



Metabolic engineering of clostridium acetobutylicum for the production of fuels and chemicals

Ngoc Phuong Thao Nguyen

► To cite this version:

Ngoc Phuong Thao Nguyen. Metabolic engineering of clostridium acetobutylicum for the production of fuels and chemicals. Microbiology and Parasitology. INSA de Toulouse, 2016. English. NNT : 2016ISAT0051 . tel-02083693

HAL Id: tel-02083693

<https://theses.hal.science/tel-02083693>

Submitted on 29 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

**En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**

Délivré par :

Institut National des Sciences Appliquées de Toulouse (INSA Toulouse)

Discipline ou spécialité :

Ingénieries microbienne et enzymatique

Présentée et soutenue par :

Ngoc-Phuong-Thao NGUYEN

le : 21/07/2016

Titre :

Metabolic Engineering of *Clostridium acetobutylicum* for the Production of
Biofuels and Chemicals

JURY

José Luis GARCÍA LÓPEZ - Professor, CSISC, Madrid

Nigel MINTON - Professor, University of Nottingham, Nottingham

Amel LATIFI - Professor, Université Aix Marseille, Marseille

Isabelle MEYNIAL-SALLES - Maître de Conférences, HDR, INSA, Toulouse

Ecole doctorale :

Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :

Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés

Directeur(s) de Thèse :

Philippe SOUCAILLE

Rapporteurs :

José Luis GARCÍA LÓPEZ et Nigel MINTON

Metabolic engineering
of *Clostridium acetobutylicum*
for the production of fuels and chemicals

**A Thesis Submitted to INSA in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy**

By

Ngoc-Phuong-Thao NGUYEN

July 2016

Examining Committee

José Luis GARCÍA LÓPEZ Professor, CSISC, Madrid

Nigel MINTON Professor, University of Nottingham, Nottingham

Amel LATIFI Professor, Université Aix Marseille, Marseille

Isabelle MEYNIAL-SALLES Maître de Conférences, HDR, INSA, Toulouse

Philippe SOUCAILLE Professor, INSA, Toulouse

“I do not know what I may appear to the world, but to myself I seem to have been only like a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay undiscovered before me”

Memoirs of the Life, Writings, and Discoveries of Sir Isaac Newton (1855)

Abstract

Currently, there is a resurgence of interest in *Clostridium acetobutylicum*, the biocatalyst of the historical Weizmann process, to produce n-butanol for use both as a bulk chemical and as a renewable alternative transportation fuel. To develop a commercial process for the production of n-butanol via a metabolic engineering approach, it is necessary to develop genetic tools and to better characterize both the primary metabolism of *C. acetobutylicum* and its regulation. This thesis describes a method of homologous recombination by replicative plasmid to delete or introduce genes in *C. acetobutylicum*. This method was successfully used to delete genes, including *CAC1502*, *CAC3535*, *CAC2879* (*upp*), to generate *C. acetobutylicum* Δ *cac1502* Δ *upp* and *C. acetobutylicum* Δ *cac1502* Δ *upp* Δ *cac3535*. These strains are readily transformable without any previous plasmid methylation and can serve as hosts for a “marker-less” genetic exchange system. The strain *C. acetobutylicum* Δ *cac1502* Δ *upp* and this homologous recombination system have been used in further engineering experiments to obtain a strain that produces butanol with high selectivity and eliminates by-products. Many genes were successfully deleted, including *ldhA*, *ctfAB*, *ptb*, and *buk*. The *hbd* gene was simultaneously removed and replaced by the *hbd1* gene from *C. kluyveri* DSM555, which encodes an NADPH-dependent enzyme. The *thlA* gene was replaced by an optimized synthetic *atoB* gene from *Escherichia coli* K12 MG1655. The mutant also contains an inactivated Rex protein reported to be an important regulator that senses the NADH/NAD⁺ ratio and represses many genes involved in the central metabolism. The final mutant *C. acetobutylicum* Δ *cac1502* Δ *upp* Δ *ctfAB* Δ *ptb300* Δ *buk* Δ *rex* *hbd::hbd1* *thlA::atoB* (*C. acetobutylicum* CAB1060) was successfully generated. This final mutant produces mainly butanol, with ethanol and traces of acetate at a molar ratio of 7:1:1. This CAB1060 strain was subjected to a new continuous fermentation process using i) *in situ* extraction of alcohols by distillation under low pressure and ii) high cell density cultures to increase the titer, yield and productivity of n-butanol production to levels that have never been previously achieved in any organism.

A second homologous recombination method using non-replicative plasmid for marker less gene modification is also described in this thesis. This method allows the simultaneous

inactivation of two genes. It has been successfully used to construct a mutant unable to produce hydrogen and useful, as a platform strain, for further engineering of *C. acetobutylicum* to continuously produce bulk chemicals and fuels. This strain has been successfully engineered to continuously produce lactate from glucose and 1, 3 propanediol from glycerol at yields, titers and productivities that have never been achieved before in *C. acetobutylicum*.

As a contribution to the characterization of the regulatory network of *C. acetobutylicum*, a novel global regulatory protein Cap0037 has been discovered. *Cap0037* and *Cap0036* are arranged in one operon located upstream of the *adhE2* gene. These two proteins are considered to be acidogenic indicators based on their significant up-regulation in acidogenesis and strong down-regulation in solventogenesis and alcohologenesis. The resultant mutant showed significantly different metabolic fluxes patterns in acidogenic (producing mainly lactate, butyrate and butanol) and alcohologenic (producing mainly butyrate, acetate and lactate) chemostat cultures but not in solventogenic or batch cultures. A quantitative transcriptomic analysis was employed for a deeper understanding of the corresponding effect. The *Cap0037::int* mutant has somewhat similar transcriptomic profiles to the *Rex*, *Per* and *Fur* mutants. The differences and similarities are discussed in detail. Moreover, *Cap0037/Cap0036* was found to be self-regulated. The DNA binding sites of *Cap0037* in the promoters' regions of the *Cap0037/Cap0036*, *adc* and *sol* operons have been experimentally determined. Using these experimental results as references, the complete putative regulon of *Cap0037* has been constructed by RegPredict.

Keywords: metabolic engineering, *C. acetobutylicum*, butanol, regulatory network, *Cap0037*, homologous recombination

Résumé

À l'heure actuelle, il y a un regain d'intérêt pour *Clostridium acetobutylicum*, le biocatalyseur du procédé Weizmann historique, pour produire le n-butanol un produit chimique de commodité et un bio-carburant alternatif et renouvelable. Pour développer un procédé commercial de production de n-butanol via une approche d'ingénierie métabolique, il était nécessaire de développer des outils génétiques permettant d'introduire de nombreuses modifications et de mieux caractériser à la fois le métabolisme primaire de *C. acetobutylicum* et sa régulation. Ce mémoire de thèse décrit un procédé de recombinaison homologue, utilisant plasmide réplcatif, pour la délétion ou l'introduction de gènes chez *C. acetobutylicum* avec une élimination facile des marqueurs utilisés. Cette méthode a été utilisée avec succès pour déléter de nombreux gènes, comprenant CAC1502, CAC3535, CAC2879 (*upp*), afin de générer *C. acetobutylicum* Δ *cac1502* Δ *upp* et *C. acetobutylicum* Δ *cac1502* Δ *upp* Δ *cac3535*. Ces souches sont facilement transformables sans méthylation préalable des plasmides et peuvent servir d'hôtes pour le système d'échange génétique sans marqueur qui a été développé. La souche de *C. acetobutylicum* Δ *cac1502* Δ *upp* et ce système de recombinaison homologue ont été utilisés dans d'autres expériences d'ingénierie pour obtenir une souche produisant du n-butanol avec une sélectivité élevée et en éliminant la plupart des co-produits. De nombreux gènes ont été délétés avec succès, dont *ldhA*, *ctfAB*, *ptb* et *buk*. Le gène *hbd* a été simultanément délété et remplacé par le gène *hbd1* de *C. kluyveri* DSM555, qui code pour une enzyme NADPH-dépendante. Le gène *thlA* a été remplacé par *atoB* un gène synthétique optimisé provenant d'*Escherichia coli* K12 MG1655. Le mutant contient également une délétion du gène codant pour la protéine Rex qui est un important régulateur du métabolisme central qui en fonction du rapport NADH / NAD⁺ réprime de nombreux gènes impliqués dans le métabolisme central. Le mutant final, *C. acetobutylicum* Δ *cac1502* Δ *upp* Δ *ctfAB* Δ *ldhA* Δ *ptb300* Δ *buk* Δ *rex* *hbd* :: *thlA* *hbd1* :: *atoB* (*C. acetobutylicum* CAB1060) a été généré avec succès. Ce dernier mutant produit principalement du n-butanol, et de faibles quantités d'éthanol et d'acétate. Cette souche CAB1060 a été utilisée dans un nouveau procédé de fermentation continu qui utilise i) l'extraction *in situ* des alcools par distillation sous pression réduite et ii) des cultures à haute densité cellulaire (et ne faisant pas intervenir de procédé membranaire) pour atteindre des

titre, rendement et productivité en n-butanol qui n'ont jamais été obtenus chez aucun micro-organisme.

Un second procédé de recombinaison homologue utilisant un plasmide non réplcatif pour la modification de gène sans marqueur est également décrit dans le présent mémoire. Cette méthode permet d'inactiver simultanément deux gènes. Il a été utilisé avec succès pour la construction d'un mutant incapable de produire de l'hydrogène et utile, comme souche plateforme, pour l'ingénierie de *C. acetobutylicum* pour produire en continu des produits chimiques de commodité et des bio-carburants. Cette souche a été modifiée avec succès pour produire en continu du lactate à partir du glucose et de 1, 3 propanediol à partir du glycérol à des rendements, des titres et des productivités qui n'ont jamais été atteints auparavant chez *C. acetobutylicum*.

En tant que contribution à la caractérisation des réseaux de régulation chez *C. acetobutylicum*, une nouvelle protéine régulatrice Cap0037 a été découverte. Les gènes *Cap0037* et *Cap0036* sont organisés en opéron et sont situés en amont du gène *adhE2*. Les deux protéines sont considérées comme des indicateurs d'acidogénèse car leur niveaux sont élevés en acidogénèse et fortement diminués en solvantogénèse et en alcoologénèse. Le mutant inactivé pour *Cap0037* a montré, lors de cultures en chimostat, des flux métaboliques significativement différents de la souche sauvage en acidogénèse (produisant principalement du lactate, du butyrate et du butanol) et en alcoologénèse (produisant principalement du butyrate, de l'acétate et du lactate), mais pas pour des cultures en solvantogénèse. Une analyse transcriptomique quantitative a été utilisée pour une meilleure compréhension des effets observés. Le mutant *Cap0037::int* a des profils transcriptomiques assez semblables aux mutants Rex, Per et Fur. Les différences et les similitudes sont discutés en détail dans le manuscrit. En outre, il a été montré que l'opéron *Cap0037 / Cap0036* était autorégulé par Cap0037. Les sites de fixation à l'ADN de Cap0037 dans les régions promotrices de *Cap0037 / Cap0036*, *adc* et de l'opéron *sol* ont été déterminées expérimentalement. À partir de ces résultats expérimentaux, le régulon putatif complet de Cap0037 a été construit par RegPredict.

Mots clés: metabolic engineering, *C. acetobutylicum*, butanol, regulatory network, Cap0037, homologous recombination

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor, Professor Philippe Soucaille, who gave me the chance six years ago so that today I am writing this Ph.D. dissertation. More important than knowledge and technical skills, he taught me to be patient, optimistic and consistent with the target. And I would like to thank him for his great support along the journey.

I would like to thank my fellow lab mates, especially CriCri (Christian Croux), without his support, I couldn't finish my Ph.D. by now. I want to thank Celine Foulquier, Minyeong Yoo, July (Julie Soula), Dr Charles Gauquelin and other current and former colleagues for their encouragement, their help, and great discussions even after work. "I owe you, guys!" And I thank Isabelle Meynial-Salles for her support and management during the years.

I would like to thank other collaborators from Metabolic Explorer, INSA, Ulm University and ClostNet program for their contribution to many parts of this dissertation.

Thank you, my Vietnamese friends, for sharing with me my good and bad time.

And most importantly, I am grateful to my family and my boyfriend, who always stay beside me and sacrifice for my sake. I finally can say, "Dear Daddy, the certificate is for you" and to my boyfriend "my dear, I am coming home!"

Table of contents

Abstract.....	3
Résumé.....	5
Acknowledgements	7
Table of contents	8
List of tables	14
List of figures.....	15
1 Chapter 1 Background	18
1.1 The history of acetone-butanol-ethanol (ABE) fermentation	18
1.2 <i>Clostridium acetobutylicum</i> strain	21
1.3 The central metabolism of <i>Clostridium acetobutylicum</i>	22
1.3.1 Glycolysis.....	24
1.3.2 Enzymes involved in electron fluxes.....	24
1.3.3 Central pathway from acetyl-CoA to butyryl-CoA	26
1.3.4 Acetate formation	28
1.3.5 Butyrate formation.....	29
1.3.6 Lactate and acetoin formation	30
1.3.7 Acetone formation.....	30
1.3.8 Ethanol and butanol formation	31
1.4 Metabolic engineering tools in <i>Clostridium acetobutylicum</i>	34
1.4.1 Shuttle vectors for gene over-expression.....	34
1.4.2 Inducible promoter/repressor systems	34
1.4.3 Reporter systems	35
1.4.4 Antisense RNA (asRNA)	35
1.4.5 Group II intron.....	36
1.4.6 Genomic mutagenesis by homologous recombination system	37

1.4.7	Random mutagenesis by mariner transposon system	41
1.5	<i>Clostridium acetobutylicum</i> as an n-butanol producer-a challenging task	42
1.6	Genome scale metabolic (GSM) models for <i>Clostridium acetobutylicum</i>	48
1.7	Regulatory network in <i>Clostridium acetobutylicum</i>	48
1.7.1	Spo0A and Sigma factors	48
1.7.2	The controversial SolR	49
1.7.3	Agr quorum sensing	49
1.7.4	Heat-shock proteins: DnaK, GroESL, and Hsp21	50
1.7.5	PhoPR	50
1.7.6	AbrB310	51
1.7.7	Per and Fur	51
1.7.8	CsrA	52
1.7.9	CcpA	52
1.7.10	Rex	52
1.7.11	AraR	53
1.7.12	Small regulatory RNA	53
2	Chapter 2 Thesis Objectives.....	55
3	Chapter 3 Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer <i>Clostridium acetobutylicum</i>.....	56
3.1	Abstract	57
3.2	Introduction	58
3.3	Results and Discussion	59
3.3.1	MGCΔ <i>cac1502</i> strain, a <i>C. acetobutylicum</i> strain that is transformable without previous in vivo plasmid methylation	59
3.3.2	Construction of the MGCΔ <i>cac1502</i> Δ <i>upp</i> strain: the first marker-free <i>C. acetobutylicum</i> strain with two deleted genes	63

3.3.3	Deletion of the <i>CA_C3535</i> gene in the <i>MGCΔcac1502Δupp</i> strain using the <i>upp/5-FU</i> system as a counter-selectable marker for the loss of plasmid	64
3.3.4	Determination of the recognition sequence of <i>Cac824II</i> encoded by <i>CA_C3535</i>	66
3.3.5	Deletion of the <i>ctfAB</i> genes in the <i>MGCΔcac1502ΔuppΔcac3535</i> to create a strain no longer producing acetone	67
3.3.6	Deletion of the <i>ldhA</i> gene in the <i>MGCΔcac1502ΔuppΔcac3535</i> to create a strain no longer producing lactate	68
3.4	Discussion	69
3.5	Conclusion	71
3.6	Experimental Procedure	71
3.6.1	Bacterial strain, plasmids and oligonucleotides	71
3.6.2	Culture and growth conditions	76
3.6.3	DNA manipulation techniques	76
3.6.4	Construction of pUC18-FRT-MLS2	76
3.6.5	Construction of pCons2.1	77
3.6.6	Construction of pCIP2-1	77
3.6.7	Construction of pREPcac15	77
3.6.8	Construction of pCIPcac15	78
3.6.9	Construction of pREPupp	78
3.6.10	Construction of pCLF1	78
3.6.11	Construction of pCons::upp	79
3.6.12	Construction of pREPcac35::upp	79
3.6.13	Construction of pREPctfAB::upp	79
3.6.14	Construction of pREPldhA::upp	80
3.6.15	Construction of pCLF::upp	80
4	Chapter 4 The Weizmann process revisited for the continuous production of n-butanol	82
4.1	Abstract	83

4.2 Main text	87
4.3 Supplementary Materials and Methods.....	95
4.3.1 Bacterial strains, plasmids and primers	95
4.3.2 Chemostat culture of recombinant strains	98
4.3.3 Continuous extractive high density cell recycle fermentation.....	98
4.3.4 Measurement of fermentation parameters	99
4.3.5 General protocol for gene deletion by homologous recombination system using replicative plasmid pSOS95-MLS ^r [167,174]	99
4.3.6 Construction of pSOS95-MLS ^r -upp-Drex-catP and <i>CABI058</i> strain	100
4.3.7 Construction of pEryupp-atoB plasmid and <i>CABI059</i> strain	100
4.3.8 Construction of pSOS95-upp-hbd1-catP-oriRepA and <i>CABI060</i> strain.....	101
4.3.9 Enzyme activity measurements	102
5 Chapter 5 Construction of a hydrogenase minus mutant of <i>Clostridium acetobutylicum</i> : a platform strain for the continuous production of fuels and chemicals	104
5.1 Main text	105
5.2 Conclusions	113
5.3 Material and Methods	113
5.3.1 Knockout of <i>hydA</i> and <i>thlA</i> using LtrB intron.....	113
5.3.2 Construction and integration by single crossing over of the pEryUppAthlA plasmid in the chromosome of <i>C. acetobutylicum</i> ΔuppΔcac1502	115
5.3.3 Construction and integration by single crossing over of the pCatUppHydA270 plasmid in the chromosome of <i>C. acetobutylicum</i> ΔuppΔcac1502	116
5.3.4 Simultaneous inactivation of <i>thlA</i> and <i>hydA</i> in <i>C. acetobutylicum</i> ΔuppΔcac1502	116
6 Chapter 6 Cap0037, a novel global regulator of <i>Clostridium acetobutylicum</i> metabolism	120
6.1 Abstract	121

6.2	Introduction.....	122
6.3	Results and Discussion	124
6.3.1	The phylogenetic tree and bioinformatics analysis of Cap0037/Cap0036	124
6.3.2	Disruption of CA_P0037 by pCUI-cap37 (189s)	126
6.3.3	Carbon and electron fluxes of the Cap37::intron mutant under different physiological conditions.....	128
6.3.4	Determination of the CA_P0037 DNA binding site (BS)	138
6.3.5	Determination of putative Cap0037 regulon	145
6.3.6	Global transcription changes in the Cap37::int mutant	151
6.4	Conclusion	172
6.5	Experimental Procedure.....	172
6.5.1	Bacterial strains, plasmids and culture media	172
6.5.2	Plasmid construction	172
6.5.3	Continuous culture.....	174
6.5.4	Isolation of total mRNA and microarray	174
6.5.5	Analytical methods.....	174
6.5.6	Southern blot analysis.....	175
6.5.7	Expression and purification of His-tagged protein.....	175
6.5.8	Electromobility shift assays (EMSAs)	175
6.5.9	DNase I protection assays (DNA Footprinting)	176
6.5.10	Bioinformatic tools.....	177
6.5.11	Microarray data accession number	178
6.6	Acknowledgement	178
6.7	Supplementary	178
7	Chapter 7 Conclusion and future perspectives	191
	Abbreviations	193
	References.....	197

Publications 215

List of tables

Table 1-1 Butanol production from lignocellulosic biomass	20
Table 1-2 Kinetic parameters of <i>C. acetobutylicum</i> Hbd [37].....	27
Table 1-3 Activities of key enzymes catalyzing butanol formation [42]	32
Table 1-4 <i>C. acetobutylicum</i> mutants and changes in product profiles [86]	46
Table 1-5 Genome scale metabolic models of <i>C. acetobutylicum</i> ATCC 824.....	48
Table 3-1 Transformation efficiencies of <i>C. acetobutylicum</i> ATCC824 and MGCΔ <i>cac1502</i> for unmethylated and methylated pCons2.1.....	63
Table 3-2 Transformation efficiencies of MGCΔ <i>cac1502</i> and MGCΔ <i>cac1502</i> Δ <i>upp</i> Δ <i>cac3535</i> for unmethylated pCons2.1.....	65
Table 3-3 Bacterial strains and plasmids used in this study	73
Table 3-4 Oligonucleotides used for PCR amplifications.	76
Table 4-1 List of strains	95
Table 4-2 List of plasmids	96
Table 4-3 List of primers	97
Table 5-1 Prediction of bio-chemical yield produced by a hydrogenase minus mutant	105
Table 5-2 List of primers	118
Table 6-1 Prediction of putative Cap0037 regulon and the correspondent relative transcript levels of those genes of Cap0037::int mutant in the three metabolic states: acidogenesis (AC), alcohologenesis (AL), and solventogenesis (SO).....	148
Table 6-2 Expression profiles of Rex regulon in the three metabolic states: acidogenesis (AC), alcohologenesis (AL) and solventogenesis (SO)	166
Table 6-3 Transcriptional profiles of selected genes in Cap0037::int mutant in the three metabolic states compared with those of oxygen exposed WT, Δ <i>PerR</i> mutant, iron limited WT and <i>Fur</i> ::int mutant	169

List of figures

Figure 1-1 Metabolic pathways of <i>C. acetobutylicum</i> . Enzymes involved in the fermentative pathways are shown in blue.	23
Figure 1-2 Electron flow in <i>C. acetobutylicum</i> [30,31].....	25
Figure 1-3 Proposed mechanism of endergonic ferredoxin reduction with NADH coupled to exergonic crotonyl-CoA reduction catalyzed by Bcd/EtfAB complex of <i>C. kluyveri</i> [43] .	28
Figure 1-4 Butanol pathway analysis under acidogenesis (A), solventogenesis (B) and alcohologenesis (C) [42]	33
Figure 1-5 Putative cross-over events during <i>spo0A</i> gene inactivation [92].....	38
Figure 1-6 DNA integration at the <i>thl</i> locus of <i>C. acetobutylicum</i> [93].	39
Figure 1-7 I- <i>SceI</i> based-homologous recombination system [95]	40
Figure 1-8 Single-crossover homologous recombination [97]	41
Figure 3-1 General diagram representing gene replacement via allelic exchange at the <i>Y</i> locus, and excision of the MLS^R marker by the FLP recombinase to create an unmarked <i>Y</i> deletion mutant.	61
Figure 3-2 Gene replacement via allelic exchange at the <i>ca_c1502</i> , <i>upp</i> and <i>ca_c3535</i> loci.	62
Figure 3-3 Growth curves of <i>C. acetobutylicum</i> ATCC824 and <i>MGCΔcac1502 Δupp Δcac3535</i> at pH 4.5 in SM medium	65
Figure 3-4 Digestion properties of recombinant <i>Cac824II</i> as compared to commercial <i>AcuI</i> (New England Biolabs).....	67
Figure 3-5 Solvent and acid production of <i>MGCΔcac1502 Δupp Δcac3535</i> , <i>MGCΔcac1502 Δupp Δcac3535Δ ctfAB::Em^R</i> , and <i>MGCΔcac1502 Δupp Δcac3535Δ ldhA::Em^R</i> in batch culture at pH 4.5 in SM medium	69
Figure 4-1 Optimization of <i>n</i> -butanol production in <i>Clostridium acetobutylicum</i>	84
Figure 4-2 Schematic representation of the extractive continuous fermentation process	86
Figure 4-3 Product ratios in engineered <i>C. acetobutylicum</i> strains and wild-type	90

Figure 4-4 Butanol, hydrogen and electron yields in the engineered <i>C. acetobutylicum</i> and the wild-type	91
Figure 4-5 Enzymatic assays on crude extracts of <i>C. acetobutylicum</i> mutants.....	92
Figure 4-6 Chemostat culture of <i>C. acetobutylicum</i> CAB1060.....	93
Figure 4-7 Continuous high cell density bioreactor with <i>in situ</i> extraction of the alcohols (n-butanol + ethanol) by distillation under low pressure	94
Figure 5-1 Construction of <i>hydA::int(s)</i> and <i>thlA ::int(s)</i> mutants.....	107
Figure 5-2 Schematic representation of the different steps for the simultaneous inactivation of <i>hydA</i> and <i>thlA</i>	108
Figure 5-3 PCR analysis of intermediate (with plasmids integrated in 3') and the final hydrogenase mutant <i>C. acetobutylicum</i> $\Delta upp\Delta cac1502 \Delta thlA\Delta hydA270$	110
Figure 5-4 Metabolism of <i>C. acetobutylicum</i> $\Delta upp\Delta cac1502 \Delta thlA\Delta hydA270$	111
Figure 5-5 Continuous production of lactate by the <i>C. acetobutylicum</i> $\Delta upp\Delta cac1502 \Delta thlA\Delta hydA270\Delta pSOL$	112
Figure 5-6 Continuous production of 1,3-propanediol by the <i>C. acetobutylicum</i> $\Delta upp\Delta cac1502 \Delta thlA\Delta hydA270 \Delta pSOL (pSPD5)$	113
Figure 6-1 Phylogenetic trees of Cap0036 (A), Cap0037 (B) sequences from <i>Clostridium acetobutylicum</i> and neighbor proteins from other bacteria.	125
Figure 6-2 Construction of Cap37:: <i>int</i> mutant.....	127
Figure 6-3 Metabolic fluxes of mutant Cap0037:: <i>int</i> versus control strain in different metabolic states (A) Acidogenesis (B) Alcohologenesis (C) Solventogenesis.....	134
Figure 6-4 Electron fluxes of mutant Cap0037:: <i>int</i> versus control strain in different metabolic states (A) Acidogenesis (B) Alcohologenesis (C) Solventogenesis.....	135
Figure 6-5 Relative transcript levels of central metabolic genes of Cap0037:: <i>int</i> mutant in acidogenesis, alcohologenesis and solventogenesis conditions.	136
Figure 6-6 The butanol pathway in the Cap0037:: <i>int</i> mutant (MT) and the wild-type (WT) in acidogenesis (A), alcohologenesis (B) and solventogenesis (C).....	137

Figure 6-7 Analysis of Cap0037 binding to the promoter region of the <i>CA_P0037/CA_P0036</i> operon.....	140
Figure 6-8 EMSAs using the promoter region of <i>adc</i> , <i>sol</i> operon and Cap0037 protein.	141
Figure 6-9 DNA footprinting of Cap0037 interacting with <i>adc</i> and <i>sol</i> promoter.....	142
Figure 6-10 DNA sequence analysis of the promoter regions of (A), <i>adc</i> , and (B) <i>sol</i> operon.....	144
Figure 6-11 Putative Cap0037 binding sequence. (A) alignment and (B) sequence logo showing frequencies of residues.....	147
Figure 6-12 Statistical analysis of transcriptomics of <i>C. acetobutylicum</i> Cap0037::int mutant in acidogenesis, alcohologenesis and solventogenesis.....	153
Figure 6-13 Expression profiles of genes involved in Arginine (A), Histidine and Tryptophan (B) and Cystein (C) transport and metabolism in acidogenesis, alcohologenesis and solventogenesis.....	158
Figure 6-14 Expression profiles of genes involved in nucleic acid metabolism in acidogenesis, alcohologenesis and solventogenesis.....	161
Figure 6-15 Expression profiles of genes involved in cation transport, glycerol lipid transport and metabolism in acidogenesis, alcohologenesis and solventogenesis.....	162
Figure 6-16 Expression profiles of genes involved in sporulation in acidogenesis, alcohologenesis and solventogenesis state.....	164

Chapter 1 Background

1.1 The history of acetone-butanol-ethanol (ABE) fermentation

In the early 19th century, the scarcity of natural rubber raised interest in synthetic rubber production. The best source for the production of butadiene was butanol or isoamyl alcohol. In 1912, Weizmann stated that butanol or isoamyl alcohol was critical in the production of synthetic rubber. Between 1912 and 1914, he succeeded in the isolation of a strain called BY [1], later named *Clostridium acetobutylicum*. The BY strain was able to produce good yields of butanol and acetone (AB fermentation). AB fermentation by *C. acetobutylicum* is also more recently termed acetone-butanol-ethanol (ABE) fermentation.

During World War I, acetone became a chemical in high demand as a solvent for nitrocellulose in the manufacture of cordite for the British army. Butanol became an unwanted by-product of the fermentation [2]. When the war ended, the demand for acetone decreased. Butanol and its ester, butyl acetate, were found to be ideal solvents for the production of a quick drying lacquer in the automobile industry, which was expanding rapidly. From 1920 to 1950, AB fermentation plants were constructed in the USA, Australia, Canada, China, India, Japan, and many other countries. Great efforts were made to find substitute substrates to optimize AB production, such as molasses, grain, and corncob hydrolysate. The metabolic pathways of acetone and butanol synthesis were of great interest, but even today, these pathways are not completely understood [2].

In the early 1950s, ABE fermentation declined for two reasons. First, the petrochemical industry grew rapidly, and the fermentation process could not compete with the chemical process in production yields or cost. Secondly, molasses began to be used to feed cattle, increasing the feedstock cost [2].

Now, with the negative impacts of fossil fuels on the climate and its exhaust, biofuels are the key to future clean energy. Thus, ABE fermentation is now attracting renewed interest, especially for n-butanol. Many research groups have studied fermentation with *C. acetobutylicum* and other Clostridial strains to produce cellulosic butanol from abundant agricultural feedstocks, which is a promising strategy for producing biofuels at a low cost (examples in Table 1-1). Others have worked extensively on the cellular functioning and metabolic engineering of *C. acetobutylicum* to further optimize the strain.

Strain	Feedstock	Butanol Yield		Butanol Productivity (g/l/h)	Condition	References
		g/g	g/L			
<i>C. acetobutylicum</i> KF 158795	Glycerol	0.35	13.6	0.14	Batch, 30 L, optimized	[3]
<i>C. acetobutylicum</i> ATCC 824	Sweet sorghum juice	0.40	6.5	0.07	Batch	[4]
<i>C. acetobutylicum</i> GX01	Cassava starch, soybean meal		18.8		Batch, 30 L	[5]
<i>C. acetobutylicum</i> TSH1 and <i>B. cereus</i> TSH2	Glucose	0.18	11.0		Micro aerobic, batch	[6]
<i>C. celevecrescens</i> N3 and <i>C. acetobutylicum</i> ATCC 824	Cellulose	0.15	3.7		Batch	[7]
<i>C. acetobutylicum</i> GX01	Alkali pretreated sugarcane bagasse	0.22	14.2	0.24	Fed-batch	[8]
<i>C. acetobutylicum</i> ABE 1201	Sweet sorghum juice	0.24	13.9	0.18	Immobilized cells, gas stripping	[9]
<i>C. beijerinckii</i> B592	Pinewood hydrolysate		11.6	0.19	Batch	[10]

<i>C. cellulolyticum</i> pM9	Crystalline cellulose	0.18		Batch	[11]
<i>C. sporogenes</i> BE01	Pretreated rice straw	4.46	0.05	Batch	[12]
<i>C. beijerinckii</i> TISTR 1461	Pretreated corncob	6.26		Batch	[13]

Table 1-1 Butanol production from lignocellulosic biomass

1.2 *Clostridium acetobutylicum* strain

The *Clostridium* genus belongs to the phylum *Firmicutes* and contains approximately 100 species identified by 16S rRNA analysis [14]. It is a diverse group of Gram-positive, rod shaped anaerobes. The *Clostridium* genus can be divided into pathogens and non-pathogens. Some well-known pathogens include *C. botulinum*, *difficile*, *tetani*, and *perfringens*, causing botulism, diarrhea, tetanus and gas gangrene, respectively. Extensive research is in progress addressing the impacts of those diseases. The genus *Clostridium* also includes organisms of importance to detoxification, remediation, and the carbon cycle, such as acetogens, which use the Wood-Ljungdahl pathway to generate huge quantities of acetate through fixing 20% of the CO₂ on earth [15]. A large number of *Clostridium* strains have potential for industrial applications. For example, *C. thermocellum* can utilize cellulose as its sole carbon source to produce ethanol [16]; *C. kluyveri* is under study to produce polyester poly 4-hydroxybutyric acid [17]; and *C. acetobutylicum* have been using for the production of ABE.

The strain *C. acetobutylicum* ATCC 824 is one of the best studied solventogenic clostridia. It was isolated in 1954 from the soil of a garden in Connecticut [18]. The ATCC 824 strain was found to be closely related to the historical Weizmann strain [19]. The genome of the strain was mapped in the early 1990s. The strain has a 4 Mb chromosome with 11 ribosomal operons and harbors a megaplasmid, pSOL1, approximately 200 kb in size that carries the genes involved in solvent formation [19]. The construction of a genomic map was expected to trigger the development of engineering tools to optimize the strain and has been useful for better understanding of the degeneration mechanism (i.e., loss of the strain's ability to produce solvent). *C. acetobutylicum* DSM792 is another widely used strain that is identical to ATCC 824. *C. acetobutylicum* can consume various types of monosaccharides, disaccharides, starch, inulin, pectin, whey, and xylan but not crystalline cellulose [19].

The growth of *C. acetobutylicum* in a batch culture occurs in three different growth phases. The exponential growth phase produces biomass, acids, ATP, and hydrogen. The accumulation of acids decreases the medium pH to a threshold that is toxic to the cells. Thus, cell growth is inhibited, and the cells enter a stationary phase. Solventogenesis is triggered in the transition to the stationary phase. Acids are partially re-assimilated, and ABE solvents are produced at a ratio of 3:6:1. At the end of the stationary phase, the *C. acetobutylicum* cells sporulate.

C. acetobutylicum can also be cultivated in continuous cultures and, under phosphate limitation, maintained in stable metabolic states [20–24]:

- Acidogenesis: when grown at neutral pH on glucose.
- Solventogenesis: when grown at low pH on glucose.
- Alcohologenesis (production of ethanol and butanol, but not acetone): when grown at neutral pH under conditions of high NAD(P)H availability [20]. There are multiple ways to achieve high intracellular NAD(P)H: cultivating the culture either under iron limitation or by sparging in carbon monoxide (to lower the electron flow towards hydrogen production), the addition of methyl viologen or neutral red and the replacement of part of the glucose with glycerol [20–24].

Steady-state continuous cultures provide better homogenous cells, which are useful for quantitative fluxomics, transcriptomics and proteomics research.

1.3 The central metabolism of *Clostridium acetobutylicum*

The primary pathway leading to the formation of acids and solvents is shown in Figure 1-1.

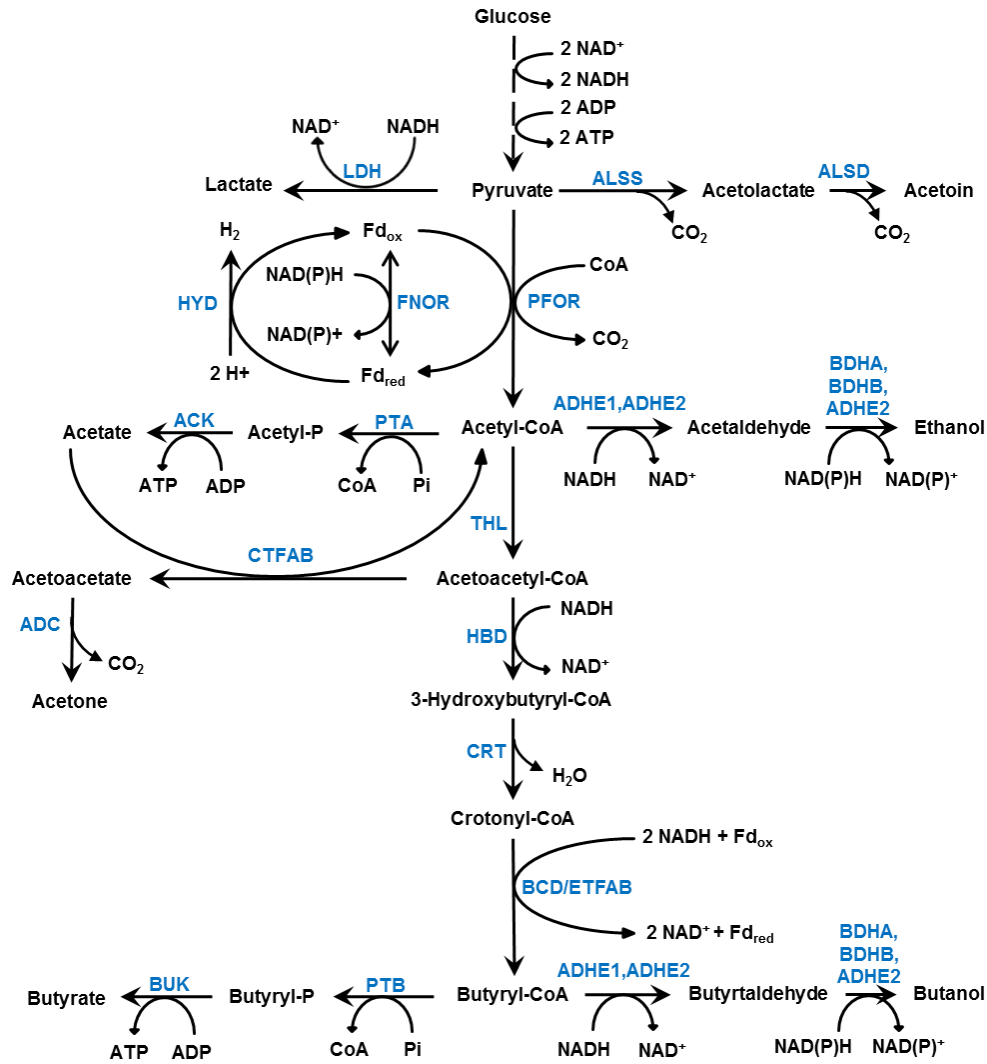


Figure 1-1 Metabolic pathways of *C. acetobutylicum*. Enzymes involved in the fermentative pathways are shown in blue.

LDH: lactate dehydrogenase (CAC0267), ALSS: acetolactate synthase (CAC3652), ALSD: acetolactate decarboxylase (CAC2967), PFOR: pyruvate ferredoxin oxidoreductase (CAC2229), HYD: hydrogenase (CAC0028), FNOR: NAD(P) ferredoxin oxidoreductase, ACK: acetate kinase (CAC1743), PTA: phosphotransacetylase (CAC1742), BUK: butyrate kinase (CAC3075), PTB: phosphotransbutyrylase (CAC3076), THL: thiolase (CAC2873), HBD: hydroxybutyryl-coA dehydrogenase (CAC2708), CRT: crotonase (CAC2712), BCD: butyryl coA dehydrogenase (CAC2711), ETFAB: electron transfer flavoprotein subunit A/B (CAC2709/2710), ADHE1: aldehyde alcohol dehydrogenase (CAP0162), ADHE2 (CAP0035), BDHA/B: butanol dehydrogenase (CAC3298/3299), ADC: acetoacetate decarboxylase (CAP0165), CTFA/B: coA-transferase subunit A/B (CAP0163/0164).

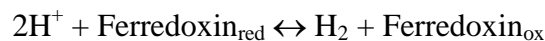
1.3.1 Glycolysis

Hexose sugars such as glucose, sucrose and maltose can be metabolized completely via the Embden-Meyerhof glycolytic pathway. Pentose sugars such as xylose, ribose and arabinose are metabolized via the pentose phosphate pathway. Glucose is the most preferred hexose [25]. EMP produces 2 moles of pyruvate, 2 moles of ATP and 2 moles of NADH per mole of hexose.

Pyruvate ferredoxin oxidoreductase (PFOR-CAC2229) catalyzes the decarboxylation of pyruvate to acetyl-CoA. Pfor from *C. acetobutylicum* was purified and characterized by Meinecke B et al. [26]. The purified protein possessed a specific activity of 25 U/mg. The monomer molecular weight was 123 kDa. One mole of enzyme monomer contains 0.39 mole of thiamine pyrophosphate, 4.23 moles of iron and 0.91 moles of sulfur [26]. The pH optimum was at pH 7.5, and the temperature optimum was at 60°C. The apparent K_m values for pyruvate and coenzyme A were 322 μ M and 3.7 μ M, respectively [26]. The reaction catalyzed by Pfor in glycolysis reduces oxidized ferredoxin and releases CO₂. The re-oxidation of reduced ferredoxin occurs through the activity of hydrogenase (encoded by *hydA*, CA_C0028) in acidogenesis and of NAD(P) ferredoxin oxidoreductase in solventogenesis.

1.3.2 Enzymes involved in electron fluxes

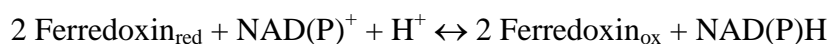
Hydrogenase (HYD, EC 1.12.7.1) transfers electrons from electron donors, such as ferredoxin and flavodoxin, to protons.



There are two Fe-Fe hydrogenases (HydA-CAC0028 and HydB-CAC3230) and one Ni-Fe hydrogenase in *C. acetobutylicum* [27]. HydA fused with Strep-tag was homologously expressed and characterized [28]. For hydrogen evolution activity, both k_{cat} and k_{cat}/K_m were higher for reduced ferredoxin than for reduced methyl viologen, 5 and 250 times, respectively. When using flavodoxin (encoded by CAC0587) to replace ferredoxin (CAC0303) as an electron donor, the catalytic efficiency k_{cat}/K_m was $5.5 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ relatively higher compared to $4 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ for ferredoxin [27]. For hydrogen uptake activity, the catalytic efficiency (k_{cat}/K_m) was measured 100 fold higher for ferredoxin compared to methyl viologen as electron acceptor [27]. Maturation of hydrogenase requires the biosynthesis and insertion of the catalytic iron-sulfur cluster, the H-cluster. The maturation proteins HydEF and HydG were coexpressed with HydA or HydB in *E. coli* by King et al

[29]. HydB exhibits H₂ evolutionary activity 10 fold less than HydA. As a result, HydA is the main enzyme responsible for the regeneration of reduced ferredoxin in *C. acetobutylicum*.

Ferredoxin NAD(P)H oxidoreductase (FNOR, EC 1.18.1.3(2)) activities were measured in crude extracts of *C. acetobutylicum* cells at different states [21,30,31].



In the solventogenic state, in addition to hydrogenase, which regenerates reduced ferredoxin, there are also ferredoxin NAD⁺ reductase and ferredoxin NADP⁺ reductase, which also regenerate reduced ferredoxin coupling by regenerating NAD⁺/NADP⁺ to produce NADH/NADPH, which are required in the formation of butanol and ethanol. In the acidogenic state, ferredoxin NADH reductase activity was found to regenerate the NADH produced by EMP (Figure 1-2).

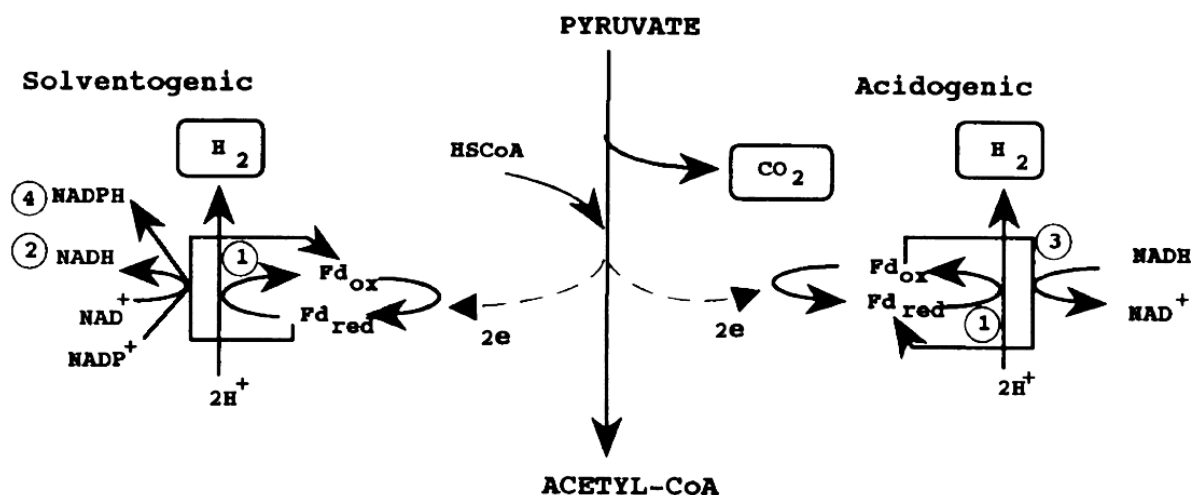


Figure 1-2 Electron flow in *C. acetobutylicum* [30,31]

1-Hydrogenase, 2-ferredoxin NAD⁺ reductase, 3-ferredoxin NADH reductase, 4-ferredoxin NADP⁺ reductase

1.3.3 Central pathway from acetyl-CoA to butyryl-CoA

Thiolase (THL, E.C. 2.3.1.19) is an important enzyme that condenses two acetyl-CoA (C2) into one acetoacetyl-CoA (C4). There are two thiolase encoding genes in *C. acetobutylicum*: *thlA* and *thlB* [32]. ThlA was demonstrated to be primarily involved in the production of butyric acid and in AB formation. The *thlB* gene was found to be poorly expressed in both acidogenic and solventogenic cells [32]. ThlA from *C. acetobutylicum* ATCC 824 was purified and characterized [33,34]. It possessed high activity in the physiological range of internal pH from 5.5 to 7.0. In the condensation direction, the thiolase (ThlA) was very sensitive regarding the CoA-SH, which was inhibited by micro molar levels of CoA-SH. ATP and butyryl-CoA were also shown to be thiolase's inhibitors. The native enzyme consisted of four subunits, each had a molecular weight of 44 kDa. The K_m value toward acetyl-CoA substrate was 0.27 mM, measured at 30°C and a pH of 7.4. The K_m value toward acetoacetyl-CoA substrate was 0.032 mM (at 30°C and pH 8.0), which is much lower than for acetyl-coA [34]. When thiolase (ThlA) was mutated at three positions, R133G, H156N, and G222V, it showed significantly less sensitivity to (50 μ M) CoA-SH [35]. When this mutated ThlA was expressed in *C. acetobutylicum*, there was a clear delay in ethanol and acetone formation. Additionally, the ethanol and butanol titers were increased by 46% and 18%, respectively, while the final acetone concentrations were not affected. The crystal structures of ThlA in its reduced/oxidized states were studied [36], and ThlA was found to be regulated by redox-switch modulation through a reversible disulfide bond between two catalytic cysteine residues, Cys88 and Cys378. The ThlA^{V77Q/N153Y/A286K} mutant, which cannot form disulfide bonds, exhibits higher activity than wild-type ThlA and enhances butanol production upon over-expression [36].

