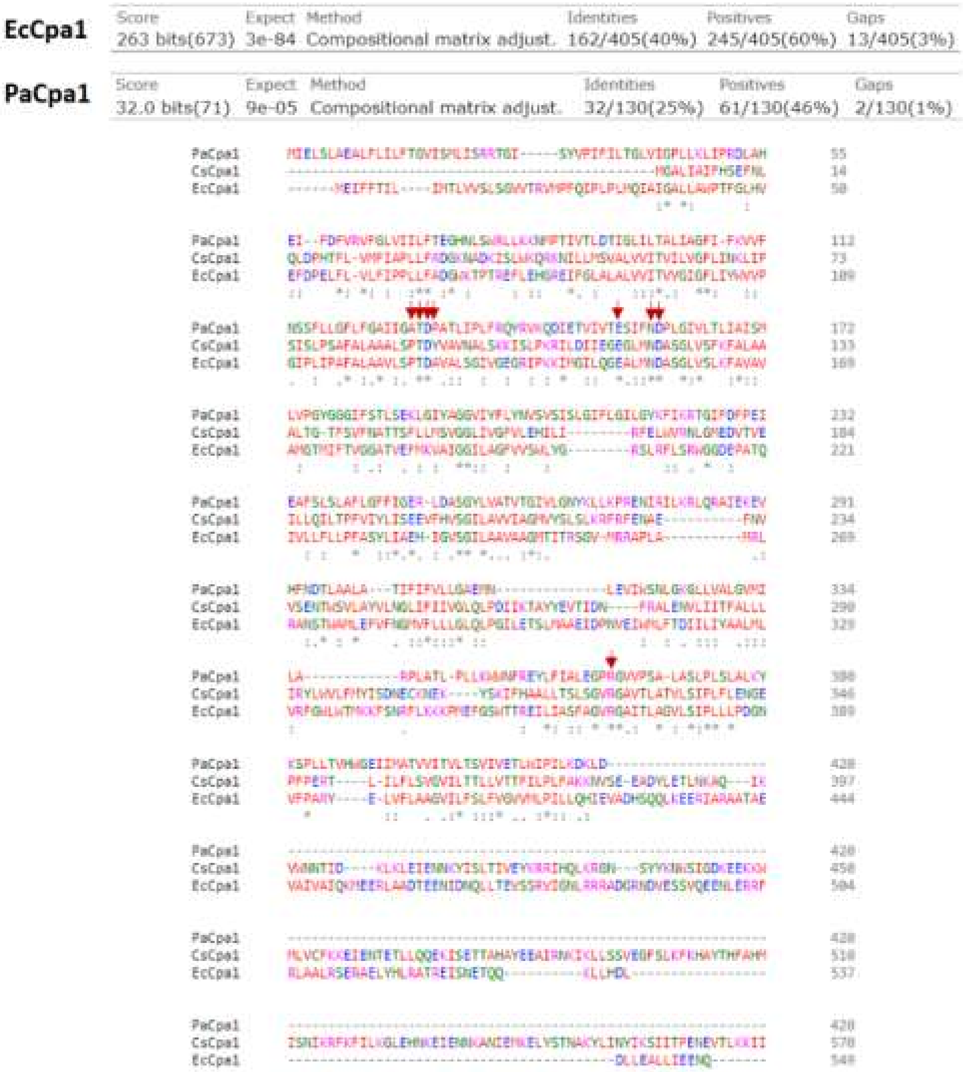
Figure 4.5 Comparative modelling of *E. coli* vs *C. saccharoperbutylacetonicum* NhaA
Comparison of *C. saccharoperbutylacetonicum* NhaA (in red) and NhaA (in blue) protein models. RMSD = 0.169. Software: PyMol.

4.2.2.1.2 Cpa1

The sequence for *C. saccharoperbutylacetonicum* N1-4HMT Cpa1 (CsCpa1) was obtained from Uniprot (https://www.uniprot.org/uniprot/M1MNX0) and NCBI Reference Sequence: AGF57903.1. *E. coli* Cpa1 (EcCpa1) was obtained by performing a protein blast from the NCBI, reference sequence: EEC7718274.1. The *Pyrococcus abyssi* NhaP (PaCpa1) was obtained through the SWISS_MODEL by Expasy code 4cza.1.

EcCpa1 belongs to the CPA1 family; it consists of 638 amino acid residues, with the amino terminus exposed to the exterior and the carboxy terminus exposed to the cytoplasm. The closest template protein with an experimental structure is the NhaP from *P. abyssi* (pdb: 4cza); however, their sequence similarity (21.73 %) is insufficient to give a reliable 3D model. *P. abyssi* consists of 420 amino acid residues, 13 transmembranes, with the N terminal in the exterior and the C terminal in the cytoplasm (Matsuoka *et al.*, 2022). Figure 4.6 shows the sequence alignment of *C. saccharoperbutylacetonicum* Cpa1*, E. coli* Cpa1 and *P. abyssi* NhaP. *E. coli* Cpa1 is similar to *C. saccharoperbutylacetonicum,* but its structure has not been experimentally identified. It is unknown if the large missing C-terminal section in *P. abyssi* are other domains on *E. coli* and *C. saccharoperbutylacetonicum,* more research on these two proteins are needed. Cpa1 belongs to the CPA1 family of electroneutral antiporters (Figure 4.7). A glutamate in the fifth position of the eight highly conserved residues is only characteristic among CPA1s (Masrati *et al.*, 2018). The eight highly conserved residues for CPA1 members are indicated in

Figure 4.6 with red arrows and Figure 4.7 with yellow circles; these residues are presented in *C. saccharoperbutylacetonicum* N1-4HMT.



EcCpa1

Score	Expect	Method	Identities	Positives	Gaps
263 bits(673)	3e-84	Compositional matrix adjust.	162/405(40%)	245/405(60%)	13/405(3%)

PaCpa1

Score	Expect	Method	Identities	Positives	Gaps
32.0 bits(71)	9e-05	Compositional matrix adjust.	32/130(25%)	61/130(46%)	2/130(1%)

Figure 4.6 Sequence alignment of *P. abyssi* NhaP, *C. saccharoperbutylacetonicum* Cpa1 and *E. coli* Cpa1
Sequence alignment of *C. saccharoperbutylacetonicum* Cpa1 (CsCpa1) vs *E. coli* Cpa1 (EcCpa1) and *P. abyssi* NhaP(PaCpa1). The numbers on the right refer to the protein residue numbers. (*) Indicates a single, fully conserved residue. (:) Indicates conservation between groups of strongly similar properties. (.) Indicates conservation between groups of weakly similar properties. (-) Indicates gaps in the sequence. Alignment obtained using Clustal Omega from EMBL-EBI. Red arrows indicate the eight highly conserved residues for CPA1 antiporters.

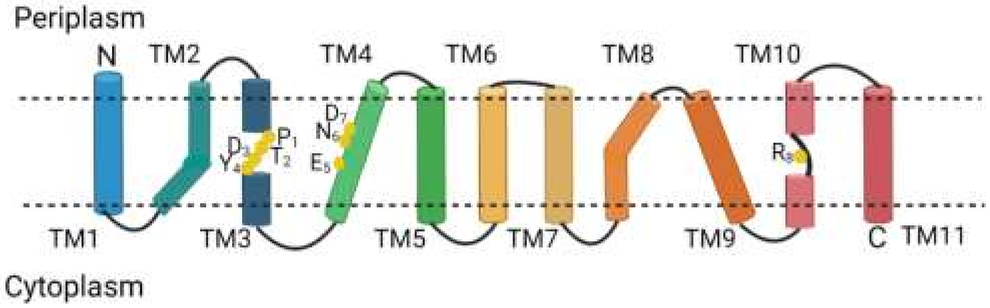


Figure 4.7 Schematic two-dimensional representation of the Cs-Cpa1
Based on the secondary and transmembranal prediction, it is believed that Cs-Cpa1 contains 11 transmembranes numbered TM-1 through TM-11. The amino terminal (N) is exposed to the exterior, and the carbon terminal is exposed to the interior. The yellow spheres are the motif residues. X_1-X_4 distinguish between different CPA groups within each of the 2 main clades. In the 5th position, CPA1 accommodates only glutamate. 6th and 7th comprise the ND motif, and the 8th position features arginine.

The absence of D_6 is a characteristic of an electroneutral antiporter, as D_6 and D_7 are considered the proton carriers, with D_7 being the primary carrier (Masrati *et al.*, 2018). CsNhaA contains both D_6 and D_7, whereas CsCpa1 only contains D_7; with this information, it can be inferred that CsNhaA is electrogenic (CPA2 member) because for 1 Na^+ exchanges 2 H^+ (DD), and CsCpa1 is electroneutral (CPA1 member) because for 1 Na^+ exchanges 1 H^+ (ND).

Both proteins CsNhaA and CsCpa1 were going to be looked at because despite CsNhaA looking more efficient for being electrogenic, it has been reported to be active at alkaline pH, not acidic (Padan, 2014, Padan *et al.*, 2004); whereas Cpa1 is more active in the acidic pH (Calinescu *et al.*, 2016).

Predicting the 3D structure helps develop further experiments such as mutagenesis (Breda *et al.,* 2007, Schwede, 2013). Gerchman *et al.* (1993) performed experiments by mutating histidine codons to arginine codons to identify the pH sensor of EcNhaA; this could be performed on CsCpa1 for the same purpose. Another experiment could be to test if CsCpa1 can be transformed into an electrogenic antiporter by changing the Asparagine for Aspartate like CsNhaA.

4.2.2.2 GlpF

The overexpression of Na^+/H^+ antiporters for proton extrusion can increase the osmotic pressure because of the increased levels of Na^+ ions within the cytoplasm. Clostridial GlpF may be able to alleviate osmotic stress; also, it has observed an apparent role in mitigating butanol stress which suggests it could have a protective role on the shift from acidogenic to the solventogenic phase. For these two reasons, GlpF was selected for further characterisation.

Osmoregulation in bacterial, plant, and animal cells requires the presence of membrane channels specific for water and small non-ionic solutes. Phylogenetic analysis revealed the existence of two subfamilies clusters: the aquaporins (AQPs) and glycerol facilitators (GLPs). All members of these two families share the same basic structure consisting of a tandem repeat, each of which was predicted to span the lipid bilayer three times (Figure 4.8). Their sequences share an internal repeat, two Asn-Pro-Ala (NPA motifs) near the centre (Stahlberg *et al.*, 2000). The two NPA motifs line the pore forming the "hourglass" model that works as a selectivity filter (Unger, 2000). The structure of GlpF crystallised as a symmetric homotetrameric arrangement of four channels with three glycerol molecules in each. Each tetramer is circled by a hydrophobic surface matching the dimensions of the lipid bilayer (Fu *et al.*, 2000). Overall, the structure of GlpF and AQP1 are very similar, although they have different environments created by their side chains and residues. In GlpF, the conserved Asn of the NPA motif provides hydrogen bonds during glycerol passage, but they are not involved in the initial discrimination of glycerol from other solutes. Trp48, Phe200 and Arg206 are the residues responsible for the selectivity. The Arg side chain forms hydrogen bonds with two of the glycerol's hydroxyl groups. At the same time, the carbon backbone packs tightly against a hydrophobic wedge formed by the two aromatic side chains, excluding other linear polyols while forcing the glycerol molecules to enter the selectivity filter in a single file (Unger, 2000).

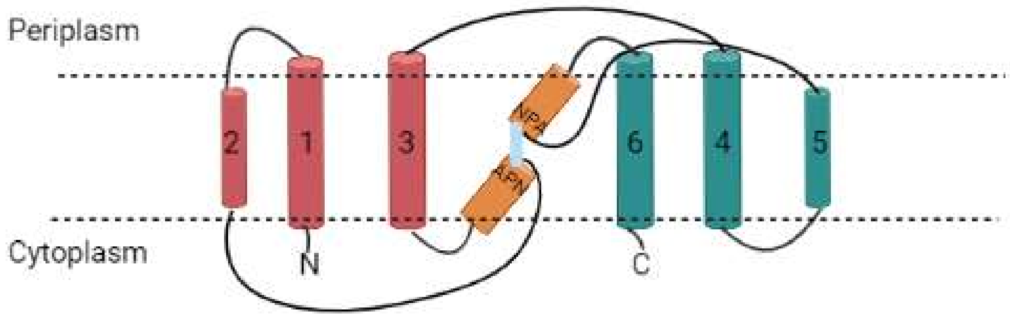
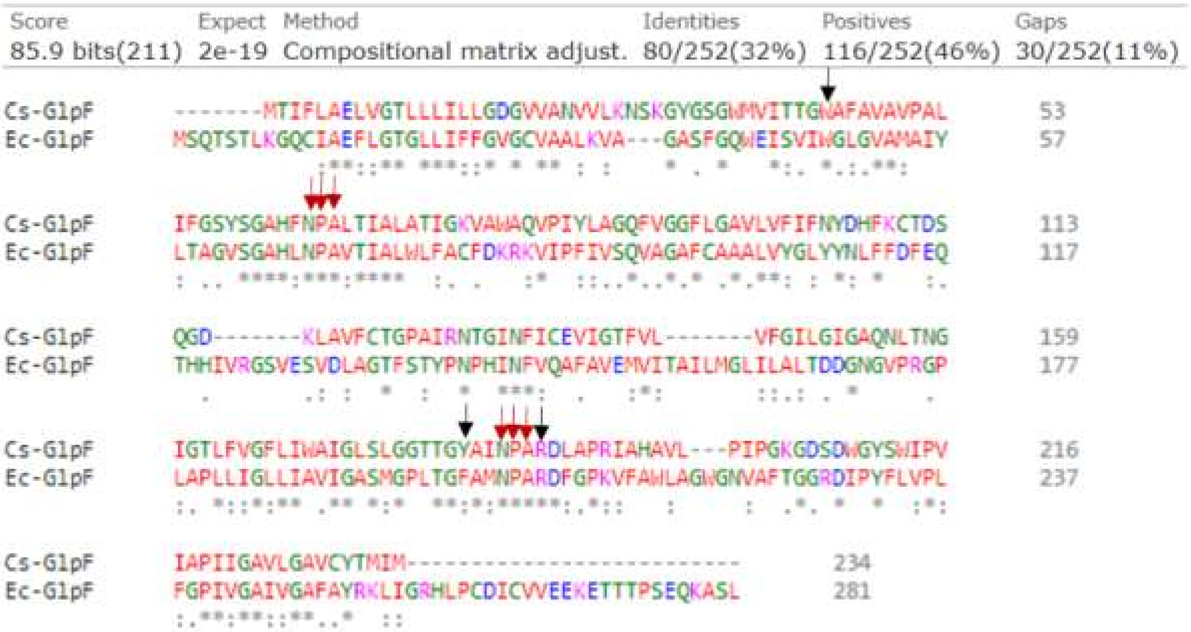


Figure 4.8 Schematic representation of aquaporin
The general structure is a tandem repeat of two three-helix bundles (TM 1-3 and TM 4-6) that are related by a pseudo-two-fold symmetry axis (marked blue) parallel to the plane of the membrane. A unique aquaporin feature is that the two-pore helices (oranges) form an additional transmembrane span. TM 2 and 5 need to be longer to traverse the membrane completely. The 'NPA motif' is the aquaporin family's functionally important signature sequence; at the end of each of the two-pore helices in the middle of the membrane and works as a selectivity filter. Created with biorender.com

The sequence for *C. saccharoperbutylacetonicum* N1-4HMT GlpF (CsGlpF) was obtained from Uniprot (https://www.uniprot.org/uniprotkb/M1MSM5/entry) and NCBI Reference sequence: WP_01539098.1. The *E. coli* GlpF (EcGlpF) was obtained doing a protein blast using the NCBI Blast tool NCBI reference sequence: WP_047608243.

E. coli (Ec) GlpF has been well characterised. In order to obtain the predicted conformation of *C. saccharoperbutylacetonicum* GlpF, first, an alignment between EcGlpF and CsGlpF using Clustal

Omega from EMBL-EBI (McWilliam *et al.*, 2013) was performed (Figure 4.9). Both sequences have a 32 % similarity meaning that the 3D model would be reliable (≥ 30 % is regarded as reliable). In Figure 4.9, the NPA motifs are indicated with red arrows, and the conserved residues characteristic of GlpF are indicated with black arrows. Interestingly the Phe200, an aromatic residue in charge of discriminating other solutes from glycerol, is changed for another aromatic residue, Tyr182. Tyrosine is also aromatic, partially hydrophobic and very similar to phenylalanine; the difference is that Tyr contains a reactive hydroxyl group in place of the *o*-hydrogen on the benzene ring (Betts and Russel, 2007).



Score	Expect	Method	Identities	Positives	Gaps
85.9 bits(211)	2e-19	Compositional matrix adjust.	80/252(32%)	116/252(46%)	30/252(11%)

```
Cs-GlpF     -------MTIFLAELVGTLLLILLGDGVVANVVLKNSKGYGSGWMVITTGNAFAVAVPAL     53
Ec-GlpF     MSQTSTLKGQCIAEFLGTGLLIFFGVGCVAALKVA---GASFGQWEISVIWGLGVAMAIY     57
                       :**::** ***::* * ** : :    * . *   *:. *.:.**:

Cs-GlpF     IFGSYSGAHFNPALTIALATIGKVAWAQVPIYLAGQFVGGFLGAVLVFIFNYDHFKCTDS     113
Ec-GlpF     LTAGVSGAHLNPAVTIALWLFACFDKRKVIPFIVSQVAGAFCAAALVYGLYYNLFFDFEQ     117
            : .. ****:***:****  :. .   :*  ::..*..*.* .*.**: : *: *   :.

Cs-GlpF     QGD-------KLAVFCTGPAIRNTGINFICEVIGTFVL-------VFGILGIGAQNLTNG     159
Ec-GlpF     THHIVRGSVESVDLAGTFSTYPNPHINFVQAFAVEMVITAILMGLILALTDDGNGVPRGP     177
             .       .: :  *  :  *  ***:  .    :*:        ::.: . *        .

Cs-GlpF     IGTLFVGFLIWAIGLSLGGTTGYAINPARDLAPRIAHAVL---PIPGKGDSDWGYSWIPV     216
Ec-GlpF     LAPLLIGLLIAVIGASMGPLTGFAMNPARDFGPKVFAWLAGWGNVAFTGGRDIPYFLVPL     237
            :. *::*:** .** *:*  **:*:*****:.*::    :     :  .*. *  *  :*:

Cs-GlpF     IAPIIGAVLGAVCYTMIM--------------------------     234
Ec-GlpF     FGPIVGAIVGAFAYRKLIGRHLPCDICVVEEKETTTPSEQKASL     281
            :.**:**::**..*  ::
```

Figure 4.9 Sequence alignment between *C. saccharoperbutylacetonicum* (Cs) and *E. coli* (Ec) GlpF sequences
Sequence alignment of *C. saccharoperbutylacetonicum* GlpF (CsGlpF) vs *E. coli* GlpF (EcGlpF). The numbers on the right refer to protein residue numbers. (*) Indicates a single, fully conserved residue. (:) Indicates conservation between groups of strongly similar properties. (.) Indicates conservation between groups of weakly similar properties. (-) Indicates gaps in the sequences. Alignment obtained by Clustal Omega from EMBL-EBI. The NPA motif is indicated with red arrows. The residues that are characteristic of GlpF are indicated with black arrows.

Next, using SWISS-MODEL from Expasy software, the predicted conformation was generated based on the experimentally determined protein structure of Ec-GlpF (pdb: 1fx8). The co-ordinate file from SWISS-MODEL was displayed using the molecular visualisation program of PyMol to generate good-quality images (Figure 4.10). The protein consists of 234 amino acid residues with the amino and carboxy terminus exposed to the cytoplasm. The model is arranged in 8 transmembrane segments.

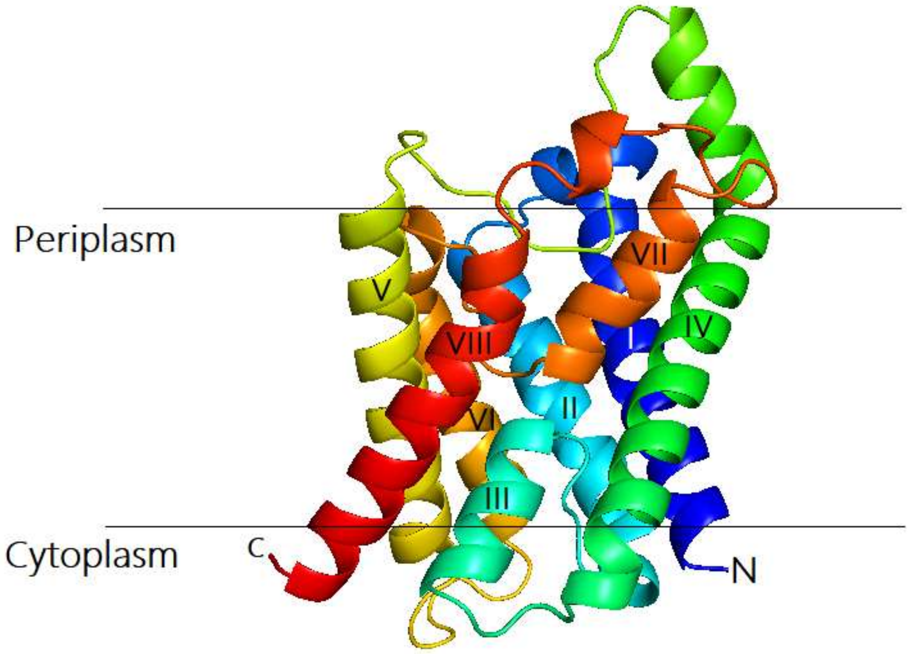


Figure 4.10 Overall architecture of *C. saccharoperbutylacetonicum* GlpF
Stereo view of a ribbon representation viewed parallel to the membrane (grey lines). The transmembrane domains are labelled with Roman numerals. Amino and carboxy termini are indicated by N, and C. Protein obtained using SWISS-MODEL and PyMol.

The Ramachandran plot, which visualises energetically favoured regions for backbone dihedral angles against amino acid residues in protein structures, gave a 93.72 % of favoured residues. Each point represents a protein residue, and the coloured regions represent favoured and allowed regions. A percentage of residues falling in the allowed areas higher than 90 % indicates that the structure is reliable. The CsGlpF structure obtained with the EcGlpF is reliable on that basis.

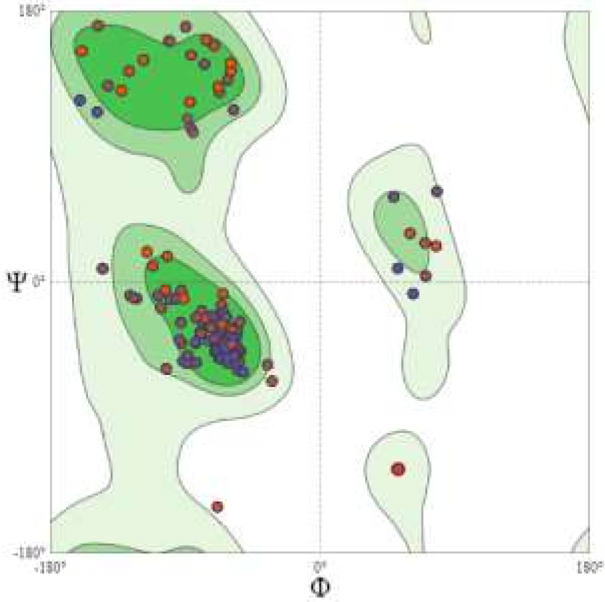


Figure 4.11 Ramachandran plot for the *C. saccharoperbutylacetonicum* N1-4HMT GlpF protein
Each red point represents a residue of the Cs GlpF protein. The plot obtained with SWISS-MODEL shows that most of the φ/ψ torsion angles were found in the favoured or additionally allowed regions (green coloured regions).

Figure 4.12 shows the superimposition of the two 3D models of *C. saccharoperbutylacetonicum* GlpF and *E. coli* GlpF, highlighting the NPA motif and the characteristic residues for GlpF using PyMol. The root mean square deviation (RMSD), which is the extent to which the main chains of the template and the target protein overlap if it is less than 0.5, means that the target and the template are essentially identical. For GlpF protein, the RSMD was 0.147, which determines that the quality of the model is very good.

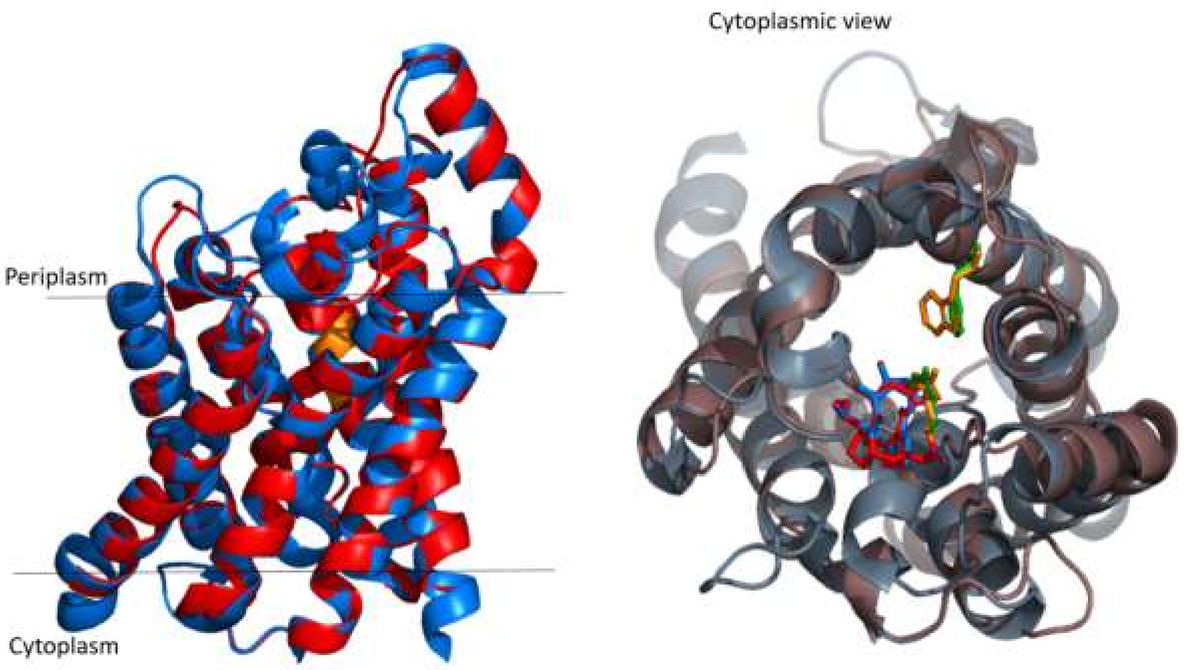


Figure 4.12 Comparative modelling of *E. coli* vs *C. saccharoperbutylacetonicum* GlpF
Comparison of *C. saccharoperbutylacetonicum* GlpF (in red) and *E. coli* GlpF (in blue) protein models. On the right side, the cytoplasmic view is shown with the NPA motif as red sticks for *Clostridium,* and blue for *E. coli.* The characteristic residues for GlpF are shown in orange sticks for *Clostridium,* and green for *E. coli*. The orange part that can be observed is the same orange residue on the right side. RMSD = 0.147. Made with PyMol.

Predicting the 3D structure is required to understand the protein's mechanism of activity and regulation (Padan *et al.*, 2004). It could be tested whether the change of tyrosine for phenylalanine in CsGlpF had an effect on the protein. Based on the sequence alignment, homology modelling and the validation of the homology modelling, GlpF modelling gave more significant values, so it can be inferred that role CsGlpF would be closer to EcGlpF since the design resembles the EcGlpF structure of pore-lining α-helices.

4.2.3 Expression and characterisation of NhaA and Cpa1 in *E. coli*

As a first step to obtaining purified proteins to characterise the function of the antiporters, the individual genes were cloned in pTZ19R/U. The vector pET17b and *E. coli* BL21 (DE3) were chosen for their expression because pET plasmids are medium copy vectors designed for cloning and expressing on *E. coli* proteins. The expression is induced by providing a source of T7 RNA polymerase (Novagen, 1999). BL21 (DE3) is a common laboratory strain for recombinant protein production; it does not contain the Ion protease nor the outer membrane protease OmpT, reducing the degradation of heterologous proteins. It carries the T7 RNA polymerase gene under the control of the *lac*UV5 promoter (Jeong *et al.*, 2015). The PBD records several membrane proteins produced in BL21 (DE3) and pET17b (Dilworth *et al.*, 2018). A brief schematic representation of the cloning strategy is represented in Figure 4.13. Based on their nucleotide sequence, three primers for each protein were designed (Table 2.3); one had two restriction enzymes to insert the gene in the cloning and expression plasmid. The second primer had a 3’ restriction enzyme, and the third primer was a 3’ with a His-tag for the C-terminal of the protein and one restriction enzyme. With these primers, the corresponding genes were amplified. Double digestion of the amplicons and the vector was performed to ligate them and transform DH5α *E. coli* cells. Ampicillin-positive colonies were chosen for double digestion and sequencing. The cloning strategy with and without the His-Tag was performed separately. This strategy was chosen because having a his-tag on the C-terminal would facilitate the purification through an affinity column and their identification with western blot and the anti-his-tag antibody kit. Because the his-tag might cause changes in the membrane folding, it was also decided to express the protein without his-tag, potentially obtaining it in a more native conformation.

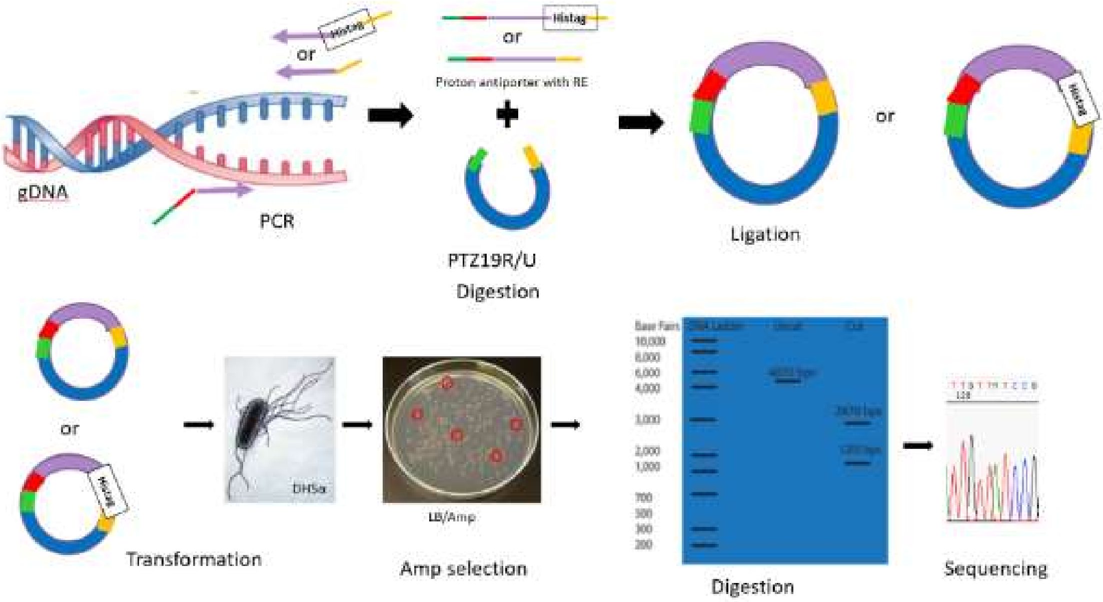
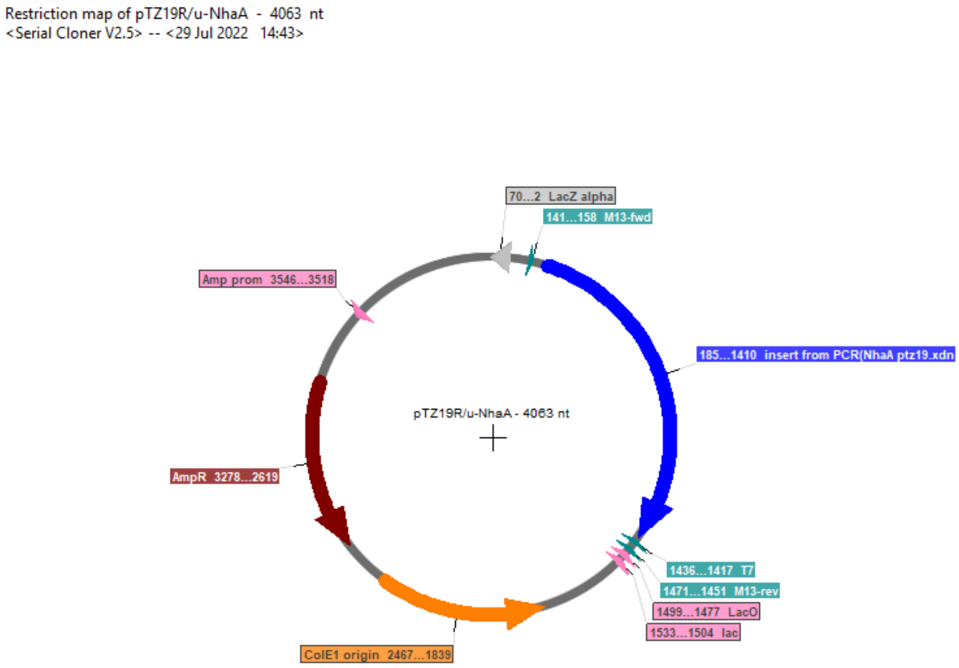

Histag
or
gDNA
PCR
Histag
or
Proton antiporter with RE
PTZ19R/U
Digestion
Ligation
or
Histag
or
Histag
Transformation
DH5α
LB/Amp
Amp selection
Base Pairs
10,000
8,000
6,000
4,000
3,000
2,000
1,000
700
500
300
200
Digestion
Sequencing

Figure 4.13 Cloning strategy.
Schematic representation of the cloning strategy of proton antiporters. The three primers are represented in purple arrows. RE = the restriction enzymes represented in green, red and yellow lines and rectangles. The box written as Histag is the hexahistidine tag. Amp = ampicillin.

The amplification, digestion and ligation were virtually performed using Serial Cloner v2.6 and the schematic representation of the pTZ19R/U-NhaA-His (top) and pTZ19R/U-Cpa1 (bottom) plasmid are depicted in Figure 4.14.


Restriction map of pTZ19R/u-NhaA - 4063 nt
<Serial Cloner V2.5> -- <29 Jul 2022 14:43>
pTZ19R/u-NhaA - 4063 nt
70...2 LacZ alpha
141...158 M13-fwd
185...1410 insert from PCR(NhaA ptz19.xdn
Amp prom 3546...3518
AmpR 3278...2619
ColE1 origin 2467...1839
1436...1417 T7
1471...1451 M13-rev
1499...1477 LacO
1533...1504 lac

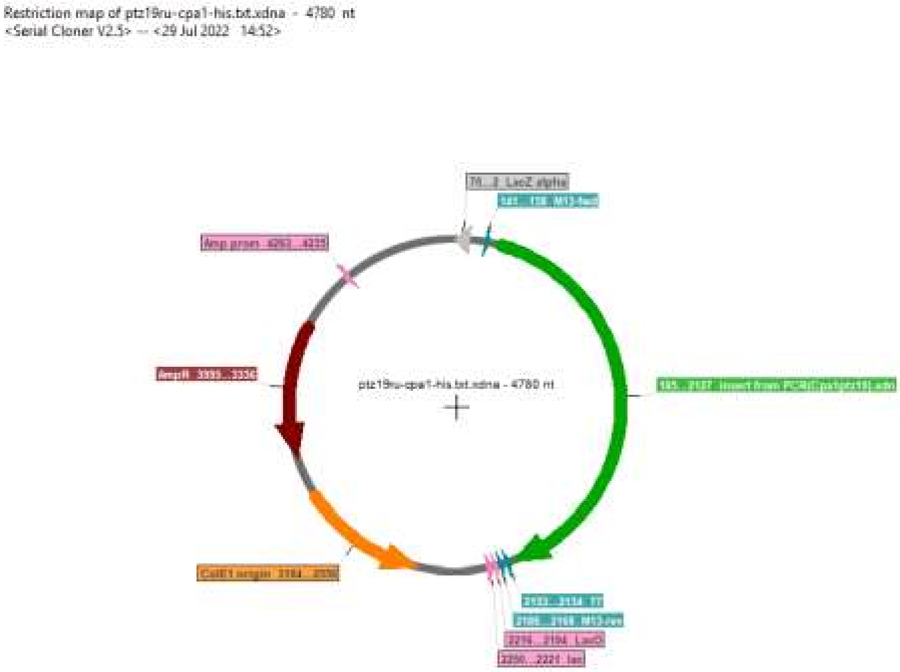

Restriction map of ptz19ru-cpa1-his.txt.xdna - 4780 nt
<Serial Cloner V2.5> -- <29 Jul 2022 14:52>
ptz19ru-cpa1-his.txt.xdna - 4780 nt
70...2 LacZ alpha
141...158 M13-fwd
185...2127 insert from PCR
Amp prom 4263...4235
AmpR 3995...3336
ColE1 origin 3184...2556
2153...2134 T7
2188...2168 M13-rev
2216...2194 LacO
2250...2221 lac

Figure 4.14 Schematic representation of the pTZ19R/U-NhaA-His and pTZ19R/U-Cpa1-His
The genes *nhaA* and *cpa1* from *C. saccharoperbutylacetonicum* N1-4HMT were virtually amplified on Serial Cloner V 2.6 with the designed primers, double digested with HindIII and BamHI and ligated into pTZ19R/U plasmid, the final construct is shown on the figure. *nhaA* gene is in blue, and the *cpa1* gene is in green. The origin of replication is in orange. Ampicillin resistance gene in red. LacZ alpha gene in grey. Promoters are in pink. M13 Forward and reverse sequences are in aqua.

4.2.3.1 Gene amplification and PTZ19R/U ligation

Both genes were amplified by PCR using the appropriate primer pairs (Table 2.3). *cpa1* had an expected size of 1917 bp and *nhaA* of 1203 bp. All products were of the expected size (Figure 4.15 and Figure 4.16). dH_2O and pET17b-GlpF were used for the negative control to confirm that the primers were specific to the *cpa1* gene.

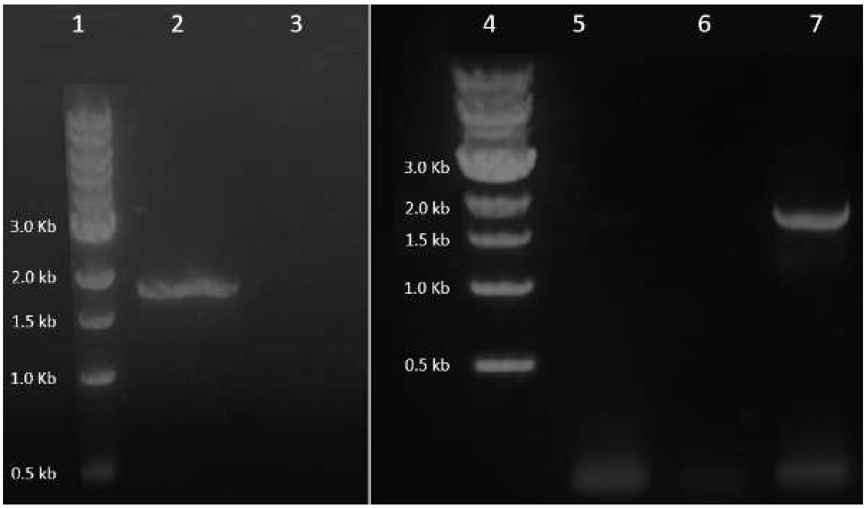


Figure 4.15 PCR products from the *cpa1* gene
PCR amplification of the *cpa1* gene from *C. saccharoperbutylacetonicum* N1-4HMT was performed using the specific primers with and without His tag. 5 µL of PCR products were run on 1% agarose gel. Lane 1 DNA ladder. 2) c*pa1* with His tag. 3) Empty lane. 4) DNA ladder. 5) pET17b-GFP plasmid for negative control. 6) dH_2O negative control. 7) c*pa1* without His tag.

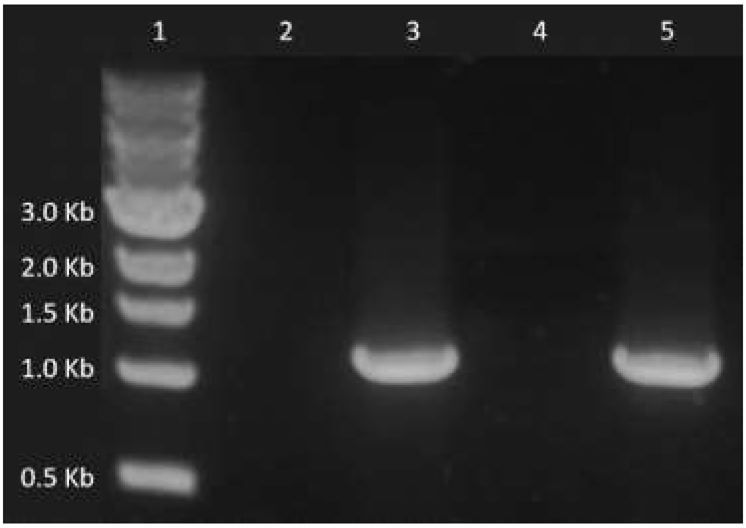


Figure 4.16 PCR products from the *nhaA* gene
PCR amplification of the *nhaA* gene from *C. saccharoperbutylacetonicum* N1-4HMT was performed using the specific primers with and without His tag. 5 uL of PCR products were run on 1% agarose gel. Lane 1. DNA ladder. 2) dH_2O negative control. 3) *nhaA* without His-tag. 4) dH_2O negative control. 5) *nhaA* with His tag

After PCR amplification, the fragments were purified, double digested (*Hind*III and *BamH*I) and ligated into pTZ19R digested with the same enzymes. *E. coli* DH5α cells were transformed with these ligation mixtures. Transformants were selected based on ampicillin resistance. To establish whether the plasmid had the desired *cpa1* or *nhaA* insert, double digestions (*Hind*III and *BamH*I) were undertaken from a random selection of colonies following plasmid isolation (Figure 4.17). Plasmids from positive colonies: 5 (*cpa1*-Histag colony 3), 9 (*nhaA*-Histag colony 2), and 12 (*nhaA* colony 2), which contained an insert of the expected size were sent for sequencing to confirm that the plasmids carried the desired insert (Appendix).

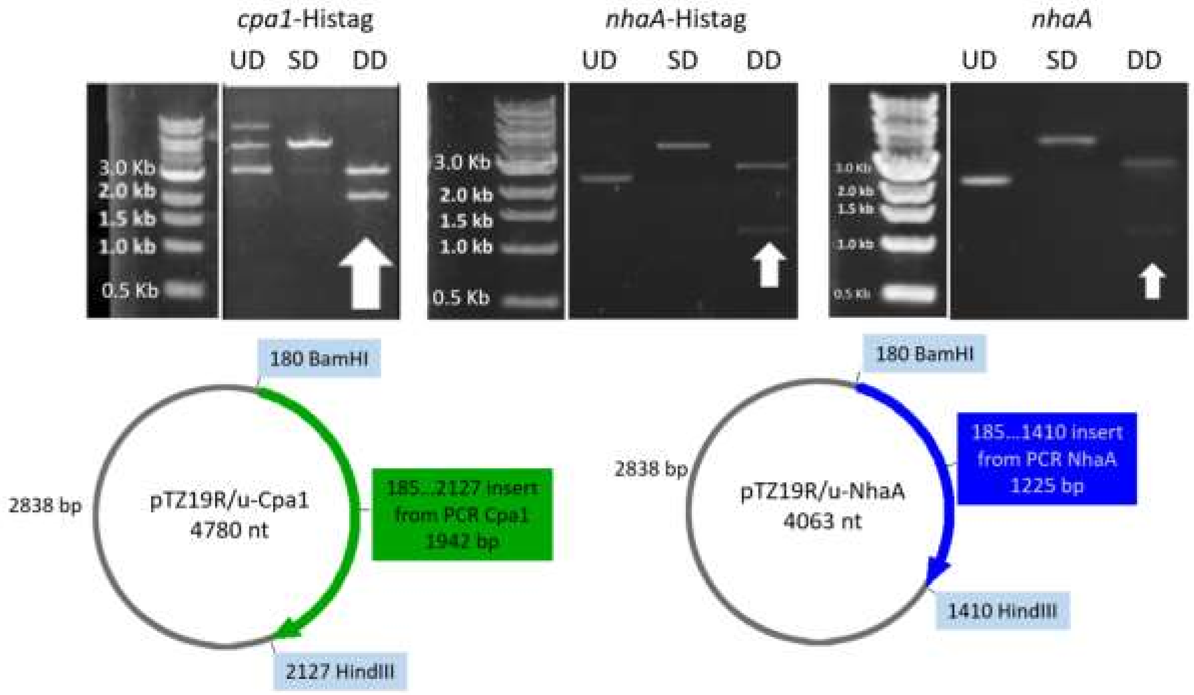
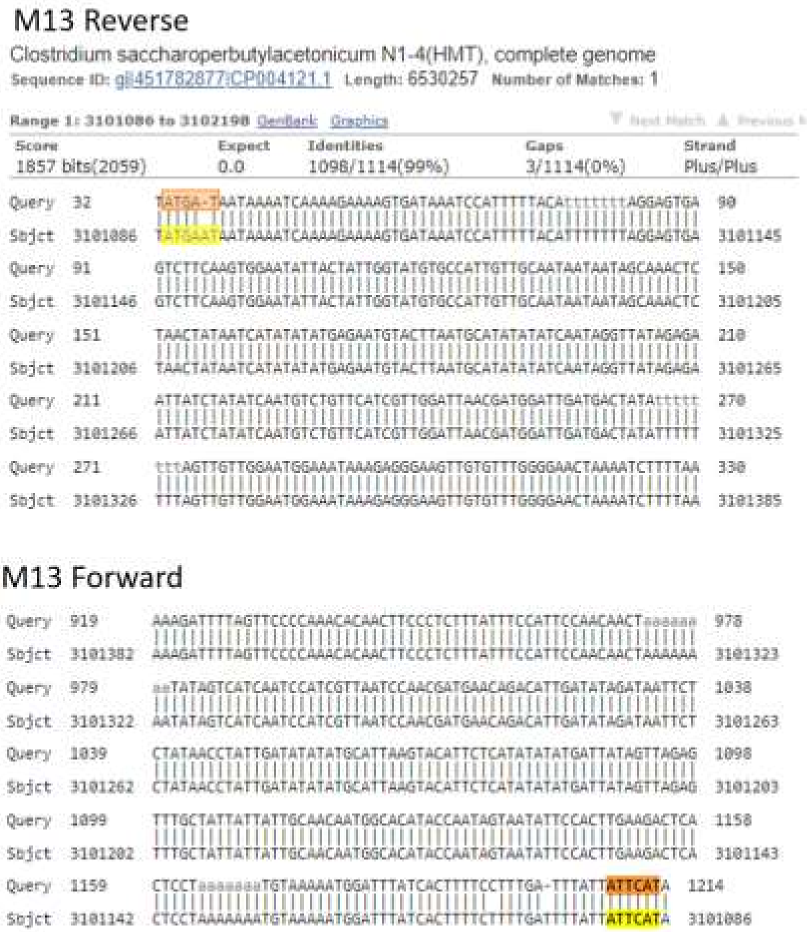

cpa1-Histag
UD SD DD
3.0 Kb
2.0 kb
1.5 kb
1.0 kb
0.5 Kb
nhaA-Histag
UD SD DD
3.0 Kb
2.0 kb
1.5 kb
1.0 kb
0.5 Kb
nhaA
UD SD DD
3.0 Kb
2.0 kb
1.5 kb
1.0 kb
0.5 Kb
180 BamHI
2838 bp
pTZ19R/u-Cpa1
4780 nt
185...2127 insert from PCR Cpa1 1942 bp
2127 HindIII
180 BamHI
2838 bp
pTZ19R/u-NhaA
4063 nt
185...1410 insert from PCR NhaA 1225 bp
1410 HindIII

Figure 4.17 Double digestion from a random selection of colonies
Ampicillin-positive colonies were digested with *Hind*III only (SD) and double digested with *Hind*III and *BamH*I (DD). 5 µL of the digestion were run on 1 % agarose gel. The positive colonies of *cpa1*-Histag, *nhaA*-Histag and *nhaA* are shown. The white arrows indicate the inserts of the expected size. On the bottom, the plasmid maps and the restriction sites are shown.

A BLAST alignment (Altschul *et al.*, 1997) using sequencing data obtained from SourceBioscience was also performed to confirm the sequence of the plasmids (Figure 4.18). The sequencing was done with the M13 forward and reverse primers. Only *nhaA* is exemplified; in both cases, it gave 99 % similarity; however, the missing nucleotide on the reverse looked correct in the forward and vice versa. The chromatogram was also checked. It was assumed the errors were normal since they appeared at the beginning or at the end of the chromatogram, where the signal contained much noise. Therefore the cloned genes were considered to be correct.



M13 Reverse

Clostridium saccharoperbutylacetonicum N1-4(HMT), complete genome

Sequence ID: gi|451782877|CP004121.1 Length: 6530257 Number of Matches: 1

Range 1: 3101086 to 3102198 GenBank Graphics

```
Score            Expect   Identities        Gaps         Strand
1857 bits(2059)  0.0      1098/1114(99%)    3/1114(0%)   Plus/Plus

Query  32       TATGA-TAATAAAATCAAAAGAAAAGTGATAAATCCATTTTTACAtttttttAGGAGTGA  90
                ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101086  TATGAATAATAAAATCAAAAGAAAAGTGATAAATCCATTTTTACATTTTTTTAGGAGTGA  3101145

Query  91       GTCTTCAAGTGGAATATTACTATTGGTATGTGCCATTGTTGCAATAATAATAGCAAACTC  150
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101146  GTCTTCAAGTGGAATATTACTATTGGTATGTGCCATTGTTGCAATAATAATAGCAAACTC  3101205

Query  151      TAACTATAATCATATATATGAGAATGTACTTAATGCATATATATCAATAGGTTATAGAGA  210
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101206  TAACTATAATCATATATATGAGAATGTACTTAATGCATATATATCAATAGGTTATAGAGA  3101265

Query  211      ATTATCTATATCAATGTCTGTTCATCGTTGGATTAACGATGGATTGATGACTATAttttt  270
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101266  ATTATCTATATCAATGTCTGTTCATCGTTGGATTAACGATGGATTGATGACTATATTTTT  3101325

Query  271      tttAGTTGTTGGAATGGAAATAAAGAGGGAAGTTGTGTTTGGGGAACTAAAATCTTTTAA  330
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101326  TTTAGTTGTTGGAATGGAAATAAAGAGGGAAGTTGTGTTTGGGGAACTAAAATCTTTTAA  3101385
```

M13 Forward

```
Query  919      AAAGATTTTAGTTCCCCAAACACAACTTCCCTCTTTATTTCCATTCCAACAACTaaaaaa  978
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101382  AAAGATTTTAGTTCCCCAAACACAACTTCCCTCTTTATTTCCATTCCAACAACTAAAAAA  3101323

Query  979      aaTATAGTCATCAATCCATCGTTAATCCAACGATGAACAGACATTGATATAGATAATTCT  1038
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101322  AATATAGTCATCAATCCATCGTTAATCCAACGATGAACAGACATTGATATAGATAATTCT  3101263

Query  1039     CTATAACCTATTGATATATATGCATTAAGTACATTCTCATATATATGATTATAGTTAGAG  1098
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101262  CTATAACCTATTGATATATATGCATTAAGTACATTCTCATATATATGATTATAGTTAGAG  3101203

Query  1099     TTTGCTATTATTATTGCAACAATGGCACATACCAATAGTAATATTCCACTTGAAGACTCA  1158
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3101202  TTTGCTATTATTATTGCAACAATGGCACATACCAATAGTAATATTCCACTTGAAGACTCA  3101143

Query  1159     CTCCTaaaaaaaTGTAAAAATGGATTTATCACTTTTCCTTTGA-TTTATTATTCATA  1214
                ||||||||||||||||||||||||||||||||||||| ||||| |||||||||||||
Sbjct  3101142  CTCCTAAAAAAATGTAAAAATGGATTTATCACTTTTCTTTTGATTTTATTATTCATA  3101086
```

Figure 4.18 Sequence alignment of *nhaA*

The image represents the first nucleotides of the *nhaa* aligned with the sequence obtained with M13 primers forward and reverse to confirm the sequence of the colonies. Query is the nucleotides of the *nhaA* sequence encoded by the positive double digested colony. The nucleotide sequence of *C. saccharoperbutylacetonicum* N1-4HMT is Sbjct. The alignment gives a 99% similarity between sequences.