β -Hydroxybutyryl dehydrogenase (HBD, E.C.1.1.1.35), catalyzing acetoacetyl-CoA hydration to (S) β -hydroxybutyryl-CoA, was purified and characterized [37]. The reduction of acetoacetyl-CoA to (S) β -hydroxybutyryl-CoA occurred over a wide pH range, between 4 and 10, with pH_{opt} at 5. In contrast, the reverse reaction occurred only within a small pH range from 9.5 to 10.5, with a maximum at pH 10. Hbd specificities were examined and are reviewed in [37]. Hbd from *C. acetobutylicum* is NADH-dependent. Hbd retained 50% activity in the presence of 20% (v/v) ethanol, 9% (v/v) butanol and 10% (v/v) isobutanol [37]. When Hbd was inactivated, ethanol production increased, whereas C4 products such as butyrate and butanol were significantly reduced [38].

Hbd from *C. butyricum* (CbHbd) has better catalytic efficiency than Hbd from *C. acetobutylicum* (K_{cat}/K_m of 1.2×10^8 ($M^{-1}s^{-1}$), which is 343-fold higher) [39]. The crystal structure of *C. butyricum* Hbd was solved. The mutated *C. butyricum* Hbd^{K50A/K54A/L232Y} has a 4.9-fold increase in K_{cat}/K_m compared to the wild-type enzyme. Hbd1 from *C. kluyveri* is NADP-dependent and possesses K_m values of 50 μM for acetoacetyl-CoA and 70 μM for NADPH₂ [40].

	V_{max} (U/mg)	K_m (mM)	K_{cat} (s ⁻¹)	K_{cat}/K_m (M ⁻¹ s ⁻¹)
Acetoacetyl CoA	75.7	0.11	38.6	3.5×10^5
3-hydroxybutyryl CoA	53.5	0.29	27.1	9.4×10^4
NADH	61.5	0.14	31.4	2.3×10^5
NAD+	55.1	0.27	28.1	1.0×10^5
NADPH	3.31	0.45	1.7	3.8×10^3

Table 1-2 Kinetic parameters of *C. acetobutylicum* Hbd [37]

Crotonase (CRT, EC 4.2.1.150) reversibly converts (S)-3-hydroxybutyryl-CoA to crotonyl-CoA by eliminating a proton and a hydroxyl group, leading to the formation of a double bond at C2 and C3 of the enoyl moiety. The crotonase crystal structure was determined, showing that the protein forms a hexamer through the dimerization of two trimers. Phe143 and Phe233 are key residues in the binding pocket structure that accommodates crotonyl-CoA as a substrate [41].

Butyryl-CoA dehydrogenase enzyme (BCD, EC 1.3.99.2) catalyzes the reduction of crotonyl-CoA to butyryl-coA. BCD is encoded by *bcd-etfB-etfA* (CAC2711-CAC2710-CAC2709). The enzyme complex was homologously expressed and purified in *C. acetobutylicum* and was demonstrated to be a strictly NADH-dependent enzyme with ferredoxin required for the reaction [42]. Under fully coupled conditions, 1 mole of ferredoxin is reduced by 2 moles of NADH and 1 mole of crotonyl-CoA, which is similar to BCD of *C. kluyveri* [42,43].

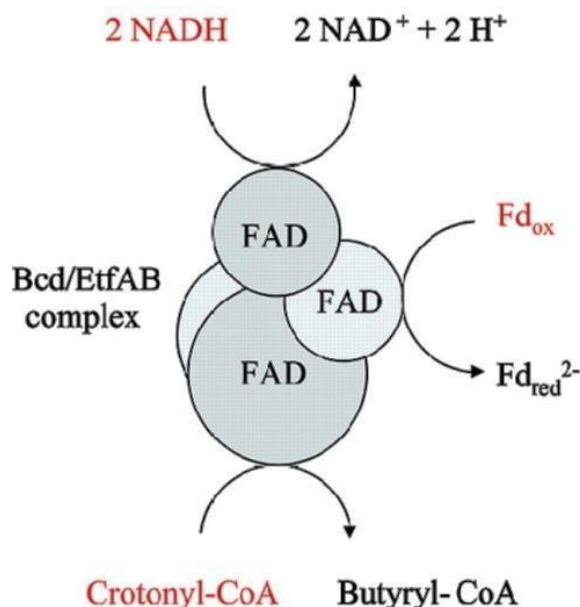


Figure 1-3 Proposed mechanism of endergonic ferredoxin reduction with NADH coupled to exergonic crotonyl-CoA reduction catalyzed by Bcd/EtfAB complex of *C. kluyveri* [43]

The five ORFs, ordered *crt-bcd-etfB-etfA-hbd*, were found to be transcribed as a transcriptional unit, and the promoter's consensus sequences were identified upstream of the *crt* gene [44]. When over-expressing any two of *crt*, *bcd*, and *hbd*, butanol production increased by at least two fold, resulting in an increase in the final butanol titer of 28% [45]. When the butanol pathway genes, including acetyl-CoA acetyltransferase (*thlA*), 3-hydroxybutyryl-CoA dehydrogenase (*hbd*), crotonase (*crt*), butyryl-CoA dehydrogenase (*bcd*, *etfA*, *etfB*) and butyraldehyde dehydrogenase (*adhe2*), were cloned and expressed in *E. coli* EG03(Δ *frdABCD* Δ *ldhA* Δ *ackA* Δ *pflB* Δ *adhE* Δ *pdhR::pflBp6-aceEF-lpd* Δ *mgsA*), butanol (1,254 mg/L) was produced from glucose under anaerobic conditions [46].

1.3.4 Acetate formation

During acidogenesis, acetyl-CoA is converted to acetate in a two-step process by a phosphotransacetylase (*pta*, CA_C1742) and an acetate kinase (*ack*, CA_C1743). Acetate kinase (ATP:phosphotransferase (ACK), EC 2.7.2.1) from *C. acetobutylicum* DSM 1731 was purified and characterized [47]. Ack protein exists in a dimeric form, and each subunit has a molecular weight of 43 kDa. Ack has high specificity for acetate and ATP. The maximal activity was obtained when Mn^{2+} was added to the reaction as a divalent cation. The K_m values for acetate, ATP, acetyl phosphate and ADP were 73, 0.37, 0.58 and 0.71 mM,

respectively. The *ack* and *pta* genes were found to be arranged in an operon. Primer extension analysis identified a single transcriptional start site located 70 bp upstream of the start codon of the *pta* gene. When *ack* and *pta* were over-expressed in *C. acetobutylicum*, the final ratios of acetate to other major products were higher, and there was a greater proportion of C2-derived compared to C4-derived products [48].

1.3.5 Butyrate formation

Butyryl-CoA is converted to butyrate by phosphotransbutyrylase (*ptb*, CA_C3076) and butyrate kinase (*buk*, CA_C3075). The *ptb* and *buk* genes are organized in an operon. The transcription start point was identified 57 bp upstream of the *ptb* start codon, and the putative Rho-independent terminator structure was identified 26 bp downstream from the *buk* stop codon [49,50].

Phosphotransbutyrylase (PTB, EC 2.3.1.19) catalyzes butyryl-CoA to butyryl-phosphate. Ptb of *C. acetobutylicum* was purified and characterized [50,51]. It was found to be very sensitive to pH shifts within the range of 5.5-7 and to be inhibited at pH below 6.0, suggesting a regulatory role of internal pH on Ptb catalytic activity [50,51]. In the reverse catalytic reaction, Ptb was less sensitive to pH variation [51]. The enzyme exhibited high activity toward butyryl-CoA, isovaleryl-CoA and valeryl-CoA [51]. The K_m values for butyryl-CoA, phosphate, butyryl phosphate, and CoASH were 0.11, 14, 0.26, and 0.077 mM, respectively, at 30°C and pH 8.0 [51].

Butyrate kinase (BUK, E.C. 2.7.2.7) catalyzes the reversible formation of butyrate from butyryl-P. Two *buk* genes have been characterized in *C. acetobutylicum* [52–54]. The primary butyrate kinase gene, *bukI*, is encoded in an operon with the *ptb* gene, while *bukII* is a single gene [54]. BukI is a dimer with subunit molecular weight of 39 kDa, with a specific activity of 402 UI/mg [53]. Disruption of the *bukI* gene reduced butyrate kinase, not all enzyme activity was eliminated [55]. BukI has a wide substrate range and binds butyrate very tightly. BukI can act on butyrate, valerate, isobutyrate, propionate, isovalerate, and vinyl acetate with highly specific activities. When binding butyrate, BukI has a K_m of 14 mM [56]. The second *bukII* was expressed, purified in *E. coli* and characterized [52]. The purified BukII has a subunit molecular weight of 43 kDa and functions as a dimer of 80 kDa [56]. BukII can act on a limited substrate range, with only the activity on butyrate retained among the substrates tested. BukII has an extremely lower binding affinity for butyrate (K_m of 620 mM) compared to BukI [54].

1.3.6 Lactate and acetoin formation

Lactate dehydrogenase (LDH, EC 1.1.1.27) catalyzes the reduction of pyruvate. The reaction is irreversible, NADH-dependent and activated specifically by fructose-1,6-bisphosphate [57,58]. The purified enzyme existed in tetramers of 159kDa, had optimum pH at 5.8 [57]. Expression of *C. acetobutylicum* Ldh in a ldh-deficient *E. coli* strain recovered the production of lactate [58].

Acetoin formation-related genes (*alsSD*, *alsR*) have been studied [59,60]. In many bacteria, acetoin production is controlled by AlsR, which binds and regulates the promoter of the *alsSD* operon. In *Bacillus subtilis*, when AlsR was disrupted, the *alsSD* operon was not expressed [61]. In *C. acetobutylicum*, the regulatory of acetoin production was different. The genes *alsS* and *alsD* do not share the same operon. The disruption of AlsR did not change either acetoin production or the expression of *alsS* and *alsD*. The over-expression of *alsS* did not improve acetoin formation [60]. However, Liu et al has successfully increased acetoin formation by over-expressing *alsD* in an *adc*-deficient mutant [59].

1.3.7 Acetone formation

Acetoacetyl-CoA is converted to acetone in a two-step process via the intermediary acetoacetate. The two steps require a CoA-transferase and an acetoacetate decarboxylase. The CoA-transferase consists of 2 different subunits encoded by *ctfA* (*CA_P0163*) and *ctfB* (*CA_P0164*) genes, which are part of the *sol* operon. The acetoacetate decarboxylase gene (*adc*, *CA_P0165*) is located outside the *sol* operon and is transcribed in the opposite direction by its own promoter [62,63]. These two operons (*adc* and *sol*) were found to be induced at the onset of solventogenesis [64].

CoA transferase (CTFAB, EC 2.3.8.9) was purified and characterized by Wiesenborn DP et al. [65]. The purified enzyme is a hetero-tetramer with subunit molecular weights of 23 and 25 kDa. The K_m values at pH 7.5 and 30°C for acetate, propionate and butyrate were 1200, 1000 and 660 mM, respectively. The K_m values for acetoacetyl-CoA were from 7 to 56 μ M, depending on the acid co-substrate. The acetate and butyrate conversion reactions are inhibited by the physiological acetone and butanol levels in vitro. High K_m values and the inhibitory effects of acetone and butanol are the two regulatory factors for CoA transferase activity in vivo.

It was experimentally confirmed that acetate re-assimilation occurred via the Adc/CtfAB pathway. In contrast, butyrate was not re-consumed via Adc/CtfAB but via Ptb/Buk because, despite the interruption of Adc and CtfA, the mutant was able to consume butyrate. However,

the mechanism of the production and re-consumption of butyrate via Ptb/Buk remains unknown [66].

Acetoacetate decarboxylase enzyme (ADC, EC 4.1.1.4) catalyzed the decarboxylation of acetoacetate to acetone. The expression of the *adc* gene was reported to be up-regulated by Spo0A protein [67]. However, the deletion of the Spo0A box did not fully eliminate the enzyme activity, suggesting additional regulatory factors for the pathway [67].

1.3.8 Ethanol and butanol formation

Ethanol and butanol are synthesized from acetyl-coA and butyryl-coA, respectively. The conversion occurs in the following two steps: the first reduction of acetyl-CoA/butyryl-CoA to acetaldehyde/butyraldehyde and the second reduction of the aldehydes to ethanol/butanol. The *C. acetobutylicum* genome includes two bifunctional alcohol/aldehyde dehydrogenases (*adhE1*-CAP0162 and *adhE2*-CAP0035) and three alcohol dehydrogenases (*bdhA*-CAC3299, *bdhB*-CAC3298 and *bdhC*-CAC3392).

The first bifunctional alcohol/aldehyde dehydrogenase is encoded by *adhE1* and is part of the *sol* operon [68]. The lack of pSOL1 mega-plasmid, which carries the *sol* operon and *adc*, completely abolished solvent production [69]. The inactivation of *adhE1* drastically decreased butanol formation, indicating that this gene product is crucial for butanol production [68]. The second bifunctional alcohol/aldehyde dehydrogenase, encoded by the *adhE2* gene, is highly expressed in alcohologenic cells and poorly expressed in acidogenic cells [42]. Purified AdhE2 protein exhibited NADH-dependent butyraldehyde and butanol dehydrogenase activities. Two transcriptional start sites were determined, at 160 bp and 215 bp upstream of the *adhE2* start codons [70].

Two NADH-dependent butanol dehydrogenase genes *bdhA* (*bdhI*) and *bdhB* (*bdhII*) have been purified and cloned [71]. Genes *bdhA* and *bdhB* were found to be induced or derepressed prior to significant butanol production in controlled pH 5.0 batch fermentations [72].

Recently, AdhE1, AdhE2, BdhA, BdhB, and BdhC were heterologously expressed in *E. coli* and purified as tag free proteins for biochemical characterization by Yoo et al. [42]. AdhE1 was found to exhibit only an NADH-dependent aldehyde dehydrogenase activity and to have lost its alcohol dehydrogenase activity. AdhE2 is an NADH-dependent bifunctional aldehyde/alcohol dehydrogenase activity, as described by Fontaine et al [70]. All three, BdhA/B/C, were reported to be NADPH-dependent butanol dehydrogenase enzymes (Table 1-3).

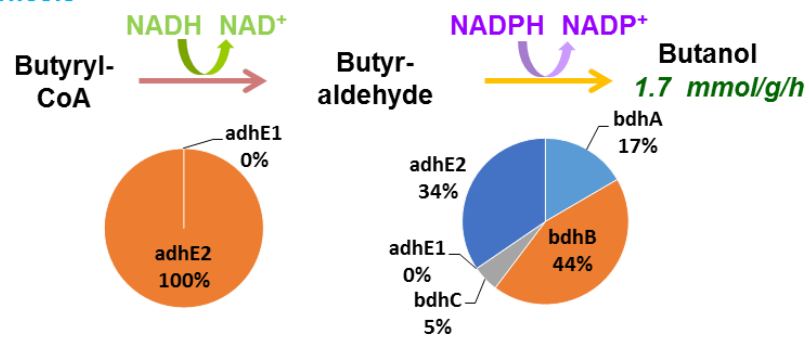
ORF	Gene name	Enzyme activity	Activity (U/mg)	
			NADH	NADPH
CAC3299	bdhA	Butanol dehydrogenase	0.15	2.57
CAC3298	bdhB	Butanol dehydrogenase	0.18	2.95
CAC3392	bdhC	Butanol dehydrogenase	0.24	2.21
CAP0162	adhE1	Butanol dehydrogenase	0.04	ND
CAP0035	adhE2	Butanol dehydrogenase	4.80	0.12
CAP0162	adhE1	Butyraldehyde dehydrogenase	2.27	0.08
CAP0035	adhE2	Butyraldehyde dehydrogenase	2.50	0.07

Table 1-3 Activities of key enzymes catalyzing butanol formation [42]

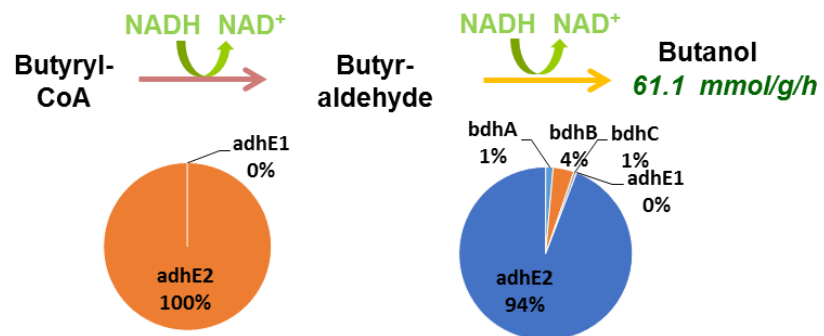
ND: not detected

Based on transcriptomic, proteomic and biochemical data, Yoo et al. have proposed an activity distribution of the five enzymes for two-step catalysis in the butanol pathway (Figure 1-4). In summary, the first step was performed by the aldehyde dehydrogenase activity of AdhE1 or AdhE2, whereas the second step was mainly performed by alcohol dehydrogenase activity of BdhB or AdhE2, depending on the fermentation states.

Acidogenesis



Alcohologenesis



Solventogenesis

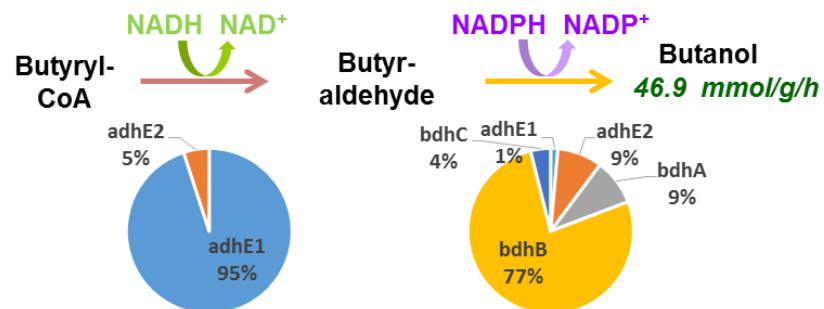


Figure 1-4 Butanol pathway analysis under acidogenesis (A), solventogenesis (B) and alcohologenesis (C) [42]

1.4 Metabolic engineering tools in *Clostridium acetobutylicum*

Great efforts have been made worldwide to develop engineering tools to construct novel strains for different purposes, such as to reduce by-products, produce isopropanol-butanol-ethanol, improve butanol tolerance, engineer for cheap substrate utilization, and so on. This section will summarize the tools developed for *C. acetobutylicum* and their applications in engineering *C. acetobutylicum* strains.

1.4.1 Shuttle vectors for gene over-expression

Previously, vector pFNK1 harboring the replicon derived from pIM13 of *B. subtilis*, considered as a *B. subtilis/C. acetobutylicum* shuttle vector, was used to over-express the *adc* and *ptb* genes. Plasmid pFNK1 was functional as long as the corresponding enzyme activities were elevated [73]. The difficulty was observed in the transformation of *C. acetobutylicum* ATCC 824 with vectors containing DNA from *E. coli* plasmids. A type II restriction endonuclease named *Cac824I* frequently recognizes the sequence 5'-GCNGC-3' and cuts *ColE1* plasmids. Plasmid pAN1 and pAN2, expressing the ϕ 3TI methyltransferase gene of *B. subtilis* phage ϕ 3TI, are used to protect plasmids for easy transformation into *C. acetobutylicum* [74]. Recently, other common shuttle vectors such as pSOS94, pIMP1, pMTL80000 and their derivatives have been used for gene over-expression [73,75–77].

A new study has developed a new shuttle plasmid, designated as pLK1-MCS, that can replicate in both *C. acetobutylicum* and *E. coli* [78]. The plasmid pLK1-MCS, carrying either the pIM13 or the pAM β 1 replicon, is more stable than previously reported plasmids (specifically, pGS1-MCS and pMTL500E) in the absence of antibiotic selective pressure. The transformation efficiency of pLK1-MCS into *C. acetobutylicum* was similar to that of other shuttle plasmids. The plasmid pLK1-MCS carrying the *adhE1-ctfAB-adc* operon compensated in the solvent-deficient *C. acetobutylicum* ML1 due to the loss of the megaplasmid pSOL1. The transformant restored anaerobic solvent production.

1.4.2 Inducible promoter/repressor systems

Several inducible promoter/repressor systems in *C. acetobutylicum* have been reported. A xylose-inducible gene expression system was constructed in *C. acetobutylicum*, based on the *Staphylococcus xylosus* promoter (*xyIA*)-repressor (*xyIR*) regulatory system [79]. A *lacI* expression system driven by the *fdx* promoter from *C. pasteurianum* can be induced in *C.*

acetobutylicum by adding IPTG [74]. Recently, an anhydrotetracycline-inducible Pcm-2tetO1 promoter was reported, allowing reliable stringency and a broad range of inducibility [5]. The most recent system reported was the arabinose inducible system, consisting of AraR from *Bacillus subtilis* and the L-arabinose sensitive-phosphoketolase promoter from *C. acetobutylicum* [80].

1.4.3 Reporter systems

Several reporter systems have been developed to study the expression of genes in *C. acetobutylicum*. The genes *lacZ* from *Thermoanaerobacterium thermosulfurigenes* (encoding β -galactosidase) and *lucB* from *Photinus pyralis* (encoding luciferase) were used as reporter genes in *C. acetobutylicum* to study the promoters of the genes required for solventogenesis and acidogenesis, including *bdhA*, *bdhB*, *adc*, *sol*, and *ptb-buk*. Both systems allowed the detection of differences in promoter strength of up to 100-fold. The luciferase assay could be performed much faster and further optimized for virtual measurement [81]. However, the β -galactosidase from *T. thermosulfurogenes* EM1 LacZ was stable in the exponential phase of growth and degraded in the late exponential and stationary phases [76]. This behavior can limit the use of the LacZ reporter system. The gene *gusA* from *E. coli* was also developed for use as a sensitive reporter system and tested to study the expression of *adc*, *ptb*, *hydA* in *C. acetobutylicum* [79].

1.4.4 Antisense RNA (asRNA)

The asRNA technique employs the small single-stranded RNAs complementary to the target mRNA to inhibit its translation by base pairing. Antisense RNAs offer many advantages for gene down-regulation. First, the asRNA technique is fast. Second, the asRNA strategy can avoid the hidden difficulty of lethal mutations because it does not totally inhibit protein expression. Third, the down-regulation of multiple genes can be conducted by expressing multiple asRNAs on the same plasmid. Lastly, the expression of asRNA can be cleverly controlled using my inducible promoters [82].

As an example of asRNA application, studied by [82], 102-nucleotide asRNA with 87% complementarity to the butyrate kinase (*buk*) gene has been used successfully to reduce Buk activity by 90%. The *ptb* expression was also reduced. The strain produced 50% higher acetone and 30% higher butanol. When using asRNA with 96% complementarity against the *ptb* gene, the strain exhibited only 70%-80% lower Ptb and Buk activity. This strain produced significantly lower butanol and acetone but higher lactate. Butyrate and acetate formation

remained unchanged [82]. In *C. acetobutylicum* ATCC 824 which expressed asRNA against *ctfB* (*ctfb*-asRNA), acetone production was minimal; butyrate and acetate formation increased whereas butanol and ethanol production decreased [83]. When *adhE1* was co-expressed with *ctfb*-asRNA, the butanol titer was restored to the level of the control strain. Ethanol was significantly increased, whereas acetone production remained 63% of its value in the control strain. Butyrate peaks were lower by 49% and 21% compared to the expressing-only *ctfb*-asRNA strain and control strain, respectively. In contrast, peak and final acetate levels in this double mutant were 11% and 68% higher, respectively, compared to the acetate levels from the expressing-only *ctfb*-asRNA strain and control strain. Dramatic transcriptomic changes were observed; however, the regulatory mechanisms remained unknown [83].

1.4.5 Group II intron

The ClosTron employs a group II intron (*L1. LtrB*) from *Lactococcus lactis* and a retro-transposition-activated marker (RAM) to selectively insert DNA into the targeted site to inactivate the target gene [74,84,85]. It also contains flippase recognition target (FRT) sites, which flank the marker cassette. After obtaining the mutant, the second plasmid carrying functional flippase is introduced into the mutant to excise the marker cassette from the genome. The marker-free mutant then could be used for another ClosTron round to inactivate another gene [84]. The tool is functional in various Clostridial species. The disadvantage of using ClosTron used to inactivate a gene is the polar effects it may cause to the downstream genes, which complicate the subsequent analysis of mutant phenotypes [86,87]. On the other hand, the excision of the marker cassette still leaves FRT's scar in the genome. Many FRT scars would be problematic for the next round of flippase excision due to the combination of different identical FRTs existed.

Other alternatives of ClosTron vector were demonstrated [88,89]. Because intron insertion happens frequently, stable and reproducible, it does not necessarily require selective pressure. An alternative group II intron not containing selective marker (plasmid pSY6) has been successfully used [88]. This strategy is time-saving due to avoiding marker excision. The other alternatives, in which erythromycin and thiamphenicol resistant genes were used on the plasmid backbone, provide flexibility for multiple genes knock-out constructions [89].

The intron insertion can be targeted to both strands: the sense and antisense strand. An intron inserted into the antisense strand cannot be spliced and promotes an unconditional disruption. In contrast, intron inserted into sense strand that can be spliced in the presence of

functional *LI*. *LtrB*, in this case, produces a conditional disruption [90]. As a result, when using conditional disruption, curing the plasmid after obtaining the mutant is a critical step. However, this step was not emphasized and was not paid enough consideration in many Clostridia-related publications.

As an example of using ClosTron technology, a large-scale targeted mutagenesis in *C. acetobutylicum* ATCC 824 was conducted by Cooksley et al. [91]. Ten single knock-out mutants were generated, including phosphotransbutyrylase (*ptb*), acetate kinase (*ack*), acetoacetate decarboxylase (*adc*), CoA transferase (*ctfA/ctfB*), alcohol/aldehyde dehydrogenases 1 and 2 (*adhE1*, *adhE2*), butanol dehydrogenases A and B (*bdhA*, *bdhB*) and a putative alcohol dehydrogenase (*CAP0059*) [91]. The inactivation of the main hydrogenase (*hydA*) and thiolase (*thl*) could not be achieved. Phenotypic characterization showed that *bdhA* and *bdhB* mutants did not affect solvent production, whereas inactivation of the previously uncharacterized alcohol dehydrogenase (encoded by *CAP0059*) caused an increase in acetone, butanol, and ethanol formation. Other mutants showed predicted phenotypes: *adc*, *ctfA* and *ctfB* mutants did not produce acetone, *ctfA* and *ctfB* mutants could not re-assimilate acids, and the *ack* mutant drastically reduced acetate formation. The *adhE1* mutant produced very little solvent. The mutant *ptb* did not restore its parental phenotype in the complementary experiment, which was caused by a frameshift mutation in the *thlA* gene. Therefore, the study also insisted on performing genetic complementation tests and Southern blots to confirm single ClosTron insertion and detect mutations elsewhere in the genome.

1.4.6 Genomic mutagenesis by homologous recombination system

1.4.6.1 Homologous recombination by replicative plasmid

The first time homologous recombination by a replicative plasmid was used was to knock-out *Spo0A* (Figure 1-5, [92]). This first system uses no counter-selectable markers. The second cross-over event did not occur at the expected homologous region but accidentally at a ten-nucleotide homologous sequence (ACGACCAAAA). However, the *spo0A* gene was finally interrupted by an insertion of *MLS^r* gene and a small fragment from the pETSPO plasmid backbone.

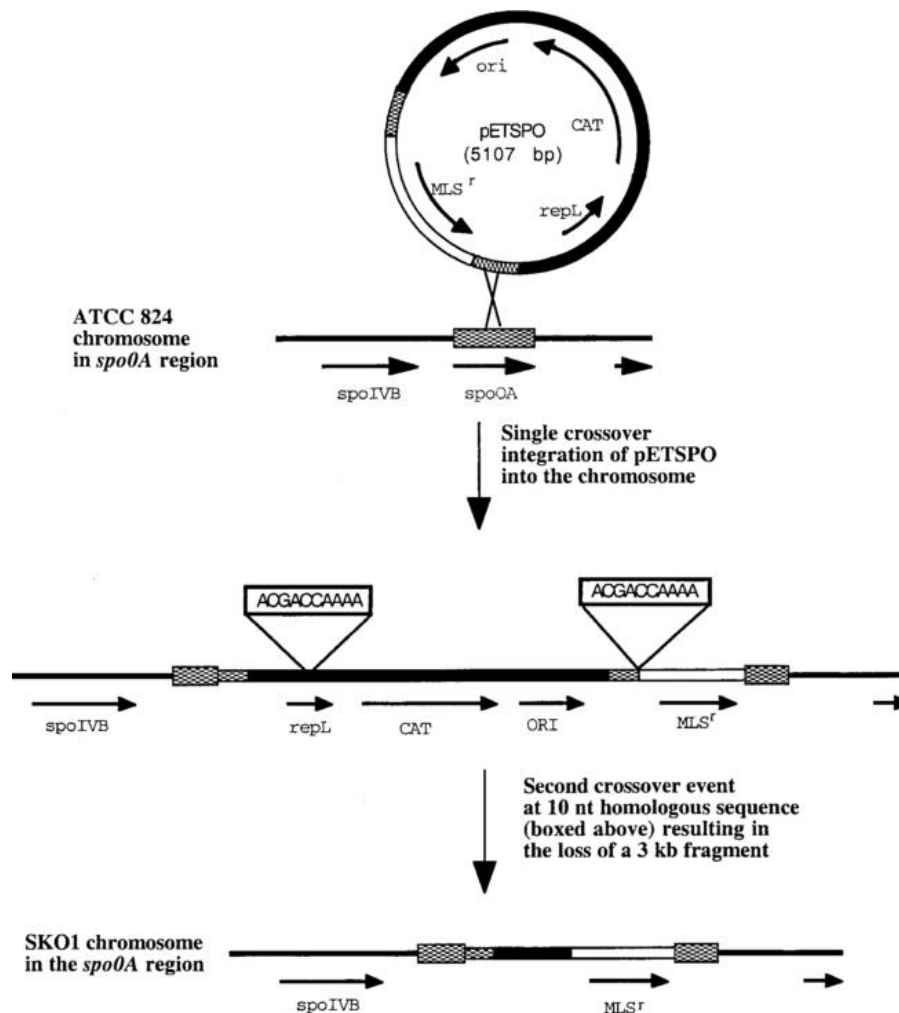


Figure 1-5 Putative cross-over events during *spo0A* gene inactivation [92]

A homologous recombination system to integrate DNA into *C. acetobutylicum*, called allele couple exchange (ACE), was developed. The ACE system is based on the use of a bidirectional marker, both positively and negatively selectable, namely *pyrE* in *C. acetobutylicum*, which encodes the pyrimidine biosynthesis enzyme orotate phosphoribosyl transferase. Cells that contain *pyrE* can be selected on growth medium lacking uracil, as it is required for uracil biosynthesis. Conversely, cells that lack *pyrE* can be selected on the growth medium supplemented with 5-FOA (5-Fluoroorotic acid), as this compound is highly toxic only to cells containing this pathway [93]. To use *pyrE* as a selective marker on the plasmid, the native *pyrE* gene was inactivated by deleting 300 bp from the 3' end of the gene. The system also employs asymmetric homology arms. The long arm directs the first integration event. Afterward, at the short arm, a double cross-over event is facilitated by applying selective pressure. The ACE allowed the rapid insertion of heterologous DNA, without inherent limits on size or complexity, into the genome [93].

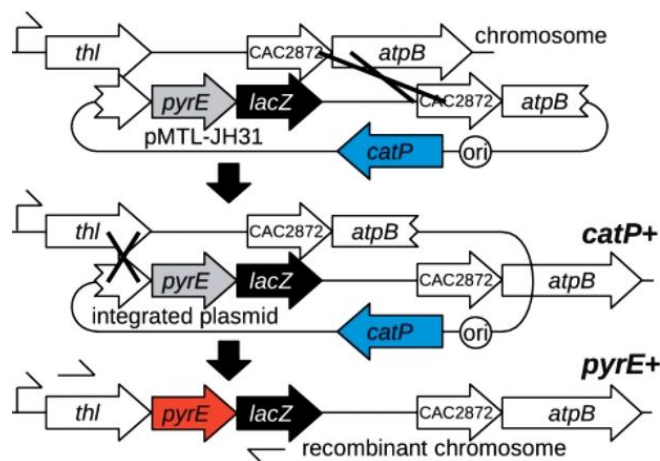


Figure 1-6 DNA integration at the *thl* locus of *C. acetobutylicum* [93].

Figure 1-6 illustrates the ACE technique. Plasmid pMTL-JH31 was used. The first recombination event occurred at the long homologous arm between the pMTL-JH31 plasmid and CAC2872/*atpB* on the chromosome. Single-crossover mutants were selected on plate containing thiamphenicol. The second recombination event to excise the plasmid occurred at the short homologous arm between pMTL-JH31 and the 3'-end of *thl* when growing the single cross-over mutant on medium lacking uracil.

The second alternative ACE system uses *CodA* as a counter-selectable marker [87]. *CodA* encodes cytosine deaminase (EC 3.5.4.1), which converts 5-fluorocytosine (5'-FC) into 5-fluorouracil (5'-FU), which via uracil phosphoribosyltransferase (EC 2.4.2.9) encoded by *upp* in *C. acetobutylicum*, followed by a series of steps causing the permanent inhibition of thymidylate synthase and misincorporation of fluorinated nucleotides into DNA and RNA [87]. As a result, the introduction of a functional *CodA* into *C. acetobutylicum* results in a strain sensitive to the presence of 5'-FC. Most recently, a toxic gene *mazF* (codon optimized from *E. coli* encoding mRNA interferase protein) under the control of a lactose-inducible promoter (from *C. perfringens*) has also been used as a counter-selectable marker [94]

1.4.6.2 Homologous recombination system based on I-SceI endonuclease

This technique uses a replicative plasmid containing two homologous regions on either side of the target sequence and I-SceI recognition sites (flanking the homologous cassette). The plasmid was integrated into the *Clostridium* chromosome, resulting in single-crossover integrants. Afterward, the second plasmid carrying the synthetic gene (*sceC*) encoding the I-SceI enzyme was introduced into the single crossover mutant, resulting in double-stranded breaks at the I-SceI recognition sites. This double-strand break event stimulates homologous recombination and yields double-crossover mutants. The method was applied to delete two genes (*adc* and *glcG*) from *C. acetobutylicum* ATCC 824 and one gene (*adc*) from *C.*

beijerinckii NCIMB 8052 and to introduce point mutations into *xylR* of *C. beijerinckii* NCIMB 8052 [95].

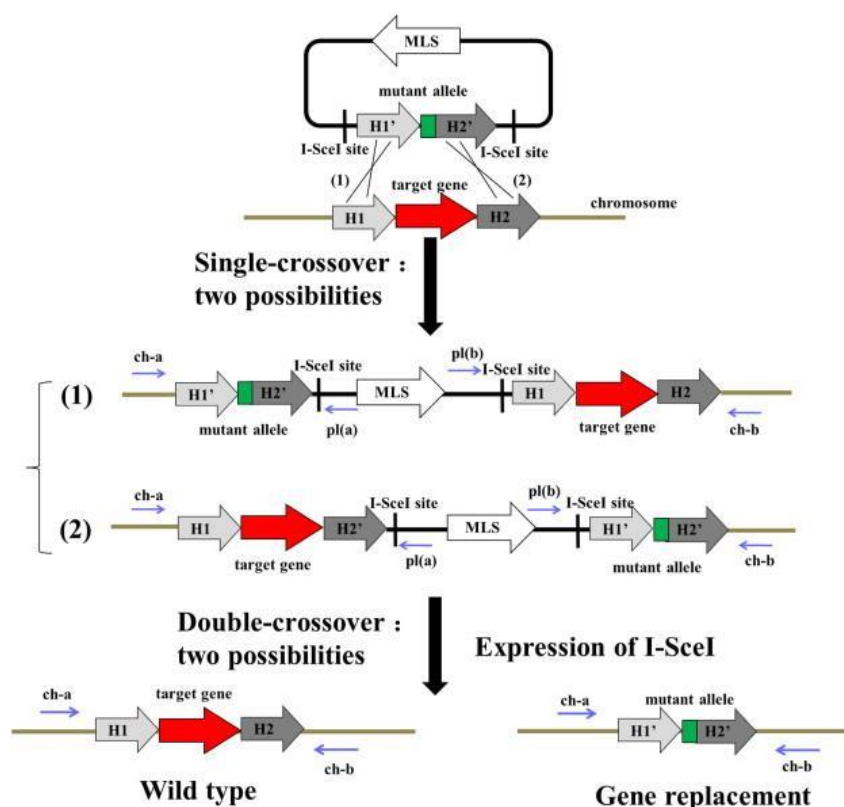


Figure 1-7 I-SceI based-homologous recombination system [95]

1.4.6.3 Non-replicative plasmid for gene inactivation

The first development of a non-replicative plasmid for gene inactivation in *C. acetobutylicum* was published [55,96] for the inactivation of the *ack*, *buk* and *adhE1* genes. The non-replicative plasmid was integrated into the target gene by a single recombination event, inactivating genes. The integrant was quite stable for 25 generations [96]. The segregation frequency (rate of plasmid loss per generation) was 35×10^2 [96]. The resulting integrants of *ack*, *buk* and *adhE1* decreased acetate, butyrate and butanol production, respectively (Table 1-4).

Another genome-engineering tool was recently developed based on single-crossover homologous recombination and the flippase-mediated excision event [97]. The small non-replicable plasmid pHKO1 was designed for efficient integration into the *C. acetobutylicum* genome. The plasmid pHKO1, containing an antibiotic resistance gene, can be excised in vivo by Flp recombinase expressed from the second plasmid pSHL-FLP. The excision leaves a flippase recognition target (FRT) sequence in the middle of the targeted gene. The second plasmid pSHL-FLP plasmid, carrying the Flp recombinase gene and a segregationally

unstable pAM β 1 replicon, could rapidly cure the mutant [97]. However, the technique leaves a second copy of the designed homologous region (approximately 650 bp, Figure 1-8) next to the FRT scar, which raises a concern about the stability of the mutant, as those two copies could randomly recombine to return to the wild-type genotype without any selective pressure.

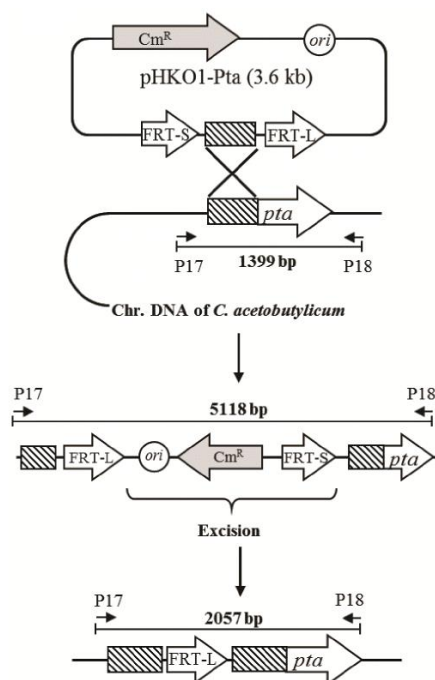


Figure 1-8 Single-crossover homologous recombination [97]

1.4.6.4 Mutagenesis by recombineering technique

The recombineering method employs the mechanism of RecT-mediated homologous recombination. The putative *recT* gene (CPF0939) from *C. perfringens* has been found to be functional in *C. acetobutylicum*. The short complementary oligonucleotides carrying point mutations are introduced and recombined by RecT protein to create mutations in the targeted locus [98].

1.4.7 Random mutagenesis by mariner transposon system

In the forward genetic manipulation, a mariner transposon system was developed to produce highly efficient random mutagenesis in *C. acetobutylicum* ATCC 824. The mini-transposon, which was under a xylose inducible promoter and incorporated with the *catP* gene, can be efficiently trans-positioned into the *C. acetobutylicum* genome. Sporulation-deficient clones were screened among 200 transposon clones. One was found with transposon inserted into the well-known function *spo0A* [99].

1.5 *Clostridium acetobutylicum* as an n-butanol producer-a challenging task

Lütke-Eversloh [86] has summarized the metabolic engineering results in *C. acetobutylicum* up to 2014 (Table 1-4). An updated table is presented in this thesis.

Great efforts have been made to engineer *C. acetobutylicum* to eliminate its by-products such as acetate, butyrate, and acetone (Table 1-4). It is quite difficult to compare those results because the strains were cultivated differently, with different medium components and pH regulation. However, we can see the general impact from the different data. The inactivation of *pta* or *ack* neither totally abolished acetate production nor had a significant impact on other product titers. In contrast, the inactivation of *ptb* or *buk* could minimize butyrate production and improved the ethanol or butanol titers. These mutants also produced higher levels of acetate. The double inactivation of an accurate butyrate biosynthesis pathway and the presence of acetate did not eliminate all acid production: BEKW, BBKW, and PEKW exhibited a decrease in acetate production with no change or an increase in butyrate production. In the quadruple HCBKW or the quintuple HYCBKW mutant, *pta* and *buk* were inactivated; nevertheless, those mutants proved to be acid producers. To the best of my knowledge, the effort to construct a mutant that stops producing acids completely has not yet been successful.

Table 1-4 also shows that inactivating the acetone pathway based on inactivating *ctfA* or *ctfB* resulted in a significantly greater reduction of acetone formation compared to even deactivated *adc*. Those acetone negative strains produced more acetate, butyrate, and lactate. Ethanol and butanol production was reduced.

An interesting study [100] was performed on *C. acetobutylicum* M5, which has lost the pSOL1 megaplasmid that carries the main solvent producing genes (*adc*, *ctfAB*, *adhE1*, *adhe2*). This *C. acetobutylicum* M5 produced only butyrate, acetate and a small amount of ethanol. When over-expressing *adhE1* (under *ptb* promoter) in this M5 strain, butanol was produced at the wild-type level, while butyrate decreased due to the increased flux from butyryl-CoA to butanol. Ethanol and acetate production also increased. To improve butanol production, *thlA* and *adhE1* were co-expressed, but the resulting strain did not produce more butanol as expected. The two single mutants AKKO and BKKO (*ack* knock-out and *buk* a) were constructed from M5. These two mutants produced 60% less acetate and butyrate, respectively. When over-expressing *adhE1* in the AKKO strain, acetate and butyrate were unexpectedly shifted to slightly lower values than M5 complementing *adhE1*. Ethanol and butanol concentration were approximately two-fold lower than for M5 expressing *adhE1*.

What was observed here was the production of butanol (by a pathway that consumes NADH) coupled with the acetate pathway (which produces NADH) in different strain backgrounds of *C. acetobutylicum*. NAD(P)H ferredoxin oxidoreductase (FNOR) regenerates NAD⁺ and oxidizes ferredoxin to provide NADH for solventogenesis. It has been stated that because FNOR could be tightly controlled, the NADH production had to come from the additional acetate production to obtain a redox balance. Consequently, it would be challenging to construct a high butanol-producing strain without acids and acetone production, as we do not yet know the activating mechanism to increase butanol production [66,100].

There are also other approaches that can help to increase butanol selectivity. Recently, the inactivation of histidine kinase enzyme (encoded by *CAC3319*) has shown improved butanol tolerance and the production of 44% more butanol [101]. Previously, the over-expression of phosphofructokinase (*pfkA*) and pyruvate kinase (*pykA*) increased the butanol titer (30%) and ethanol titer (85%) due to the boosted ATP and NADH concentrations and the enhanced butanol tolerance [102]. In a new publication [103], the biotin uptake and synthesis genes (*CAC1360-CAC1362* and *CAC0210*) were over-expressed and shown to improve the ABE titers to 30% in the free biotin media. Over-expression of the C4 pathway (*hbd*, *thlA*, *crt*, *bcd*, *adhE1*, and *ctfAB*) in a strain containing the *adc* gene replaced by *gshAB* (from *E. coli*, involved in glutathione biosynthesis) increased the butanol titer from 5.17 g/L in the control to 14.8 g/L in the mutant [104].

Mutant strain	Inactivated gene	Pathway	Product pattern						Reference
			Acetate	Butyrate	Lactate	Acetone	Ethanol	Butanol	
PJC4PTA ^{a,d}	<i>pta</i>	Acetate biosynthesis	-	+	uk	0	0	0	[55]
<i>pta::int(80)</i> ^b	<i>pta</i>	Acetate biosynthesis	0	0	0	0	0	0	[66]
EKW ^b	<i>pta</i>	Acetate biosynthesis	-	0	uk	+	+	+	[89]
WUR AK ^c	<i>ack</i>	Acetate biosynthesis	-	0	0	0	0	0	[105]
Cac-ack349a::CT ^b	<i>ack</i>	Acetate biosynthesis	-	0	+	0	+	+	[91]
AKW ^b	<i>ack</i>	Acetate biosynthesis	-	+	uk	0	0	0	[89]
PJC4BK ^{a,d}	<i>buk</i>	Butyrate biosynthesis	+	-	uk	-	+	+	[55]
824 (buk-) ^b	<i>buk</i>	Butyrate biosynthesis	uk	uk	uk	0	+	+	[88]
BKW ^b	<i>buk</i>	Butyrate biosynthesis	-	-	uk	+	+	+	[89]
<i>ptb::int(87)</i> ^c	<i>ptb</i>	Butyrate biosynthesis	+	--	0	0	++	-	[106]
Cac-ptb50a::CT ^b	<i>ptb</i>	Butyrate biosynthesis	++	--	++	--	++	-	[91]
PKW ^b	<i>ptb</i>	Butyrate biosynthesis	0	0	uk	0	0	+	[89]
BEKW ^b	<i>pta</i> and <i>buk</i>	Acetate, butyrate biosynthesis	-	0	uk	-	0	+	[89]
BBEKW ^b	<i>pta</i> , <i>buk</i> and <i>bukII</i>	Acetate, butyrate biosynthesis	-	0	uk	-	0	-	[89]
PEKW ^b	<i>pta</i> and <i>ptb</i>	Acetate, butyrate biosynthesis	-	+	uk	-	0	-	[89]
2018adc ^b	<i>adc</i>	Acetone biosynthesis	++	0	uk	-	-	-	[107]

acd::int(180) ^b	<i>adc</i>	Acetone biosynthesis	++	+	0	-	-	-	[66]
Cac-adc180s::CT ^b	<i>adc</i>	Acetone biosynthesis	+	+	+	-	0	-	[91]
pta::int(80)ΔRAM-adc::int(180) ^b	<i>pta and adc</i>	Acetate, acetone biosynthesis	-	++	++	-	-	-	[66]
ctfA::int(352) ^b	<i>ctfa</i>	Acetone biosynthesis	++	0	0	--	-	-	[66]
Cac-ctfA(398s) ^b	<i>ctfa</i>	Acetone biosynthesis	+	++	++	--	0	-	[91]
Cac-ctfB(132s) ^b	<i>ctfb</i>	Acetone biosynthesis	+	++	++	--	0	-	[91]
CKW ^b	<i>ctfb</i>	Acetone biosynthesis	+	+	uk	--	-	-	[89]
CEKW ^b	<i>pta and ctfb</i>	Acetate, acetone biosynthesis	-	+	uk	--	-	-	[89]
pta::int(80)ΔRAM-ctfA::int(352) ^b	<i>pta and ctfa</i>	Acetate, acetone biosynthesis	-	++	++	--	-	-	[66]
CBEKW ^b	<i>pta, buk and ctfb</i>	Acetate, acetone, butyrate biosynthesis	-	+	uk	--	-	-	[89]
hbd::int(69) ^b	<i>hbd</i>	Butanol, butyrate biosynthesis	-	--	0	-	++	--	[38]
PJC4AAD ^{b,d}	<i>adhE1</i>	Alcohol biosynthesis	+	+	uk	--	+	-	[96]
Cac-adhE1468s::CT ^b	<i>adhE1</i>	Alcohol biosynthesis	+	++	0	-	-	-	[91]
Cac-adhE2696s::CT ^b	<i>adhE2</i>	Alcohol biosynthesis	0	0	0	0	0	0	[91]
Cac-bdhA459s::CT ^b	<i>bdha</i>	Alcohol biosynthesis	0	0	0	0	0	0	[91]
Cac-bdhB475a::CT ^b	<i>bdhb</i>	Alcohol biosynthesis	0	0	0	0	0	0	[91]
Cac0059-261s::CT ^b	<i>Cac0059</i>	Alcohol biosynthesis	-	-	+	+	+	+	[91]

HCEKW ^c	<i>pta, ctfb</i> and <i>adhE1</i>	Acetate, acetone, alcohol biosynthesis	-	++	uk	--	-	--	[108]
HCBEKW ^c	<i>pta, buk, ctfb</i> and <i>adhE1</i>	Acetate, butyrate, acetone, alcohol biosynthesis	+	++	++	--	--	-	[108]
HYCBEKW ^c	<i>pta, buk, ctfb, adhE1</i> and <i>hydA</i>	Acetate, butyrate, acetone, alcohol, hydrogen biosynthesis	+	++	+	--	--	-	[108]
SolRH ^{c,d}	putative repressor of solventogenic genes <i>solR</i>	Solventogenesis	+	0	uk	+	+	+	[109]
824 (solR-) ^b	<i>solR</i>	Solventogenesis	uk	uk	uk	+	+	+	[88]
Rex::int(95) ^b	redox sensing transcriptional regulator <i>rexA</i>	Solventogenesis	0	-	0	-	++	+	[110]
HKKO ^c	histidine kinase		+	-	uk	+	+	+	[101]
Pta ::int(80) ^e	<i>pta</i>	Acetate biosynthesis	--	0	uk	0	-	-	[111]
Adc ::int(180) ^e	<i>adc</i>	Acetone biosynthesis	+	--	uk	--	-	-	[111]
Adc ::int(80)ΔRAM pta ::int(80) ^e	<i>adc</i> and <i>pta</i>	Acetate, acetone biosynthesis	+	-	uk	--	-	-	[111]
Adc::int(180)	<i>adc</i>	Acetone biosynthesis	++	++	uk	--	-	-	Acetoin -, [59]
Adc::int(180) pSY8-alsD	<i>adc</i>	Acetone biosynthesis	++	++	uk	--	-	-	Acetoin ++, [59]

Table 1-4 *C. acetobutylicum* mutants and changes in product profiles [86]

+, increased, ++, highly increased (>100%); 0, unchanged; -, reduced; --, none; uk, unknown

a, pH 5.5 controlled batch; b, without pH controlled batch; c, pH 5.0 controlled batch; d, knock out mutant obtained by integration plasmids, all other mutants were generated using mobile group II intron; e, pH 4.5 chemostat.

1.6 Genome scale metabolic (GSM) models for *Clostridium acetobutylicum*

Genome-scale metabolic models of *Clostridium acetobutylicum* have been reconstructed, updated and improved by many groups. The two primary models constructed were iJ432 [112] and Sanger [113], published in 2008. The iJ432 and Sanger's models were reported to successfully predict metabolic fluxes in acidogenesis and solventogenesis. The Sanger model could also predict the loss of butyric acid production in the *buk⁻* strain and the loss of solvents in the degenerated M5 strain. The iCac490 model [114], integrating more reactions, and thermodynamic constraints based on enzyme specificity could simulate the fluxes of acetate, lactate, butyrate, acetone, butanol, ethanol, CO₂ and H₂ in the wild-type strain. The iCac802 [115] in 2014 was another improvement in terms of the gene, reaction and metabolite numbers. In addition, this model integrated transcriptomic data regarding metabolic changes in response to butanol and butyrate stresses. The latest model, iCac967 [42], has updated the biochemical characteristics of 15 key metabolic enzymes for an accurate fluxomics data [42].

Strain	Metabolic model			Reference
	Genes	Reactions	Metabolites	
<i>C. acetobutylicum</i> ATCC 824	432	502	479	iJL432 [112]
<i>C. acetobutylicum</i> ATCC 824	473	522	422	Sanger [113]
<i>C. acetobutylicum</i> ATCC 824	490	794	707	iCac490 [114]
<i>C. acetobutylicum</i> ATCC 824	802	1462	1137	iCac802 [115]
<i>C. acetobutylicum</i> ATCC 824	967	1231	1058	iCac967 [42]

Table 1-5 Genome scale metabolic models of *C. acetobutylicum* ATCC 824

1.7 Regulatory network in *Clostridium acetobutylicum*

1.7.1 Spo0A and Sigma factors

In *C. acetobutylicum*, Spo0A is a transcriptional regulator that positively controls sporulation and solvent production. The inactivation of Spo0A (mutant SKO1) also drastically decreases solvent formation [92], as the SKO1 mutant cannot septate and form endospores. In the strain over-expressing Spo0A (824(pMSPOA)), butanol concentration was higher than in the control strain, and sporulation was enhanced and accelerated [92]. Spo0A has a conserved C-terminal in both *C. beijerinckii* and *C. acetobutylicum*. Spo0A binding motifs (TGNCGAA) were found in the promoter regions of the *adc* and *ptb* genes. Gel retardation confirmed the direct control of *adc* and *ptb* gene expression in *C. acetobutylicum* [67].

The sporulation process in *C. acetobutylicum* begins with the phosphorylation of Spo0A. The phosphorylated SpoA-P then activates other sigma factors such as σE , σF , σG , and σK , which are already well known. The σE inactivation mutant did not produce granulose but did produce solvents dependent on cell states, whereas the σG inactivation mutant produced granulose and exhibited a wild-type-like solventogenesis [116]. The factor σK is important in the early stage, prior to Spo0A expression, and also in the late stage for σG expression. The deletion of σK rendered the mutant unable to sporulate or produce solvents [117]. When σF was disrupted, the expression of other sigma factors, including σE and σG , was not detected, and differentiation was blocked [118]. Al-Hinai et al. provided an excellent comparative summary of the sporulation programs in *B. subtilis* and *C. acetobutylicum* [119]. The incomplete sporulation program in *C. acetobutylicum*, however, has been shown to be different from the one in *B. subtilis*.

1.7.2 The controversial SolR

Nair et al [120] claimed that SolR (located 663 bp upstream of the *adhE* start codon) was a putative binding transcriptional repressor that negatively regulates solventogenic metabolism. It was induced in acidogenesis and suppressed in solventogenesis. In the *solR*-inactivated strain, the expression of the solventogenic genes is induced earlier and maintained throughout the solventogenic phase. Nair's conclusion has proven to be controversial. Thormann et al [121] reported that during the construction of the DNA clone, the *solR* knock-out phenotype also contained the binding site of a regulator, possibly SpoA, that affected the *sol* operon expression.

1.7.3 Agr quorum sensing

This study was conducted by Steiner et al [122]. In the *C. acetobutylicum* genome, there is an *agr* locus, *agrBDCA*, known to form a cyclic peptide-based quorum sensing system in *Staphylococci*. In *Staphylococci*, *agrBD* is necessary for the production of a peptide signal. The peptide signal accumulates extracellularly and is sensed by a two-component system encoded by *agrCA*. To understand the role of *agrBDCA* in *C. acetobutylicum*, the ClosTron technology was used to generate the *agrB*, *agrC* and *agrA* mutants, and their phenotypes were characterized. Those mutants and the wild-type displayed similar growth and no obvious differences in solvent formation under the conditions tested. However, on agar-solidified medium, spore formation was strongly affected, particularly in the *agrA* and *agrC* mutants. Similarly, granulose accumulation was almost undetectable in colonies of the *agrB*, *agrA*, and *agrC* mutants. AgrBD complementation in the *agrB* mutant restored granulose and spore

formation. In addition, when synthetic cyclic peptides based on the *agrD* sequence were added exogenously into the *argB* mutant culture, the parental phenotype was also restored. These findings support the hypothesis that *agr*-dependent quorum sensing is involved in the regulation of sporulation and granulose formation in *C. acetobutylicum* [122].

1.7.4 Heat-shock proteins: DnaK, GroESL, and Hsp21

The so-called heat-shock proteins, including DnaK and GroEL (two highly conserved molecular chaperones), were induced by a temperature upshift (from 30 to 42°C) [123]. The *dnaK* operon includes four genes: *orfA* (a heat shock gene with an unknown function), *grpE*, *dnaK*, and *dnaJ*. The *groE* operon includes *groES* and *groEL*. A conserved inverted repeat was found upstream of *dnaK* and *groEL* and may be implicated as a *cis*-acting regulatory site. Those heat-shock genes were expressed at the beginning of solvent formation and after heat stress. The mRNA levels of heat-shock genes and solvent-producing genes reached a maximum at the same time during the metabolic shift [123]. In addition, over-expression of the heat shock proteins (GroES and GroEL) allowed an increase in the final solvent concentration, presumably by stabilizing solventogenic enzymes [124]. Therefore, the heat-shock response in *C. acetobutylicum* might be part of a global regulatory network including different stress responses, such as heat shock, metabolic switching and sporulation [123].

The small heat shock protein Hsp21 of *C. acetobutylicum* was identified as a rubrerythrin-like protein. The transcription of its encoding *rbr3AB* operon was induced by various environmental stress conditions: heat, oxidative stress, pH of the growth medium, addition of NaCl or butanol, and incubation temperature [125].

1.7.5 PhoPR

The *phoPR* gene locus of *C. acetobutylicum* comprises two genes, *phoP* and *phoR*. These PhoPR proteins are predicted to represent a response regulator and sensor kinase of a phosphate-dependent two-component regulatory system [126]. It was shown that a significant increase in the mRNA transcript levels of both *phoP* and *phoR* was observed when external phosphate concentrations dropped below 0.3 mM. The autophosphorylation of PhoR and phosphorylation of PhoP were shown in vitro. An electromobility shift assay proved that there was a specific binding of PhoP to the promoter region of the phosphate-regulated *pst* operon (comprising five genes, *pstS*, *pstC*, *pstA*, *pstB*, and *phoU*) of *C. acetobutylicum* [126].

1.7.6 AbrB310

SinR and AbrB are known to be transcriptional regulators involved in sporulation initiation in *Bacillus subtilis*. The role of SinR and AbrB in *C. acetobutylicum* was studied by Scotcher et al [127]. There were three highly similar homologs to *B. subtilis abrB* found in *C. acetobutylicum*: *abrB310*, *abrB1941* and *abrB3647*. The promoters of *abrB1941* and *abrB3647* were inactive, whereas the *abrB310* promoter was strongly active. The expression of *abrB310* was elevated at the beginning of solventogenesis. Moreover, the promoter of *abrB310* contains a putative OA box. When *abrB310* expression was repressed by asRNA, the acids (acetate and butyrate) accumulated to approximately twice their normal concentration. Acetone and butanol production were delayed and decreased. It was also found that sporulation in this mutant was delayed, but the morphology of the sporulating cells and spores was normal. Therefore, it was proposed that AbrB310 may act as a regulator of the transition between the acidogenic and solventogenic states. Regarding SinR, a single homolog to *B. subtilis* SinR was found. When the *C. acetobutylicum* strain was transformed with plasmids to increase or decrease SinR expression, no significant differences were observed either in growth or in acid or solvent production compared to the control strain [127].

1.7.7 Per and Fur

Peroxide repressor (PerR)-homologous protein was identified as a key repressor that plays an important role in oxidative stress defense [128,129]. In vivo, PerR from *C. acetobutylicum* acts as a functional peroxide sensor. Transcriptomic data showed that genes involved in the detoxification system were regulated in PerR mutant, including reverse rubrerythrins (*rbr3A-rbr3B*), desulfoferrodoxin (*dfx*), rubredoxin (*rd*), NADH-dependent rubredoxin oxidoreductase (*NROR*) and the oxygen-reducing flavodiiron proteins (FDPs, encoded by *fprA1* and *fprA2*). Only a few targets of direct PerR regulation were identified. They include two highly expressed genes, *gapN* and *CODH*, that are putatively involved in the central energy metabolism. Under exposure to O₂, *C. acetobutylicum* also activates the repair and biogenesis of DNA and Fe-S clusters. The genes that were downregulated when exposed to O₂ were the ones involved in butyrate formation [129].

In the same family as PerR, a ferric uptake regulator (Fur) that helps *C. acetobutylicum* sense and respond to the availability of iron was identified and characterized [130]. The Fur mutant was constructed and showed a slow-growing phenotype and an enhanced sensitivity to oxidative stress, but no dramatic change in the fermentation pattern was observed. However, the mutant overproduced riboflavin. The operon *ribDBAH* responsible for riboflavin synthesis

was significantly upregulated. Iron limitation and inactivation of Fur also affected the expression of several genes involved in energy metabolism. Two genes, encoding a lactate dehydrogenase (*ldh*) and a flavodoxin (*fld*), were highly induced [130].

1.7.8 CsrA

Carbon storage regulator (CsrA, encoded by *CAC2209*) largely involved in regulating multiple pathways, including flagella assembly, oligopeptide transport, phosphotransferase transport systems, stage III sporulation and the central carbon metabolism [131]. The disruption of CsrA resulted in a decrease in solvent production. Genes involved in iron uptake and riboflavin synthesis were also reported to be regulated by CsrA in this study.

1.7.9 CcpA

Catabolite control protein (CcpA) mediates the utilization of hexose and pentose sugar in *C. acetobutylicum*. When CcpA was mutated at V302N, the utilization of xylose was greatly improved [132]. When this mutant was coupled with the over-expression of *sol* genes (*ctfA*, *ctfB* and *adhE1*), 30% more xylose was consumed compared to the wild-type when fermented in a mixture of glucose and xylose. EMSA experiments confirmed His-CcpA binding to the promoters of *xylAB*, *sol*, and *CAC1353* (PTS system transporter subunit IIC). The mutated CcpA-V302N exhibited lower affinities to those promoters than the wild type [132].

The RegA protein of *C. acetobutylicum*, a homolog of the catabolite control protein CcpA in *B. subtilis*, has been found to inhibit the degradation of starch by the *C. acetobutylicum* *staA* gene product in *E. coli* [133].

1.7.10 Rex

A redox-sensing transcriptional repressor (Rex) was recently found to modulate its DNA-binding activity in response to the NADH/NAD⁺ ratio and to participate in the solventogenic shift of *C. acetobutylicum* [110,134]. High NADH/NAD⁺ prevents the binding of Rex to DNA [110]. Rex binding sites were found in the promoters of different operons: *adhE2*, *ldh*, *crt*, *thlA*, *asrT*, *ptb*, and *nadA* [110,134], and those operons were upregulated in the Rex mutant. Among them, *ldhI* was the most strongly induced gene, followed by *adhE2*. Genes encoding proteins involved in detoxification, including the reverse rubrerythrins (*rbr3A-rbr3B*), desulfoferrodoxin (*dfx*), rubredoxin (*rd*), NADH-dependent rubredoxin oxidoreductase (*nror*) and the oxygen reducing flavoproteins (*fprA1* and *fprA2*), were down regulated by 2 to 18-fold compared to the wild type. The Rex mutant produced high amounts

of ethanol and butanol and a significantly lower amount of acetone and hydrogen than the wild type [110,134].

1.7.11 AraR

The L-arabinose utilization pathway has been studied extensively in bacteria, including Gram-positive *B. subtilis*. The permease encoded by *araE* transports L-arabinose into the cells. Subsequent genes involved in the conversion of L-arabinose belong to the operon *araABDLMNPQ-abfA*. The *AraR* gene encodes the regulatory protein for the L-arabinose metabolism in *B. subtilis* [135,136], which negatively controls the expression of the *ara* regulon. AraR exhibits a DNA binding domain at the N-terminal region and an effector binding domain at the C-terminal [136]. In the absence of arabinose, AraR represses the expression of the *ara* regulons. In the AraR mutant, there was a dramatic increase in the expression of the genes involved in arabinose utilization (two *araEs*, *araD*, two *araAs* and *araK*), the genes in the pentose phosphate pathway (*ptk*, *tal* and *tkl*), and genes involved in arabinoside degradation (*epi* and *arb43*). EMSA experiments confirmed the affinity of AraR for the promoter of *araK*, *araE*, *araD*, *ptk* and *araR* itself. In the presence of L-arabinose, AraR was released from the AraR-DNA complex [136].

1.7.12 Small regulatory RNA

Small noncoding RNAs are widespread in all kingdoms of life, where they serve to regulate and tune gene expression, subsequently influencing a wide range of cellular processes that include environmental stress responses and virulence processes in pathogens [137]. Many sRNAs have been described in Gram-negative bacteria, but our understanding of their role in Gram-positive bacteria has been slower [137]. Hundreds of sRNAs have recently been identified in Clostridia. Chen et al [138] have developed a method to identify potential sRNAs in the *Clostridium* genus. By genomics comparison and the in-silico prediction of rho-independent terminators and promoters, they have predicted sRNAs in the following 21 Clostridial genomes: *Clostridium acetobutylicum*, *C. beijerinckii*, eight *C. botulinum* strains, *C. cellulolyticum*, *C. difficile*, two *C. kluyveri* strains, *C. novyi*, three *C. perfringens* strains, *C. phytofermentans*, *C. tetani*, and *C. thermocellum*. Most of the predicted sRNAs were found in noncoding regions. Seven of these sRNAs have Shine-Dalgarno (SD) sequences, and some possess a start codon downstream of the SD sequences. Thirty sRNAs (of 113 predicted) of *C. acetobutylicum* and 21 sRNAs of *C. botulinum* were validated by northern blot analysis and quantitative reverse transcription PCR (qRT-PCR). A conserved, novel sRNA was found to be responsive to the antibiotic clindamycin [138]. Lately, Venkataramanan et al. [139] have

used RNA deep sequencing to examine the sRNome of *C. acetobutylicum* that is responsive to butanol and butyrate and have identified 46 additional sRNAs. Under metabolite stress, these 159 sRNAs display divergent expression patterns. The expression of sRNAs affecting transcriptional (6S, S-box, and solB) and translational (tmRNA and SRP-RNA) processes and RNA chaperone Hfq were found to be stress-related [139].

Chapter 2 Thesis Objectives

This PhD thesis has two main parts, as follows: the metabolic engineering of *C. acetobutylicum* for the production of fuels and chemicals and understanding the regulatory role of an uncharacterized Cap0037 protein.

The metabolic engineering part has three objectives:

- Detail the homologous recombination technique by replicative plasmid that has been used successfully to construct *C. acetobutylicum* mutants.
- Construct a *C. acetobutylicum* mutant that produces n-butanol with high selectivity and yield on glucose. By the homologous recombination method, the metabolic pathways leading to by-products including lactate, butyrate and acetone were eliminated. The carbon fluxes were guided to n-butanol formation by replacing *hbd* with *hbd1* and *thla* with *atoB* genes. The final strain CAB1060 also contains inactivated Rex protein which globally regulates metabolism in response to the intracellular NADH/NAD⁺ concentration. This CAB1060 mutant was then evaluated for its ability to produce butanol at high yield and selectivity in a continuous bioreactor coupling alcohol extraction by vacuum distillation to a cell recycling in order to improve productivities.
- Construct a *C. acetobutylicum* platform strain unable to produce hydrogen and useful for further engineering to produce chemicals and fuels. In this part, a new homologous recombination system using non-replicative plasmid was developed.

In the second part, the Cap0037 protein of unknown function was primarily studied. To do so, a *Cap0037* inactivated mutant was constructed using group II intron. A thorough comparison of the *Cap0037::int* mutant and the wild-type strain was analyzed from the transcriptomics, fluxomics, and proteomics aspects. In addition, DNA binding activity of Cap0037 was also studied. The potential Cap0037 regulon was predicted by bio-informatics tools.

Chapter 3 Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*

Christian Croux^{1#}, Ngoc-Phuong-Thao Nguyen^{1#}, Jieun Lee², Céline Raynaud³, Florence Saint-Prix¹, Maria Gonzalez-Pajuelo¹, Isabelle Meynial-Salles¹ and Philippe Soucaille^{1,3*}

Published in Biotechnology for Biofuels, 2016 Feb 2; 9:23. doi: 10.1186/s13068-016-0432-2

3.1 Abstract

Clostridium acetobutylicum is a Gram-positive, spore-forming, anaerobic bacterium capable of converting various sugars and polysaccharides into organic acids (acetate and butyrate) and solvents (acetone, butanol, and ethanol). *Clostridium acetobutylicum* has become of interest in postgenomic era due to the complete sequencing and annotation of its genome, supplying a wealth of information on its protein machinery. This global knowledge has prompted new approaches to genetic analysis, functional genomics, and metabolic engineering to develop an industrial strain for the production of biofuels and bulk chemicals. The method used in this study to knock-out or knock-in genes in *C. acetobutylicum* combines the use of an antibiotic resistance gene for the deletion or replacement of the target gene, the subsequent elimination of the antibiotic resistance gene with the flippase (FLP) recombinase system from *Saccharomyces cerevisiae*, and a *C. acetobutylicum* strain that lacks *upp*, which encodes uracil phosphoribosyl transferase, for subsequent use as a counter-selectable marker. A replicative vector containing i) a pIMP13 origin of replication from *Bacillus subtilis* that is functional in *Clostridia*, ii) a replacement cassette consisting of an antibiotic resistance gene (*MLSR*) flanked by two FRT sequences, and iii) two sequences homologous to selected regions around target DNA sequence was first constructed. This vector was successfully used to consecutively delete the *Cac824I* restriction endonuclease encoding gene (*CA_C1502*) and the *upp* gene (*CA_C2879*) in the *C. acetobutylicum* ATCC824 chromosome. The resulting *C. acetobutylicum* Δ *cac1502**Uupp* strain is marker-less, readily transformable without any previous plasmid methylation and can serve as the host for the “marker-less” genetic exchange system. The third gene, *CA_C3535*, shown in this study to encode for a type II restriction enzyme (*Cac824II*) that recognizes the *CTGAAG* sequence, was deleted using an *upp*/5-FU counter-selection strategy to improve the efficiency of the method. The restriction-less marker-less strain and the method was successfully used to delete two genes (*ctfAB*) on the pSOL1 mega-plasmid and one gene (*ldhA*) on the chromosome to get strains no longer producing acetone or L-lactate. The restriction-less, marker-less strain described in this study, as well as the marker-less genetic exchange coupled with positive selection, will be useful for functional genomic studies and for the development of industrial strains for the production of biofuels and bulk chemicals.

3.2 Introduction

In recent years, *Clostridium acetobutylicum* ATCC824 has been of interest in the postgenomic era due to the complete sequencing and annotation of its genome [19], supplying a wealth of information regarding its protein machinery. This global knowledge has prompted new approaches to genetic analysis, functional genomics, and metabolic engineering in order to develop industrial strains for the production of biofuels and bulk chemicals.

To this end, several reverse genetic tools have been developed for *C. acetobutylicum* ATCC 824, including a gene inactivation system based on non-replicative [55,96] and replicative plasmids [92–94,140] and the group II intron gene inactivation system [74,88]. Among these methods, only the method developed by Al-Hinai et al.[94] allows for *in frame* deletions and/or the introduction of genes at their normal chromosomal context without an antibiotic marker remaining. This system is made of two parts. The first part is a replicative vector containing (1) a pIMP13 origin of replication from *Bacillus subtilis* functional in *Clostridia*, (2) a replacement cassette consisting of an antibiotic resistance gene (*Th^R*) flanked by two FRT sequences, (3) two sequences homologous to the selected regions around the target DNA sequence, and (4) a codon-optimized *mazF* toxin gene from *Escherichia coli* under the control of a lactose-inducible promoter from *Clostridium perfringens* to allow for the positive selection of double-crossover allelic exchange mutants. The second part is a plasmid system with inducible segregational instability, enabling efficient deployment of the FLP-FRT system to generate marker-less deletion or integration mutants.

In 2006, our group patented a marker-less, in-frame deletion method [141] similar to the two-part method published by Al-Hinai et al. [94] in 2012. The first part of our method is based on the same replicative plasmid and the same replacement cassette, but it uses the uracil PRTase *upp*/5-fluorouracil (5-FU) system as a counter-selection strategy. The second part is based on a plasmid carrying (1) the FLP-FRT system to generate marker-less deletion and (2) the uracil PRTase *upp*/5-FU system to select for the plasmid loss after marker excision. This method was successfully used by the Metabolic Explorer Company to develop and patent an industrial recombinant strain of *C. acetobutylicum* for *n*-butanol production. As this method has not been described in detail and to make it available to and usable by the scientific community, we report how this method was developed and its use to create a restriction-less, marker-less strain of *C. acetobutylicum*. We show that this strain lacking *upp* (CA_C2879,

encoding the uracil-phosphoribosyl-transferase), *CA_C1502* encoding *Cac824I* and *CA_C3535* encoding *Cac824II* (the second type II restriction enzyme) can be transformed by non-methylated DNA at very high efficiency and can be used for rapid gene knock-in and knock-out using the *upp/5-FU* counter-selectable system for both functional genomic and metabolic engineering of *C. acetobutylicum*. This strain and the method were further used to delete three genes *ctfAB* and *ldhA* to create strains no longer producing acetone and lactate, respectively.

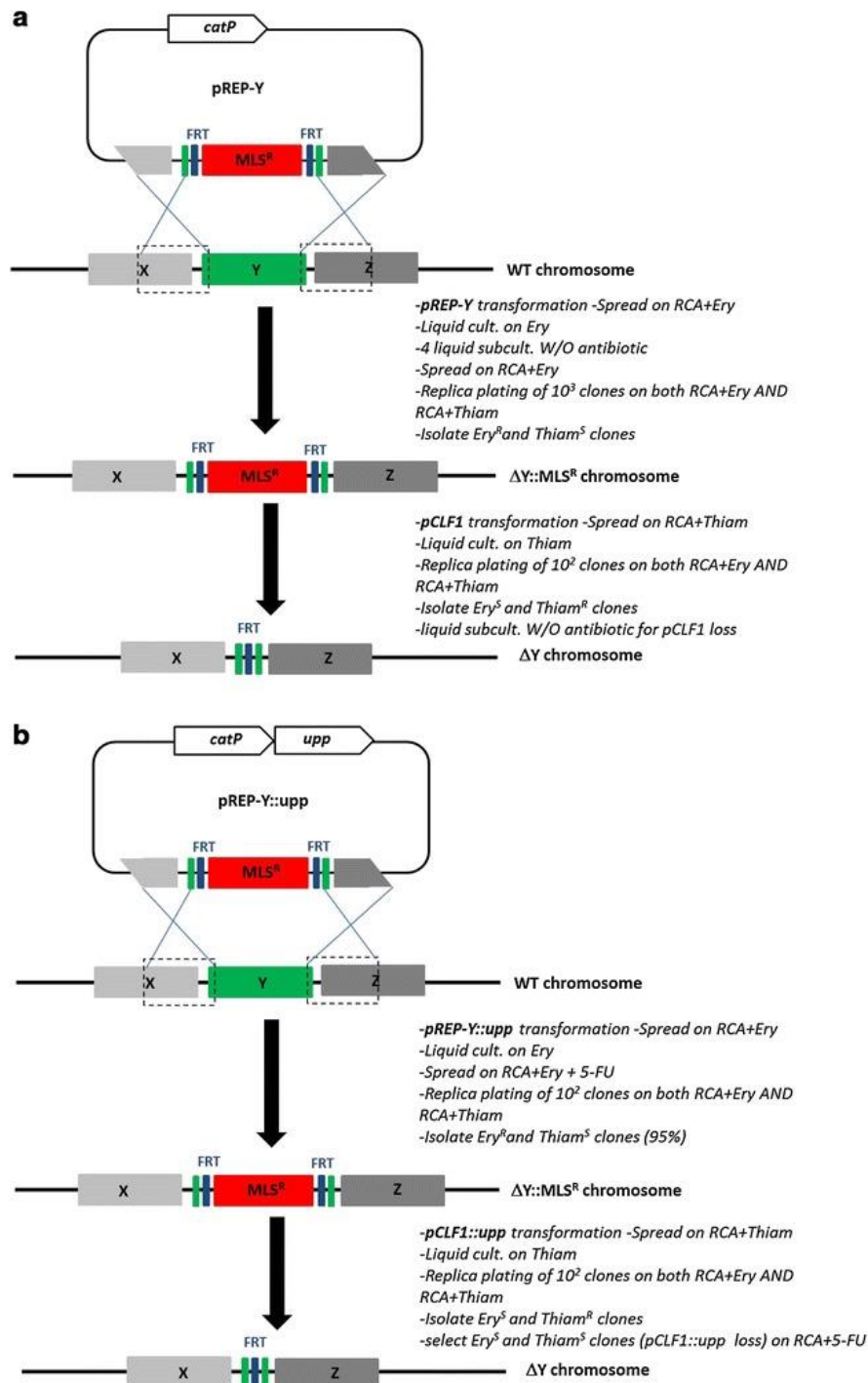
3.3 Results and Discussion

3.3.1 *MGCΔcac1502* strain, a *C. acetobutylicum* strain that is transformable without previous in vivo plasmid methylation

Cac824I, the type II restriction endonuclease encoded by *CA_C1502*, is a major barrier to the electrotransformation of *C. acetobutylicum* with *E. coli*–*C. acetobutylicum* shuttle vectors. The *Cac824I* restriction endonuclease recognition sequence 5'-GCNGC3', where N can be any nucleotide, occurs infrequently in *C. acetobutylicum* DNA because of the high A + T DNA content (72 % A + T), but the sequence occurs frequently in *E. coli* plasmids. No methyltransferase that can be used in vitro to protect DNA from restriction by *Cac824I* is commercially available. Prior to the transformation of *C. acetobutylicum*, shuttle plasmids have to be methylated in vivo by transformation into *E. coli* ER2275 (pAN1) expressing the *Bacillus subtilis* phage ϕ 3TI methyltransferase, which protects the shuttle plasmids from digestion by the Clostridial endonuclease *Cac824I* [142]. This step is time consuming and may be a drawback if the genes to be transferred to *C. acetobutylicum* are toxic when expressed in *E. coli*. Therefore, a *C. acetobutylicum* strain deficient for this particular restriction system would be valuable for efficient electrotransformation without previous treatment of the plasmid to be transformed.

To delete the *Cac824I* encoding gene, the first step is the construction of a shuttle vector carrying the replacement cassette. The *CA_C1502* replacement cassette was cloned into the *Bam*HI site of the pCons2-1 and pCIP2-1 to generate the pREPCac15 and pCIPCac15 plasmids, respectively. The difference between these two plasmids is the origin of replication. The pREPCac15 contains a pIMP13 origin of replication from *B. subtilis* (rolling circle mechanism) functional in *Clostridia*, whereas pCIPCac15 contains the origin of replication of the pSOL1 megaplasmid (θ replication mechanism). The pREPCac15 and pCIPCac15 plasmids

were methylated in vivo in *E.coli* ER2275 (pAN1) and were used to transform *C. acetobutylicum* ATCC824 by electroporation. After selection on plates for clones resistant to erythromycin at 40 µg/ml, one colony of each transformant was cultured for 24 h in liquid SM with erythromycin and was then subcultured four times in liquid 2YTG medium without antibiotic (Figure 3-1a). To select integrants that lost the pREPcac15 or pCIPcac15 plasmids, 10³ erythromycin resistant clones were replica plated on both RCA with erythromycin and RCA with thiamphenicol at 50 µg/ml. Whereas several colonies resistant to erythromycin and sensitive to thiamphenicol were obtained with pREPcac15 transformants, no such colonies were obtained with the pCIPcac15 transformants, which indicates that the θ replication mechanism of pCIPcac15 is less favorable for promoting double-crossover in *C. acetobutylicum* than a rolling circle mechanism. The genotype of clones with the desired phenotype was checked by PCR (polymerase chain reaction) analysis (Figure 3-2a). The $\Delta cac1502::mIs^R$ strain, which had lost the pREPcac15, was isolated. This strain was transformed with the pCLF1 plasmid expressing the *FLP1* gene of *S. cerevisiae* encoding for the FLP recombinase. The expression of *FLP1* was under the control of the promoter and RBS (ribosome binding site) from the thiolase gene from *C. acetobutylicum*. After transformation and selection for resistance to thiamphenicol at 50 µg/ml, one colony was cultured in liquid SM with thiamphenicol. One hundred thiamphenicol resistant clones were replica plated on both RCA with erythromycin and RCA with thiamphenicol. The genotype of the clones with erythromycin sensitivity and thiamphenicol resistance was checked by PCR analysis with primers CAC 0 and CAC 5 (Figure 3-2a). Two successive 24-h liquid cultures of the $\Delta cac1502$ strain were conducted in the absence of antibiotics to remove pCLF1. The $\Delta cac1502$ strain that lost pCLF1 was isolated according to its sensitivity to both erythromycin and thiamphenicol. This strain was called MGC $\Delta cac1502$.



The *boxed regions of X and Z genes* represent approximately the regions of homology incorporated into the replicative plasmid used for the double-crossover event (~ 1 kbp each). **a** Initial strategy used for the construction of the *MGCΔcac1502* and *MGCΔcac1502Δupp* strains, **b** counter-selection strategy with the 5-FU/*upp* system used for the construction of the *MGCΔcac1502ΔuppΔcac3535* strain.

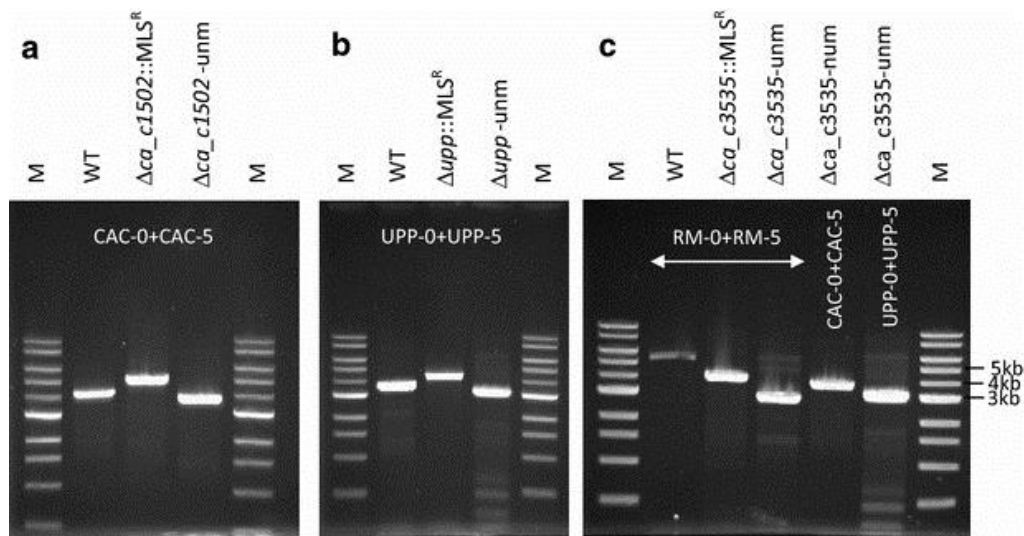


Figure 3-2 Gene replacement via allelic exchange at the *ca_c1502*, *upp* and *ca_c3535* loci.

PCR confirmation of the different double-crossover deletion mutants using external primers annealing to the chromosome upstream and downstream of each deletion cassette used in the different pREP plasmids: **a** *Δca_c1502* deletion mutants with CAC-0 + CAC-5 primers, **b** *Δupp* deletion mutants with UPP-0 + UPP-5 primers, **c** *Δca_c3535* deletion mutants with RM-0 + RM-5 primers (*lanes 1, 2 and 3*). For each experiment, *lanes 2 and 3* refer to before and after excision of the *MLSR* marker by the FLP recombinase, respectively, giving finally an unmarked deletion mutant (*Δ-unm*). The previous unmarked deletions in the final *Δca_c1502 Δupp Δca_c3535* were confirmed with CAC-0 + CAC-5 (**c, lane 4**) and UPP-0 + UPP-5 (**c, lane 5**) primers. *Lane M*, 1 kb DNA ladder (0.5–10 kb) (NEB)

The efficiency of transformation of this strain with methylated and unmethylated pCons2.1 plasmid was evaluated and compared to the wild-type strain. Both strains can be transformed with methylated pCons2.1 with similar efficiency, but only MGCΔ*cac1502* can be transformed efficiently with unmethylated DNA (Table 3-1).

	<i>C. acetobutylicum</i> ATCC824	MGCΔ <i>cac1502</i>
Unmethylated pCons2.1	0	$0.79 (\pm 0.24) \times 10^4$
Methylated pCons2.1	$0.46 (\pm 0.11) \times 10^4$	$0.58 (\pm 0.18) \times 10^4$

Table 3-1 Transformation efficiencies of *C. acetobutylicum* ATCC824 and MGCΔ*cac1502* for unmethylated and methylated pCons2.1.

Values are expressed in number of transformants per μg DNA. Mean values and standard deviations from three independent experiments are given 25 ng pCons2.1 was used in each experiment.

The following deletions described in this manuscript were conducted in this strain without previous in vivo plasmid methylation.

3.3.2 Construction of the MGCΔ*cac1502*Δ*upp* strain: the first marker-free *C. acetobutylicum* strain with two deleted genes

To develop a positive screening of integrants, we used the “*upp*/5-FU as counter selection marker” system. The *C. acetobutylicumupp* gene (CA_C2879) encodes uracil phosphoribosyl-transferase (UPRTase), which catalyzes the conversion of uracil into UMP, thus allowing the cell to use exogenous uracil [143]. The pyrimidine analog 5-fluoro uracil (5-FU) can be converted by UPRTase into 5-fluoro-UMP, which is metabolized into 5-fluoro-dUMP, an inhibitor of thymidylate synthetase, toxic for the cell. The use of the *upp* expression cassette as a counter-selection marker is linked to the construction of a *C. acetobutylicum* strain deleted for the *upp* gene, thus resistant to 5-FU.

To delete *upp*, the *upp* replacement cassette was cloned into the *Bam*HI site of pCons2-1 to generate the plasmid pREPupp. The plasmid pREPupp was used to transform the MGCΔ*cac1502* strain by electroporation without previous in vivo methylation. After

selection on plates for clones resistant to erythromycin at 40 µg/ml, one colony was cultured for 24 h in liquid SM with erythromycin and was then subcultured in liquid 2YTG medium without antibiotic (Figure 3-1a). To select integrants having lost the pREP_{upp} plasmid, 10³ erythromycin resistant clones were replica plated on both RCA with erythromycin and RCA with thiamphenicol. The genotype of the clones resistant to erythromycin and sensitive to thiamphenicol was determined by PCR analysis (Figure 3-2b). The *GCA*Δ*cac1502*Δ*upp*::*mls*^R strain that lost pREP_{upp} was isolated. When the resistance to 5-FU was analyzed, we showed that this strain was resistant to up to 1 mM 5-FU compared to 50 µM for the *MGC*Δ*cac1502* strain. This strain was then transformed with the pCLF1 plasmid, and selection of *MGC*Δ*cac1502*Δ*upp* strain with sensitivity to both erythromycin and thiamphenicol was performed, as previously described for the *MGC*Δ*cac1502* strain (Figure 3-2b).

3.3.3 Deletion of the *CA_C3535* gene in the *MGC*Δ*cac1502*Δ*upp* strain using the *upp*/5-FU system as a counter-selectable marker for the loss of plasmid

The *CA_C3535* gene encodes *Cac824II*, a potentially bi-functional enzyme carrying both a type II restriction endonuclease and methylase activities. To delete *CA_C3535*, the *CA_C3535* replacement cassette was cloned into the *Bam*HI site of the pCons::upp to generate the plasmid pREP*cac3535*::upp. The plasmid pREP*cac3535*::upp was used to transform the *C. acetobutylicum* *MGC*Δ*cac1502*Δ*upp* strain by electroporation without previous in vivo methylation.

After plate selection for clones resistant to erythromycin at 40 µg/ml, 100 transformants were replica plated on RCA with erythromycin, RCA with thiamphenicol and RCA with 5-FU at 400 µM (Figure 3-1b). All transformants were resistant to erythromycin and thiamphenicol and were sensitive to 5-FU compared to the parental strain, which was resistant to 5-FU. This result demonstrates that the expression of the *upp* gene carried by pREP*cac*::upp confers sensitivity to 5-FU.

To select for Δ*cac3535*::*Em*^R integrants that lost the pREP*cac3535*::upp plasmid, erythromycin- and 5-FU-resistant clones were selected on RCA plates containing erythromycin and 5-FU from 100 µl of a liquid culture of the *MGC*Δ*cac1502*Δ*upp* (pREP*cac3535*::upp) strain. Approximately 500 colonies were obtained, and 100 of them were replica plated on both RCA with erythromycin and RCA with thiamphenicol. Most of the clones (95 %) were resistant to erythromycin and sensitive to thiamphenicol. Four clones

were checked by PCR analysis (Figure 3-2c) All four clones had the correct phenotype, and one of the clones was selected as the MGC $\Delta cac1502\Delta upp\Delta cac3535::m1s^R$ strain. This strain was then transformed with pCLF::upp, a derivative of the pCLF1 plasmid that also carries the *upp* gene, in order for the positive selection of plasmid loss after the excision of the *m1s^R* marker. After the first selection of clones resistant to thiamphenicol and sensitive to erythromycin, a second selection of clones resistant to 5-FU and sensitive to thiamphenicol was performed to obtain the MGC $\Delta cac1502\Delta upp\Delta cac3535$ strain that was control by PCR (Figure 3-2c) for the presence of all the marker-less deletions. Finally, when compared to *C. acetobutylicum* ATCC824 wild-type, the growth of the restriction-less marker-less strain in MS medium at pH 4.5 (Figure 3-3) was shown to be unaffected by the different deletions.

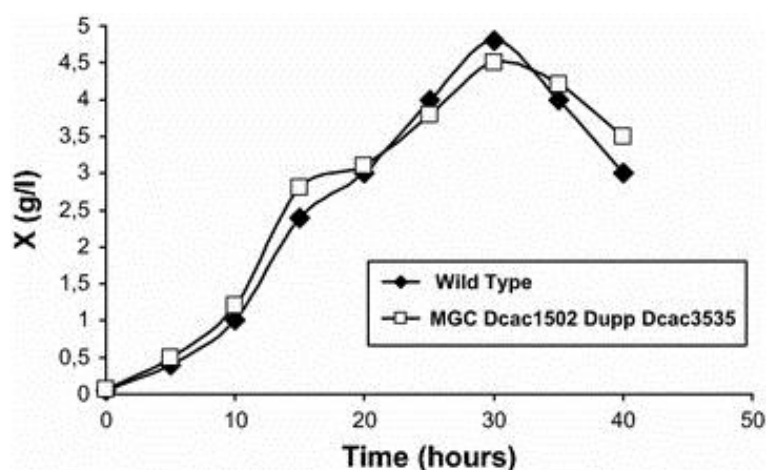


Figure 3-3 Growth curves of *C. acetobutylicum* ATCC824 and MGC $\Delta cac1502$ Δupp $\Delta cac3535$ at pH 4.5 in SM medium

The unmethylated pCons2.1 plasmid was used to evaluate the transformation efficiency of the MGC $\Delta cac1502$ and the MGC $\Delta cac1502\Delta upp$ $\Delta cac3535$ strains. The transformation efficiency of MGC $\Delta cac1502\Delta upp\Delta cac3535$ for unmethylated pCons2.1 was ~eightfold higher than that of MGC $\Delta cac1502$ (Table 3-2)

	MGC $\Delta cac1502$	MGC $\Delta cac1502\Delta upp$ $\Delta cac3535$
Unmethylated pCons2.1	$0.79 (\pm 0.24) \times 10^4$	$6.1 (\pm 3.2) \times 10^4$

Table 3-2 Transformation efficiencies of MGC $\Delta cac1502$ and MGC $\Delta cac1502\Delta upp\Delta cac3535$ for unmethylated pCons2.1.

Values are expressed in number of transformants per μg DNA. Mean values and standard deviations from three independent experiments are given. 25 ng pCons2.1 was used in each experiment.