4.2.3.2 *Ligation into expression vector pET17b*

In order to express and purify the membrane transporters, constructs were inserted into pET17b vectors. Plasmids containing an insert of the expected size (1917 bp for *cpa1* and 1203 bp for *nhaA*) were purified and double digested with *Nhe*I and *BamH*I for *cpa1* and *Nde*I and *BamH*I for *nhaA* and ligated into double digested *Nhe*I, *BamH*I or *Nde*I, *BamH*I pET17b vectors. *E. coli* BL21 was transformed and selected by ampicillin resistance. Positive colonies were screened with double digestion. No positive colony with an insert of the expected size was found. Several attempts were performed, trying different vector:insert ratios, ligation product, and new enzymes, and in none of them a positive colony was recovered.

Membrane proteins are naturally expressed at low levels. Therefore, the overexpression of membrane proteins can be toxic to the cells. Also, the space in the membrane is limited compared with the proteins expressed in the cytosol or secreted. Wu *et al.,* (2013) observed that overexpression of *nhaA* or *nhaR* with a strong promoter induced lethality. In this project, no positive colonies of *E. coli* cells transformed with the pET17b and the proton antiporter were obtained. Inducible expression systems can solve this toxicity (Hardy *et al.,* 2016). pET systems contain components of the T7 promoter. *E. coli* BL21 strains had inserted into their bacterial chromosome the T7 RNA polymerase under the control of the inducible lac UV5 operon, where transcription is induced by IPTG and should be absent without an inducer. However, low recombinant protein expression is observed during bacterial growth without induction, a phenomenon known as leaking (Briand *et al.,* 2016). This is because efficient transcription initiation requires the presence of cyclic AMP (cAMP), which is strongly influenced by the carbon source present in the medium. Mediums free of glucose forced cells to use different carbon sources such as glycerol, increasing cAMP levels and inducing the *lac* operon (Novy and Morris, 1999).

4.2.3.3 Re-defining strategy. Amplification of genes and ligation into pET15b and pBR322

Due to the challenges above, two different cloning strategies were chosen. The first was to use the expression vector pET15b because this plasmid contains a *lac* operator sequence and the coding sequence for the *lac* repressor (*lacI)*. The *lac* repressor acts at the *lacUV5* in the host chromosome and the T7*lac* promoter in the vector, repressing the transcription of the T7 RNA polymerase and eliminating the leakiness (Novagen 1992). A second strategy was to follow the Wu *et al.* (2013b) strategy; use their endogenous *nhaA* promoter in the promoterless pBR322. For both strategies, it was also decided to increase the histidine tag to 12 histidines (primers Table 2.3) to enable effective downstream processing with SMALPs which show reduced interaction with Ni^{2+} IMAC columns (Broecker *et al.*, 2017). Amplification of both genes was performed (Figure 4.19). *cpa1* plus its promoter had an expected size of 2104 bp, and *nhaA* plus its promoter 1501 bp. *cpa1* gave a band of the expected size.

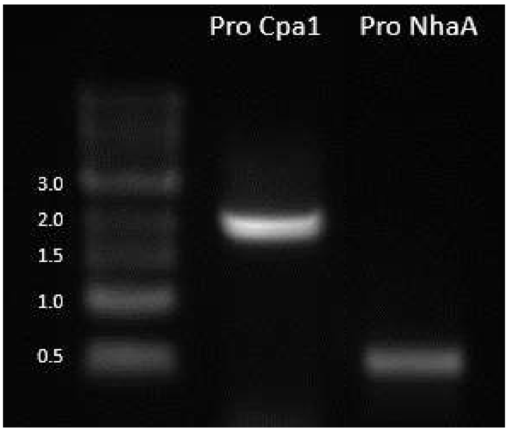


Figure 4.19 PCR products of *cpa1* and *nhaA* with their promoters

PCR amplification of the *cpa1* and *nhaA* genes from *C. saccharoperbutylacetonicum* N1-4HMT was performed using specific primers. 5 µL of PCR products were run on 1 % agarose gel. Lane 1 is the DNA ladder. Lane 2: *cpa1* plus its promoter (Pro Cpa1). Lane 3: *nhaA* plus its promoter (Pro NhaA).

nhaA has an approximate size of 1203 bp; *nhaA* with its promoter (Pro-NhaA) has an approximate size of 1464 bp. Pro-NhaA PCR did not show bands of the expected size. Primer annealing is a critical step in the PCR, as is the step where the primers bind to the target DNA for amplification. A temperature too high can reduce the efficiency since primer annealing is reduced. On the other hand, a low temperature can also reduce the yield of the desired product since primers can partially anneal to sequences other than the intended target. As recommended by the DNA polymerase providers, if there is no or low product, a decrease in the annealing temperature (Rychlik *et al.*, 1990) and an increase of genomic DNA template are needed for the PCR optimisation. It was decided to reduce the annealing temperature of the primers to 50 ˚C, although this will increase nonspecific amplification of the primers and to increase the concentration of the gDNA (Figure 4.20).

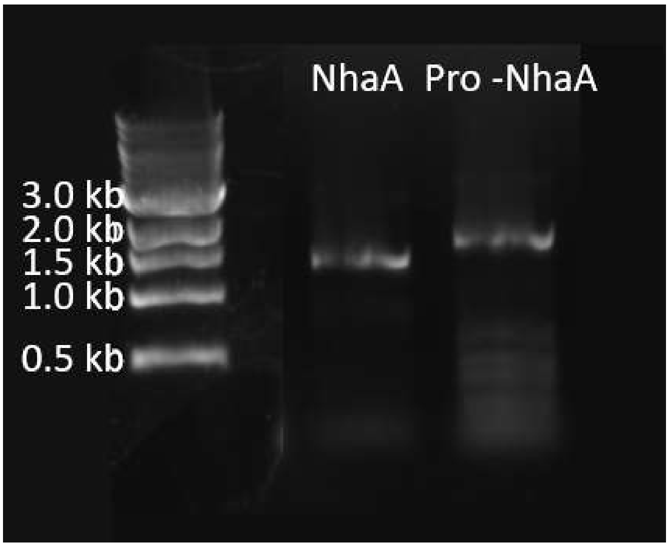


Figure 4.20 PCR products of NhaA and promoter-NhaA
PCR amplification of the *nhaA* and promoter-*nhaA* gene from *C. saccharoperbutylacetonicum* N1-4HMT was performed using specific primers. 5 µL of PCR products were run on 1 % agarose gel. Lane 1 is the DNA ladder. Lane 2: *nhaA* gene (NhaA). Lane 3: Pro- NhaA is the *nhaA* gene plus its promoter.

The same strategy was repeated to obtain the *cpa1* gene without the promoter, to insert it in pET15b. Different concentrations of gDNA were tested, and dimethyl sulfoxide (DMSO) was added to the PCR mix (Figure 4.21). DMSO is an additive that can enhance the separation of the double-stranded DNA during the denaturation step and promote specificity.

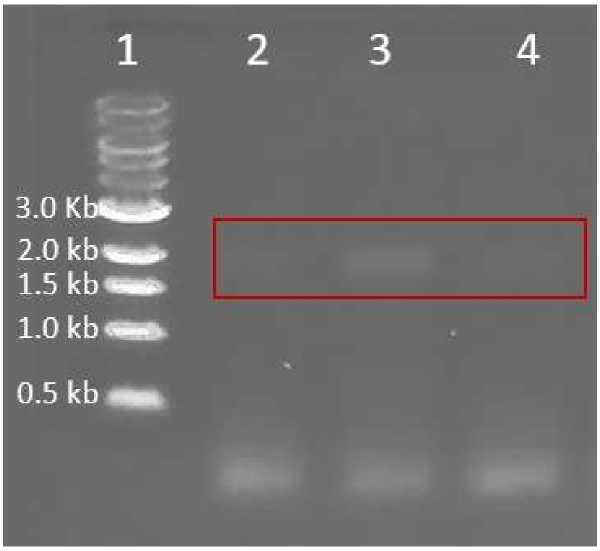


Figure 4.21 PCR products of *cpa1* from different concentrations of gDNA
PCR amplification of the *cpa1* gene from *C. saccharoperbutylacetonicum* N1-4HMT was performed using specific primers. Different concentrations of gDNA were tested. 1. DNA ladder. 2. 1 µL of gDNA. 3. 2 µL of gDNA. 4. 4 µL of gDNA. 5 µL of PCR products were run on 1 % agarose gel. The red box indicates the area where the expected band size is.

cpa1 has an expected size of 1953 bp. Very faintly, a band of the expected size in all the lanes can be observed, but the one observed with a higher concentration was lane 3, where 2 µL of gDNA were used. Once all *nhaA* and *cpa1* without the promoter were amplified, they could be inserted in pET15b plasmids. Therefore, pET15b plasmid, *nhaA* and *cpa1* products were double digested with *Nco*I and *BamH*I (Figure 4.22).

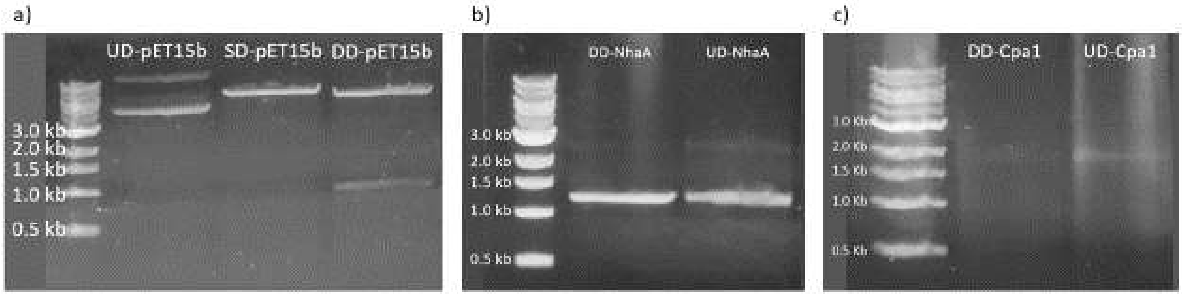


Figure 4.22 Double digestion of pET15b, *nhaA* and *cpa1*
PET15b plasmid, NhaA and Cpa1 PCR products were double digested with *Nco*I and *Bam*HI. a) pET 15b. b) NhaA. c) Cpa1. UD = undigested. SD = single digested. DD = double digested. 20 µL of the digested products were run on 1 % agarose gel.

The plasmid pET15b was obtained from an already made pET155b-Ec-GlpF-myc which contains the *E. coli* GlpF gene and has an expected size of 983 bp, and the insert has been removed from the plasmid in the double digestion of the plasmid (Figure 4.22.a) DD-pET15b). Because only 10 bp are removed from the PCR products by digestion, no difference can be observed between the size of the undigested (UD) and double digested (DD) (Figure 4.22. b) and c)). However, based on the results obtained from the double digestion of the plasmid, it is assumed that it also worked for the PCR products. The double digestion of *cpa1* was very faint; therefore, the whole volume was used for the following steps. A ligation of pET15b-NhaA and pET15b-Cpa1 was performed, followed by a transformation of *E.coli* BL21 with the ligated plasmids. No colonies were obtained from the ligation.

pBR322 plasmid, promoter-NhaA and promoter-Cpa1 PCR products were double digested with *Hind*III and *BamH*I to be ligated (Figure 4.23). One of the buffers used for the agarose gel must have been contaminated since the agarose gel does not look clear under UV light. From the gel image, much noise is observed. It is complicated to determine if the vector's double digestion worked as the cut band of ~ 350 bp cannot be observed. All the reagents used for the DNA gel and electrophoresis buffer were re-made from fresh to repeat the double digestion.

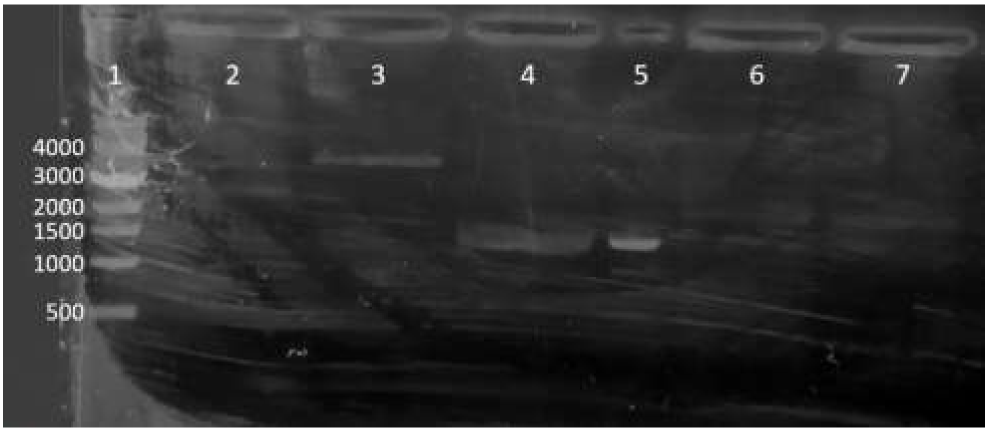


Figure 4.23 Double digestion of pBR322, promoter-NhaA and promoter-Cpa1
pBR322 plasmid, promoter-NhaA and promoter-Cpa1 PCR products were double digested with *Hind*III and *BamH*I. 1. DNA Ladder. 2. Undigested pBR322. 3. Double digested pBR322. 4. Double digested promoter-NhaA. 5. Undigested promoter-NhaA. 6. Double digested promoter-Cpa1. 7. Undigested promoter-Cpa1. 20 µL of the digested products were run on 1 % agarose gel.

The remainder of the work to obtain the purified *C. saccharoperbutylacetonicum* N1-4(HMT) NhaA and Cpa1 were not possible due to the COVID-19 pandemic preventing laboratory access.

4.2.4 Expression and characterisation of pET17b-CsGlpF in *E. coli*

From an already constructed and confirmed pET17b-CsGlpf-Histag plasmid, 50 µL of *E. coli* BL21 were transformed and selected by ampicillin resistance. Positive colonies were screened with double digestion using the *Nde*I and *Hind*III (Figure 4.24). The three colonies screened gave an insert of the expected size.

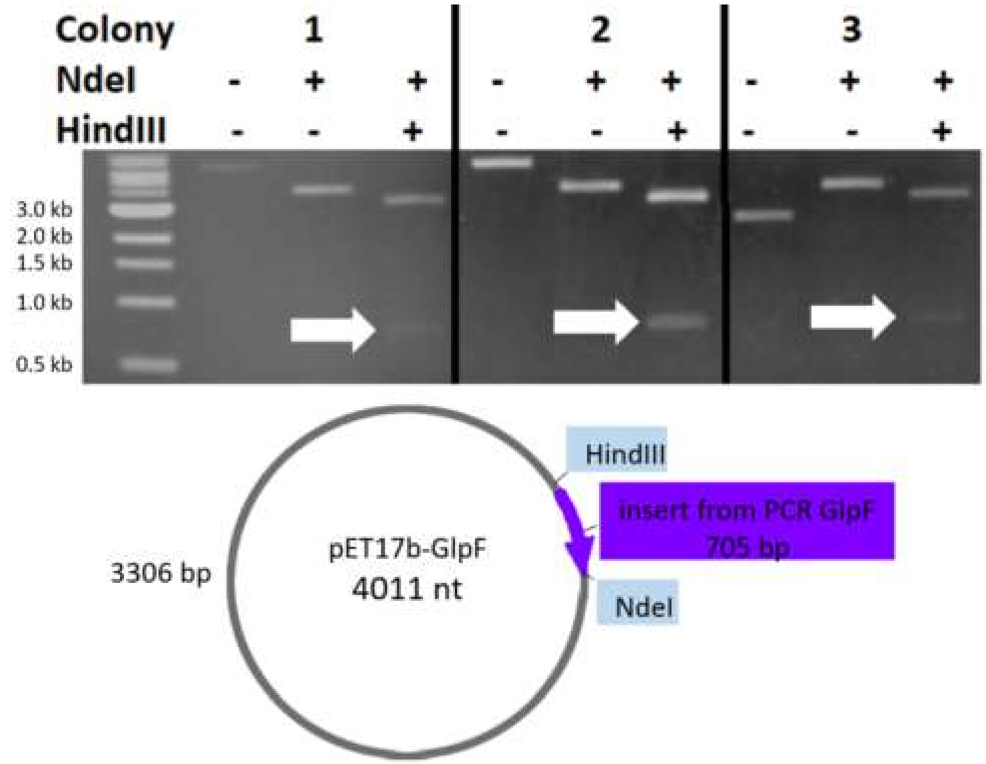


Figure 4.24 pET17-CsGlpF Double digestion from random colonies
Ampicillin-positive colonies were digested with *Nde*I and double digested with *Nde*I, and *Hind*III. 5 µL of the digestion were run on 1 % agarose gel. 1. pET17b-CsGlpF-Histag colony 1. 2. pET17b-CsGlpF-Histag colony 2. 3. pET17b-CsGlpF-Histag colony 3. The white arrows indicate inserts of the expected size. The plasmid map and the restriction sites are also shown.

After the cells were transformed and confirmed, expression of CsGlpF was induced. pET17b contains a bacteriophage T7 promoter, and BL21 *E. coli* cells contain a T7 RNA polymerase regulated by a *lac* operon which can be induced with IPTG. First, different temperatures and IPTG concentrations were screened to optimise protein expression. CsGlpF was expressed in 5 mL of LB medium at room temperature, 27 °C and 37 °C for 2, 4 and 24 hrs with 0, 0.25, 0.5, 0.75, and 1 mM of IPTG. After the induction, the pellets were recovered and frozen. The following day the pellets were resuspended in a 6 x loading buffer, boiled at 90 °C for 10 min and run on a 12 % SDS-PAGE gel (Figure 4.25). Membrane proteins aggregate when boiled; this step was eliminated from further experiments. Because no DNAse was added, all the samples ran with a smear; nevertheless, two conditions were chosen. At 37 ˚C for 4 h without IPTG because a band was clearly seen (top gel Figure 4.25) and to observe the effect of temperature and IPTG concentration on CsGlpF expression, the condition of 4 h at room temperature with 0.75 mM were also chosen.

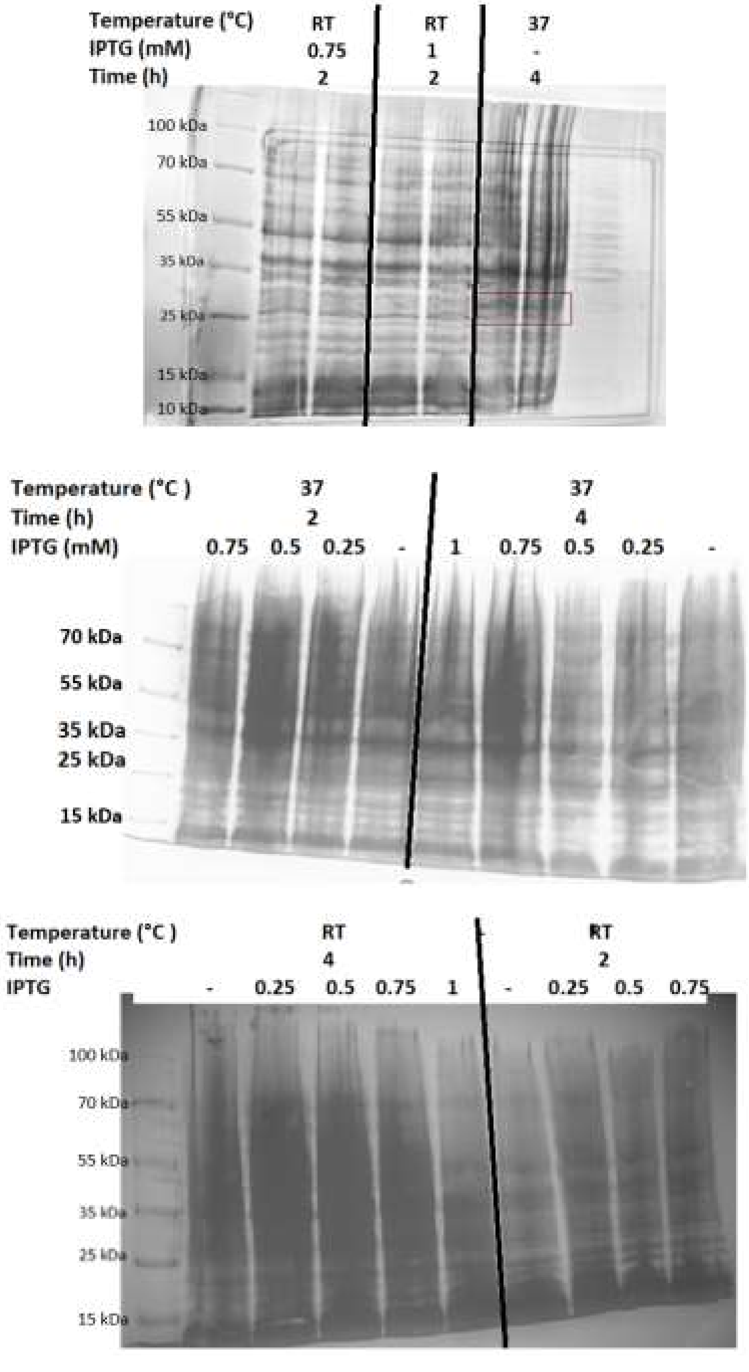

Temperature (°C)
IPTG (mM)
Time (h)
RT
0.75
2
RT
1
2
37
-
4
100 kDa
70 kDa
55 kDa
35 kDa
25 kDa
15 kDa
10 kDa
Temperature (°C)
Time (h)
IPTG (mM)
37
2
0.75
0.5
0.25
-
37
4
1
0.75
0.5
0.25
-
70 kDa
55 kDa
35 kDa
25 kDa
15 kDa
Temperature (°C)
Time (h)
IPTG
RT
4
-
0.25
0.5
0.75
1
RT
2
-
0.25
0.5
0.75
100 kDa
70 kDa
55 kDa
35 kDa
25 kDa
15 kDa

Figure 4.25 SDS-PAGE CsGlpF induction
CsGlpF expression was induced at different temperatures, IPTG concentrations and induction times. 20 µL of the boiled samples were loaded on a 12 % SDS-PAGE gel. RT=room temperature. The red box indicates the bands that have the expected size of CsGlpF.

CsGlpF expression was induced in a 5 mL culture under the conditions described above. Membrane fractions were extracted with a French press, solubilised with SMA, and affinity chromatography was performed for their purification. A Western blot with an anti-His antibody was performed to detect the His-tag on the samples (Figure 4.26). Different bands of 62, 49 and 38 kDa were observed in the 1st and 2nd elution fractions (lanes 5 and 6, red rectangles) at 37 °C without IPTG, suggesting that

CsGlpF is in different oligomeric forms. As reported in section 4.2.2.2, all aquaporins assemble into stable tetramers *in vivo* and *in vitro*; tetramerisation might contribute to increased *in vivo* stability (Klein *et al.*, 2019). Observation of GlpF oligomeric forms agrees with what has been reported. No other bands were observed in the rest of the lanes, suggesting that GlpF was soluble and purified.

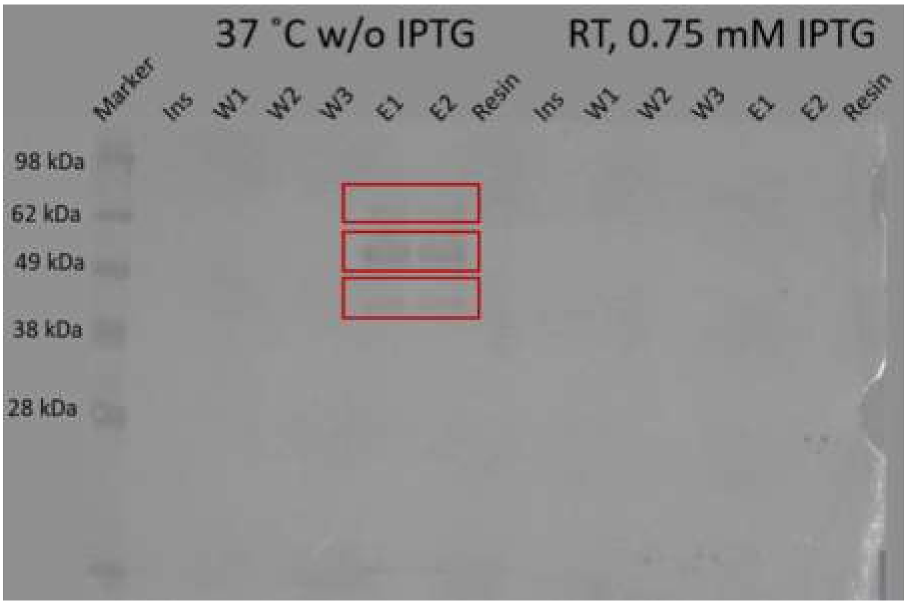


Figure 4.26 Western blot of CsGlpF His-tag
10 µL of sample fractions were separated by a 12 % SDS-PAGE and analysed by Western blot using antibodies directed against His-tag. 37 ˚C w/o IPTG: induction at 37 °C without IPTG. RT, 0.75 mM IPTG: induction at room temperature with 0.75 mM IPTG. Ins: insoluble fraction. W1: 1st wash. W2: 2nd wash. W3: 3rd wash. E1: 1st elution. E2: 2nd elution.

No expression of GlpF was observed at Room Temperature and 0.75 mM of IPTG. The expression of GlpF with the two conditions (RT and 0.75 mM of IPTG and 37 ˚C without IPTG) was scaled up to 1 L of culture (Figure 4.27, Figure 4.28*)*. The SDS-PAGE is shown in Figure 4.27, where the ~28 kDa bands of the insoluble fractions were seen. It is unclear, but the band observed at ~49 kDa might only be observed on the insoluble fractions. A band below 49 kDa was observed in the first elution at 37 ˚C (3rd lane), which might be a contaminant since it was not observed in the Western blot. In the Western blot (Figure 4.28), a band of about ~49 kDa was observed in all lanes suggesting that GlpF was in an oligomeric form (red arrow). Also, bands below ~28 kDa were observed in the insoluble fractions at both conditions. The difference between what is observed in the SDS-PAGE and a Western is because Western blots are more sensitive than SDS-PAGE; therefore, less amount of protein can be observed by Western blot.

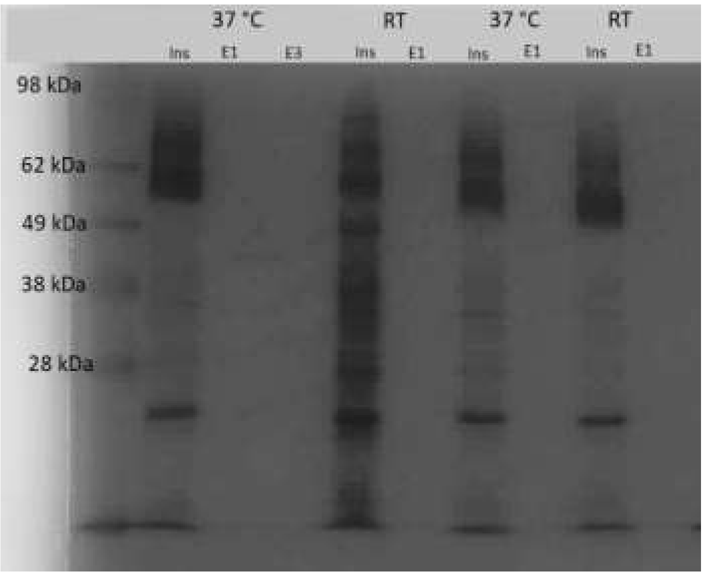

37 °C
RT
37 °C
RT
Ins
E1
E3
Ins
E1
Ins
E1
Ins
E1
98 kDa
62 kDa
49 kDa
38 kDa
28 kDa

Figure 4.27 SDS-PAGE gel of GlpF His-tag of 1L of starting culture
A 12 % SDS-PAGE separated 10 µL of sample fractions. Ins are the insoluble fraction, and EX are the elution fraction 1 is the first and 3 is the 3rd. 37 ˚C belongs to the culture grown at 37 ˚C without IPTG. RT are the culture grown at Room Temperature with 0.75 mM of IPTG.

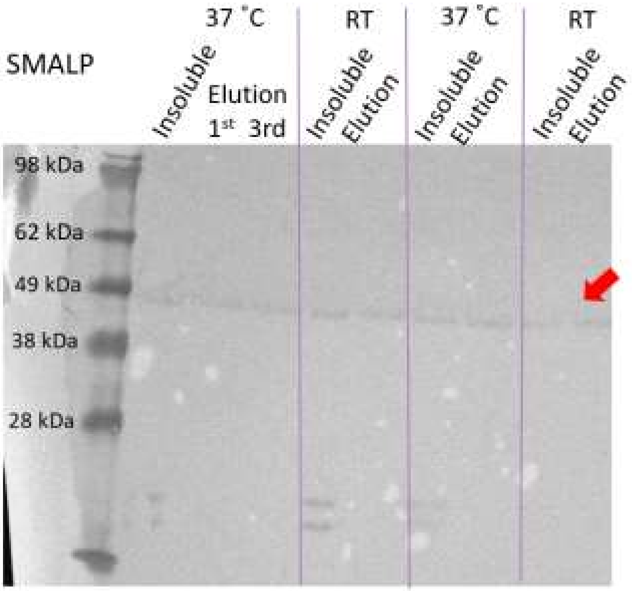

SMALP
37 ˚C
RT
37 ˚C
RT
Insoluble
Elution
1st 3rd
Insoluble
Elution
Insoluble
Elution
Insoluble
Elution
98 kDa
62 kDa
49 kDa
38 kDa
28 kDa

Figure 4.28 Western blot of GlpF His-tag of 1 L of starting culture
10 µL of sample fractions were separated by a 12 % SDS-PAGE and analysed by Western blot using antibodies directed against His-tag. 37 ˚C are the insoluble and elution fractions at 37°C without IPTG. RT are the insoluble and elution fractions at Room Temperature with 0.75 mM of IPTG. The red arrow indicates the band shown in all lanes at ~49 kDa, which suggests GlpF in oligomeric form.

GlpF was again induced at 37 °C without IPTG in 1 L culture; for its extraction and purification, the detergent DDM was used (Figure 4.29). A band of the expected monomeric size (~28 kDa) was observed in the insoluble, 1st and 3rd elution fractions and the resin. This expression was insufficient to continue with the characterisation. It was therefore decided to see if a change of plasmid to pET15b would give a better expression.

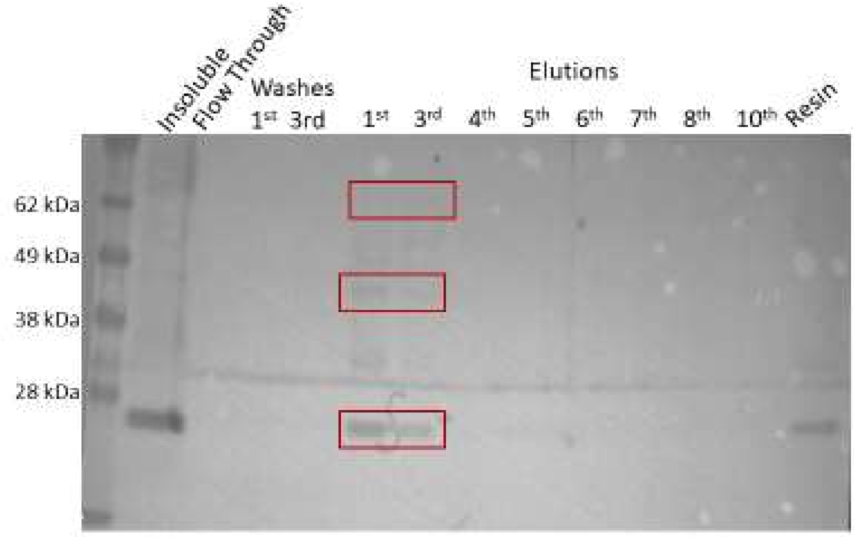


Figure 4.29 Western blot of GlpF Hist-tag of 1 Lt of starting culture and DDM for its extraction
GlpF was induced at 37 °C without IPTG. 10 µL of sample fractions were separated by a 12 % SDS-PAGE and analysed by Western blot using antibodies directed against His-tag. The red boxes indicate GlpF in monomeric and oligomeric forms in the 1st and 3rd elution fractions.

4.2.5 Expression of pET15b-Ec-GlpF-myc

Because the level of expression of GlpF was insufficient to continue with further experiments, a pET15-Ec-Glpf-myc plasmid with *E. coli* GlpF was used to test if pET15b is better than pET17b for membrane protein expression. *E. coli* GlpF was induced at 37 °C with 1 mM IPTG in 1 L culture; for its extraction and purification, the detergent OG was used (Figure 4.30). OG detergent was used as Borgnia and Agre (2000) used it to characterise GlpF, and the vector construction was based on the plasmid reported in their study. A band of the expected size (28 kDa) was observed in the SDS-PAGE and elution fractions in the Western blot. Bands of higher size ~40 kDa were faintly observed, suggesting oligomerisation of GlpF.

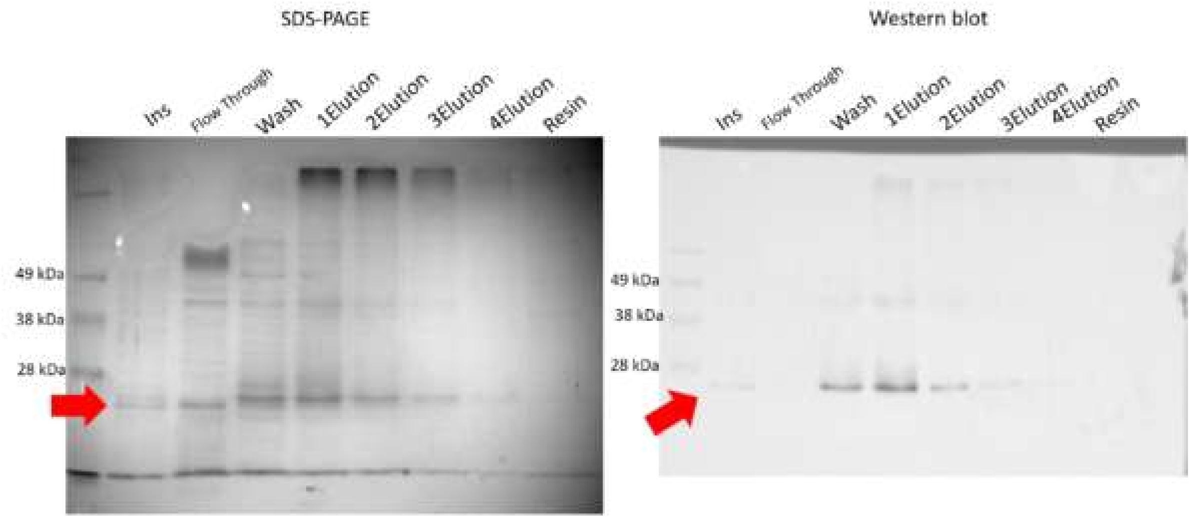

SDS-PAGE
Ins
Flow Through
Wash
1Elution
2Elution
3Elution
4Elution
Resin
49 kDa
38 kDa
28 kDa
Western blot
Ins
Flow Through
Wash
1Elution
2Elution
3Elution
4Elution
Resin
49 kDa
38 kDa
28 kDa

Figure 4.30 Ec-GlpF expression in SDS and Western blot
Ec-GlpF was induced and purified with OG. 10 µL of sample fractions were separated by 12 % SDS-PAGE in duplicates. One gel was stained with Coomassie blue (SDS-Page), and the other was analysed using an antibody directed against His-Tag by Western blot. 1st Lane is the Ladder. Ins: insoluble fraction. Red arrows indicate the height where the band of GlpF should appear.

The expression of *E. coli* GlpF was sufficient to continue with the experiments. However, Ec-GlpF has been fully characterised, and this research aims to characterise *C. saccharoperbutylacetonicum* N1-4HMT GlpF. Therefore, it was decided to replace Ec-GlpF with Cs-GlpF in the pET15b-Ec-GlpF-myc vector and increase the his-tag to 12 histidines to allow efficient purification in SMALPs. Because His_6-tags do not properly bind to Ni^{2+} when they SMALPs are used (Broecker *et al.*, 2017), and more stringent washing steps can be used with longer polyhistidine tags (Bornhorst and Falke, 2000).

4.2.6 Construction of pET15b-Cs-GlpF-Histag

Clostridium saccharoperbutylacetonicum N1-4HMT *glpF* was amplified via PCR from gDNA (Figure 4.31), and a band of the expected size (742 bp) was observed.

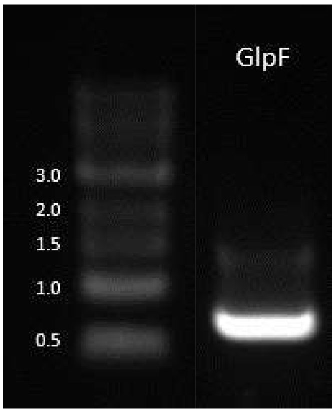

GlpF
3.0
2.0
1.5
1.0
0.5

Figure 4.31 *C. saccharoperbutylacetonicum* N1-4HMT *glpF* PCR
PCR amplification of the *glpF* gene from *C. saccharoperbutylacetonicum* N1-4HMT was performed using specific primers. 5 µL of PCR products were run on 1 % agarose gel. Lane 1 is the DNA ladder. Lane 2: *glpF*

The PCR product was purified using the Monarch® PCR & DNA cleanup kit, and then the purified products and the plasmid pET15b were double digested with *Nco*I and *BamH*I (Figure 4.32). The double digested vector (DD-pET15b) and Cs*glpF* insert (DD-GlpF) were gel extracted and purified using the Monarch ® DNA gel extraction kit. In the DD-pET15b, the Ec*glpF* was seen at ~ 1000 bp. Ec*glpF* was going to be replaced with the Cs*glpF* insert (DD-GlpF). The DD-pET15b and DD-GlpF were ligated using the T4 DNA ligase from (NEB, USA).

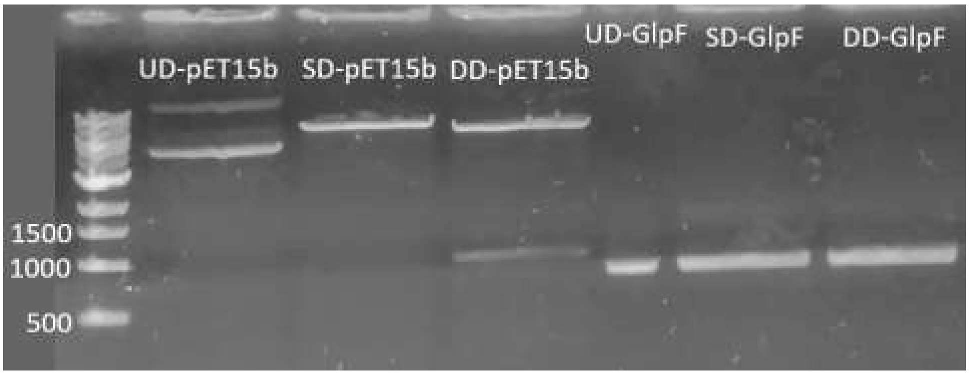


Figure 4.32 Double digestion of pET15b and GlpF
PET15b plasmid and Cs*glpF* PCR products were double digested with *Nco*I and *BamH*I. UD = undigested. SD = single digested with *Nco*I. DD = double digested with *Nco*I and *BamH*I. 20 µL of the digested products were run on 1 % agarose gel.

E. coli BL21 was transformed with the ligated pET15b-CS-GlpF. Colonies were obtained from the transformation plates at 1:3 and 1:5 vector:insert ratio. One colony was observed in the negative plate. It was decided to do a colony-PCR, using specific primers for Cs*glpf* amplification (Table 2.3), in three transformed colonies and on the colony of the negative plate (Figure 4.33). The three transformed colonies gave a band of the expected size of ~ 741 bp, whereas the negative colony did not give a band. These experiments confirm that the transformation worked and that the experiments to express GlpF from *C. saccharoperbutylacetonicum* N1-4HMT in *E. coli* can continue.

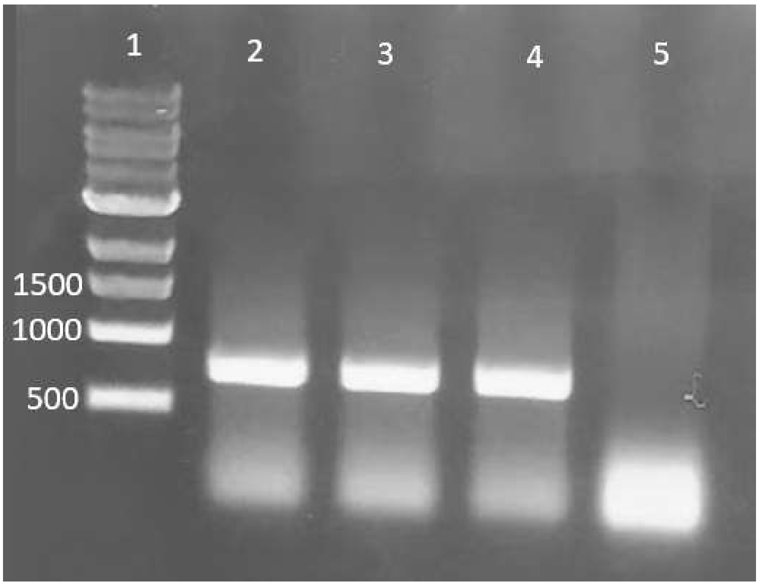


Figure 4.33 Colony PCR from pET15b-Cs-GlpF transformation.
Colony PCR was performed on three colonies obtained from the transformation plates (lanes 2-4) and on the colony observed on the negative plate (lane 5). Lane 1: DNA ladder. 5 µL of the colony PCR were run on 1 % agarose gel

4.2.7 pET15-Cs-GlpF-Histag expression

GlpF was expressed in 1 L of culture; after it reached an OD_{600} of 0.5 (approximately 2 h), it was induced with 1 mM of IPTG for 4 hours at 37 ˚C. Membranes were extracted with French Press, and the culture was divided in two. Half of the membrane fractions were solubilised with SMA 2000 and the other half with the Octyl-β-glucoside (OG) detergent. For both procedures, affinity chromatography was performed for their purification. Western blot was performed to detect the His-tag on the samples (Figure 4.34). Several bands were observed in the SDS-PAGE fractions of the OG detergent (Figure 4.34.a)). The bands that remained throughout the elutions were above 98 kDa, around 50 kDa and 36 kDa, which correspond with previous observations that suggest that GlpF is in oligomeric form. There are fewer and fainter bands above 98, around 50 and slightly below 36 kDa on the SMA side. With the Western blot (Figure 4.34. b)), a single line is observed at 50 kDa in the OG section, which could be a contaminant in one of the buffers since the line looks continuous. On the SMA blot, a strong band is observed above 50 kDa in the 1st elution fraction, with a less intense band below 36 kDa. The ≥ 50 kDa band slowly fades off with the elutions, and the ≤ 36 kDa seems to disappear at the 5th elution. CsGlpF has been successfully expressed and purified using SMA.

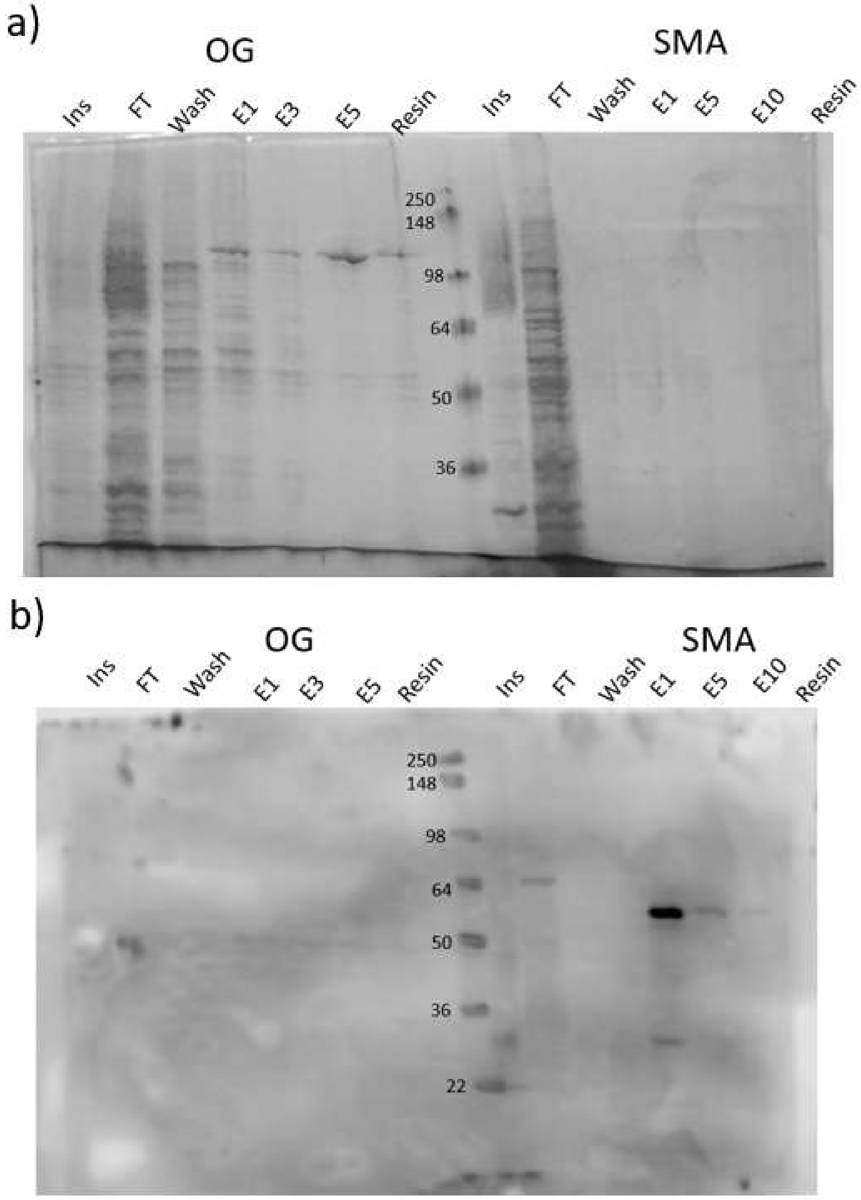

a)
OG
SMA
Ins
FT
Wash
E1
E3
E5
Resin
Ins
FT
Wash
E1
E5
E10
Resin
250
148
98
64
50
36
b)
OG
SMA
Ins
FT
Wash
E1
E3
E5
Resin
Ins
FT
Wash
E1
E5
E10
Resin
250
148
98
64
50
36
22

Figure 4.34 Cs-GlpF solubilisation with OG or SMA
Cs-GlpF was induced and purified with OG, and SMA. 10 µL of sample fractions were separated by 12 % SDS-PAGE in duplicates. a) One gel was stained with Coomassie blue (SDS-PAGE), and b) the other was analysed by Western blot using antibodies directed against His-Tag. Ins: insoluble fraction. FT: Flow-through. EX: the number of elution fractions.

Despite the results obtained with SMA being as expected, the concentrations still needed to be increased to proceed. It is unclear why the bands that appeared on the SDS-PAGE gel on the OG side disappeared in the Western blot. It was not a transfer problem, antibody, or chemiluminescence activity since the protein ladder and the bands on the SMA side can be clearly seen. There is no information about OG interfering or masking the His-tag. Different attempts to improve these results were performed, such as increasing the concentration of imidazole with OG and re-running the flow through in the column. No attempt gave better results (data not shown).

DDM is considered a good starting detergent for membrane protein purification due to its mild denaturation properties (Kotov *et al.*, 2019). It is the most successful detergent for crystallising α-helical IMPs (Yang *et al.*, 2014). It was decided to try the detergent DDM (Figure 4.35). Galka *et al.* (2008) reported that GlpF in DDM exists predominantly as a tetramer. In the SDS-PAGE gel (Figure 4.35. a)), one band was observed at 98 kDa of the first elution. In the second elution, clearer bands were observed at 98, 64, around 50 and 36 kDa. In the Western blot (Figure 4.35. b)), strong bands are observed below 64 kD in the second and third elution. Slightly below 36 kD, a strong band is observed in the second elution, and about half of the intensity is observed in the third elution at the same molecular weight. Faint bands are observed at 98 kDa in the second and third elution. Compared to what Galka *et al.* (2008) reported, DDM solubilised CsGlpF is not predominantly observed as a tetramer but definitely as an oligomer. Comparing the results with SMA, OG and DDM, DDM is the one that has given the higher density bands on the western blot that can also be observed in the SDS-PAGE.

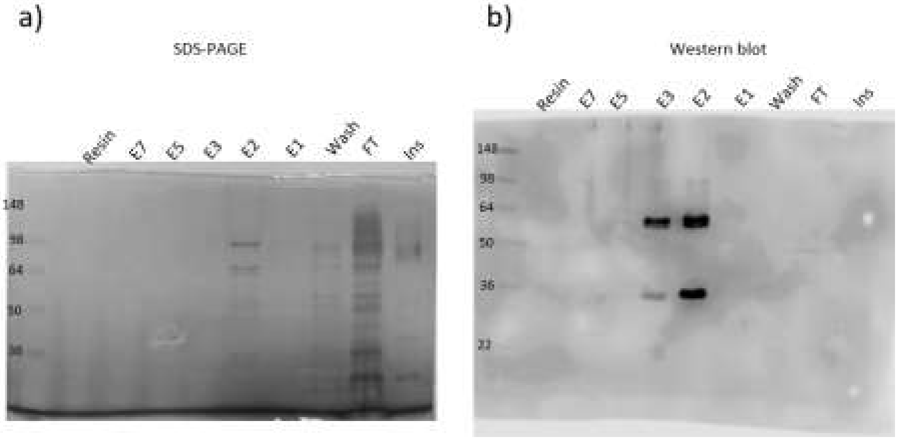


Figure 4.35 Cs-GlpF solubilisation with DDM
Cs-GlpF was induced and purified with DDM. 10 µL of sample fractions were separated by 12 % SDS-PAGE in duplicates. a) One gel was stained with Coomassie blue (SDS-PAGE), and b) the other was analysed by Western blot using antibodies directed against His-Tag. Ins: insoluble fraction. FT: Flow-through. EX: the number of elution fractions.

This experiment was meant to be repeated and scaled up to 4 L to continue characterising the protein. However, the remainder of the work was impossible due to the COVID-19 pandemic preventing laboratory access.

To summarise, a common problem in producing organic acids is the acidification of the cytoplasm due to their undissociation inside the cell, releasing protons. Proton antiporters can alleviate H^+ and Na^+ accumulation because of their capacity to pump them out of the cells. The thesis project aimed to explore whether proton antiporters protect against product toxicity and increase organic acid production. No reports have been found that examine their potential for tolerance engineering. The chapter aimed to choose two antiporters from Clostridia, express them and purify them in *E. coli* for further characterisation to evaluate their potential for tolerance engineering. The *C. saccharoperbutylacetonicum* antiporters were identified. Based on the characteristics of the antiporters to export protons in exchange for Na^+, the Na^+:H^+ antiporter A (NhaA) and the monovalent cation:proton antiporter 1 (Cpa1) were chosen. Their sequence was identified, and only the protein modelling from NhaA was obtained. Cpa1 does not have a homology protein with a protein that has been experimentally identified. The genes from NhaA and Cpa1 from *C. saccharoperbutylacetonicum* N1-4HMT were amplified with and without their own promoter, adding a His-Tag at the C-terminus. The amplification was ligated into PTZ19R/U, pET17b, pET15b and pBR322 and appropriate host *E. coli* cells were transformed. Positive colonies were obtained only from PTZ19R/U.

The *C. saccharoperbutylacetonicum* N1-4HMT, glycerol facilitator (GlpF), was chosen for two reasons. 1) Its overexpression has been observed during the fermentation's acidogenic phase, suggesting a protective role against the production of organic acids in the fermentation. 2) The overexpression of two proton antiporters, NhaA and Cpa1, was going to be tested, and the osmolarity of the cell is likely to change; GlpF has a role in the control of osmotic balance. *E. coli* GlpF has been fully characterised, but no attempt has been reported in *C. saccharoperbutylacetonicum* N1-4HMT, which can give valuable information that could help improve the stress resistance of Clostridial fermentation for large-scale productions. The plasmid pET15b-Cs-GlpF-Histag was successfully constructed, and *E. coli* BL21 cells were transformed. The detergents OG, DDM, and the polymer SMA were used to express and purify GlpF. DDM gave the best results.

The intention was to obtain *E. coli* strains that expressed Cs-NhaA or Cs-Cpa1 in order to purify the proton antiporters and for Cs-GlpF to scale up to 4 L to obtain enough protein and proceed with its characterisation. Their function could be tested in liposomes, measuring the passage of H^+ and Na^+ with Cs-NhaA and Cs-Cpa1 and glycerol passage in Cs-GlpF. However, due to the closure of the laboratories during the Covid-19 pandemic, it was impossible to continue with these experiments.