3.3.4 Determination of the recognition sequence of *Cac824II* encoded by *CA_C3535*

CA_C3535 encoded a 993 amino acid protein with a calculated molecular mass of 116,842 Da. The amino acid sequence analysis revealed high similarities with two restriction endonucleases: *AcuI* from *Acinetobacter calcoaceticus* SRW4 [144] and *Eco57I* from *E. coli* RFL57 [145] with 44 and 46 % identity, respectively. Both enzymes belong to the IIg family of restriction enzymes and possess both a restriction and methylase activity. To heterologously express the *AcuI*-encoding gene in *E. coli* [144], it was necessary to first express the *AcuIM* methylase-encoding gene because the methylase activity of *AcuI* was not sufficient to protect DNA against its restriction activity. We applied the same strategy for the expression of *CA_C3535*-encoding *Cac824II*: we cloned into the pSOS2K2 gene and expressed in *E. coli* the *CA_C3534* gene that encodes a putative methylase and that is located immediately downstream of *CA_C3535* gene in the *C. acetobutylicum* chromosome. The pSC-CAC3534 plasmid expressing *CA_C3534* has three *AcuI* recognition sites, but when we tried to digest it with *AcuI*, it was completely protected from the activity of this enzyme. To express, purify and determine the recognition sequences of *Cac824II*, we cloned *CA_C3535* in the pPAL vector using the *E. coli* BL21-AI cells containing the pSC-CAC3534 plasmid as host. The *Cac824II* endonuclease was purified, and its activity towards unmethylated pCons2.1 in the presence of SAM was determined. *AcuI* recognizes the 5'-CTGAAG-3' sequence and cuts the pCons2.1 plasmid two times, resulting in two fragments of 2411- and 882-bp. Figure 3-4 shows that *Cac824II* gives the same restriction pattern as *AcuI*. To confirm that the *AcuI* and *Cac824II* recognition sequences were identical, pCons2.1 was digested by 50 μg of *Cac824II* in the presence of 1 U of *AcuI*. Figure 3-4 shows that the restriction pattern was unchanged, which definitively confirms that *AcuI* and *Cac824II* are isoschizomers.

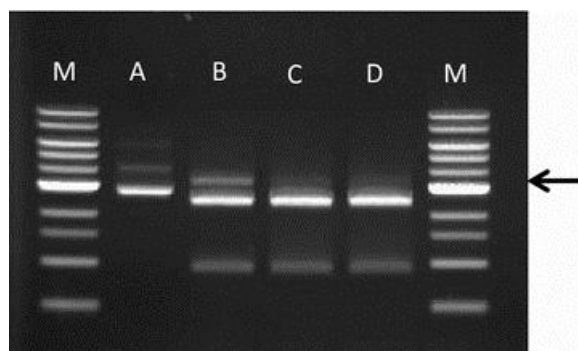


Figure 3-4 Digestion properties of recombinant *Cac824II* as compared to commercial *AcuI* (New England Biolabs).

250 ng of unmethylated pCONS2.1 plasmid were incubated for 1 h at 37 °C in a reaction volume of 20 μ L containing 50 mM potassium acetate, 20 mM Tris–acetate (pH 7.9), 10 mM Magnesium acetate, 100 μ g/mL BSA and 0.04 mM S-adenosyl-methionine with (A) No enzyme, (B) purified *Cac824II* (50 μ g), (C) *AcuI* (5U), and (D) purified *Cac824II* (50 μ g) + *AcuI* (5U). Lanes M, 1 kb DNA ladder (0.5–10 kbp, NEB). Reactions products were electrophoresed on a 0.8 % agarose gel. An *arrow* indicates the incomplete digestion product remaining after incubation with *Cac824II*.

3.3.5 Deletion of the *ctfAB* genes in the MGC Δ *cac1502* Δ *upp* Δ *cac3535* to create a strain no longer producing acetone

The *ctfAB* genes (*CA_P0163*–*CA_P0164*) located on the pSOL1 megaplasmid encodes an acetoacetyl-CoA:acyl CoA-transferase involved in the first specific step of acetone formation [42]. To delete *ctfAB*, the *ctfAB* replacement cassette was cloned into the *Bam*HI site of the pCons::upp to generate the plasmid pREPctfAB::upp. The plasmid pREPctfAB::upp was used to transform the *C. acetobutylicum* MGC Δ *cac1502* Δ *upp* Δ *cac3535* strain by electroporation without previous in vivo methylation and cell containing the plasmid were selected on RCA plate with erythromycin at 40 μ g/ml. To select for Δ *ctfAB*::*Em*^R integrants that lost the pREPctfAB::upp plasmid, erythromycin- and 5-FU-resistant clones were selected on RCA plates containing erythromycin and 5-FU from 100 μ L of a liquid culture of the MGC Δ *cac1502* Δ *upp* Δ *cac3535* (pREPctfAB::upp) strain. Approximately 500 colonies were obtained, and 50 of them were replica plated on both RCA with erythromycin and RCA with

thiamphenicol. Most of the clones (90 %) were resistant to erythromycin and sensitive to thiamphenicol. Four clones were checked by PCR analysis (with primers CTF-0 and CTF-5 located outside of the *ctfAB* replacement cassette and primers CTF-D and CTF-R located inside of *ctfAB*). All four clones had the correct phenotype, and one of the clones was selected as the MGC $\Delta cac1502\Delta upp\Delta cac3535\Delta ctfAB::mIs^R$ strain. The fermentation profile of this strain was compared to the MGC $\Delta cac1502\Delta upp\Delta cac3535$ control strain during batch fermentation at pH 4.5 (Figure 3-5). The production of acetone was totally abolished but the production of acetic acid was increased more than sixfold while butyric acid was only slightly increased, proving that the acetoacetyl-CoA:acetyl CoA-transferase is mainly involved in the consumption of acetic acid.

3.3.6 Deletion of the *ldhA* gene in the MGC $\Delta cac1502\Delta upp\Delta cac3535$ to create a strain no longer producing lactate

The *ldhA* genes (CA_C0267) encodes a lactate dehydrogenase involved in the last step of l-lactate formation [42]. To delete *ldhA*, the *ldhA* replacement cassette was cloned into the *Bam*HI site of the pCons::upp to generate the plasmid pREPldhA::upp. The plasmid pREPldhA::upp was used to transform the *C. acetobutylicum* MGC $\Delta cac1502\Delta upp\Delta cac3535$ strain by electroporation without previous in vivo methylation and cell containing the plasmid were selected on RCA plate with erythromycin at 40 μ g/ml. To select for $\Delta ldhA::Em^R$ integrants that lost the pREPldhA::upp plasmid, erythromycin- and 5-FU-resistant clones were selected on RCA plates containing erythromycin and 5-FU from 100 μ l of a liquid culture of the MGC $\Delta cac1502\Delta upp\Delta cac3535$ (pREPldhA::upp) strain. Approximately 500 colonies were obtained, and 50 of them were replica plated on both RCA with erythromycin and RCA with thiamphenicol. Most of the clones (80 %) were resistant to erythromycin and sensitive to thiamphenicol. Four clones were checked by PCR analysis (with primers LDH-0 and LDH-5 located outside of the *ldhA* replacement cassette and primers LDH-D and LDH-R located inside of *ldhA*). All four clones had the correct phenotype, and one of the clones was selected as the MGC $\Delta cac1502\Delta upp\Delta cac3535\Delta ldhA::mIs^R$ strain. The fermentation profile of this strain was compared to the MGC $\Delta cac1502\Delta upp\Delta cac3535$ control strain during batch fermentation at pH 4.5 (Figure 3-5). The production of l-lactate was totally abolished proving that *ldhA* encodes the main l-lactate dehydrogenase of *C. acetobutylicum*.

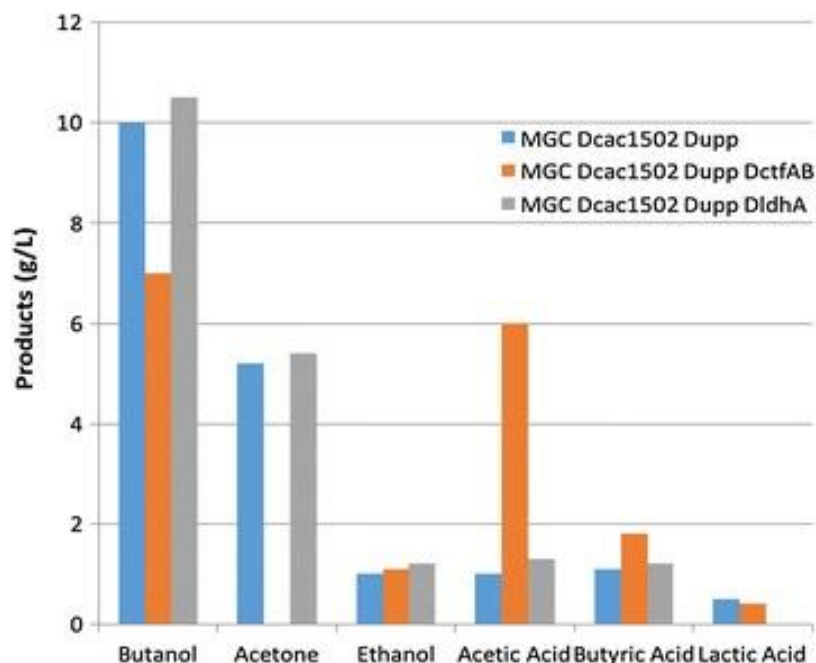


Figure 3-5 Solvent and acid production of MGCΔcac1502 Δupp Δcac3535, MGCΔcac1502 Δupp Δcac3535Δ *ctfAB*::Em^R, and MGCΔcac1502 Δupp Δcac3535Δ *ldhA*::Em^R in batch culture at pH 4.5 in SM medium

3.4 Discussion

We developed a simple and efficient method to create mutations in the *Clostridium acetobutylicum* chromosome. This method is based on the use of (1) a replicative plasmid, (2) a deletion cassette containing both DNA sequences with homology to the flanking region of the target gene (to delete it) and an antibiotic resistance gene surrounded by FRT sequences (as an excisable marker), and (3) the *upp* gene, which encodes the uracil-phosphoribosyl-transferase, as a counter-selectable marker.

A plasmid that replicates via a rolling circle mechanism was more efficient in terms of double cross over frequency than a plasmid that replicates through a theta mechanism. This result is in agreement with previous findings in *Bacillus subtilis* showing that plasmid replication through a rolling circle mechanism favors recombination between homologous sequences [146,147].

The deletion cassette can be rapidly constructed through a three-step procedure using pre-constructed building blocks. After a fusion PCR and TOPO cloning of the product, a predesigned antibiotic resistance gene surrounded by two FRT sites in direct repeats is

inserted. The *upp* gene is located on the plasmid outside of the deletion cassette. This allows the positive selection of clones that have lost the plasmid and integrated the deletion cassette by a double recombination event. We demonstrate here that this event occurs at a frequency of 10^{-5} , which means that without the selection procedure, it would be much more difficult to isolate the correct deletion mutant by replica plating alone. Once the deletion cassette is integrated into the chromosome, the expression of the *flp* recombinase allows (1) the excision of the antibiotic marker for a clean in-frame deletion of the targeted gene (without polar effect) and (2) consecutive gene deletions. Such a strategy was previously applied to marker-less gene deletion in *E. coli* [148] and *Mycobacterium smegmatis* [149]. The plasmid expressing the FLP recombinase-encoding gene was further improved by coexpressing the *upp* gene to use it as a positive selection for the plasmid loss after excision of the *MLS^R* marker. A similar tool was developed by Al-Hinai et al. [94] using a plasmid (1) that expresses the FLP recombinase-encoding gene and (2) that has an inducible segregational instability to promote the plasmid loss.

In this study *Cac824II* (encoded by *CA_C3535*), the second type II restriction enzyme of *C. acetobutylicum* predicted by REBASE [100], was biochemically characterized and it was demonstrated that it is an isoschizomer of *AcuI* [144] recognizing the 5'-CTGAAG-3' sequence. It was also shown that *Cac824II* methylase (encoded by *CA_C3534*) protects DNA against restriction by *Cac824II* and *AcuI* by probably methylating one of the adenine in the 5'-CTGAAG-3' sequence. Two *Cac824II* restriction sites are present (in the *ampR* gene and in the *colE1* origin of replication) in most the shuttle vector use to transform *C. acetobutylicum* and it was then justified to construct a marker-less strain deleted from *CA_C3535*. The transformation efficiency of *MGCΔcac1502ΔuppΔcac3535* for unmethylated pCons2.1 was much higher (~eightfold higher) than that of *MGCΔcac1502* and it will be an interesting strain to develop new genetic tools based on suicide vectors [100].

The restriction-less marker-less strain and the method was successfully used to delete two genes (*ctfAB*) on the pSOL1 megaplasmid and one gene (*ldhA*) on the chromosome to get strains no longer producing acetone or l-lactate. This work demonstrate that (1) *ctfAB* encode an acetoacetyl-CoA:acetate CoA-transferase that coupled acetone formation to acetate consumption and (2) *ldhA* encodes the main lactate dehydrogenase in *C. acetobutylicum* although a second gene *ldhB* (*CA_C3552*) is also present [42]. A strain with clostron

inactivated *ctfAB* genes was previously constructed [91]. From the physiological analysis of this mutant and with the help of a mathematical model [150], it was demonstrated that butyrate was mainly reconsumed by the phosphotransbutyrylase-butyrate-kinase pathway and not by the acetoacetyl-CoA:acetate CoA-transferase in agreement with the data presented in our study.

3.5 Conclusion

The restriction-less, marker-less strain and the genome modification method presented here become simple and convenient tools that are useful for research groups involved in functional genomic studies of *C. acetobutylicum* and for further metabolic engineering of this strain to produce bulk chemicals and biofuel. As a demonstration of the efficiency of the method, we constructed two strains unable to produce l-lactate or acetone. Furthermore, this method was successfully used by the Metabolic Explorer Company to develop and patent an industrial recombinant strain of *C. acetobutylicum* for *n*-butanol production [151] at high yield.

3.6 Experimental Procedure

3.6.1 Bacterial strain, plasmids and oligonucleotides

The bacterial strain and plasmids used in this study are listed in Table 3-3. The specific oligonucleotides used for PCR amplification were synthesized by Eurogentec Table 3-4.

Strain or plasmid	Relevant characteristics ^a	Source or reference ^b
Bacterial strains		
<i>E. coli</i>		
TOP10		Invitrogen
ER2275	RecA ⁻ McrBC ⁻	NEB
<i>C. acetobutylicum</i>		
ATCC824	Wild-type	ATCC
MGCΔ <i>cac1502</i>	Δ <i>CA_C1502</i>	This study
MGCΔ <i>cac1502</i> Δ <i>upp</i>	Δ <i>CA_C 1502</i> Δ <i>CA_C 2879</i>	This study
MGCΔ <i>cac1502</i> Δ <i>upp</i> Δ <i>cac3535</i>	Δ <i>CA_C 1502</i> Δ <i>CA_C 2879</i> Δ <i>CA_C 3535</i>	This study
MGCΔ <i>cac1502</i> Δ <i>upp</i> Δ <i>cac3535</i> Δ <i>ctfAB</i>	Δ <i>CA_C 1502</i> Δ <i>CA_C 2879</i> Δ <i>CA_C 3535</i> Δ <i>CA_P0162-3</i>	This study
MGCΔ <i>cac1502</i> Δ <i>upp</i> Δ <i>cac3535</i> Δ <i>ldhA</i>	Δ <i>CA_C 1502</i> Δ <i>CA_C 2879</i> Δ <i>CA_C 3535</i> Δ <i>CA_C 0267</i>	This study
Plasmids		
pAN1	Cm ^r , φ3TI, p15A origin	[152]
pKD4	Ap ^r Km ^r	[148]
pETSPO	Cm ^r MLS ^r	[92]
pUC18	Ap ^r	Fermentas
pUC18-FRT-MLS2	Ap ^r MLS ^r	This study
pCons2-1	Cm ^r	This study
pCR-BluntII-TOPO	Zeo ^r Km ^r	Invitrogen
pCIP2-1	Cm ^r	This study
pREP <i>cac15</i>	Cm ^r MLS ^r Δ <i>CA_C1502</i>	This study

pCIPcac15	Cm ^r MLS ^r Δ CA_C1502	This study
pREPupp	Cm ^r MLS ^r Δ upp	This study
pCP20	Ap ^r Cm ^r <i>FLP</i>	[153]
pSOS95	Ap ^r MLS ^r , acetone operon, <i>repL</i> gene, ColE1 origin	[154]
pCLF1	Cm ^r <i>FLP</i>	This study
pCR4-TOPO-Blunt	Ap ^r Km ^r	Invitrogen
pCons::upp	Cm ^r MLS ^r <i>upp</i>	This study
pREPCAC3535::upp	Cm ^r MLS ^r <i>upp</i> Δ CA_C3535	This study
pREPctfAB::upp	Cm ^r MLS ^r <i>upp</i> Δ ctfAB	This study
pREPl dhA::upp	Cm ^r MLS ^r <i>upp</i> Δ ldhA	This study

Table 3-3 Bacterial strains and plasmids used in this study

^a *RecA*[−] homologous recombination abolished, *McrBC*[−] lacking methylcytosine-specific restriction system, *Cm*^r chloramphenicol resistance, *Ap*^r ampicillin resistance, *MLS*^r macrolide lincosamide and streptogramin B resistance, *Zeo*^r zeomycin resistance, ϕ 3TI ϕ 3TI methyltransferase, *repL* Gram-positive origin of replication from pIM13

^b *NEB* New England BioLabs, *ATCC* American Type Culture Collection (Rockville, MD)

Primer name	Oligonucleotide sequence
PKD4.1	ctggcgcctgagtgcttgcggcagcgtgagggg
PKD4.2	agcccggggatctcatgctggagttcttcgcc
FRT-MLSR-F	tacaggccttgagcgattgtgtaggctggagc
FRT-MLSR-R	aacaggcctgggatgtaacgcactgagaagccc
PCONSAccI	ccgggggtaccgtcgacctgcagcc
PCONSEcoRI	gaattccgcgagctcggtagccggc
ORI3-D	ccatcgatgggggtcatgcatcaatactatcccc
ORI4-R	gcttcctgttttaatacctttcgg
FLP1-D	aaaaggatccaaaaggagggtataaaatgccacaatttggtatattatgtaaacaccacct
FLP1-R	aaatggcgccgcgtacttatatgcgtctatttatgtaggatgaaaggta
REP-UPP-F	aaaacagctgggaggaatgaaataatgagtaaagttacac
REP-UPP-R	aaaacagctgtattttgtaccgaataatctatctccagc
CAC 1	aaaggatccatgcacactcataaatttactgtaggaagtctg
CAC 2	ggggaggcctaaaaaggggggtcccaaataatattgccatagtaaccacc
CAC 3	cccccttttaggcctcccctcgaacttattagaatgattaagattccgg
CAC 4	aaaggatcctcattaaatttctccattttaagcctgtc
CAC 0	gtgatataatttctttaaatggaggaggatctg
CAC 5	gccgttaatagacattataattccattggc
CAC-D	gaattcttaaaaatatttgatcattaagcgg
CAC-R	gttgattggaatctttgttattatttctccc
UPP 1	aaaaggatcctcctgatctattaattcttgatgaaccc
UPP 2	ggggaggcctaaaaagggggattgcataataaaaagggtgaaaaataaatttcag
UPP 3	cccccttttaggcctcccctatttctcctccattgtattttttctatttg

UPP 4	aaaaggatccgctattatgaataggtaaataagtcagctgg
UPP 0	aatacaagcaaagagaataggctatgtgcc
UPP 5	aatacaagcaaagagaataggctatgtgcc
UPP-D	ggcatatgaagtaacaagagaaatgcagc
UPP-R	ataatctatctccagcatctccaagacc
RM3535 1	aaaaggatccgcagctttctggaaggactacggcg
RM3535 2	ggggaggcctaaaaagggggcatttacttatggtacggtcacccc
RM3535 3	cccccttttaggcctccccgtcttataaaagtaatttatcaaaggcatcaaggc
RM3535 4	aaaaggatccctaactctctaaacgttacaatagtaatgcgc
RM3535 0	cacattgtcatttataaaagtcctaggg
RM3535 5	gtagtaattccaacttcaactcttgccac
RM3535-D	cttagaatagctgatattgcttgccg
RM3535-R	agcatctctttaatgattctccgg
CTF1	aaaaggatcccagacactataatagctttaggtggtacccc
CTF2	ggggaggcctaaaaagggggattataaaaagtagttgaaatatgaaggtttaaggttg
CTF3	cccccttttaggcctccccatatccaatgaacttagacctagggctg
CTF4	aaaaggatccgtgttataatgtaaataataaataaggactagaggcg
CTF0	taccacetttttcacgcttggtgcgg
CTF5	tatttaaagaggcattatcaccagagcg
LDH1	aaaaggatccgcttataaaatttgaaagaggaagttgtg
LDH2	ggggaggcctaaaaagggggtagaaatcttataaaattctctatagagcccatc
LDH3	cccccttttaggcctccccggtataaagacctaactccaagggtggaggctaggtc
LDH4	aaaaggatccccattgtggagaatattccaaagaagaaaataattgc
LDH0	cagaaggcaagaatgtattaagcggaaatgc

LDH5	cttcccattatagctcttattcacattaagc
Cac3535-d- <i>SpeI</i>	aaaactagtatgaatgatattaaaatagctttgaaaaaattggtgac
Cac3535-R- <i>BamHI</i>	aaaaggatccctacaaattatatatatctgttaccatgcctc

Table 3-4 Oligonucleotides used for PCR amplifications.

3.6.2 Culture and growth conditions

C. acetobutylicum was maintained as spores in synthetic medium (SM) as previously described [22,155]. Spores were activated by heat treatment at 80 °C for 15 min. All *C. acetobutylicum* strains were grown in anaerobic conditions at 37 °C in SM, in *Clostridium* growth medium (CGM) [34] in 2YTG [152] or in reinforced clostridial medium (RCM) (Fluka). Solid media were obtained by adding 1.5 % agar to the liquid media. Media were supplemented, when required, with the appropriate antibiotic in the following concentrations: for *C. acetobutylicum*, erythromycin at 40 µg/ml and thiamphenicol at 50 µg/ml; for *E. coli*, erythromycin at 200 µg/ml and chloramphenicol at 30 µg/ml. Transformations of *C. acetobutylicum* were conducted by electroporation, as previously described [142]. 5-FU was purchased from Sigma, and stock solutions were prepared in DMSO (dimethyl sulfoxide).

3.6.3 DNA manipulation techniques

Total genomic DNA from *C. acetobutylicum* was isolated as previously described [152]. Plasmid DNA was extracted from *E. coli* with the QIAprep kit (Qiagen, France). Pfu DNA Polymerase (Roche) was used to generate PCR products for cloning, and Taq Polymerase (New England BioLabs) was used for screening colonies by PCR with standard PCR protocols employed for all reactions. DNA restriction and cloning were performed according to standard procedures [156]. Restriction enzymes and Quick T4 DNA ligase were obtained from New England BioLabs (Beverly, MA) and were used according to the manufacturer's instructions. DNA fragments were purified from agarose gels with the QIAquick gel purification kit (Qiagen, France).

3.6.4 Construction of pUC18-FRT-MLS2

Inverse PCR was performed using the pKD4 plasmid [148] as a template and oligonucleotides PKD4.1 and PKD4.2 as primers to amplify the plasmid region with the FRT sites but without the kanamycin resistance marker. This blunt end fragment was later ligated to the *MLS^r* gene obtained after a *HindIII* digestion of the pETSPO plasmid [92] and Klenow

treatment. The corresponding plasmid (pKD4-Ery1) was then used as a template to amplify by PCR the macrolide lincosamide streptogramin B resistance (*MLS^r*) gene, functional in *Clostridia* and flanked by two FRT sites and two *StuI* sites, using the oligonucleotides FRT-MLSR-F and FRT-MLSR-R as primers. This fragment was directly cloned into the *SmaI* digested pUC18 to generate the pUC18-FRT-MLS2 plasmid.

3.6.5 Construction of pCons2.1

Inverse PCR was performed using the pETSPO plasmid [92] as a template and oligonucleotides PCONS*AccI* (mutating a *BamHI* site) and PCONS*EcoRI* as primers. The PCR product, containing a pIMP13 *B. subtilis* origin of replication functional in *Clostridia* (rolling circle mechanism of replication) and a *catP* gene conferring resistance to thiamphenicol was phosphorylated and ligated to yield the pCons0 plasmid. This plasmid was then digested with *BamHI* to remove the *spoA* cassette, and the DNA fragment was purified and ligated to generate the pCons2-1 plasmid.

3.6.6 Construction of pCIP2-1

The pIMP13 origin of replication from pCons2-1 was replaced by the origin of replication of the pSOL1 megaplasmid. The origin of replication of pSOL1 was amplified by PCR using *C. acetobutylicum* total DNA as a template and oligonucleotides ORI3-D and ORI3-R as primers. This PCR product was cloned into the pCR-BluntII-TOPO vector, and the resulting plasmid was digested by *EcoRI* to obtain the 2.2 kb *EcoRI* fragment containing the origin of replication of pSOL1. The pCons2-1 plasmid was digested by *EcoRI*, and the 2.4 kb fragment was ligated to the 2.2 kb *EcoRI* fragment to generate the plasmid pCIP2-1.

3.6.7 Construction of pREP*cac15*

Two DNA fragments surrounding *cac1502* were amplified by PCR using *C. acetobutylicum* total DNA as the template and two pairs of oligonucleotides as primers. Using the primers pairs CAC 1 and CAC 2 or CAC 3 and CAC 4, 1493 and 999-bp DNA fragments were obtained, respectively. Both primers CAC 1 and CAC 4 introduce a *BamHI* site, whereas primers CAC 2 and CAC 3 have complementary 5' extended sequences that introduce a *StuI* site. DNA fragments CAC 1–CAC 2 and CAC 3–CAC 4 were joined in a PCR fusion with primers CAC 1 and CAC 4, and the resulting fragment was cloned into the pCR4-TOPO-Blunt vector to generate pTOPO::*cac15*. At the unique *StuI* site of pTOPO::*cac15*, the 1372-

bp *StuI* fragment of pUC18-FRT-MLS2 carrying the antibiotic resistance *MLS^r* gene with FRT sequences on both sides was introduced. The *cac1502* replacement cassette obtained after *BamHI* digestion of the resulting plasmid was cloned into the *BamHI* site of the pCons2-1 to generate the plasmid pREPcac15.

3.6.8 Construction of pCIPcac15

The *cac1502* replacement cassette above was cloned into the *BamHI* site of the pCIP2-1 to generate the plasmid pCIPcac15.

3.6.9 Construction of pREPupp

Two DNA fragments upstream and downstream of *cac2879* were amplified by PCR using total DNA from *C. acetobutylicum* as the template and two pairs of oligonucleotides as primers. With the primer pairs UPP 1–UPP 2 and UPP 3–UPP 4, 1103- and 1105-bp DNA fragments were obtained, respectively. Both primers UPP 1 and UPP 4 introduce a *BamHI* site, whereas primers UPP 2 and UPP 3 have 5' extended sequences that introduce a *StuI* site. DNA fragments UPP 1–UPP 2 and UPP 3–UPP 4 were joined in a PCR fusion with primers UPP 1 and UPP 4, and the resulting fragment was cloned into pCR4-TOPO-Blunt vector to generate pTOPO::upp. At the unique *StuI* site of pTOPO::upp, the 1372-bp *StuI* fragment of pUC18-FRT-MLS2 carrying the antibiotic resistance *MLS^r* gene with FRT sequences on both sides was introduced. The *upp* replacement cassette obtained after *BamHI* digestion of the resulting plasmid was cloned into the *BamHI* site of the pCons2-1 to generate the plasmid pREPupp.

3.6.10 Construction of pCLF1

The *FLP1* gene was amplified by PCR using the pCP20 plasmid [153] as a template and oligonucleotides FLP1-D and FLP1-R as primers. These primers introduced *BamHI* and *SfoI* restriction sites on the ends of the PCR product. After a *BamHI*–*SfoI* double digestion, the PCR product was cloned into the *BamHI*–*SfoI* sites of the pSOS95 expression vector to generate the pEX-FLP1 plasmid. The 1585-bp *SalI* fragment of pEX-FLP1 containing the *FLP1* expression cassette was cloned into the *SalI* site of pCons2-1 to generate the pCLF1 plasmid.

3.6.11 Construction of pCons::upp

The *upp* gene with its own ribosome binding site (RBS) was amplified by PCR from *C. acetobutylicum* total DNA with the oligonucleotides REP-UPP-F and REP-UPP-R as primers. The 664-bp PCR product was digested by *Pvu*II and was cloned into pCons2.1, digested by *Bcg*I and treated with T4 DNA polymerase to generate the pCons::*upp* plasmid. In this way, the *upp* gene was located just downstream of the *catP* gene to construct an artificial operon with *upp* expressed under the control of the *catP* promoter.

3.6.12 Construction of pREPcac35::upp

Two DNA fragments upstream and downstream of *CA_C3535* were amplified by PCR using the total DNA from *C. acetobutylicum* as a template and two pairs of oligonucleotides as primers. With the primer pairs RM3535 1 and RM3535 2 or RM3535 3 and RM3535 4, 1044- and 938-bp DNA fragments were obtained, respectively. Both primers RM3535 1 and RM3535 4 introduce a *Bam*HI site, whereas primers RM3535 2 and RM3535 3 have 5' extended sequences that introduce a *Stu*I site. DNA fragments RM3535 1-RM3535 2 and RM3535 3-RM3535 4 were joined in a PCR fusion with primers RM3535 1 and RM3535 4, and the resulting fragment was cloned into the pCR4-TOPO-Blunt vector to generate pTOPO::cac3535. At the unique *Stu*I site of pTOPO::cac3535, the 1372-bp *Stu*I fragment of pUC18-FRT-MLS2 carrying the antibiotic resistance *MLS*^r gene with FRT sequences on both sides was introduced. The *CA_C3535* replacement cassette obtained after *Bam*HI digestion of the resulting plasmid was cloned into the *Bam*HI site of the pCons::*upp* to generate the plasmid pREPcac3535::*upp*.

3.6.13 Construction of pREPctfAB::upp

Two DNA fragments upstream and downstream of *ctfAB* (*CA_P0162-CA_P0163*) were amplified by PCR using the total DNA from *C. acetobutylicum* as a template and two pairs of oligonucleotides as primers. With the primer pairs CTF 1 and CTF 2 or CTF 3 and CTF 4, 1144- and 1138-bp DNA fragments were obtained, respectively. Both primers CTF 1 and CTF 4 introduce a *Bam*HI site, whereas primers CTF 2 and CTF 3 have 5' extended sequences that introduce a *Stu*I site. DNA fragments CTF 1-CTF 2 and CTF 3-CTF 4 were joined in a PCR fusion with primers CTF 1 and CTF 4, and the resulting fragment was cloned into the pCR4-TOPO-Blunt vector to generate pTOPO::ctfAB. At the unique *Stu*I site of pTOPO::ctfAB, the 1372-bp *Stu*I fragment of pUC18-FRT-MLS2 carrying the antibiotic

resistance *MLS*^r gene with FRT sequences on both sides was introduced. The *ldhA* replacement cassette obtained after *Bam*HI digestion of the resulting plasmid was cloned into the *Bam*HI site of the pCons::upp to generate the plasmid pREPctfAB::upp.

3.6.14 Construction of pREPldhA::upp

Two DNA fragments upstream and downstream of *ldhA* (CA_C0267) were amplified by PCR using the total DNA from *C. acetobutylicum* as a template and two pairs of oligonucleotides as primers. With the primer pairs LDH 1 and LDH 2 or LDH 3 and LDH 4, 1135- and 1161-bp DNA fragments were obtained, respectively. Both primers LDH 1 and LDH 4 introduce a *Bam*HI site, whereas primers LDH 2 and LDH 3 have 5' extended sequences that introduce a *Stu*I site. DNA fragments LDH 1-LDH 2 and LDH 3-LDH 4 were joined in a PCR fusion with primers LDH 1 and LDH 4, and the resulting fragment was cloned into the pCR4-TOPO-Blunt vector to generate pTOPO::ldhA. At the unique *Stu*I site of pTOPO::ldhA, the 1372-bp *Stu*I fragment of pUC18-FRT-MLS2 carrying the antibiotic resistance *MLS*^r gene with FRT sequences on both sides was introduced. The *ldhA* replacement cassette obtained after *Bam*HI digestion of the resulting plasmid was cloned into the *Bam*HI site of the pCons::upp to generate the plasmid pREPldhA::upp.

3.6.15 Construction of pCLF::upp

The 1585-bp *Sal*I fragment of pEX-FLP1 containing the *FLP1* expression cassette was cloned into the *Sal*I site of pCons::upp to generate the pCLF::upp plasmid.

3.6.15.1 *Cac3535* expression and purification

For the general cloning methods of restriction endonuclease genes in *E. coli*, the first step to clone and express the recombinant CA_C3535 gene into *E. coli* was to pre-protect the host genomic DNA against the restriction activity of the *Cac3535* bi-functional enzyme. The CA_C3534 methylase-encoding gene was thus amplified by PCR with Phusion DNA polymerase using *C. acetobutylicum* ATCC824 total genomic DNA as the template and *Cac3534*-d-*Age*I and *Cac3534*-R-*Pvu*I as primers. After digestion with *Age*I and *Pvu*I, the resulting 1748-bp fragment was then cloned into pAH105 [157] a pSC101 derivative, that has been previously digested with *Age*I and *Pac*I, resulting in the pSC-CAC3534 plasmid. In this construct, the CA_C3534 gene expression was placed under the control of the pGI 1.6 promoter [158].

The *E. coli* BL21-AI strain (Invitrogen) was then transformed by the pSC-CAC3534 plasmid to give the BL21-AI-3534 strain. This strain, with host genomic DNA protected against the restriction activity of the *Cac3535* bi-functional enzyme, was finally used as the host strain for the *CA_C3535* gene over-expression using the T7-based expression system (see below). The *Cac3535* protein was expressed in *E. coli* BL21 AI-3534 and was purified using the Profinity eXact Protein Purification System, following the recommendations of the manufacturer (Biorad). The *CA_C3535* gene was amplified by PCR with Phusion DNA polymerase using *C. acetobutylicum* ATCC824 total gDNA as the template and *Cac3535*-d-*SpeI* and *Cac3535*-R-BamHI as primers. The resulting 3002 bp fragment was cloned into the Zero Blunt TOPO vector (Invitrogen) to generate the TOPO-CAC3535 plasmid. After verification by DNA sequencing, the 2988-bp *SpeI*-*BamHI* fragment from the latter plasmid was then introduced into the pPAL7 vector previously digested with the same enzymes to give the final pPAI-3535-I_2.4 plasmid.

After transformation, *E. coli* BL21-AI-3534 cells harboring the pPAI-3535-I_2.4 plasmid were grown in TB medium in the presence of 50 µg/ml carbenicillin and 100 µg/ml Spectinomycin at 37 °C to an OD550 ~ 0.45 and were then induced with 500 µM IPTG for 4 h at 37 °C. After centrifugation, the cell lysate was obtained by sonicating the resuspended pellet in bind/wash buffer (0.1 M sodium phosphate buffer, pH 7.2).

The tag-free *Cac3535* protein was prepared using the Profinity eXact protein purification system, according to the standard protocol. After the Profinity Exact mini-spin column was bound by the protein and washed, the proteolytic activity of the affinity matrix was activated by applying two column volumes of room temperature 0.1 M sodium phosphate buffer, pH 7.2, containing 0.1 M sodium fluoride. The column was incubated for 30 min to allow for the cleavage of the tag from the protein; then, the tag-free protein was released from the mini-spin column by centrifugation. The tag-free *Cac3535* purified protein retains a Thr-Ser linker at its N-Terminus, ensuring optimal binding and cleavage during the purification steps (“Imprecise Fusion protein”).

Chapter 4 The Weizmann process revisited for the continuous production of n-butanol

Ngoc-Phuong-Thao Nguyen^{1,2,3}, Celine Raynaud⁴, Isabelle Meynial-Salles^{1,2,3} & Philippe Soucaille^{1,2,3,4} *

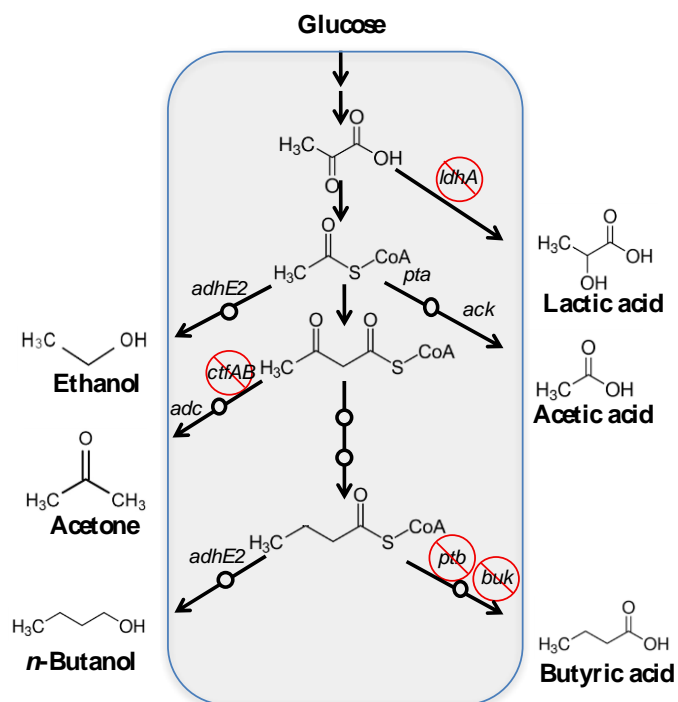
¹Université de Toulouse, INSA, UPS, INP, LISBP, Toulouse, France; ²INRA, UMR792, Toulouse, France; ³CNRS, UMR5504, Toulouse, France; ⁴Metabolic Explorer, Biopôle Clermont-Limagne, Saint Beauzire, France

4.1 Abstract

The industrial production, by *Clostridium acetobutylicum*, of *n*-butanol for use as both a chemical and an alternative transportation fuel, is not currently economical due to low yield, titer and productivity. We used an advanced metabolic engineering approach to engineer *C. acetobutylicum* to produce *n*-butanol from glucose at a high yield: the deletion of genes encoding unwanted pathways was combined to the debottlenecking of the *n*-butanol pathway for maximizing both the yield of alcohol production and the *n*-butanol to ethanol ratio. We also designed a new continuous fermentation process using i) *in situ* extraction of alcohols by distillation under low pressure and ii) high cell density cultures to increase the titer, yield and productivity of *n*-butanol production to levels that have never been previously achieved in any organism. This process provides a means to produce *n*-butanol at performance levels that are now compatible with a commercial process.

One Sentence Summary: Advanced metabolic engineering and process design were combined to develop an economical bioprocess to produce *n*-butanol at high yield, titer and productivity.

(A)



(B)

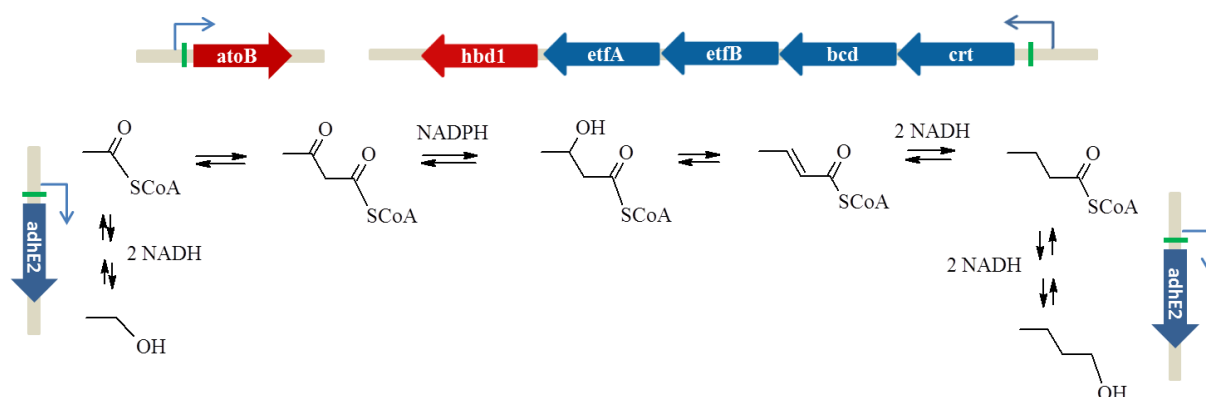


Figure 4-1 Optimization of *n*-butanol production in *Clostridium acetobutylicum*

(A) Elimination of by-products formation (strain CAB1057) by gene deletion. *ldhA*, lactate dehydrogenase; *ctfAB*, acetoacetyl-CoA transferase; *adc*, acetoacetate decarboxylase; *ptb*, phosphotransbutyrylase; *buk*, butyrylate kinase; *adhE2*, bifunctional aldehyde and alcohol dehydrogenase. (B) Optimization of acetyl-CoA conversion to *n*-butanol (Strain CAB1060) using a chimeric pathway derived from three different microorganisms. Green boxes are Rex binding sites. *atoB*, acetoacetyl-CoA thiolase/synthase from *E. coli*; *hbd1*,

NADP⁺-dependent 3-hydroxybutyryl-CoA dehydrogenase from *C. kluyveri*; *ldhA*, L-lactate dehydrogenase; *crt*, crotonase; *bcd-etfAB*, butyryl-CoA dehydrogenase.

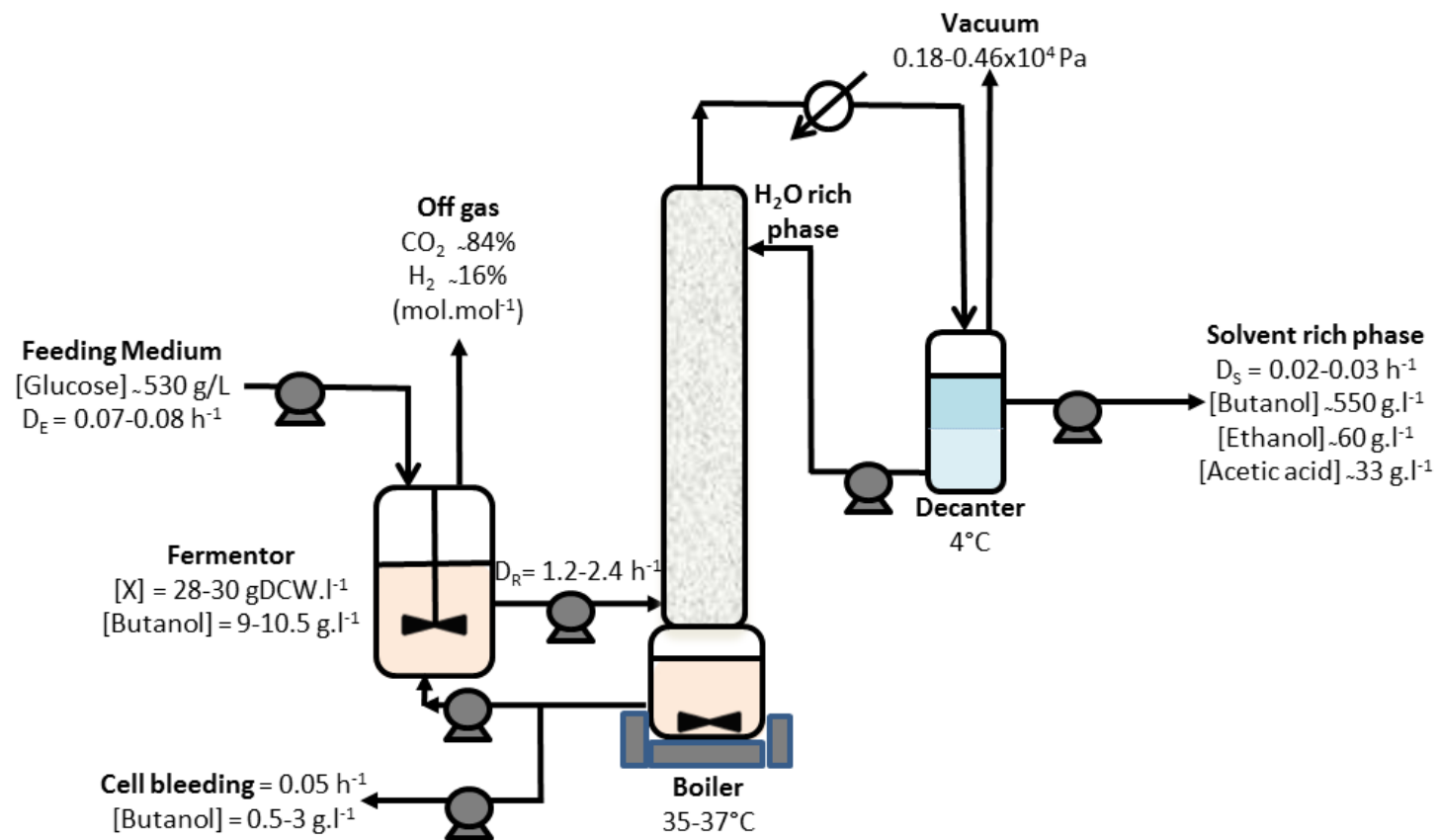


Figure 4-2 Schematic representation of the extractive continuous fermentation process

4.2 Main text

The Weizmann process for acetone and butanol production by *Clostridium acetobutylicum* was the second largest fermentation process (after ethanol) of enormous industrial, social and historical importance [2,159,160]. Beyond its use during the First World War to produce acetone for smokeless gunpowder (cordite) manufacturing, it became a process used worldwide to produce these two industrial solvents from a variety of renewable substrates [2,8,15,22]. Its demise in the early 1960s was the result of the superior petrochemical-process economics due to the low yield, titer and productivity of the fermentation [161] and the inability to use continuous process technologies (as in the case of petrochemical processes) due to the loss of *C. acetobutylicum*'s capacity to produce solvents, known as « degeneration » [69,162]. Today, there is a resurgence of interest in *C. acetobutylicum* for the production of i) fuels after chemical transformation of the solvents mixture [163,164] or ii) n-butanol as the main product for use both as a chemical and an alternative petrol fuel [165,166]. However, to make this biological production economical, all of the limitations of the Weizmann process would have to be resolved as follows: the strain should primarily produce n-butanol at a yield close to the theoretical yield (because the wild-type strain can produce n-butanol and acetone but also ethanol, acetate, butyrate or lactate), concentrated sugar solutions should be fermented to limit waste water treatment costs and the process should be continuous with productivity at least equal to that of the best corn wet milling ethanol plants using yeast as a biocatalyst.