5. Results III – Genetic modifications of *C. saccharoperbutylacetonicum* N1-4HMT and its characterisation for organic acid production

5.1 Introduction

Three strategies were tested in this chapter to determine whether *C. saccharoperbutylacetonicum* N1-4HMT can produce high yields of organic acids (butyric and lactic). The first was to test the overexpression of NhaA and Cpa1, using a plasmid-based method to create +NhaA and +Cpa1 strains. The second strategy was to develop a GlpF knock-out strain (ΔGlpF) using CLEAVE^{TM} technology; a way to elucidate the role of protein is to disrupt the chromosomal gene. The third strategy was to use response surface methodology (RSM), first to evaluate the effects of temperature, pH and salt concentration on the wild-type strain (WT); and then the effects of salt, lactic and butyric acid on all of the strains (WT, +NhaA, +Cpa1 and ΔGlpF) for organic acid production. Finally, an existing knock-out strain that has disrupted membrane lipid gene synthesis was used for small bottle screening to show how lipid genes can be attractive targets for organic acid production and tolerance engineering.

Proton antiporters can alleviate H^+ and Na^+ accumulation because of their capacity to pump them out of cells. Wu *et al.* (2013b) reported that overexpression of the proton antiporters *nhaA* and *nhaR* in *E. coli* increased lactate production by 25%. There are no reports of Cpa1 being tested to improve organic acid tolerance in bacteria. However, its overexpression in plants increases salt tolerance (Ma *et al.*, 2017). In this chapter, overexpression of NhaA and Cpa1 was evaluated in *C. saccharoperbutylacetonicum* N1-4HMT to assess its impact on organic acid.

The glycerol facilitator, GlpF, regulates the flow of glycerol and water in response to osmotic stress without disturbing the electrochemical gradient across the membrane (Chakrabarti *et al.*, 2004). Overexpression of GlpF has been observed during the acidogenic phase of *C. acetobutylicum* (Dong *et al.*, 2016, Mao *et al.*, 2010). This could be to glycerol accumulation (as a protective compatible solute) during butanol production in solventogenic Clostridia (Patakova *et al.*, 2022). To elucidate whether overexpression of GlpF has a protective mechanism, the expression of GlpF was prevented using CLEAVE^{TM} technology, and the resulting GlpF knock-out strain (ΔGlpF) was evaluated for the production of organic acids compared to the WT.

Another goal of this chapter was to determine whether it is possible to block the solventogenic switch to produce organic acids (butyric and lactic acid) as the primary fermentation product. One strategy was the addition of sodium chloride concentrations (NaCl) based on Maddox *et al.* (1995), who observed that high salt concentrations could divert the fermentation from being solventogenic to acidogenic.

5.2 Results

5.2.1 Generation of the genetically modified *C. saccharoperbutylacetonicum* N1-4HMT strains.

Herein, overexpression and knock-out constructs have the prefixes "+" and "Δ", respectively.

5.2.1.1 Overexpression of cation/proton antiporters in C. saccharoperbutylacetonicum N1-4HMT

Cation proton antiporters are crucial to maintaining cation and pH homeostasis, preventing cytoplasmic acidification. To investigate their role as organic acid protectors (proton extruders), NhaA and Cpa1 antiporters were overexpressed in the acidogenic and solventogenic producer *C. saccharoperbutylacetonicum* N1-4HMT. From the sequences previously obtained for chapter 4, CsNhaA and CsCpa1 sequences were synthesised by GeneArt (Thermo Fisher, USA). For quality assurance, the resulting synthesised DNA was investigated using BLAST and had 100 % identity with WP_015392945.1 for NhaA and AGF57903.1 for Cpa1. Then, NhaA and Cpa1 were subcloned in the plasmid pMTL82151 under the weak promoter pSec (Bio-Cleave Ltd), and N1-4HMT cells were transformed by electroporation. Colony PCR was performed to identify positive colonies. Empty vector pMTL82151 was used for negative control. The primers used were provided by Bio-Cleave Ltd, bind to a region of the pMTL82151 outside the multiple cloning site, and can amplify inserted sequences. Overloading of the sample resulted in very intense bands. pSec-NhaA had an expected size of 1415 bp; all colonies had an insert of the expected size (Figure 5.1). pSec-Cpa1 had an expected size of 2129 bp; it also appears that all colonies had an insert of the expected size (Figure 5.2).

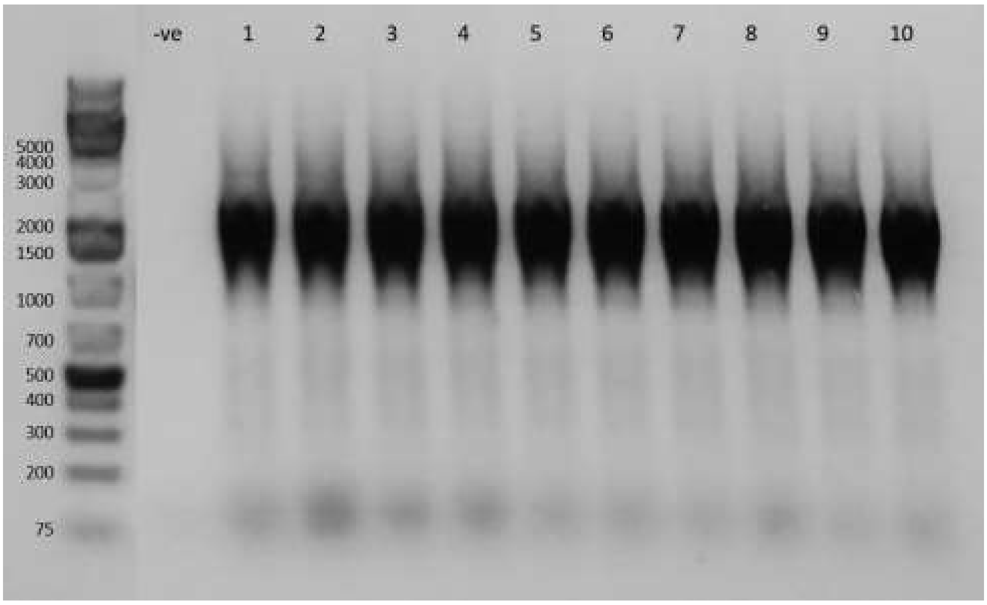

-ve
1
2
3
4
5
6
7
8
9
10
5000
4000
3000
2000
1500
1000
700
500
400
300
200
75

Figure 5.1 pSec-NhaA colony PCR
Colony PCR of pMTL82151-pSec-*nhaA* inserts from *C. saccharoperbutylacetonicum* N1-4HMT colonies using Bio-Cleave Ltd primers that amplify regions of the multiple cloning site. 10 µL of PCR products were run on 1 % agarose gel. Lane 1: Ladder. Lanes 2: -ve, negative control. Lanes 3-12: +NhaA colonies. Lanes between the ladder and –ve are not shown since they contain bands of DNA genes that do not belong to this project.

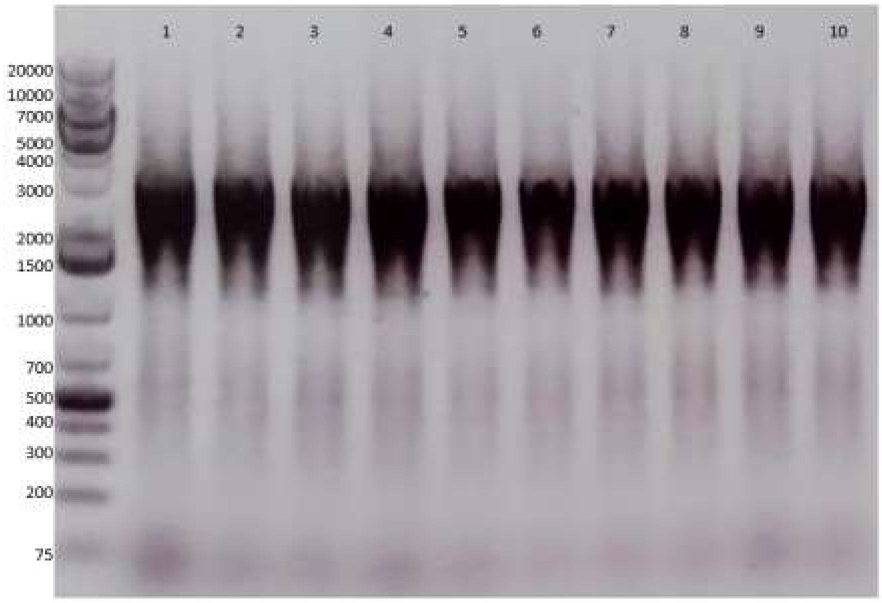

1
2
3
4
5
6
7
8
9
10
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
400
300
200
75

Figure 5.2 Colony PCR from pSec-Cpa1
Colony PCR of pMTL82151-pSec-*cpa1* inserts from *C. saccharoperacetobutylicum* N1-4HMT colonies using Bio-Cleave Ltd primers that amplify regions of the multiple cloning site. 10 µL of PCR products were run on 1 % agarose gel. Lane 1: Ladder. Lanes 2-11: +Cpa1 colonies.

Three colonies from each transformation were randomly chosen to conserve as glycerol stocks that were stored at -80 ˚C. One colony of the three was chosen for a small bottle screening.

5.2.1.2 *CLEAVE^{TM} technology*

Gene inactivation in solventogenic Clostridia has historically been inefficient, tedious, and slow. However, Jenkinson and Krabben (2015) developed CLEAVE^{TM} technology for use in *C. saccharoperbutylacetonicum* N1-4HMT, allowing microbial engineering to be exploited, and increasing its industrial potential. CLEAVE^{TM} technology works using the endogenous CRISPR-Cas mechanism. They created a strain that once a site is specifically targeted within the wild type using a targeting vector that mimics the native CRISPR-Cas direct repeats clusters, the microbe cannot repair the DNA cleavage, killing the cell. Combined with a homologous recombination step, they improved the selection method for microbes containing the desired modifications (Atmadjaja *et al.*, 2019).

5.2.1.2.1 *CLEAVE^{TM} knock-out procedure*

Figure 5.3 represents the two steps needed for knocking out a gene using CLEAVE^{TM} technology. A homologous recombination procedure is the first step, using the pMTL82151 plasmid to introduce the truncated gene in the genome. The second step is using the S3V9 plasmid to activate the CRISPR/Cas mechanism and eliminate the cells that were not modified.

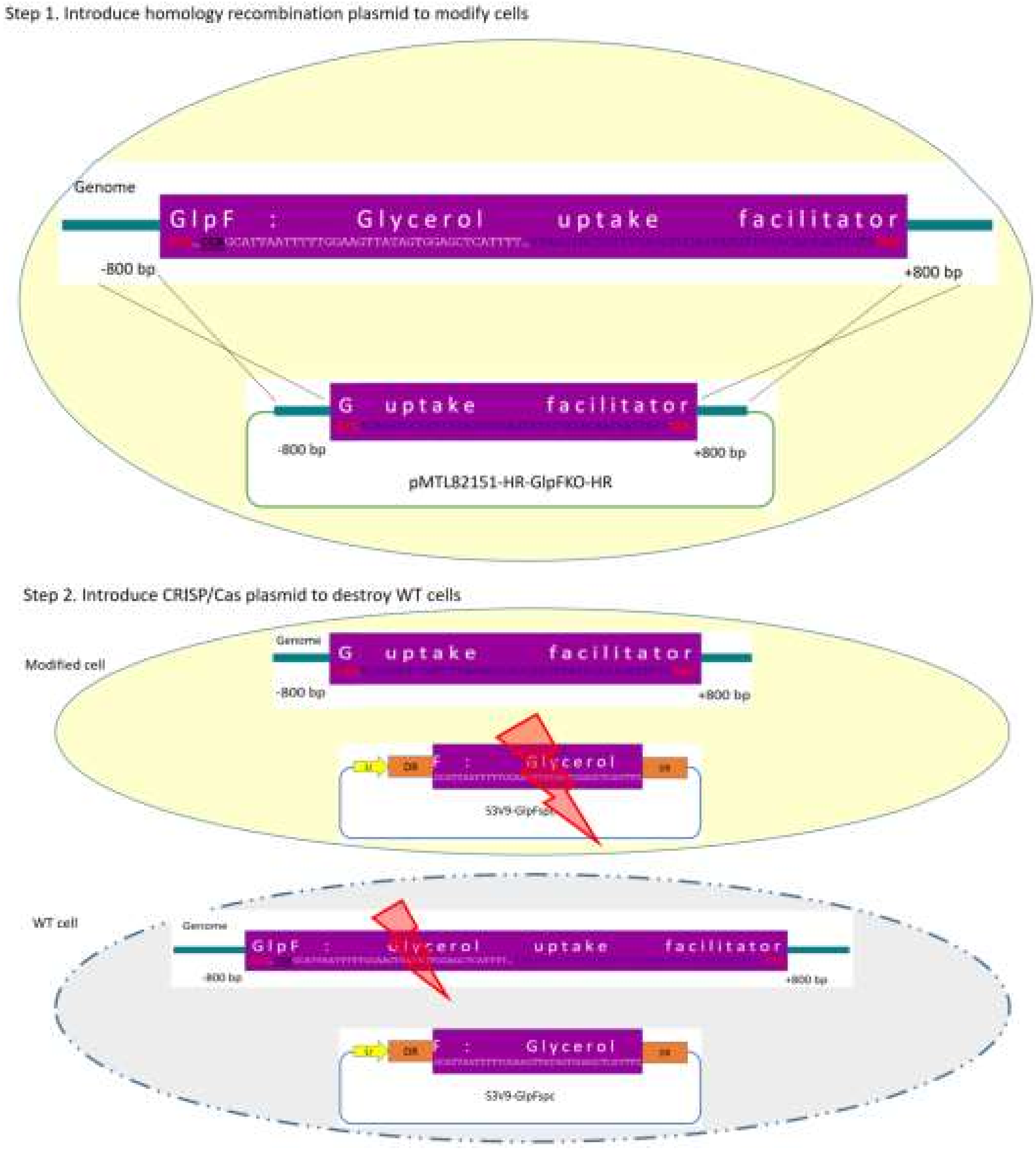

Step 1. Introduce homology recombination plasmid to modify cells
Genome
GlpF : Glycerol uptake facilitator
-800 bp
+800 bp
G uptake facilitator
-800 bp
+800 bp
pMTL82151-HR-GlpFKO-HR
Step 2. Introduce CRISP/Cas plasmid to destroy WT cells
Modified cell
Genome
G uptake facilitator
-800 bp
+800 bp
DR
F : Glycerol
DR
S3V9-GlpFspc
WT cell
Genome
GlpF : Glycerol uptake facilitator
-800 bp
+800 bp
DR
F : Glycerol
DR
S3V9-GlpFspc

Figure 5.3 CLEAVE knock-out procedure
The two steps of the CLEAVE technology. The ovals represent the bacteria; beige is alive, and grey is dead. Step 1 is the homology recombination step, which involves the target gene in the genome and the homology arm vector pMTL82151. Step 2 involves the S3V9 with the spacer to induce the CRISPR/Cas mechanism of the cell to destroy the cells that were not modified. The red lightning indicates what gets lost in the cell; the plasmid in the left modified cell and the genome in the right WT cell.

5.2.1.2.2 *Plasmid design*

To generate the deletion of a gene in Clostridial cells with the CLEAVE^{TM} technology, two vectors must be designed, the homologous arms (in pMTL82151) and the spacer plasmid for CRISPR (S3V9 [Bio-Cleave Ltd Property]) (Figure 5.4). The homologous arm (HR) plasmid in pMTL82151 must contain ~800 bp down and upstream of the targeted gene and the coding sequence for the gene's first and last 15 amino acids. The spacer plasmid in S3V9 must contain the Leader sequence – Direct repeats (identified from CRISPR/Cas array in N1-4HMT) – Spacer – Direct repeats. The leader sequence and direct repeats are Bio-Cleave Ltd's intellectual property. The spacer is a sequence of ~30 bp in the deleted sequence next to a Protospacer Adjacent Motif (PAM). For *C. saccharoperbutylacetonicum* N1-4HMT, the PAM has been identified as CCN. The PAM site must be absent in the modified gene version so CLEAVE^{TM} will target the WT copy only (Atmadjaja *et al.*, 2019).

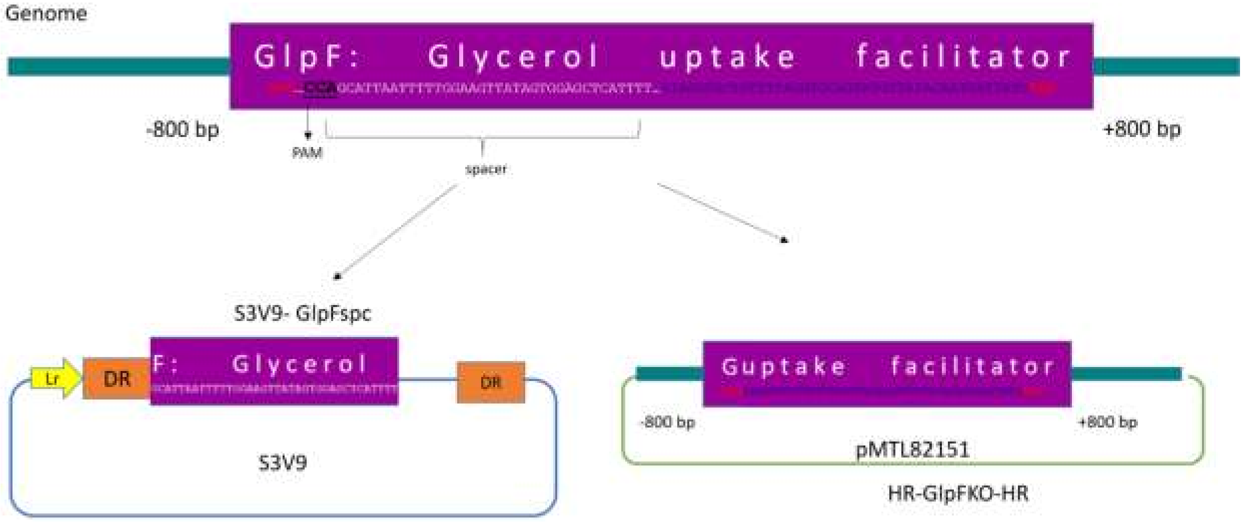
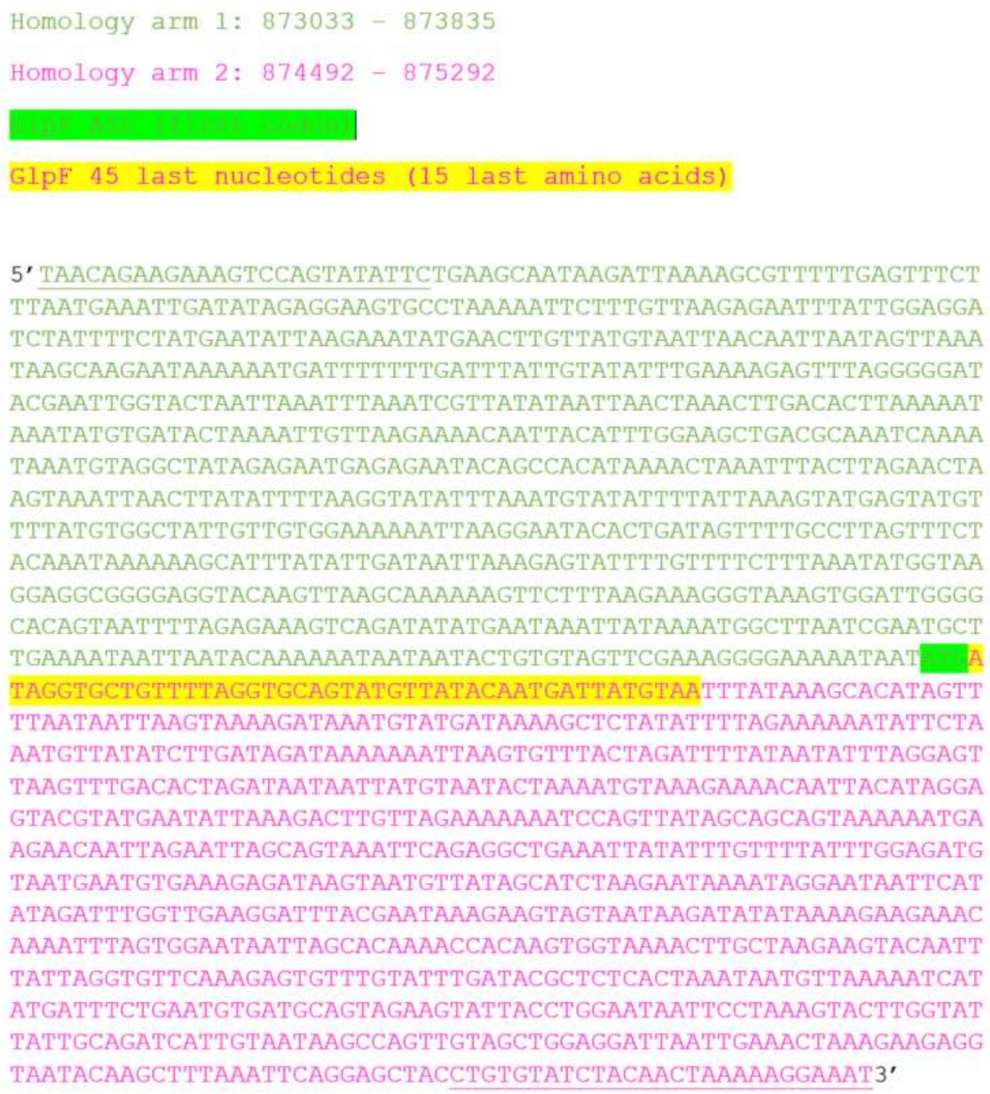


Figure 5.4 Design of CLEAVE knock-out plasmids
The top part represents the genome of N1-4HMT, and the bottom part the two vectors. The GlpF gene is represented as the purple box. The first and last codons of the gene are represented in red. pMTL82151 vector represents the down and upstream 800 bp of GlpF plus its first and last 12 codons, also represented as "G uptake facilitator". In black and underlined are the PAM sequence nucleotides in the genome GlpF. The spacer is the 20 nucleotides next to the PAM, represented as "F: Glycerol". PAM is not included in the S3V9 vector, and the spacer is not included in the pMTL82151 vector. Lr: Leader sequence. DR: Direct repeats.

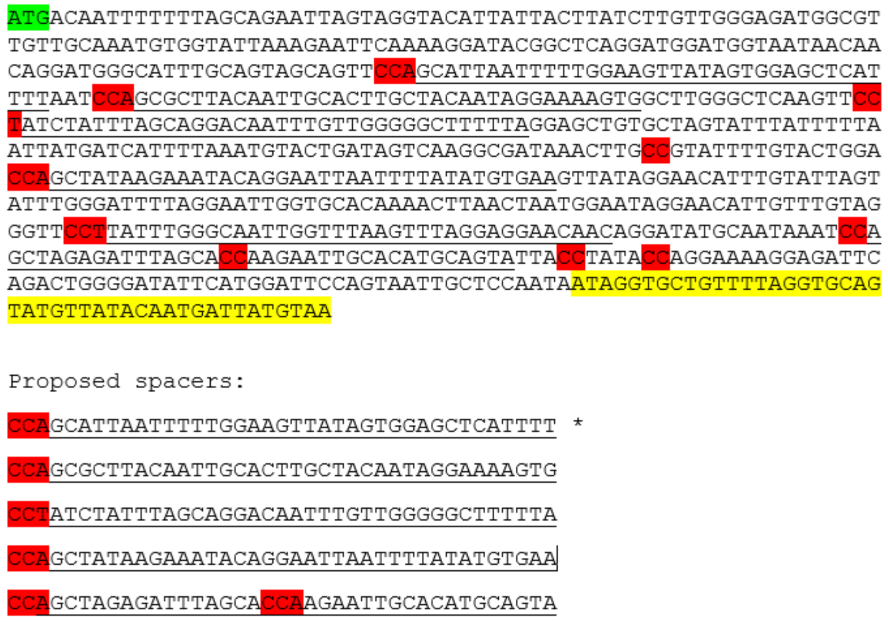
For the generation of the HR-GlpFkO-HR plasmid, the CsGlpF sequence previously obtained in Chapter IV and the complete *C. saccharoperbutylacetonicum* N1-4HMT genome (CP004121.1) (ENA browser) were used. The CsGlpF gene is located at 873,833 – 874,534. The homology regions were chosen as follows: Homology arm 1: 873, 033 – 873 835 and Homology arm 2: 874,492 – 875, 292. For the GlpF knock-out, the chosen approach was to keep the first codon of GlpF and the 15 last codons of the protein to keep the open reading frame of the whole genome. Figure 5.5 contains the homology arms (HR) before and after the GlpF sequence and the GlpF truncated gene (the first 3 and last 45 nucleotides, first and 15 last codons). This sequence was synthesised by GeneArt (Thermo Fisher, USA) and cloned into pMTL82151. The homologous recombination will occur with the genome directed by the homology arms, replacing the native GlpF sequence with the GlpF truncated sequence.



```
Homology arm 1: 873033 - 873835

Homology arm 2: 874492 - 875292

GlpF ATG (first codon)

GlpF 45 last nucleotides (15 last amino acids)


5’TAACAGAAGAAAGTCCAGTATATTCTGAAGCAATAAGATTAAAAGCGTTTTTGAGTTTCT
TTAATGAAATTGATATAGAGGAAGTGCCTAAAAATTCTTTGTTAAGAGAATTTATTGGAGGA
TCTATTTTCTATGAATATTAAGAAATATGAACTTGTTATGTAATTAACAATTAATAGTTAAA
TAAGCAAGAATAAAAAATGATTTTTTTGATTTATTGTATATTTGAAAAGAGTTTAGGGGGAT
ACGAATTGGTACTAATTAAATTTAAATCGTTATATAATTAACTAAACTTGACACTTAAAAAT
AAATATGTGATACTAAAATTGTTAAGAAAACAATTACATTTGGAAGCTGACGCAAATCAAAA
TAAATGTAGGCTATAGAGAATGAGAGAATACAGCCACATAAAACTAAATTTACTTAGAACTA
AGTAAATTAACTTATATTTTAAGGTATATTTAAATGTATATTTTATTAAAGTATGAGTATGT
TTTATGTGGCTATTGTTGTGGAAAAAATTAAGGAATACACTGATAGTTTTGCCTTAGTTTCT
ACAAATAAAAAAGCATTTATATTGATAATTAAAGAGTATTTTGTTTTCTTTAAATATGGTAA
GGAGGCGGGGAGGTACAAGTTAAGCAAAAAAGTTCTTTAAGAAAGGGTAAAGTGGATTGGGG
CACAGTAATTTTAGAGAAAGTCAGATATATGAATAAATTATAAAATGGCTTAATCGAATGCT
TGAAAATAATTAATACAAAAAATAATAATACTGTGTAGTTCGAAAGGGGAAAAATAATATGA
TAGGTGCTGTTTTAGGTGCAGTATGTTATACAATGATTATGTAATTTATAAAGCACATAGTT
TTAATAATTAAGTAAAAGATAAATGTATGATAAAAGCTCTATATTTTAGAAAAAATATTCTA
AATGTTATATCTTGATAGATAAAAAAATTAAGTGTTTACTAGATTTTATAATATTTAGGAGT
TAAGTTTGACACTAGATAATAATTATGTAATACTAAAATGTAAAGAAAACAATTACATAGGA
GTACGTATGAATATTAAAGACTTGTTAGAAAAAAATCCAGTTATAGCAGCAGTAAAAAATGA
AGAACAATTAGAATTAGCAGTAAATTCAGAGGCTGAAATTATATTTGTTTTATTTGGAGATG
TAATGAATGTGAAAGAGATAAGTAATGTTATAGCATCTAAGAATAAAATAGGAATAATTCAT
ATAGATTTGGTTGAAGGATTTACGAATAAAGAAGTAGTAATAAGATATATAAAAGAAGAAAC
AAAATTTAGTGGAATAATTAGCACAAAACCACAAGTGGTAAAACTTGCTAAGAAGTACAATT
TATTAGGTGTTCAAAGAGTGTTTGTATTTGATACGCTCTCACTAAATAATGTTAAAAATCAT
ATGATTTCTGAATGTGATGCAGTAGAAGTATTACCTGGAATAATTCCTAAAGTACTTGGTAT
TATTGCAGATCATTGTAATAAGCCAGTTGTAGCTGGAGGATTAATTGAAACTAAAGAAGAGG
TAATACAAGCTTTAAATTCAGGAGCTACCTGTGTATCTACAACTAAAAAGGAAAT3’
```

Figure 5.5 HR- GlpF knock-out construct for genome integration
The HR-Knockout contains a homology arm (HR) 1, which is -800 bp away from the GlpF gene (green font), the AUG start nucleotides from GlpF (first codon, green highlighting), the last 45 nucleotides of GlpF including the stop codon (15 last amino acids, pink font yellow highlighting) and the second homology arm (HR) +800 bp from GlpF gene (pink font). This sequence is cloned into pMTL82151 for homologous recombination.

For the S3V9 plasmid, five possible PAM sequences were identified (Figure 5.6). In the CRISPR-Cas system, the spacer is the sequence Cas will recognise for the double strand cut. Therefore, it should only be present in the WT cells and not in the knock-out gene. The bacteria recognises the PAM sequence as foreign DNA, which means it cannot be included in the spacer. Bio-Cleave Ltd has software that evaluates the spacer sequences and selects the adequate sequence that will have higher efficiency in its identification and double-strand cut.



```
ATGACAATTTTTTTAGCAGAATTAGTAGGTACATTATTACTTATCTTGTTGGGAGATGGCGT
TGTTGCAAATGTGGTATTAAAGAATTCAAAAGGATACGGCTCAGGATGGATGGTAATAACAA
CAGGATGGGCATTTGCAGTAGCAGTTCCAGCATTAATTTTTGGAAGTTATAGTGGAGCTCAT
TTTAATCCAGCGCTTACAATTGCACTTGCTACAATAGGAAAAGTGGCTTGGGCTCAAGTTCC
TATCTATTTAGCAGGACAATTTGTTGGGGGCTTTTTAGGAGCTGTGCTAGTATTTATTTTTA
ATTATGATCATTTTAAATGTACTGATAGTCAAGGCGATAAACTTGCCGTATTTTGTACTGGA
CCAGCTATAAGAAATACAGGAATTAATTTTATATGTGAAGTTATAGGAACATTTGTATTAGT
ATTTGGGATTTTAGGAATTGGTGCACAAAACTTAACTAATGGAATAGGAACATTGTTTGTAG
GGTTCCTTATTTGGGCAATTGGTTTAAGTTTAGGAGGAACAACAGGATATGCAATAAATCCA
GCTAGAGATTTAGCACCAAGAATTGCACATGCAGTATTACCTATACCAGGAAAAGGAGATTC
AGACTGGGGATATTCATGGATTCCAGTAATTGCTCCAATAATAGGTGCTGTTTTAGGTGCAG
TATGTTATACAATGATTATGTAA

Proposed spacers:

CCAGCATTAATTTTTGGAAGTTATAGTGGAGCTCATTTT *

CCAGCGCTTACAATTGCACTTGCTACAATAGGAAAAGTG

CCTATCTATTTAGCAGGACAATTTGTTGGGGGCTTTTTA

CCAGCTATAAGAAATACAGGAATTAATTTTATATGTGAA

CCAGCTAGAGATTTAGCACCAAGAATTGCACATGCAGTA
```

Figure 5.6 PAM sequences identified in GlpF gene for CLEAVE technology
Above is the GlpF sequence with the identified protospacer adjacent motif (PAM) sequences (red highlighting) and the proposed spacer sequences within it (underlined). The spacer sequence is a ~ 40 nt length adjacent upstream by the PAM (CCN sequence). The Met codon (green highlighting) and the 15 last amino acids, including the stop codon (yellow highlighting), are shown. The spacer sequence cannot be inside the truncated gene sequence. The PAM sequence cannot be in the spacer sequence. Bio-Cleave Ltd analysed the proposed spacers and, with the help of software, chose the most adequate, denoted by a "*".

The chosen spacer was synthesised by GeneArt synthesis (Thermo Fisher, USA) flanked by *Xho*I and *Zra*I restriction sites and was subcloned in the S3V9 plasmid (Figure 5.7).

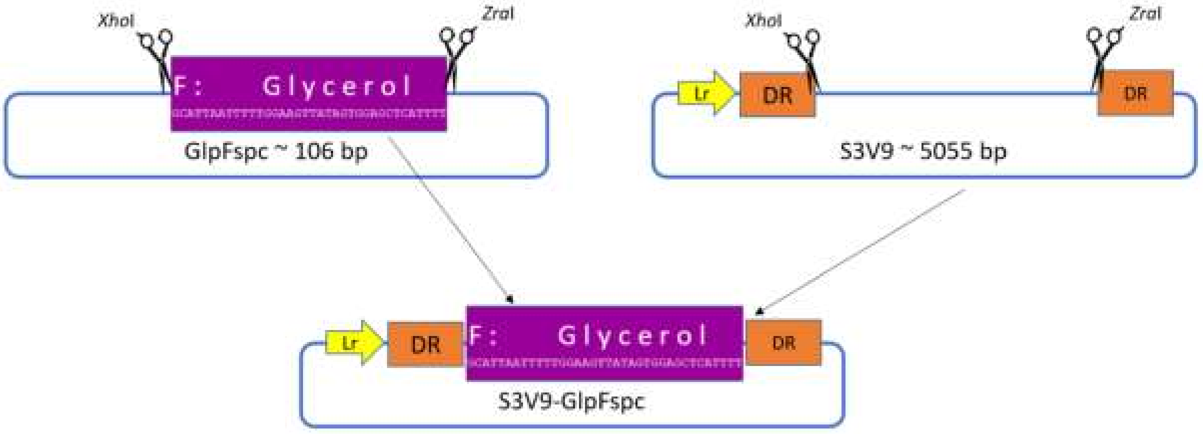


Figure 5.7 Spacer vector construction
Schematic representation of the spacer vector construction. The GlpF spacer (purple box "F: Glycerol") was synthesised in a standard backbone (GlpFspc vector) and needed to be ligated into the SV39 vector (Bio-Cleave Ltd, property), which contains the Leader sequence (yellow arrow Lr) and the Direct repeats (orange box DR). The scissors indicate double digestion. The bottom drawing indicates how the plasmid SV39-GlpFspc is going to be ligated.

5.2.1.2.3 *CLEAVE^{TM} technology step 1, pMTL82151 construction and bacteria transformation for the homologous recombination*

Once the pMTL82151 HR-GlpFKO-HR plasmid was synthesised, *C. saccharoperbutylacetonicum* N1-4HMT cells were transformed with electroporation (Figure 5.3, step 1). Colony PCR was performed to identify positive colonies (Figure 5.8). The primers were provided by Bio-Cleave Ltd, designed to amplify part of the Multiple Cloning Site of the pMTL82151 vector plus the possible insert. Water was used as the negative control. A lot of the sample was loaded, giving very intense bands. HR-GlpFKO-HR had an expected size of 1614 bp; it appears that only colony 10 had an insert of the expected size. It needs to be clarified why the rest of the colonies gave a band of a larger size. However, the bands do not represent the complete GlpF gene and truncated GlpF gene since the amplification was targeted to the pMTL82151 vector, not the GlpF sequence, and GeneArt synthesised the plasmid. pMTL82151 vectors containing full GlpF were not possible. Since one colony gave an expected size band, this was not further investigated. Colony 10 was chosen for glycerol stock and to continue with the CLEAVE technology procedure.

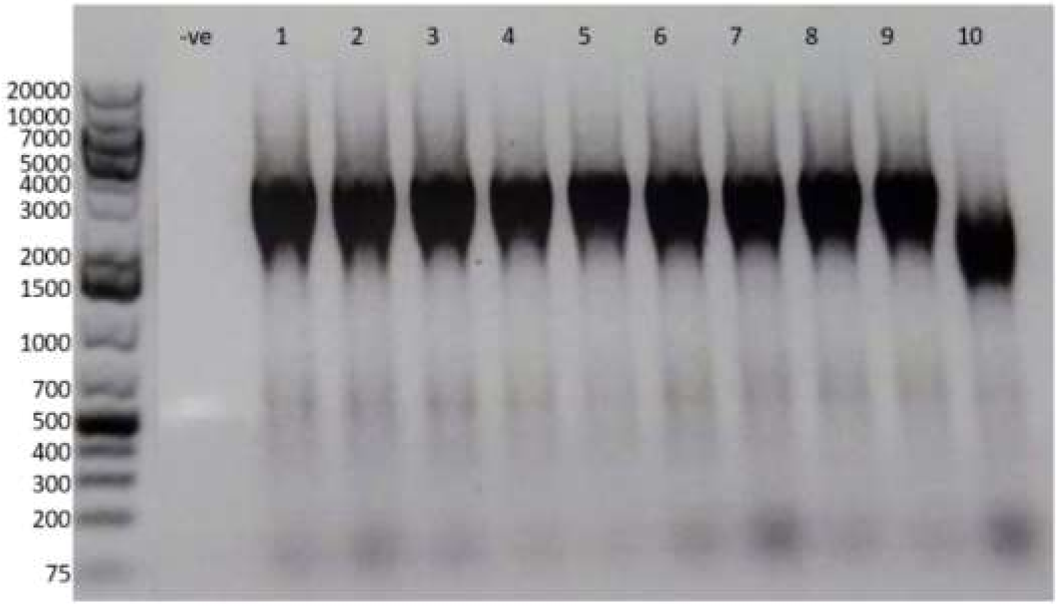


Figure 5.8 HR-GlpFKO-HR colony PCR
Colony PCR of pMTL82151 inserts from *C. saccharoperbutylacetonicum* N1-4HMT colonies using Bio-Cleave Ltd primers. 10 µL of PCR products were run on 1% agarose gel. Lane –ve: negative control. Lane 1-10: HR-GlpFKO-HR colonies. Lanes between the ladder and –ve are not shown since it contains bands of DNA genes that do not belong to this project.

5.2.1.2.4 *Spacer vector construction*

GeneArt (Thermo Fisher, USA) synthesised the spacer with the direct repeats and sent in a backbone plasmid. It needed to be ligated into the S3V9 vector (Bio-Cleave Ltd, property) (Figure 5.7). Therefore, the spacer and S3V9 vector were double digested with *Xho*I and *Zra*I (Figure 5.9). GlpF spacer had an expected size of 106 bp (red box) and S3V9 of 5055 bp.

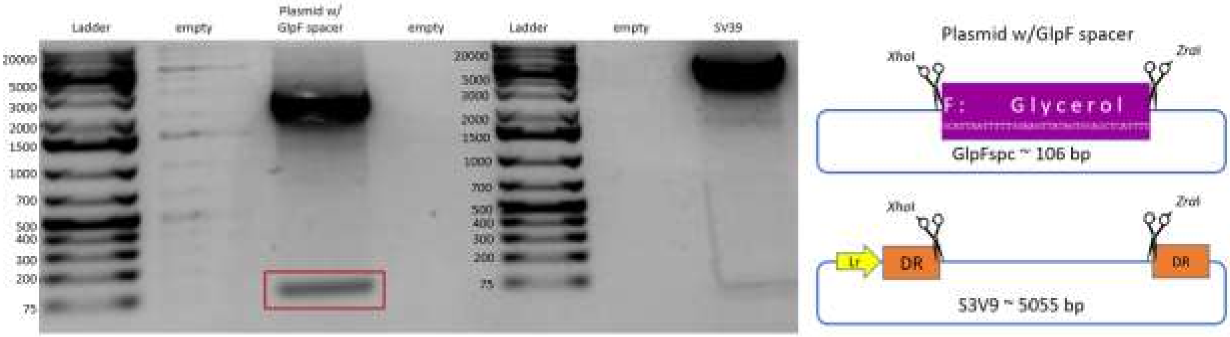


Figure 5.9 GlpF spacer (plasmid w/GlpF spacer) and S3V9
1000 ng of BC7 (GlpF spacer) and S3V9 were double digested with *Xho*I and *Zra*I, as shown in Figure 5.7. 20 µL of the digestion were run on 1% agarose gel. Lane 1 and 5: Ladder. Lane 3: GlpF Spacer indicated by a red box. Lane 7: S3V9. Lane 2,4,6: empty lanes. On the right, the plasmid maps are shown.

The double digested GlpF spacer and S3V9 vector bands were gel purified, ligated and transformed in *E. coli* DH10b cells. Colony PCR was performed to identify positive colonies. An empty vector was used as a positive control (Figure 5.10). A band of ~700 bp was expected. The primers were designed by Bio-Cleave Ltd and should amplify the leader sequence, direct repeats and the spacer (Figure 5.7). Based on size difference, colony 3 was chosen, glycerol stock was made, and the plasmid was purified and sent to sequencing. Because the bands were smaller than expected, the remaining colonies could be S3V9 plasmids where the GlpF spacer was not ligated. To test the efficiency of the spacer, *C. saccharoperbutylacetonicum* N1-4HMT strains were transformed by electroporation with the S3V9-GlpFspc plasmid. No colonies were expected to grow because the spacer targeted wild-type cells that contained the complete GlpF gene, as represented in Figure 5.3, step two. Alongside, cells were transformed with the S3V9 empty vector. 24 colonies were obtained in the S3V9-GlpFspc, compared to 565 colonies from the positive control (S3V9 empty vector), which means the GlpF spacer is not 100 % efficient. After the CLEAVE^{TM} technology is performed, more colonies must be tested to confirm that the knock-out was achieved since some WT colonies will survive.

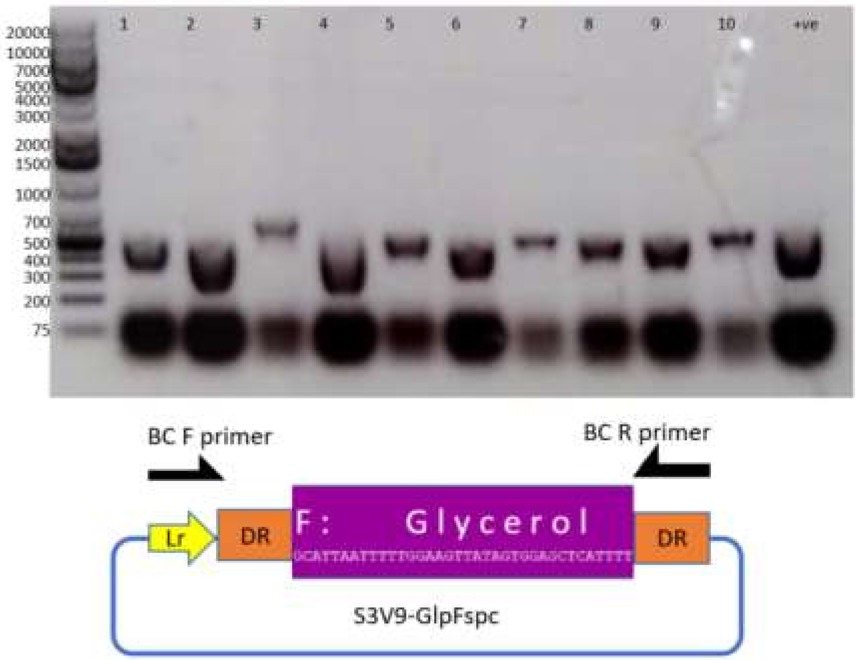


Figure 5.10 S3V9-GlpFspacer colony PCR
Colony PCR of *E. coli* colonies transformed with S3V9-GlpFspc plasmid, using Bio-Cleave Ltd primers. 10 µL of PCR products were run on 1 % agarose gel. Lane 1: Ladder. Lanes 2-11: S3V9-GlpF spc colonies. Lane 12: +ve, positive control, S3V9 empty vector. At the bottom, the plasmid map and where the primers anneal are shown; a band of ~ 700 bp is expected if the GlpF spacer was correctly inserted. BC F primer: Biocleave forward primer. BC R primer: Biocleave reverse primer. Lr: leader sequence. DR: direct repeats. F: Glycerol: GlpF spacer.

Bio-Cleave Ltd kindly completed the remaining work to produce ΔGlpF strain due to the COVID-19 pandemic and the resulting lack of laboratory access. This involved subculturing the *C. saccharoperbutylacetonicum* N1-4HMT transformed with HR-GlpFKO plasmid (Figure 5.8, colony 10) to induce homologous recombination, as depicted in Figure 5.3, step 1. The cells were then transformed with the S3V9-GlpF spacer construct (Figure 5.10, colony 3) to eliminate the cells that did not achieve the recombination and still contained the complete GlpF gene (Figure 5.3. Step 2). Three colonies were chosen, but only one colony was sent for sequencing. For ΔGlpF sequencing confirmation, primers were designed to amplify ~ -150 bp away from the homology arm 1 and ~150 bp of the homology arm 2 (Figure 5.11).

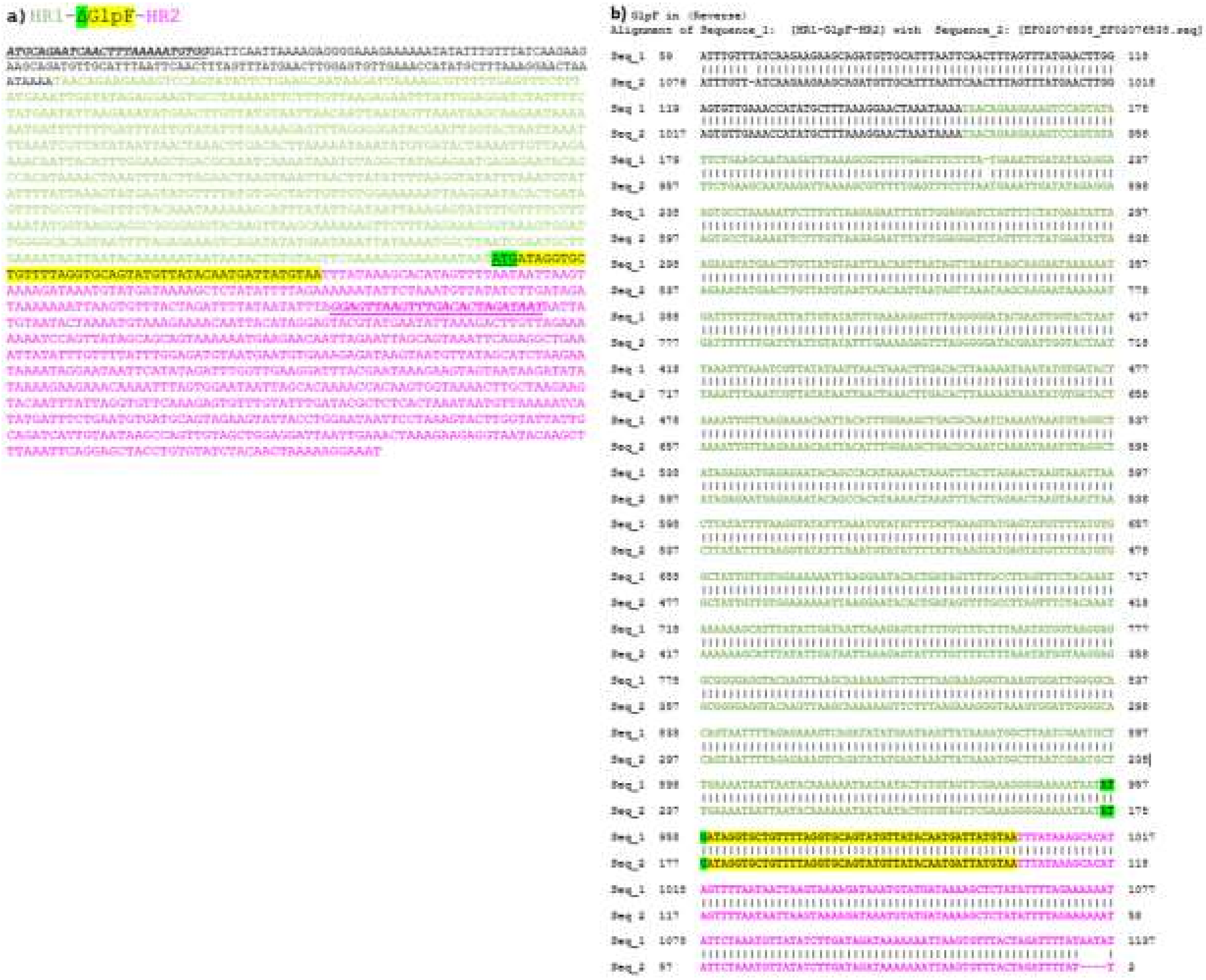


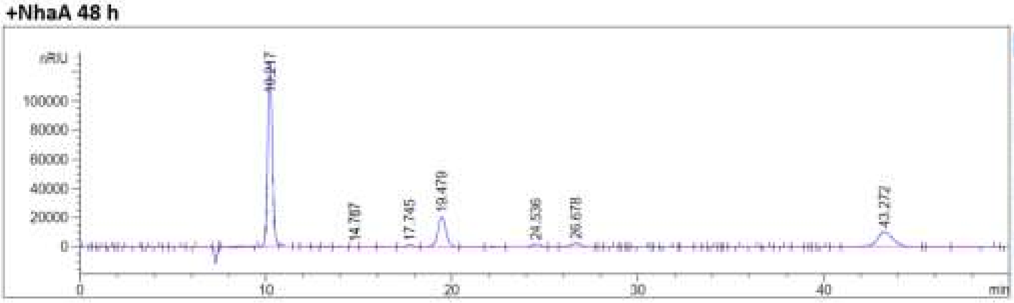
Figure 5.11 Primer design for ΔGlpF confirmation and sequence alignment
a) The sequence with the ~ 150 bp (black font) before the homology arm 1 (green font), the GlpF Met codon (green highlighting), the last 15 GlpF amino acids (yellow highlighting) and the homology arm 2 (pink font) are shown. The sequences of the primers designed to confirm the GlpF knock-out are underlined. The forward primer is ~150 bp before the homology arm 2, and the reverse is ~ 150 bp after the stop codon in the homology arm 2. b) The sequence alignment of the HR-ΔGlpF-HR sequence a) (Seq_1) and the sequencing obtained from the chosen colony (seq_2). The coloured code is kept for easy identification. It can be seen that the sequencing contains only the Met codon and the last 15 amino acids, which confirm that GlpF has been truncated.

Once confirmed that GlpF was successfully knocked out, the strain was stored at -80 ˚C and sent to Aston.

From these experiments, three *C. saccharoperbutylacetonicum* N1-4HMT mutant strains were obtained. +NhaA and +Cpa, with a plasmid-based method, overexpressed the proton antiporters CsNhaA and CsCpa1, respectively. Moreover, ΔGlpF, which by CLEAVE^{TM} technology, knocked out the GlpF gene. The fermentation characteristics of these mutant strains were investigated using a small bottle screening.

5.2.2 Characterisation of the +NhaA, +Cpa1, ΔGlpF and WT strains using small bottle screening

An initial characterisation of the +NhaA, +Cpa1, ΔGlpF and WT *C. saccharoperbutylacetonicum* N1-4HMT strains was performed via small bottle screenings in 60 mL of TYIR medium containing 50 g/L glucose and supplemented with 30 mM MES for pH control. Cells were initially grown in seed culture before being used to inoculate small bottles of 54 mL, as described in section 2.3.2. The experiment was carried out over 48 h at 32 ˚C. Samples of 1.5 mL were taken at 0, 6, 24, and 48 h, centrifuged, filtered and sent for HPLC with 1:4 dilution with HPLC water. An HPLC example is shown in Figure 5.12



Compound	Retention time
Glucose	10.000 ± 0.3
Lactic acid	14.500 ± 0.3
Acetic acid	17.600 ± 0.3
Unknown	19.200 ± 0.3
Ethanol	24.700 ± 0.3
Butyric acid	25.900 ± 0.3
Acetone	26.800 ± 0.3
Butanol	43.000 ± 0.3

Figure 5.12 HPLC example of +NhaA strain at 48 h
Chromatographic data for the analysis of +NhaA strain at 48 h. Analysis was performed using Bio-rad Aminex HPX087H column. Mobile phase 5 mM H_2SO_4 at 0.6 mL/min; 20 µL injection volume. Run time: 50 min. HAc: acetic acid. EtOH: ethanol. BuOH: butanol.

A comparison of WT with the +NhaA and +Cpa1 strains is shown in Figure 5.13. Two-way ANOVA was performed using GraphPad Prism 8.1. In the growth curve at 48 h, the +NhaA and +Cpa1 grew significantly less than the WT *C. saccharoperbutylacetonicum* N1-4HMT. +Cpa1 also grew significantly less than +NhaA. Interestingly, with sugar consumption, in the end, at 48 h, there was no significant difference. At 6 h, the pH of +Cpa1 was significantly higher than WT, which could mean that acid production was slightly slower. Although it is not significantly different, +NhaA produced the highest amount of acetic acid, which correlates with +NhaA producing the highest amount of acetone. +NhaA acetone production was significantly increased compared to the acetone produced by +Cpa1 only. The WT strain produced the highest amount of butyric acid at 6 h; + Cpa1 took longer to reassimilate it. However, the three strains at 48 h consumed all the produced butyric acid, and no significant differences were found. Peak concentrations of butanol in + NhaA and + Cpa1 strains were similar to those of the WT cultures at approximately 9.32 ± 0.51 g/L; nevertheless, +NhaA produced significantly more butanol (10.06 ± 0.6 g/L) than +Cpa1 (8.09 ± 0.72 g/L) but not than WT. The most considerable discrepancies between strains were for lactic acid measurement observed at 24 h: 0.71 ± 0.17 g/L for the + Cpa1 strained compared to 0.41 ± 0.08 and 0.38 ± 0.1 g/L for the + NhaA and WT strain, but these differences were not significant. The ethanol concentrations produced by the + NhaA strain were significantly higher during the whole fermentation. This is interesting since butanol and ethanol are produced by the same enzyme, aldehyde-alcohol dehydrogenase (AAD). However, its primary role is the formation of butanol rather than ethanol (Zhao *et al.*, 2016). In this case, butanol production did not seem to affect ethanol production. Nair and Papoutsakis (1994) suggested that independent ethanol formation genes reside elsewhere on the chromosome since ethanol production was unaffected by mutations that lead to the loss of butanol and acetone. Here it might have been the same case; another gene was involved in ethanol production. + Cpa1 strain also produced a higher ethanol concentration during the first 24 h, but this was not significant. At the end, the ethanol production was 1.3 ± 0.12, 0.98 ± 0.9 and 0.92 ± 0.11 g/L for + NhaA, + Cpa1, and WT strain, respectively.

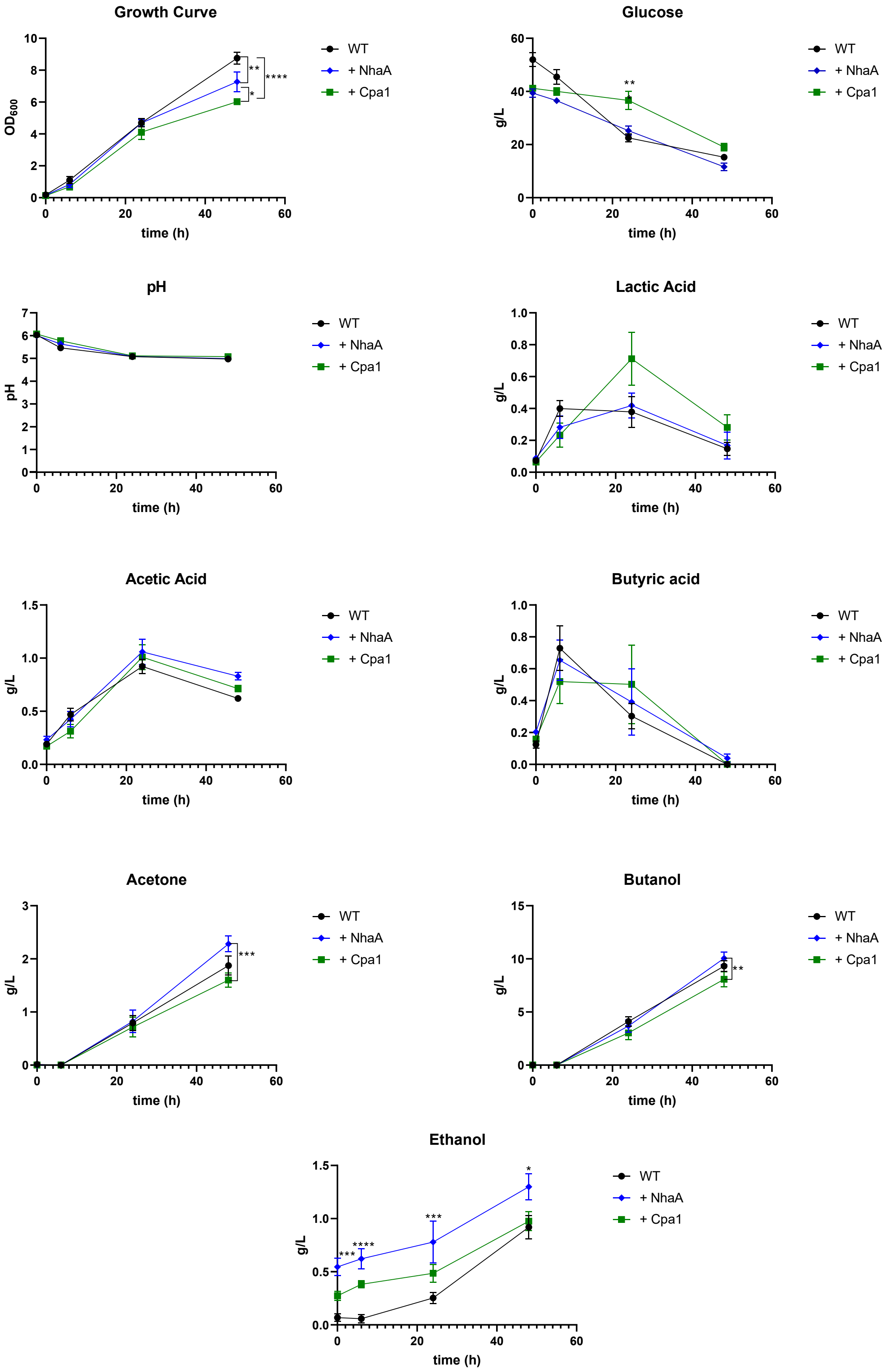

Growth Curve
OD600
time (h)
**
*
****
Glucose
g/L
**
pH
Lactic Acid
Acetic Acid
Butyric acid
Acetone
***
Butanol
**
Ethanol
***
****
***
*
WT
+ NhaA
+ Cpa1
0
20
40
60

Figure 5.13 OD600, pH and HPLC measurements of glucose, organic acids and solvents from the small bottle screening of +NhaA, +Cpa1 and WT *C. saccharoperbutylacetonicum* N1-4HMT strains
C. saccharoperbutylacetonicum N1-4HMT +NhaA (blue diamonds), +Cpa1 (green squares) and WT (black circles) grown on 50 g/L glucose in TYIR media. 30 mM MES was supplemented as a pH buffering agent. The growth curve was obtained by measuring the OD at 600 nm. pH was determined using a pHmeter, and HPLC analysed the other parameters. 2-way ANOVA was performed to calculate significant differences between strains. The * indicates a significant difference between the samples at the indicated time points. * = <0.05, **** = <0.0005.

The comparison between ΔGlpF and WT strain is shown in Figure 5.14. Two-way ANOVA was performed using GraphPad Prism 8.1. In the growth curve at 48 h, ΔGlpF grew significantly less than the WT and +NhaA strains. This might be a consequence of the glucose consumption at 24 h, which was also significantly lower, although, at 48 h, the glucose consumption was similar to the other strains. At 6 h, the pH of ΔGlpF was significantly higher than WT, which could mean that acid production was slower than WT. In acetic and butyric acid production, no significant difference was found. However, the butyric acid production peak took 18 h more in ΔGlpF. ΔGlpF lactic acid at 24 h was significantly higher than all of the strains with a concentration of 1.8 ± 0.18 g/L in comparison with 0.71 ± 0.17, 0.41 ± 0.08 and 0.38 ± 0.1 g/L that corresponds to +Cpa1, +NhaA and WT respectively. At 48 h ΔGlpF, lactic acid was still significantly higher than with WT and +NhaA. The final concentrations were 0.51 ± 0.12, 0.28 ± 0.8, 0.17 ± 0.08 and 0.15 ± 0.04 g/L that correspond to ΔGlpF, +Cpa1, +NhaA and WT, respectively. About ~ 70 % of the ΔGlpF lactic acid was reassimilated, which could be why ethanol production was similar to WT. However, acetone and butanol were significantly less at 48 h than WT and +NhaA strains.

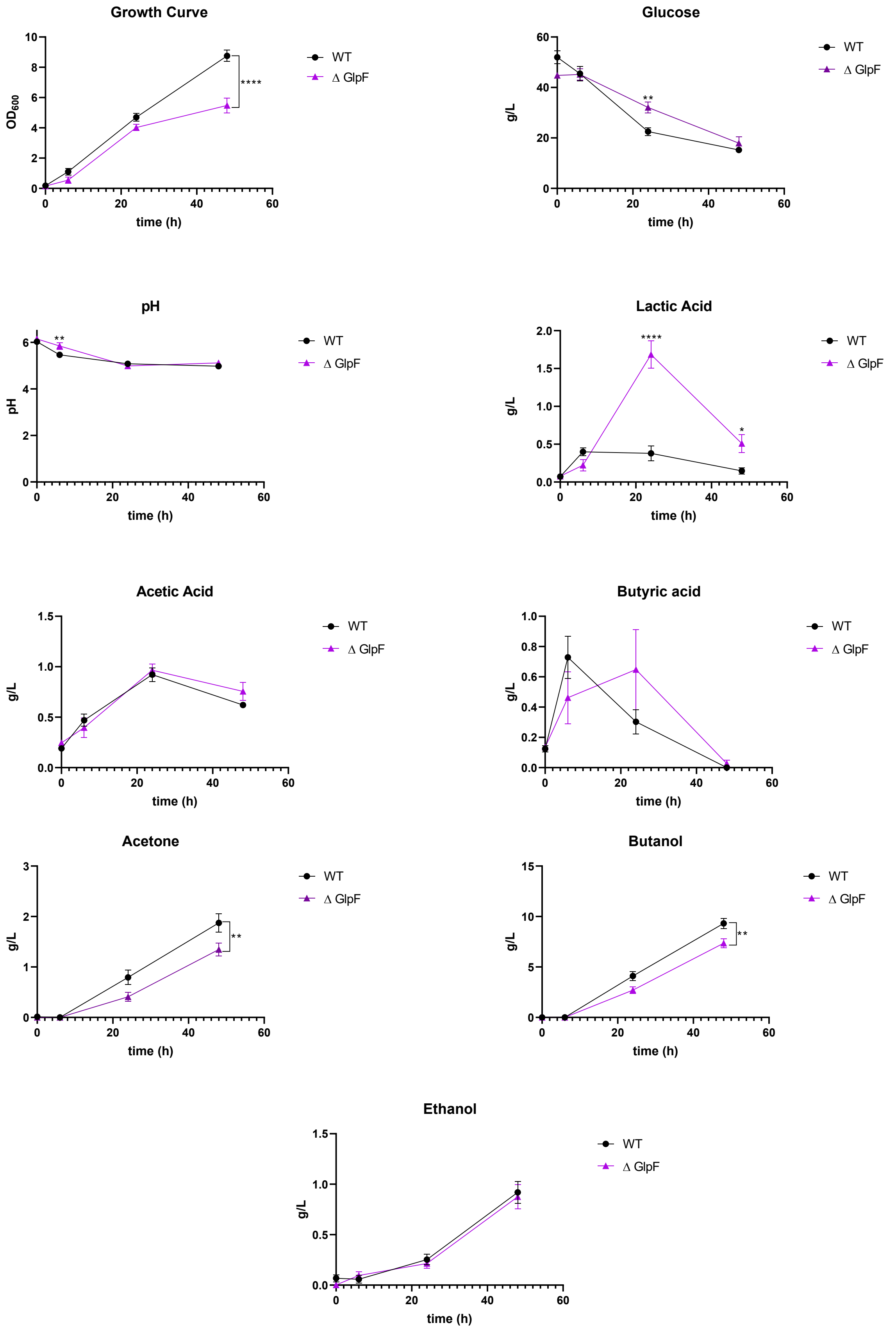

Growth Curve
OD_{600}
10
8
6
4
2
0
0
20
40
60
time (h)
****
WT
Δ GlpF
Glucose
g/L
60
40
20
0
**
time (h)
WT
Δ GlpF
pH
pH
6
4
2
0
**
time (h)
WT
Δ GlpF
Lactic Acid
g/L
2.0
1.5
1.0
0.5
0.0
****
*
time (h)
WT
Δ GlpF
Acetic Acid
g/L
1.5
1.0
0.5
0.0
time (h)
WT
Δ GlpF
Butyric acid
g/L
1.0
0.8
0.6
0.4
0.2
0.0
time (h)
WT
Δ GlpF
Acetone
g/L
3
2
1
0
**
time (h)
WT
Δ GlpF
Butanol
g/L
15
10
5
0
**
time (h)
WT
Δ GlpF
Ethanol
g/L
1.5
1.0
0.5
0.0
0
20
40
60
time (h)
WT
Δ GlpF

Figure 5.14 OD600, pH and HPLC measurements of glucose, organic acids and solvents from the small bottle screening of ΔGlpF and WT *C.saccharoperbutylacetonicum* N1-4HMT

C. saccharoperbutylacetonicum N1-4HMT ΔGlpF (purple triangle) and WT (black circles) grown on 50 g/L glucose in TYIR media. 30 mM MES was supplemented as a pH buffering agent. The growth curve was obtained by measuring the OD at 600 nm. pH was determined using a pHmeter, and HPLC analysed the remaining parameters. 2-way ANOVA was performed to calculate significant differences between strains. The * indicates a significant difference between the samples at the indicated time points. * = <0.05, **** = <0.0005. Error bars = SEM. n = 9.

In summary, +Cpa1 and ΔGlpF can produce more lactic acid and delay butyric acid reassimilation at 24 h. However, at 48 h, butyric acid was fully reassimilated, and most lactic acid was also reassimilated almost to reach WT values. Interestingly the overproduction of lactic acid was not reflected in their solvent production, and these two strains produced lower concentrations of solvents than WT. It might have been because there seemed to be a delay of 18 h compared to WT, and if the small bottle screening was extended for 72 h or more, an increase in solvent production might have been observed. Zhao *et al.* (2016) reported in *C. acetobutylicum* ATCC821 strains a 12 h delay for the solventogenesis phase in high sodium concentrations compared with the control cells. They believed the cell's metabolism and resources were diverted towards managing membrane transporters and channels, investing significant resources to stay functional. This could explain ΔGlpF strains; without the osmoregulator channel, the cell metabolism might have been diverted to manage its absence, delaying the solventogenic phase. These results agreed with GlpF having a role in the shift from the acidogenesis to the solventogenic phase but is not essential since its absence seemed to slow down the process.

5.2.3 Percentage yield and stoichiometry.

The economic feasibility of any fermentation is determined by the maximal possible: 1) conversion of the sugars, 2) product yield and selectivity, 3) reactor productivity, and 4) product separation and purification efficiency. Economically speaking, the most attractive fermentation is the one with the highest product selectivity and substrate utilisation, the one that can convert the majority of the sugar units to the desired product and produces none or very little of the by-products. Determination of the maximal theoretical yield and calculating the % yield (w/w) are of both fundamental and practical importance (Papoutsakis, 1984). Theoretical yield is the maximum amount of product given the amount of substrate. Actual yield is the amount of product that is actually formed. The per cent yield is the ratio of the actual yield to the theoretical yield expressed as a percentage (Agnew and Alviar-Agnew, 2019).

In theory, 1 mol of glucose can be converted into 1 mol of butyric acid (0.49 g/g) or 2 mols of lactic acid (1.00 g/g) in the acidogenesis phase. However, the actual yield is significantly decreased due to the production of biomass, assimilation of organic acids and the formation of other carbohydrates (Li *et al.*, 2020). For a future techno-economic analysis and to present the data for future scaling-up of the process, the mass balances, theoretical, actual and percentage yield were calculated.

For the following calculations, the stoichiometry equations used are shown in Table 5.1. The stoichiometry of Clostridia is very complex because it is a multiproduct fermentation. A full kinetic model was outside the aims of the project. However, Table 5.1 contains the relevant stoichiometric equations to quantify mass yields that can be used to calculate the economics of the ABE fermentation (Gapes, 2000)

Table 5.1 Stoichiometric reaction equations of ABE fermentation
The simplified stoichiometry reactions from glucose to one of the products from *C. saccharoperbutylacetonicum* are shown.

	Stoichiometric Reaction Equations
1 Butanol	$C_6H_{12}O_6 \rightarrow C_4H_{10}O$ (butanol) + $2CO_2 + H_2O$
2 Acetone	$C_6H_{12}O_6 + H_2O \rightarrow C_3H_6O$ (acetone) + $3CO_2 + 4H_2$
3 Ethanol	*$C_6H_{12}O_6 \rightarrow 2C_2H_6O$ (ethanol) + $2CO_2$*
4 Butyric acid	$C_6H_{12}O_6 \rightarrow C_4H_8O_2$ (butyric acid) + $2CO_2 + 2H_2$
5 Acetic acid	*$C_6H_{12}O_6 \rightarrow 3C_2H_4O_2$ (acetic acid)*
6 Lactic acid	*$C_6H_{12}O_6 \rightarrow 2C_3H_6O_3$ (lactic acid)*
7 Cell growth	$C_6H_{12}O_6 + 1.1429\ NH_4^+ \rightarrow 5.7143CH_{1.8}O_{0.5}N_{0.2}$ (biomass) + $0.2857HCO_3^- + 2.2857H_2O + 1.42H^+$

(Gapes, 2000, Merwe, 2010, Wresta *et al.*, 2021)

The theoretical yield calculation was based on the 50 g/L of glucose added to the media. A representation of how the butanol theoretical yield was calculated is shown.

$$\begin{aligned}&\text{Butanol theoretical yield}\\&\quad = 50\ \text{g}\ C_6H_{12}O_6 * \frac{1\ \text{mol}\ C_6H_{12}O_6}{180.156\ \text{g}\ C_6H_{12}O_6} * \frac{1\ \text{mol}\ C_4H_{10}O}{1\ \text{mol}\ C_6H_{12}O_6} * \frac{74.121\ \text{g}\ C_4H_{10}O}{1\ \text{mol}\ C_4H_{10}O}\\&\quad = 20.57\ \text{g}\ C_4H_{10}O\end{aligned}$$

Table 5.2 and Figure 5.15 summarised the different strains' theoretical, actual and % yields. Two-way ANOVA was performed, and no significant difference was found between the strains. From here, it can be observed that sugar is primarily converted to solvents, mainly butanol, in all strains. +NhaA is the one with higher % yields for solvents, and ΔGlpF with the lowest. ΔGlpF has the highest percentage yield of lactic acid. Overproduction of +NhaA does increase lactic acid production by ~ 20 % in comparison with WT strains; this result agrees with the one reported by Wu *et al.* (2013b). The strains did not improve organic acid production since their main end-products were still solvents, and the organic acids were produced in very low quantities.

Table 5.2 Comparison of the theoretical, actual and % yield of the different organic acids and solvents
The calculations of the theoretical (g/L), actual (g/L), and the percentage yield (% w/w) of the different strains: WT, +NhaA, +Cpa1 and ΔGlpF from the 48 h small bottle screening based on 50 g/L of glucose are shown.

		WT		+NhaA		+Cpa1		ΔGlpF	
Component	**Theoretical yield (g/L)**	**Actual yield (g/L)**	**% yield (w/w)**	**Actual yield (g/L)**	**% yield (w/w)**	**Actual yield (g/L)**	**% yield (w/w)**	**Actual yield (g/L)**	**% yield (w/w)**
Acetic acid	50	0.62 ± 0.02	1.24 ± 0.05	0.83 ± 0.02	1.66 ± 0.04	0.71 ± 0.04	1.43 ± 0.07	0.75 ± 0.09	1.51 ± 0.17
Lactic acid	50	0.14 ± 0.04	0.29 ± 0.09	0.17 ± 0.08	0.34 ± 0.17	0.28 ± 0.08	0.56 ± 0.17	0.51 ± 0.12	1.02 ± 0.23
Acetone	16.12	1.87 ± 0.17	11.62 ± 1.07	2.28 ± 0.15	14.16 ± 0.92	1.6 ± 0.13	9.93 ± 0.83	1.35 ± 0.13	8.35 ± 0.81
Butanol	20.57	9.32 ± 0.51	45.32 ± 2.47	10.06 ± 0.60	48.90 ± 2.90	8.09 ± 0.72	39.33 ± 3.50	7.35 ± 0.45	35.73 ± 2.20
Ethanol	25.57	0.92 ± 0.11	3.59 ± 0.44	1.29 ± 0.07	5.08 ± 0.26	0.98 ± 0.10	3.81 ± 0.40	0.88 ± 0.10	3.43 ± 0.40
Biomass	58.80	0.15 ± 0.01	0.26 ± 0.01	0.13 ± 0.01	0.22 ± 0.03	0.11 ± $3.42e^{-3}$	0.18 ± 0.01	0.10 ± 0.01	0.16 ± 0.01

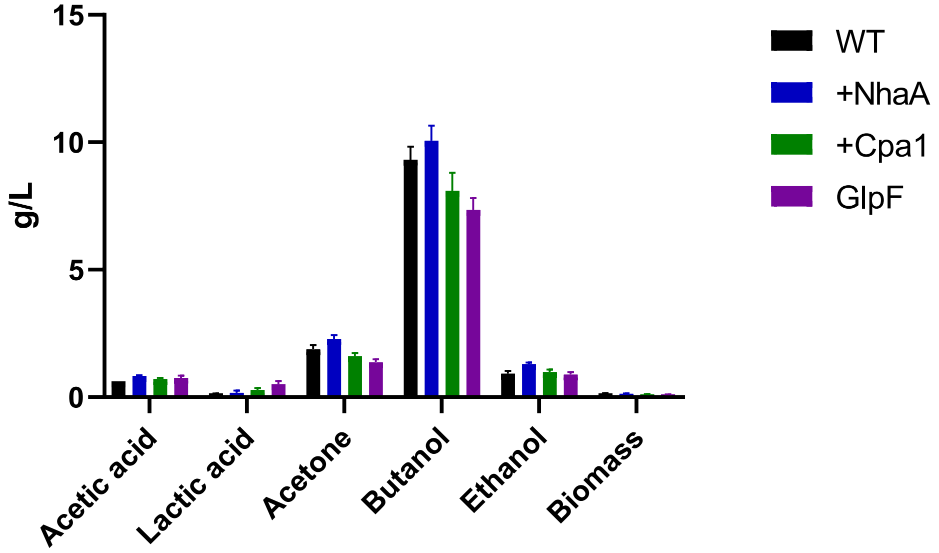
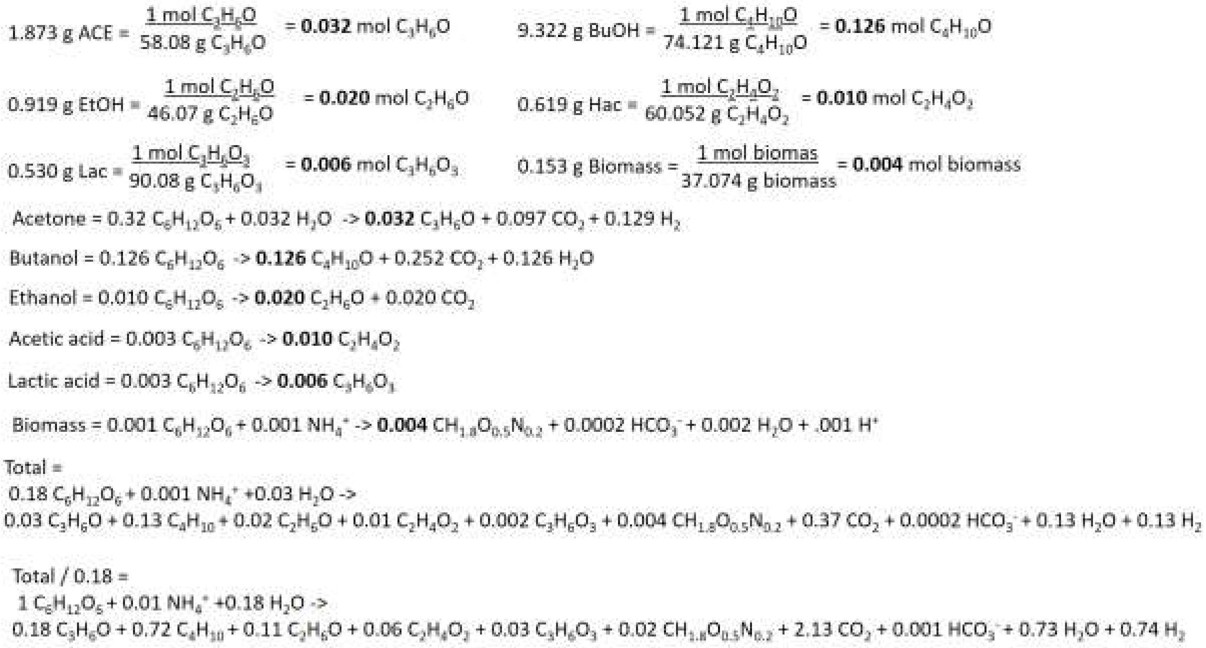


Figure 5.15 Yield of organic acids, solvents and biomass of the different strains
The actual yield (g/L) from **Table** 5.2 of acetic and lactic acid, acetone, butanol, ethanol and biomass from WT (black columns), +NhaA (blue columns), +Cpa1 (green columns) and ΔGplF (purple columns) were graphed. After 48 h of a small bottle screening based on 50 g/L of glucose, the data was obtained, and HPLC analysed the samples. Error bars = SEM, n=9. 2-way ANOVA was performed to calculate significant differences between strains.

Carbon is the element of most significant interest since the ultimate goal is to convert the carbon residing in biomass to a high-valuable chemical and reduce the carbon footprint. The carbon cost, which is the cost of greenhouse emissions (a tonne of carbon dioxide (tCO_2) emission), can be a great concern in the marketing and viability of a green product (Ricke *et al.*, 2018, Miller *et al.*, 2011). A carbon mass balance is a simple and rapid check on the consistency of reported or claimed experimental results and serves as a first-level tool for comparing processes. The carbon flux helps understand the carbon distribution and to highlight potential strategies to alter the metabolism to redistribute it towards the desired product (Pfromm *et al.*, 2010). The process has been simplified for ease of solving and implementation in process modelling. Stoichiometric coefficients are fundamental parameters in biological process modelling, which can help predict the process. The general equation for calculating carbon balances is shown below:

Glucose + NH_4^+ + H_2O → Acetone + Butanol +Ethanol + Acetic acid + Butyric acid + Lactic acid + Biomass + CO_2 + H_2 + H_2O +HCO_3^-

The moles of carbon were calculated using the grams obtained from each product's HPLC analysis, as explained on Figure 5.16.



$1.873\ \text{g ACE} = \frac{1\ \text{mol}\ C_3H_6O}{58.08\ \text{g}\ C_3H_6O} = \mathbf{0.032}\ \text{mol}\ C_3H_6O$

$9.322\ \text{g BuOH} = \frac{1\ \text{mol}\ C_4H_{10}O}{74.121\ \text{g}\ C_4H_{10}O} = \mathbf{0.126}\ \text{mol}\ C_4H_{10}O$

$0.919\ \text{g EtOH} = \frac{1\ \text{mol}\ C_2H_6O}{46.07\ \text{g}\ C_2H_6O} = \mathbf{0.020}\ \text{mol}\ C_2H_6O$

$0.619\ \text{g Hac} = \frac{1\ \text{mol}\ C_2H_4O_2}{60.052\ \text{g}\ C_2H_4O_2} = \mathbf{0.010}\ \text{mol}\ C_2H_4O_2$

$0.530\ \text{g Lac} = \frac{1\ \text{mol}\ C_3H_6O_3}{90.08\ \text{g}\ C_3H_6O_3} = \mathbf{0.006}\ \text{mol}\ C_3H_6O_3$

$0.153\ \text{g Biomass} = \frac{1\ \text{mol biomas}}{37.074\ \text{g biomass}} = \mathbf{0.004}\ \text{mol biomass}$

Acetone = $0.32\ C_6H_{12}O_6 + 0.032\ H_2O \rightarrow \mathbf{0.032}\ C_3H_6O + 0.097\ CO_2 + 0.129\ H_2$

Butanol = $0.126\ C_6H_{12}O_6 \rightarrow \mathbf{0.126}\ C_4H_{10}O + 0.252\ CO_2 + 0.126\ H_2O$

Ethanol = $0.010\ C_6H_{12}O_6 \rightarrow \mathbf{0.020}\ C_2H_6O + 0.020\ CO_2$

Acetic acid = $0.003\ C_6H_{12}O_6 \rightarrow \mathbf{0.010}\ C_2H_4O_2$

Lactic acid = $0.003\ C_6H_{12}O_6 \rightarrow \mathbf{0.006}\ C_3H_6O_3$

Biomass = $0.001\ C_6H_{12}O_6 + 0.001\ NH_4^+ \rightarrow \mathbf{0.004}\ CH_{1.8}O_{0.5}N_{0.2} + 0.0002\ HCO_3^- + 0.002\ H_2O + .001\ H^+$

Total =
$0.18\ C_6H_{12}O_6 + 0.001\ NH_4^+ + 0.03\ H_2O \rightarrow$
$0.03\ C_3H_6O + 0.13\ C_4H_{10} + 0.02\ C_2H_6O + 0.01\ C_2H_4O_2 + 0.002\ C_3H_6O_3 + 0.004\ CH_{1.8}O_{0.5}N_{0.2} + 0.37\ CO_2 + 0.0002\ HCO_3^- + 0.13\ H_2O + 0.13\ H_2$

Total / 0.18 =
$1\ C_6H_{12}O_5 + 0.01\ NH_4^+ + 0.18\ H_2O \rightarrow$
$0.18\ C_3H_6O + 0.72\ C_4H_{10} + 0.11\ C_2H_6O + 0.06\ C_2H_4O_2 + 0.03\ C_3H_6O_3 + 0.02\ CH_{1.8}O_{0.5}N_{0.2} + 2.13\ CO_2 + 0.001\ HCO_3^- + 0.73\ H_2O + 0.74\ H_2$

Figure 5.16 Stoichiometry calculations of WT strain
The calculations to obtain the stoichiometry formula of WT strain are shown. First, the gr obtained from the HPLC analysis at 48 h were converted to moles. The moles were then applied to the formulas from **Table** 5.1 to calculate the moles of sugar, water or ammonium needed to produce the final moles and to calculate the moles of carbon dioxide, dihydrogen, water and bicarbonate that were produced alongside the organic acids, solvents and biomass. Then all the moles were added and normalised based on the glucose moles.

WT:

$$C_6H_{12}O_{6\,(glucose)} + 0.005\ NH_4^+ + 0.186\ H_2O \rightarrow 0.186\ C_3H_6O_{(acetone)} + 0.727\ C_4H_{10}O_{(butanol)} + 0.115\ C_2H_6O_{(etanol)} + 0.06\ C_2H_4O_{2\,(acetic\ acid)} + 0.009\ C_3H_6O_{3\,(lactic\ acid)} + 0.024\ CH_{1.8}O_{0.5}N_{0.8\,(biomass)} + 2.13\ CO_2 + 0.749\ H_2 + 0.737\ H_2O + 0.001\ HCO_3^-$$

From this formula, the number of carbons of each element was multiplied by the moles produced at 48 h of fermentation, e.g., glucose had 6 carbons, and for the WT strain, acetone had 0.186 * 3 = 0.558 carbons. The carbon distribution of every strain is depicted in Figure 5.17. From this image, it can be observed that most of the carbon goes to butanol and carbon dioxide in every strain.

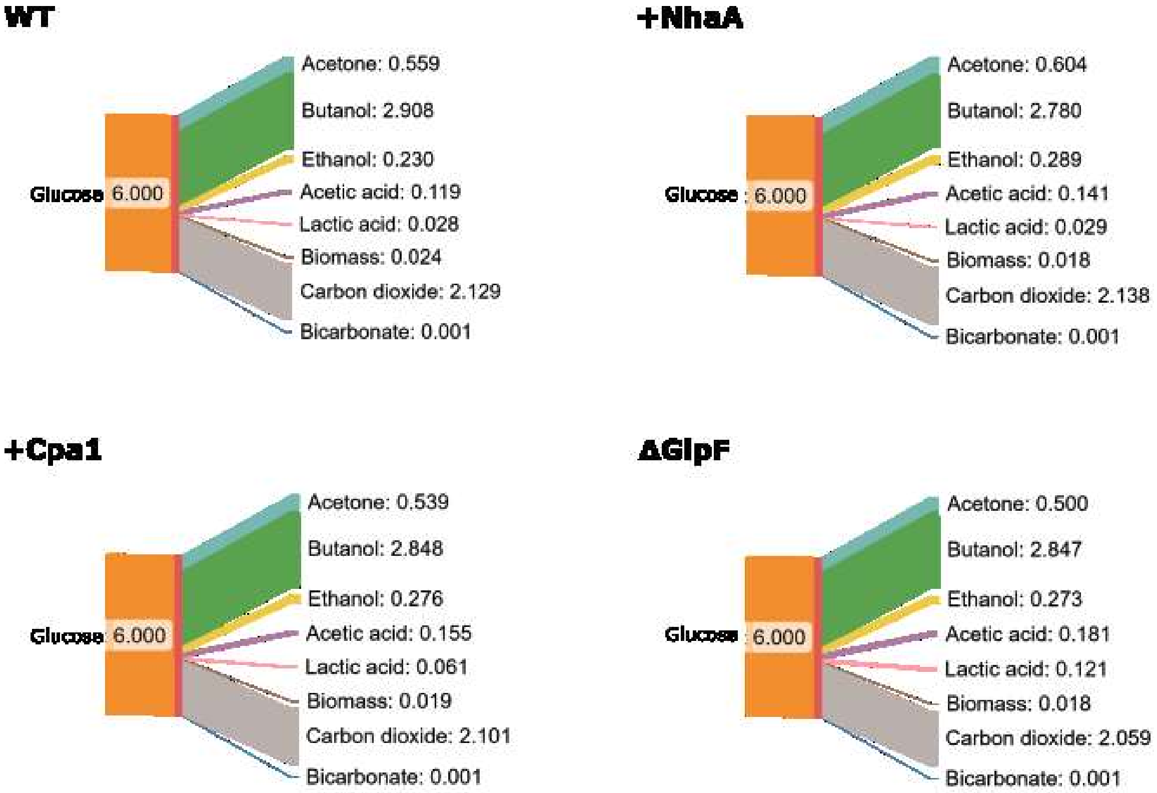


Figure 5.17 Carbon balance of WT, +NhaA, +Cpa1 and ΔGlpF after 48 h of fermentation
The carbon distribution of 1 glucose (6 carbon) after 48 h of fermentation at 32 ˚C is depicted for every strain. Made with SankeyMATIC.

These stoichiometries can be used for indications of up or down pathway regulations under the circumstances where the data were collected. Because the experiments were performed in small bottles, several parameters, such as carbon dioxide, were not measured. The development of a complete metabolic balance fell outside this project’s scope and is recommended for future use.

Up to this point, the production of organic acids as an end-product had yet to be achieved in any of the tested Clostridial strains.

5.2.4 Optimisation of WT *C. saccharoperbutylacetonicum* N1-4HMT batch fermentation for lactic acid production using the response surface method

The ability to run small-scale fermentation experiments is critical to better understanding the process of ABE fermentation, as it replicates a scaled-down version of what is seen in an industrial setting. The fermentation facility enables: (i) the characterisation of various growth conditions and feedstocks for industrial settings; (ii) the characterisation of any genetic engineering, the effect on the life cycle and ABE production (Monaghan, 2019).

As part of the scholarship obtained by CONACyT, a short research visit was carried out at the *Instituto Tecnologico y de Estudios Superiores de Occidente*, ITESO, a University in the state of Jalisco, Mexico. With the aim to increase the production of lactic and butyric acid and to establish the optimal conditions for batch fermentation of the strains, it was decided to perform a response surface method (RSM). RSM has been applied to analyse, optimise and evaluate the interactive effects of independent factors in numerous bioprocesses (Amani *et al.*, 2011). Using mathematical and statistical analysis, RSM provides an optimum condition for the target parameters. The optimum conditions can be obtained by fitting the complex relationships between independent variables in a small area with a simple quadratic polynomial model in which a statistically acceptable result can be predicted using a smaller number of experimental tests (Yang *et al.*, 2019). For the first part, the RSM was only performed on the WT strain; if successful, the optimisation would be performed on the other strains. Because it is known that low pH stimulates the solventogenic phase (Fernandez-Naveira *et al.*, 2019) and that salt concentration diverts the ABE fermentation from solventogenesis to acidogenesis (Maddox *et al.*, 1995); the factors to test using the RSM were temperature, pH and salt concentrations as the independent variables. The temperature was chosen because it impacts membrane fluidity, which can affect the production of organic acids or solvent tolerance. The responsible variables were growth, pH, acetic acid, butyric acid, lactic acid, acetone, butanol and ethanol. All samples collected were sent for commercial HPLC analysis to determine glucose, acetic acid, butyric acid, lactic acid, acetone, butanol and ethanol since the University does not have one. Lactic acid was detected at the University using their YSI equipment.

Using the software Design-Expert, 24 experimental runs were given (Table 5.3); the temperature range was between 27 and 42 ˚C, pH from 4.5 to 7.0 and salt (NaCl) % (w/v) from 0 to 4. Three more runs were added (25-27, depicted in bold), which are three replicates of the conditions used by the company, Bio-Cleave Ltd, and have been approved by the company as the best condition for butanol production. Replicates are different experimental runs with the same settings; their importance is that they are usually needed to calculate the variance of the measurements and validate the significance of the experiment. One of the advantages of performing an RSM is that its design contains the minimum replicates needed to estimate the effect of each factor and validate the method. For this experiment, the replicates were runs 5 with 6 and 19 with 20. It was decided to run the company conditions in triplicates.

Table 5.3 RSM runs given by Design-expert software with the factors of temperature, pH and NaCl on WT strains.
The 24 runs with the different conditions given by the Design-expert software are shown here. 25-27 runs are 3 replicates of the conditions used by Bio-Cleave Ltd to maximise butanol production.

	Factor 1	**Factor 2**	**Factor 3**
Run	**A: Temperature**	**B: pH**	**C: NaCl**
	°C	pH	% (w/v)
1	42.00	5.70	4.00
2	42.00	7.00	0.00
3	42.00	4.50	0.00
4	42.00	7.00	2.08
5	34.50	5.75	2.00
6	34.50	5.75	2.00
7	27.00	5.81	2.20
8	27.00	7.00	1.08
9	27.00	5.68	4.00
10	27.00	4.50	0.00
11	42.00	4.50	4.00
12	42.00	5.75	0.00
13	42.00	4.50	1.82
14	42.00	7.00	4.00
15	34.50	7.00	4.00
16	34.50	5.63	1.92
17	34.50	7.00	0.00
18	34.50	4.50	4.00
19	34.50	5.75	2.00
20	34.50	5.75	2.00
21	27.00	4.50	4.00
22	27.00	6.00	0.00
23	27.00	4.50	1.82
24	27.00	7.00	4.00
25	**32**	**6.5**	**0**
26	**32**	**6.5**	**0**
27	**32**	**6.5**	**0**

WT *C. saccharoperbutylacetonicum* N1-4HMT strains were grown on small bottles in 60 mL of TYIR media containing 50 g/L, supplemented with 5 g/L of $CaCO_3$ for pH control on different concentrations of salt. The experiment was carried out for 48 h.

Due to the COVID-19 pandemic and the closure of companies, the sample analysis by HPLC was impossible to perform. Therefore, the optimisation of WT strains on batch fermentation was based on lactic acid production alone. The results were analysed in the Design-Expert software. The 3D response surface model graphs show the interaction between the two factors. It is found that the steeper the curve, the more significant the effect was, and vice versa. From this analysis, it can be noted that pH had a more significant effect than the other factors.

Regarding the single factor level, the order of influence can be drawn as follows: pH > temperature > salt (Figure 5.18). Moreover, the shape of the contour, the lines below the graph, and whether it is elliptical or circular means that the interaction between the corresponding variables was significant or not. In Figure 5.18.a, the contour is rounder indicating a significant interaction between pH and temperature; compared with the other graphs b) and c), it does not appear to be an interaction between salt and pH or salt and temperature.

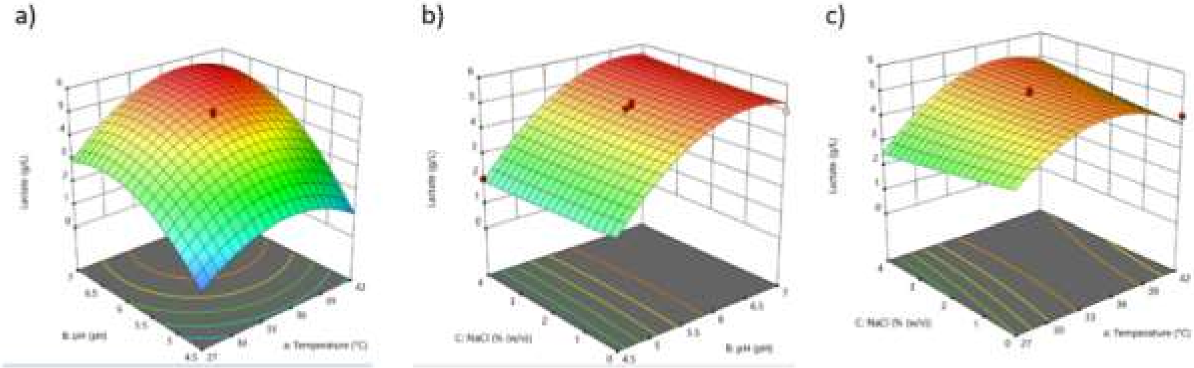


Figure 5.18 3D response surface plot of the WT strain for lactic acid production
The graphs show the effect of each independent variable with their interaction with the production of lactic acid. a) pH vs temperature. b) NaCl vs pH. c) NaCl vs temperature. Red indicates a higher production, and blue indicates minimal lactic acid production.

The Design-Expert software predicted the optimised fermentation conditions for lactic acid. It was decided to maximise the variables to choose the highest temperature, pH and salt concentration tested that would produce the highest amount of lactic acid to test the limits and tolerance of the WT strain. The software predicted that maximising the variables, a maximum of 5.91 g/L of lactic acid could be obtained under the following conditions: temperature of 42 ˚C, pH of 7 and 4% (w/v) NaCl. Mattiello-Franisco *et al.* (2021) reported that in *C. acetobutylicum* and *C. beijerinckii,* acidogenesis is favoured at a pH below 6.0, and the major lactic acid production was obtained when the pH was reduced to below 5.0. However, Drahokoupil and Patáková (2020) determined in *C. beijerinckii* NRRL B-598 that at pH 7 a high concentration of butyric acid was achieved simultaneously were minimal solvent production and no sporulation. It would have been better to have to full profile of organic acids and solvents. Based on the lactic acid production, these results correlate better with Drahokoupil and Patáková.

The verification experiments were conducted under the optimum conditions from RSM studies. The suggested conditions by the software (temperature of 42 ˚C, pH of 7 and 4 % (w/v) NaCl) were followed in two batch fermentations of 2 L. One fermenter was kept with the control conditions the company Bio-Cleave Ltd gave: temperature of 32 ˚C, pH of 6.3 and 0 % NaCl. The fermentation was carried out for 72 h. The fermenters with the suggested RSM conditions (temperature of 42 ˚C, pH of 7 and 4% of NaCl) gave at 72 hrs a production of lactic acid of 5.93 and 5.97 g/L in comparison with the control fermenter, which gave 5.75 g/L. These results validate the parameters given by the software. Optimisation of *C. saccharoperbutylacetonicum* N1-4HMT to increase lactic acid production can be obtained with RSM. The use of RSM offers several benefits. Because it is designed to test several variables simultaneously, it reduces the number of required experimental runs, which is time efficient and cost-effective. It is easier to visualise the influence of each variable and if there is a significant interaction between them. It is easy to adjust the optimisation of different responses. In this case, only lactic acid was measured. However, suppose butyric and solvent production could have been quantified. In that case, we could have modelled the shift between acidogenesis and solventogenesis and predicted the different conditions to obtain organic acids or solvents. Nevertheless, this experiment demonstrated that it is possible to modify the conditions of *C. saccharoperbutylacetonicum* N1-4HMT fermentation to increase lactic acid production.

The idea was to return and test all the strains in the fermenters. However, the COVID-19 pandemic made it impossible, and the continuation with further characterisation of the strains had to be carried out at Aston University.

5.2.5 Response surface method for optimisation of WT, +NhaA, +Cpa1 and ΔGlpF for organic acid production

A response surface method (RSM) was used to increase the different strains' organic acid production. Based on Maddox *et al.* (1995), where the addition of salt concentrations prevented the shift to solventogenesis, salt was one of the factors to be tested. It is known that the company Bio-Cleave Ltd reuses the water used for the fermentation; therefore, a small concentration of organic acids is presented at the start of the fermentation. Because the concentration of organic acids is one of the signals to shift to the solventogenic phase, initial butyric acid and lactic acid concentrations will also be tested to identify if reusing the water is advantageous for organic acid production.

The factors to test were butyric acid (g/L), lactic acid (g/L) and salt % (w/v) as the independent variables. Growth, pH, acetic acid, butyric acid, lactic acid, acetone, butanol and ethanol were the response variable. Samples were collected at 0, 6, 24, and 48 h and analysed by HPLC to determine glucose, acetic acid, butyric acid, lactic acid, acetone, butanol and ethanol concentration. Growth was determined by measuring the OD at 600 nm. The pH was determined using a pHmeter. +NhaA, +Cpa1, ΔGlpF and WT *C. saccharoperbutylacetonicum* N1-4HMT strains were grown in small bottles in 60 mL TYIR media containing 50 g/L, supplemented with 0.30 mM MES for pH control and different concentrations of lactic acid, butyric acid and salt. The different conditions were chosen based on articles that indicated that when 4 g/L of butyric acid or 5 g/L of lactic acid were added to the media, an increase in butanol production was observed (Oshiro *et al.*, 2010, Al-Shorgani *et al.*, 2018). 3 % (w/v) of NaCl inhibited growth in *Clostridium acetobutylicum* (Maddox *et al.*, 1995). An initial test was performed in all strains adding these three factors at once: 4 g/L butyric acid, 5 g/L lactic acid and 3 % (w/v) NaCl. None of the cells could grow; these conditions were too extreme for the bacteria to survive. It might have been due to the low pH obtained at < 4.5 and the low concentration of bacteria at the start of the screening. Because the proton antiporters have a role as pH sensors, they regulate the internal pH extruding protons; it was decided not to control the pH and to use a maximum of 2 g/L of butyric acid, 2.5 g/L of lactic acid and 2 % (w/v) of NaCl. Design-Expert 13 software was used for the quadratic model suggesting 21 runs. A response surface is a geometrical representation of a response variable plotted as a function of the independent variables. A three-level factorial design has a centre point included for each independent variable along with the high and low points, requiring three experiments for each independent variable. The RSM reduce the number of experiments required and generate models that fit the responses to quadratic or cubic equations (Wagner *et al.*, 2014). This design is fitted in quadratic equations. Table 5.4 shows the different runs used for the experiment. Cells were initially grown in the seed culture before being used for the inoculation of the small bottles, as described in section 2.3.2. The experiment was carried out over 48 h at 32 ˚C. Run 17, depicted in bold, is the control condition. The runs where the strains did not grow are indicated with an “x” in the table.

+Cpa1 and ΔGlpF strains grew with a media containing 2 % (w/v) of NaCl, whereas WT and +NhaA did not. Zhao *et al.* (2018) determined that a concentration of 1 % (w/v) of NaCl inhibits biomass growth and ABE production in *C. acetobutylicum* ATCC 824. No cell growth was observed at 2.5 %, which agrees with WT *C. saccharoperbutylacetonicum* N1-4HMT. This does not agree with the previous experiment in Mexico, where cells grew at 4 % (w/v); however, only lactic acid was measured. Results with +Cpa1 were as expected since it has been demonstrated in plants that its overexpression allows them to grow in salted conditions (Ma *et al.*, 2017). The +NhaA strain was expected to grow, although Wu *et al.* (2013b) also observed that overexpression of *nhaA* did not improve salt tolerance in *E. coli*. On the other hand, ΔGlpF was not expected to grow, as GlpF is involved in osmoregulation and its overexpression in *S. cerevisiae* significantly increased salt tolerance (Liu *et al.*, 2015). The ΔGlpF strain had the GlpF gene knock-out; therefore, its osmoregulation and salt tolerance should be disrupted. WT behaved as expected, although, in previous experiments, WT was able to grow in 4 % NaCl (w/v). The rest of the runs where the strains did not grow were those with a higher concentration of the factors, which indicates that overexpression of +NhaA or +Cpa1 did not increase organic acid tolerance. On this aspect, WT and ΔGlpF behaved as expected. An error must have happened for WT in run 11 since they grew in the replicate runs (1, 7, 9, 14, 16).

Table 5.4 RSM runs and different conditions are given by Design-expert 13 software to test the factors of lactic acid (g/L), butyric acid (g/L) and salt % (w/v) as the independent variables.
Run 17, depicted in bold, is the control condition. The runs where the strains did not grow are indicated with an “x” in the table. Depicted in red are the runs where non of the strains grow.

Run	Lactic acid (g/L)	Butyric acid (g/L)	Salt % (w/v)	WT	+NhaA	+Cpa1	ΔGlpF
1	1.25	1	1				
2	1.25	1	0				
3	0	2	0				
4	2.5	0	2	x	x	x	x
5	1.25	0	1				
6	1.25	1	2.7	x	x	x	x
7	1.25	1	1				
8	0	2	2	x	x	x	x
9	1.25	1	1				
10	0	0	2	x	x		
11	1.25	1	1	x			
12	2.5	0	0				
13	0	1	1				
14	1.25	1	1				
15	3.4	1	1	x	x	x	x
16	1.25	1	1				
17	**0**	**0**	**0**				
18	1.25	2.7	1	x	x	x	x
19	2.5	2	2	x	x	x	x
20	2.5	2	0	x	x	x	x
21	0	0	1				

The results were also analysed in Design-Expert software. There was no factor with a significant influence on the different parameters or strains. As mentioned before, the steeper the curve was, the more significant the effect was, and vice versa. In this case, there was no determinant factor like pH for lactic acid production in the previous study. However, adding salt did prevent the shift to the solventogenic phase (Table 5.5, Figure 5.19). However, no significant difference was found. At 1 % (w/v) NaCl, the +NhaA strain produced the highest amount of lactic acid and the lowest concentration of solvents. +Cpa1 produced the highest amount of butyric acid. It has been reported that acetoacetyl-CoA:acetate (butyrate) CoA-transferase, which is solely responsible for acetate and butyrate conversion into acetate-CoA and butyrate-CoA is inhibited by sodium ions, which it seems reasonable that a high sodium environment the extracellular levels of organic acids remains higher during the solventogenesis phase (Zhao *et al.*, 2016).

Interestingly at 2 % (w/v), ΔGlpF produced 5.05 g/L of lactic acid, 0.29 g/L of acetic acid, 0.20 g/L of butyric acid and 0.15 g/L of ethanol. Although +Cpa1 survived at 2 % (w/v) NaCl, there was minimal organic acid and solvents production. Because this experiment was performed to evaluate the relationship between three independent variables (lactic acid, butyric acid and salt) and no variable had a significant effect or interaction, it is necessary to repeat individual runs for statistics purposes. Due to the time left for the project, this was not feasible.

Table 5.5 Comparison of the theoretical, actual and % yield of the different organic acids and solvents of the different strains grown at 1 % (w/v) of NaCl
The calculations of the theoretical (g/L), actual (g/L), and the percentage yield (% w/w) of the different strains: WT, +NhaA, +Cpa1 and ΔGlpF from the 48 h small bottle screening based on 50 g/L of glucose and 1 % (w/v) of NaCl are shown.