Here, we used an advanced metabolic engineering approach to construct a stable strain for the continuous production of n-butanol at high practical yield. In the *CAB1057* strain, we have deleted all of the genes coding for unwanted metabolic pathways (Figure 4-1A), namely *ptb* and *buk* that code for the phospho-transbutyrylase and the butyrate kinase involved in the last two steps of butyrate formation [49–51], *ctfAB* that code for the acetoacetyl-CoA transferase involved in the first specific step of acetone formation [167] and *ldhA* that codes for the lactate dehydrogenase involved in the last step of L-lactate formation [167]. When *CAB1057* was evaluated in chemostat culture, the expected phenotype was obtained for the deleted pathways but the butanol yield was not improved (Figure 4-3) as this mutant produces more acetate and ethanol than the wild-type strain. The redox-sensing transcriptional repressor (Rex) has been found to modulate its DNA-binding activity in response to the NADH/NAD⁺ ratio and to repress the expression of several operons including *thlA*, *crt-bcd*-

etfAB-hbd, and *adhE2* [110,134]. As *thlA* and *crt-bcd-etfAB-hbd* code for the enzymes converting acetyl-CoA to butyryl-CoA and *adhE2* codes for the enzyme that reduces butyryl-CoA and acetyl-CoA to butanol and ethanol, respectively, the *rexA* gene, encoding Rex, was deleted in *CAB1057* (to yield *CAB1058*) with the objective of improving the C4/C2 ratio and the flux of alcohol formation. The yield of alcohol formation by *CAB1058* was significantly improved, associated with a four-fold decrease of the hydrogen yield as the electrons from reduced ferredoxin were mainly used for NAD(P)H formation (Figure 4-4). However, the higher alcohol yield was due only to an increase in the ethanol yield as the butanol yield remained almost unchanged and, consequently, the C4/C2 ratio was not improved. When the thiolase (catalyzed by ThlA) and 3-hydroxy-butyryl-CoA dehydrogenase (catalyzed by Hbd) activities were measured in *CAB1057* and *CAB1058*, ThlA remained unchanged while Hbd was 1.5 fold higher, indicating that the *rexA* deletion has the expected effect on the expression of the *crt-bcd-etfAB-hbd* operon coding for the enzymes that convert acetoacetyl-CoA to butyryl-CoA (Figure 4-5) but not on *thlA*. ThlA is one of the most abundant intracellular protein [42] representing almost 5% of the cytosolic protein. Furthermore, the condensation reaction by ThlA was shown to be very sensitive to CoA-SH, with micro molar levels totally inhibiting the enzyme [34] and it was also found to be regulated by a redox-switch through a reversible disulfide bond that is formed between two catalytic cysteine residues, Cys88 and Cys378 [36]. As it might be difficult to further overexpressed ThlA and as its biochemical properties might also contributes to the low C4/C2 ratio in *CAB1058* (Figure 4-3), ThlA was replaced by AtoB (Figure 4-1B), a thiolase from *E. coli* that has a higher catalytic efficiency, is less sensitive to CoA-SH [34,168,169] and not subject to a redox-switch. For this purpose, a synthetic *atoB* gene was designed, codon harmonized [170,171] for *C. acetobutylicum*, and used to replace in frame *thlA* in *CAB1058* (to yield *CAB1059*). In *CAB1059*, the butanol to ethanol ratio (mol/mol) only increased from 0.59 to 0.84 (Figure 4-3), while the thiolase activity increased 1.7 fold (Figure 4-5) suggesting that the C4/C2 ratio is not controlled by the thiolase level or by the levels of the other enzymes of the pathway and might be thermodynamically controlled. The condensation of two acetyl-CoA to acetoacetyl-CoA catalyzed by the thiolase is thermodynamically unfavorable [35]. Therefore, it is important that acetoacetyl-CoA is efficiently reduced by Hbd to pull the reaction [35]. Hbd is an NADH-dependent enzyme [37] and, as it likely works near the thermodynamic equilibrium, the acetoacetyl-CoA/3-hydroxybutyryl-CoA ratio will be dependent on the NADH/NAD⁺

ratio. As it has been demonstrated in solventogenic *C. acetobutylicum* cells that the NADPH/NADP⁺ ratio is at least 70 times higher than the NADH/NAD⁺ ratio [172], it would be advantageous to replace Hbd with a strictly NADPH enzyme to decrease the acetoacetyl-CoA/3-hydroxybutyryl-CoA ratio and potentially improve the butanol/ethanol ratio. Such an enzyme has been identified and characterized in *Clostridium kluyveri* [40]. In strain *CAB1060*, we have replaced *hbd* by *hbd1* from *C. kluyveri* (Figure 4-1B). A drastic improvement in the butanol/ethanol ratio was observed with an almost 6-fold increase (Figure 4-3). Enzyme assays on crude extract confirm that in the genetic modification of *CAB1060*, the NADH-dependent 3-hydroxybutyryl-CoA dehydrogenase activity was very low while the NADPH-dependent activity was high; the reverse was observed for *CAB1059* (Figure 4-5). *CAB1060* produces n-butanol at high yield (34 wt%) that, to the best of our knowledge, has never been achieved with any microorganism. Furthermore, when the stability of *CAB1060* was evaluated in chemostat culture, we could maintain stable n-butanol (between 9 and 10 g/l) and biomass concentrations (Figure 4-6) and only slight changes in the ethanol/acetate ratio were observed (data not shown).

Having a stable strain that produces n-butanol at a high yield in a continuous culture is a prerequisite, but it is clearly not sufficient for a commercial process as both the titer and the productivities are too low. To improve those two parameters, we envisioned a simple continuous high cell density bioreactor with *in situ* extraction of the alcohols (n-butanol + ethanol) by distillation under low pressure ($0.18\text{-}0.46 \times 10^4$ Pa, for a boiler temperature regulated between 35 and 37°C) without using any microfiltration membrane (Figure 4-2). After an initial batch phase without distillation, this setup was run in a continuous mode by i) starting vacuum distillation and ii) feeding the reactor with a medium at a high concentration of glucose (~500 g/l) while maintaining full cell retention until reaching an OD 600 nm of 90-100 (Figure 4-7), which corresponds to ~28-30 g/l cell dry weight. As shown in Figure 4-7, the targeted cell density was reached in approximately 60 hours after vacuum distillation and feeding began, at which point a cell bleeding was implemented at a dilution rate of 0.05 h^{-1} to maintain a constant cell density. The concentration of butanol in the fermentor was between 9 and 10.5 g/l while the average concentration of n-butanol was 550 g/l with concurrent average ethanol and acetate titers of 60 and 33 g/l. Starting at hour 60 and for 180 hours afterwards (i.e., up to 240 hours of continuous fermentation time), the volumetric productivities of butanol and ethanol remained essentially constant, along with the glucose volumetric

consumption. The average volumetric productivities of butanol and ethanol were 14 and 1.5 g/l/hr, respectively, and the average glucose volumetric consumption rate was 39.4 g/l/hr. This resulted in an average butanol yield of 35 wt% (~83% of the theoretical maximum) and butanol + ethanol yield of 39% [173]. Such high values of n-butanol productivities, titer and yield have never been obtained and compare very well with the performances of the best corn wet milling continuous industrial ethanol processes that display ethanol titers of approximately 100 g/l, yield of 46 wt% and productivity of 10 g/l/hr [174]. If scaled-up at the industrial scale, the technology presented here should drastically decrease both the operational and capital expenditure costs of the historical Weizmann process and make the production of n-butanol economical.

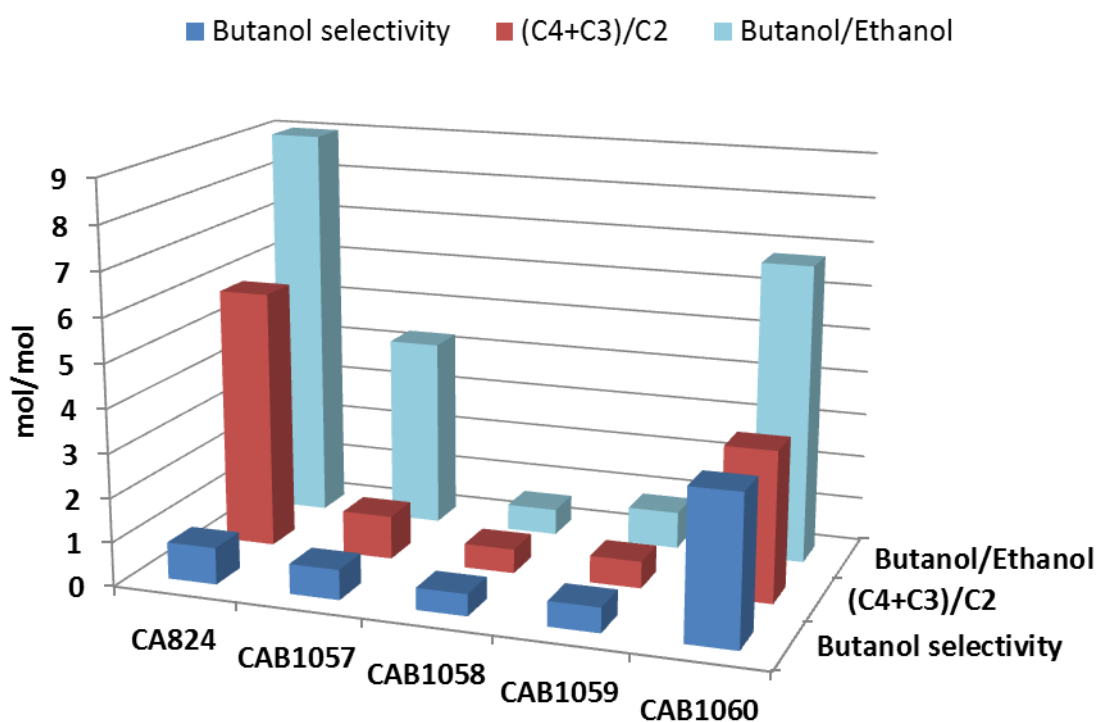


Figure 4-3 Product ratios in engineered *C. acetobutylicum* strains and wild-type

All strains were evaluated in phosphate limited chemostat cultures at pH 4.4 for wild type and 5.0 for all the other strains. Product ratios (mol/mol). $(C4+C3)/C2 = (\text{Butanol} + \text{Butyric acid} + \text{Acetone}) / (\text{Ethanol} + \text{Acetic acid})$. Butanol selectivity = Butanol / (Σ of all products except butanol).

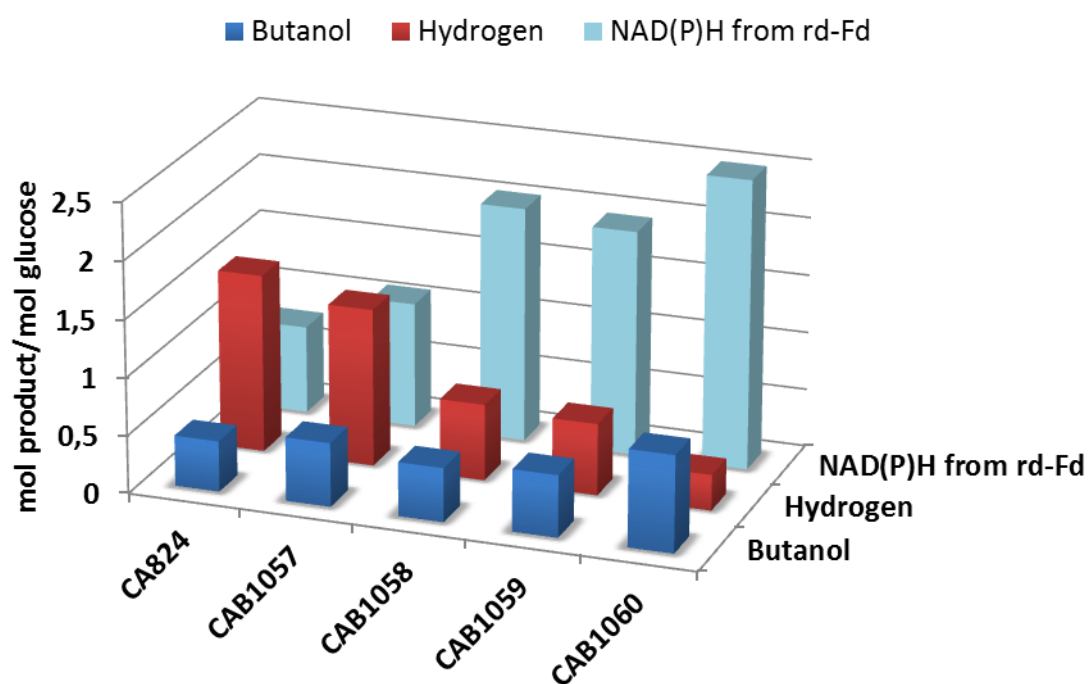


Figure 4-4 Butanol, hydrogen and electron yields in the engineered *C. acetobutylicum* and the wild-type

All strains were evaluated in phosphate limited chemostat cultures at pH 4.4 for wild type and 5.0 for all the other strains.

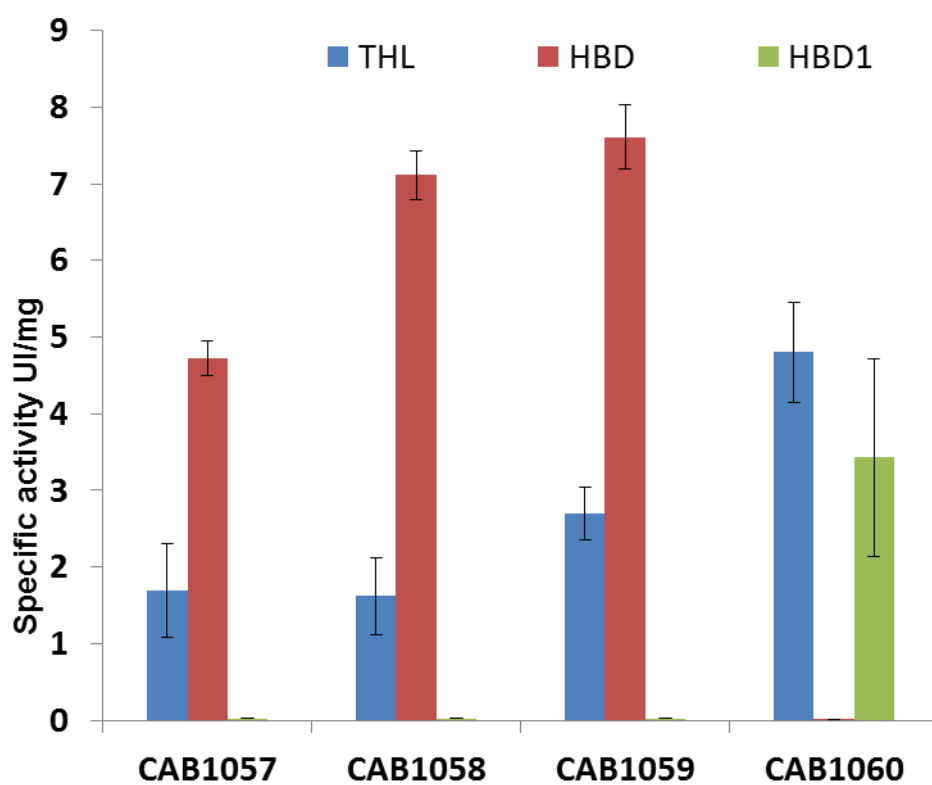


Figure 4-5 Enzymatic assays on crude extracts of *C. acetobutylicum* mutants

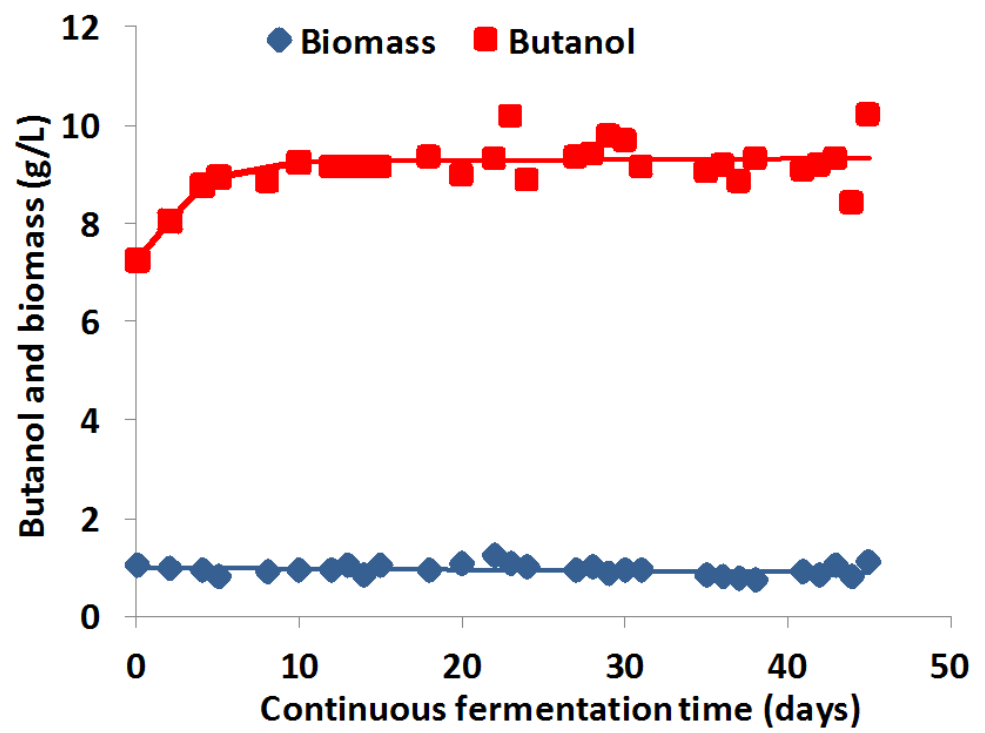


Figure 4-6 Chemostat culture of *C. acetobutylicum* CAB1060

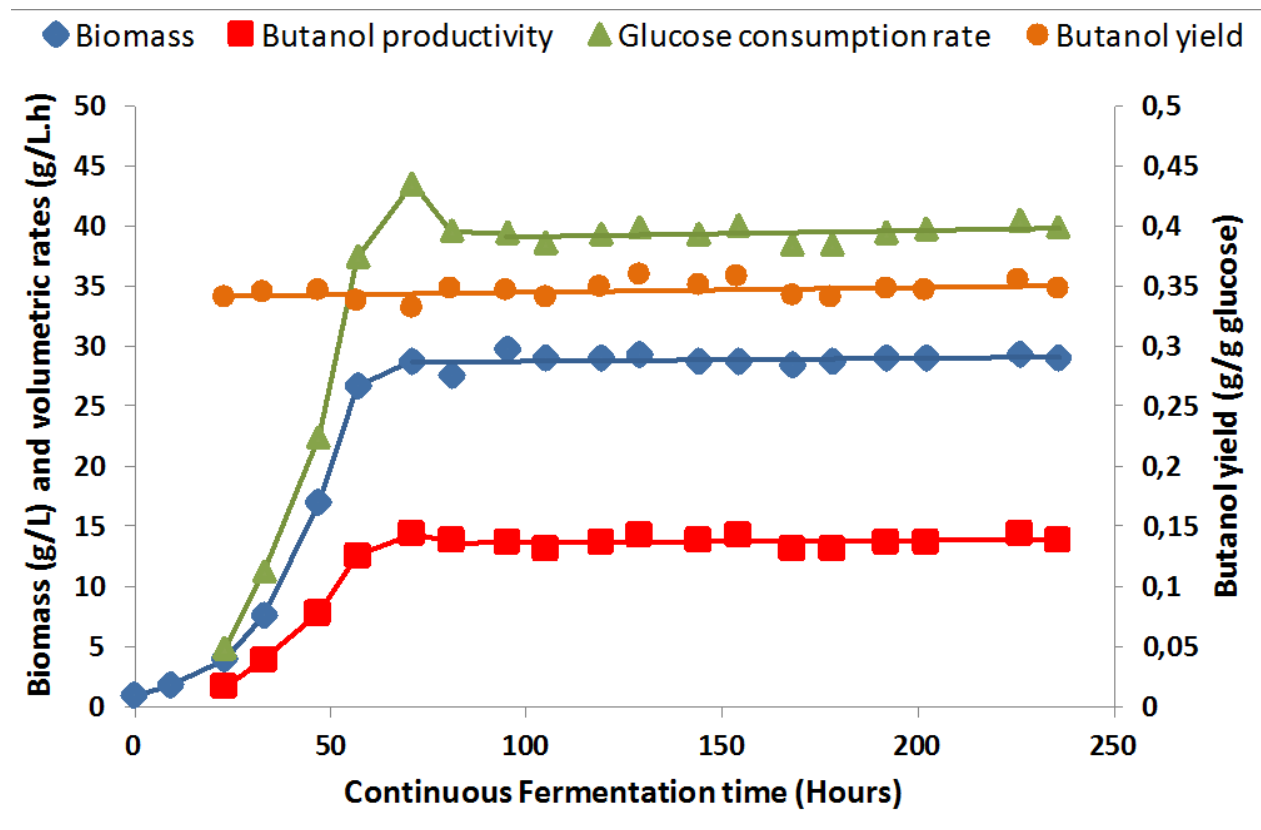


Figure 4-7 Continuous high cell density bioreactor with *in situ* extraction of the alcohols (n-butanol + ethanol) by distillation under low pressure

4.3 Acknowledgments

This work was financially supported by the European Community's Seventh Framework Program "CLOSTNET" (PEOPLE-ITN-2008-237942) (to TN) and by Metabolic Explorer Company.

4.4 Supplementary Materials and Methods

4.4.1 Bacterial strains, plasmids and primers

The bacterial strain and plasmids used in this study are listed in (Table 4-1 and 4-2). The specific oligonucleotides used for PCR amplification were synthesized by Eurogentec (Table 4-3).

Table 4-1 List of strains

Strain	Relevant characteristics	Source
<i>Clostridium acetobutylicum</i> ATCC824	Wildtype	ATCC
<i>Clostridium kluyveri</i> DSM555		DSMZ
MGC Δ <i>cac1502</i> Δ <i>upp</i>	Δ CAC 1502 Δ CAC 2879	[42,167]
MGC Δ <i>cac1502</i> Δ <i>upp</i> Δ <i>cac3535</i> Δ <i>ctfAB</i>	Δ CAC 1502 Δ CAC 2879 Δ CAC 3535 Δ CA_P0162-3	[167]
CAB1057	Δ CAC1502 Δ <i>upp</i> Δ <i>ctfAB</i> Δ <i>ptb</i> Δ <i>buk</i> Δ <i>ldh</i>	This study
CAB1058	Δ CAC1502 Δ <i>upp</i> Δ <i>ctfAB</i> Δ <i>ptb</i> Δ <i>buk</i> Δ <i>ldh</i> Δ <i>rexA</i>	This study
CAB1059	Δ CAC1502 Δ <i>upp</i> Δ <i>ctfAB</i> Δ <i>ptb</i> Δ <i>buk</i> Δ <i>ldh</i> Δ <i>rexA</i> Δ <i>thlA::atoB</i>	This study
CAB1060	Δ CAC1502 Δ <i>upp</i> Δ <i>ctfAB</i> Δ <i>ptb</i> Δ <i>buk</i> Δ <i>ldh</i> Δ <i>rexA</i> Δ <i>thlA::atoB</i> Δ <i>hbd::hbd1</i>	This study

Table 4-2 List of plasmids

Plasmid	Characteristics	Source
pSOS95-MLSr	Ap ^r MLS ^r , <i>repL</i> gene, ColE1	[175]
pCLF1	Cm ^r <i>FLP</i>	[167]
pSOS95	Ap ^r MLS ^r , <i>repL</i> , ColE1, acetone operon	[154]
pSOS95-atoBs	Ap ^r MLS ^r , <i>repL</i> , ColE1, <i>atoB</i> operon	[171]
pREPctfAB::upp	Cm ^r MLS ^r <i>upp</i> Δ <i>ctfAB</i>	[167]
pREPldhA::upp	Cm ^r MLS ^r <i>upp</i> Δ <i>ldhA</i>	[167]
pEryupp-atoB	MLS ^r <i>upp</i> Δ <i>thlA</i> <i>atoB</i>	This study, [171]
pEryUppAthlA	MLS ^r <i>upp</i> Δ <i>thlA</i>	[171]
pSOS95atoBs	MLS ^r <i>upp</i> <i>atoB</i>	[171]
pSOS95-upp-hbd1-catP-oriRepA	Ap ^r MLS ^r , Δ <i>hbd</i> :: <i>hbd1</i> , upp, Cm ^r , <i>repL</i> gene, <i>repA</i> gene	This study
pSOS95-MLSr-upp-Drex-catP	Ap ^r MLS ^r , Δ <i>rex</i> , upp, Cm ^r , <i>repL</i> , colE1	This study [171]
pSOS2K2		

Table 4-3 List of primers

Primers	Sequence
Primer1b	TTTTGGATCCGCTGTTTTACTTGGACATAATACTGAAAAA ATGTC
Primer2	TAAACAGACCTCCCTAATTAATTATTAGCAGCTTTAACTT GAGCTATTAATTCTGG
Primer3	TTAGGGAGGTCTGTTTAATGAGTATTAAGTGTAGCGGT TTAGGTA
Primer4	TTTTGCCGGCTTTTAGGCCTTTAATAAGCGAAGAATCCTTT TCCTGATTTTC
Primer5	TTTTTGGCCATTTTAGGCCTTCAAATAAGTTTACAAGAAT CCCCATTATCA
Primer6	TTTLAGATCTTTATAGCTGTATATTATTCTCTCTTGCATAA TAAA
Primer1-hbd1-sacII	AAAAGCCGCGGTGATCCGCTGTTTTACTTGGACATAATAC TGAAAAAATGTCAAAGG
Primer6-hbd6-sacII	AAAAACCGCGGTGATCTTTATAGCTGTATATTATTCTCTCT TGCATAATAAAACAAATATTGCTGAGC
oriRepA-XbaI-F	AAAAATCTAGAGAGCTTGGCACCCAGCCTG
oriRepA-5'-R	GCATTGGTAACTGTCAGACCAATTCCCGACAGTAAG
Amp3'-F	GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC
Amp-ScaI-R	TTTTTTAGTACTCACCAGTCACAGAAAAGCATC
etfB-3-F	GACAGGGAGAAGTTATTGATAAGC
cac2706-3-Rb	GGCTTCTTAACACATAAAAGTACC
catP-3D	GGAAGGAAAGCCAAATGCTCCGG
catP-5R	GTACAAGGTACACTTGCAAAGTAGTGG
Atob-1	ACTTATGAAATAGATTGAAATGGTTTATCTGTTACCCCGT AGGATCCTCGAAGGTCGACCA

Atob-2	CTAATTTATAATTCTACAGAGTTATTTTAAACAATACTTTT TAGTTTAGTCGTTCTATTACCA
Rex 1	AAGGAAAAAAGCGGCCGCAAGCTTACAAAGTGCTACA CGGGTTTTTTGCCC
Rex 2	CATTACCGTACTAATCTCGGCTTTTTTCGCGATTGAC
Rex 3	CAAAAGTGATATTAAGTAAGAGGTGGTCAAATCGCGA AAAAGCCGAGATTAGTACGGTAATG
Rex 4	AAGGAAAAAAGCGGCCGCAAATCTTTCTGTTTCATCA ATTTCTGC
Rex 0	GCATCTAGGAAATATCTATCATGAG
Rex 5	GAAAGAATAACTCCTGTAGTACCGC

4.4.2 Chemostat culture of recombinant strains

The conditions described previously by Girbal et al. [30] were used for the phosphate-limited continuous culture of *C. acetobutylicum* fed at 30 g/l glucose at a pH of 5.2.

4.4.3 Continuous extractive high density cell recycle fermentation

A 10% inoculum of mid-exponential phase (OD_{600nm} of 0.8-1.5) CAB1060 culture was used to inoculate a 1 L glass bioreactor with a 0.5 L working volume of anaerobic synthetic medium [176] with 60 g/l glucose. Once the cultures reached an OD_{600nm} of 3, extractive fermentation using low pressure distillation was initiated. Fermentation pH was controlled from dropping below 5.0 using 4M NH_4OH . Low pressure distillation was accomplished by sending the whole broth to a 3 L Pyrex boiler (0.3 L working volume) equipped [177] to maintained at a constant temperature (35-37 °C) and a constant low pressure ($0.18-0.46 \times 10^4$ Pa) above the boiler. The residence time of the broth in the boiler was between 0.25 and 0.5 hour. The distillate was condensate and demixed in a butanol rich and a water rich phase; the water rich phase was recycled to the distillation column while the butanol rich was collected. The fermenter was fed exponentially with a feed medium (see the composition below) containing high glucose concentration (530 g/l) and kept in complete cell retention until reaching an OD_{600nm} of 90 – 100 (28-30 g/l cell dry weight), at which time a harvest was started at a dilution rate of $0.05h^{-1}$ and the volume of the fermenter was kept constant by the

addition of feed medium. The feed medium has the following composition (per liter of distilled water): glucose, 530g; KH₂PO₄, 1.8g; KCl, 5.7g; MgSO₄·7H₂O, 3.6g; FeSO₄·7H₂O, 0.5g; NH₄Cl, 18 g; Biotin, 0.18 mg; PAB, 0.14g; Yeast Extract, 4g.

4.4.4 Measurement of fermentation parameters

Biomass concentration was determined both by the DCW method after centrifugation (16,000 g, 5 min, room temperature), two washes with Milli-Q water, and drying under vacuum at 80 °C. The concentrations of glucose, glycerol, acetate, butyrate, lactate, pyruvate, acetoin, acetone, ethanol, and butanol were determined based on high-performance liquid chromatography (HPLC), as described by Dusséaux et al. [178], except that the concentration of H₂SO₄ was changed to 0.5 mM, as required for mobile phase optimization.

4.4.5 General protocol for gene deletion by homologous recombination system using replicative plasmid pSOS95-MLS^r

Gene disruption in specified chromosomal loci was carried out by homologous recombination as previously described [167] with the modifications of patent application WO2008/040387 [151]. Deletion cassettes consisting of the upstream homologous region-FRT-catP-FRT-downstream homologous region were cloned in the pSOS95-MLS^r-upp plasmid. The pSOS95-MLS^r-upp plasmids with the different deletion cassettes were used to transform the successive recipient *C. acetobutylicum* strains which were constructed from the MGC*Acac1502Δupp* platform strain [167]. As a general procedure, thiamphenicol-resistant transformants were selected on CGM 50g/L Glucose 0.1M MES pH 6.1 (CGMMG) plates containing thiamphenicol (10 µg/mL) (Tm). One colony was cultured for 24 hours in liquid CGMMG with thiamphenicol (10 µg/mL) and 100 µL of undiluted culture was plated on CGMMG with thiamphenicol (10 µg/mL) and 5-FU (1 mM). Colonies resistant to both thiamphenicol and 5-FU were replica plated on both CGMMG with thiamphenicol (5 µg/mL) and CGMMG with clarithromycin (40 µg/mL) to select clones where 5-FU resistance is also associated with clarithromycin sensitivity. The genotype of the clones resistant to thiamphenicol and sensitive to clarithromycin was checked by PCR.

The catP resistance cassette was removed by transforming the strain with plasmid pCLF1 (16) expressing the *flp1* gene encoding the Flp recombinase from *S. cerevisiae*. One colony of a clarithromycin resistant clone was cultured on liquid CGMMG without clarithromycin and the culture was plated on CGMMG with 5-FU (1 mM). Colonies resistant to 5-FU were

replica plated on CGMMG with clarithromycin (40 µg/mL), CGMMG with thiamphenicol (10 µg/mL) and CGMMG without antibiotic to select clones. The 5-FU resistance was associated with the loss of the pCLF1 vector and the thiamphenicol sensitivity to the elimination of the catP marker. The genotype of sensitive clones was checked by PCR.

CAB1057 was constructed from the MGCΔ*cac1502*Δ*upp* platform strain by successively applying this protocol to the deletion of *ptb-buk*, *ctfAB* and *ldhA*

4.4.6 Construction of pSOS95-MLSr-upp-Drex-catP and CAB1058 strain

To delete the rex regulator (encoded by *CA_C2713*), protocol 4.4.5 was used except that couples of primers Rex 1-Rex 2 and Rex 3-Rex 4 were used to amplify the 1135 bp upstream homologous region and the 1138 bp downstream homologous region of the rex locus respectively. Both primers Rex 1 and Rex 4 introduced NotI sites while primers Rex 2 and Rex 3 have a complementary region which introduced a NruI site. DNA fragments Rex 1-Rex 2 and Rex 3-Rex 4 were joined in a PCR fusion experiment with primers Rex 1 and Rex 4 and the resulting fragment was cloned in pCR4-TOPO-Blunt (Invitrogen) to yield pTOPO:Rex. At the unique NruI site of pTOPO:Rex, the antibiotic resistance catP gene (Cm/Tm) with FRT sequences on both sides was introduced from the 1230 bp StuI fragment of pUC18-FRT-Pptb-catP. The Rex deletion cassette obtained after NotI digestion of the resulting plasmid was cloned into pSOS95-MLSr-upp at the NotI site to yield the pSOS95-MLSr-upp-Δrex-catP plasmid.

The pSOS95-MLSr-upp-Drex-catP plasmid was used to transform by electroporation the *Clostridium acetobutylicum* CAB1057 strain. The genotype of clones resistant to thiamphenicol and sensitive to clarithromycin was checked by PCR analysis (with primers Rex 0 and Rex 5 located outside of the Rex deletion cassette). The ATCC824 Δrex-catP^R strain having lost pSOS95-MLSr-upp-Δrex-catP vector was isolated.

The thiamphenicol resistance of the above strain was removed according to 4.4.5. The genotype of the ATCC824 Δrex strain having lost the pCLF1 plasmid was checked by PCR analysis with primers Rex 0 and Rex 5.

4.4.7 Construction of pEryupp-atoB plasmid and CAB1059 strain

A synthetic *atoB* operon coding for the AtoB thiolase of *E. coli* was synthesized as follow. A synthetic *atoB* gene was designed with both an optimized synthetic RBS

(translation initiation rate was calculated to be 60217, a value similar to the translation initiation rate of *thlA*) and a harmonized codon usage for *C. acetobutylicum*. The synthetic gene was synthesized by Life technology (*atoBs*) with a *BamHI* site in 5' and a *SfoI* site in 3' and was cloned in a *BamHI*-*SfoI* digested pSOS95 [154] to yield pSOS95atoBs.

The synthetic *atoB* gene was PCR amplified with the Phusion DNA polymerase with pSOSatoBs as a template and the Atob-1 and Atob-2 oligonucleotides as primers. This DNA fragment was directly cloned into the *StuI* linearized pEryUppAthlA (for the construction of pEryUppAthlA, see Chapter 5 or [171]) by homologous recombination using the GeneArt® Seamless Cloning & Assembly kit to yield the pEryUppAtoB plasmid.

The pEryUppAtoB plasmid was used to transform by electroporation to *CAB1058* strain. After selection on Petri plates for clones having inserted the pEryUppAtoB plasmid by homologous recombination (resistant to erythromycin 20 µg/ml), two colonies were cultured for 24 hours in CGM Glucose MES 0.1M medium. Appropriate dilutions were plated on CGM Glucose MES 0.1M (CGMMG) with 5-FU 1mM. To select integrants having excised and lost pEryUppAtoB, 5-FU resistant clones were replica plated on both CGMMG + 5FU and CGMMG with erythromycin at 40 µg/ml. To identify clones having lost pEryUppAtoB and possessing an *atoB* insertion, (clones resistant to 5-FU and sensitive to erythromycin) were checked by PCR analysis (with primers Thl-0 and Thl-5 located outside of the *AthlA* locus (Chapter 5) and Atob-1 and Atob-2 primers binding specifically to *atoB* gene). The clones had *thlA* replaced by *atoB* were isolated to obtain pure clones to obtain *CAB1059* strain.

4.4.8 Construction of pSOS95-upp-hbd1-catP-oriRepA and CAB1060 strain

To amplify the two homologous regions flanking the CDS region of the *hbd* gene on the genome of *C. acetobutylicum* ATCC824, two couples of primers (primer 1-hbd1-sacII/primer 2 and primer 5/primers 6-hbd1-sacII) were used. Two fragments of 909 bp and 989 bp were obtained. Primer 3 and primer 4 were used to amplify the *hbd1* gene from genomic DNA of *C. kluyveri* DSM555 to obtain the third fragment of 886 bp. Fragment 1 and 3 was fused by overlapping PCR using primer 1-hbd1-sacII and primer4 to obtain the fourth fragment of 1778 bp. After that, this fourth fragment was fused with the second fragment by primer 1-hbd1-sacII and primer 6-hbd1-sacII to produce the fifth fragment of 2692 bp.

The *colE1* origin on pSOS95-MLS^r [175] was replaced by *repA* origin. The *repA* origin was amplified from pSOS2KS by two primers oriRepA-XbaI-F and oriRepA-5'R to obtain a PCR product of 2296 bp. The partial ampicillin resistant gene from pSOS95-upp was amplified by Amp3'-F and Amp-ScaI-R to obtain a PCR product of 573bp. The two resultant PCR fragment was then fused by PCR using oriRepA-XbaI-F and Amp-ScaI-R. This 2849 bp fragment was introduced to pSOS95-upp to replace *repL* fragment by digestion with *XbaI* and *ScaI* restriction enzymes and ligation to obtain pSOS95-upp-*repA*. After that, the homologous fragment containing *hbd1* CDS was introduced to this pSOS95-upp-ori*repA* by the *SacII* site to obtain pSOS95-*hbd1*-ori*repA* plasmid. The *catP* cassette (flanking by FRT sequences) was inserted into the *StuI* site carried by primer 4 and primer 5 in the homologous region to generate the final plasmid pSOS95-*hbd1*-*catP*-ori*RepA*.

The final plasmid pSOS95-*hbd1*-*catP*-ori*RepA* was used to transform into *CAB1059* strain and followed protocol 4.4.5 for the deletion. The transformants on plates CGMMG Thiamphenicol 10µg/ml were selected and checked by PCR for the single integration by two couples of primers *etfB*-3-F and *catP*-3D or *cac2706*-3Rb and *catP*-5R. The integrants were spread on plates CGMMG containing Thiamphenicol 10 µg/ml and 5-FU 1mM to promote the second integration. Clones that were Thiamphenicol resistant and Clarithromycin sensitive were selected and checked for double cross-over events by PCR using two external primers *etfB*-3-F and *cac2706*-3Rb. The mutants were isolated to obtain *CAB1060-catP* strain. After that, pCLF1 [167] was introduced into *CAB1060-catP* strain to remove *catP* cassette to generate the marker-less strain CAB1060.

4.4.9 Enzyme activity measurements

Cell crude extract of *C. acetobutylicum* mutants was prepared in 100 mM Tris-HCl (pH 7.6), 5 mM dithiothreitol, 2.5% glycerol. The cell suspension was sonicated in an ultrasonic disintegrator (Vibracell 71434; Bioblock, Illkirch, France), at 0°C in four cycles of 30 s with 2-min intervals between each cycle. Cell debris was removed by three centrifugations at $13,000 \times g$ for 5 min . The supernatant was loaded to PD10 column.

Thiolase activity was measured in the reverse thiolytic cleavage direction. The reaction containing 0.1 M Tris-HCl pH8 ; 0.2 mM CoASH; 60 µM Acetoacetyl-CoA; 10 mM MgCl₂; 1 mM DTT and diluted crude extracts was incubated at 30°C and followed the decrease of acetoacetyl-coA at 303nm as described [35,179]. The enzyme activity was calculated by the

difference in the slope values of the sample and control and using the molar extinction coefficient of $14,000\text{ M}^{-1}\text{cm}^{-1}$.

HBD activity was measured as described by [179,180]. The reaction containing 0.1mM MOPS (pH 7.0) buffer, 1mM DTT, 0.1mM acetoacetyl-CoA, and 0.15mM NADH, and crude extract at different dilution was incubated at 35°C. The decreasing of NADH was followed spectrophotometrically at 345nm. The enzyme activity was calculated by the difference in the slope values of the sample and control and using the molar extinction coefficient of $6,220\text{ M}^{-1}\text{cm}^{-1}$.

HBD1 activity was measured as described by [40]. Reaction components were 0.1M potassium phosphate, 25mM potassium citrate, 0.15mM NADPH, 0.6mM Acetoacetyl-coA and crude extract. The assay was done at 35°C and the decrease of NADPH was followed at 340nm. The molar extinction coefficient of $6,220\text{ M}^{-1}\text{cm}^{-1}$ was used to calculate the enzyme activity.

Chapter 5 Construction of a hydrogenase minus mutant of *Clostridium acetobutylicum* : a platform strain for the continuous production of fuels and chemicals

Soucaille Philippe, Nguyen Ngoc-Phuong-Thao, Percheron Benjamin, Croux Christian,
Meynial-Salles Isabelle

5.1 Main text

In this chapter we want to answer the following question: what bio-chemicals could be continuously produced at high yield if we could get a hydrogenase minus mutant of *Clostridium acetobutylicum*? For this purpose, simulations have been performed using a Genome Scale Model [42] of *Clostridium acetobutylicum* recently developed in our group. The parameters were a growth rate of 0.05 h^{-1} and a specific carbon consumption rate of $50 \text{ mC.g}^{-1}.\text{h}^{-1}$ that can be obtained in a phosphate (0.75 mM Phosphate in the fed) and carbon (1000 mM Carbon in the fed) limited chemostat culture run at a dilution rate of 0.05 h^{-1} in synthetic medium [22,181]. All the simulations were done considering that no hydrogen was produced. The results of the simulation for five chemicals are presented in the following Table 5-1.

	Lactate	Ethanol	n-Butanol	Isobutanol	1,3-Propanediol
Titer (g/l)	28.5	14.7	11.8	11.8	18.1
Productivity ($\text{g.l}^{-1}.\text{h}^{-1}$)	1.42	0.73	0.59	0.59	0.90
Yield (g/g)	0.95	0.49	0.39	0.39	0.6
Substrate	Glucose	Glucose	Glucose	Glucose	Glycerol
Glycerol produced	YES	YES	YES	YES	NO

Table 5-1 Prediction of bio-chemical yield produced by a hydrogenase minus mutant

The conclusions of these simulations are that if a hydrogenase minus strain could be obtained i) it would be advantageous in term of yield and stability of the strain in continuous culture as the metabolism would be redox constraint and ii) if glucose is used as a carbon source, glycerol will have to be produced to regenerate the NAD^+ from the excess NADH associated to biomass formation.

Three hydrogenases are potentially involved in hydrogen production: two Fe-Fe hydrogenases, HydA (encoded by *CA_C0028*) and HydB (encoded by *CA_C3230*), and one Ni-Fe hydrogenase, HupSL (encoded by *CA_P0141–0142*). It was previously shown in our group [42] that in phosphate limited chemostat culture, the *hydB* and the *hupSL* genes are not expressed, nor were the HydB and HupSL proteins detected by quantitative proteomic analysis. As HydA is the only hydrogenase, we then asked ourselves the question: Is it possible to obtain a *hydA*⁻ mutant of *C. acetobutylicum*? A previous study from the group of

Nigel Minton [91] using the Clostron as a tool to inactivate *hydA*, reported that they could not obtain insertion of an intron in *hydA* suggesting that it might be an essential gene for *C. acetobutylicum*. To unambiguously establish the essentiality of *hydA*, we used another intron-based method developed in our laboratory [182] based on the pCONS::upp-intron vector for the 1) insertion of a marker-less group II intron and 2) positive selection for the loss of the plasmid in the presence of 5-FU. Use of the pCONS::upp-intron in *C. acetobutylicum* Δ *cac1502* Δ *upp* was not successful for the disruption of *hydA* with a group II intron with an antisense orientation but we were able to disrupt *hydA* with a group II intron having a sense orientation (Figure 5-1). Furthermore, the integrants were unable to lose the plasmid even in the presence of 5-FU: all 5-FU-resistant clones possessed either a deletion of or a mutation in the *upp* gene, leading to survival in the presence of 5-FU even in the presence of the plasmid. This demonstrates that the *ltrA* gene located on the plasmid has to be functional to splice the intron in the *hydA* mRNA, thus allowing expression of *hydA* and growth of the microorganism. All the results obtained from *hydA* disruption experiments clearly demonstrate that *hydA* is essential in *C. acetobutylicum*.

What are the hypothesis that might explain why *hydA* is an essential gene?

- A *hydA*⁻ mutant cannot redirect its carbon flux (down regulation of the acetate, butyrate, acetone, acetoin pathways) to adjust to the change in the electron flux and produce only butanol, ethanol and lactate.
- The glycerol production pathway is not sufficiently well expressed.

Regarding the first hypothesis, having a thiolase negative mutant that would no longer produced butyrate, acetone and butanol would potentially help in getting a *hydA* mutant that would only produce ethanol and lactate as major products and glycerol as a minor one.

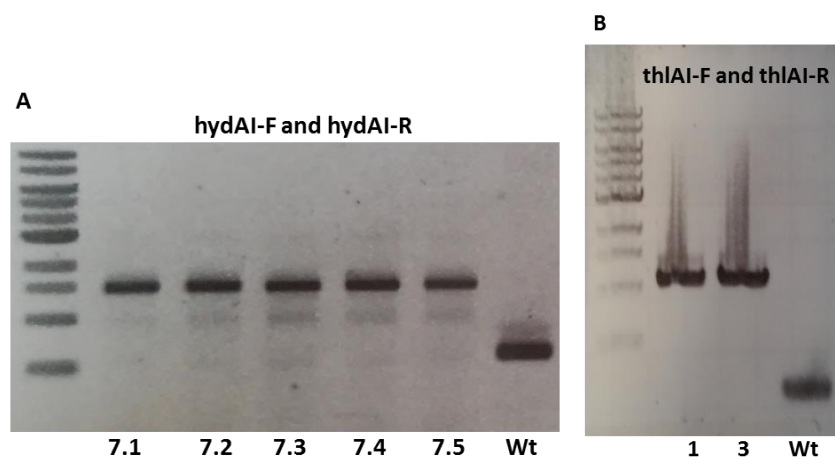


Figure 5-1 Construction of *hydA*::int(s) and *thlA* ::int(s) mutants.

(A) PCR on genomic DNA extraction of *hydA*::int(s) mutants (MT) and wildtype (WT) strains with *hydA*-F and *HydA*-R couple of primers to confirm intron insertion. (B) PCR on genomic extraction of *thlA* ::int(s) mutant (MT) and wildtype (WT) strains with *ThlA*-F and *ThlA*-R couple of primers to confirm intron insertion.

Two thiolase encoding genes have been identified: *thlA* (CA_C2873) and *thlB* (CA_P0078). It was previously shown in our group [42] that in phosphate limited chemostat culture, the *thlB* gene is poorly expressed and consequently the level of ThlB protein detected by quantitative proteomic analysis was less than 1% the level of ThlA. As ThlA is then the main expressed thiolase, we then asked ourselves the question: Is it possible to obtain a *thlA* mutant of *C. acetobutylicum*? A previous study from the group of Nigel Minton [91] using the Clostron as a tool to inactivate *thlA*, reported that they could not obtain insertion of an intron in *thlA* suggesting that it might also be an essential gene for *C. acetobutylicum*. To unambiguously establish the essentiality of *thlA*, we used the same approach as for *hydA* using the pCONS::upp-intron vector. Use of the pCONS::upp-intron in *C. acetobutylicum* Δ *cac15* Δ *upp* was not successful for the disruption of *thlA* with a group II intron with an antisense orientation but we were able to disrupt *thlA* with a group II intron having a sense orientation (Figure 5-1). Furthermore, the integrants were unable to lose the plasmid even in the presence

of 5-FU: like for *hydA*, all the 5-FU-resistant clones possessed either a deletion of or a mutation in the *upp* gene, leading to survival in the presence of 5-FU even in the presence of the plasmid. This demonstrates that the *ltrA* gene located on the plasmid has to be functional to splice the intron in the *thlA* mRNA, thus allowing expression of *thlA* and growth of the microorganism. All the results obtained from *thlA* disruption experiments clearly demonstrate that *thlA* is essential in *C. acetobutylicum*.

What are the hypothesis that might explain why *thlA* is an essential gene?

- A *thlA*⁻ mutant cannot redirect its electron flux to adjust it to the change in the carbon flux (reduce acetyl-coA to ethanol and avoid acetate production).
- The glycerol production pathway is not sufficiently well expressed

As *hydA* inactivation might be compromised by the carbon fluxes redirection and *thlA* inactivation by the electron fluxes redirection, we made the hypothesis that if we could simultaneously inactivate *hydA* and *thlA* such a mutant would be viable and would produce ethanol and lactate with minor amount of glycerol.

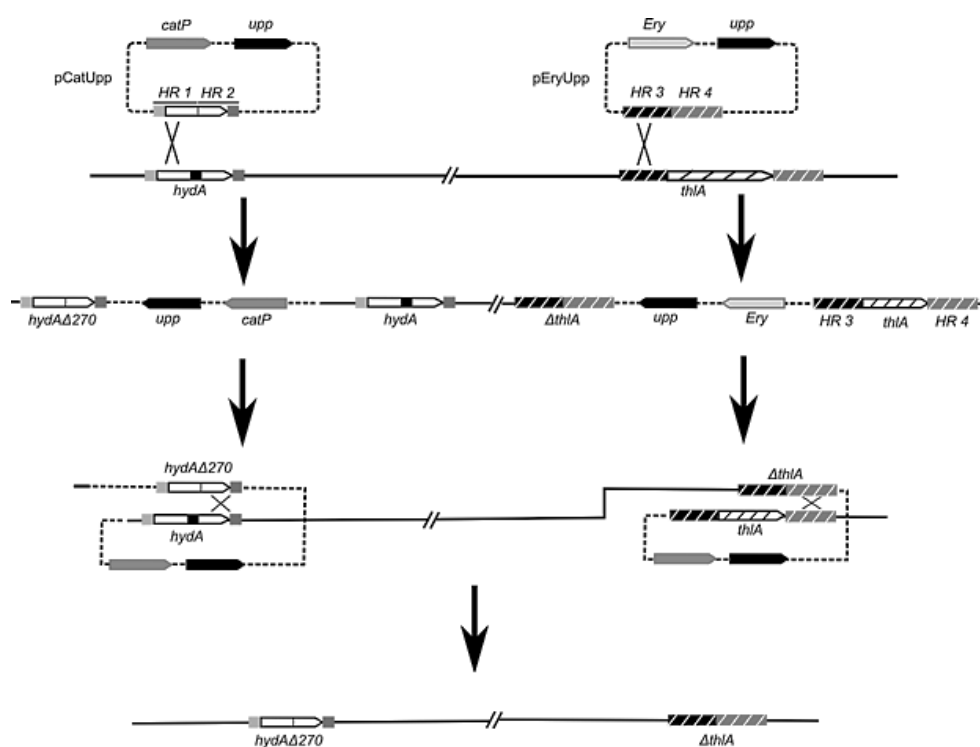


Figure 5-2 Schematic representation of the different steps for the simultaneous inactivation of *hydA* and *thlA*.

In order to simultaneously inactivate *thlA* and *hydA*, we developed the new method presented in Figure 5-2. It is based on the successive integration, by single crossing-overs, of two non-replicative plasmids possessing i) homologous region around the chromosomal regions to be deleted, ii) two different antibiotic resistance markers (*ery^R* and *th^R*) and iii) a *upp* gene that confers sensitivity to 5-FU. The advantage of the method is that the first integration of the non-replicative plasmids by single crossing over does not inactivate the targeted genes; The inactivation is only effective after the second crossing over leading to the excision of the plasmids. The simultaneous excision of the two plasmids is obtained by plating the cells on 5-FU containing CGM plates and replica plating on erythromycin and thiamphenicol CGM plates and selecting the clones which are sensitive to both antibiotics. To differentiate the clones with a wild-type or the desired genotype, PCR analysis were performed with primers outside the recombination zones. In the case of *hydA* and *thlA* it should not be possible to get single deletions mutant as it was shown above that such mutants are not viable. We applied this strategy by constructing two vectors, the pEryUppΔ*thlA* that was made from two 1kb DNA fragments surrounding *thlA* and the pCatUppΔ*hydA*270 that contains a deletion cassette made of two 1kb DNA fragments with an in frame deletion of 270 bp in a region coding for the catalytic domain of the main [Fe-Fe] hydrogenase. Integration of both plasmids was obtained both in 5' and 3' of the homologous region and a clone with integration of pCatUppΔ*hydA*270 plasmid in the 3' homologous region and of pEryUppΔ*thlA* in the 5' homologous region was chosen (Figure 5-3) and subject to 5-FU selection on CGM plates. Several clones (6 clones) with the right phenotype were obtained (resistance to 5-FU and sensitivity to erythromycin and thiamphenicol) and approximately half of them have the expected genotype (deletion of *thlA* and 270 bp of *hydA*) the others having a wild type genotype. One of the clones having the right genotype, *C. acetobutylicum*Δ*upp*Δ*cac1502* Δ*thlA*Δ*hydA*270 (Figure 5-3) was cultured in batch culture without pH regulation.

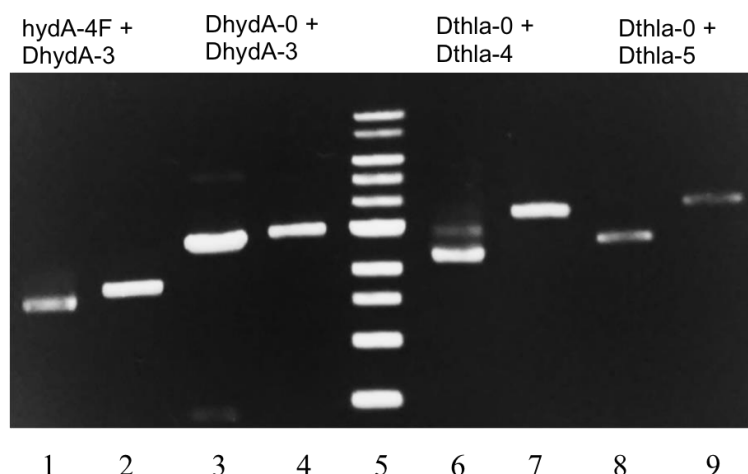


Figure 5-3 PCR analysis of intermediate (with pcatupp Δ hydA270 plasmid integrated in 3' and pEryUpp Δ thlA integrated in the 5') and the final hydrogenase mutant *C.*

***acetobutylicum* Δ upp Δ cac1502 Δ thlA Δ hydA270.**

Lane 1 and 2: integration of pCatUpp Δ hydA270 in the 3' homologous region in the intermediate strain vs genotype of wildtype. Lane 3 and 4: deletion of 270bp in *hydA* gene of the final mutant vs genotype of the wildtype. Lane 5: marker. Lane 6 and 7: integration of pEryUpp Δ thlA in the 5' homologous region in the intermediate strain vs genotype of wildtype. Lane 8 and 9: deletion of *thlA* gene of the final mutant vs genotype of the wildtype.

This strain produces ethanol and lactate as the major product and small amount of glycerol and acetate **but** H₂ (Figure 5-4)

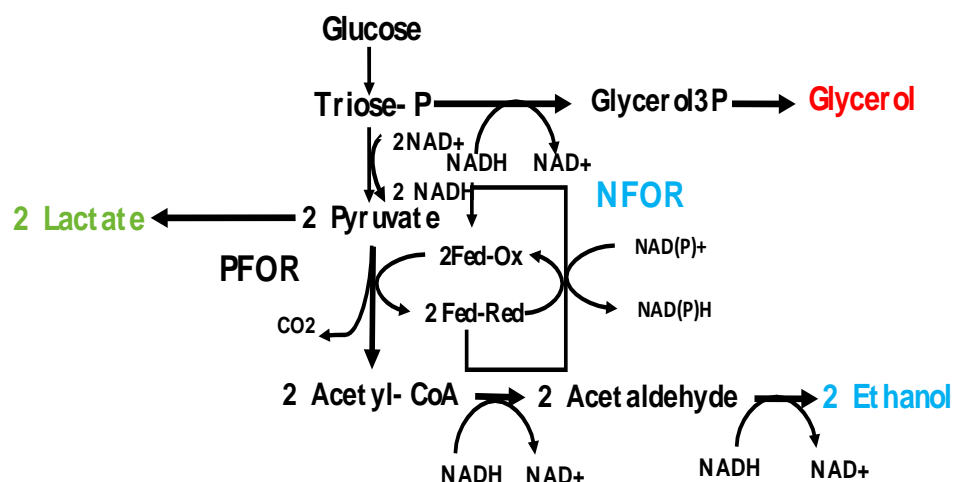
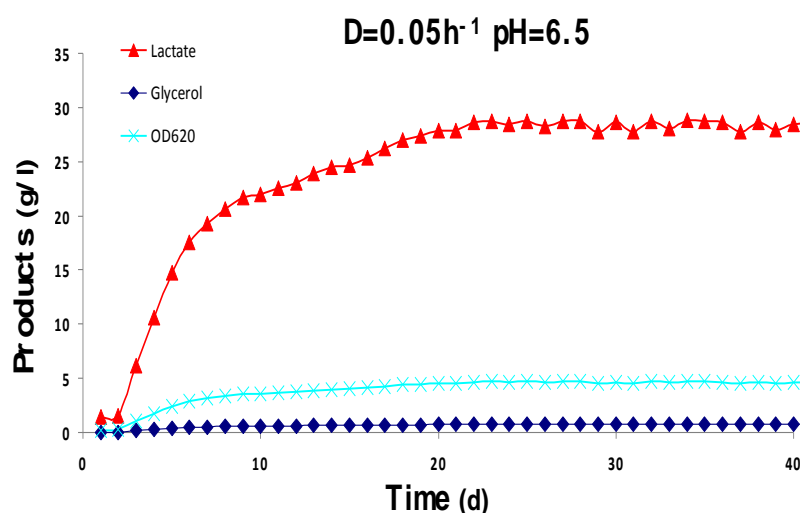


Figure 5-4 Metabolism of *C. acetobutylicum* $\Delta upp \Delta cac1502 \Delta thlA \Delta hydA270$.

Experimental conditions: Batch culture in MS glucose medium 20 g/l, Temp 35°C, initial pH 6.5, no pH regulation.

In order to obtain a strain that produces L-lactate at high yield, the pSOL1 megaplasmid was cured from the *C. acetobutylicum* $\Delta upp \Delta cac1502 \Delta thlA \Delta hydA270$ strain by 20 serial subcultures in MS glucose medium. Cells having lost the pSOL1 megaplasmid were identified on an RCA agar plate containing starch (2%) and glucose (0.2%) and after iodine staining as they do not produce halo of starch hydrolysis [183]. Approximately 1 colony out of 1000 has lost the pSOL1 megaplasmid after 20 subcultures. The *C. acetobutylicum* $\Delta upp \Delta cac1502 \Delta thlA \Delta hydA270 \Delta pSOL$ strain was obtained. This strain was cultured in batch culture without pH regulation and was shown to produce lactate as the main fermentation product with small amount of glycerol. This strain was then used for the continuous production of lactate in MS glucose medium. The results are shown in Figure 5-5. The best yield on glucose, titer and productivities obtained were respectively 0.94 g/g, 27.7 g/l and 1.37 g/l.h.



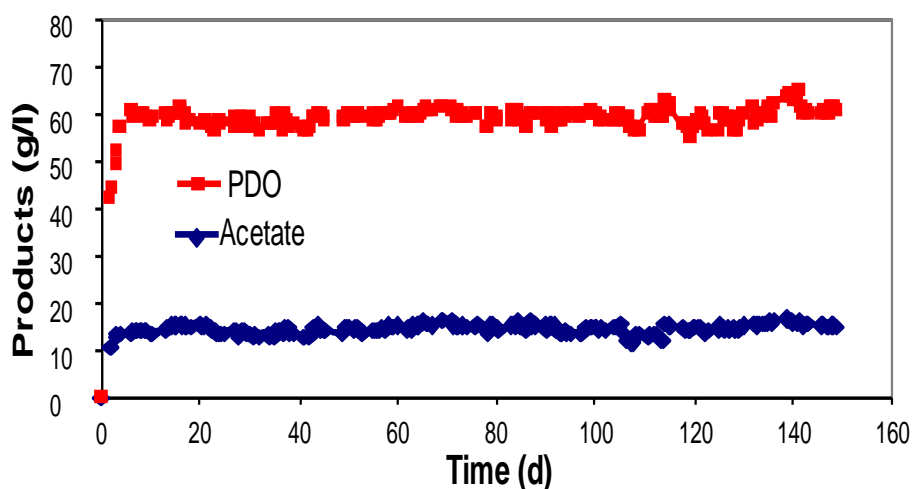
Lactate ~ 27.7g/L, yield 0.94g/g, productivity 1.37g/l.h

Figure 5-5 Continuous production of lactate by the *C. acetobutylicum*

***ΔuppΔcac1502 ΔthlAΔhydA270ΔpSOL*.**

Experimental conditions: MS glucose medium 30 g/l, pH 6.5, Temp 35°C, D=0.05h⁻¹.

The *C. acetobutylicum*ΔuppΔcac1502 ΔthlAΔhydA270 ΔpSOL strain produced lactate at high yield. In order to obtain a strain that converts glycerol to 1,3-propanediol at high yield, the pSPD5 plasmid [154] that carries the genes coding for the B12-independent 1,3-propanediol pathway from *Clostridium butyricum* was used to transform by electroporation the *C. acetobutylicum*ΔuppΔcac1502 ΔthlAΔhydA270ΔpSOL strain. After selection on Petri plates for clones containing the replicative pSPD5 plasmid (resistant to erythromycin at 40μg/ml), 10 colonies were cultured for 24 hours in liquid 2YTGlycerol medium and then sub-cultured in liquid MS Glycerol medium. The clone having the fastest growth in MS glycerol medium was designated *C. acetobutylicum*ΔuppΔcac1502 ΔthlAΔhydA270 ΔpSOL (pSPD5). This strain was used for the continuous production of 1,3-propanediol in MS glycerol medium for 150 days. The stable production of 1,3-propanediol is shown in Figure 5-6. Acetate was the only by-product. The average yield on glycerol, titer and productivities obtained were respectively 0.6 g/g, 60 g/l and 3 g/l.h.



1,3-propanediol ~ 60g/L, yield 0.6g/g, productivity 3g/l.h

Figure 5-6 Continuous production of 1,3-propanediol by the *C. acetobutylicum* $\Delta upp \Delta cac1502 \Delta thlA \Delta hydA270 \Delta pSOL$ (pSPD5).

Experimental conditions: MS glycerol medium 105 g/l, pH 6.3, Temp 35°C, D=0.05h⁻¹

5.2 Conclusions

By using the Genome Scale Model developed in our group, we demonstrated that *C. acetobutylicum* can potentially continuously produce several bio-chemicals if we manage to block hydrogen production.

By simultaneously inactivating *hydA* and *thlA*, we managed to obtain a platform strain that no longer produces hydrogen.

This platform strain was further engineered to continuously produce L-lactate and 1,3-propanediol at high yield, titer and productivities.

Further work is currently carried out to construct strains producing either, isopropanol, 1,3-butanediol or isobutene.

5.3 Material and Methods

5.3.1 Knockout of *hydA* and *thlA* using *LltrB* intron

PCR primers were designed using a computer algorithm provided as part of the TargeTron Gene Knockout System kit [184]. The *LltrB* intron obtained from pACD4K-C was retargeted to the *hupL* gene by an antisense insertion (between nucleotides 421-422), to the

hydB gene by an antisense insertion (between nucleotides 113-114) and to *hydA* by an antisense (between nucleotides 370-371) and a sense insertion (between nucleotides 60-61). A SOEing (Gene Splicing by Overlap Extension) PCR was used to assemble and amplify a 353-bp product containing the modified Intron Binding Site (IBS), Exon Binding Site 1d (EBS1d), and Exon Binding Site 2 (EBS2) sequences responsible for intron-retargeting (Table 5-2).

To inactivate a gene using TargeTron technology together with the positive selection for plasmid loss, a customized vector (pCONS::upp-intron) was constructed from the original pACD4 TargeTron vector (Sigma). The pCONS::upp -intron vector contains a pIM13 origin of replication functional in Clostridia, a *catP* gene conferring thiamphenicol resistance, a *upp* gene for 5-fluorouracil (5FU) sensibility and a LtrA ORF required for a functional expression of the group II intron. The LtrA ORF was obtained from restriction digestion of pACD4 with *XbaI* and *PshAI* and was then sub-cloned into the pSOS95 vector. The pSOS95 vector was digested with *BamHI* and *SfoI* and blunt-ended to remove the acetone formation genes, while leaving the thiolase promoter region. The LtrA ORF digest product and the linearized pSOS95 vector were ligated to create the pSOS-intron vector. The pSOS-intron vector was digested by *NsiI* and *SapI* to remove the thiolase promoter and LtrA ORF. The pCONS::upp was digested with *SapI* and blunt-ended. These fragments were ligated together to generate the pCONS::upp-intron vector.

Disruption of the *hydA* gene was carried out by the insertion of sense or antisense intron II using the pCONS::upp-intron vector. The mutated 353-bp PCR fragments, *hydA* sense and *hydA* antisense described above for ClosTron inactivation and the pCONS::upp-intron vector were digested by *BsrGI* and *HindIII* and then ligated to yield the pCONS::upp-intron *hydA* sense or the pCONS::upp-intron *hydA* antisense. Both resulting vector constructs were transferred into *E. coli* TOP10 cells (Invitrogen, USA), and the transformants were selected from LB plates supplemented with 30 µg/ml chloramphenicol. Plasmids purified from the *E. coli* transformants were then transferred into *C. acetobutylicum* RM1 Δ *ac15Aupp* on reinforced clostridial agar (RCA) plates with thiamphenicol or erythromycin supplementation [74]. *Clostridium* transformants were further screened on RCA plates supplemented with thiamphenicol (30µg/ml) and 5-fluorouracil (500 µM), a counter-selectable marker [141].

The inactivation of ThlA (encoded by CAC2873) was also done by group II intron. The insertion was chosen to be at 810|811 sense strand of the CDS. Retargeting region was

constructed by SOE PCR using the four primers: thla-810|811s-IBS primer, thla-810|811s-EBS1d primer, thla-810|811s-EBS2 primer and the EBS universal primer (designed by Clostron tool using Perutka algorithm, Table 5-2). The PCR products were digested by *Bsr*GI and *Hind*III and ligated to a double digested Targetron vector pCUI [182] by *Bsr*GI and *Hind*III to yield pCUI-Thla-int810s vector. The plasmid was then transformed to *C. acetobutylicum* RM1 Δ cac15 Δ upp [167] and selected on RCA plates containing thiamphenicol 10 μ g/ml. Intron insertion in the transformants was checked by PCR using the primers flanking 810|811 position in *thlA* CDS (ThlAI-F and ThlAI-R). The removal of pCUI-Thla-int810s vector in the mutant was done by spreading the mutants on RCA plates containing 5-FU [141]. Clones resistant to 5-FU and thiamphenicol and erythromycin sensitivity were checked by PCR for the loss of the pCUI-Thla-int810s plasmid.

5.3.2 Construction and integration by single crossing over of the pEryUpp Δ thlA plasmid in the chromosome of *C. acetobutylicum* Δ upp Δ cac1502

Two DNA fragments surrounding the *thlA* gene (CAC2872) were PCR amplified with the Phusion DNA polymerase with total DNA from *C. acetobutylicum* as template and two specific couples of oligonucleotides as primers (Table 5-2). With the couples of primers thl-1/thl-2 and thl-3/thl-4, 1025 bp and 1012 bp DNA fragments were respectively obtained. Both primers thl-1 and thl-4 introduce a *Bam*HI site while primers thl-2 and thl-3 have complementary 5' extended sequences which introduce a *Stu*I site. DNA fragments Δ thlA 1-2 and Δ thlA 3-4 were joined in a PCR fusion experiment with primers thl-1 and thl-4 and the resulting fragment was cloned in the pCR4-TOPO-Blunt vector to yield pTOPO: Δ thlA. The *thlA* replacement cassette obtained after *Bam*HI digestion of the resulting plasmid was cloned, at the *Bam*HI, site into pEryUpp to yield the pEryUpp Δ thlA plasmid.

The pEryUpp Δ thlA plasmid was used to transform by electroporation *C. acetobutylicum* Δ upp Δ cac1502. Clones (resistant to erythromycin 20 μ g/ml) having inserted the pEryUpp Δ thlA plasmid by homologous recombination were selected on Petri plates and checked by PCR analysis (with primers thl-0 and thl-5 located outside of the *thlA* replacement cassette). The *C. acetobutylicum* Δ upp Δ cac1502 $thlA::$ pEryUpp Δ thlA strain with a plasmid integration in 5' was obtained.

5.3.3 Construction and integration by single crossing over of the pCatUppHydA270 plasmid in the chromosome of *C. acetobutylicum*Δ*upp*Δ*cac1502*

A cassette creating an in frame deletion of 270 bp in a region coding for the catalytic domain of the main [Fe-Fe] hydrogenase (HydA) was created using two DNA fragments surrounding this region that were PCR amplified with the Phusion DNA polymerase with total DNA from *C. acetobutylicum* as template and two specific couples of oligonucleotides as primers (Table 5-2). With the couples of primers hyd-1- hyd-2 and hyd-3– hyd-4, 1123 bp and 1008 bp DNA fragments were respectively obtained. Both primers hyd-1 and hyd-4 introduce a *Bam*HI site while primers hyd-2 and hyd-3 have complementary 5' extended sequences. DNA fragments hyd-1- hyd-2 and hyd-3– hyd-4 were joined in a PCR fusion experiment with primers hyd-1 and hyd-4 and the resulting fragment was cloned in the pCR4-TOPO-Blunt vector to yield pTOPO::hydA270. The *hydA* replacement cassette obtained after *Bam*HI digestion of the resulting plasmid was cloned, at the *Bam*HI, site into pCatUpp to yield the pCatUppΔhydA270 plasmid.

The pCatUppΔhydA270 plasmid was used to transform by electroporation *C. acetobutylicum*Δ*upp*Δ*cac1502**thlA*::pEryUppΔ*thlA*. Clones (resistant to erythromycin 20 µg/ml and thiamphenicol 20 µg/ml) having inserted the pCatUppΔhydA270 plasmid by homologous recombination were selected on Petri plates and checked by PCR analysis (with primers Dhyd-0 and Dhyd-3 located outside of the *hydA* replacement cassette). The *C. acetobutylicum*Δ*upp*Δ*cac1502**thlA*::pEryUppΔ*thlA**hydA*:: pCatUppΔhydA270 strain with integration of both plasmids in 5' was obtained.

5.3.4 Simultaneous inactivation of *thlA* and *hydA* in *C. acetobutylicum*Δ*upp*Δ*cac1502*

To simultaneously inactivate *thlA* and *hydA*, two colonies of the *C. acetobutylicum*Δ*upp*Δ*cac1502**thlA*::pEryUppΔ*thlA**hydA*:: pCatUppΔhydA270 strain were cultured for 24 hours in liquid MS Glucose medium and then subcultured in liquid 2YTG medium without antibiotic. Appropriate dilutions were plated on RCA with 5-FU at 400 µM. To select integrants having excised and lost both the pEryUppΔ*thlA* and the pCatUppΔhydA270 plasmids, 5-FU resistant clones were replica plated on both RCA + 5FU, RCA with erythromycin 40 µg/ml and RCA with thiamphenicol at 40 µg/ml. Clones resistant

to 5-FU and sensitive to erythromycin and thiamphenicol were checked by PCR analysis (with primers Dthla-0 and Dthla-5 located outside of the *thlA* replacement cassette and primers thl-D and thl-R located inside of *thlA* and with primers DhydA-0 and DhydA-3 located outside of the *hydA* replacement cassette and primers hyd-D and hyd-R located inside of *hydA*). About half of the clones resistant to 5-FU and sensitive to erythromycin and thiamphenicol possess inactivated *thlA* and *hydA* genes. The *C. acetobutylicum* $\Delta upp \Delta cac1502 \Delta thlA \Delta hydA270$ strain was obtained.

The strain was evaluated in MS glucose batch culture [155] without pH regulation. Ethanol and lactate were the two major products with small amount of acetate and glycerol.

Table 5-2 List of primers

Name	Sequence
thla-810 811s-IBS	AAAAAAGCTTATAATTATCCTTACTTGGCGTAAAAGTGC GCCCAGATAGGGTG
thla-810 811s-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGTAA AACCTAACTTACCTTTCTTTGT
thla-810 811s-EBS2	TGAACGCAAGTTTCTAATTTTCGGTTCCAAGTCGATAGAG GAAAGTGTCT
EBS universal	CGAAATTAGAACTTGC GTTCAGTAAAC
HYDAs(IBS)	AAAAAAGCTTATAATTATCCTTAATTACCATCCTTGTGC GCCCAGATCGGGTG
HYDAs(EBS1d)	CAGATTGTACAAATGTGGTGATAACAGATAAGTCATCC TTGATAACTTACCTTTCTTTGT
HYDAs(EBS2)	TGAACGCAAGTTTCTAATTTTCGGTTGTAATCCGATAGAG GAAAGTGTCT
HYDAas(IBS)	AAAAAAGCTTATAATTATCCTTATATCTCCTGGTAGTGC GCCCAGATAGGGTG
HYDAas(EBS1d)	CAGATTGTACAAATGTGGTGATAACAGATAAGTCCTGG TAAATAACTTACCTTTCTTTGT
HYDAas(EBS2)	TGAACGCAAGTTTCTAATTTTCGGTTAGATATCGATAGAG GAAAGTGTCT
ThlaI-F	GTTACAGCTGGTAATGCATCAGG
ThlaI-R	GGTCCATATCCCATTATTGCTGGG
hydAI-F	TAAATGGCAATGAAGTGCATACAGATAAAG
hydAI-R	CATGTTGAGTCATCAAGACAAACATCATCA
Hyd-1	AAAGGATCCGTTTTTCTTAATATTTACCATATTGCACCT CCC
Hyd-2	ATATCTCTTAAGCTGTTAGTTTCCATTATAGTCATATCTG CACCAAAG
Hyd-3	CAGATATGACTATAATGGAACTAACAGCTTAAGAGAT ATTGATGCATCC
Hyd-4	AAAGGATCCCTGGTACATCAGTATACGAAACAATGCC
DhydA-0	CTGTATCCCTGCTTATAAATCTTTGAGC
DhydA-3	GTTGTTTGAGGGAGCTAATCATTCTC
Hyd-4F	CTATATACTGCACTTAGAATGTTAGGCTTTG
thl-1	AAAAGGATCCAAGCAGTTAATGAAAAGAATATTTTTAT TACAGGAAATAC
thl-2	GTTATTTTTTAACAATACTTTAGGCCTTACGGGGTAACAG ATAAACCATTTCATCTA
thl-3	AATTTAGGAGGTTAGTTAGAAGGCCTAAAGTATTGTTA AAAATAACTCTGTAGAATTATAAATTAG
thl-4	AAAAGGATCCAAGTTAACAATCATTTCTATTACGCTTTG TTTCC
Thl-D	TGGAACATTTCAAGAGAAGAACAAGATGAG

Thl-R	GCTCCTCCATTTACATTTACTTTATTCATATC
DthlA-0	ACATGGAGATACGACTACAACATTTGCTG
DthlA-5	TTCTTTTTATTGCAGTTGCATTTATTA AAAAATGC

Chapter 6 Cap0037, a novel global regulator of *Clostridium acetobutylicum* metabolism

Ngoc-Phuong-Thao Nguyen¹, Sonja Linder³, Stefanie K Flitsch³, Bettina Schiel-Bengelsdorf³, Peter Dürre³ and Philippe Soucaille^{1,2*}

¹LISBP, Université de Toulouse, CNRS, INRA, INSA, Toulouse, France

²Metabolic Explorer, Biopole Clermont-Limagne, Saint-Beauzire, France

³Institute of Microbiology and Biotechnology, University of Ulm, Ulm, Germany

* **Correspondence:** Philippe Soucaille, Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés, I.N.S.A, Toulouse, 31400, France

soucaille@insa-toulouse.fr

Keywords: transcriptomics₁, Cap0037₂, *Clostridium acetobutylicum*₃, global regulator₄, metabolism₅

6.1 Abstract

An operon comprising two genes, *CA_P0037* and *CA_P0036*, which encode proteins of unknown function that were previously shown to be highly expressed in acidogenic cells and repressed in solventogenic and alcohologenic cells, is located on the pSOL1 megaplasmid of *Clostridium acetobutylicum* upstream of *adhE2*. A *CA_P0037::int (189/190s)* mutant was successfully generated by the Targetron technique. The resultant mutant showed significantly different metabolic fluxes patterns in acidogenic (producing mainly lactate, butyrate and butanol) and alcohologenic (producing mainly butyrate, acetate and lactate) chemostat cultures but not in solventogenic or batch cultures. Transcriptomic investigation of the *CA_P0037::int (189/190s)* mutant showed that inactivation of *CA_P0037* significantly affected the expression of more than 258 genes under acidogenic conditions. Surprisingly, genes belonging to the Fur regulon, involved in iron transport (*CA_C1029-CA_C1032*), or coding for the main flavodoxin (*CA_C0587*) were the most significantly expressed genes under all conditions, whereas *fur* (coding for the ferric uptake regulator) gene expression remained unchanged. Furthermore, most of the genes of the Rex regulon, such as the *adhE2* or *ldhA* genes, and of the PerR regulon, such as *rbr3A-rbr3B* or *dfx* were over-expressed in the mutant. In addition, the whole *CA_P0037-CA_P0036* operon was highly expressed under all conditions in the *CA_P0037::int (189/190s)* mutant, suggesting a self-regulated expression mechanism. *CA_P0037* was shown to bind to the *CA_P0037-CA_P0036* operon, *sol* operon and *adc* promoters, and the binding sites were determined by DNA footprinting. Finally, a putative *CA_P0037* regulon was generated using a bioinformatics approach.

Importance Currently, there is a resurgence of interest in *Clostridium acetobutylicum* the biocatalyst of the historical Weizmann process, to produce n-butanol for use both as a bulk chemical and as a renewable alternative transportation fuel. To develop a commercial process for the production of n-butanol via a metabolic engineering approach, it is necessary to better characterize both the primary metabolism of *C. acetobutylicum* and its regulation. This study describes the discovery of Cap0037, a novel regulator which, when inactivated, drastically affects the primary metabolism of *C. acetobutylicum* under both acidogenic and alcohologenic fermentation conditions.

6.2 Introduction

Clostridium acetobutylicum is a Gram-positive, strictly anaerobic, spore-forming bacterium now considered the model organism for the study of solventogenic Clostridia [185,186]. The superiority of butanol over ethanol as an alternative biofuel has attracted research interest in *C. acetobutylicum* and other recombinant bacteria producing butanol as a major product [187].

In batch cultures, *C. acetobutylicum* shows a bi-phasic fermentation pattern: after producing acetate and butyrate during exponential growth, the microorganism switches to the formation of acetone, butanol and ethanol (ABE, at a ratio of 3:6:1) shortly before entering the stationary phase [185].

In phosphate-limited chemostat cultures, *C. acetobutylicum* can also be maintained in three different stable metabolic states [20–22,181,188] without cellular differentiation [189]: acidogenic (producing acetate and butyrate) when grown at a neutral pH with glucose; solventogenic (producing acetone, butanol, and ethanol) when grown at a low pH with glucose; and alcohologenic (forming butanol and ethanol but not acetone) when grown at a neutral pH under conditions with high NAD(P)H availability [21,181,190].

Understanding the complex regulatory network of *C. acetobutylicum* metabolism is crucial for further manipulating the genotype to obtain industrial strains, but our understanding of the cellular functioning remains very limited. Quantitative transcriptomics, proteomics and fluxomics analyses were established to characterize the three typical metabolic states of *C. acetobutylicum* in chemostat cultures [42]. Many studies on the solventogenic shift were conducted [23,191,192]. Transcriptional changes in response to metabolite stresses such as acetate, butanol and butyrate stress were also conducted to find regulons and DNA binding motifs of transcriptional factors [139,193–197].

In terms of regulatory information, some common regulators have already been studied. A peroxide repressor (PerR)-homologous protein was identified to be a key repressor that plays an important role in the oxidative stress defense [128,129]. In the same family as PerR, a ferric uptake regulator (Fur), which helps *C. acetobutylicum* sense and respond to iron

availability at multiple levels using a sophisticated system, was identified and well characterized [130]. Spo0A was found to directly control the solventogenic switch, the expression of the *adc* and *ptb* genes [67] and the activation of the sporulation program in *C. acetobutylicum* [92]. The carbon storage regulator (CsrA) was also shown to directly or indirectly regulate multiple pathways including flagella assembly, oligopeptide transport, iron uptake, phosphotransferase systems (PTS), synthesis of riboflavin, stage III sporulation and central carbon metabolism [131]. A catabolite control protein (CcpA) was mutated at V302N to improve the utilization of xylose and solvent production [132]. RegA, an homolog of CcpA, was shown to inhibit the degradation of starch by the *staA* gene product [198]. Recently, a redox sensing transcriptional repressor (Rex) has been found to modulate its DNA-binding activity in response to the NADH/NAD⁺ ratio and plays a role in the alcohologenic shift of *C. acetobutylicum* [110,134].

The regulatory network of *C. acetobutylicum* remains largely unexplored. In this study, we identified Cap0037 to be a novel self-regulated protein that globally affects *C. acetobutylicum* metabolism. *CA_P0037* is the first gene of a polycistronic operon with *CA_P0036*. This operon is located upstream of *adhE2* on the pSOL1 megaplasmid and it was significantly expressed in acidogenic cells and repressed in solventogenic and alcohologenic cells [42,199]. Furthermore, the two proteins were present at a 1:1 molar ratio and were among the 30 most abundant proteins in the cell (around 50,000 molecules per cell) [42]. Another study found that Cap0037/Cap0036 expression was slightly down-regulated in an *adc* mutant, clearly up-regulated during solventogenesis in a *pta* mutant and an *adc/pta* double mutant and strongly repressed in a *ptb* mutant independent of the cultivation conditions [111]. Interestingly, Cap0037 and Cap0036 were found to appear in multiple spots (up to five in the case of Cap0037 and up to at least four in the case of Cap0036) in a 2-D gel analysis of acidogenic cells, suggesting posttranslational modifications such as phosphorylation, acetylation, glycosylation or methylation [199].

In order to investigate the roles of these proteins in the acidogenic state of *C. acetobutylicum*, we inactivated *CA_P0037* via the mobile group II intron. The comparison between wild-type and mutants lacking *CA_P0037* was conducted at the system scale level

through a combination of quantitative fluxomics, transcriptomics and proteomics analysis. DNA footprinting was used to determine DNA binding motifs of Cap0037 on its own promoter and in the upstream region of selected genes shown to be up- or down-regulated. Putative Cap0037 binding sites across *C. acetobutylicum* genomes were analyzed via a bioinformatics approach. The regulatory role of Cap0037 and the observed phenotypes will be discussed in light of the different omics datasets. This study builds a foundation to further investigate the regulatory mechanism of this novel protein in *C. acetobutylicum*, which should be beneficial for understanding the role of homologous proteins in other Firmicutes.

6.3 Results and Discussion

6.3.1 The phylogenetic tree and bioinformatics analysis of Cap0037/Cap0036

Cap0037 and Cap0036 are proteins of unknown function. They are well conserved in other Firmicutes such as *Clostridia* (*C. scatologenes*, *C. drakei* and *C. carboxidivorans*), *Bacillus* (*B. cereus* and *B. shackletonii*) and *Geobacillus* (*G. thermoglucosidasus* and *G. caldxylosilyticus*) (Figure 6-1). All of these proteins have never been characterized.

Cap0036 was found to be homologous to some cytoplasmic proteins of the *Bacillus* species (such as *Bacillus enclensis*, *Bacillus shakletoni* and *Bacillus cereus*) and also to some bactofilin family proteins from the *Geobacillus* species (Figure 6-1A). Two bioinformatics tools, TMpred [200] and OCTOPUS [201], were employed for the prediction and identification of possible transmembrane domains in Cap0036 without success. Therefore, Cap0036 is a cytoplasmic protein rather than a membrane bactofilin family protein.

Noticeably, Cap0037 is homologous to the YgaS-like protein in *Bacillus cereus* G9241 (a putative-DNA-binding-like protein), although the relationship is distant (Figure 6-1B). A helix-turn-helix DNA binding motif (ISKKELLEITHISYGQLYRWKR) was found in the N-terminus of the Cap0037 protein sequence (by NPS@, [202]). As a result, Cap0037 is likely a novel transcriptional regulator. The 1:1 expression ratio of Cap0037 and Cap0036 [42] suggested an interaction of these two proteins that forms a system acting as an unknown regulatory mechanism in *C. acetobutylicum*. Interestingly, a transmembrane domain (LIRKLGIAICFLMLIPNEIYI) and an outer membrane region

(ENTASLVLVKVNARECLEELKSKLSL) were found at the C-terminus of Cap0037 (OCTOPUS program, [201]).

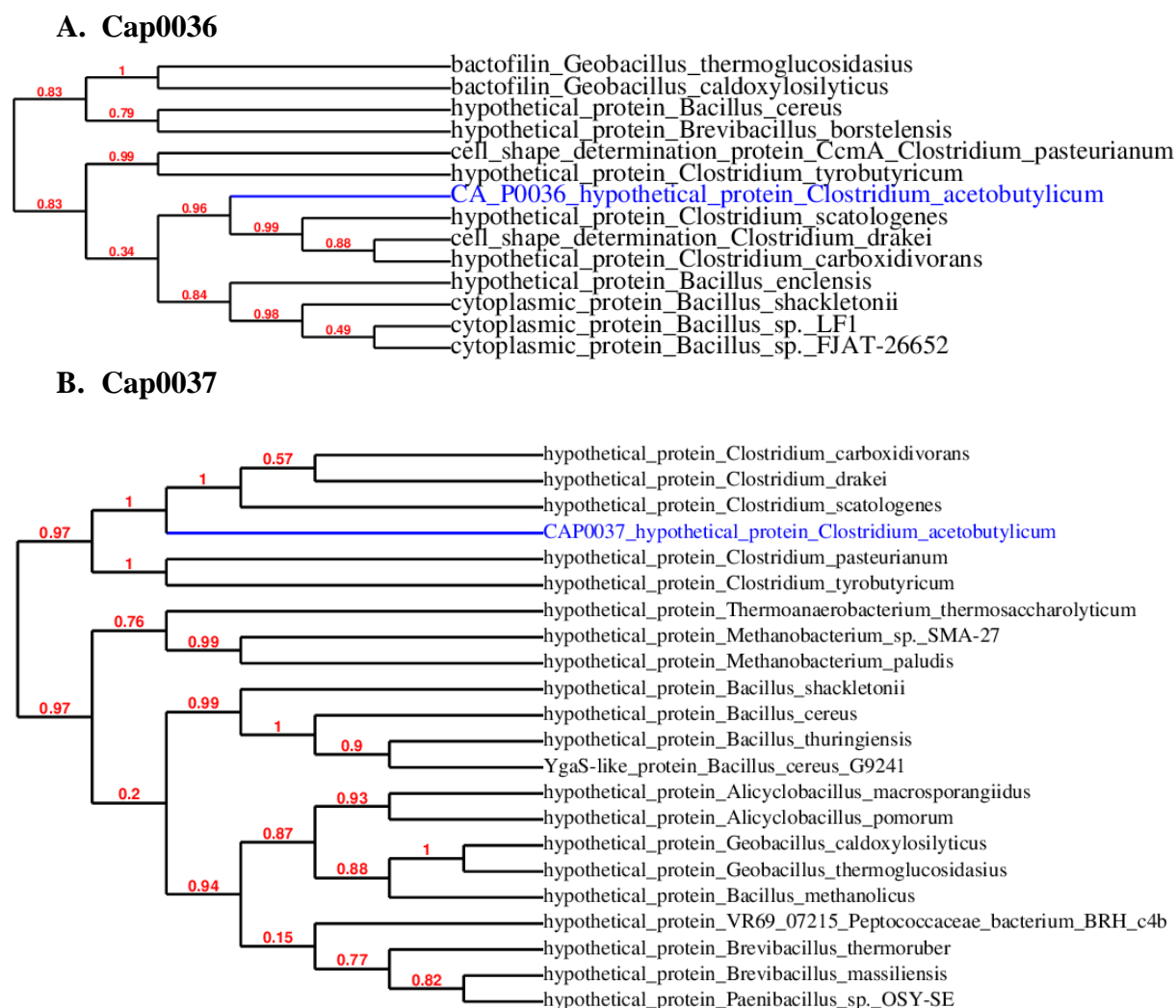


Figure 6-1 Phylogenetic trees of Cap0036 (A), Cap0037 (B) sequences from *Clostridium acetobutylicum* and neighbor proteins from other bacteria.

The neighbor proteins were selected by running BLAST on NCBI database. Cap0036 and Cap0037 from *C. acetobutylicum* were shown in blue.

6.3.2 Disruption of CA_P0037 by pCUI-cap37 (189s)

The Clostron technique was first applied to inactivate *CA_P0037*. Clostron plasmid pMTL007C-E2::cap37 (277a), targeting an insertion into the 277/278 antisense of *CA_P0037* CDS, was constructed and used as described elsewhere [84,85]. Intron insertion in *CA_P0037* was not detected in all of the obtained integrants. Therefore, we used another intron-based method developed in our laboratory [182] for knocking out *CA_P0037* based on the pCUI-cap0037 vector as mentioned in the Methods section. The mutant was successfully obtained with the use of plasmid pCUI-cap37: a 1 kbp-long intron (Figure 6-2A) was inserted at the 189/190 sense strand position of *CA_P0037*. PCR using two external primers flanking the targeted position confirmed the insertion of the intron into the gene (Figure 6-2B). After selecting for the Cap37::int mutant, the strain was subjected to 5'-FU selection to cure the pCUI-cap37 plasmid to avoid intron splicing by LtrA, whose encoding gene is present on the plasmid backbone. Southern blotting was conducted to confirm a single intron insertion into the genome (Figure 6-2C).

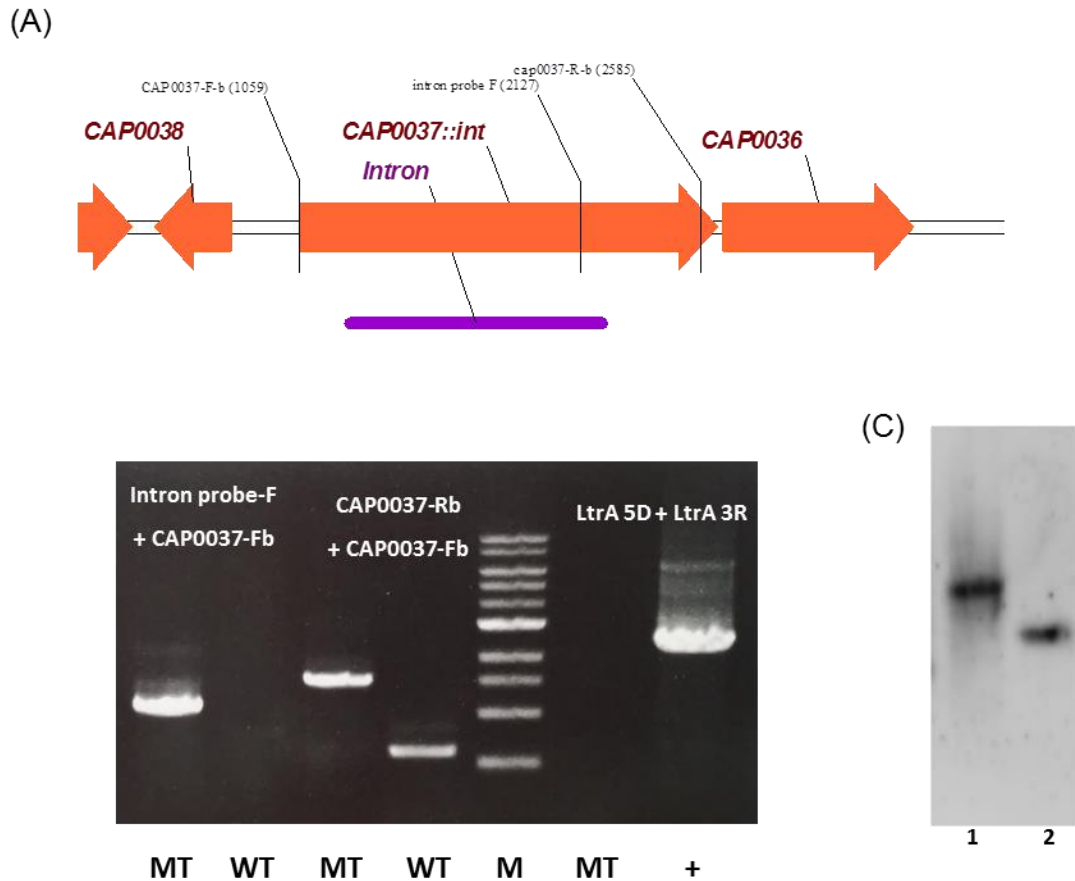


Figure 6-2 Construction of Cap37::int mutant.