		WT		+NhaA		+Cpa1		ΔGlpF	
Component	**Theoretical yield (g/L)**	**Actual yield (g/L)**	**% yield (w/w)**	**Actual yield (g/L)**	**% yield (w/w)**	**Actual yield (g/L)**	**% yield (w/w)**	**Actual yield (g/L)**	**% yield (w/w)**
Acetic acid	50	1.33	2.66	0.28	0.56	2.18	4.36	1.07	2.14
Butyric acid	24.45	2.14	8.75	1.09	4.46	4.21	17.22	2.44	9.98
Lactic acid	50	0.3	0.60	4.74	9.48	0.55	1.10	1.70	3.40
Acetone	16.12	0	0	0	0	0	0	0	0
Butanol	20.57	0.86	4.18	0	0	0.47	2.28	1.41	6.85
Ethanol	25.57	0.41	1.60	0.02	0.08	0.20	0.78	0.13	0.51
Biomass	58.80	0.07	0.12	0.13	0.03	0.03	0.06	0.04	0.06

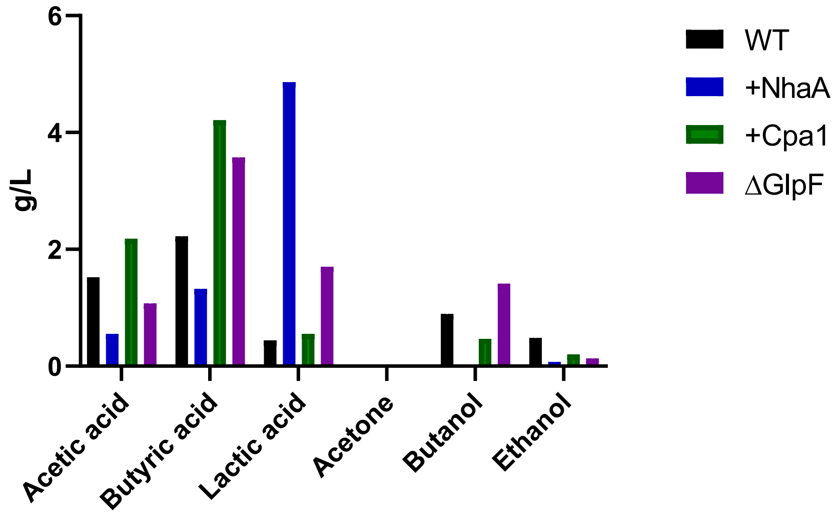


Figure 5.19 Yield of organic acids and solvents of the different strains when 1 % (w/v) of salt is added
The final yield (g/L) from RSM runs with only 1 % (w/v) of salt added (run 21) of acetic and lactic acid, acetone, butanol, ethanol and biomass from WT (black columns), +NhaA (blue columns), +Cpa1 (green columns) and ΔGplF (purple columns) were graphed. After 48 h of a small bottle screening based on 50 g/L and 1 % (w/v) of salt, the data was obtained, and HPLC analysed samples.

The general carbon balance equation was also calculated, but only the WT is shown.

WT:

$$C_6H_{12}O_{6\,(glucose)} + 0.007\ NH_4^+ \rightarrow 0.233\ C_4H_{10}O_{(butanol)} + 0.179\ C_2H_6O_{(ethanol)} + 0.489\ C_4H_8O_{2(butyric\ acid)} + 0.446\ C_2H_4O_{2\,(acetic\ acid)} + 0.067\ C_3H_6O_{3\,(lactic\ acid)} + 0.037\ CH_{1.8}O_{0.5}N_{0.8\,(biomass)} + 1.623\ CO_2 + 0.982\ H_2 + 0.248\ H_2O + 0.001\ HCO_3^-$$

The carbon distribution was calculated and is shown in Figure 5.20. From this diagram, the strain +NhaA is the one where more carbons were destined for butyric and lactic acid, has less solvent production (only ethanol) and less theoretical carbon dioxide.

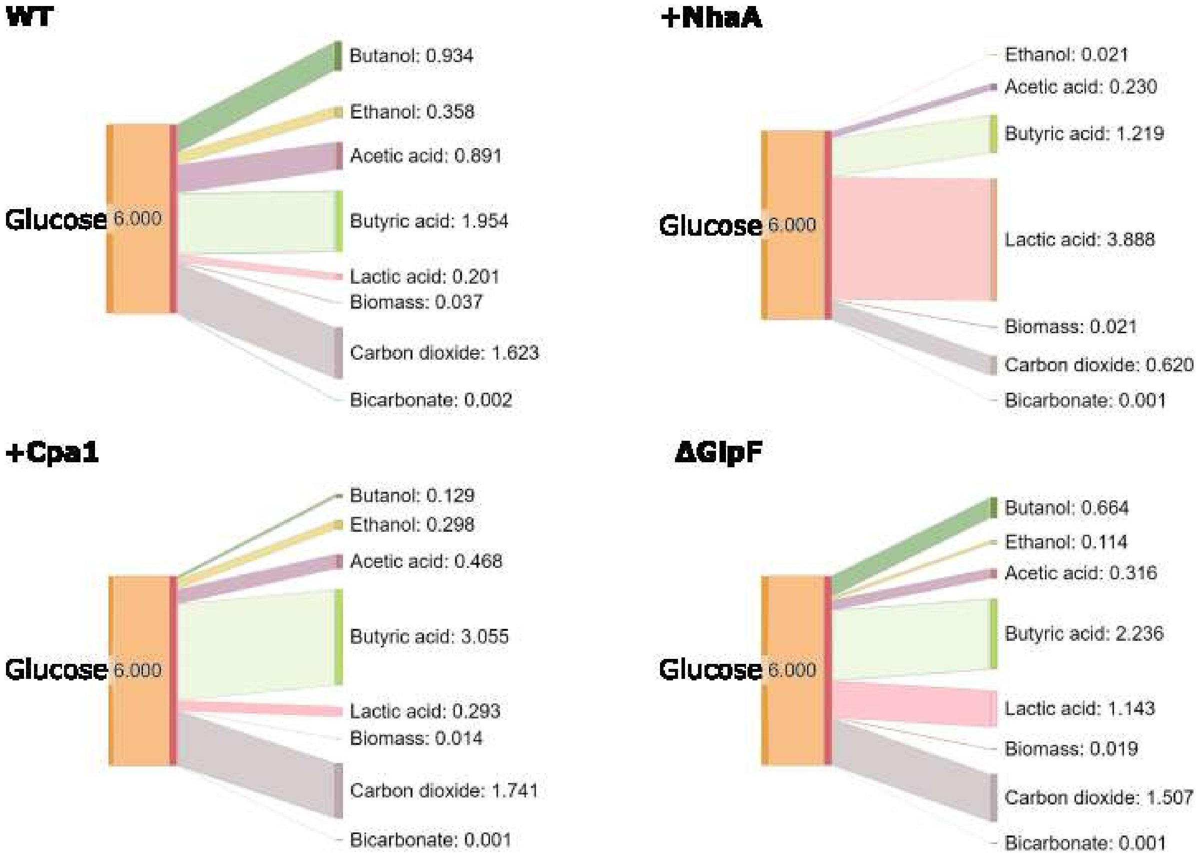
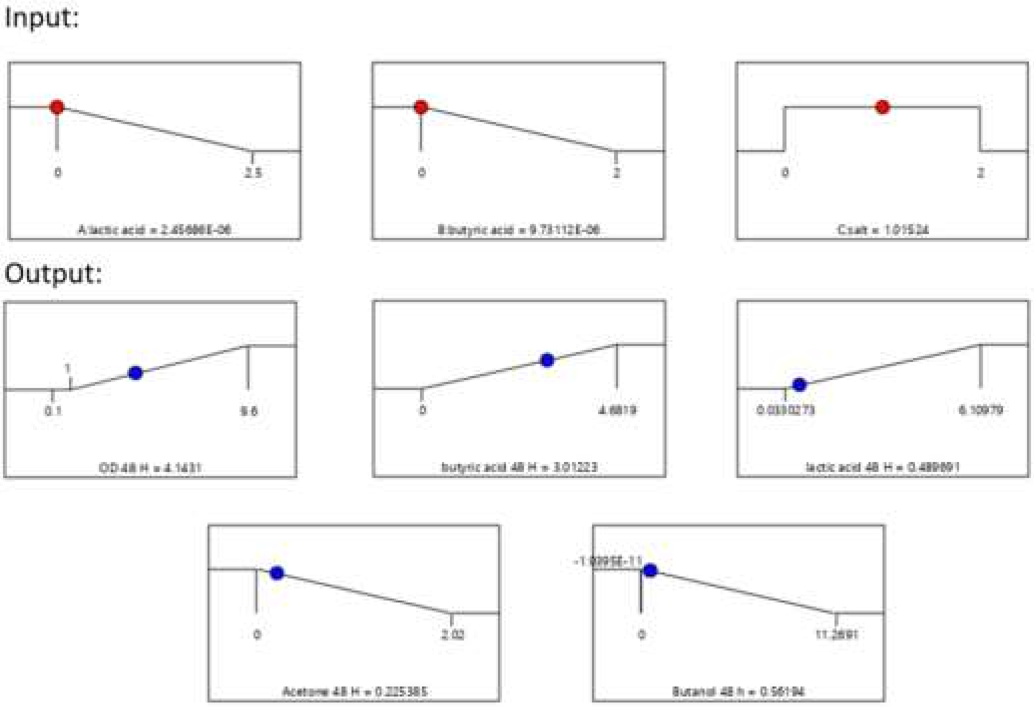


Figure 5.20 Carbon distribution of WT, +NhaA, +Cpa1, ΔGlpF strains after 48 h of fermentation in salt conditions

The carbon distribution of 1 glucose (6 carbon) after 48 h of fermentation at 32 ˚C is depicted for every strain. Made with SankeyMATIC.

Based on the software predictions, the parameters were set to maximise butyric and lactic acid and minimise solvent production. The maximum butyric acid production would be obtained with +Cpa1 and the maximum production of lactic acid with +NhaA (Figure 5.21). Ideally, the predictions obtained from the software should be tested in fermenters to validate them; however, this was not possible due to the time limit. Nevertheless, the use advantages of RSM for optimisation have been proved. Different conditions can be tested in fewer experimental runs and times. The response and significance of each variable can be easily analysed. After analysis, the optimisation of experimental conditions can be easily adjusted to the desired conditions. This type of study could also obtain the conditions to maximise butanol production.

+Cpa1


Input:
0
2.5
A:lactic acid = 2.45686E-06
0
2
B:butyric acid = 9.73112E-06
0
2
C:salt = 1.01524
Output:
1
0.1
9.6
OD 48 H = 4.1431
0
4.6819
butyric acid 48 H = 3.01223
0.0330273
6.10979
lactic acid 48 H = 0.489891
0
2.02
Acetone 48 H = 0.225385
-1.0395E-11
0
11.2691
Butanol 48 h = 0.56194

+NhaA

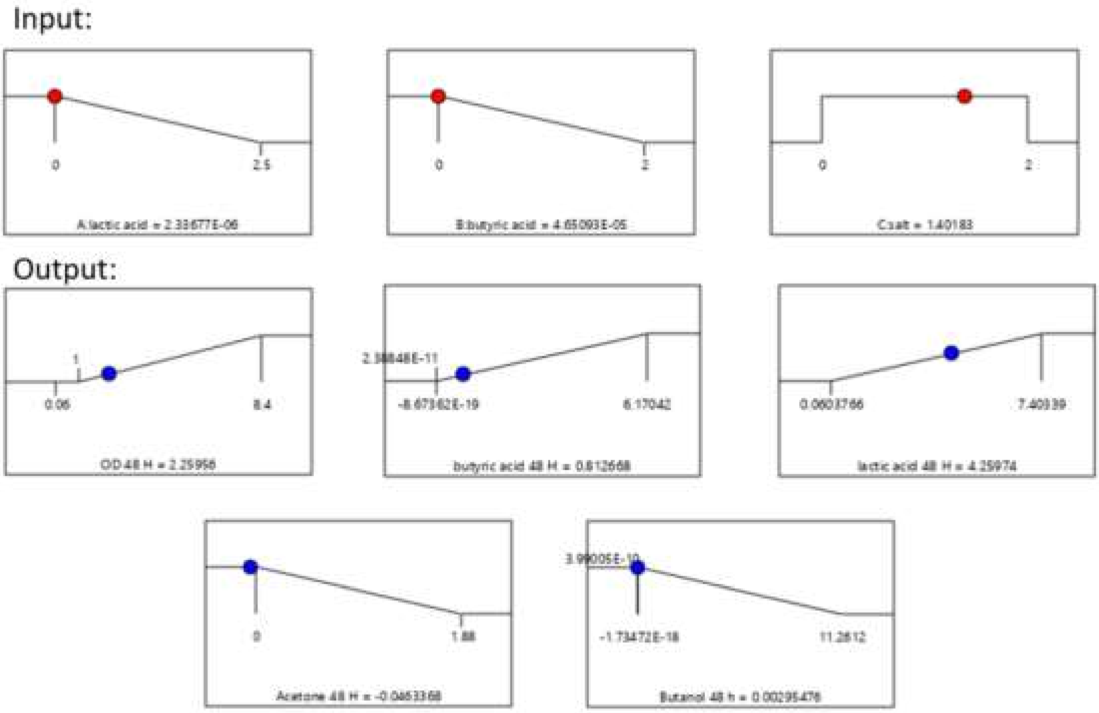

Input:
0
2.5
A:lactic acid = 2.33677E-06
0
2
B:butyric acid = 4.65093E-05
0
2
C:salt = 1.40183
Output:
1
0.06
8.4
OD 48 H = 2.25958
2.38848E-11
-8.67362E-19
6.17042
butyric acid 48 H = 0.812668
0.0603766
7.40339
lactic acid 48 H = 4.25974
0
1.88
Acetone 48 H = -0.0463368
3.99005E-10
-1.73472E-18
11.2612
Butanol 48 h = 0.00295478

Figure 5.21 Software prediction of +Cpa1 and +NhaA for organic acid production

The predictions from Design-Expert for +Cpa1 and +NhaA are shown. The inputs indicated at the top and in red are the tested factors and the initial concentrations needed to obtain the output; the values in blue of the different responses. Salt is in % (w/w). Lactic, butyric and acetic acid; acetone butanol and ethanol are in g/L. OD was measured at 600 nm.

Interestingly, adding butyric acid and/or lactic acid under salt concentration could recover the production of solvents (Figure 5.22). No significant result was obtained between the different strains. This was not ideal because, at 48 h, organic acids and solvents had a similar concentration, making the recovery of pure products more difficult and expensive. It would have been interesting to extend the fermentation of these experiments to see if, at 72 h, a complete shift to the solventogenic phase would have been observed.

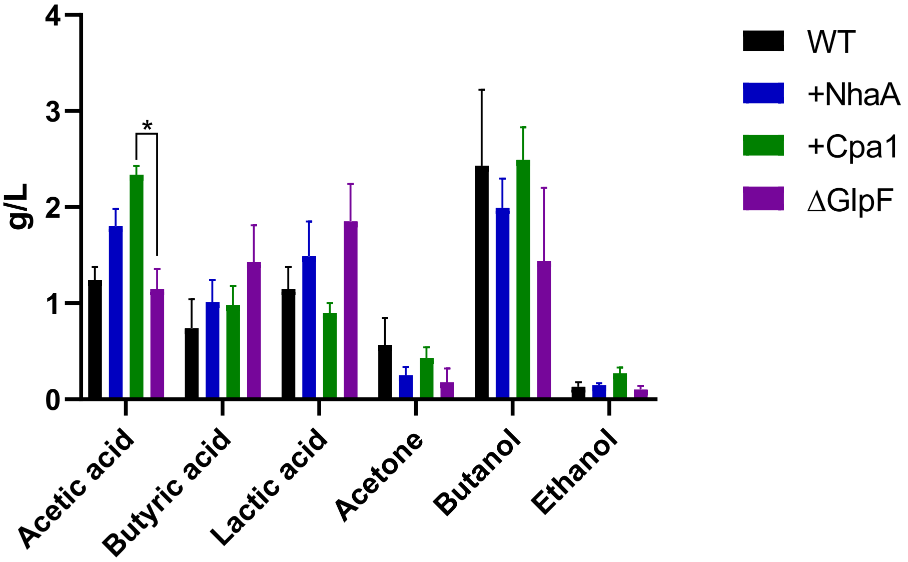


Figure 5.22 Graph of the g/L of the organic acids and solvents of the different strains when salt, butyric and lactic acid are added

The final yield (g/L) from RSM runs with salt, lactic and butyric acid added (runs 1, 7, 9, 11, 14, 16) of acetic and lactic acid, acetone, butanol, ethanol and biomass from WT (black columns), +NhaA (blue columns), +Cpa1 (green columns) and ΔGplF (purple columns) were graphed. After 48 of a small bottle screen based on 50 g/L of glucose, samples were obtained and analysed by HPLC. 2-way ANOVA was performed to calculate significant differences between strains. The * indicates a significant difference between the samples. * = <0.05. Error bars = SEM. n =6.

Oshiro *et al.* (2010) demonstrated that adding lactic acid can be used as a substrate and converted to butanol without increasing acetone production. Supplementing the fermentation medium with butyric acid has improved final ABE production (Al-Shorgani *et al.*, 2018).

To summarise, Maddox *et al.* (1995) demonstrated that adding salt prevents the strains from producing solvents. This can be attributed to the decrease in membrane permeability, which may affect nutrient uptake mechanisms. It has also been reported that acetoacetyl-CoA:acetate(butyrate)

CoA – transferase, which converts acetate and butyrate into acetate-CoA and butyrate-CoA, is inhibited by sodium ions (Zhao *et al.*, 2016). +NhaA, as previously reported, did not improve salt tolerance, but its overexpression increases lactic acid production with and without salt, but the production is higher when salt is added. +Cpa1, as previously reported, increased salt tolerance and butyric acid production when salt is added. Na^+ and H^+ are among the most prevalent ions within the cells and are essential in cell bioenergetics. An appropriate concentration of these ions within the cell is crucial for protein function. An overly high or low ion concentration is a potent stressor to the cell (Mondal *et al.*, 2021). Cation/proton antiporters act to adjust ionic balance (Ma *et al.*, 2017). It has been challenging to elucidate the exact mechanisms of CPAs. However, their activity is always believed to drive cytoplasmic pH into the desired neutral pH range. Once the neutral internal pH has been reached, they are inactive and no longer pose the risk of pH changes, excessive loss or accumulation of Na^+ (Calinescu *et al.*, 2016). Cation/proton antiporters might improve fermentations since if an optimal pH is to be maintained for the continued formation of organic acids, a base such as NaOH must be added into the system, leading to the accumulation of cations. Proton antiporters might improve organic acid production. It would have been good to test the cation/proton antiporters in fermenters with and without pH control to prove if *C. saccharoperbutylacetonicum* has a higher growth rate and accumulates more acid when a proton antiporter is overexpressed. Based on these results, +Cpa1 looked more efficient.

ΔGlpF did not behave as expected. It grew under salt concentrations and was observed with similar behaviour to the other strains. Its overexpression has been observed in the acidogenic phase, and it is believed that glycerol protects the cells during butanol production. From this experiment, GlpF might have a role in protecting the cells in the shift from acidogenesis to solventogenesis and during butanol production, but this role is not essential. More channels must be involved since, in its absence, the cells manage to produce solvents. GlpF produced the highest amount of lactic acid compared to the other strains. It would have been interesting to observe the behaviour of ΔGlpF strains longer, to 72 h, to see if reassimilation of lactic acid was delayed and if the cells managed to recover and produce similar amounts of butanol as WT at 48 h. Nevertheless, compared to the other strains, ΔGlpF is the less efficient strain since lactic acid production was alongside solvent production, making product recovery more difficult and costly. Even under salt conditions, ΔGlpF produced the highest amount of butanol at 48 h.

5.2.6. *C. saccharoperbutylacetonicum* N1-4HMT knock-out ΔPssA

ΔPssA was a strain constructed by Dr John Linney using CLEAVE^{TM} technology to delete the *pssA* gene (Figure 5.23). The phosphatidylserine synthase (PssA) represents an attractive locus for manipulating phospholipid synthesis since it is involved in lipid synthesis, forming phosphatidylserine (PS) lipids. In turn, PS lipids are converted into phosphatidylethanolamine lipids (PE), which means that regulating the expression of PssA can crudely regulate the amount of PE in the membrane. Its knock-out should lead to more PG, CL and inositol derivatives. Manipulating levels of PssA presents an excellent way to change the lipid composition of the bacteria's membrane and thus alter the related properties. This is interesting as protonated organic acids can cross the membrane and be reassimilated by the bacteria to produce the solvents. Altering the diffusion across the membrane to stop reassimilation and the shift to the solventogenic phase can be a strategy to increase the production of organic acids in solventogenic Clostridia.

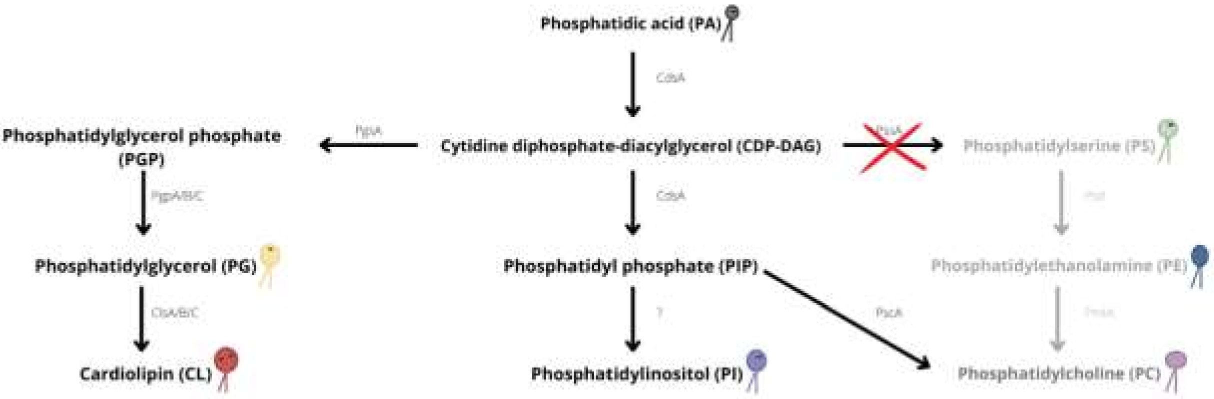
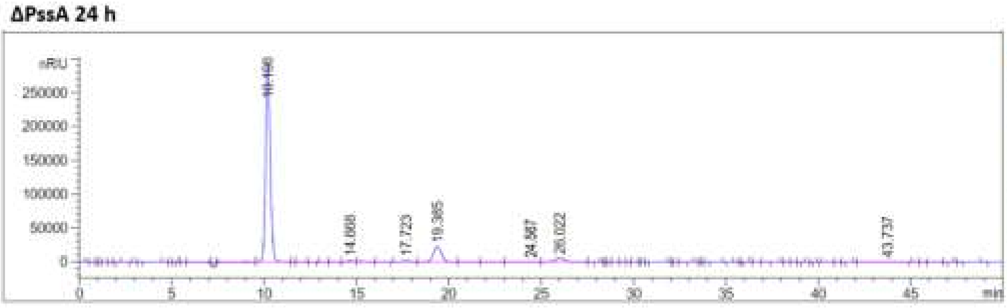
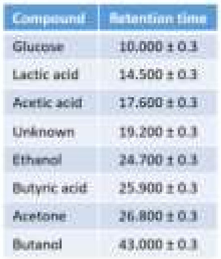

Phosphatidic acid (PA)
CdsA
Phosphatidylglycerol phosphate (PGP)
PgsA
Cytidine diphosphate-diacylglycerol (CDP-DAG)
PssA
Phosphatidylserine (PS)
PgpA/B/C
CdsA
Psd
Phosphatidylglycerol (PG)
Phosphatidyl phosphate (PIP)
Phosphatidylethanolamine (PE)
ClsA/B/C
?
PscA
Cardiolipin (CL)
Phosphatidylinositol (PI)
Phosphatidylcholine (PC)

Figure 5.23 Metabolic pathway of the phospholipid headgroup synthesis in bacteria and its alteration if PssA is knocked-out.
The metabolic pathway of the phospholipid headgroup synthesis in bacteria is shown with the route for strain engineering. A schematic diagram of the membrane phospholipids depicting the net charge (-: -1, - -: -2, otherwise, 0) and headgroup size is shown. PE and PC are neutral. PA, PS, PC and PC have small size headgroups. A red cross shows the knock-out of PssA and the headgroup lipids that will be altered on grey fond. Not all enzymes are present in Clostridia.

A small bottle screening was performed in 60 mL TYIR media with 50 g/L glucose and supplemented with 30 mM MES for pH control. Cells were initially grown in the seed culture before inoculating small bottles, as described in section 2.3.2. The experiment was carried out over 48 h at 32 ˚C. Samples were taken at 0, 6, 24 and 48 h. The product of acidogenesis and solventogenesis was analysed by HPLC. An HPLC example is shown in Figure 5.24.

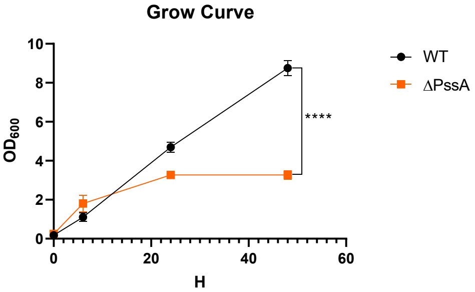


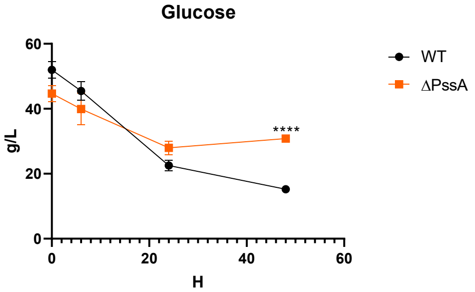


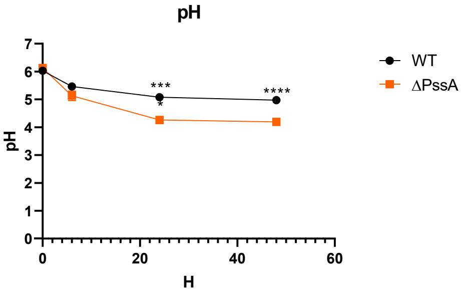
Compound	Retention time
Glucose	10.000 ± 0.3
Lactic acid	14.500 ± 0.3
Acetic acid	17.600 ± 0.3
Unknown	19.200 ± 0.3
Ethanol	24.700 ± 0.3
Butyric acid	25.900 ± 0.3
Acetone	26.800 ± 0.3
Butanol	43.000 ± 0.3

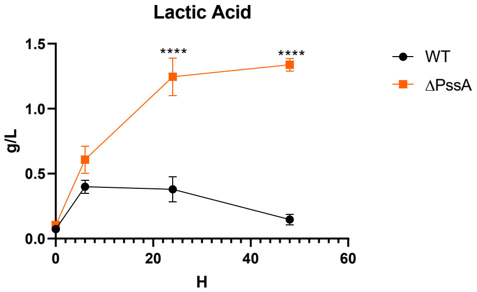
Figure 5.24 HPLC example of ΔPssA strain at 24 h
Chromatographic data for the analysis of ΔPssA strain at 24 h. Analysis was performed using Bio-rad Aminex HPX087H column. Mobile phase 5 mM H_2SO_4 at 0.6 mL/min; 20 µL injection volume. Run time: 50 min. Lac: lactic acid. HAc: acetic acid. BuAc: butyric acid.

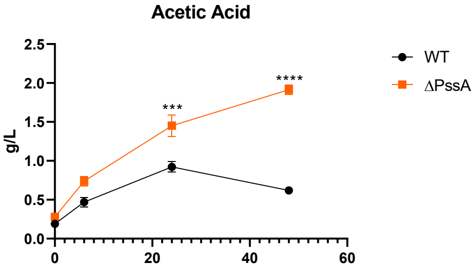
A comparison of the WT with the ΔPssA is shown in Figure 5.25. The strain was compared with the other strains (WT, +NhaA, +Cpa1 and ΔGlpF) with a two-way ANOVA. The ΔPssA strain exhibited the desired characteristics. The ΔPssA grew at 48 h significantly less than the rest, with an average OD_{600} of 3.4 compared with the 8.76 from WT. Its glucose consumption at 48 h was also significantly less than the rest of the strains correlating with its growth. At 6 h, the pH was significantly less than ΔGlpF, +NhaA and +Cpa1, and from 24 h was significantly less than all of the strains. This correlates with organic acid production. For acetic acid, at 6 h, it was significantly more than +Cpa1 and from 24 in comparison with all of the strains. Butyric acid was significantly more from 6 h and lactic acid from 24 h compared to all strains. The final concentration of organic acids was 1.91 ± 0.06 g/L of acetic acid, 5.65 ± 0.30 g/L of butyric acid, and 1.34 ± 0.05 g/L of lactic acid in comparison with WT, which gave 0.62 ± 0.02 g/L of acetic acid, 0 g/L of butyric acid and 0.15 ± 0.04 g/L of lactic acid. In the other strains, the highest butyric acid production was obtained on the +Cpa1 6.10 g/L at 24 h, 1 % NaCl (w/v) and 1 g/L of butyric acid. At 48 h, butyric acid was reassimilated, and the final concentration for +Cpa1 was 2.46 g/L. On lactic acid, the highest concentration was obtained by +NhaA with 1 % (w/v) for a final concentration of 4.86 g/L at 48 h. The lipid membrane change might have affected the reassimilation of organic acids because the ΔPssA strain failed in solvent production; 0 g/L of acetone and ethanol and 0.13 ± 0.03 g/L of butanol which was significantly less in comparison with the rest strains. WT produced 1.87 ± 0.17 g/L of acetone, 0.92 ± 0.11 g/L of ethanol and 9.32 ± 0.51 g/L of butanol at 48 h.

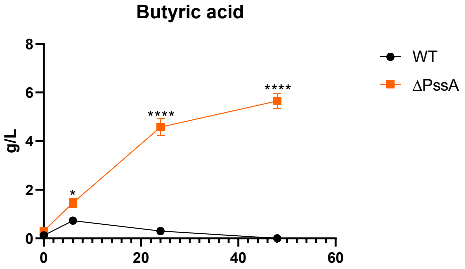

Grow Curve
OD600
0
2
4
6
8
10
0
20
40
60
H
****
WT
ΔPssA

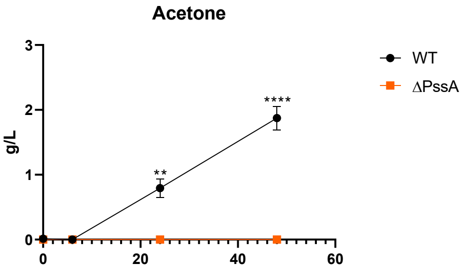

Glucose
g/L
0
20
40
60
0
20
40
60
H
****
WT
ΔPssA

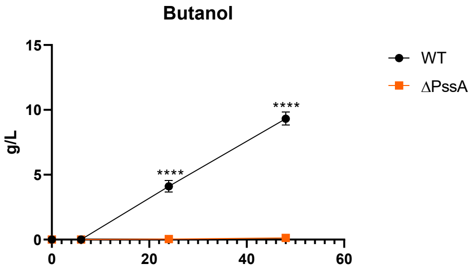

pH
pH
0
1
2
3
4
5
6
7
0
20
40
60
H
***
****
WT
ΔPssA


Lactic Acid
g/L
0.0
0.5
1.0
1.5
0
20
40
60
H
****
****
WT
ΔPssA


Acetic Acid
g/L
0.0
0.5
1.0
1.5
2.0
2.5
0
20
40
60
***
****
WT
ΔPssA


Butyric acid
g/L
0
2
4
6
8
0
20
40
60
*
****
****
WT
ΔPssA


Acetone
g/L
0
1
2
3
0
20
40
60
**
****
WT
ΔPssA


Butanol
g/L
0
5
10
15
0
20
40
60
****
****
WT
ΔPssA

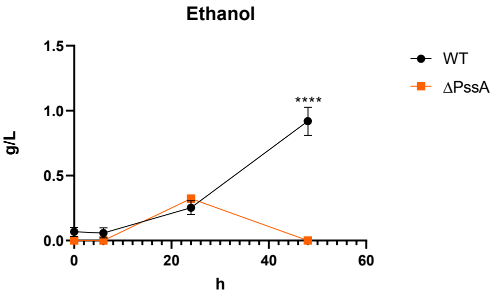

Ethanol
g/L
0.0
0.5
1.0
1.5
0
20
40
60
h
****
WT
ΔPssA

Figure 5.25 Small bottle screening measurements for ΔPssA and WT C. saccharoperbutylacetonicum N1-4HMT
C. saccharoperbutylacetonicum N1-4HMT ΔPssA (orange square) and WT (black circles) grown on 50 g/L glucose in TYIR media. 30 mM MES was supplemented as a pH buffering agent. The growth curve was obtained by measuring the OD at 600 nm. pH was calculated using a pHmeter, and HPLC analysed the rest. 2-way ANOVA was performed to calculate significant differences between strains. The (*) indicates a significant difference between the samples at the indicated time points. * = <0.05, **** = <0.0005. Error bars = SEM. n = 3.

5.2.6.1. Percentage yield and stoichiometry.

The theoretical, actual and % yield was calculated for ΔPssA and compared with the results obtained for WT based on the stoichiometric equations in Table 5.1.

$$\begin{aligned}&\text{Butyric acid theoretical yield}\\&= 50\ \mathrm{g\ C_6H_{12}O_6} * \frac{1\ \mathrm{mol\ C_6H_{12}O_6}}{180.156\ \mathrm{g\ C_6H_{12}O_6}} * \frac{1\ \mathrm{mol\ C_4H_8O_2}}{1\ \mathrm{mol\ C_6H_{12}O_6}} * \frac{88.106\ \mathrm{g\ C_4H_8O_2}}{1\ \mathrm{mol\ C_4H_8O_2}}\\&= 24.45\ \mathrm{g\ C_4H_8O_2}\end{aligned}$$

Table 5.6 Comparison of the theoretical and actual yield, % yield for butanol, acetone, ethanol, lactic acid, butyric acid and biomass between WT and ΔPssA strain
The calculations of the theoretical (g/L), actual (g/L) and the percentage yield (% w/w) of the different strains: WT and ΔPssA from the 48 h small bottle screening based on 50 g/L of glucose are shown. Error = SEM.

		WT		ΔPssA	
Component	**Theoretical yield (g/L)**	**Actual yield (g/L)**	**% yield (w/w)**	**Actual yield (g/L)**	**% yield (w/w)**
Acetic acid	50	0.62 ± 0.02	1.24 ± 0.05	1.91 ± 0.06	3.83 ± 0.12
Butyric acid	24.45	0.00	0.00	5.65 ± 0.30	23.11 ± 1.24
Lactic acid	50	0.14 ± 0.04	0.29 ± 0.09	1.34 ± 0.05	2.68 ± 0.09
Acetone	16.12	1.87 ± 0.17	11.62 ± 1.07	0.00	0.00
Butanol	20.57	9.32 ± 0.51	45.32 ± 2.47	0.13 ± 0.03	0.62 ± 0.17
Ethanol	25.57	0.92 ± 0.11	3.59 ± 0.44	0.00	0.00
Biomass	58.8	0.15 ± 0.01	0.26 ± 0.01	0.07 ± $4.50e^{-3}$	0.12 ± .01

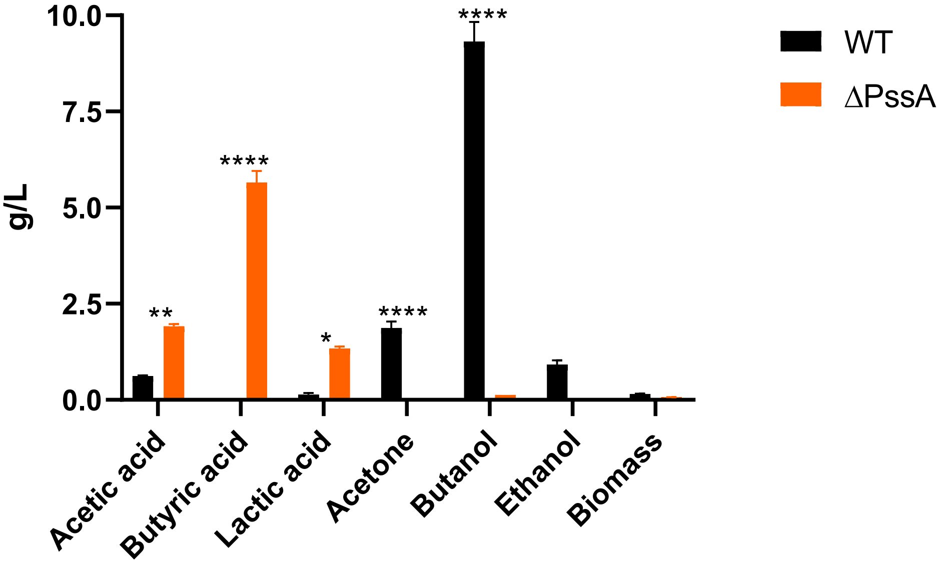


Figure 5.26 g/L of the organic acids and solvents after 48 h of fermentation from WT and ΔPssA strains
The actual yield (g/L) from Table 5.6 of acetic and lactic acid, acetone, butanol, ethanol and biomass from WT (black columns) and ΔPssA (orange columns) were graphed. After 48 of a small bottle screen based on 50 g/L of glucose, samples were obtained and analysed by HPLC. Error bars = SEM, n=3. 2-way ANOVA is performed to calculate significant differences between strains. The * indicates a significant difference between the samples at the indicated time points. * = <0.05, **** = <0.0005.

The full stoichiometry was also calculated, as described in section 5.2.3. The equation for calculating carbon balances is shown below

$$\text{Glucose} + NH_4^+ + H_2O \rightarrow \text{Acetone} + \text{Butanol} + \text{Ethanol} + \text{Acetic acid} + \text{Butyric acid} + \text{Lactic acid} + \text{Biomass} + CO_2 + H_2 + H_2O + HCO_3$$

WT:

$$C_6H_{12}O_{6\,(glucose)} + 0.005\ NH_4^+ + 0.186\ H_2O \rightarrow 0.186\ C_3H_6O_{(acetone)} + 0.727\ C_4H_{10}O_{(butanol)} + 0.115\ C_2H_6O_{(etanol)} + 0.060\ C_2H_4O_{2\,(acetic\ acid)} + 0.009\ C_3H_6O_{3\,(lactic\ acid)} + 0.024\ CH_{1.8}O_{0.5}N_{0.8\,(biomass)} + 2.13\ CO_2 + 0.749\ H_2 + 0.737\ H_2O + 0.001\ HCO_3^-$$

ΔPssA:

$$C_6H_{12}O_{6\,(glucose)} + 0.004\ NH_4^+ \rightarrow 0.020\ C_4H_{10}O_{(butanol)} + 0.378\ C_2H_4O_{2\,(acetic\ acid)} + 0.761\ C_4H_8O_{2(butyric\ acid)} + 0.176\ C_3H_6O_{3\,(lactic\ acid)} + 0.022\ CH_{1.8}O_{0.5}N_{0.8\,(biomass)} + 1.564\ CO_2 + 1.526\ H_2 + 0.029\ H_2O + 0.001\ HCO_3^-$$

From this, the carbon distribution was calculated and depicted in Figure 5.27

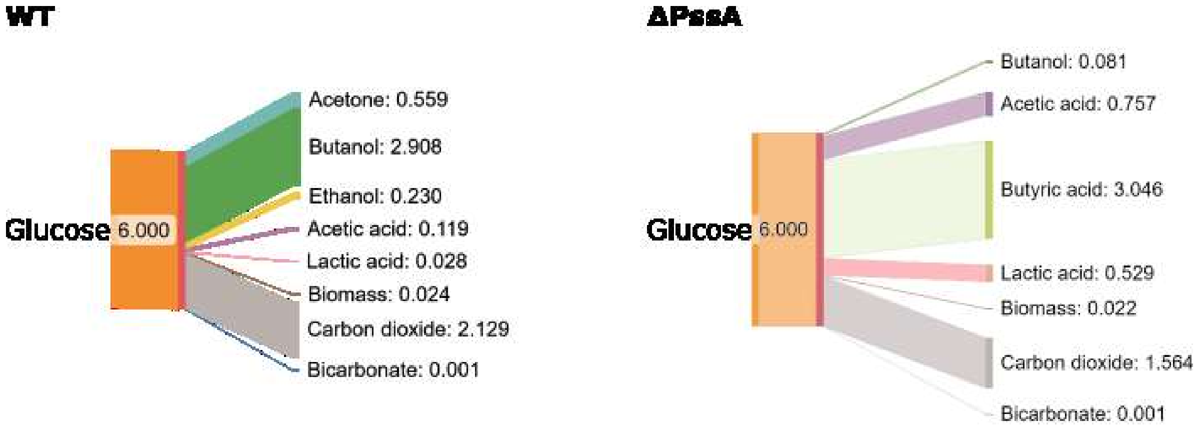


Figure 5.27 Carbon distribution of WT and ΔPssA strain after 48 h of fermentation
The carbon distribution of 1 glucose (6 carbon) after 48 h of fermentation at 32 ˚C is depicted for every strain. Made with SankeyMATIC

Based on the carbon balance, it can be confirmed that the sugar carbons go towards the production of organic acids rather than solvents. This was the behaviour expected to be seen in the ΔGlpF strain. These findings will be followed up in subsequent studies.

In summary, three strategies were applied to increase the production of organic acids. For strategy one, the plasmid-based achieved the overexpression of +NhaA and +Cpa1 on *C. saccharoperbutylacetonicum* N1-4HMT. +NhaA did not increase salt tolerance but increased lactic acid production, as reported. +Cpa1 had an increased sodium tolerance compared to WT and +NhaA, which also agrees with what has been reported. For the second strategy, the knock of GlpF was successfully achieved using the CLEAVE^{TM} technology. ΔGlpF also increased lactic acid production and had a higher salt tolerance than WT and +NhaA. This does not agree with what has been reported in *Aspergillus glaucus* and *Arabidopsis thaliana (Liu et al., 2015)*. None of these three strains improved butyric acid production after 48 h fermentation. For the third strategy, between temperature, salt concentration and pH, pH was the factor that significantly impacted lactic acid production in the WT strain. Salt concentration did prevent the shift from acidogenic to solventogenic phase. +NhaA with 1 % NaCl (w/v) produced the highest lactic acid concentration. +Cpa1 with 1 % (w/v) NaCl and 1 g/L of butyric acid produced the highest butyric acid at 24 and 48 h compared with the other strains. Response surface methodology helps test several parameters with fewer experimental runs. ΔPssA, a strain obtained with the CLEAVE^{TM} where the gene *pssA* was knocked-out, presented interesting results: very low solvent production and increased organic acid production.

6. General discussion, alternative strategies, conclusions and future work

6.1 General discussion

Organic acids are valuable building block chemicals with increasing market demand for future biorefining applications (Mattiello-Franisco *et al.*, 2021, Guan and Liu, 2020). For this project, lactic and butyric acids were chosen; lactic acid because it is in the top 10 high-value chemicals chosen by the US Department of Energy (Bozell and Petersen, 2010) and butyric acid because although it is not on the top 10, it is the precursor of butanol. Butanol is a commercially valuable solvent (Berezina *et al.*, 2009). Lactic acid is mainly produced by microorganisms (lactic acid bacteria, LAB). The majority of butyric acid is chemically produced (Jaros, 2014). Undissociated organic acids pass freely through the cell membrane and dissociate upon entry into the slightly alkaline cytoplasm releasing protons that lower internal pH and anions. High organic acid titres add protons to the culture, decreasing pH (Warnecke and Gill, 2005).

This project aimed to increase the organic acid production of *C. saccharoperbutylacetonicum* N1-4HMT by increasing its organic acid tolerance. To accomplish this aim, several approaches were taken. The first approach was to understand organic acids' effect on its membrane and the protective mechanisms the cell takes to change the lipid composition. This project developed a miniaturised IV method to use 1 mg of sample and a final assay volume of 300 µL. The miniaturised IV was used to determine the average number of double bonds per FA in *E. coli* and yeast cell membranes, obtaining results that correlate with reported values. Besides being cheap, quick, and easy to perform and interpret, the miniaturised IV method will provide additional and valuable information about the membrane composition. The idea was to perform the miniaturised IV method on *C. saccharoperbutylacetonicum* N1-4HMT lipid membranes grown at different time points to observe changes in the unsaturation of the FAs and correlate them with membrane fluidity.

The second was to choose two membrane proteins that regulate pH homeostasis and increase organic acid tolerance when overexpressed. The goal was to express them in *E. coli* and purify them to characterise their role in increasing the tolerance of *C. saccharoperbutylacetonicum* to organic acids. A third membrane protein, the glycerol facilitator (GlpF), was chosen because it is known to be overexpressed in the acidogenic phase. Glycerol can maintain intracellular water balance, and by avoiding the production of hydroxyl radicals and scavenging reactive oxygen species, glycerol can stabilise proteins and prevent oxidative damage (Zhou *et al.*, 2020). Glycerol-3-phosphate is a major substrate for the biosynthesis of glycerophospholipids, involved in the composition and fluidity of the cell membrane (Petrov *et al.*, 2021). Accumulation of glycerol and overexpression of GlpF are believed to protect the bacteria during butanol production (Patakova *et al.*, 2022, Zhou *et al.*, 2020).

The third approach was genetically modifying *C. saccharoperbutylacetonicum* N1-4HMT to overexpress the chosen proton antiporters and knock out GlpF. The WT and different mutant strains were characterised using small bottle screening. These strains were optimised with a response surface method to find the optimal conditions to increase butyric and lactic acid.

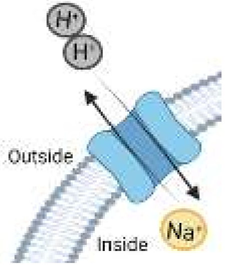
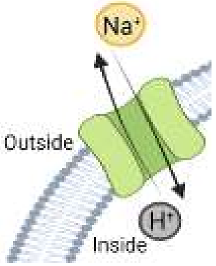
6.1.1 Are any of the strains adequate for the industrial production of organic acids?

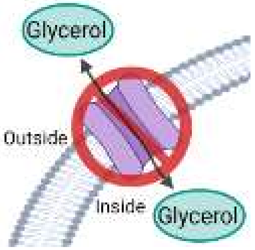
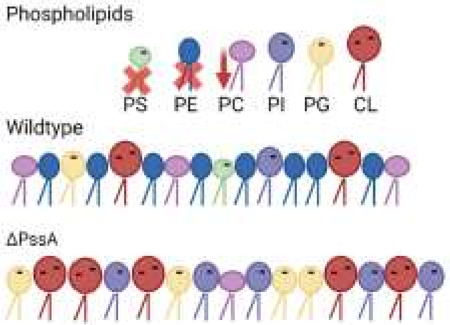
For lactic acid, *Rhizopus oryzae* has a reported titre of 231 g/L of lactic acid with a yield and productivity of 0.92 g/g glucose and 1.89 g/L/h in a fed-batch immobilised culture. However, it requires aeration and maintenance of pH above 4.5, which does not make it commercially viable (Chen and Nielsen, 2016). For butyric acid, *C. tyrobutyricum* ATCC 25755 had in 2018 the highest butyric acid titre of 86.9 g/L (Jiang *et al.*, 2018).

When designing an industrial fermentation, the practical way of performing experiments is through bottle screens because of the large number of experiments required. However, industrial production takes place in fermenters. Therefore a scale-up is required (Seletzky *et al.*, 2007). Only the WT strain was performed in batch fermentation, and only lactic acid was measured. Therefore, the strains of the project cannot be compared with what has been reported for industrial purposes. The cost of product recovery and purification under these conditions also needs to be considered for industrial purposes. An important aspect that needs to be considered in butyric acid production and not with lactic acid is that a third of the substrate in butyric acid production is lost as carbon dioxide. Metabolic engineering has also been performed to re-assimilate the CO_2 for butyrate titre (Jiang *et al.*, 2018). Because of the design of the experiments, the calculation of carbon dioxide was only done theoretically. Experiments are recommended to calculate the actual carbon dioxide produced by the different strains. **Table 6.1** summarises the different mutated clostridial strains with their key behaviours under the different tested conditions.

Table 6.1 Summary of the different mutated clostridial strains and the different tested conditions

Four *C. saccharoperbutylacetonicum* N1-4HMT mutated strains were generated. +NhaA, +Cpa1 and ΔGlpF have a schematic draw of their physiological role. ΔPssA has a schematic role of the different headgroups (headgroup size and net charge are indicated; -: -1 and - -: -2; otherwise, the net charge is zero) that are affected by this knockout and the difference between how the Wildtype and ΔPssA membrane phospholipids should be. The impact of the different strains on the organic acids or solvent production under the different tested conditions are shown as a summarised concluding remark from each strain.

Strain	**Under Biocleave conditions** 32 ˚C, pH < 6.3, 0 % NaCl	**Under salt conditions** 1 % or 2 % (w/v) NaCl	**Under acidic conditions** pH ~ 4.5	**Concluding remarks**
+NhaA  CPA2 electrogenic	• It increased lactic acid production, as reported by Wu *et al.* (2013b).	• At 1 % (w/v) NaCl, it produced the highest amount of lactic acid, 4.74 g/L. • It did not improve salt tolerance in agreement with Wu *et al.* (2013b).	• It did not improve acidic tolerance.	• For this project is the best candidate for lactic acid production.
+Cpa1  CPA1 electroneutral	• The less efficient strain for organic acid or solvent production.	• It increased salt tolerance; it grew at 2 % (w/v) NaCl, as reported by Calinescu *et al.* (2016), Paulino and Kuhlbrandt (2014).	• It did not improve acidic tolerance.	• It is an adequate candidate for salt tolerance.
ΔGlpF	• At 24 h, it produced a high amount of lactic acid (1.68 ± 0.18 g/L). Bienert *et al.* (2013) proposed a role for lactic acid uptake. • It had reduced growth and a delayed organic acid re-assimilation.	• It increased salt tolerance; it grew at 2 % (w/v) NaCl, as opposed to what Liu *et al.* (2015) reported.	• It did not improve acidic tolerance.	• For this project is the worst candidate as it produced a low amount of organic acids with a low amount of solvents in all conditions.

		• It produced the highest amount of butanol and acetone at 1 % (w/v) NaCl.		• It can be confirmed that its role is not necessary for the protection of butanol production since, in its absence, only a delay in the shift to solventogenesis and a reduction in solvent production was observed. •
ΔPssA  Phospholipids PS PE PC PI PG CL Wildtype ΔPssA	• It produced the highest amount of butyric acid, 5.64 ± 0.3 g/L. • It did not achieve solventogenesis. Bogdanov and Dowhan (1995) reported that mutants lacking PE show a defect in the active uptake of many sugars. They addressed the role of the native phospholipid environment in the assembly and function of transporters. Rowlett *et al.* (2017) observed that PE mutants had a decreased ATP content. ATP concentration is a main signal to switch from acidogenesis to solventogenesis. ΔPssA strains might not produce solvents due to a decrease in ATP.	Not tested.	Not tested.	• The ideal candidate for lactic and butyric acid production.

6.2 Alternative strategies

Because acid tolerance is complex and involves multiple genes, developing acid-tolerant strains and improving organic acid production in fermentation could be obtained through adaptive laboratory evolution (ALE), a set of environmentally induced mutations that confer growth advantages to cells. An organism is subjected to serial or continuous cultivation for many generations, which is not optimally adapted to select more fit genetic variants (Liu *et al.*, 2013). The former Green Biologics Ltd., now Bio-Cleave Ltd, had an adaptive laboratory evolution (ALE) technology to improve strain performance and tolerance; their knowledge and industrial perspective could have still been obtained. The problem with genetic engineering is that sometimes the responses obtained differ from those expected. The most significant advantage is that ALE allows phenotypic changes to be associated with a particular growth environment that leads to the selection of the traits (Dragosits and Mattanovich, 2013). The problem with ALE technology is that it is very time-consuming; it can take weeks to years. An example is Ju *et al.* (2016), who, after a year of exogenously supplying lactic acid to wild-type *Leuconostoc mesenteroides,* they obtained a production from 37.8 g/L to 76.8 g/L. One of the mutations they found was in the ATPase ε subunit, which regulates intracellular pH. This approach could adapt Clostridia for lactic or butyric acid or low pH tolerance.

Another strategy could be a Metabolic Flux Analysis, which provides essential information such as maximal yields on a substrate, recognising bottlenecks in the overproduction of the desired product, and identifying the energetic parameters. It is based on network stoichiometry and conservation of mass and does not require information regarding enzyme kinetics (Gheshlaghi *et al.*, 2007). It helps predict the essentiality of gene/reaction and the robustness of a cell under specific genetic or environmental conditions. It can evaluate carbon source utilisation capacity and maximise a particular biochemical production (Lee *et al.*, 2011). Metabolic flux analysis is essential for metabolic engineering, especially for converting the maximum amount of substrate into valuable products (Abdelaal and Yazdani, 2021). With this approach, it could be known what is needed or if it is possible to use *C. saccharoperbutylacetonicum* N1-4HMT to produce lactic and butyric acid, reducing by-products like acetic acid and carbon dioxide production. MFA has limitations in elucidating complex metabolic pathways involving pathways that converge at another point of the network, metabolic cycles and unknown network structure and cofactor utilisation. However, fluxes of intracellular metabolites can be traced with ^{13}C- labelled carbon sources and measured by gas chromatography-mass spectrometry (GC-MS) and nuclear magnetic resonance (NMR). Another limitation is because of the robustness of the cell where changes in the relative level of a particular gene, transcript or protein do not imply a corresponding change in the level of its transcript, protein activity or reaction rate in the cell because of the complicated interactions among the metabolic and regulatory networks (Yang *et al.*, 2007). Wang *et al.* (2012) performed an MFA on *C. acetobutylicum* to see if an early solventogenesis could increase yields. They observed that the oxidoreduction potential (ORP) regulation could be changed. By controlling the fermentation broth at -290 mV using air, an earlier solventogenesis was achieved and got a 35 % increment in solvent production.

A genome-scale metabolic model (GSM) could be generated for a complete approach. GSM is the gene-protein-reaction association for entire metabolic genes in an organism that are computationally described and can predict metabolic fluxes for various systems-level. GSM uses Flux balance analysis for the prediction of metabolic fluxes. The first GSM was reported in 1999 for *Haemophilus influenza* and now is currently used for strain development for chemicals and materials production, prediction of enzyme functions, drug targeting in pathogens, modelling interactions among multiple cells or organisms and understanding human diseases (Gu *et al.*, 2019). The limitation with GSMs is that when using simulated fluxes, cation should be taken. The good thing is that GSMs are open platforms; however, many published GSMs are low quality (Zhang and Hua, 2015). Salimi *et al.* (2010) developed a GSM for *C. acetobutylicum* because MFA could not predict the metabolism correctly; a GSM was necessary to describe the metabolic shift from acidogenesis to solventogenesis. GSMs of *C. acetobutylicum, C. beijerinckii, C. butyricum, C. cellulolyticum, C. ljungahlii* and *C. thermocellum* have been developed (Dash *et al.*, 2016). No GSM of *C. saccharoperbutylacetonicum* N1-4HMT has been reported. In combination with CLEAVE^{TM} technology, the metabolic changes suggested by the GSM or MFA could be performed to optimise organic acid production.

6.3 Future work

The COVID-19 pandemic prevented some experiments that would have expanded this study. Firstly, the miniaturised IV assay would have been used to test more biological membranes, and on *C. saccharoperbutylacetonicum* N1-4HMT, the WT and the different mutant strains +NhaA, +Cpa1, ΔGlpF and ΔPssA. The results would have been correlated with Mass spectrometry (MS) and the Laurdan assay described below.

The Laurdan assay is a method of assessing membrane fluidity. Laurdan is a naphthalene-based amphiphilic molecule (Yu *et al.*, 1996). Laurdan (2-dimethylamino-6-lauroylnpathlanene) is a fluorescent probe synthesised in 1979 by Gregorio Weber. Due to its physical properties, Laurdan binds tightly to the bilayer occupying the region of the glycerol backbone at the level of the ester carbonyl, being able to report water penetration into the membrane surface by displaying an emission wavelength shift. Water penetration correlates with lipid packing and membrane fluidity (Sanchez *et* *al.*, 2012). It can be measured in a fluorescent plate reader, and the analysis is also straightforward (Wenzel *et al.*, 2018). The Laurdan assay does not detect the degree of unsaturation. Therefore, a combination of the miniaturised IV assay and the Laurdan assay could be used on *E. coli, C. saccharoperbutylacetonicum* N1-4HMT and propionic bacteria; those microorganisms would be grown in the presence of different concentrations of organic acids and see if differences in membrane composition could be noticed with those two assays. The results would be complemented and correlated with MS and Thin-Layer chromatography (TLC). This would be used to determine the necessary changes in lipid membrane composition that increase tolerance to organic acids and to establish the importance of unsaturation if, as Tan *et al.* (2017) reported, an increase in unsaturation is essential for survival in an acidic environment. With these results, we could better assess the behaviour of ΔPssA and the absence of solvent production. TLC could have also been performed to observe whether the mutant strains had headgroup differences and to corroborate the unsaturation obtained by IV assay.

C. saccharoperbutylacetonicum N1-4HMT membrane proteins, CsNhaA, CsCpa1, and CsGlpF would have been expressed in *E. coli* and purified by testing the different detergents or SMA co-polymer. The proteins could have been sent for crystallography. The original idea was to reconstitute them into planar bilayers to assay their ability to facilitate organic acids, anions, protons, salt and glycerol. *In silico* modelling could be performed to predict the lactic/butyric acid interaction with membrane lipids and salt interaction with the lipids and organic acids. Particularly *in silico* modelling of ΔPssA could have improved understanding of the effect of phospholipids in the shift to solventogenesis.

For the WT and different mutant strains, it would have been interesting to test them in bioreactors, to do some kinetics and have a more industrial performance. A different approach for RSM could have been taken; the one performed in Mexico could have been repeated and performed with the other strains, keeping in mind that a higher pH and higher temperature might, in the end, increase the cost of the fermentation. In addition, the production of more strains such as ΔNhaA, ΔCpa1, +GlpF, and +PssA mutants in *C. saccharoperbutylacetonicum* N1-4HMT using the CLEAVE^{TM} technology would be of interest in understating the response to organic-acid production, and to develop more tolerant phenotypes.

Finally, a techno-economic analysis would have been performed on the different strains. Academic research needs to consider the economic part, the substrate needed, the investment in equipment, operating expenses, the CO_2 produced, and recovery and purification systems. Sometimes a process optimised at a laboratory scale is not necessarily industrially feasible. This is why a techno-economic analysis of the different strains would have been adequate to determine if the whole procedure was actually reducing Gas emissions and would have reduced costs.

6.4 Conclusions

In conclusion, microorganisms producing biofuels and building-block chemicals have the potential to be the basis of economically competitive bioprocesses, especially if the product has been optimised and the waste and carbon dioxide have been reduced. *C. saccharoperbutylacetonicum* N1-4HMT is one of the most valuable industrial organisms because of its butanol production, which, during the first stage of fermentation, the acidogenic phase, produces lactic and butyric acid. Collectively, the work in this thesis has demonstrated that *C. saccharoperbutylacetonicum* N1-4HMT can be adapted to produce organic acids as end-products. Gene editing using CLEAVE^{TM} technology is easy to perform, and the screening of positive colonies is timesaving. Proton antiporters can be adequate candidates for acid tolerance. Altering the membrane lipid composition through genetic engineering can improve the robustness of industrial strains. This work provides insights into organic acid production and metabolic engineering strategies, which will be valuable in developing acid-tolerant strains. When combined with bioprocess optimisation, improved feedstock utilisation, carbon dioxide reduction, and improved recovery and purification of the products, a sustainable and economically feasible biorefinery processing can be achieved.

7. References

ABDELAAL, A. S. & YAZDANI, S. S. 2021. Chapter 6 - Metabolic engineering for microbial cell factories. *In:* SINGH, V. (ed.) *Microbial Cell Factories Engineering for Production of Biomolecules.* Academic Press 79-94.

AGNEW, H. & ALVIAR-AGNEW, M. 2019. 8.7 Limiting Reactant, Theoretical Yield, and Percent Yield from Initial Masses of Reactants. *In:* LIBRETEXTS (ed.) *Intro Chemistry.* California: MindTouch.

AGUIAR, A., MILESSI, T. S., MULINARI, D. R., LOPES, M. S., DA COSTA, S. M. & CANDIDO, R. G. 2021. Sugarcane straw as a potential second generation feedstock for biorefinery and white biotechnology applications. *Biomass and Bioenergy,* 144. 105896.

AJALA, E. O., OLONADE, Y. O., AJALA, M. A. & AKINPELU, G. S. 2020. Lactic Acid Production from Lignocellulose – A Review of Major Challenges and Selected Solutions. *ChemBioEng Reviews,* 7. 2**,** 38-49.

AL-SHORGANI, N. K. N., KALIL, M. S., YUSOFF, W. M. W. & HAMID, A. A. 2018. Impact of pH and butyric acid on butanol production during batch fermentation using a new local isolate of *Clostridium acetobutylicum* YM1. *Saudi J Biol Sci,* 25. 2**,** 339-348.

ALEGRIA-SCHAFFER, A., LODGE, A. & VATTEM, K. 2009. Performing and optimizing Western blots with an emphasis on chemiluminescent detection. *Methods Enzymol,* 463. 573-599.

ALSAKER, K. V., PAREDES, C. & PAPOUTSAKIS, E. T. 2010. Metabolite stress and tolerance in the production of biofuels and chemicals: gene-expression-based systems analysis of butanol, butyrate, and acetate stresses in the anaerobe *Clostridium acetobutylicum*. *Biotechnol Bioeng,* 105. 6**,** 1131-1147.

ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research,* 25. 3389-3402.

AMANI, T., NOSRATI, M., MOUSAVI, S. M. & KERMANSHAHI, R. K. 2011. Study of syntrophic anaerobic digestion of volatile fatty acids using enriched cultures at mesophilic conditions. *Int. J. Environ. Sci. Tech,* 8. 1**,** 83-96.

ANSCHAU, A., CARUSO, C. S., KUHN, R. C. & FRANCO, T. T. 2017. Validation of the Sulfo-Phospho-Vanillin (Spv) Method for the Determination of Lipid Content in Oleaginous Microorganisms. *Brazilian Journal of Chemical Engineering,* 34. 1**,** 19-27.

ANYASI, T. A., JIDEANI, A. I. O., EDOKPAYI, J. N. & ANKWURU, C. P. 2017. Chapter 1: Application of organic acids in food preservation. *In:* VARGAS, C. (ed.) *Organic acids. Characteristics, Properties and synthesis.* New York: Nova publishers 1-45.

ARBER, W. 2011. 2. Bioengineering at the Interface between Science and Society. *In:* MOO-YOUNG, M. (ed.) *Comprehensive Biotechnology.* Amsterdam: Elsevier 5-10.

ATMADJAJA, A. N., HOLBY, V., HARDING, A. J., KRABBEN, P., SMITH, H. K. & JENKINSON, E. R. 2019. CRISPR-Cas, a highly effective tool for genome editing in *Clostridium saccharoperbutylacetonicum* N1-4(HMT). *FEMS Microbiol Lett,* 366. 6**,** 1-10.

BAEK, S. H., KWON, E. Y., KIM, Y. H. & HAHN, J. S. 2016. Metabolic engineering and adaptive evolution for efficient production of D-lactic acid in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol,* 100. 6**,** 2737-2748.

BAJERSKI, F., WAGNER, D. & MANGELSDORF, K. 2017. Cell Membrane Fatty Acid Composition of *Chryseobacterium frigidisoli* PB4(T), Isolated from Antarctic Glacier Forefield Soils, in Response to Changing Temperature and pH Conditions. *Front Microbiol,* 8. 677.

BAKER-AUSTIN, C. & DOPSON, M. 2007. Life in acid: pH homeostasis in acidophiles. *Trends Microbiol,* 15. 4**,** 165-171.

BAPTISTA, S. L., COSTA, C. E., CUNHA, J. T., SOARES, P. O. & DOMINGUES, L. 2021. Metabolic engineering of *Saccharomyces cerevisiae* for the production of top value chemicals from biorefinery carbohydrates. *Biotechnol Adv,* 47. 107697.

BASTIDAS-OYANEDEL, J.-R., BONK, F., THOMSEN, M. H. & SCHMIDT, J. E. 2019. The Future Perspectives of Dark Fermentation: Moving from Only Biohydrogen to Biochemicals. *Biorefinery*. 375-412.

BEREZINA, O. V., BRANDT, A., YARATOSKY, S., SCHWARZ, W. H. & ZVERLOV, V. V. 2009. Isolation of a new butanol-producing *Clostridium* strain: High level of hemicellulosic activity and structure of solventogenesis genes of a new *Clostridium saccharobutylicum* isolate. *Systematic and Applied Microbiology,* 32. 449-459.

BETTS, M. J. & RUSSEL, R. B. 2007. Amino-acid properties and consequences of subtituions. *In:* BARNES, M. R. & GRAY, I. C. (eds.) *Bioinformatics for Geneticists.* Chichester, UK: John Wiley & Sons Ltd. 311-342.

BIENERT, G. P., DESGUIN, B., CHAUMONT, F. & HOLS, P. 2013. Channel-mediated lactic acid transport: a novel function for aquaglyceroporins in bacteria. *Biochem J,* 454. 3**,** 559-570.

BLANKSBY, S. J. & MITCHELL, T. W. 2010. Advances in mass spectrometry for lipidomics. *Annu Rev Anal Chem (Palo Alto Calif),* 3. 433-465.

BLIGH, E. G. & DYER, W. J. 1959. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology,* 37. 8**,** 911-917.

BOGDANOV, M. & DOWHAN, W. 1995. Phosphatidylethanolamine is required for in vivo function of the membrane-associated lactose permease of Escherichia coli. *J Biol Chem,* 270. 2**,** 732-739.

BONZOM, P. M. A., NICOLAOU, A., ZLOH, M., BALDEO, W. & GIBBONS, W. A. 1999. NMR lipid profile of *Agaricus bisporus*. *Phytochemistry,* 50. 1311-1321.

BOOTH, I. R. 1985. Regulation of Cytoplasmic pH in Bacteria. *Microbiological Reviews,* 49. 4**,** 359-378.

BORGNIA, M. J. & AGRE, P. 2000. Reconstitution and functional comparison of purified GlpF and AqpZ, the glycerol and water channels from *Escherichia coli PNAS,* 98. 5**,** 2888-2893.

BORNHORST, J. A. & FALKE, J. J. 2000. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol,* 326. 245-254.

BOZELL, J. J. & PETERSEN, G. R. 2010. Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's "Top 10" revisited. *Green Chemistry,* 12. 4**,** 539-554.

BRIGHT-GAERTNER, E. & PROULX, P. 1971. Metabolism of phosphoglycerides in *Escherichia coli* during growth at 37 C and during a cold-induced lag phase. *Biochimica et Biophysica acta,* 270. 40-49.

BROECKER, J., EGER, B. T. & ERNST, O. P. 2017. Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs. *Structure,* 25. 2**,** 384-392.

CALINESCU, O., LINDER, M., WOHLERT, D., YILDIZ, O., KUHLBRANDT, W. & FENDLER, K. 2016. Electrogenic Cation Binding in the Electroneutral Na+/H+ Antiporter of *Pyrococcus abyssi*. *J Biol Chem,* 291. 52**,** 26786-26793.

CHAKRABARTI, N., TAJKHORSHID, E., ROUX, B. T. & POMÈS, R. 2004. Molecular Basis of Proton Blockage in Aquaporins. *Structure,* 12. 1**,** 65-74.

CHANROJ, S., WANG, G., VENEMA, K., ZHANG, M. W., DELWICHE, C. F. & SZE, H. 2012. Conserved and diversified gene families of monovalent cation/h(+) antiporters from algae to flowering plants. *Front Plant Sci,* 3. 25.

CHEN, T. Y. 2005. Structure and function of clc channels. *Annu Rev Physiol,* 67. 809-839.

CHEN, X. 2021. Thriving at low pH: adaptation mechanisms of Acidophiles. *Acidophiles-Fundamentals and Application,* Intechopen.

CHEN, Y. & NIELSEN, J. 2016. Biobased organic acids production by metabolically engineered microorganisms. *Curr Opin Biotechnol,* 37. 165-172.

CHENG, C., BAO, T. & YANG, S. T. 2019. Engineering *Clostridium* for improved solvent production: recent progress and perspective. *Appl Microbiol Biotechnol,* 103. 14**,** 5549-5566.

CHO, C., JANG, Y.-S., MOON, H. G., LEE, J. & LEE, S. Y. 2015. Metabolic engineering of clostridia for the production of chemicals. *Biofuels, Bioproducts and Biorefining,* 9. 2**,** 211-225.

COBAN, H. B. 2020. Organic acids as antimicrobial food agents: applications and microbial productions. *Bioprocess Biosyst Eng,* 43. 4**,** 569-591.

COLLAS, F., KUIT, W., CLEMENT, B., MARCHAL, R., LOPEZ-CONTRERAS, A. M. & MONOT, F. 2012. Simultaneous production of isopropanol, butanol, ethanol and 2,3-butanediol by *Clostridium acetobutylicum* ATCC 824 engineered strains. *AMB Expresss,* 2. 45**,** 1-10.

COOKSLEY, C. M., ZHANG, Y., WANG, H., REDL, S., WINZER, K. & MINTON, N. P. 2012. Targeted mutagenesis of the *Clostridium acetobutylicum* acetone-butanol-ethanol fermentation pathway. *Metab Eng,* 14. 6**,** 630-641.

COVERT, T., GREENSTONE, M. & KNITTEL, C. R. 2016. Will We Ever Stop Using Fossil Fuels? *Journal of Economic Perspectives,* 30. 1**,** 117-138.

CZAMARA, K., MAJZNER, K., PACIA, M. Z., KOCHAN, K., KACZOR, A. & BARANSKA, M. 2015. Raman spectroscopy of lipids: a review. *Journal of Raman Spectroscopy,* 46. 1**,** 4-20.

DA SILVA, T. L., GOUVEIA, L. & REIS, A. 2014. Integrated microbial processes for biofuels and high value-added products: the way to improve the cost effectiveness of biofuel production. *Appl Microbiol Biotechnol,* 98. 3**,** 1043-1053.

DASH, S., NG, C. Y. & MARANAS, C. D. 2016. Metabolic modeling of clostridia: current developments and applications. *FEMS Microbiol Lett,* 363. 4.

DE JONG, E. & JUNGMEIER, G. 2015. Biorefinery Concepts in Comparison to Petrochemical Refineries. *Industrial Biorefineries & White Biotechnology.* 3-33.

DILWORTH, M. V., PIEL, M. S., BETTANEY, K. E., MA, P., LUO, J., SHARPLES, D., POYNER, D. R., GROSS, S. R., MONCOQ, K., HENDERSON, P. J. F., MIROUX, B. & BILL, R. M. 2018. Microbial expression systems for membrane proteins. *Methods,* 147. 3-39.

DITU, L.-M. & GHEORGHE, I. 2017. Introduction in Soft Chemistry and Food Fermentation. *Soft Chemistry and Food Fermentation.* London: Academic Press 1-19.

DONG, H., ZHAO, C., ZHANG, T., LIN, Z., LI, Y. & ZHANG, Y. 2016. Engineering *Escherichia coli* cell factories for n-Butanol production. *Adv Biochem Eng Biotechnol,* 155. 141-163.

DRAGOSITS, M. & MATTANOVICH, D. 2013. Adaptive laboratory evolution - principles and applications for biotechnology. *Microb Cell Fact,* 12. 64**,** 1-17.

DRAHOKOUPIL, M. & PATÁKOVÁ, P. 2020. Production of butyric acid at constant pH by a solventogenic strain of *Clostridium beijerinckii*. *Czech Journal of Food Sciences,* 38. No. 3**,** 185-191.

DU, D., VAN VEEN, H. W., MURAKAMI, S., POS, K. M. & LUISI, B. F. 2015. Structure, mechanism and cooperation of bacterial multidrug transporters. *Curr Opin Struct Biol,* 33. 76-91.

DÜRRE, P. 2005. *Handbook on Clostridia,* Boca Raton, CRC Press.

ECOVIA INTELLIGENCE. 2022. *Which trends offer opportunities or pose threats on the European market of natural ingredients for cosmetics* [Online]. https://www.cbi.eu/market-information/natural-ingredients-cosmetics/trends.

ELSLIGER, M. A. & WILSON, I. A. 2012. 1.8 Structure Validation and Analysis. *In:* EGELMAN, E. H. & DYSON, H. J. (eds.) *Comprehensive Biophysics.* Amsterdam: Elsevier 116-135.

EMERY, L., WHELAN, S., HIRSCHI, K. D. & PITTMAN, J. K. 2012. Protein Phylogenetic Analysis of Ca(2+)/cation Antiporters and Insights into their Evolution in Plants. *Front Plant Sci,* 3. 1.

FENG, J., ZONG, W., WANG, P., ZHANG, Z. T., GU, Y., DOUGHERTY, M., BOROVOK, I. & WANG, Y. 2020. RRNPP-type quorum-sensing systems regulate solvent formation, sporulation and cell motility in *Clostridium saccharoperbutylacetonicum*. *Biotechnol Biofuels,* 13. 84**,** 1-16.

FERNANDEZ-NAVEIRA, A., VEIGA, M. C. & KENNES, C. 2019. Effect of salinity on C1-gas fermentation by *Clostridium carboxidivorans* producing acids and alcohols. *AMB Express,* 9. 110**,** 1-11.

FERNANDEZ, L. 2022. Total revenue of the chemical industry worldwide from 2005 to 2021. *In:* STATISTA (ed.). https://www.statista.com/statistics/302081/revenue-of-global-chemical-industry/.

FERNANDO, S., ADHIKARI, S., CHANDRAPAL, C. & MURALI, N. 2006. Biorefineries: Current staturs, challenges and future direction. *Energy & Fuels,* 20. 4**,** 1727-1737.

FINEFIELD, J. M., SHERMAN, D. H., KREITMAN, M. & WILLIAMS, R. M. 2012. Enantiomeric natural products: occurrence and biogenesis. *Angew Chem Int Ed Engl,* 51. 20**,** 4802-4836.

FONTAINE, L., MEYNIAL-SALLES, I., GIRBAL, L., YANG, X., CROUX, C. & SOUCAILLE, P. 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. *J Bacteriol,* 184. 3**,** 821-830.

FRAZZETTO, G. 2003. White biotechnology. *EMBO reports,* 4. 9**,** 835-837.

FU, D., LIBSON, A., MIERCKE, L. J. W., WEITZMAN, C., NOLLERT, P., KRUCINSKI, J. & STROUD, R. M. 2000. Structure of a glycerol-conducting channel and the basis for its selectivity. *Science,* 290. 481-486.

FURRER, E. M., RONCHETTI, M. F., VERREY, F. & POS, K. M. 2007. Functional characterization of a NapA Na(+)/H(+) antiporter from *Thermus thermophilus*. *FEBS Lett,* 581. 3**,** 572-578.

GALKA, J. J., BATURIN, S. J., MANLEY, D. M., KEHLER, A. J. & O'NEIL, J. D. 2008. Stability of the glycerol facilitator in detergent solutions. *Biochemistry,* 47. 3513-3524.

GAPES, J. R. 2000. The economics of Acetone-Butanol Fermentation: Theoretical and Market considerations. *J. Mol. Microbiol. Biotechnol.,* 2. 1**,** 27-32.

GERCHMAN, Y., OLAMI, Y., RIMON, A., TAGLICHT, D., SCHULDINER, S. & PADAN, E. 1993. Histidine-226 is part of the pH sensor of NhaA, a Na^+/H^+ antiporter in *Escherichia coli*. *Biochemistry,* 90. 1212-1216.

GHESHLAGHI, R., SCHARER, J. M., MOO-YOUNG, M. & CHOU, C. P. 2009. Metabolic pathways of clostridia for producing butanol. *Biotechnol Adv,* 27. 6**,** 764-781.

GHESHLAGHI, R., SCHARER, J. M., MOO-YOUNG, M. & DOUGLAS, P. L. 2007. Metabolic flux analysis for optimizing the specific growth rate of recombinant Aspergillus niger. *Bioprocess Biosyst Eng,* 30. 6**,** 397-418.

GINKEL, G. V., LANGEN, H. V. & LEVINE, Y. K. 1989. The Membrane fluidity concept revisited by polarized fluorescence spectroscopy on different model membranes containing unsaturated lipids and sterols. *Biochimie,* 71. 23-32.

GONZALEZ-PEÑAS, H., LU-CHAU, T., EIBES, G. & LEMA, J. M. 2020. Energy requirements and economics of ABE extractive fermentation: a solvent-based comparative assessment. *Authorea*.

GOYAL, D., MISHRA, S. & DANTU, P. K. 2020. Microbial technologies for biorefineries: Current research and future applications. *In:* YADAV, A. N., RASTEGARI, A.A., YADAV, N., GAUR, R. (EDS) (ed.) *Biofuels Production - Sustainability and Advances in Microbial Bioresources. Biofuel and Biorefinery Technologies.* Springer, Cham 229 - 250.

GRIMMLER, C., JANSSEN, H., KRAUSSE, D., FISCHER, R. J., BAHL, H., DURRE, P., LIEBL, W. & EHRENREICH, A. 2011. Genome-wide gene expression analysis of the switch between acidogenesis and solventogenesis in continuous cultures of *Clostridium acetobutylicum*. *J Mol Microbiol Biotechnol,* 20. 1**,** 1-15.

GRUPE, H. & GOTTSCHALK, G. 1992. Physiological Events in *Clostridium acetobutylicum* during the Shift from Acidogenesis to Solventogenesis in Continous Culture and Presentation of a Model for Shift Induction. *Applied and Environmental Micrngiology,* 58. 12**,** 3896-3902.

GU, C., KIM, G. B., KIM, W. J., KIM, H. U. & LEE, S. Y. 2019. Current status and applications of genome-scale metabolic models. *Genome Biol,* 20. 1**,** 121.

GUAN, N. & LIU, L. 2020. Microbial response to acid stress: mechanisms and applications. *Appl Microbiol Biotechnol,* 104. 1**,** 51-65.

GURTLER, J. B. & MAI, T. L. 2014. PRESERVATIVES | Traditional Preservatives – Organic Acids. *In:* C.A, B. & ROBINSON, R. K. (eds.) *Encyclopedia of Food Microbiology.* Second edition ed.: Academic Press 119-130.

HAGVE, T.-A. 1988. Effects of unsaturated fatty acids on cell membrane functions. *Scandinavian Journal of Clinical and Laboratory Investigation,* 48. 5**,** 381-388.

HARRIS, L., BLANK, L., DESAI, R., WELKER, N. & PAPOUTSAKIS, E. 2001. Fermentation characterization and flux analysis of recombinant strains of *Clostridium acetobutylicum* with an inactivated *solR* gene. *Journal of Industrial Microbiology & Biotechnology,* 27. 322-328.

HASHIMOTO, M. & HOSSAIN, S. 2018. Fatty acids: From membrane Ingredients to signaling molecules. *In:* (ED)., V. W. (ed.) *Biochemistry and Health Benefits of Fatty Acids.* London: IntechOpen 1-20.

HEUX, S., MEYNIAL-SALLES, I., O'DONOHUE, M. J. & DUMON, C. 2015. White biotechnology: State of the art strategies for the development of biocatalysts for biorefining. *Biotechnol Adv,* 33. 8**,** 1653-1670.

HILLMANN, F., FISCHER, R. J., SAINT-PRIX, F., GIRBAL, L. & BAHL, H. 2008. PerR acts as a switch for oxygen tolerance in the strict anaerobe Clostridium acetobutylicum. *Mol Microbiol,* 68. 4**,** 848-860.

HINGSAMER, M. & JUNGMEIER, G. 2019. Biorefineries. *In:* LAGO, C., CALDES, N. & LECHON, Y. (eds.) *The Role of Bioenergy in the Bioeconomy.* Amsterdam: Academic Press 179-222.

HO, B. K. & BRASSEUR, R. 2005. The Ramachandran plots of glycine and pre-proline. *BMC structural Biology,* 5. 14**,** 1-11.

HUANG, I. H., WATERS, M., GRAU, R. R. & SARKER, M. R. 2004. Disruption of the gene (spo0A) encoding sporulation transcription factor blocks endospore formation and enterotoxin production in enterotoxigenic *Clostridium perfringens* type A. *FEMS Microbiol Lett,* 233. 2**,** 233-240.

IEA. 2021. *Chemicals* [Online]. Paris. Available: https://www.iea.org/reports/chemicals.

ITO, M., GUFFANTI, A. A., ZEMSKY, J., IVEY, D. M. & KRULWICH, T. A. 1997. Role of the *nhaC*-Encoded Na+/H+ antiporter of Alkaliphilic *Bacillus firmus* OF4. *Journal of Bacteriology,* 179. 12**,** 3851-3857.

JAROS, A. M. 2014. Butyric acid production from renewable resources. *Dissertation to Michigan State University*.

JENKINSON, E. R. & KRABBEN, P. 2015. *Targeted mutations*. PCT/GB2015/051.

JEONG, H., KIM, H. J. & LEE, S. J. 2015. Complete Genome Sequence of Escherichia coli Strain BL21. *Genome Announc,* 3. 2.

JIANG, L., FU, H., YANG, H. K., XU, W., WANG, J. & YANG, S. T. 2018. Butyric acid: Applications and recent advances in its bioproduction. *Biotechnol Adv,* 36. 8**,** 2101-2117.

JIANG, Y., XU, C., DONG, F., YANG, Y., JIANG, W. & YANG, S. 2009. Disruption of the acetoacetate decarboxylase gene in solvent-producing Clostridium acetobutylicum increases the butanol ratio. *Metab Eng,* 11. 4-5**,** 284-291.

JONES, C. M., HERNANDEZ LOZADA, N. J. & PFLEGER, B. F. 2015. Efflux systems in bacteria and their metabolic engineering applications. *Appl Microbiol Biotechnol,* 99. 22**,** 9381-9393.

JOSEPH, R. C., KIM, N. M. & SANDOVAL, N. R. 2018. Recent Developments of the Synthetic Biology Toolkit for *Clostridium*. *Front Microbiol,* 9. 154.

JU, S. Y., KIM, J. H. & LEE, P. C. 2016. Long-term adaptive evolution of *Leuconostoc mesenteroides* for enhancement of lactic acid tolerance and production. *Biotechnol Biofuels,* 9. 240**,** 1-12.

KALAIR, A., ABAS, N., SALEEM, M. S., KALAIR, A. R. & KHAN, N. 2020. Role of energy storage systems in energy transition from fossil fuels to renewables. *Energy Storage,* 3. 135**,** 1-27.

KELL, D. P., MW. RODGER G AND MORRIS JG. 1981. On the permeability to weak acids and bases of the cytoplasmic membrane of *Clostridium Pasteurianum*. *Biochemical and Biophysical Research Communications,* 99. 1**,** 81-88.

KLEIN, N., TREFZ, M. & SCHNEIDER, D. 2019. Covalently Linking Oligomerization-Impaired GlpF Protomers Does Not Completely Re-establish Wild-Type Channel Activity. *Int J Mol Sci,* 20. 4.

KNOTHE, G. 2002. Structure indices in FA chemistry. How relevant is the iodine value? *Journal of American Oil Chemists' Society,* 79. 847-854.

KOMESU, A., OLIVEIRA, J. A. R. D., MARTINS, L. H. D. S., MACIEL, M. R. W. & FILHO, R. M. 2017. Lactic acid production to purification: a review. *bioResources,* 12. 2**,** 4364-4383.

KORDI, M., SALAMI, R., BOLOURI, P., DELANGIZ, N., ASGARI LAJAYER, B. & VAN HULLEBUSCH, E. D. 2022. White biotechnology and the production of bio-products. *Systems Microbiology and Biomanufacturing,* 2. 3**,** 413-429.

KOTOV, V., BARTELS, K., VEITH, K., JOSTS, I., SUBHRAMANYAM, U. K. T., GUNTHER, C., LABAHN, J., MARLOVITS, T. C., MORAES, I., TIDOW, H., LOW, C. & GARCIA-ALAI, M. M. 2019. High-throughput stability screening for detergent-solubilized membrane proteins. *Sci Rep,* 9. 1**,** 10379.

KUBIAK, P., LEJA, K., MYSZKA, K., CELIŃSKA, E., SPYCHAŁA, M., SZYMANOWSKA-POWAŁOWSKA, D., CZACZYK, K. & GRAJEK, W. 2012. Physiological predisposition of various *Clostridium* species to synthetize 1,3-propanediol from glycerol. *Process Biochemistry,* 47. 9**,** 1308-1319.

KUDO, H. & SASAKI, Y. 2019. Intracellular pH Determination for the Study of Acid Tolerance of Lactic Acid Bacteria. *Methods Mol Biol,* 1887. 33-41.

KUFAREVA, I. & ABAGYAN, R. 2012. Methods of protein structure comparison. *Methods Mol Biol,* 857. 231-257.

KWON, S. W., PAARI, K. A., MALAVIYA, A. & JANG, Y. S. 2020. Synthetic Biology Tools for Genome and Transcriptome Engineering of Solventogenic *Clostridium*. *Front Bioeng Biotechnol,* 8. 282**,** 1-9.

KYRIAKIDIS, N. B. & KATSILOULIS, T. 2000. Calculation of iodine value from measurements of fatty acid methyl esters of some oils: Comparison with the relevant American Oil Chemist Society method. *Journal of the American Oil Chemists' Society,* 77. 1235-1238.

LEBARBE, T. 2013. Synthesis of novel "green" polyesters from plant oils: application to the rubber-toughening of poly(L-lactide). .

LEE, J. 2021. Lessons from Clostridial Genetics: Toward Engineering Acetogenic Bacteria. *Biotechnology and Bioprocess Engineering,* 26. 6**,** 841-858.

LEE, J., SANDS, Z. A. & BIGGIN, P. C. 2016. A Numbering System for MFS Transporter Proteins. *Front Mol Biosci,* 3. 21.

LEE, S. Y., PARK, J. M. & KIM, T. Y. 2011. Application of metabolic flux analysis in metabolic engineering. *Methods Enzymol,* 498. 67-93.

LEVI, P. G. & CULLEN, J. M. 2018. Mapping Global Flows of Chemicals: From Fossil Fuel Feedstocks to Chemical Products. *Environ Sci Technol,* 52. 4**,** 1725-1734.

LEWIS, R. N. & MCELHANEY, R. N. 2013. Membrane lipid phase transitions and phase organization studied by Fourier transform infrared spectroscopy. *Biochim Biophys Acta,* 1828. 10**,** 2347-2358.

LI, Q., CHEN, J., MINTON, N. P., ZHANG, Y., WEN, Z., LIU, J., YANG, H., ZENG, Z., REN, X., YANG, J., GU, Y., JIANG, W., JIANG, Y. & YANG, S. 2016a. CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. *Biotechnol J,* 11. 7**,** 961-972.

LI, Q. Z., JIANG, X. L., FENG, X. J., WANG, J. M., SUN, C., ZHANG, H. B., XIAN, M. & LIU, H. Z. 2016b. Recovery Processes of Organic Acids from Fermentation Broths in the Biomass-Based Industry. *J Microbiol Biotechnol,* 26. 1**,** 1-8.

LI, S., HUANG, L., KE, C., PANG, Z. & LIU, L. 2020. Pathway dissection, regulation, engineering and application: lessons learned from biobutanol production by solventogenic clostridia. *Biotechnol Biofuels,* 13. 39.

LI, X., SCHULER, M. A. & BERENBAUM, M. R. 2002. Jasmonate and salicylate induce expression of herbivore cytochrome P450 genes. *Nature,* 419. 6908**,** 712-715.

LI, Y. & CUI, F. 2010. Microbial Lactic Acid Production from Renewable Resources. 211-228.

LIU, X.-B., GU, Q.-Y. & YU, X.-B. 2013. Repetitive domestication to enhance butanol tolerance and production in *Clostridium acetobutylicum* through artificial simulation of bio-evolution *Bioresour Technol,* 130. 638-643.

LIU, X. D., WEI, Y., ZHOU, X. Y., PEI, X. & ZHANG, S. H. 2015. *Aspergillus glaucus* Aquaglyceroporin Gene glpF Confers High Osmosis Tolerance in Heterologous Organisms. *Appl Environ Microbiol,* 81. 19**,** 6926-6937.

LIZHI, H., TOYODA, K. & IHARA, I. 2008. Dielectric properties of edible oils and fatty acids as a function of frequency, temperature, moisture and composition. *Journal of Food Engineering,* 88. 2**,** 151-158.

MA, Y. C., AUGE, R. M., DONG, C. & CHENG, Z. M. 2017. Increased salt tolerance with overexpression of cation/proton antiporter 1 genes: a meta-analysis. *Plant Biotechnol J,* 15. 2**,** 162-173.

MADDOX, I. S., QURESHI, N. & ROBERTS-THOMSON, K. 1995. Production of Acetone-Butanol-Ethanol from concentrated substrates using *Clostridium acetobutylicum* in an integrated fermentation-product removal process. *Process biochemistry,* 30. 3**,** 209-215.

MAO, S., LUO, Y., ZHANG, T., LI, J., BAO, G., ZHU, Y., CHEN, Z., ZHANG, Y., LI, Y. & MA, Y. 2010. Proteome reference map and comparative proteomic analysis between a wild type *Clostridium acetobutylicum* DSM 1731 and its mutant with enhanced butanol tolerance and butanol yield. *J Proteome Res,* 9. 6**,** 3046-3061.

MASRATI, G., DWIVEDI, M., RIMON, A., GLUCK-MARGOLIN, Y., KESSEL, A., ASHKENAZY, H., MAYROSE, I., PADAN, E. & BEN-TAL, N. 2018. Broad phylogenetic analysis of cation/proton antiporters reveals transport determinants. *Nat Commun,* 9. 4205**,** 1-14.

MATSUOKA, R., FUDIM, R., JUNG, S., ZHANG, C., BAZZONE, A., CHATZIKYRIAKIDOU, Y., ROBINSON, C. V., NOMURA, N., IWATA, S., LANDREH, M., ORELLANA, L., BECKSTEIN, O. & DREW, D. 2022. Structure, mechanism and lipid-mediated remodeling of the mammalian Na(+)/H(+) exchanger NHA2. *Nat Struct Mol Biol,* 29. 2**,** 108-120.

MATTIELLO-FRANISCO, L., VIEIRA, F. M., PEIXOTO, G. & MOCKAITIS, G. 2021. Lactic acid production by *clostridium acetobutylicum* and *clostridium beijerinckii* under anaerobic conditions using a complex substrate. *BioRxiv*. 1-18.

MAVIS, R. D. & VAGELOS, P. R. 1972. The Effect of Phospholipid Fatty Acid Composition on Membranous Enzymes in Escherichia coli. *Journal of Biological Chemistry,* 247. 3**,** 652-659.

MAYANK, R., RANJAN, A. & MOHOLKAR, V. S. 2013. Mathematical models of ABE fermentation: review and analysis. *Crit Rev Biotechnol,* 33. 4**,** 419-447.

MCWILLIAM, H., LI, W., ULUDAG, M., SQUIZZATO, S., PARK, Y., BUSO, N., COWLEY, A. & LOPEZ, R. 2013. *Analysis tool web services from the EMBL-EBI* [Online].

MERWE, A. B. V. D. 2010. *Evaluation of different process designs for biobutanol production from sugarcane molasses.* Master of Science in Engineering (Chemical engineering), University of Stellenbosch.

MICHALIK, B., BIEL, W., LUBOWICKI, R. & JACYNO, E. 2014. Chemical composition and biological value of proteins of the yeast Yarrowia lipolytica growing on industrial glycerol. *Canadian Journal of Animal Science,* 94. 1**,** 99-104.

MILLER, C., FOSMER, A., RUSH, B., MCMULLIN, T., BEACOM, D. & SUOMINEN, P. 2011. 3.17 Industrial Production of Lactic Acid. *In:* MOO-YOUNG, M. & BUTLER, M. (eds.) *Comprehensive Biotechnology.* Elsevier Science 179-188.

MOLS, M. & ABEE, T. 2011. *Bacillus cereus* responses to acid stress. *Environ Microbiol,* 13. 11**,** 2835-2843.

MONAGHAN, T. I. 2019. *Optimising solvent production in Clostridium saccharoperbutylacetonicum N1-4(HMT).* PhD in Microbiology, University of Kent.

MONDAL, R., RIMON, A., MASRATI, G., BEN-TAL, N., FRIEDLER, A. & PADAN, E. 2021. Towards Molecular Understanding of the pH Dependence Characterizing NhaA of Which Structural Fold is Shared by Other Transporters. *J Mol Biol,* 433. 19**,** 1-12.

MOON, H. G., JANG, Y. S., CHO, C., LEE, J., BINKLEY, R. & LEE, S. Y. 2016. One hundred years of clostridial butanol fermentation. *FEMS Microbiol Lett,* 363. 3.

MOREIN, S., ANDERSSON, A., RILFORS, L. & LINDBLOM, G. 1996. Wild-type *Escherichia coli* cells regulate the membrane lipid composition in a "window" between gel and non-lamellar structures. *J Biol Chem,* 271. 12**,** 6801-6809.

NAGARAJU, S., DAVIES, N. K., WALKER, D. J., KOPKE, M. & SIMPSON, S. D. 2016. Genome editing of *Clostridium autoethanogenum* using CRISPR/Cas9. *Biotechnol Biofuels,* 9. 219**,** 1-8.

NAIR, R. V. & PAPOUTSAKIS, E. T. 1994. Expression of plasmid-encoded *aad* in *clostridium acetobutylicum* M5 restores vigorous butanol production. *Journal of Bacteriology,* 176. 18**,** 5843-5846.

NANDA, S., GOLEMI-KOTRA, D., MCDERMOTT, J. C., DALAI, A. K., GOKALP, I. & KOZINSKI, J. A. 2017. Fermentative production of butanol: Perspectives on synthetic biology. *N Biotechnol,* 37. Pt B**,** 210-221.

NEW HOPE NETWORK STAFF 2022. The top natural products industry trends and 2022 predictions. *In:* NEWS, S. S. (ed.) *organic & natural.* https://www.supermarketnews.com/organic-natural/top-natural-products-industry-trends-and-2022-predictions.

NOVAGEN 1999. pET System Manual. 8th Edition ed. US & Canada: Novagen.

ODOOM, W. & EDUSEI, V. O. 2015. Evaluation of Saponification value, Iodine value and Insoluble impurities in Coconut oils from Jomoro District in the Western Region of Ghana. *Asian Journal of Agriculture and Food Sciences,* 3. 5**,** 494-499.

OSHIRO, M., HANADA, K., TASHIRO, Y. & SONOMOTO, K. 2010. Efficient conversion of lactic acid to butanol with pH-stat continuous lactic acid and glucose feeding method by *Clostridium saccharoperbutylacetonicum*. *Appl Microbiol Biotechnol,* 87. 3**,** 1177-1185.

PADAN, E. 2014. Functional and structural dynamics of NhaA, a prototype for Na(+) and H(+) antiporters, which are responsible for Na(+) and H(+) homeostasis in cells. *Biochim Biophys Acta,* 1837. 7**,** 1047-1062.

PADAN, E., TZUBERY, T., HERZ, K., KOZACHKOV, L., RIMON, A. & GALILI, L. 2004. NhaA of *Escherichia coli,* as a model of a pH-regulated Na+/H+antiporter. *Biochim Biophys Acta,* 1658. 1-2**,** 2-13.

PALMA, M., GUERREIRO, J. F. & SA-CORREIA, I. 2018. Adaptive Response and Tolerance to Acetic Acid in *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii*: A Physiological Genomics Perspective. *Front Microbiol,* 9. 274**,** 1-16.

PANDA, S. K., SAHU, L., BEHERA, S. K. & RAY, R. C. 2019. Research and Production of Organic Acids and Industrial Potential. *In:* MOLINA, G. (ed.) *Bioprocessing for Biomolecules Production.* Hoboken, NJ: Wiley 195-209.

PAPAGIANNI, M. 2011. 1.09 Organic Acids. *In:* MOO-YOUNG, M. (ed.) *Comprehensive Biotechnology.* Amsterdam: Elsevier 109-120.

PAPOUTSAKIS, E. T. 1984. Equations and calculations for fermentations of butyric acid bacteria. *Biotechnology and Bioengineering,* 26. 174-187.

PARDESHI, D. S. 2020a. Evaluation of determination of iodine value of extracted groundnut seed oils by using mercuric acetate as accelerator. *International Journal of Scientific Development and Research,* 5. 9**,** 59-63.

PARDESHI, S. 2020b. Evaluation of determination of iodine value of extracted groundnut seed oils by using mercuric acetate as accelerator. *International Journal of Scientific Development and Research,* 5. 9**,** 59-63.

PARENT, K. E. 2003. Cleaning up with atom economy. *In:* INSTITUTE, G. C. (ed.). American Chemical Society.

PATAKOVA, P., BRANSKA, B., VASYLKIVSKA, M., JURECKOVA, K., MUSILOVA, J., PROVAZNIK, I. & SEDLAR, K. 2022. Transcriptomic studies of solventogenic clostridia, Clostridium acetobutylicum and Clostridium beijerinckii. *Biotechnol Adv,* 58. 107889.

PATEL, A., ANTONOPOULOU, I., ENMAN, J., ROVA, U., CHRISTAKOPOULOS, P. & MATSAKAS, L. 2019. Lipids detection and quantification in oleaginous microorganisms: an overview of the current state of the art. *BMC Chemical Engineering,* 1. 1.

PAULINO, C. & KUHLBRANDT, W. 2014. pH- and sodium-induced changes in a sodium/proton antiporter. *Elife,* 3. e01412.

PEI, H., WANG, C., WANG, Y., YANG, H. & XIE, S. 2019. Distribution of microbial lipids at an acid mine drainage site in China: Insights into microbial adaptation to extremely low pH conditions. *Organic Geochemistry,* 134. 77-91.

PESSOA, M. G., VESPERMANN, K. A. C., PAULINO, B. N., BARCELOS, M. C. S., PASTORE, G. M. & MOLINA, G. 2019. Newly isolated microorganisms with potential application in biotechnology. *Biotechnology Advances,* 37. 2**,** 319-339.

PETROV, K., ARSOV, A. & PETROVA, P. 2021. Butanol Tolerance of *Lactiplantibacillus plantarum*: A Transcriptome Study. *Genes (Basel),* 12. 2.

PFROMM, P. H., AMANOR-BOADU, V., NELSON, R., VADLANI, P. & MADL, R. 2010. Bio-butanol vs bio-ethanol: a technical and economic assessment for corn and switchgrass fermented by yeast or Clostridium acetobutylicum. *Biomass & Bioenergy,* 34. 515-524.

PIETERSE, B., LEER, R. J., SCHUREN, F. H. J. & WERF, M. J. V. D. 2005. Unraveling the multiple effects of lactic acid stress on *Lactobacillu plantarum* by transcription profiling. *Microbiology,* 151. 3881-3894.

PIPER, P. W. 2011. Resistance of yeasts to weak organic acid food preservatives. *Adv Appl Microbiol,* 77. 97-113.

POEHLEIN, A., SOLANO, J. D. M., FLITSCH, S. K., KRABBEN, P., WINZER, K., REID, S. J., JONES, D. T., GREEN, E., MINTON, N. P., DANIEL, R. & DURRE, P. 2017. Microbial solvent formation revisited by comparative genome analysis. *Biotechnol Biofuels,* 10. 58.

POMERANZ, Y. & MELOAN, C. E. 2002. *Food Analysis Theory and Practice* New York, Chapman & Hall.

PRAMANIK, J. & KEASLING, J. D. 1997. Stoichiometric Model of *Escherichia coli* metabolism: Incorporation of Growth-Rate Dependent Biomass Composition and Mechanistic Energy Requirements. *Biotechnology and Bioengineering,* 56. 4**,** 398 - 421.

QI, Y., LIU, H., CHEN, X. & LIU, L. 2019. Engineering microbial membranes to increase stress tolerance of industrial strains. *Metab Eng,* 53. 24-34.

RACHWALSKI, M., VERMUE, N. & RUTJES, F. P. 2013. Recent advances in enzymatic and chemical deracemisation of racemic compounds. *Chem Soc Rev,* 42. 24**,** 9268-9282.

RAO, S., SAYEEDA, M., PRAKASH, T., AP, P., IMRAN, S. & RAVI, L. 2020. Construction of Computational 3d Structures of Protein Drug Targets of *Mycobacterium Tuberculosis*. *Asian Journal of Pharmaceutical and Clinical Research*. 82-85.

REPORTLINKER. 2022. *Global Biorefinery Market 2022 - 2026* [Online]. Available: https://www.reportlinker.com/p05483101/?utm_source=GNW.

RETTNER, R. E. & SAIER, M. H., JR. 2010. The autoinducer-2 exporter superfamily. *J Mol Microbiol Biotechnol,* 18. 4**,** 195-205.

RICKE, K., DROUET, L., CALDEIRA, K. & TAVONI, M. 2018. Country-level social cost of carbon. *Nature Climate Change,* 8. 10**,** 895-900.

ROTHNIE, A. J. 2016. Detergent-Free Membrane Protein Purification. *Methods Mol Biol*. 261-268.

ROWLETT, V. W., MALLAMPALLI, V. K. P. S., KARLSTAEDT, A., DOWNHAN, W., TAEGTMEYER, H., MARGOLIN, W. & VITRAC, H. 2017. Impact of membrane phospholipid alterations in *Escherichia coli* on cellular function and bacterial stress adaptation. *J Bacteriol,* 199. 13.

RUMIN, J., BONNEFOND, H., SAINT-JEAN, B., ROUXEL, C., SCIANDRA, A., BERNARD, O., CADORET, J. P. & BOUGARAN, G. 2015. The use of fluorescent Nile red and BODIPY for lipid measurement in microalgae. *Biotechnol Biofuels,* 8. 42.

RYCHLIK, W., SPENCER, W. J. & RHOADS, R. E. 1990. Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Research,* 18. 21**,** 6409-6412.

SAHASRABUDHE, N. A. & SANKPAL, N. V. 2001. Production of organic acids and metabolites of fungi for food industry. *Applied mycology and biotechnology,* 1. 387-425.

SALIMI, F., ZHUANG, K. & MAHADEVAN, R. 2010. Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing. *Biotechnol J,* 5. 7**,** 726-738.

SANCHEZ, S. A., TRICERRI, M. A. & GRATTON, E. 2012. Laurdan generalized polarization fluctuations measures membrane packing micro-heterogeneity *in vivo*. *Proc Natl Acad Sci U S A,* 109. 19**,** 7314-7319.

SANKARAN, J. & WOHLAND, T. 2020. Fluorescence strategies for mapping cell membrane dynamics and structures. *APL Bioeng,* 4. 020901**,** 1-15.

SAUER, M., PORRO, D., MATTANOVICH, D. & BRANDUARDI, P. 2008. Microbial production of organic acids: expanding the markets. *Trends Biotechnol,* 26. 2**,** 100-108.

SCHRODINGER, LLC 2015. The PyMOL Molecular Graphics System, Version 1.8.

SCIENTIFIC LABORATORY SUPPLIES. 2022. *Accuris Smart Reader^{TM} 96* [Online]. Available: https://www.scientificlabs.co.uk/product/MIC1121#overview.

SCIENTIFIC, T. 2022. *Agilent BioTek 800 Ts Absrobance Reader* [Online]. Available: https://www.fishersci.co.uk/shop/products/800-ts-absorbance-reader-6/p-7112442#?keyword=plate%20reader.

SELETZKY, J. M., NOAK, U., FRICKE, J., WELK, E., EBERHARD, W., KNOCKE, C. & BUCHS, J. 2007. Scale-up from shake flasks to fermenters in batch and continuous mode with Corynebacterium glutamicum on lactic acid based on oxygen transfer and pH. *Biotechnol Bioeng,* 98. 4**,** 800-811.

SHANTHA, N. C. & NAPOLITANO, G. E. 1998. Lipid analysis using thin-layer chromatography and Iatroscan. *Lipid analysis in Oils and Fats*. 1-33.

SINGH, G., JEYASEELAN, C., BANDYOPADHYAY, K. K. & PAUL, D. 2018. Comparative analysis of biodiesel produced by acidic transesterification of lipid extracted from oleaginous yeast Rhodosporidium toruloides. *3 Biotech,* 8. 10**,** 434.

SINGH, S., SRIVASTAVA, P., DEEP, S., ASHISH, SHUKLA, P., RIZVI, H. & SINHA, R. 2022. Fermentation strategies for organic acid production. *In:* VERMA, P. (ed.) *Industrial Microbiology and Biotechnology.* Singapore: Springer 379-425.

SOHLENKAMP, C. & GEIGER, O. 2016. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol Rev,* 40. 1**,** 133-159.

STAHLBERG, H., BRAUN, T., DE GROOT, B., PHILIPPSEN, A., BORGNIA, M. J., AGRE, P., KUHLBRANDT, W. & ENGEL, A. 2000. The 6.9-A structure of GlpF: a basis for homology modeling of the glycerol channel from *Escherichia coli*. *J Struct Biol,* 132. 2**,** 133-141.

STEWART, J. C. M. 1980. Colorimetric dtermination of phospholipids with ammonium ferrothiocyanate. *Analytical Biochemistry,* 104. 10-14.

STILLWELL, W. 2016. Chapter 4: Membrane Lipids. *In:* STILLWELL, W. (ed.) *An Introduction to Biological Membranes.* Second Edition ed.: Elsevier 49-62.

STROUD, R. M., NOLLERT, P. & MIERCKE, L. 2003. The glycerol facilitaotr GlpF, aquaporin family *Emerald Biostructures,* 63. 291-316.

TAN, Z., KHAKBAZ, P., CHEN, Y., LOMBARDO, J., YOON, J. M., SHANKS, J. V., KLAUDA, J. B. & JARBOE, L. R. 2017. Engineering Escherichia coli membrane phospholipid head distribution improves tolerance and production of biorenewables. *Metab Eng,* 44. 1-12.

TAN, Z., YOON, J. M., NIELSEN, D. R., SHANKS, J. V. & JARBOE, L. R. 2016. Membrane engineering via trans unsaturated fatty acids production improves Escherichia coli robustness and production of biorenewables. *Metab Eng,* 35. 105-113.

THONGCHUL, N., TOLIENG, V. & PRASITCHOKE, P. 2019. Process for producing lactic acid or its salts from fermentation using thermotolerance bacillus bacteria. *United states Patent Application Publication*. US 2019/0100779 A1.