(A) DNA configuration of mutant with group II intron inserted at 189/190s. (B) PCR on genomic extraction of mutant (MT) and wild-type (WT) by different couples of primers to confirm intron insertion (by Intron probe-F and CAP0037-Fb, CAP0037-Rb and CAP0037-Fb) and to detect the presence of pCUI1-cap37(189s) (by LtrA 5D and LtrA 3R). (C) Southern blot for intron insertion in Cap37::int mutant. The intron probe was hybridized to *Eco*RI digested genomic DNA of Cap37::int mutant (lane 1) and to the positive control (lane 2) with expected size of 4388 bps and 1700bps, respectively.

6.3.3 Carbon and electron fluxes of the Cap37::intron mutant under different physiological conditions

The Cap0037::intron mutant was first evaluated under acidogenic conditions and normalized metabolic fluxes compared to previously published data for the control strain. The mutant strain underwent a profound change in its metabolism: butyrate and acetate normalized fluxes decreased 2.5- and 5.7-fold, respectively, while its lactate and butanol normalized fluxes increased 85- and 9.7-fold, respectively, compared to the control strain (Figure 6-3A). Ethanol normalized fluxes, on the other hand, remained unchanged. Such drastic flux changes were accompanied by a 2.9-fold decrease in hydrogen production and by other changes in electron fluxes (Figure 6-4), which are described in detail below. In a complementation experiment, the Cap0037::intron mutant was transformed with the pSOS95-Cap0037 plasmid to express the *CA_P0037* gene under the control of the *thlA* promoter. When this strain was evaluated in acidogenic chemostat culture, in the presence of clarithromycin (40 µg/ml) for the maintenance of the plasmid, the normalized metabolic fluxes were not significantly different from the wild-type strain (Figure 6-3A), demonstrating that the Cap0037::intron mutant phenotype was only attributable to *CA_P0037* inactivation. The production of butanol by the Cap0037::intron strain under acidogenesis is explained by the higher *adhE2* expression (~50-fold higher than the control strain, with 21 mRNA molecules/cell) (Figure 6-5), and although the expression of the *sol* operon (*ctfA*, *ctfB*, *adhE1*) was increased ~6-fold, the amount of mRNA molecules remained low (0.5) compared to *adhE2* (21). The increase in lactate formation was associated with an 83-fold increase in *ldhA* expression (Figure 6-5, Dataset S1). For the Cap0037::intron mutant, the acetate flux decreased 5.7-fold compared to the control strain, although *pta-ack* (*CA_C1742–CA_C1743*) expression was unchanged (Figure 6-5, Dataset S1). Thus, flux is controlled at the enzyme level via a decrease in the acetyl-CoA pool, probably due to the 2-fold higher expression of most of the genes (*CA_C2708–CA_C2712*) coding for the enzymes converting acetoacetyl-CoA to butyryl-CoA.

The alcohologenesis state (pH 6.3, in mixture of glucose and glycerol) was the second most affected state following the disruption of *CA_P0037* (Figure 6-3B). Although all of the glucose provided in the medium was consumed, the normalized flux of glycerol consumption

was reduced more than 15-fold compared to the control strain. While the control strain produced mainly ethanol (normalized flux of 8.6 %), butyrate (normalized flux of 25.1 %) and butanol (normalized flux of 61.1 %), the mutant produced mainly acids: butyrate (normalized flux of 63.6 %), acetate (normalized flux of 21.9 %) and lactate (normalized flux of 11.8 %) (Figure 6-3B). The ethanol and butanol fluxes were 1.8-fold and 9.8-fold lower than the parental strain, respectively. In the mutant, the alcohologenesis state was barely induced, and it behaved like the control strain in the acidogenesis state. Such drastic flux changes were accompanied by a 1.4-fold increase in hydrogen production and by other changes in electron fluxes (Figure 6-4), which are described in detail below. The lower glycerol consumption of the Cap0037::intron strain is explained by the lower expression (4–6-fold decrease) of the gene cluster coding for glycerol transport and utilization (*CAC1319–CAC1323*). The increase in lactate formation was associated with a 32-fold increase in *ldhA* expression and the lower production of butanol was associated with a 10-fold lower *adhE2* expression. On the other hand, the butyrate flux increase was not associated with a change in the expression of *ptb-buk* (*CA_C3074–CA_C3075*), which remained unchanged (Figure 6-5, Dataset S1). Thus, flux is controlled at the enzyme level via an increase in the butyryl-CoA pool, probably due to the 10-fold lower expression levels of *adhE2*.

The mutant strain was then evaluated under solventogenic conditions and compared to previously published data for the control strain. The Cap0037::intron mutant was not able to consume all glucose and therefore the titers of products were lower, but the fluxes were not significantly different from the wild-type strain (Figure 6-3C). Both the *sol* operon and *adhE2* were more expressed in the mutant (2- and 4-fold, respectively), although the amount of mRNA molecule for *adhE1* (12) was much higher than for *adhE2* (~1). Surprisingly, the expression of *adc* was down-regulated 3-fold in the mutant, although the amount mRNA molecules per cell (4.3) remained high.

The butanol pathway was analyzed for three different conditions in the Cap0037::intron mutant by calculating the contribution of each of the five enzymes potentially involved in each of the two steps to the fluxes (see [42] for the calculation). Under acidogenesis, *adhE2* (21 mRNA molecules per cell, 5×10^4 protein molecules per cell) was more highly expressed

than *adhE1* (0.5 mRNA molecules per cell, 4×10^3 protein molecules per cell), and AdhE2 was the main contributor (93% of the flux) to the conversion of butyryl-CoA to butyraldehyde in the Cap0037::intron mutant (Figure 6-6A). The butanol dehydrogenase flux is primarily attributable to AdhE2 (92% of the flux), BdhB (5% of the flux), BdhA (2% of the flux) and BdhC (1% of the flux), in decreasing order of activity. Under solventogenesis, *adhE1* (12 mRNA molecules per cell, 10^5 protein molecules per cell) was more expressed than *adhE2* (1 mRNA molecules per cell, 2.2×10^3 protein molecules per cell), and *adhE1* was the main contributor (98% of the flux) to the conversion of butyryl-CoA to butyraldehyde in the Cap0037::intron mutant (Figure 6-6C). Similar to the control strain, the butanol dehydrogenase flux in the Cap0037::intron mutant is primarily due to BdhB (82% of the flux), AdhE2 (7%), BdhA (5% of the flux) and BdhC (3% of the flux), in decreasing order of activity. Under alcohologenesis, although the butanol flux decreased 9.8-fold, the Cap0037::intron mutant behaved the same as the control strain in terms of enzyme contribution, as AdhE2 was mainly responsible for both the conversion of butyryl-CoA to butyraldehyde (97% of the flux) and the conversion of butyraldehyde to butanol (70% of the flux).

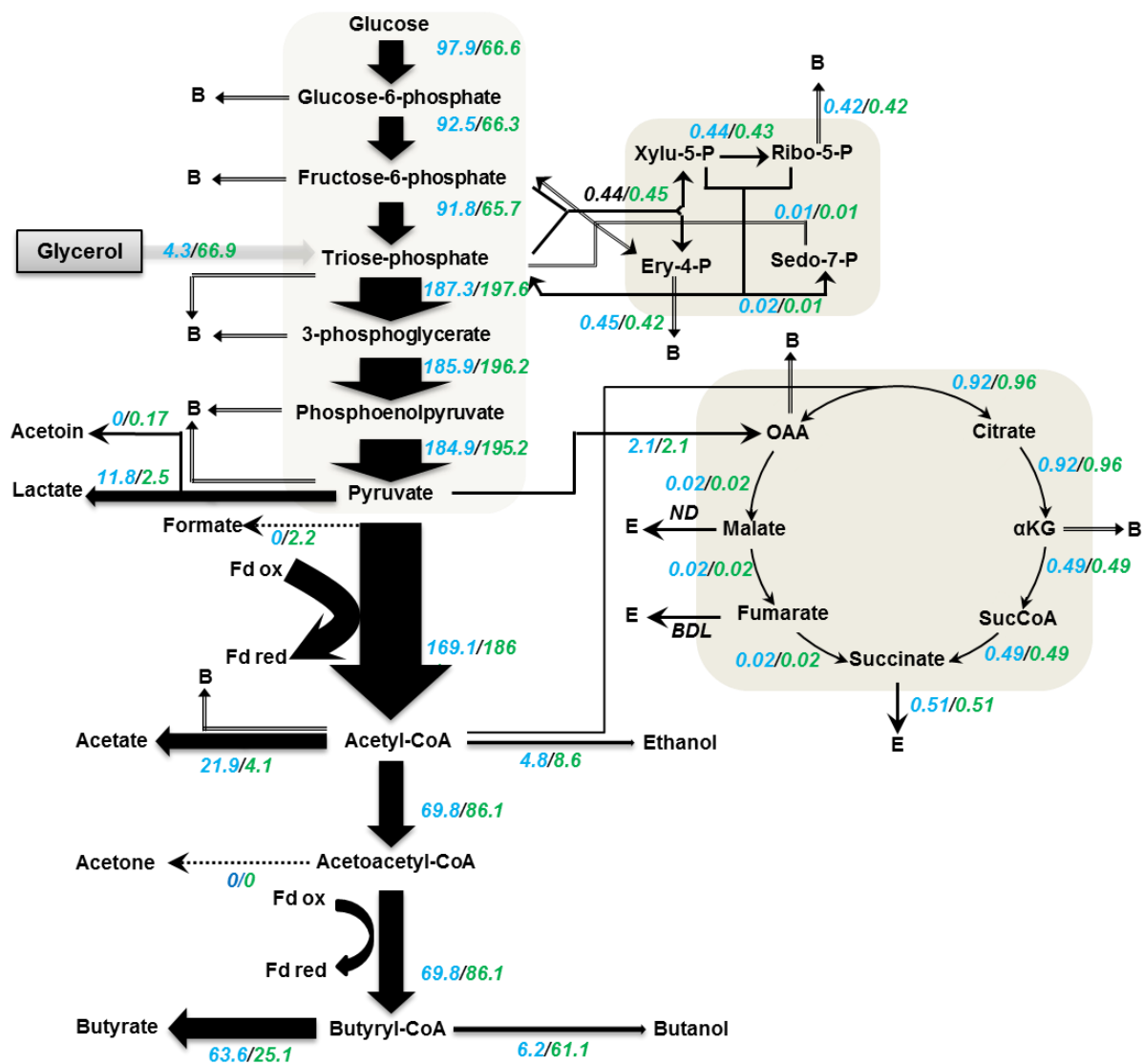
The normalized electron fluxes were analyzed for the Cap0037::intron mutant under three different conditions (Figure 6-4). In acidogenesis, the primary use of reduced ferredoxin was switched from hydrogen to NADH production in response to the high expression of *adhE2* and *ldhA* in the mutant. In acidogenesis, the normalized hydrogen production fluxes decreased ~3-fold while the normalized fluxes of NADH production from reduced ferredoxin increased 2.4-fold. The decrease in hydrogen production was not attributable to the expression of *hydA* (*CA_C0028*), which remained unchanged, but to the lower flux of reduced ferredoxin production, as due to high *ldhA* expression lactate became the major product. The 2-fold increase in the normalized Ferredoxin-NAD⁺ reductase fluxes might be explained by the 500-fold higher expression of *CA_C0587* encoding the major flavodoxin and the 1.5-fold lower expression of *CA_C0303* encoding the major ferredoxin, as well as the possibility that reduced flavodoxin might be a better substrate than ferredoxin for the Ferredoxin-NAD⁺ reductase. In solventogenesis, the normalized electron fluxes were not significantly changed in agreement

with the conserved product pattern. In alcohologenesis, butyrate was the main fermentation product of the Cap0037::intron mutant, and the normalized hydrogenase flux increased 1.4-fold while the normalized Ferredoxin-NAD⁺ reductase fluxes decreased 2.6-fold. The high normalized hydrogen production and low normalized Ferredoxin-NAD⁺ reductase fluxes and normalized butanol fluxes can be explained by the 5-fold down-regulation of *CA_C3486* encoding a potential multimeric flavodoxin and the proposed physiological role [42] in alcohologenic culture of the wild-type strain, providing a better substrate for the Ferredoxin NAD⁺ reductase than for hydrogenase.

The metabolic map illustrates the central carbon metabolism of *E. coli*. Key features include:

- Glycolysis/Gluconeogenesis:** The main pathway from Glucose to Pyruvate, involving intermediates like Glucose-6-phosphate, Fructose-6-phosphate, Triose-phosphate, 3-phosphoglycerate, Phosphoenolpyruvate, and Pyruvate. Fluxes are shown in blue, green, and red boxes.
- Biosynthesis:** Various metabolites are diverted for biomass (B) production, including Glucose-6-phosphate, Fructose-6-phosphate, Triose-phosphate, 3-phosphoglycerate, Phosphoenolpyruvate, Pyruvate, Acetyl-CoA, and Butyryl-CoA.
- Catabolism:** Pyruvate can be converted to Lactate, Acetoin, or Formate. Acetyl-CoA can be converted to Ethanol or Acetone. Butyryl-CoA can be converted to Butanol.
- TCA Cycle:** Pyruvate enters the TCA cycle via Acetyl-CoA, producing intermediates like Citrate, αKG, Succinyl-CoA, Succinate, Fumarate, Malate, and Oxaloacetate (OAA). Fluxes are shown in blue, green, and red boxes.
- Other Pathways:** Pyruvate can be converted to Xylulose-5-phosphate (Xylul-5-P) and Sedoheptulose-7-phosphate (Sedo-7-P), which are involved in gluconeogenesis. Pyruvate can also be converted to Ethanol and Acetoin.
- Legend:**
 - B:** Biomass
 - BDL:** below detection limit
 - E:** Excretion
 - ND:** not determined

B. Alcohologensis



C. Solventogenesis

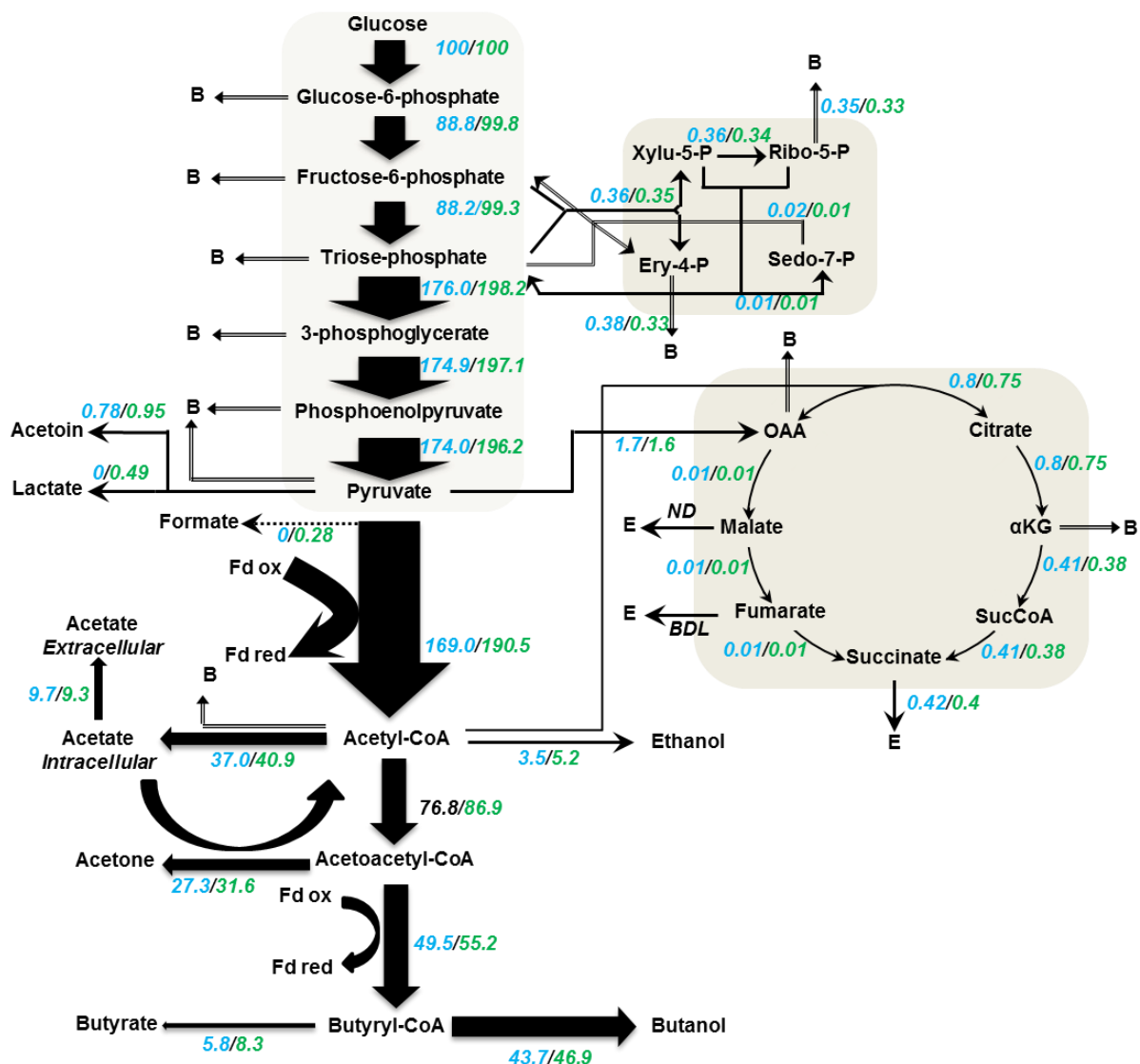


Figure 6-3 Metabolic fluxes of mutant Cap0037::int versus control strain in different metabolic states (A) Acidogenesis (B) Alcohologenes (C) Solventogenesis.

All values are normalized to the flux of the initial carbon source (mmol/gDCW/h). Glucose flux is normalized as 100 for acidogenesis and solventogenesis, and the sum of glucose and half of the glycerol normalized as 100 for alcohologenes. The values of corresponding mutant are shown in blue letter, and those of control strain are shown in green letter. The control data was from [42]

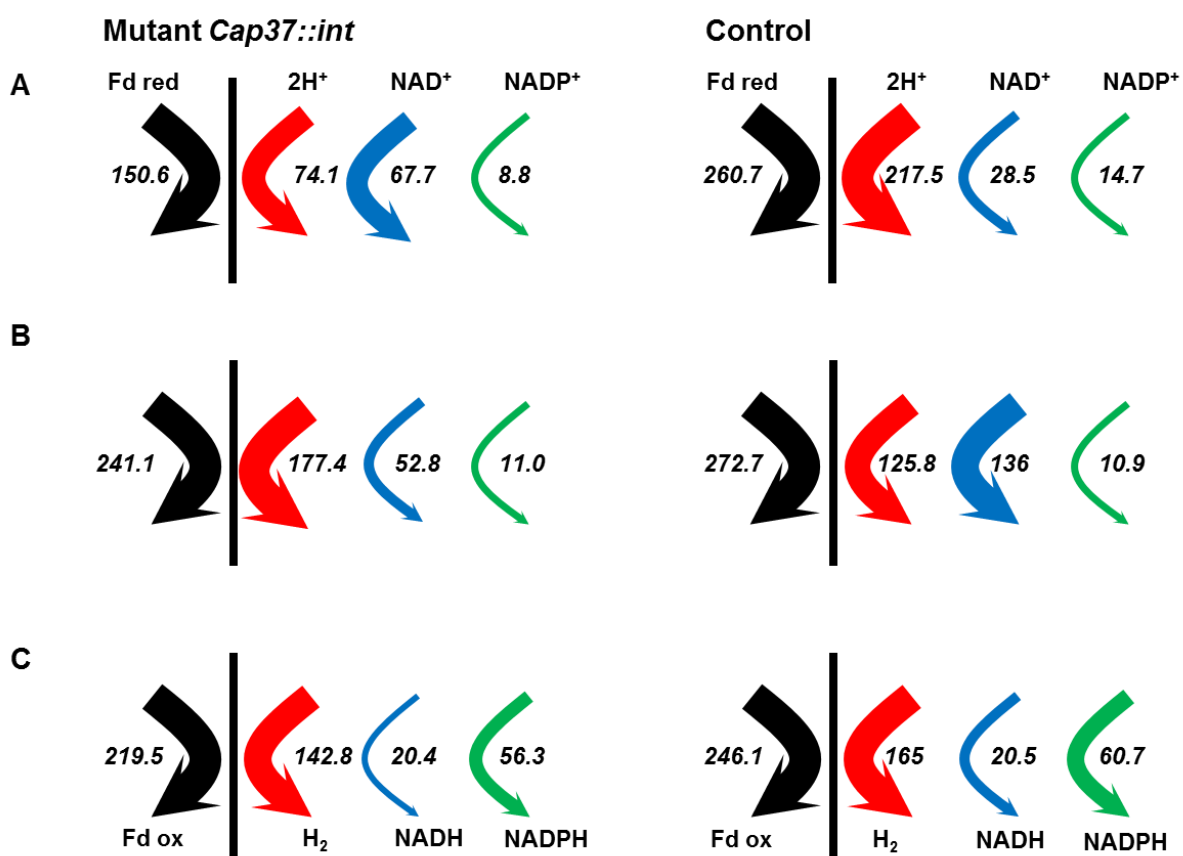


Figure 6-4 Electron fluxes of mutant *Cap0037::int* versus control strain in different metabolic states (A) Acidogenesis (B) Alcohologenesis (C) Solventogenesis.

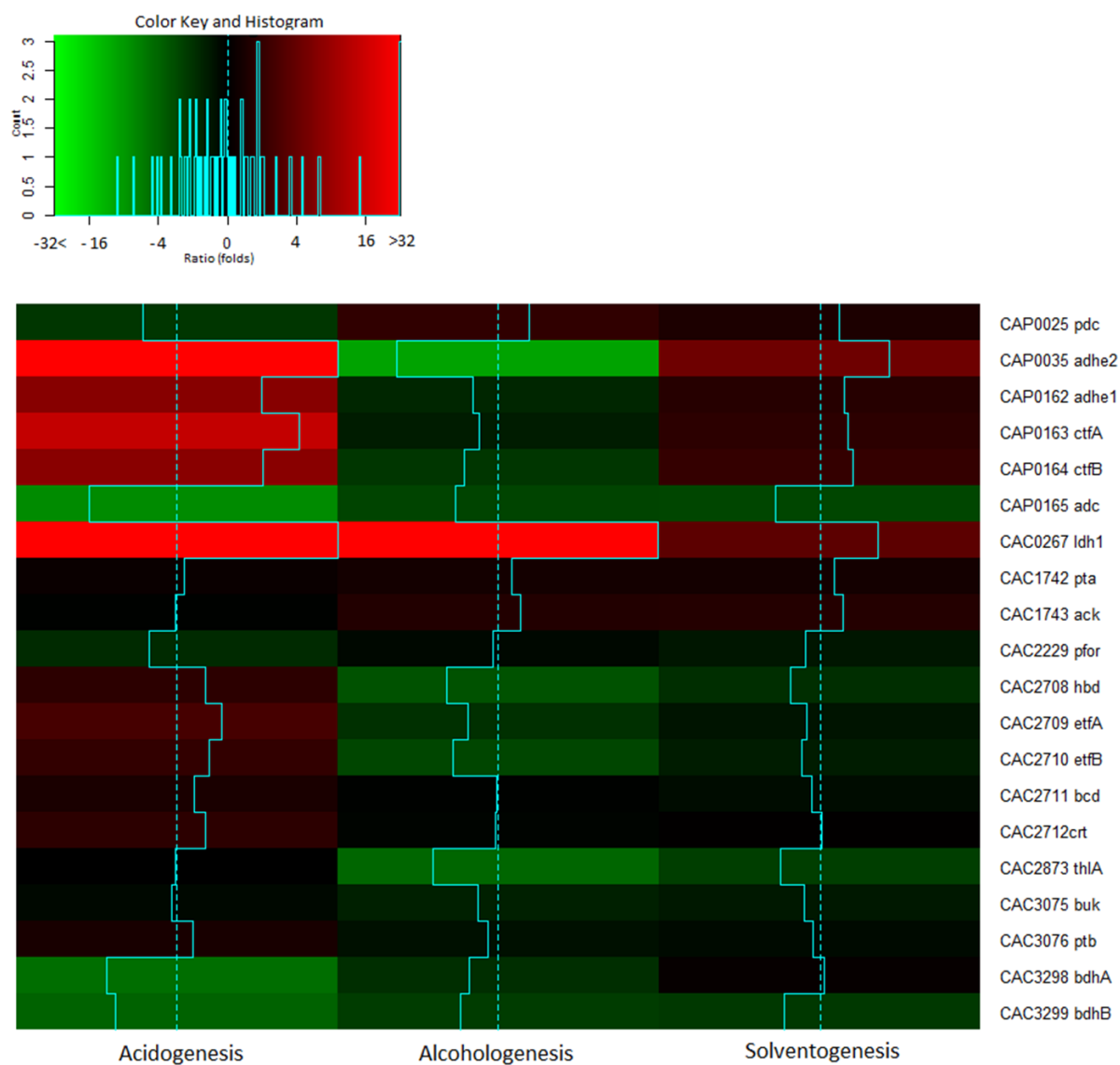


Figure 6-5 Relative transcript levels of central metabolic genes of Cap0037::int mutant in acidogenesis, alcohologenesis and solventogenesis conditions.

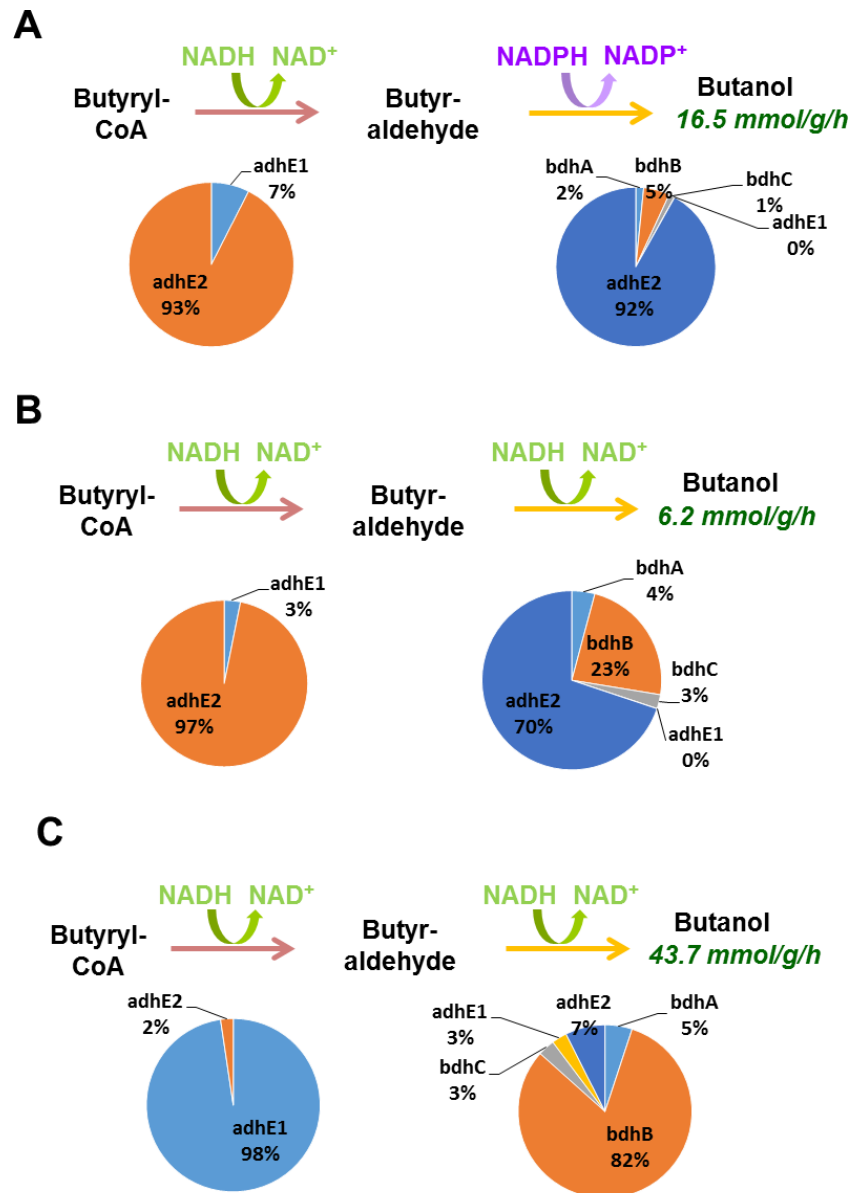


Figure 6-6 The butanol pathway in the Cap0037 ::int mutant (MT) and the wild-type (WT) in acidogenesis (A), alcohologenesis (B) and solventogenesis (C)

6.3.4 Determination of the CA_P0037 DNA binding site (BS)

To study the possible regulatory function of Cap0037 in central metabolism, we analyzed the ability of Cap0037 to bind to the promoter regions of several genes coding for key enzymes of the central metabolism of *C. acetobutylicum* that were either up-regulated (CA_P0037/CA_P0036, *ldhA*, the *sol* operon, *fldI* (CA_C0587)) or down-regulated (*adc*) in the Cap0037::int mutant under acidogenesis. A shift was only observed for the promoter regions of CA_P0037/CA_P0036, the *sol* and *adc* operon (Figure 6-7 and Figure 6-8), indicating that the up-regulation of *ldhA* and *fldI* (CA_0587) was an indirect effect of CA_P0037 inactivation (data for *ldhA* and *fldI* promoters are not shown).

We used DNA footprinting to determine the DNA binding sites of Cap0037 in the promoter regions of CA_P0037, the *sol* operon and *adc* (Figure 6-7 and Figure 6-9). Cap0037 was found to bind to the 5'-ATATTTTCATATAAAT-3' sequence overlapped the putative -10 region of the CAP0037-CAP0036 operon promoter (determined by BPROM tool by [203]), suggesting that Cap0037 might repress its own transcription. This hypothesis is in agreement with the 50.7- and 68.6-fold higher expression levels of the CA_P0037- CA_P0036 operon observed in alcohologenesis and solventogenesis of the Cap0037::intron mutant. Two binding sites were found in the *adc* promoter region. The first was 5'-ATAAGTTTATATAAAT-3', located upstream of the -35, which contains three mismatches compared to the binding site found in the CA_P0037 promoter. Cap0037 could potentially bind to this sequence to activate the expression of *adc*, which fit with the transcriptomic data for Cap0037::int mutant, in which *adc* was down-regulated under all conditions. The second binding site in the *adc* promoter region, 5'-TAATGTAAATATAAAT-3', is less conserved compared to the CAP0036-CAP0037 operon promoter binding with five mismatches, and it is located between the -10 and +1 promoter region suggesting that it might be also involved in the repression of *adc* under conditions that have not been characterized yet. It is worth noting that a Spo0A~P binding box was found upstream of the -35 region of the *adc* promoter [67]. This box is only six base pairs downstream of the first Cap0037 binding site (Figure 6-10). Hence, there might be some synergy between Spo0A~P and Cap0037 for the activation of *adc*.

In the *sol* promoter region, Cap0037 binds to the 5'-TAAATATACTGATAAT-3' sequence located upstream the Spo0A~P and the -35 region of the *sol* operon promoter (Figure 6-10). Interestingly this Cap0037 binding site overlapped the R1 sequence (TCTAAATATA) that has been proposed [204] to be the binding site of an additional activator protein, which acts in concert with Spo0A~P to activate the transcription of the *sol* operon. Binding of CAP0037 would prevent the binding of this activator and would explain the up regulation of the *sol* operon in the Cap0037::int mutant in acidogenesis and solventogenesis (Figure 6-5, Dataset S1).

Based on these experimentally validated binding sequences, we can search for putative Cap0037 binding sites across the genome using the RegPredict tool.

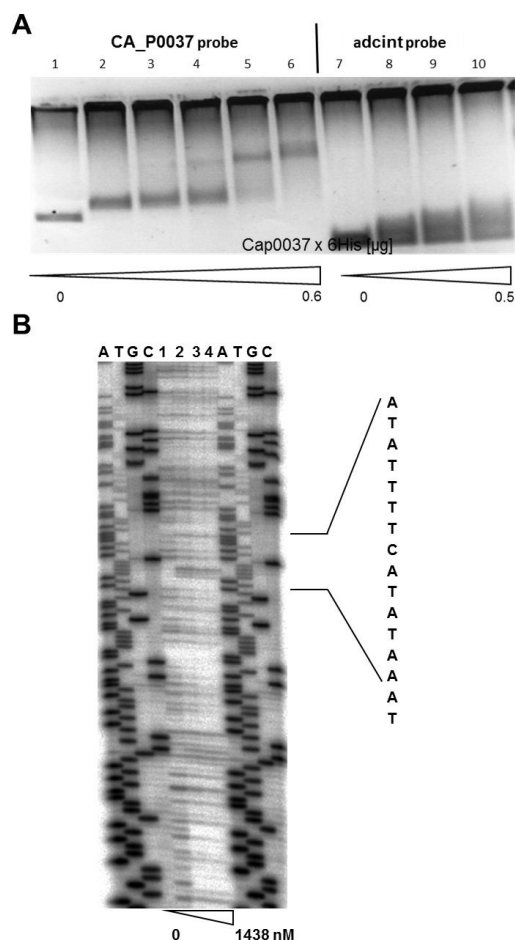


Figure 6-7 Analysis of Cap0037 binding to the promoter region of the *CA_P0037/CA_P0036* operon.

(A) EMSAs using the promoter region of the *CA_P0037/CA_P0036* operon and Cap0037 protein. Lanes 1-6: 0/0.1/0.2/0.3/0.5/0.6 μg protein; lanes 7-10: 0/0.1/0.2/0.5 μg protein. (B) DNAase I protection assay (DNA Footprinting) of Cap0037 interacting with *CA_P0037/CA_P0036* promoter region (probe 85). End-labeled DNA fragment carrying the promoter region of the *CA_P0037/CA_P0036* operon was incubated with different concentrations of Cap0037, subjected to DNAase I cleavage, and analyzed on sequencing gel. Sequencing reaction was performed with plasmid pDrive_86kurz. Line 1-4: 0-1438 nM Cap0037. The assigned region on the right side indicates the region protected by Cap0037 and the respective sequence.

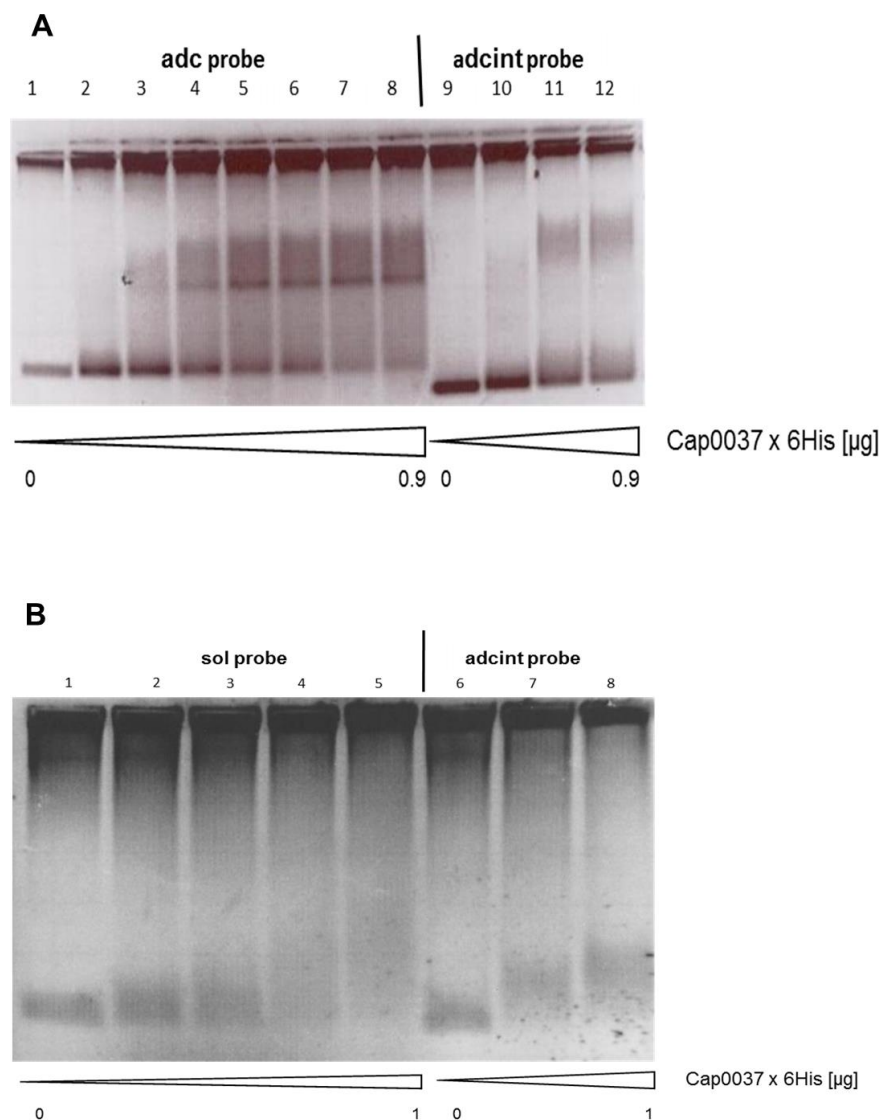


Figure 6-8 EMSAs using the promoter region of *adc*, *sol* operon and Cap0037 protein.

(A) EMSAs using the promoter region of the *adc* operon and Cap0037 protein. Lanes 1-8: 0/0.2/0.3/0.5/0.6/0.7/0.8/0.9 µg protein; lanes 9-12: 0/0.2/0.6/0.9 µg protein. (B) EMSAs using the promoter region of the *sol* operon and Cap0037 protein. Lanes 1-5: 0/0.28/0.42/0.7/1 µg protein; lanes 6-8: 0/0.42/1 µg protein.

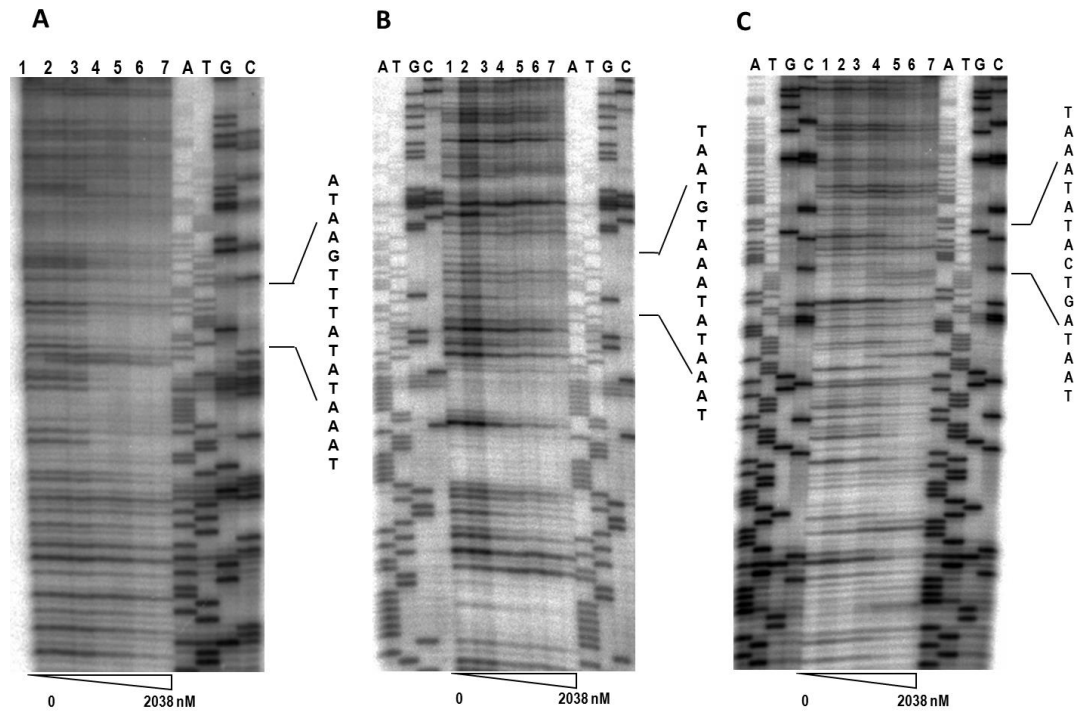


Figure 6-9 DNA footprinting of Cap0037 interacting with *adc* and *sol* promoter

(A) DNAase I protection assay I (DNA Footprinting) of Cap0037 interacting with *adc* promoter region (probe 144). End-labeled DNA fragment carrying the promoter region of the *adc* operon was incubated with different concentrations of Cap0037, subjected to DNAase I cleavage, and analyzed on sequencing gel. Sequencing reaction was performed with plasmid pDrive_144.

(B) DNA footprinting of Cap0037 interacting with *adc* promoter region (transcription start) (probe *adcStart*). End-labeled DNA fragment carrying the promoter region of the *adc* operon (transcription start) was incubated with different concentration of Cap0037, subjected to DNAase I cleavage, and analyzed on sequencing gels. Sequencing reaction was performed with plasmid pDrive_*adcStart*.

(C) DNA Footprinting of Cap0037 interacting with *sol* promoter region (probe *sol*). End-labeled DNA fragment carrying the promoter region of the *sol* operon was incubated with different concentration of Cap0037, subjected to DNAase I cleavage, and analyzed on sequencing gels. Sequencing reaction was performed with plasmid pDrive_*sol*. Line 1-7: 0-2308 nM Cap0037. The assigned region on the right side indicates the region protected by Cap0037 and the respective sequence.

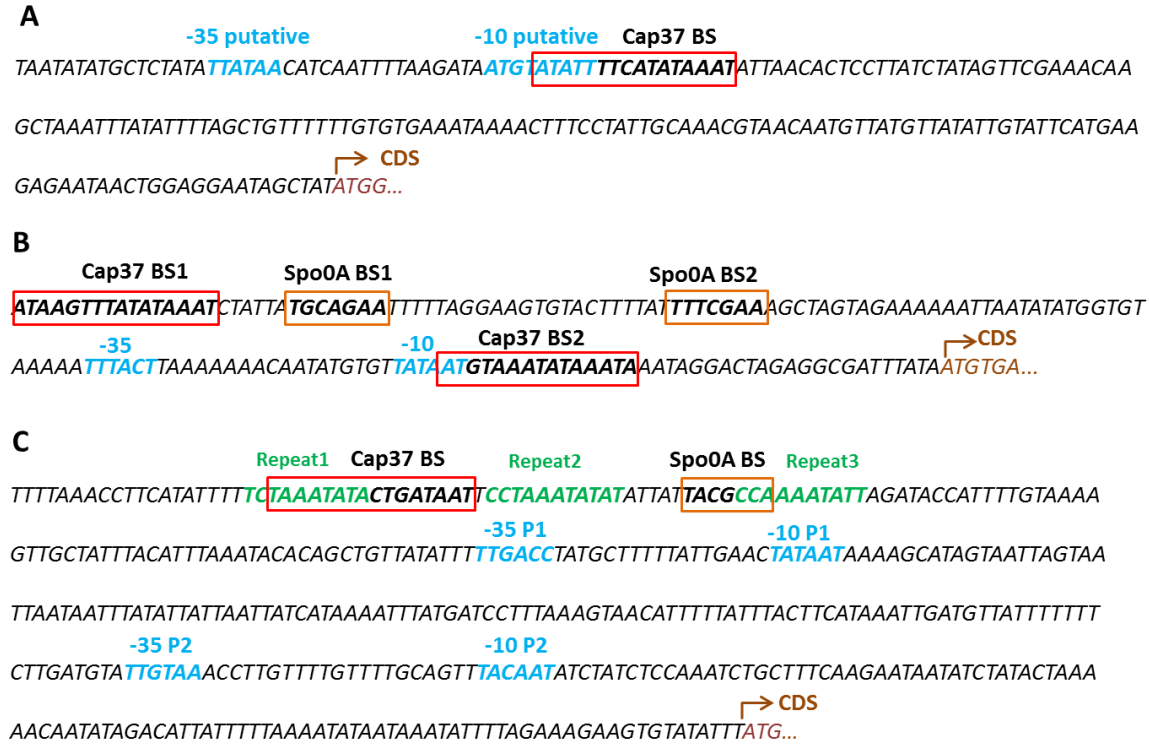


Figure 6-10 DNA sequence analysis of the promoter regions of (A), *adc*, and (B) *sol* operon.

Cap0037 binding sites (Cap0037 BS) determined in this study are shown in red boxes. Spo0A binding sites (Spo0A BS) are shown in orange boxes [121,205]. Promoter -35 and -10 elements are shown in blue letters [205,206]. Putative -35 and -10 sequences of CAP0037/CAP0036 operon were analyzed by BPROM tool. Three imperfect repeats on *sol* promoter are shown in green letters [205]. Coding sequences are in brown letters.

6.3.5 Determination of putative Cap0037 regulon

The 5'-ATATTTTCATATAAAT-3' sequence and the conserved motif (Figure 6-11) were used as references for the construction of a putative Cap0037 regulon by whole *C. acetobutylicum* genome analysis using RegPredict. The sequences that scored with values higher than 5.0 (which allow 3 mismatches or less) were selected and further analyzed for their relative position to the CDS of the corresponding gene. In addition to the three Cap0037 binding sites previously identified in the electromobility shift assays, eight other putative Cap0037 binding sites were discovered (Table 6-1). Of these eight sites, the transcription of the genes potentially regulated by Cap0037 binding was analyzed for the Cap0037::int mutant and compared to the wild-type in the three metabolic states.