TIEFENBACHER, K. F. 2017. Technology of Main Ingredients—Sweeteners and Lipids. *Wafer and Waffle.* 123-225.

TOMAS, C. A., WELKER, N. E. & PAPOUTSAKIS, E. T. 2003. Overexpression of groESL in Clostridium acetobutylicum results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. *Appl Environ Microbiol,* 69. 8**,** 4951-4965.

TRAN, V. G. & ZHAO, H. 2022. Engineering robust microorganisms for organic acid production. *J Ind Microbiol Biotechnol,* 49. 1-8.

TREFZ, M., KELLER, R., VOGT, M. & SCHNEIDER, D. 2018. The GlpF residue Trp219 is part of an amino-acid cluster crucial for aquaglyceroporin oligomerization and function. *Biochim Biophys Acta Biomembr,* 1860. 4**,** 887-894.

TUMMALA, S. B., JUNNE, S. G. & PAPOUTSAKIS, E. T. 2003. Antisense RNA downregulation of coenzyme A transferase combined with alcohol-aldehyde dehydrogenase overexpression leads to predominantly alcohologenic *Clostridium acetobutylicum* fermentations. *J Bacteriol,* 185. 12**,** 3644-3653.

UNGER, V. M. 2000. Fraternal twins: AQP1 and GlpF. *Nature structural biology,* 7. 12**,** 1082-1084.

VAN MARIS, A. J., KONINGS, W. N., VAN DIJKEN, J. P. & PRONK, J. T. 2004. Microbial export of lactic and 3-hydroxypropanoic acid: implications for industrial fermentation processes. *Metab Eng,* 6. 4**,** 245-255.

VANDENBERGHE, L. P. S., KARP, S. G., DE OLIVEIRA, P. Z., DE CARVALHO, J. C., RODRIGUES, C. & SOCCOL, C. R. 2018. Solid-State Fermentation for the Production of Organic Acids. *Current Developments in Biotechnology and Bioengineering.* 415-434.

VLAB.AMRITA.EDU. 2011. *Estimation of Iodine Value of Fats and oils* [Online]. Available: https://vlab.amrita.edu/?sub=3&brch=63&sim=1111&cnt=1.

WAGNER, J. R., MOUNT, E. M. & GILES, H. F. 2014. Design of Experiments. *Extrusion.* 291-308.

WALTHER, J., GODAWAT, R., HWANG, C., ABE, Y., SINCLAIR, A. & KONSTANTINOV, K. 2015. The business impact of an integrated continuous biomanufacturing platform for recombinant protein production. *J Biotechnol,* 213. 3-12.

WANG, Q., VENKATARAMANAN, K. P., HUANG, H., PAPOUTSAKIS, E. T. & WU, C. H. 2013. Transcription factors and genetic circuits orchestrating the complex, multilayered response of *Clostridium acetobutylicum* to butanol and butyrate stress. *BMC systems biology,* 7. 120**,** 1-17.

WANG, S., DONG, S., WANG, P., TAO, Y. & WANG, Y. 2017. Genome editing in *C. saccharoperbutylacetonicum* N1-4 with the CRISPR-Cas9 system. *Appl Environ Microbiol,* 83. 10**,** 1-16.

WANG, S., ZHU, Y., ZHANG, Y. & LI, Y. 2012. Controlling the oxidoreduction potential of the culture of *Clostridium acetobutylicum* leads to an earlier initiation of solventogenesis, thus increasing solvent productivity. *Appl Microbiol Biotechnol,* 93. 3**,** 1021-1030.

WARNECKE, T. & GILL, R. T. 2005. Organic acid toxicity, tolerance, and production in *Escherichia coli* biorefining applications. *Microb Cell Fact,* 4. 25**,** 1-8.

WEBB, C. 2011. 2.01 Introduction. *In:* MOO-YOUNG, M. (ed.) *Comprehensive Biotechnology.* Amsterdam: Elsevier 1-3.

WENDISCH, V. F., BOTT, M. & EIKMANNS, B. J. 2006. Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for biotechnological production of organic acids and amino acids. *Current Opinion in Microbiology,* 9. 3**,** 268-274.

WENZEL, M., VISCHER, N. O. E., STRAHL, H. & HAMOEN, L. W. 2018. Assessing Membrane Fluidity and Visualizing Fluid Membrane Domains in Bacteria Using Fluorescent Membrane Dyes. *Bio Protoc,* 8. 20**,** 1-20.

WILLSON, B. J., KOVACS, K., WILDING-STEELE, T., MARKUS, R., WINZER, K. & MINTON, N. P. 2016. Production of a functional cell wall-anchored minicellulosome by recombinant *Clostridium acetobutylicum* ATCC 824. *Biotechnol Biofuels,* 9. 109**,** 1-22.

WOHLERT, D., KUHLBRANDT, W. & YILDIZ, O. 2014. Structure and substrate ion binding in the sodium/proton antiporter PaNhaP. *Elife,* 3. e03579.

WOLRAB, D., PETERKA, O., CHOCHOLOUŠKOVÁ, M. & HOLČAPEK, M. 2022. Ultrahigh-performance supercritical fluid chromatography / mass spectrometry in the lipidomic analysis. *TrAC Trends in Analytical Chemistry,* 149.

WOOD, J. M. 2015. Bacterial responses to osmotic challenges. *J Gen Physiol,* 145. 5**,** 381-388.

WRESTA, A., WIDYARANI, R., BOOPATHY, R. & SETIADI, T. 2021. Thermodynamic approach to estimating reactions and stoichiometric coefficients of anaerobic glucose and hydrogen utilization. *Engineering Reports,* 3. 6.

WU, C., ZHANG, J., WANG, M., DU, G. & CHEN, J. 2012. Lactobacillus casei combats acid stress by maintaining cell membrane functionality. *J Ind Microbiol Biotechnol,* 39. 7**,** 1031-1039.

WU, L., FAN, Z., GUO, L., LI, Y., CHEN, Z.-L. & QU, L.-J. 2005. Over-expression of the bacterial nhaA gene in rice enhances salt and drought tolerance. *Plant Science,* 168. 2**,** 297-302.

WU, M., TONG, S., WALTERSPERGER, S., DIEDERICHS, K., WANG, M. & ZHENG, L. 2013a. Crystal structure of Ca2+/H+ antiporter protein YfkE reveals the mechanisms of Ca2+ efflux and its pH regulation. *Proc Natl Acad Sci U S A,* 110. 28**,** 11367-11372.

WU, X., ALTMAN, R., EITEMAN, M. A. & ALTMAN, E. 2013b. Effect of overexpressin *nhaA* and *nhaR* on sodium tolerance and lactate production in *Escherichia coli*. *Journal of Biological Engineering,* 7. 3**,** 1-7.

XU, K., LEE, Y. S., LI, J. & LI, C. 2019. Resistance mechanisms and reprogramming of microorganisms for efficient biorefinery under multiple environmental stresses. *Synth Syst Biotechnol,* 4. 2**,** 92-98.

XU, M., ZHAO, J., YU, L., TANG, I. C., XUE, C. & YANG, S. T. 2015. Engineering *Clostridium acetobutylicum* with a histidine kinase knockout for enhanced n-butanol tolerance and production. *Appl Microbiol Biotechnol,* 99. 2**,** 1011-1022.

XU, Y., ZHAO, Z., TONG, W., DING, Y., LIU, B., SHI, Y., WANG, J., SUN, S., LIU, M., WANG, Y., QI, Q., XIAN, M. & ZHAO, G. 2020. An acid-tolerance response system protecting exponentially growing *Escherichia coli*. *Nat Commun,* 11. 1**,** 1-13.

XU, Z. & JIANG, L. 2011. Butyric Acid. *In:* MOO-YOUNG, M. (ed.) *Comprehensive Biotechnology.* Amsterdam: Elsevier 235-243.

YANG, M., AN, Y., ZABED, H. M., GUO, Q., YUN, J., ZHANG, G., AWAD, F. N., SUN, W. & QI, X. 2019. Random mutagenesis of *Clostridium butyricum* strain and optimization of biosynthesis process for enhanced production of 1,3-propanediol. *Bioresour Technol,* 284. 188-196.

YANG, S.-T. 2007. Chapter 1. Bioprocessing - from Biotechnology to Biorefinery. *In:* YANG, S.-T. (ed.) *Bioprocessing for value-added products from renewable resources.* USA: Elsevier B.V. 1-24.

YANG, S.-T., LIU, X. & ZHANG, Y. 2007. Chapter 4 - Metabolic Engineering - Applications, Methods, and Challenges. *In:* YANG, S.-T. (ed.) *Bioprocessing for value-added products from renewable resources.* Elsevier 73-118.

YANG, Z., WANG, C., ZHOU, Q., AN, J., HILDEBRANDT, E., ALEKSANDROV, L. A., KAPPES, J. C., DELUCAS, L. J., RIORDAN, J. R., URBATSCH, I. L., HUNT, J. F. & BROUILLETTE, C. G. 2014. Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains. *Protein Sci,* 23. 6**,** 769-789.

YU, W., SO, P. T. C., FRENCH, T. & GRATTON, E. 1996. Fluorescence generalized polarization of cell membranes: a two-photon scanning microscopy approach. *Biophysical Journal,* 70. 626-636.

ZAKATAEVA, N. P., ALESHIN, V. V., TOKMAKOVA, I. L., TROSHIN, P. V. & LIVSHITS, V. A. 1999. The novel transmembrane *Escherichia coli* proteins involved in the amino acid efflux. *FEBS Letters,* 452. 228-232.

ZHANG, C. & HUA, Q. 2015. Applications of Genome-Scale Metabolic Models in Biotechnology and Systems Medicine. *Front Physiol,* 6. 413.

ZHAO, J., QUAN, C., JIN, L. & CHEN, M. 2018. Production, detection and application perspectives of quorum sensing autoinducer-2 in bacteria. *J Biotechnol,* 268. 53-60.

ZHAO, X., CONDRUZ, S., CHEN, J. & JOLICOEUR, M. 2016. A quantitative metabolomics study of high sodium response in *Clostridium acetobutylicum* ATCC 824 acetone-butanol-ethanol (ABE) fermentation. *Sci Rep,* 6. 28307**,** 1-13.

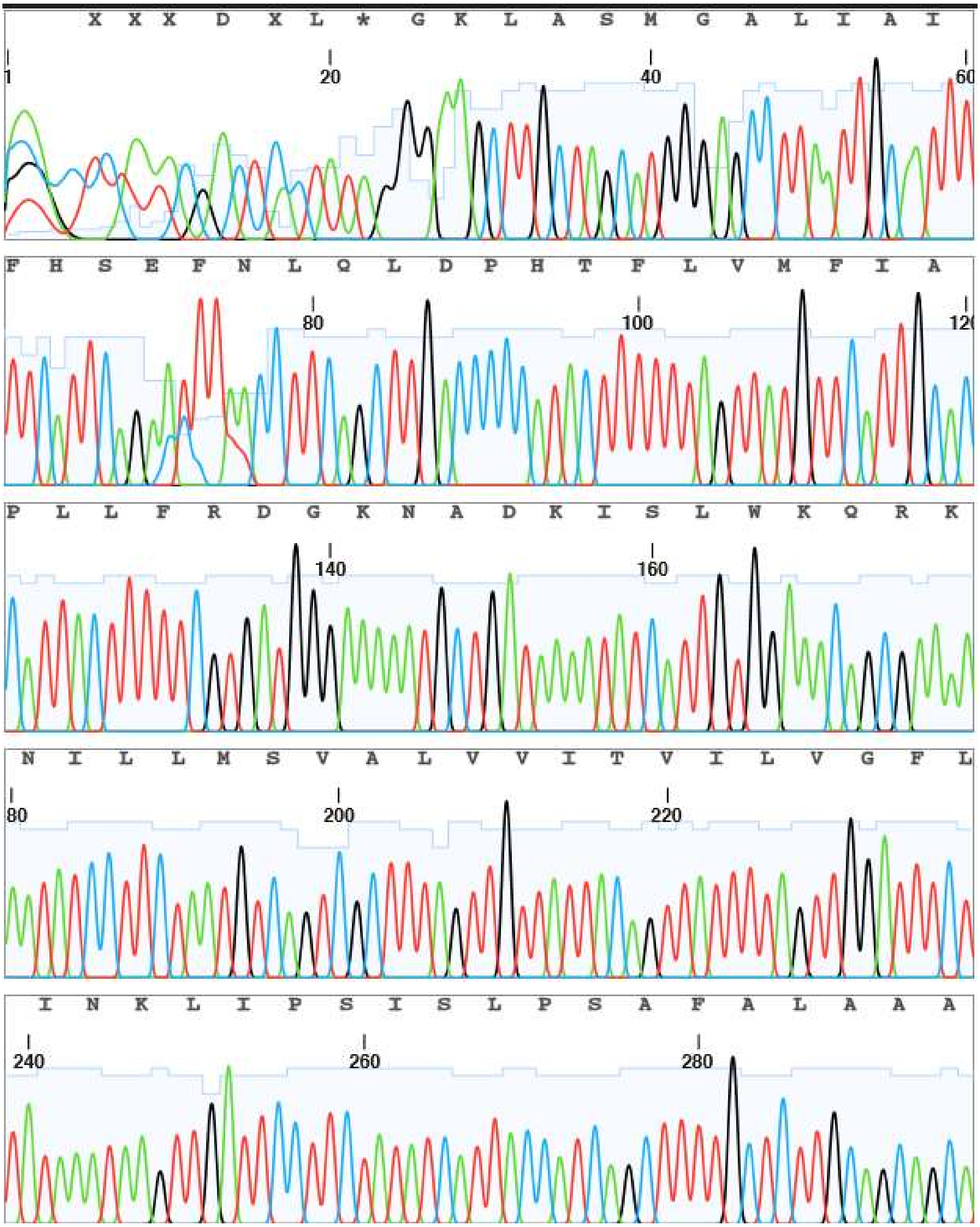
ZHOU, A., CAO, Y., ZHOU, D., HU, S., TAN, W., XIAO, X., YU, Y. & LI, X. 2020. Global transcriptomic analysis of *Cronobacter sakazakii* CICC 21544 by RNA-seq under inorganic acid and organic acid stresses. *Food Res Int,* 130. 108963.

ZHU, Y. & YANG, S.-T. 2003. Adaptation of *Clostridium tyrobutyricum* for Enhanced Tolerance to butyric acid in a fibrous-bed bioreactor. *Biotechnol. Prog.,* 19. 365-372.

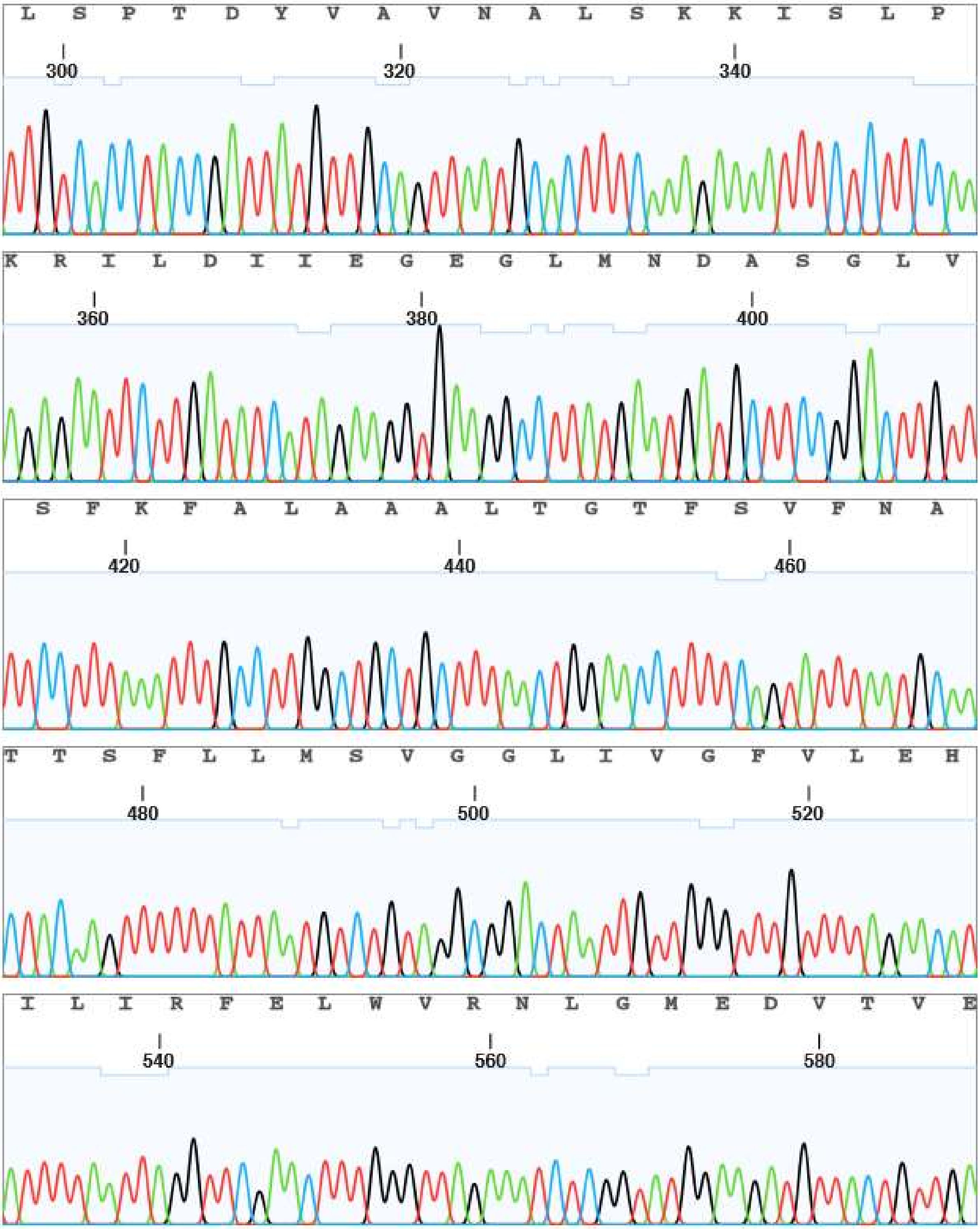
8. Appendix

8.1 Cpa1 sequencing

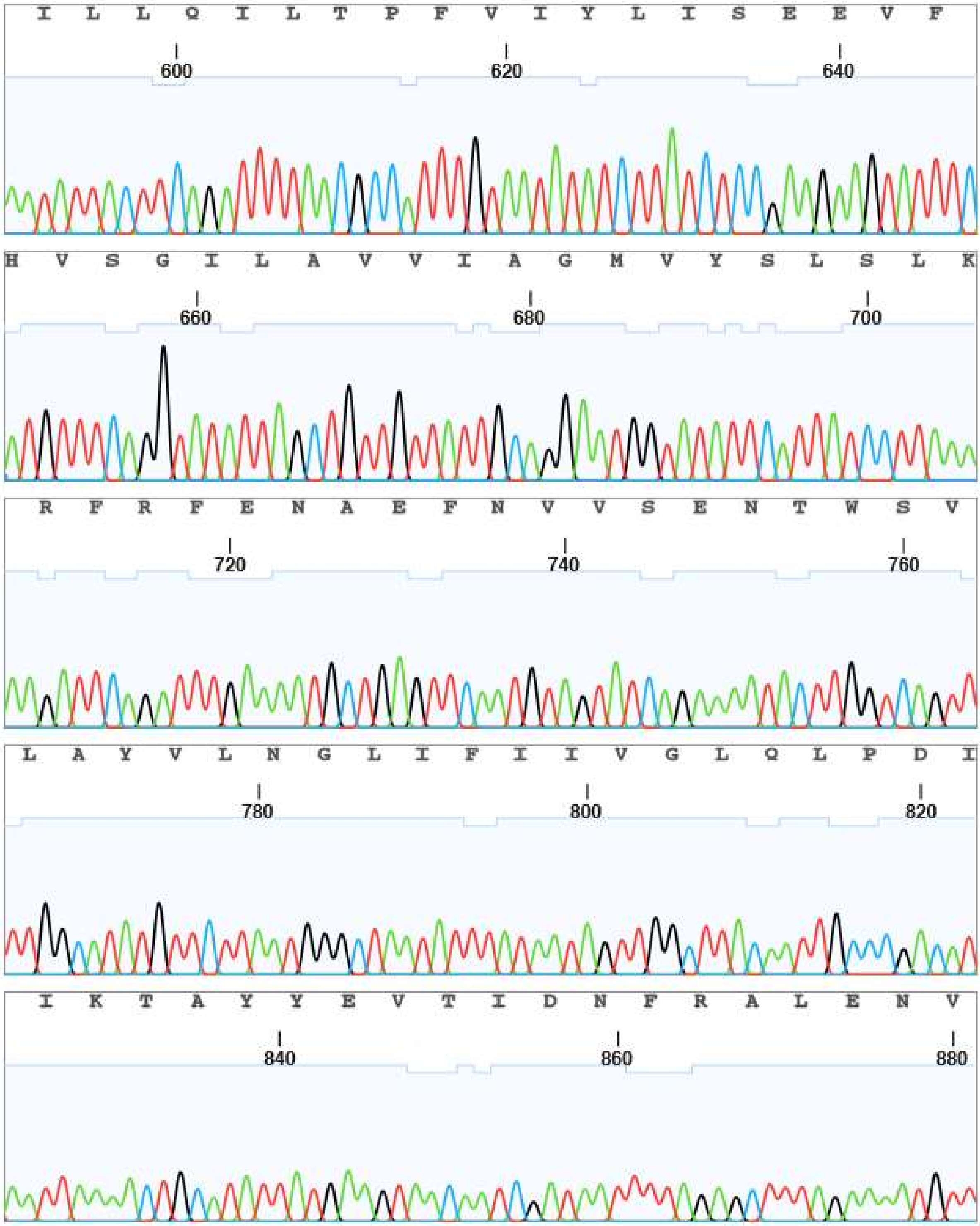
```
>AGF57903.1 Na+/H+ antiporter [Clostridium saccharoperbutylacetonicum N1-
4(HMT)]
MGALIAIFHSEFNLQLDPHTFLVMFIAPLLFRDGKNADKISLWKQRKNILLMSVALVVITVILVGFLINK
LIPSISLPSAFALAAALSPTDYVAVNALSKKISLPKRILDIIEGEGLMNDASGLVSFKFALAAALTGTFS
VFNATTSFLLMSVGGLIVGFVLEHILIRFELWVRNLGMEDVTVEILLQILTPFVIYLISEEVFHVSGILA
VVIAGMVYSLSLKRFRFENAEFNVVSENTWSVLAYVLNGLIFIIVGLQLPDIIKTAYYEVTIDNFRALEN
VLIITFALLLIRYLWVLFMYISDNECKNEKYSKIFHAALLTSLSGVRGAVTLATVLSIPLFLENGEPFPE
RTLILFLSVGVILTTLLVTTFILPLFAKKNVSEEADYLETLNKAQIKVWNNTIDKLKLEIENNKYISLTI
VEYKRRIHQLKRGNSYYKNWSIGDKEEKKWMLVCFKKEIENTETLLQQEKISETTAHAYEEAIRNKIKLL
SSVEGFSLKFKHAYTHFAHMISNIKRFKFILKGLEHNKEIENNKANIEMKELYSTNAKYLINYIKSIITP
ENEVTLKKIILYYQGMCWLGSKSILEKYNKYEKKKADIKAMQLERIIIQSLFENGDINWHIASELRKNLN
YIESDILR

>CP004121.1:4492649-4494565 Clostridium saccharoperbutylacetonicum N1-
4(HMT), complete genome
ATGGGAGCCTTAATTGCAATTTTTCATTCAGAATTTAACCTTCAGCTTGACCCCCATACTTTTTTAGTTA
TGTTCATTGCACCATTACTTTTTCGTGATGGGAAAAATGCTGATAAAATATCATTGTGGAAACAGCGAAA
AAATATCCTTCTAATGTCAGTCGCTTTAGTTGTTATTACAGTTATTTTAGTTGGATTTCTTATAAATAAG
TTGATTCCTTCTATATCATTACCATCAGCTTTTGCTCTTGCAGCAGCATTGTCACCTACCGATTATGTTG
CAGTTAATGCACTTTCAAAGAAAATTTCTCTTCCAAAGAGAATTCTTGATATCATAGAAGGTGAAGGCCT
TATGAATGATGCTTCCGGACTTGTTTCCTTTAAATTTGCCTTGGCTGCTGCTTTAACTGGAACCTTTTCA
GTATTTAATGCAACTACAAGTTTTTTATTAATGTCTGTAGGCGGACTAATTGTTGGGTTTGTTTTAGAAC
ATATTTTAATTAGGTTCGAACTTTGGGTTAGAAATCTCGGTATGGAAGATGTTACTGTTGAAATATTACT
TCAGATTTTAACGCCATTTGTAATATATCTTATCTCCGAAGAAGTATTTCATGTTTCAGGTATATTAGCT
GTTGTTATTGCAGGAATGGTATATTCATTATCCTTAAAAAGATTCAGATTTGAAAATGCTGAGTTCAATG
TAGTATCAGAAAATACTTGGTCAGTTTTGGCATATGTTCTTAATGGGCTAATATTTATAATAGTTGGCTT
ACAATTGCCCGACATAATTAAAACTGCATATTATGAAGTAACAATCGATAATTTTAGAGCTTTAGAAAAT
GTTTTAATAATAACTTTTGCTTTACTTCTAATCAGATATTTATGGGTACTATTTATGTATATTAGCGATA
ATGAGTGTAAAAACGAGAAATATAGTAAAATATTTCACGCTGCTCTTTTAACTTCTCTTTCGGGAGTAAG
AGGAGCTGTAACCTTAGCTACAGTCCTTTCAATTCCCTTGTTTCTTGAAAATGGTGAACCATTTCCTGAA
AGAACACTAATTTTATTTTTATCCGTTGGAGTAATTTTAACAACCCTTCTAGTAACAACTTTTATACTAC
CTCTTTTTGCAAAGAAAAATGTATCTGAGGAAGCAGATTATTTAGAAACCTTAAATAAAGCTCAAATAAA
GGTATGGAATAATACAATAGATAAACTAAAATTAGAAATTGAAAATAATAAATATATTTCTTTAACCATA
GTTGAATATAAACGCCGTATCCACCAATTAAAACGTGGAAATTCCTATTATAAAAATTGGAGCATAGGTG
ATAAAGAAGAAAAAAAGTGGATGCTTGTATGTTTTAAAAAAGAAATTGAAAATACAGAAACTCTATTGCA
GCAAGAGAAAATATCTGAAACAACAGCGCATGCTTATGAAGAAGCTATAAGAAATAAAATCAAACTACTC
TCTTCTGTTGAAGGTTTTTCATTAAAATTTAAGCACGCTTACACACATTTTGCTCATATGATTTCTAATA
TAAAGAGATTTAAATTTATTTTGAAGGGATTAGAGCATAATAAGGAAATTGAGAATAATAAGGCTAATAT
AGAAATGAAAGAGCTTTATTCAACTAATGCAAAATATCTAATTAACTACATCAAGTCAATTATCACTCCT
GAGAATGAAGTTACTTTGAAAAAAATAATTTTATATTACCAAGGAATGTGCTGGCTTGGAAGCAAATCTA
TCTTAGAAAAATATAATAAATATGAGAAAAAGAAAGCTGACATTAAAGCAATGCAATTAGAAAGAATTAT
TATACAATCTCTATTTGAAAATGGGGATATAAACTGGCATATAGCCTCTGAACTCAGAAAAAATTTAAAT
TATATTGAAAGCGATATTTTAAGATAA
```


X X X D X L * G K L A S M G A L I A I
1
20
40
60
F H S E F N L Q L D P H T F L V M F I A
80
100
12
P L L F R D G K N A D K I S L W K Q R K
140
160
N I L L M S V A L V V I T V I L V G F L
80
200
220
I N K L I P S I S L P S A F A L A A A
240
260
280

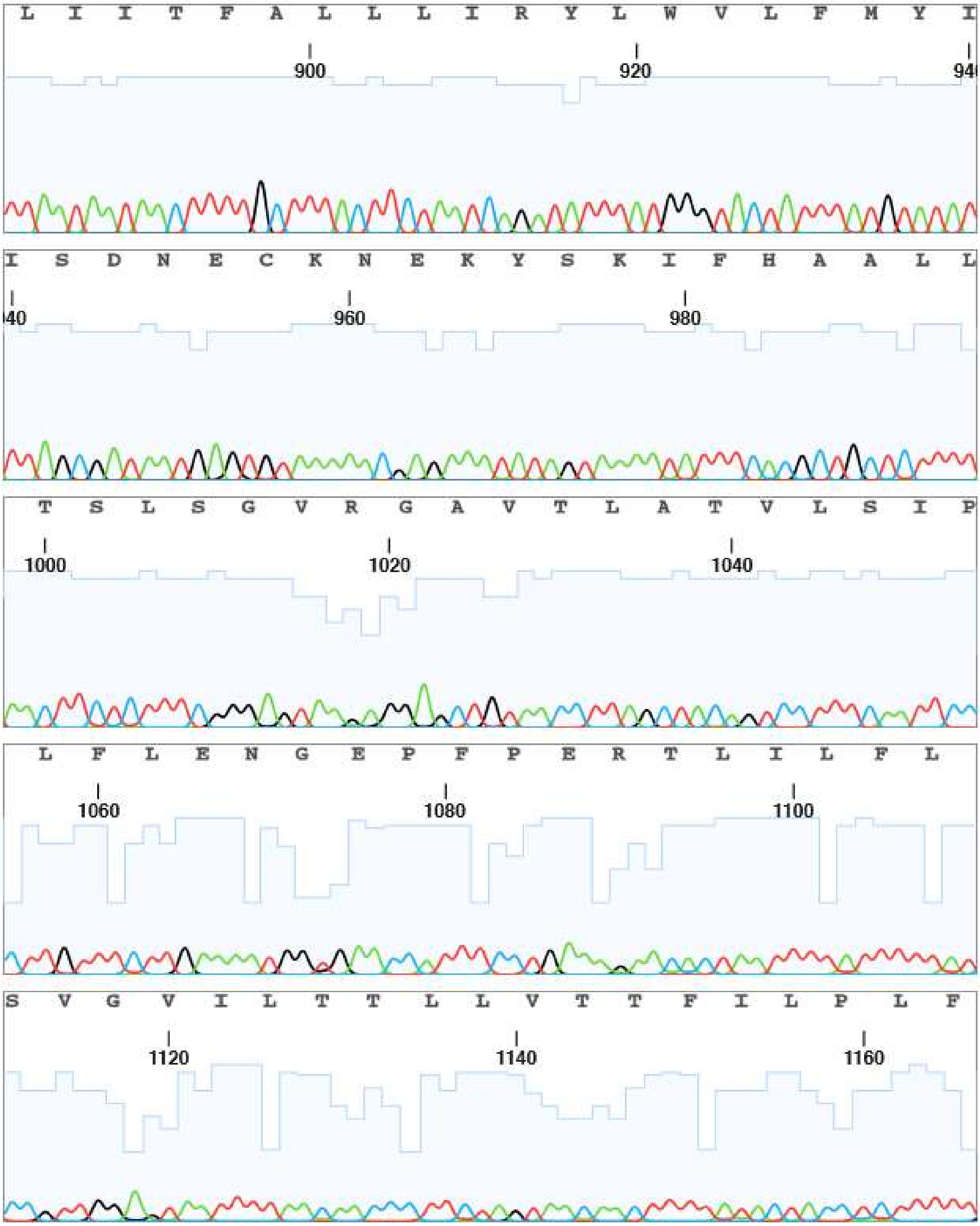
450676501_cpa11_M13R_B02.ab1 450676501_cpa11_M13R (18/04/2019) Page 2 / 5



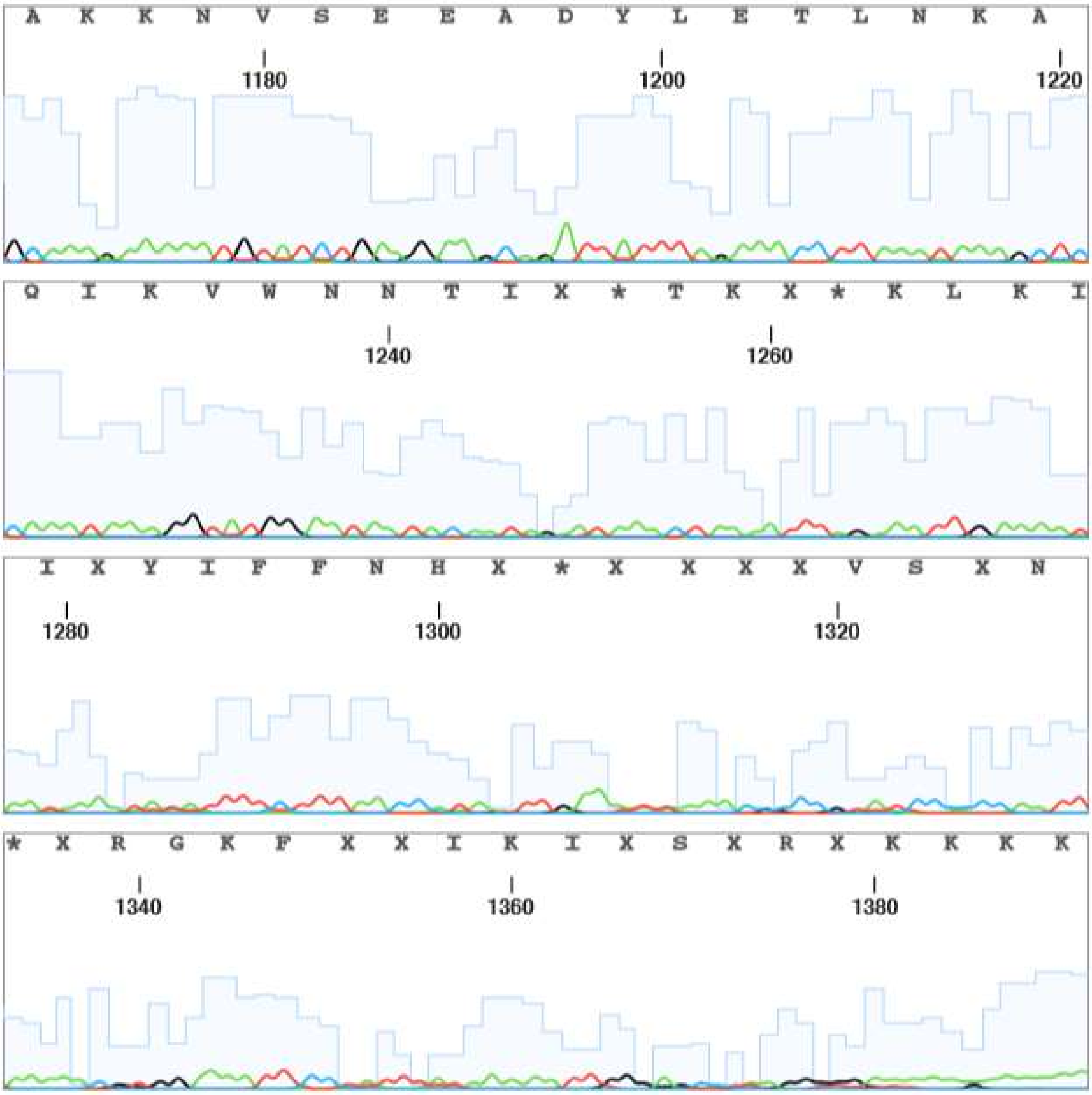
450676501_cpa11_M13R_B02.ab1 450676501_cpa11_M13R (18/04/2019) Page 3 / 5



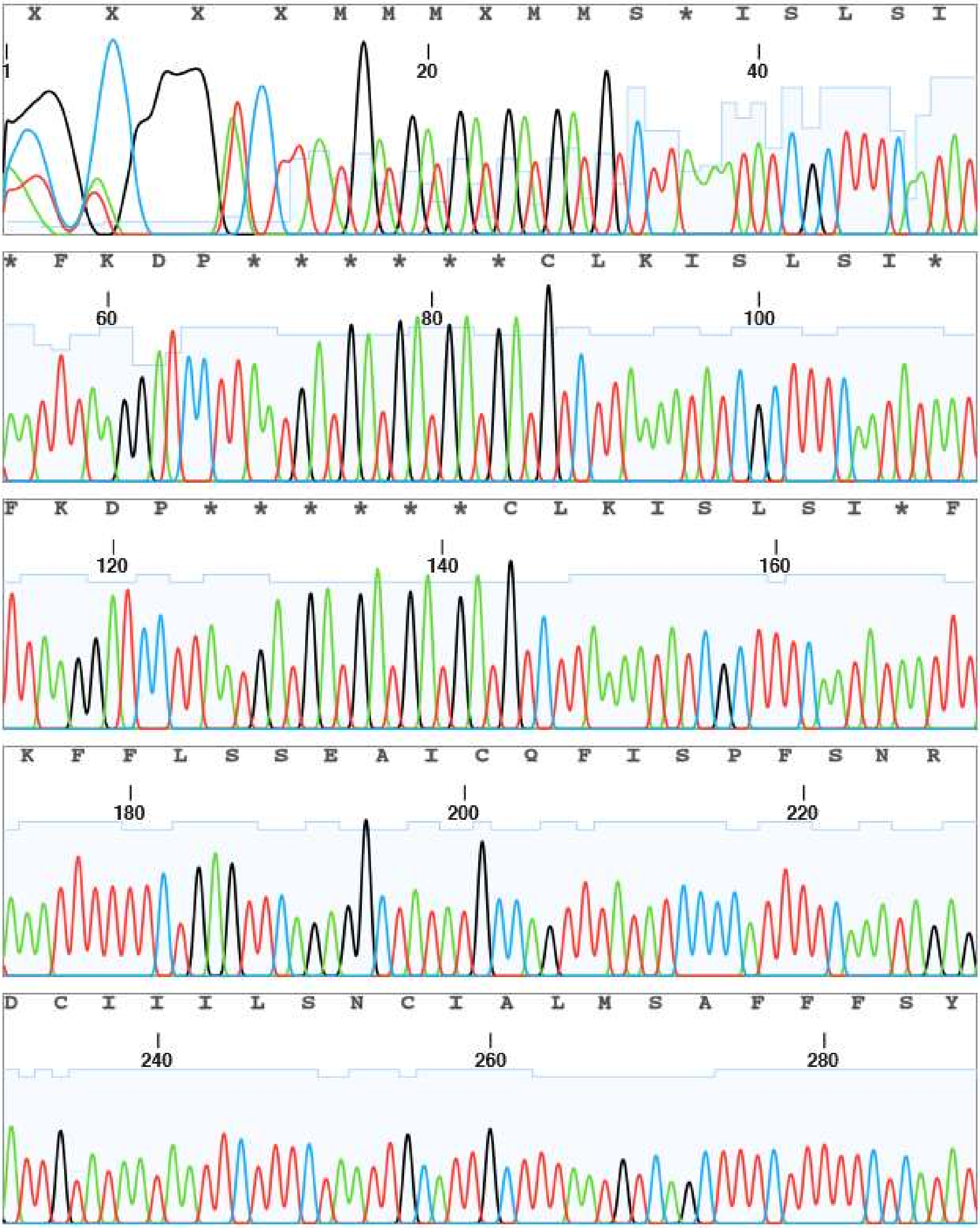
450676501_cpa11_M13R_B02.ab1 450676501_cpa11_M13R (18/04/2019) Page 4 / 5

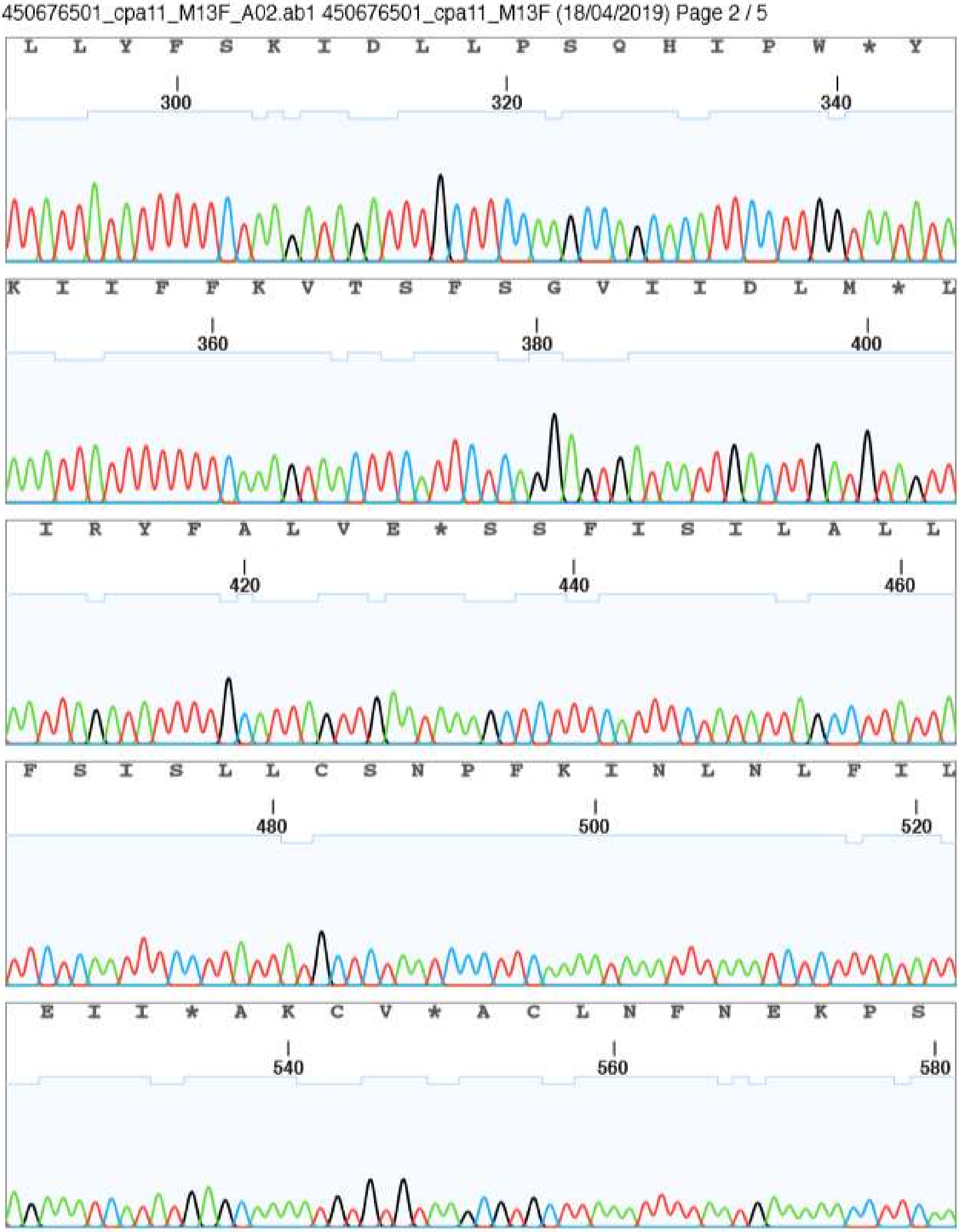


450676501_cpa11_M13R_B02.ab1 450676501_cpa11_M13R (18/04/2019) Page 5 / 5

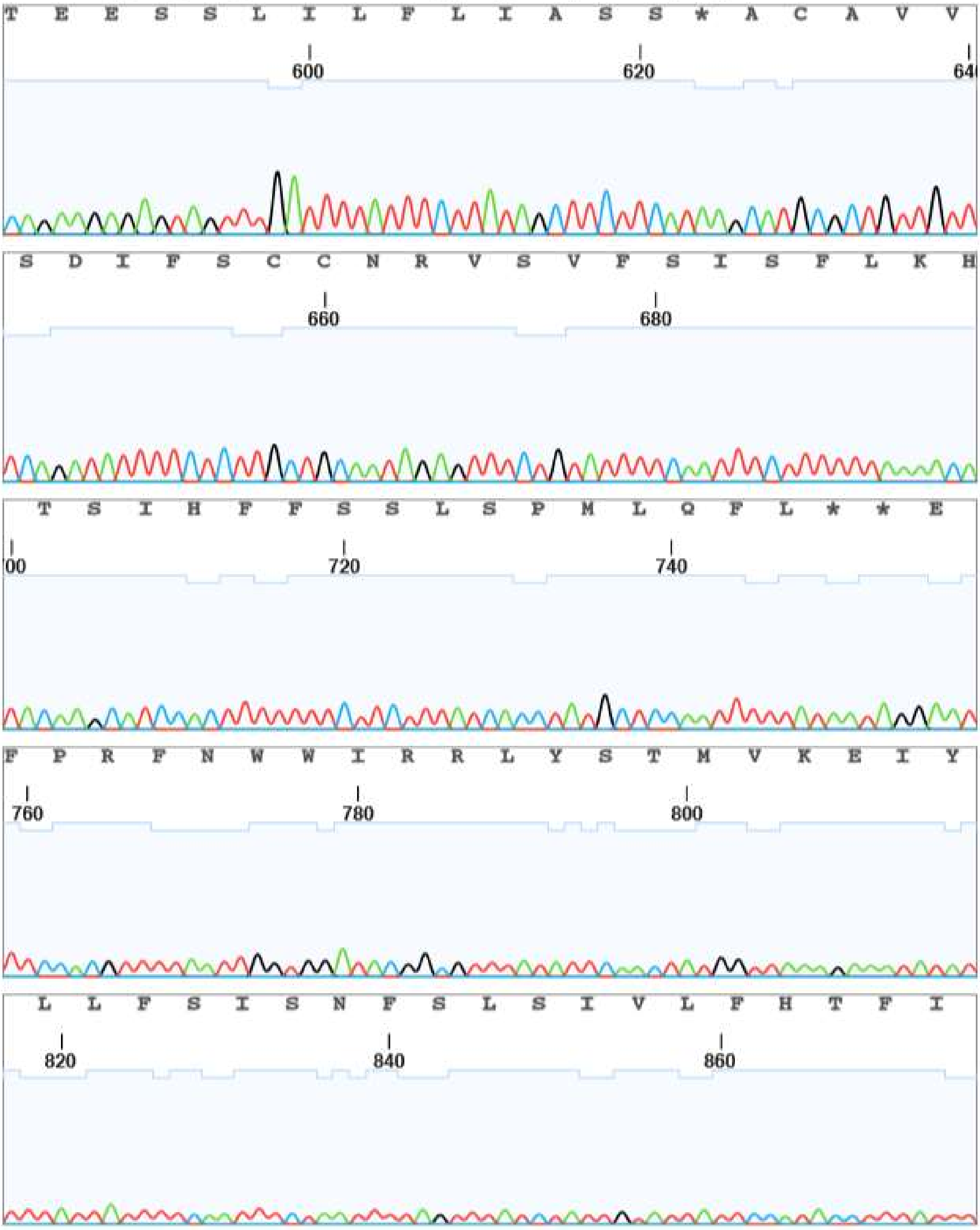


450676501_cpa11_M13F_A02.ab1 450676501_cpa11_M13F (18/04/2019) Page 1 / 5

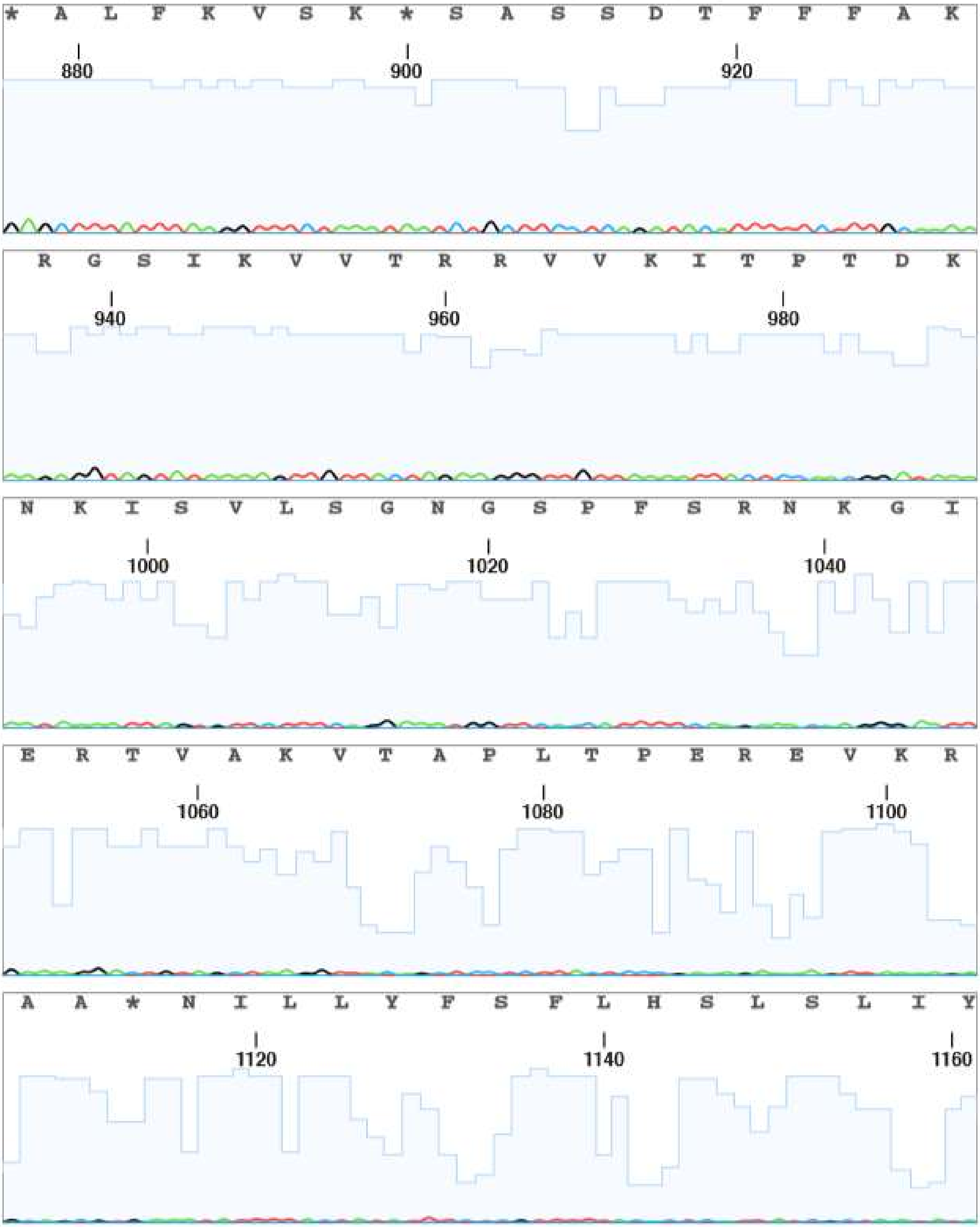



450676501_cpa11_M13F_A02.ab1 450676501_cpa11_M13F (18/04/2019) Page 2 / 5

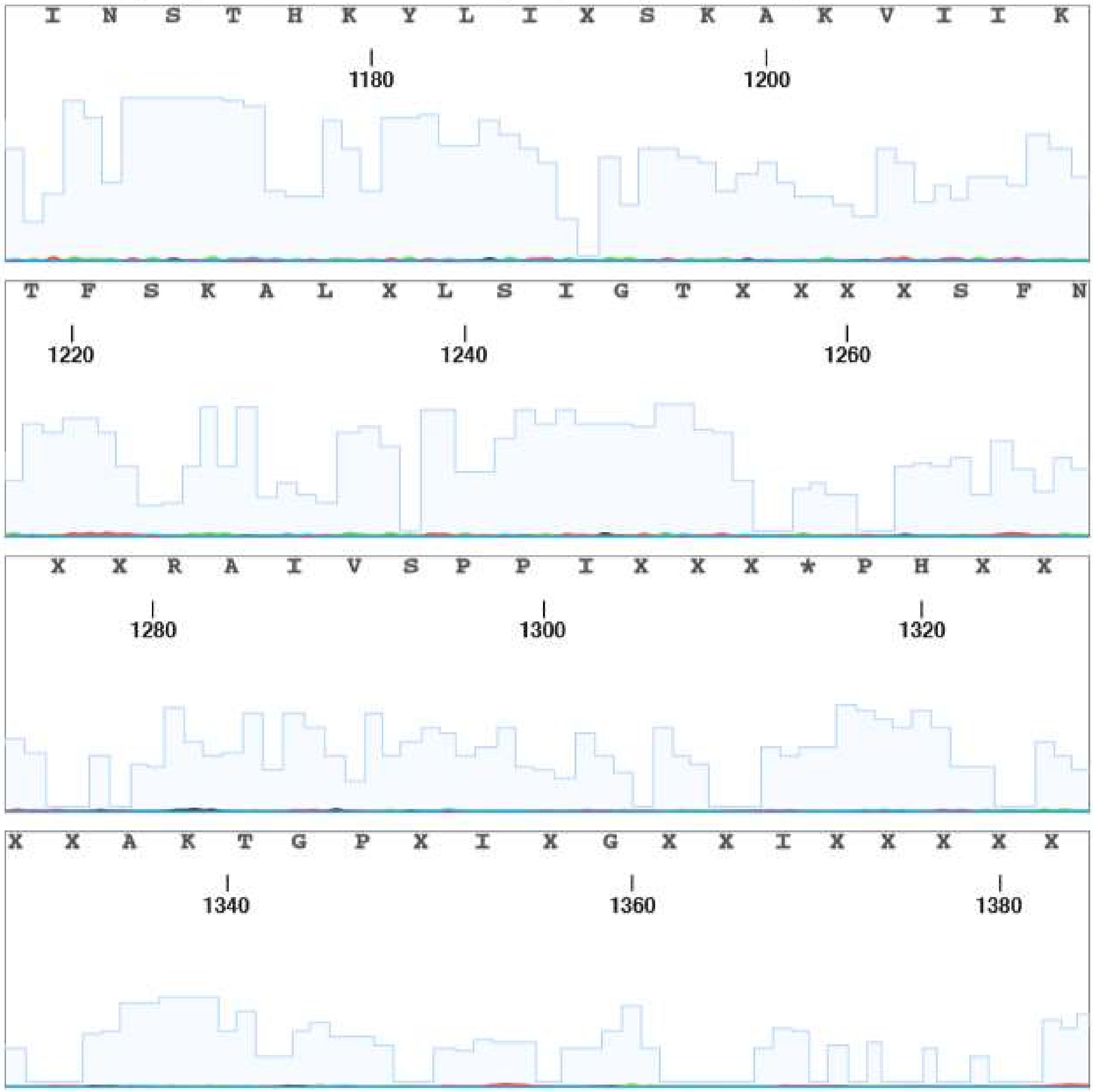
450676501_cpa11_M13F_A02.ab1 450676501_cpa11_M13F (18/04/2019) Page 3 / 5



450676501_cpa11_M13F_A02.ab1 450676501_cpa11_M13F (18/04/2019) Page 4 / 5



450676501_cpa11_M13F_A02.ab1 450676501_cpa11_M13F (18/04/2019) Page 5 / 5



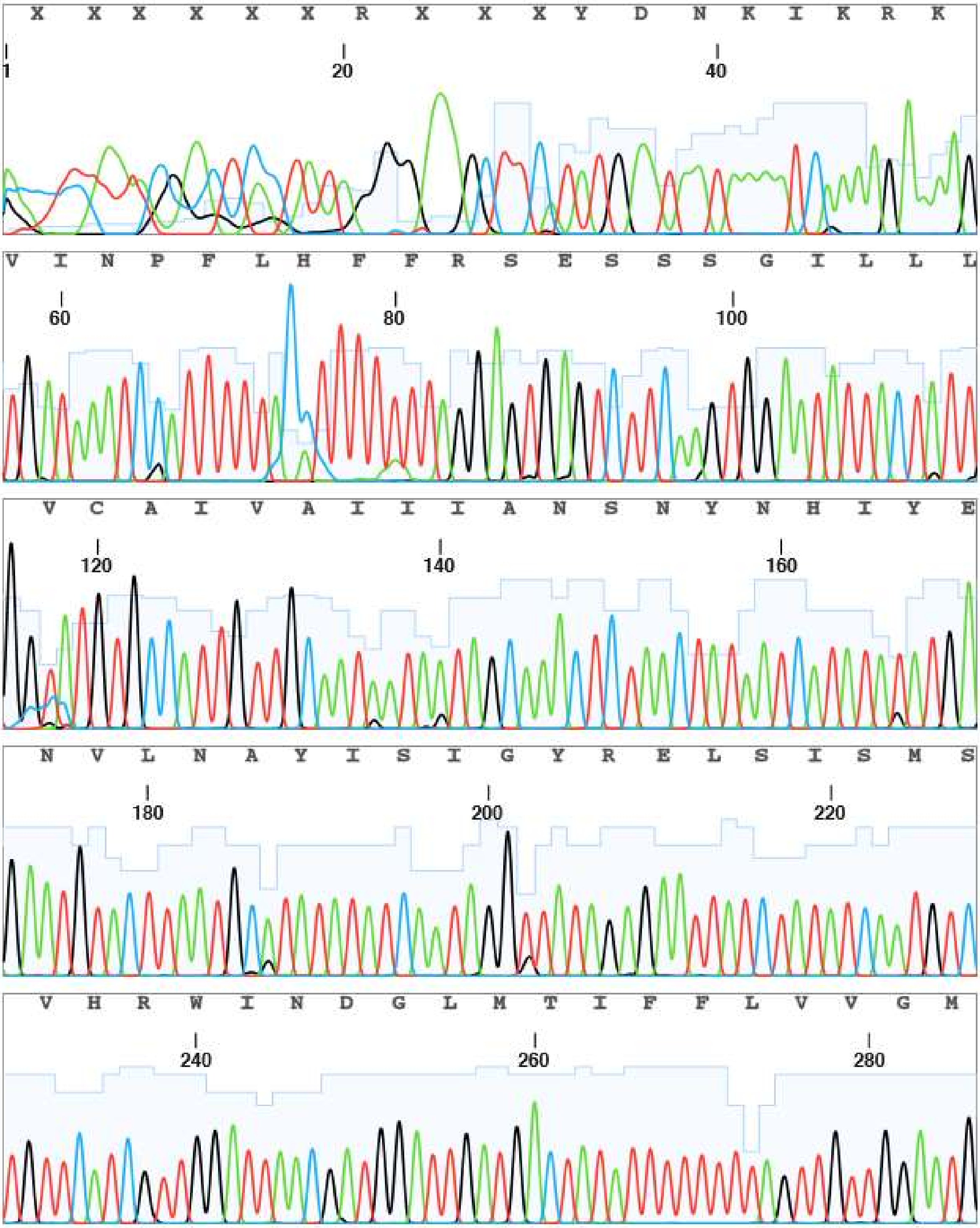
8.2 NhaA sequence

```
>AGF56626.1 Na(+)/H(+) antiporter NhaA 2 [Clostridium
saccharoperbutylacetonicum N1-4(HMT)]
MNNKIKRKVINPFLHFFRSESSSGILLLVCAIVAIIIANSNYNHIYENVLNAYISIGYRELSISMSVHRW
INDGLMTIFFLVVGMEIKREVVFGELKSFKRTILPISAAIGGMVVPAIIYVLFNIGEPTISGWGIPMATD
IAFALGMLSLVAKNAPKGIVVFLTALAIVDDLGAILVIAIFYNSQISWLALIMGLVIFISILLANKLKVK
YVSVYIILGILLWLCFLKSGVHATIAGVLLGMALPIGENIHKFKSSMLYKFEHKLTSWSSYVIMPIFAFA
NSGIEISFETFSKDIFSPVSLGIIFGLFIGKQVGIFVTSYLLVKFKIAKFPASVTKRHLYGASVLGGIGF
TMSIFVSSLSFTDMEILSMAKICIMIASVLSALYGTIVFKVISFKNKRIA
```

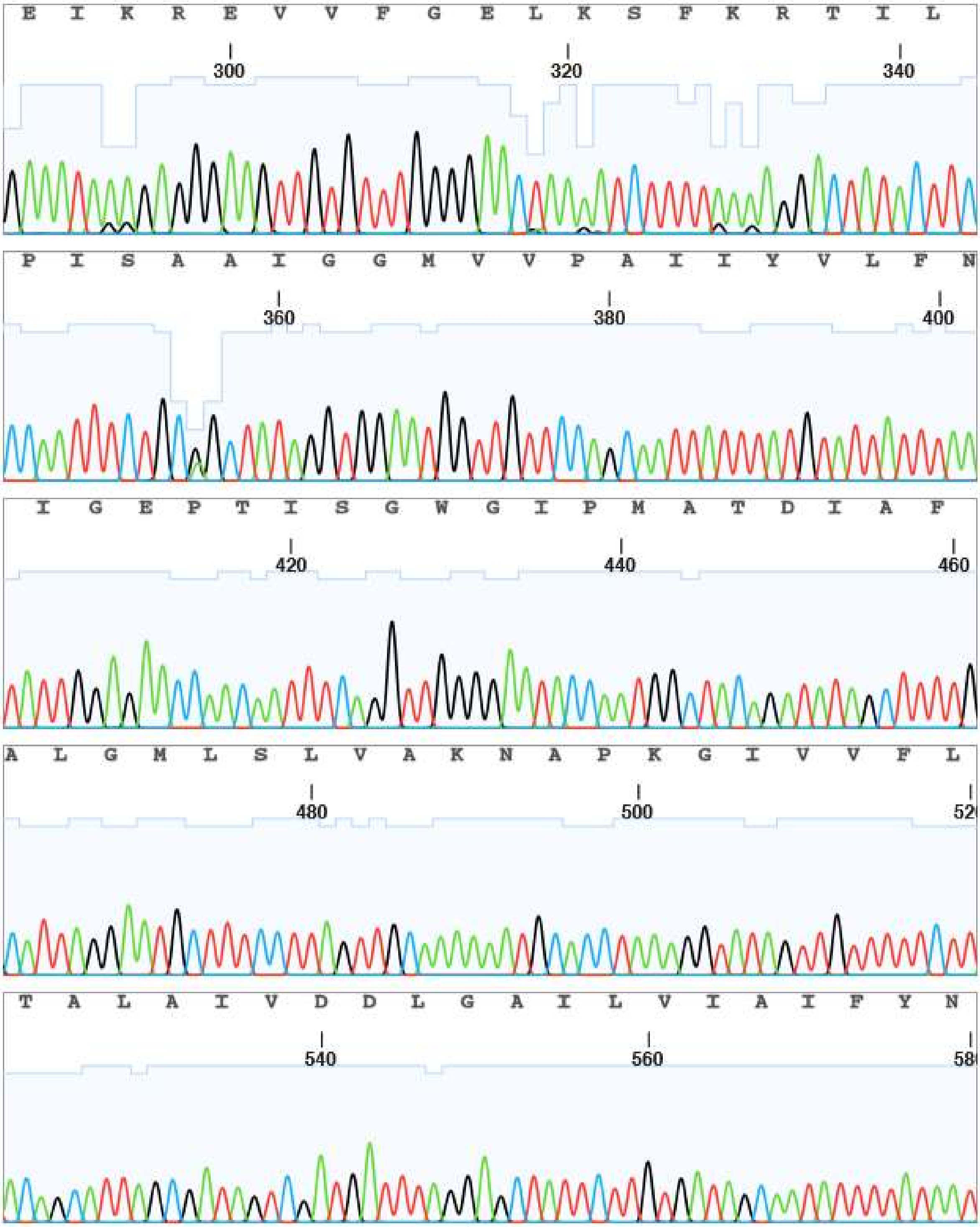
```
>CP004121.1:3101087-3102289 Clostridium saccharoperbutylacetonicum N1-
4(HMT), complete genome
ATGAATAATAAAATCAAAAGAAAAGTGATAAATCCATTTTTACATTTTTTTAGGAGTGAGTCTTCAAGTG
GAATATTACTATTGGTATGTGCCATTGTTGCAATAATAATAGCAAACTCTAACTATAATCATATATATGA
GAATGTACTTAATGCATATATATCAATAGGTTATAGAGAATTATCTATATCAATGTCTGTTCATCGTTGG
ATTAACGATGGATTGATGACTATATTTTTTTTAGTTGTTGGAATGGAAATAAAGAGGGAAGTTGTGTTTG
GGGAACTAAAATCTTTTAAAAGGACTATACTTCCAATTTCTGCGGCTATAGGTGGAATGGTTGTTCCAGC
AATTATTTATGTATTATTTAATATTGGAGAACCAACAATTTCAGGTTGGGGAATACCAATGGCTACAGAT
ATAGCTTTTGCATTAGGAATGCTTTCCTTAGTTGCAAAAAATGCACCTAAAGGTATAGTTGTTTTTCTTA
CAGCATTAGCGATAGTCGATGATTTAGGAGCTATTCTTGTGATAGCAATATTTTATAATAGCCAAATTTC
ATGGCTTGCTCTTATTATGGGGCTAGTGATATTTATAAGCATTTTGCTAGCAAATAAATTAAAAGTTAAG
TATGTATCAGTTTATATTATTTTGGGCATATTGCTTTGGTTATGTTTTTTGAAATCAGGAGTACATGCGA
CAATAGCTGGAGTACTACTTGGAATGGCACTTCCTATTGGAGAAAATATTCATAAATTTAAATCTTCAAT
GTTATATAAGTTCGAACATAAATTAACTTCTTGGTCAAGTTATGTAATAATGCCAATTTTTGCTTTTGCA
AATTCAGGGATAGAAATTAGTTTTGAAACTTTTTCTAAGGACATATTTTCGCCAGTAAGTTTAGGGATTA
TATTTGGACTCTTTATTGGAAAACAGGTTGGAATATTTGTAACTTCTTATTTATTAGTCAAATTTAAAAT
TGCAAAGTTTCCAGCAAGTGTAACTAAAAGACATTTATATGGAGCAAGTGTTTTGGGGGGGATTGGTTTT
ACAATGTCAATATTTGTTTCATCGTTATCTTTTACTGACATGGAGATACTATCAATGGCCAAAATATGTA
TAATGATTGCTTCAGTATTATCAGCATTATATGGAACAATAGTATTTAAAGTTATCAGTTTTAAGAATAA
AAGAATTGCCTAA
```

pNhaA

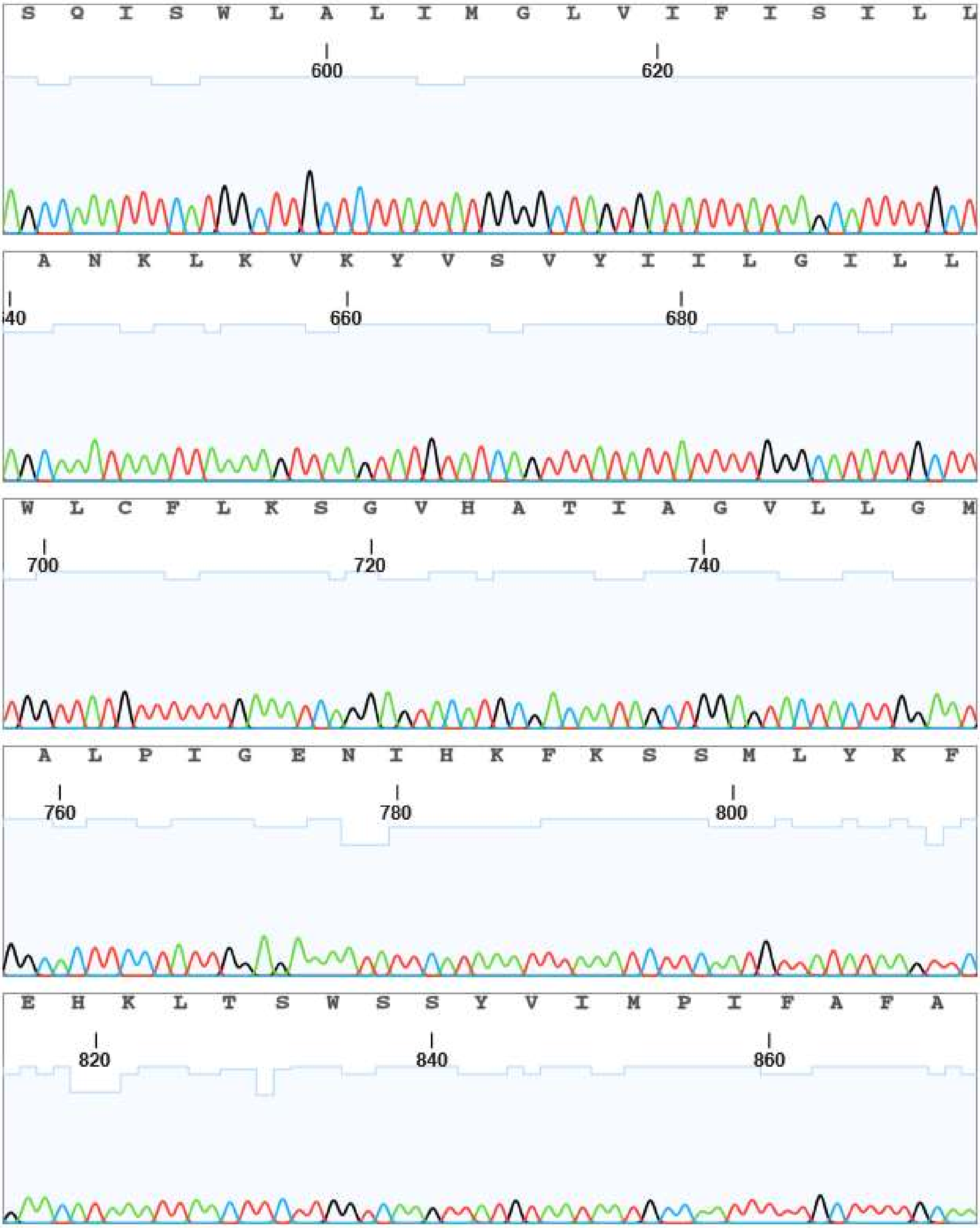
451863801_pNhaA_M13R_D09.ab1 451863801_pNhaA_M13R (20/06/2019) Page 1 / 7

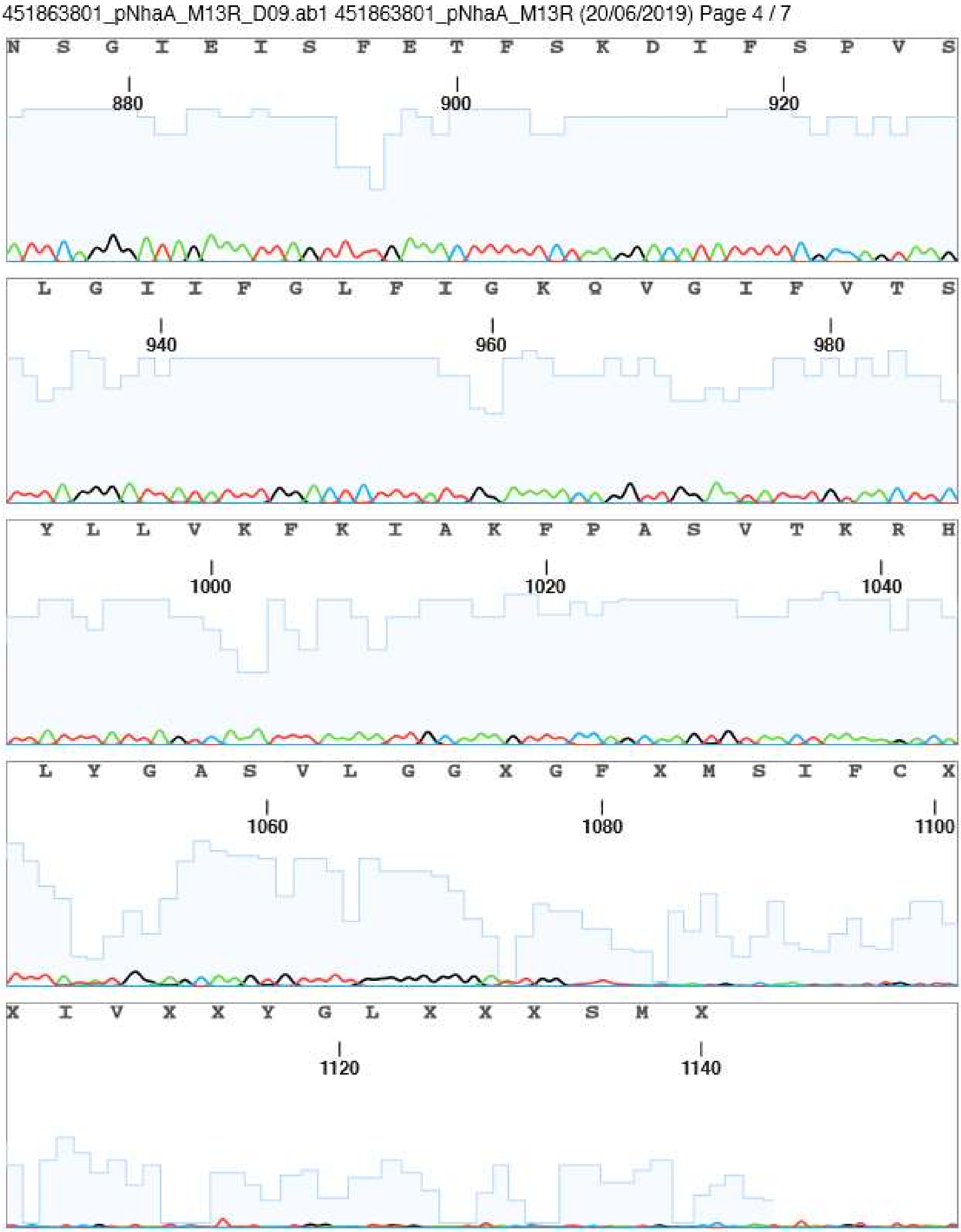


451863801_pNhaA_M13R_D09.ab1 451863801_pNhaA_M13R (20/06/2019) Page 2 / 7



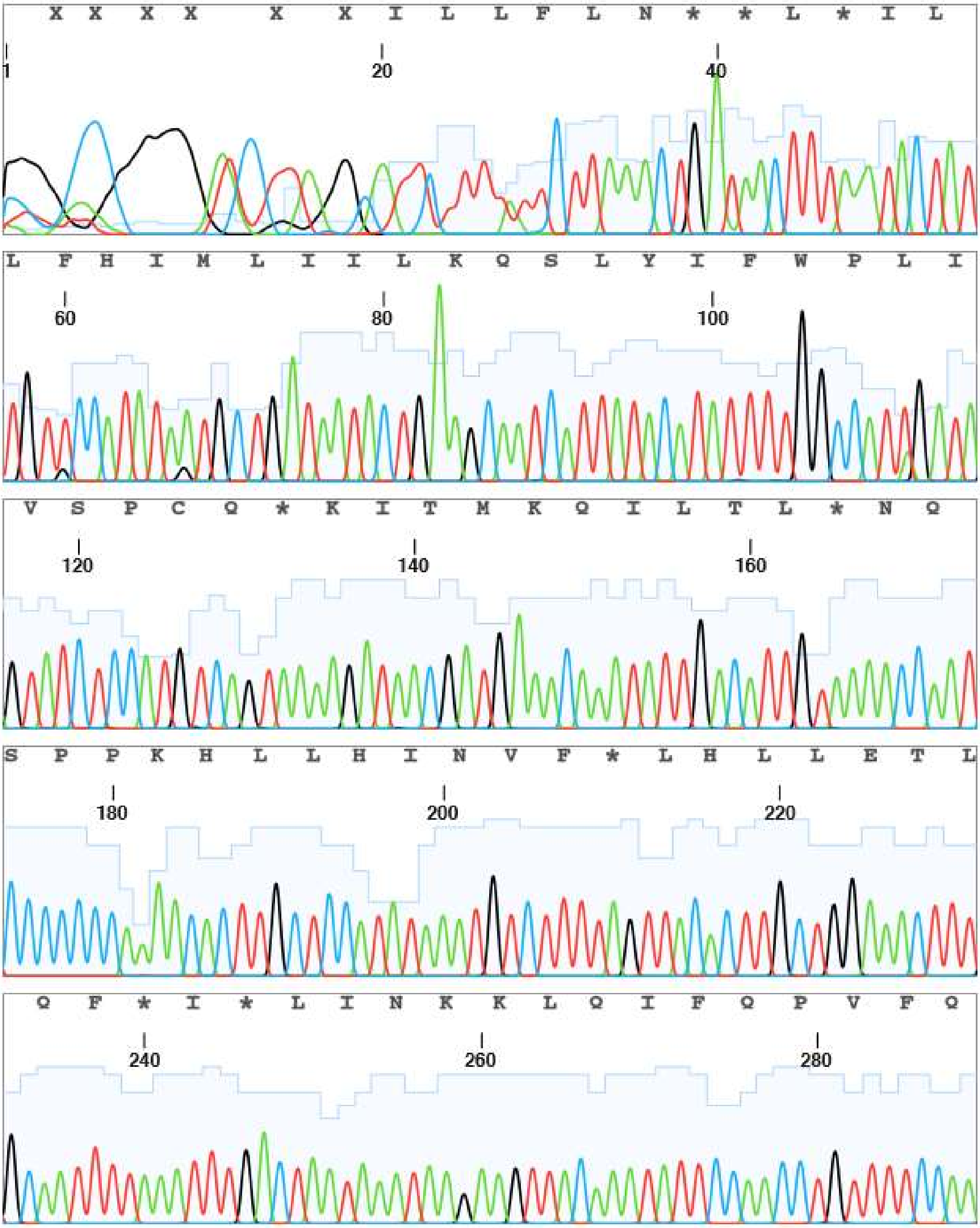
451863801_pNhaA_M13R_D09.ab1 451863801_pNhaA_M13R (20/06/2019) Page 3 / 7



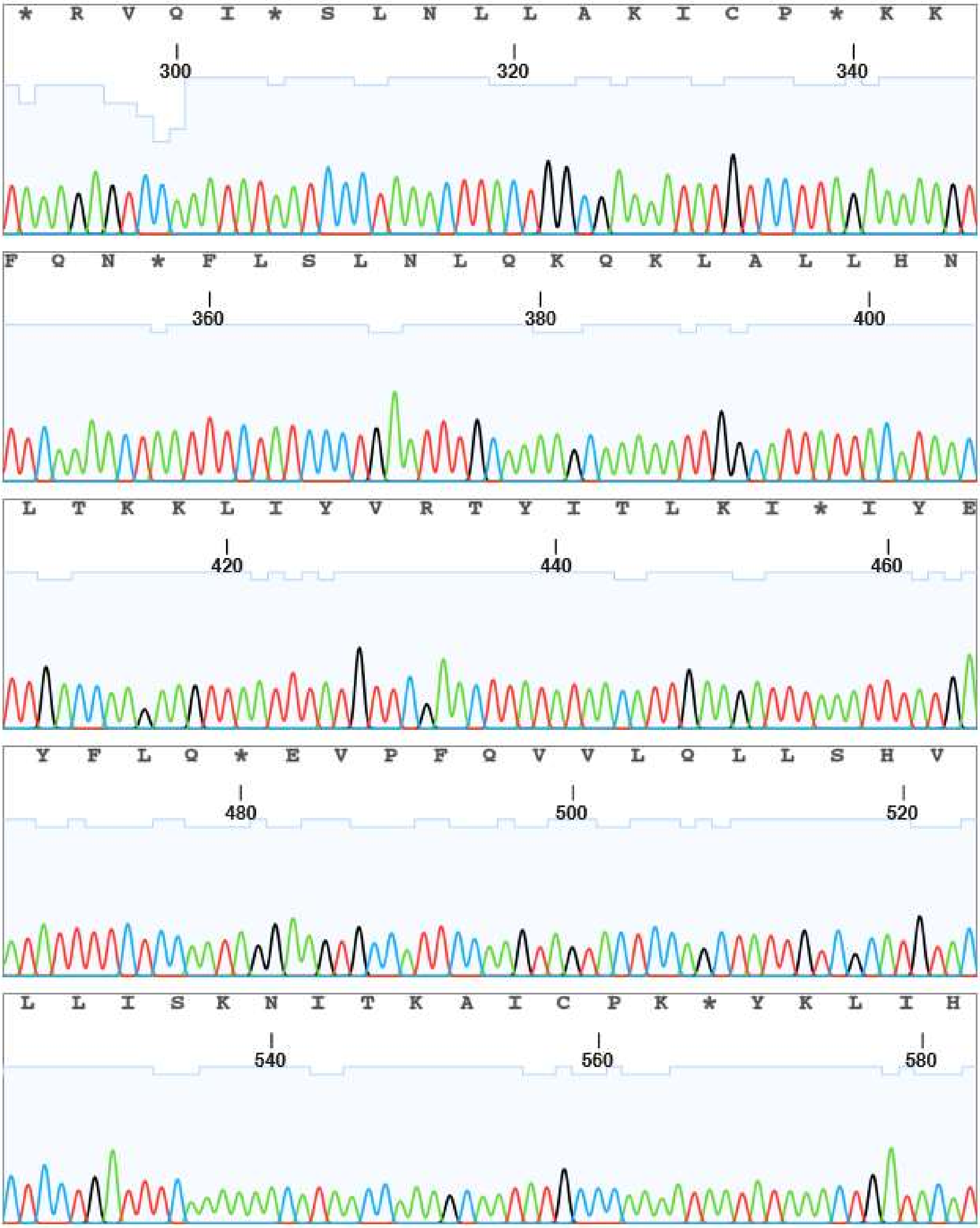


451863801_pNhaA_M13R_D09.ab1 451863801_pNhaA_M13R (20/06/2019) Page 4 / 7

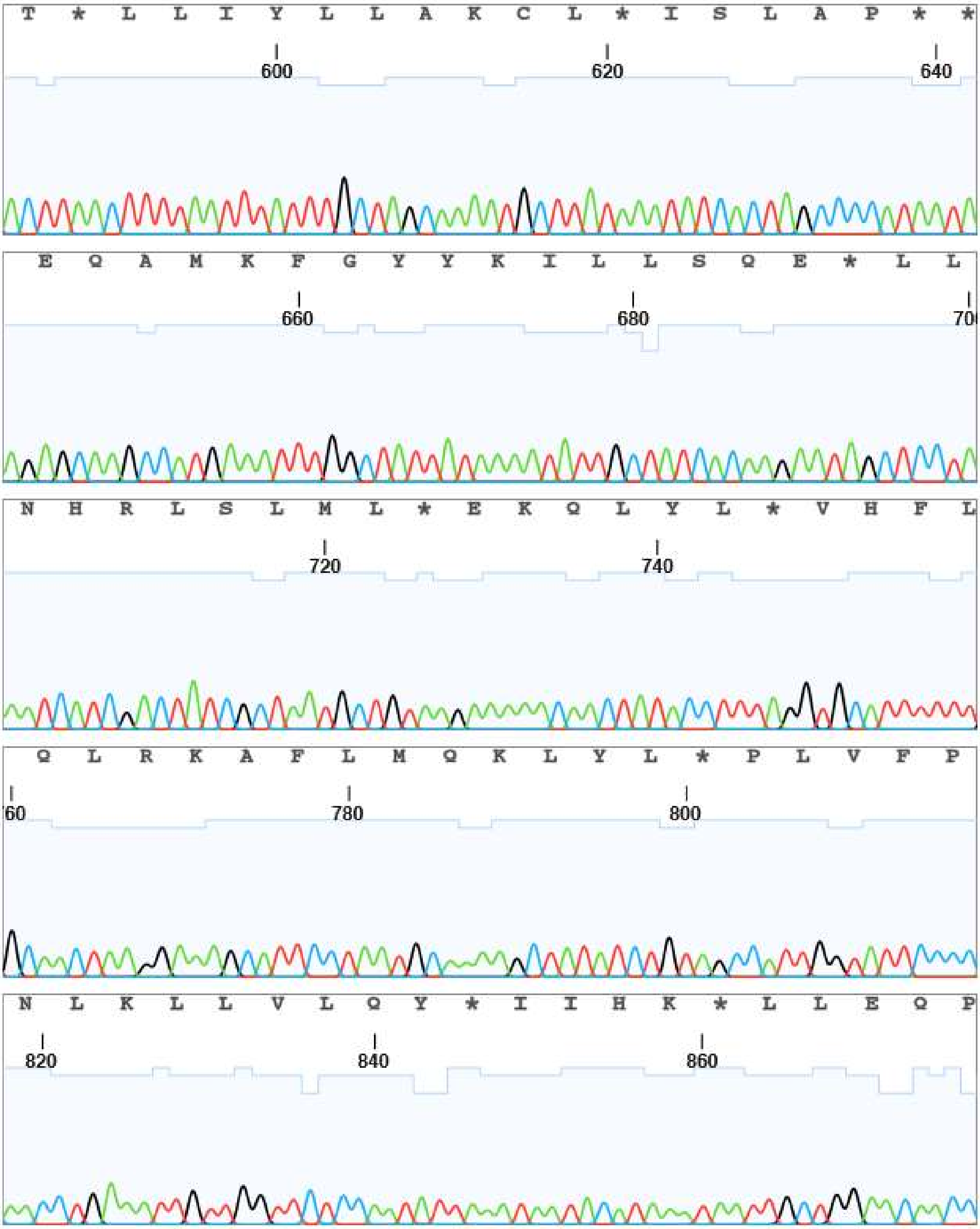
451863801_pNhaA_M13F_C09.ab1 451863801_pNhaA_M13F (20/06/2019) Page 1 / 7



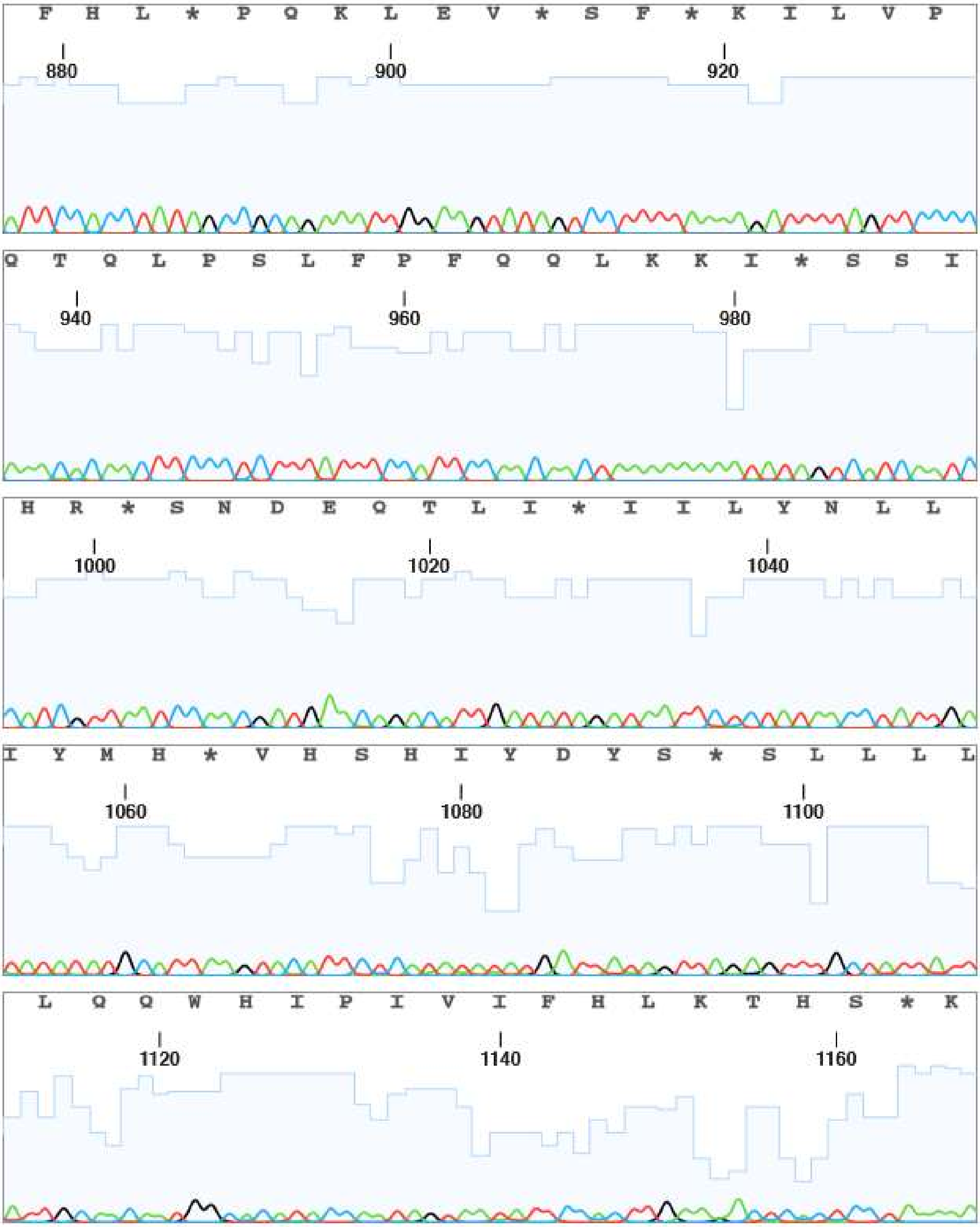
451863801_pNhaA_M13F_C09.ab1 451863801_pNhaA_M13F (20/06/2019) Page 2 / 7

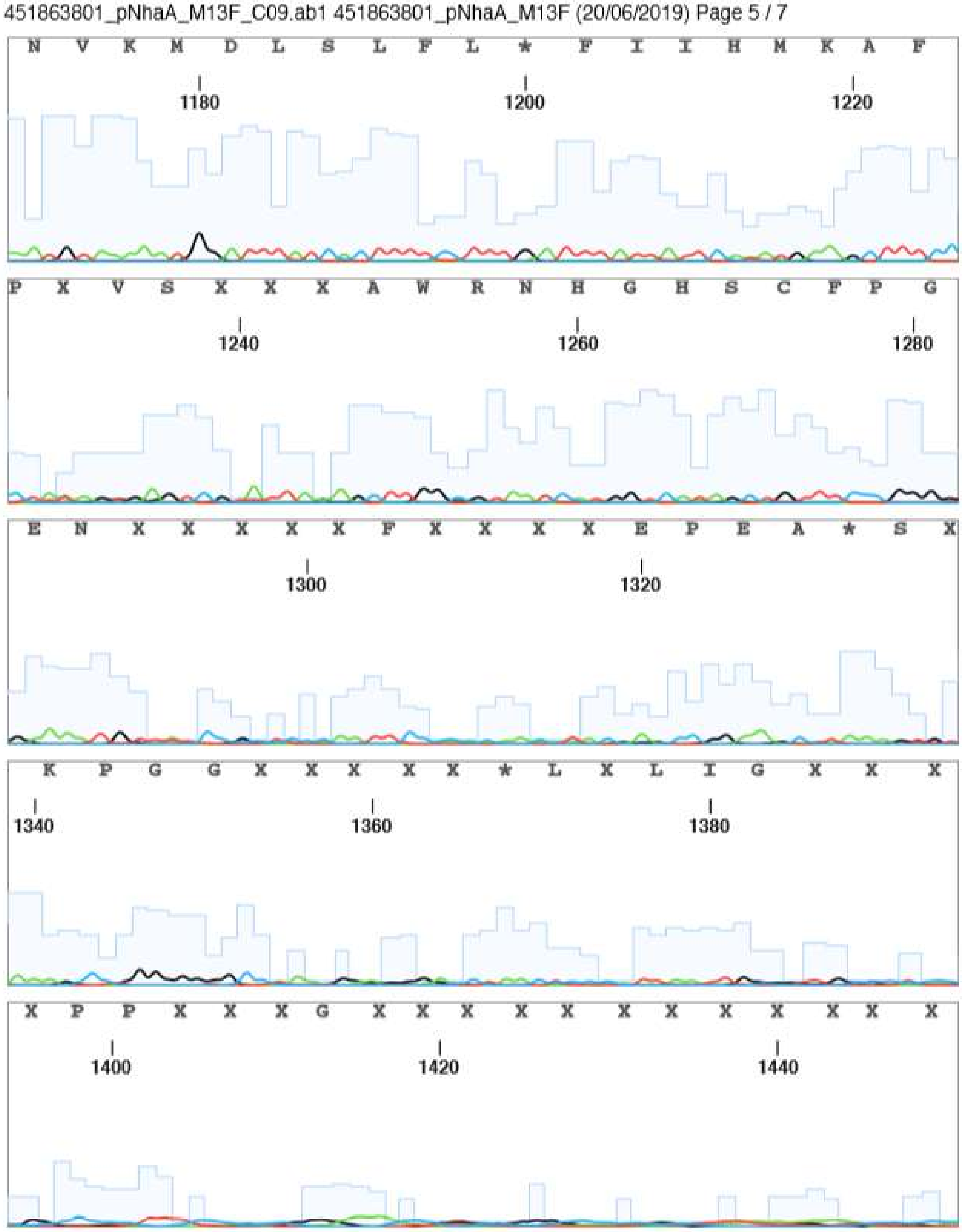


451863801_pNhaA_M13F_C09.ab1 451863801_pNhaA_M13F (20/06/2019) Page 3 / 7



451863801_pNhaA_M13F_C09.ab1 451863801_pNhaA_M13F (20/06/2019) Page 4 / 7

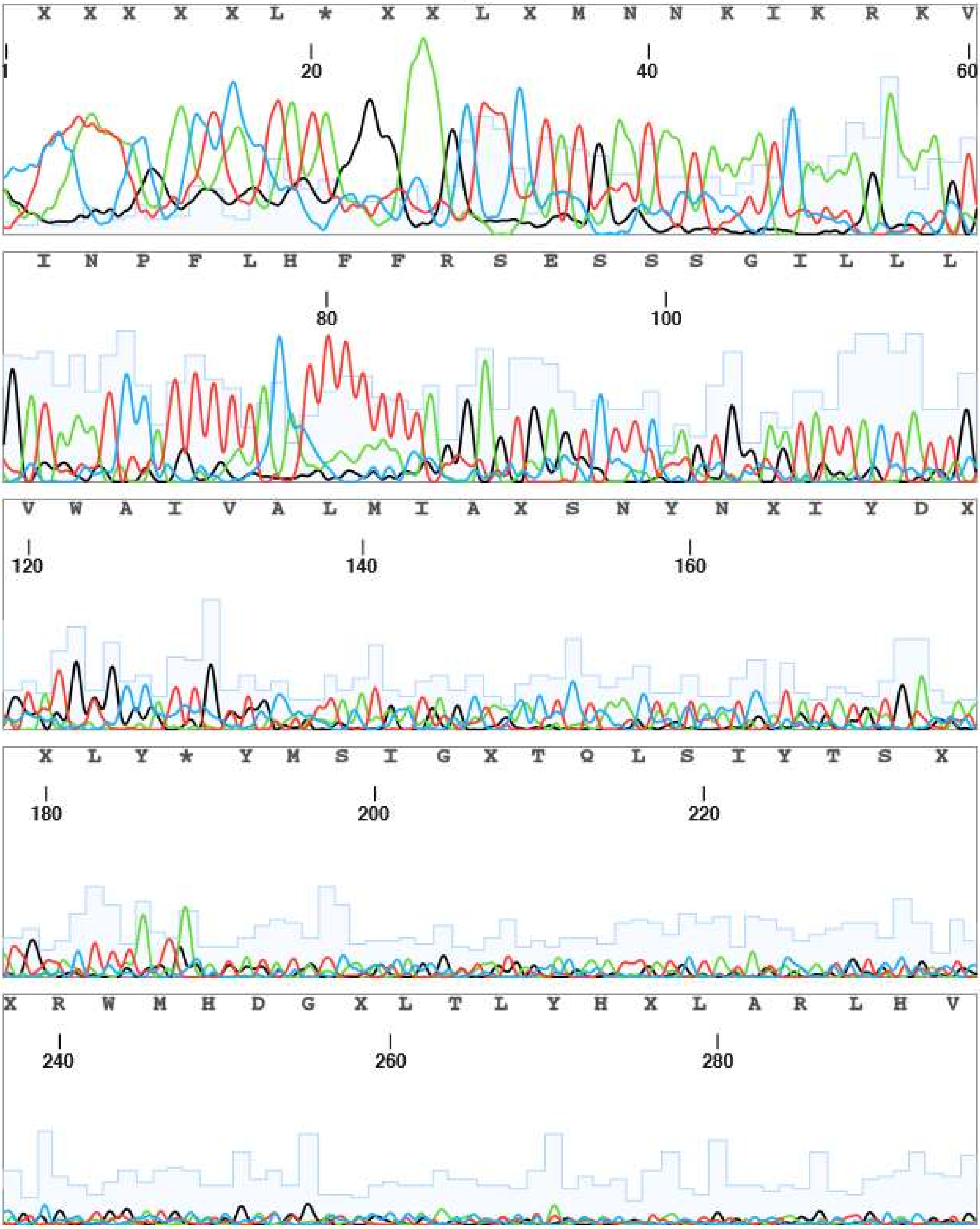




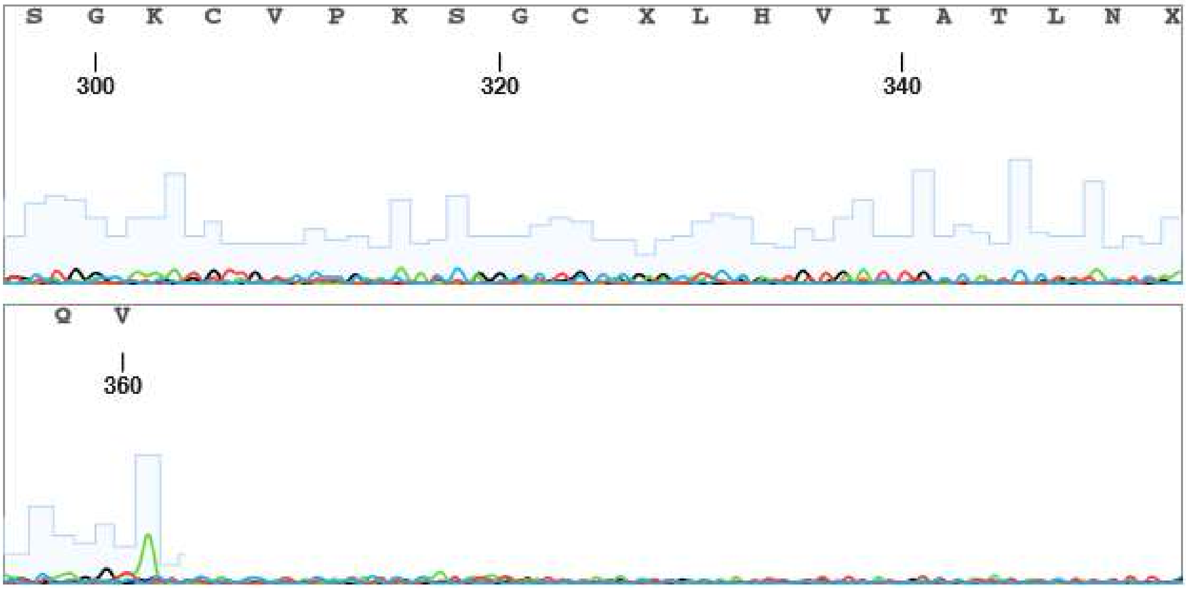
451863801_pNhaA_M13F_C09.ab1 451863801_pNhaA_M13F (20/06/2019) Page 5 / 7

pNhaA-Histag

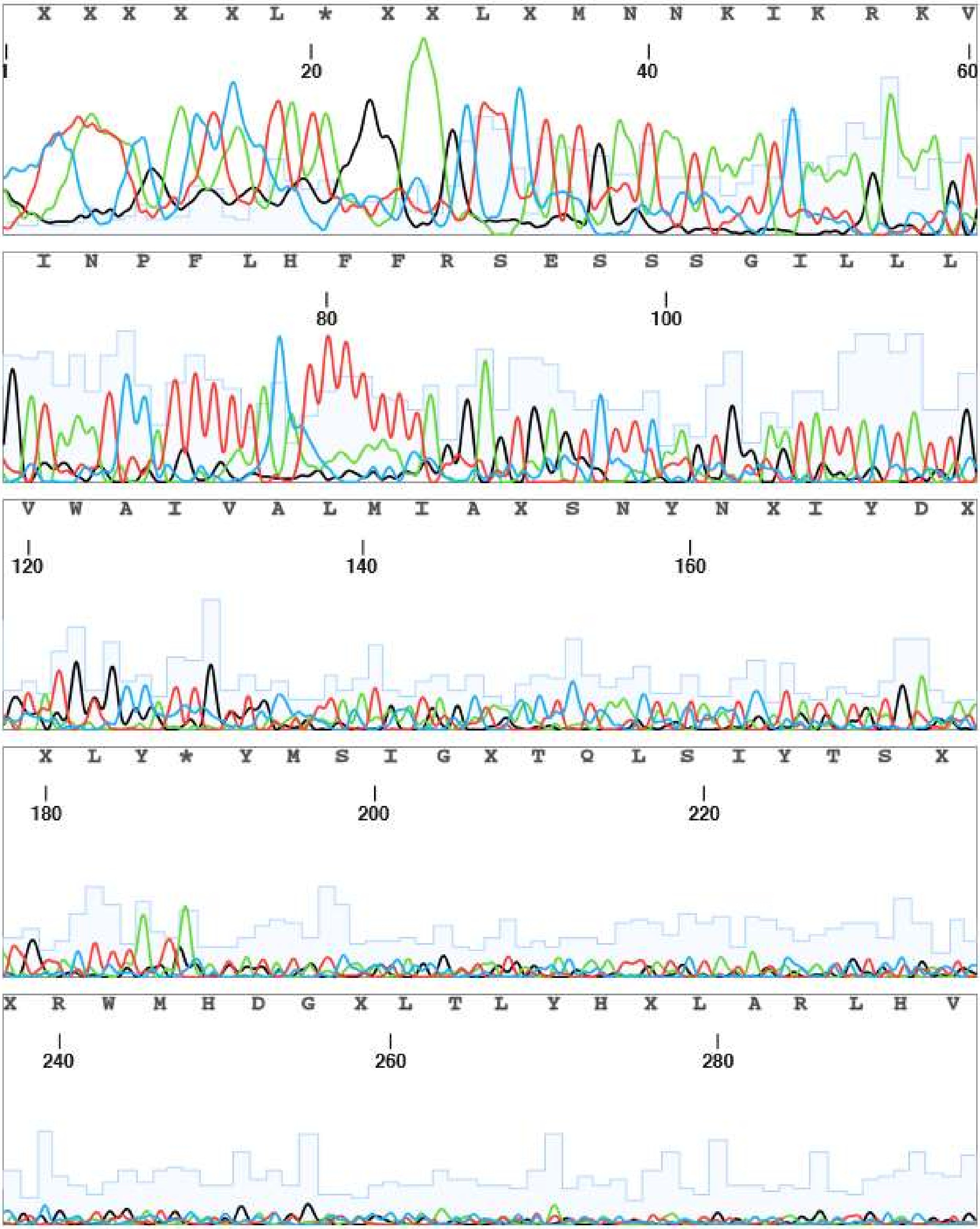
451863801_pNhaAH_M13R_B09.ab1 451863801_pNhaAH_M13R (20/06/2019) Page 1 / 6

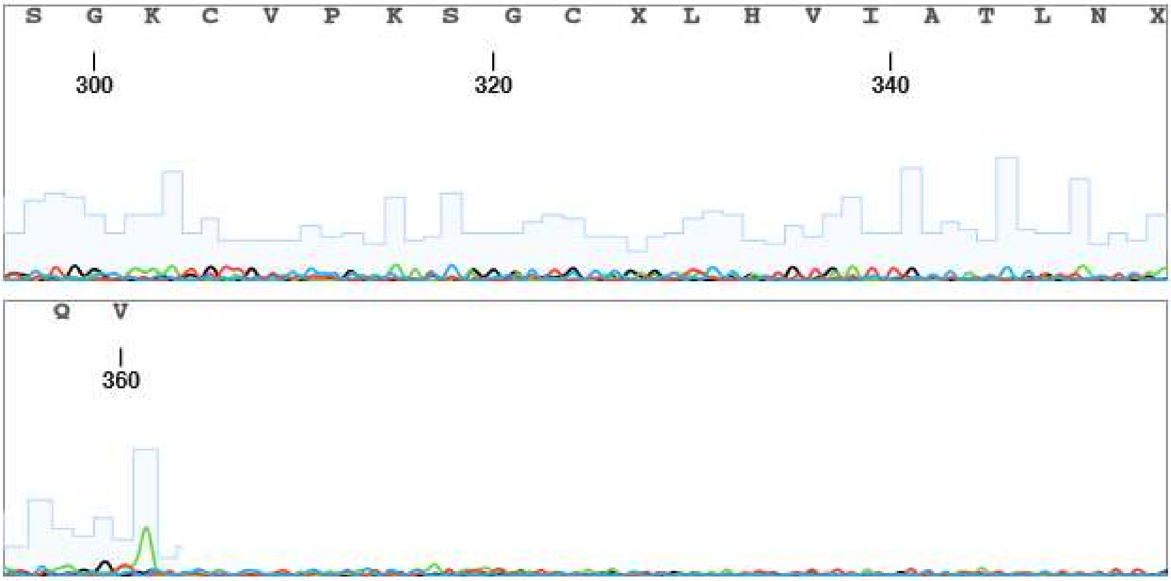


451863801_pNhaAH_M13R_B09.ab1 451863801_pNhaAH_M13R (20/06/2019) Page 2 / 6



451863801_pNhaAH_M13R_B09.ab1 451863801_pNhaAH_M13R (20/06/2019) Page 1 / 6



451863801_pNhaAH_M13R_B09.ab1 451863801_pNhaAH_M13R (20/06/2019) Page 2 / 6