The first putative Cap0037 binding site was located between the -35 and -10 promoter regions of an operon (*CA_C1029-CA_C1032*) encoding an iron transport system of the Feo family. This operon was strikingly up-regulated under all conditions in the Cap0037::int mutant. Previously, this operon was shown to be controlled by Fur (encoded by *CA_C1682*), the ferric uptake regulator, and was found to be highly induced by iron limitation or in the *fur*::int mutant [130]. There is one Fur box located downstream of the -10 promoter region of *CA_C1029-CA_C1032*. The inhibition of transcription via Fur occurs through the polymerization of the Fur-Fe²⁺ complex in the operator region, subsequently blocking RNA polymerase access to the promoter. During iron starvation, iron dissociates from the metallo-regulatory protein, resulting in dissociation of the complex and the alleviation of the transcriptional repression of genes involved in the acquisition of iron [207]. Here, we determined the putative Cap0037 binding position between the -35 and -10 promoter region of *CA_C1029-CA_C1032*. The expression of the Fur encoding gene was unchanged in the Cap0037::int mutant under all continuous cultures conditions. Thus, Cap0037 might be the second repressor of this iron transport system, raising the question of how Cap0037 and/or Cap0037 together with Fur carry out a regulatory role.

The second putative Cap0037 binding site was located 7 bp upstream of the -35 promoter region of an operon (*CA_C2777-CA_C2778*) encoding a glutaredoxin and a rubredoxin. This

operon was up-regulated under all conditions in the Cap0037::int mutant. Previously, this operon was shown to be controlled by PerR (encoded by *CA_C1682*), the peroxide repressor, and was found to be induced by oxygen exposure or in a Δ *perR* mutant [129]. There is one PerR box located downstream of the -10 promoter region of *CA_C2277-CA_C2278*. Thus, in addition to the iron transport discussed above, which could be regulated by the Cap0037 and Fur proteins, we have potentially identified another operon that can be controlled by two regulators, Cap0037 and PerR.

The third and the fourth putative Cap0037 binding sites were in the promoter region of two genes encoding potential transcriptional regulators of the MarR family for *CA_C3283* and of a CRO repressor-like DNA binding protein for *CA_C3645*. The Cap0037 binding site was located downstream of the TSP of *CA_C3283*, which is part of a three-gene operon (*CA_C3283-CA_C3281*) including two genes encoding a potential ABC transporter. All three genes were up-regulated in the Cap0037::int mutant under acidogenesis and alcohologenesis. The regulatory effects of *Cac3283* on the transcription of the two ABC transporter encoding genes and the role of this ABC transporter are still unknown. *CA_C3645* is expressed as a monocistronic operon, and the Cap0037 binding site was located downstream of the TSP. Its transcription was up-regulated under acidogenesis only in the Cap0037::int mutant.

The last four putative Cap0037 binding sites were located in the promoter region of a three-gene operon (*CA_C0984-CA_C0986*) encoding membrane proteins of unknown function, a three-gene operon (*CA_C0460-CA_C0462*) involved in butanoate metabolism, a gene *CA_C3231* encoding a potential HAD (haloacid dehalogenase-like) type of phosphatase and the *CA_C0091* gene (encoding ketol-acid reductoisomerase, involved in valine, leucine and isoleucine biosynthesis). However, the transcription levels of all of these genes were poorly affected in the Cap0037::int mutant (Table 6-1).

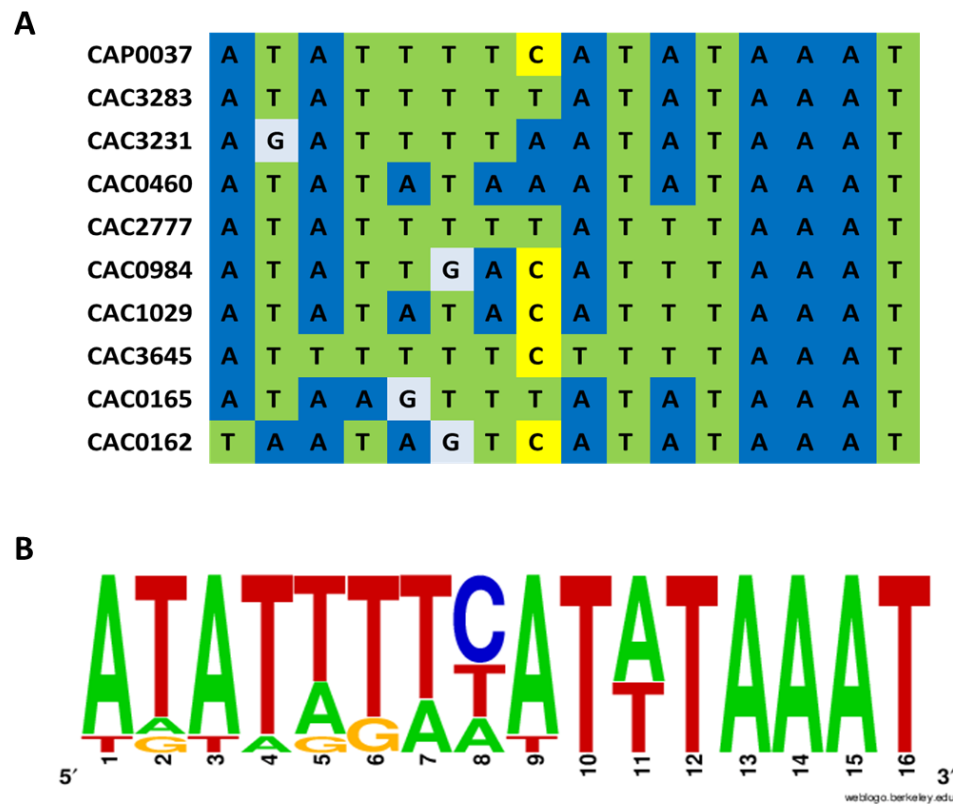


Figure 6-11 Putative Cap0037 binding sequence. (A) alignment and (B) sequence logo showing frequencies of residues

Table 6-1 Prediction of putative Cap0037 regulon and the correspondent relative transcript levels of those genes of Cap0037::int mutant in the three metabolic states: acidogenesis (AC), alcohologenesis (AL), and solventogenesis (SO). Letters in orange are mismatched nucleotides compared to the binding box *ATATTTTCATATAAAT* in Cap0037 promoter.

ORF	Putative binding sequence	Genes	Function	Fold regulation		
				AC	AL	SO
CAP0037/0036	<i>ATATTTTCATATAAAT</i>	CAP0037	Uncharacterized protein	0.962	50.716	68.558
		CAP0036	Uncharacterized protein	1.116	45.455	59.483
CAC3283/3282/3281	<i>ATATTTTATATAAAT</i>	CAC3283	Transcriptional regulator, MarR/EmrR family	2.515	2.151	1.223
		CAC3282	ABC-type multidrug/protein/lipid transport system, ATPase component	4.146	2.427	1.858
		CAC3281	ABC-type multidrug/protein/lipid transport system, ATPase component	3.030	1.759	1.054
CAC3231	<i>AGATTTTATATAAAT</i>	CAC3231	Predicted phosphatase, HAD superfamily	1.515	1.419	1.285
CAC2777/2778	<i>ATATTTTATTAAAT</i>	CAC2777	Glutaredoxin	7.201	3.447	1.623

		CAC2778	Rubredoxin	8.474	4.304	2.002
CAC0091	<i>GTTTTTTTATATAAAT</i>	CAC0091	Ketol-acid reductoisomerase	1.157	0.584	0.746
CAC0460/0461/0462	<i>ATATATAAATATAAAT</i>	CAC0460	CBS-domain containing protein, YHDP <i>B.subtilis</i> ortholog	1.758	1.066	1.280
		CAC0461	Mercuric resistance operon regulatory protein, MerR family	1.562	1.177	1.047
		CAC0462	Trans-2-enoyl-CoA reductase	1.677	1.085	1.300
CAP0165	<i>ATAAGTTTATATAAAT</i>	CAP0165	Acetoacetate decarboxylase	0.150	0.397	0.382
CAC0984/0985/0986	<i>ATATTGACATTAAAT</i>	CAC0984	ABC transporter, ATP-binding protein	1.098	0.559	2.064
		CAC0985	ABC transporter, permease component	0.802	0.452	2.207
		CAC0986	Lipoprotein, attached to the cytoplasmic membrane, NLPA family	0.976	0.714	2.681
CAC1029/1030/1031/1032	<i>ATATACATTAAAT</i>	CAC1029	FeoA-like protein, involved in iron transport	195.62	243.32	76.236
		CAC1030	FeoA-like protein, involved in iron transport	252.12	283.60	83.150

		CAC1031	FeoB-like GTPase, responsible for iron uptake	220.99	337.12	68.679
		CAC1032	Predicted transcriptional regulator	318.15	452.71	82.493
CAC3645	<i>ATTTTTCTTTTAAAT</i>	CAC3645	CRO repressor-like DNA- binding protein	2.422	0.702	0.798

6.3.6 Global transcription changes in the Cap37::int mutant

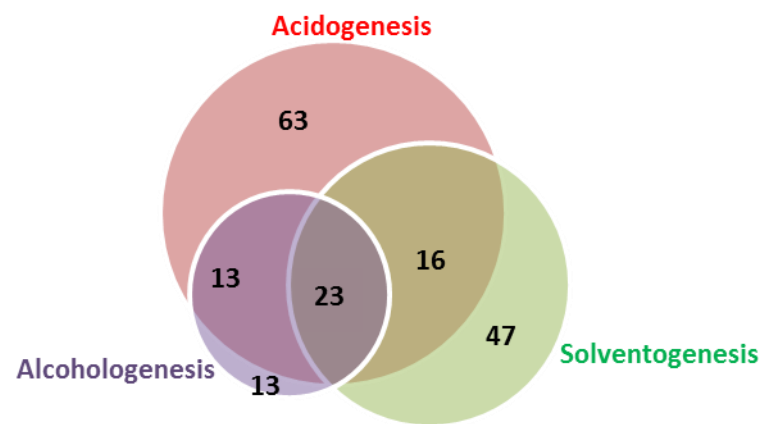
6.3.6.1 Overview

The assessment of global transcriptional changes is summarized in Figure 6-12. To filter the data for only significant results, the same criteria used to compare the wild-type strain under different physiological conditions [42] were used to compare the mutant to the control strain. The first criterion was > 4.0 -fold higher expression or > 4.0 -fold lower expression of the Cap37::int genes compared to the control strain under the same physiological conditions, and the second criterion was > 0.2 mRNA molecules per cell in at least one of the two strains being compared.

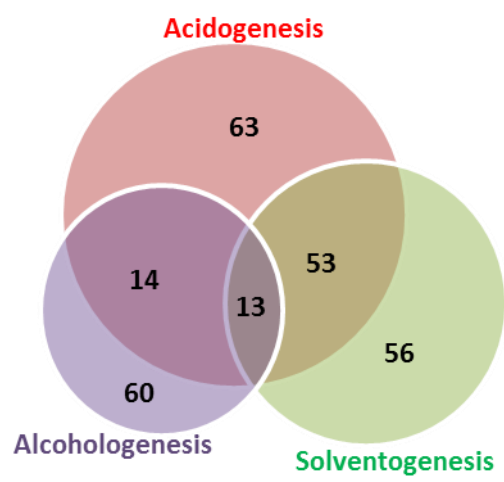
Venn diagrams (Figure 6-12A and B) show the number of significantly up-regulated and down-regulated genes in all three metabolic states of the mutant. There are 23 genes up-regulated and 13 genes down-regulated under all conditions. Noticeably, the up-regulated genes in alcohologenesis were also generally up-regulated in acidogenesis. The down-regulated genes in the three states were diverse.

Figure 6-12C and D statically categorized the up- and down-regulated genes into COGs (Cluster Orthologous Groups, data from <http://eggnogdb.embl.de> and <http://www.microbesonline.org>). There were many genes belonging to the “not determined” (N.D.) group, as orthologs could not be found in the Gram-positive bacteria database. There was also a significant number of genes belonging to the “function unknown proteins” (S) COG, especially in the down-regulated gene set. However, the graphs show the diversity of genes that were affected by Cap0037 inactivation. In the up-regulated gene set, C, E, F, H, M, P, R, S, and N.D. COGs were enriched. The most enriched groups were the amino acid transport and metabolism and inorganic ion transport and metabolism groups. The down-regulated genes were divided across the C, G, H, I, K, M, N, S, T, and N.D COGs. The most enriched group was G, which is the carbohydrate transport and metabolism group, while the S and N.D. groups were not very informative. The results will be discussed in more detail below.

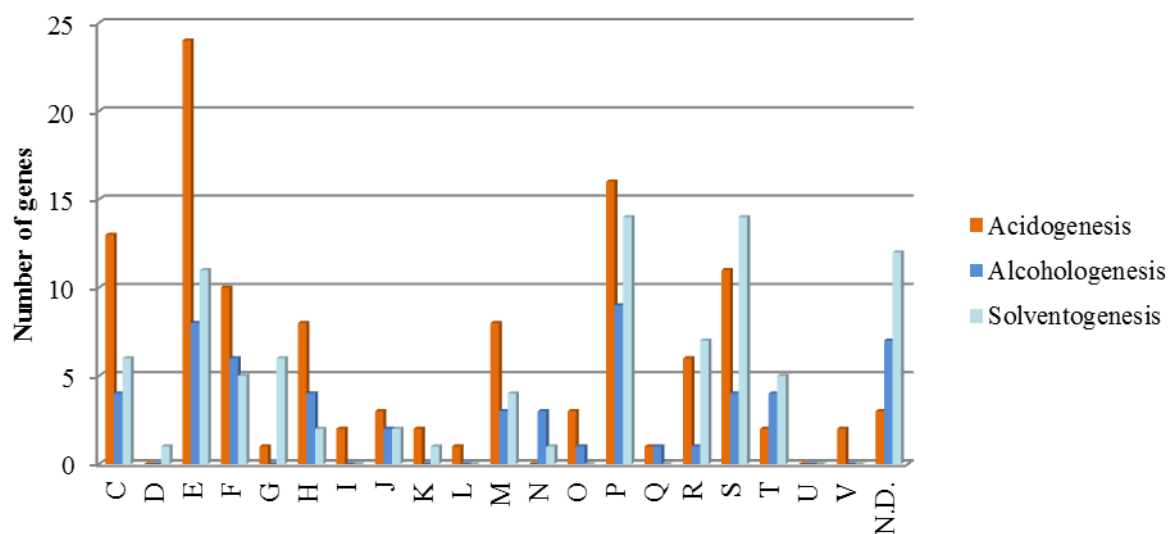
(A)



(B)



(C)



(D)

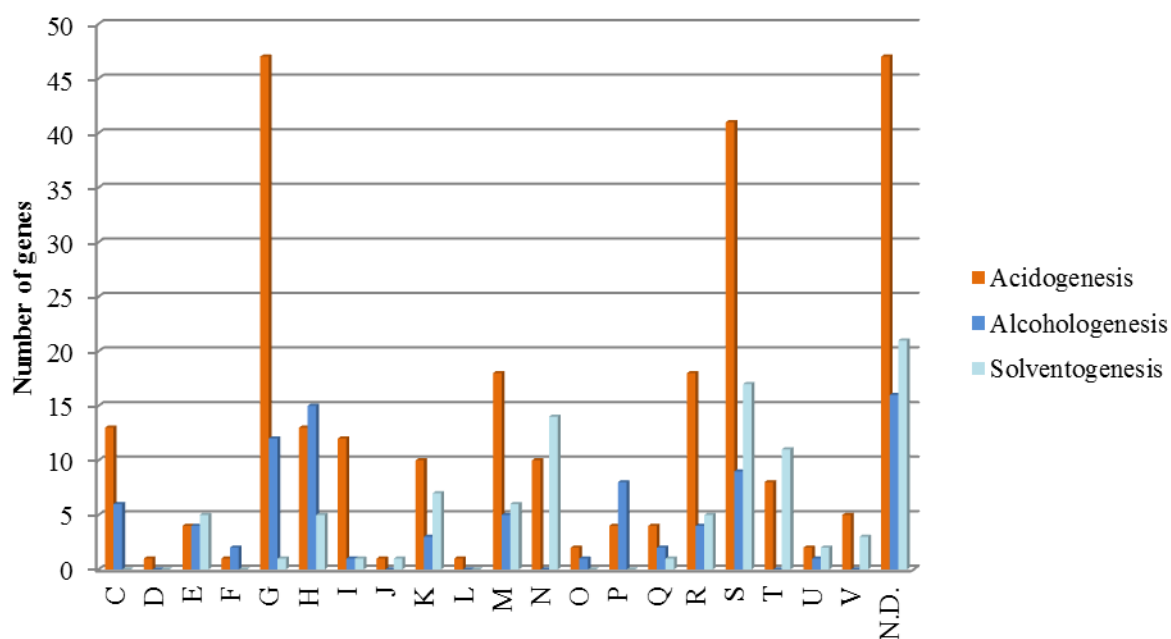


Figure 6-12 Statistical analysis of transcriptomics of *C. acetobutylicum* Cap0037::int mutant in acidogenesis, alcohologenesis and solventogenesis.

Venn's diagram of the comparison of upregulated genes (A) and of down regulated genes (B) in the three metabolic states. Distribution into COG functional group of upregulated genes (C) and down regulated genes (D) in the three metabolic states. COG C: energy production and conversion; COG D: cell division and chromosome partitioning; COG E: amino acid

transport and metabolism; COG F: nucleotide transport and metabolism; COG G: carbohydrate transport and metabolism; COG H: coenzyme metabolism; COG I: lipid metabolism; COG J: translation, ribosomal structure and biogenesis; COG K: transcription; COG L: DNA replication, recombination and repair; COG M: cell envelope biogenesis and outer membrane; COG N: cell mobility secretion; COG O: post translational modification, protein turn over, chaperones; COG P: inorganic ion transport and metabolism; COG Q: secondary metabolites biosynthesis, transport and catabolism; COG R; general function prediction only; COG S: unknown function; COG T: signal transduction mechanisms; COG U: intracellular trafficking, secretion and vesicular transport; COG V: defense mechanism; N.D. not determined.

6.3.6.2 Amino acid transport and metabolism

6.3.6.2.1 Arginine transport and metabolism

The genes involved in arginine biosynthesis, *argDBCJ* (CA_C2388-CA_C2391), *argGH* (CA_C0973-CA_C0974) and *carA* (CA_C2645), were up-regulated in acidogenesis and expressed differentially in solventogenesis and alcohologenesis. These genes were predicted to be regulated by ArgR; however, the expression of *argR* was unchanged in the mutant. In addition, there were many other genes predicted to be regulated by ArgR [193] whose expression was unchanged or only slightly changed. This information can be found in Figure 6-13A. The arginine synthesis pathway was found to be up-regulated in response to butyrate stress, especially at low concentrations, or to butanol stress at high concentrations in *C. acetobutylicum* [193]. The arginine synthesis pathway was also induced in response to oxidative stress in the fungus *Candida albicans* [208] but repressed in *C. acetobutylicum* in the oxygen-exposed wild-type strain and in a *per* inactivated mutant [129].

6.3.6.2.2 Histidine biosynthesis

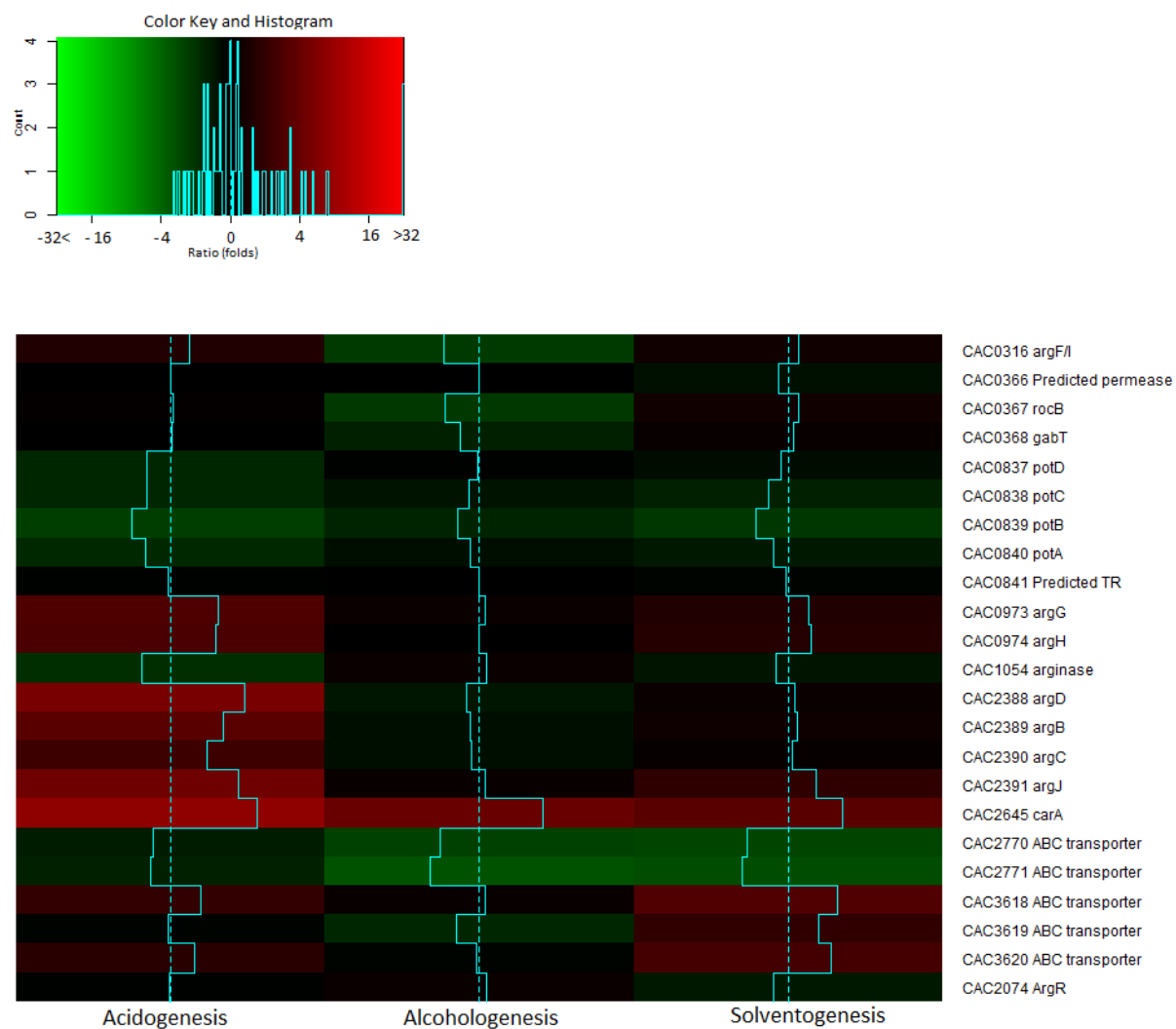
Genes involved in the histidine biosynthesis CA_C0935-CA_C0944 were up-regulated more than 2-fold in acidogenesis but only slightly affected under the other conditions (Figure 6-13B). In *S. typhimurium*, to synthesize one histidine molecule, 41 ATP equivalents are required [209]. Thus, histidine biosynthesis is associated with a high energy cost for the cells. No regulatory protein was found to control the transcription regulation of biosynthesis genes in *E. coli* and *S. typhimurium* [210]. Additionally, the expression levels of histidine

biosynthesis genes in *E. coli*, *S. typhimurium* and *Corynebacterium glutamicum* were stimulated by the two effector molecules guanosine tetraphosphate (ppGpp) and guanosine pentaphosphate (pppGpp), which accumulated when the cells responded to amino acid starvation [209–211]. However, histidine biosynthesis regulation appears to be different in *C. acetobutylicum*. It was found that CA_C2675 encodes a protein orthologous to the HisR of *Staphylococcaceae* and that a *cis*-element in the upstream region of CA_C0935- CA_C0944 was highly similar to the DNA binding motif of HisR [193]. Using RegPredict [212], a HisR regulon in *C. acetobutylicum* was proposed to contain the operon CAC0935-0944 and *hisC* (CA_C3031). Those genes also responded differently to different levels of butanol and butyrate stresses [193]. In addition, induction of the histidine biosynthesis pathway contributed to acid tolerance in *Lactococcus casei* [213].

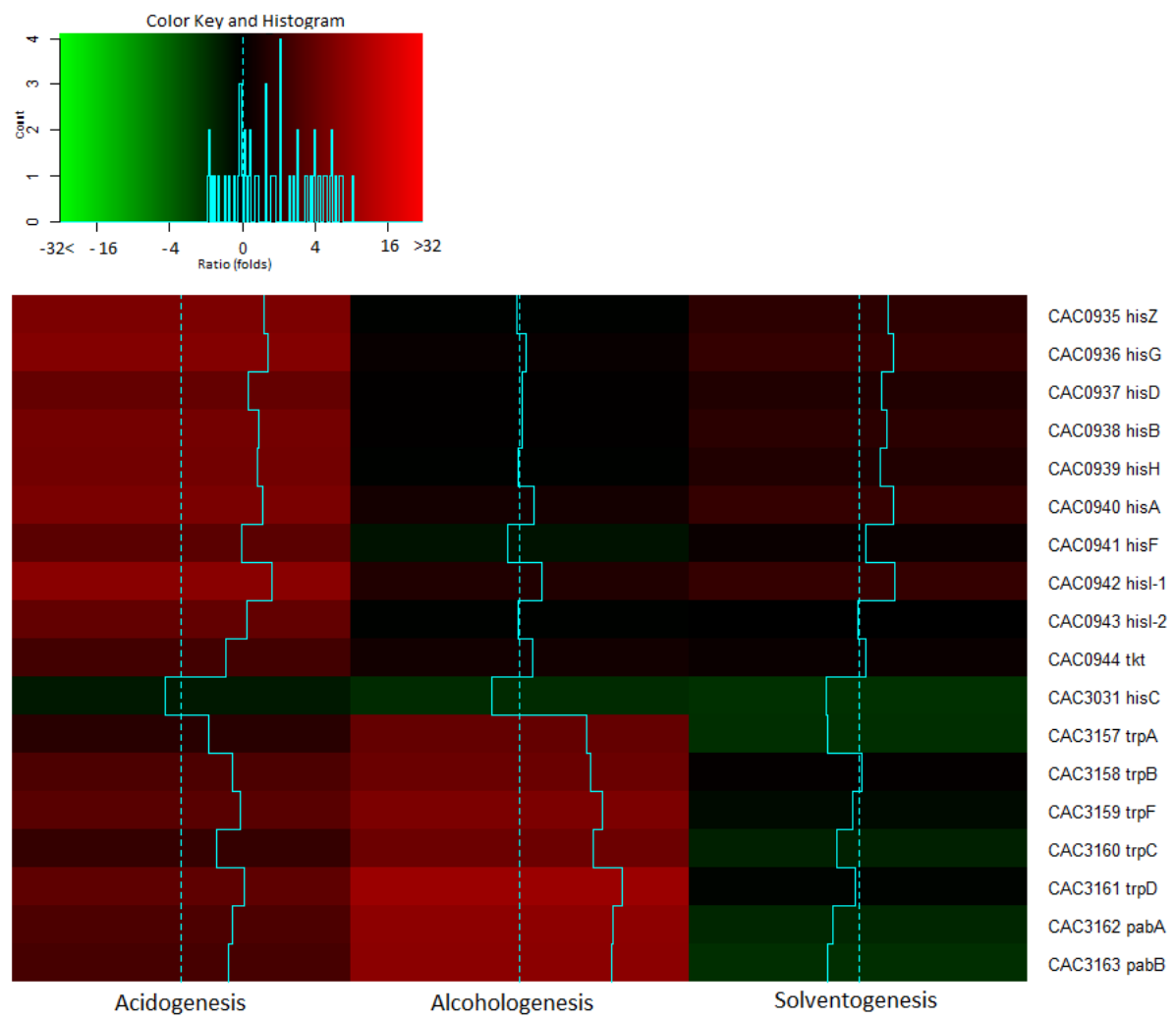
6.3.6.2.3 Tryptophan biosynthesis

The tryptophan biosynthesis operon *pabBA/trpDCFBA* (CA_C3163-CA_C3157), which encodes proteins that convert chorismate into tryptophan, was up-regulated > 2-fold in acidogenesis and alcohologenesis and was slightly down-regulated in solventogenesis. Previously, the tryptophan biosynthesis operon was reported to be down-regulated by the addition of butyrate during the continuous culture of a *ptb* mutant [111] or in batch culture under butyrate stress [193,197] (Figure 6-13B).

(A)



(B)



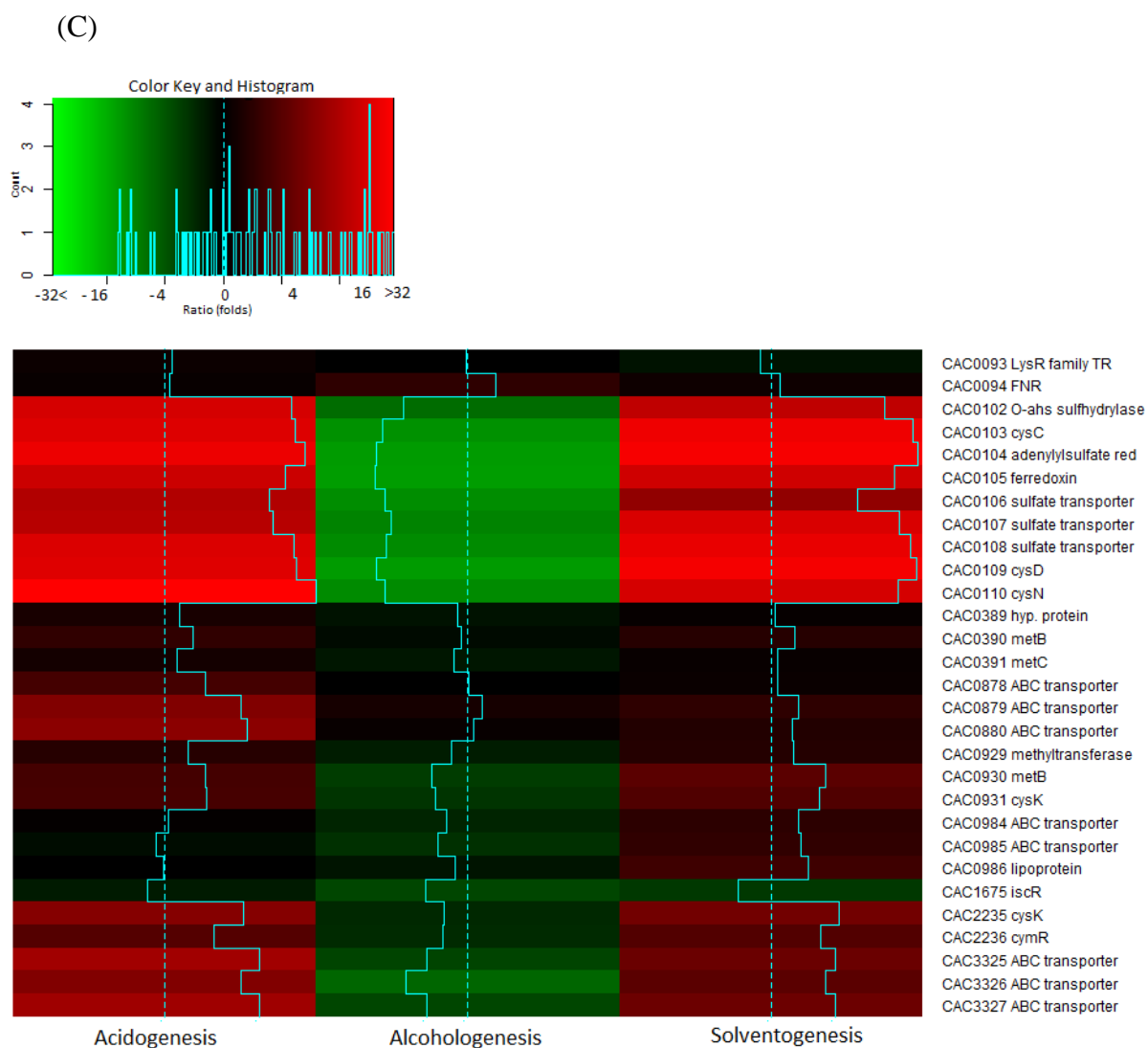


Figure 6-13 Expression profiles of genes involved in Arginine (A), Histidine and Tryptophan (B) and Cystein (C) transport and metabolism in acidogenesis, alcohologenesis and solventogenesis

6.3.6.2.4 The putative CymR and IscR regulons

Two putative CysK proteins (encoded by *CA_C0931* and *CA_C2235*) involved in L-cysteine biosynthesis from O-acetyl-L-serine (OAS) were up-regulated in acidogenesis and solventogenesis in the *Cap37::int* mutant (Figure 6-13C). In *Bacillus subtilis*, CysK plays important roles in the response to environmental changes by signaling cysteine availability to a CymR repressor via sensing the concentration of its substrate OAS [214]. CysK forms a complex with CymR and stabilizes the CymR-DNA interaction. In the presence of cysteine, the pool of OAS is low and the formation of CymR-CysK occurs and represses the transcription. In the absence of cysteine, the pool of OAS is high, and the CymR-CysK complex dissociates and de-represses the transcription [214]. In *C. acetobutylicum*, the protein encoded by *CA_C2236* had more sequence similarity to CymR than the protein encoded by *CA_C1675*, whereas the protein encoded by *CA_C1675* showed more similarity to IscR than the protein encoded by *CA_C2236* [193]. However, the IscR and CymR DNA binding motifs had a significant number of sequence similarities. Thus, more experimental data is necessary to distinguish and validate the functions of these two proteins. In addition, in the *Cap37::int* mutant, the expression of the putative CymR-encoding gene was up-regulated in acidogenesis and solventogenesis, whereas the putative IscR-encoding gene was down-regulated under all conditions. All other putative genes belonging to the CymR regulon predicted by RegPredict [193] were up-regulated, especially the *CA_C0102-CA_C0110* operon, which was significantly up-regulated in acidogenesis (up to 58-fold) and solventogenesis (up to 50-fold) and was significantly down-regulated in alcohologenesis (more than 7-fold). This *CA_C0102-CA_C0110* operon has been found to be up-regulated in response to butanol/butyrate stress (Wang et al., 2013, Alsaker et al., 2010); under alcohologenesis in wild-type strains [42] and in the Δptb strain [111]. The two putative cysteine ABC transporter operons belonging to the CymR regulon (*CA_C0878-CA_C0880* and *CA_C3325-CA_C3327*) were also up-regulated in acidogenesis. However, the expression patterns of these two operons were slightly different under solventogenesis and alcohologenesis conditions. Genes potentially belonging to the CymR regulon, such as *CA_C0985-CA_C0986*, which is involved in the D-methionine transport system, did not exhibit remarkable changes.

6.3.6.3 The up-regulation of genes involved nucleic acid transport and metabolism

Genes involved in pyrimidine biosynthesis, such as *pyrDZFIB* (CA_C2650-2654) and *carAB* (CA_C2644-CA_C2645), were strikingly up-regulated under all three conditions (between 17- and 100-fold) Figure 6-14. Those operons are shown to be up-regulated after n-butanol pulse in wild-type *C. acetobutylicum* [194] but down-regulated under acetate and butyrate stresses[197]. The putative uracil permease-encoding gene, *pyrP* (CA_C2112), was slightly up-regulated under all conditions.

Genes involved in purine biosynthesis, *purECFMNHD* (CA_C1390-1396) and *purQ/J* (CA_C1655), were up-regulated more than 2-fold in acidogenesis but insignificantly down-regulated in solventogenesis and alcohologenesis. Purine metabolism was reported to be strongly up-regulated in response to high butyrate stress levels, suggesting its involvement in acid resistance [193] Those genes were up-regulated under low butanol stress conditions but down-regulated under high butanol stress conditions [193,194]. In *C. acetobutylicum* cells exposed to oxygen, this purine biosynthesis pathway was also up-regulated in order to repair damaged DNA [129].

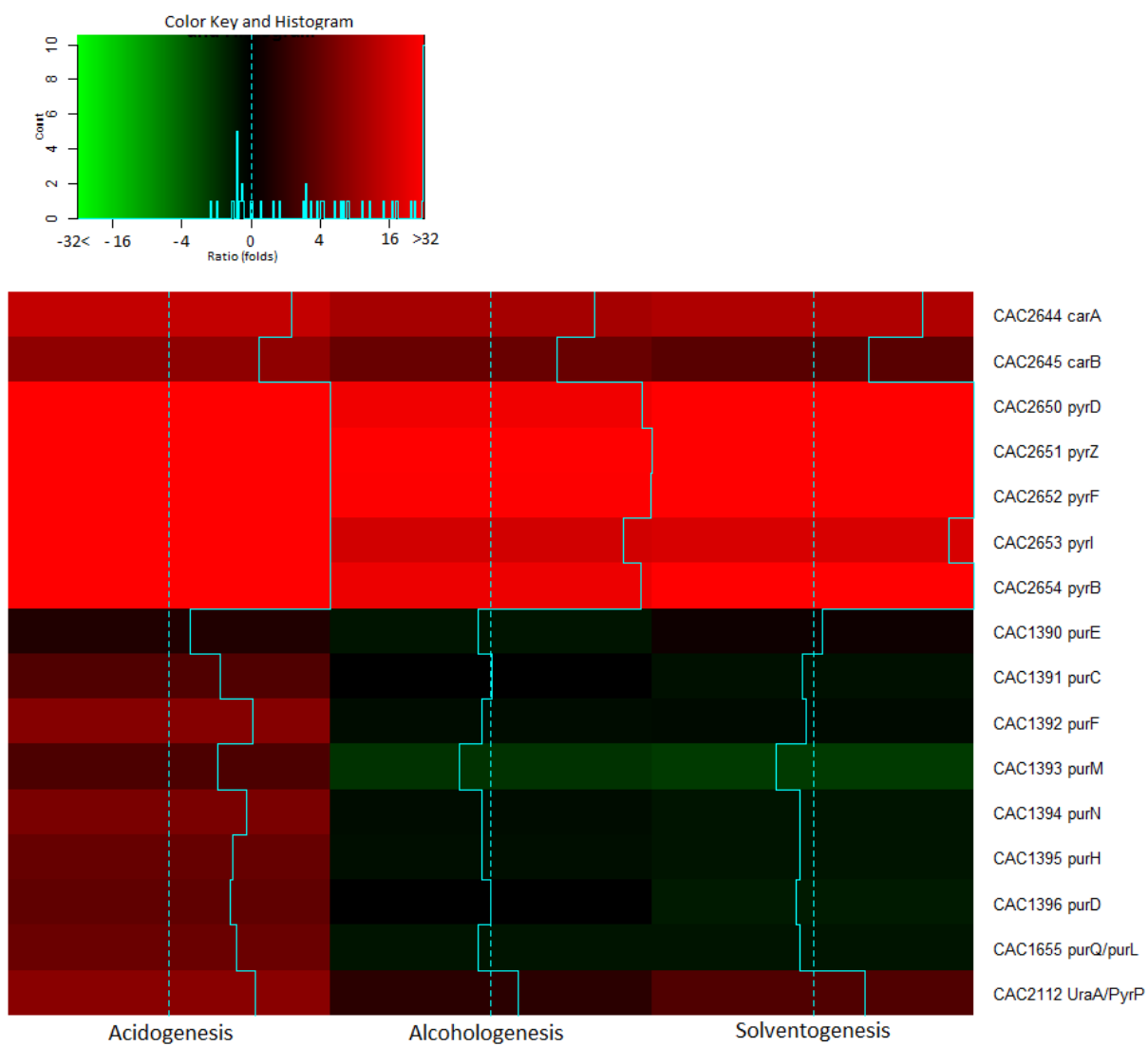


Figure 6-14 Expression profiles of genes involved in nucleic acid metabolism in acidogenesis, alcohologenesis and solventogenesis

6.3.6.4 The expression of cation transporter-encoding genes

Three potassium transport-related genes, *kdpX* (CA_C3679), *kdpC* (CA_C3681) and *kdpD* (CA_C3680), were significantly up-regulated in acidogenesis whereas the sodium efflux pump encoding the operon consisting of *natA* (CA_C3551) and *natB* (CA_C3550) was down-regulated (Figure 6-15). Those cation transporters were previously found to be up-regulated in response to acetate and butyrate stresses [197].

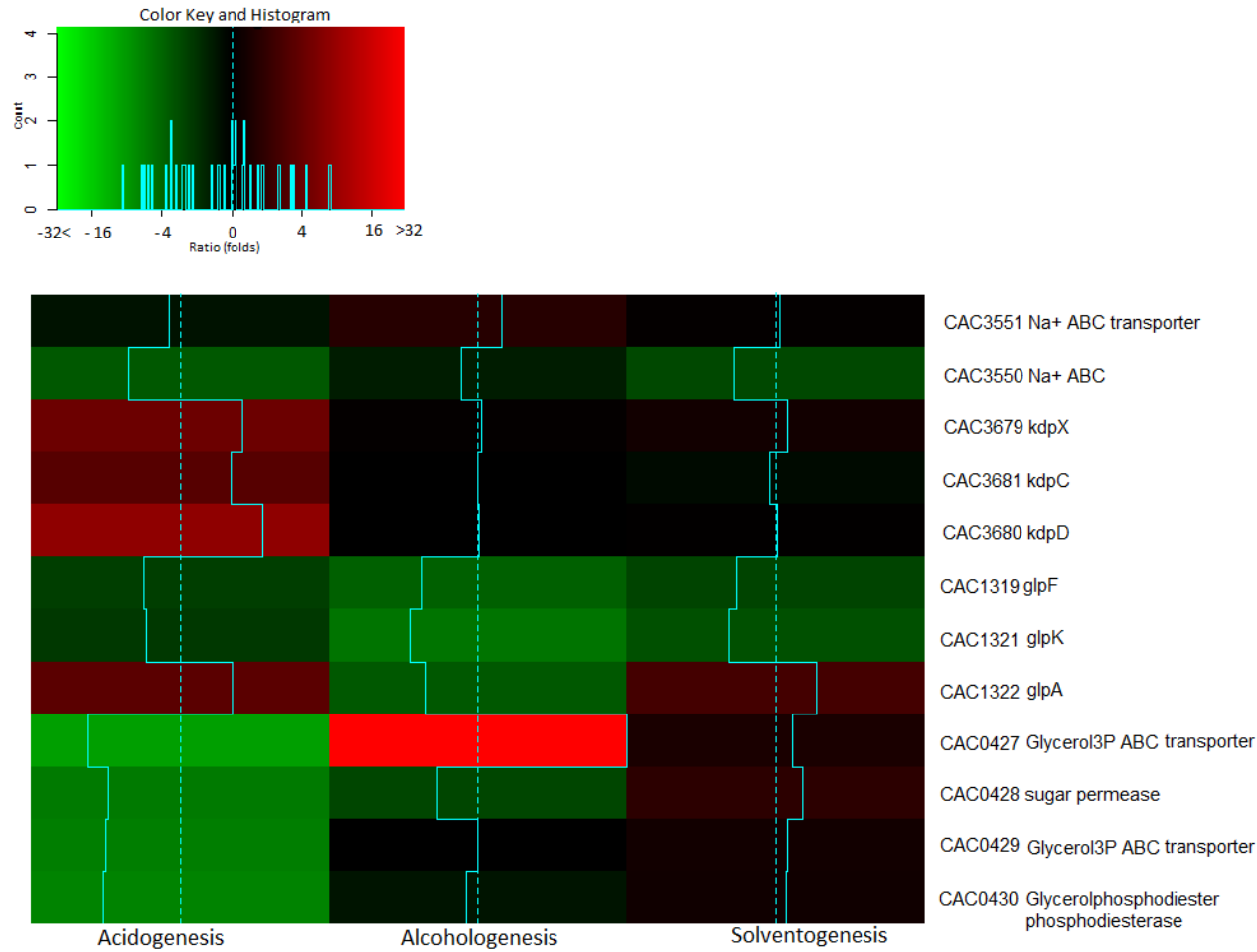


Figure 6-15 Expression profiles of genes involved in cation transport, glycerol lipid transport and metabolism in acidogenesis, alcohologenesis and solventogenesis

6.3.6.5 Glycerol lipid metabolism

The Cap37::int mutant was only able to consume low amounts of glycerol in alcohologenesis. The transcriptomic data reveals significantly down-regulated genes involved in glycerol uptake and the oxidation pathway (CA_C1319-CA_C1322). Glycerol-3-P ABC transporter-encoding genes (CA_C0427-CA_C0430) were significantly down-regulated in acidogenesis and alcohologenesis but remained unchanged in solventogenesis (Figure 6-15).

6.3.6.6 Motility and chemotaxis

The cell motility genes encode products responsible for the chemotactic response and flagella assembly. Table S- shows the expression changes in cell motility related genes in the Cap37::int mutant. It was reported that in Clostridia and Bacilli, those genes were down-regulated in cells undergoing sporulation processes [215]. Response regulators *CheA*, *CheB*, *CheC*, *CheD*, *CheR*, *CheV*, *CheY*, *CheW*, and *CheX*, the methyl-accepting chemotaxis proteins and the genes involved in flagella assembly, including *FlgGEDBCKM* and *FliDEFGHJKRSPQYZ*, were down-regulated in the Cap37::int mutant under acidogenesis and solventogenesis; however, they were differentially regulated under alcohologenesis. These motility and chemotaxis systems were also down-regulated in the *buk*-knockout mutant and in the *spo0A*-over-expressing strain [195,216].

6.3.6.7 Sporulation

In *Bacillus subtilis*, AbrB inhibits sporulation by repressing *sigH*, subsequently suppressing the expression of *spo0A* [94]. In *C. acetobutylicum*, the role of AbrB is unclear. There are many *abrB* annotated genes in the *C. acetobutylicum* genome. The *abrB310* (CA_C0310) was the only gene found to be actively transcribed based on the chloramphenicol acetyltransferase reporter gene analysis. When this gene is knocked down by antisense RNA, solventogenesis and differentiation are delayed [127]. In our Cap37::int mutant, *abrB310* was significantly down-regulated under acidogenesis and alcohologenesis conditions. However, the other *abrB* (CA_C3647) gene was significantly up-regulated under these conditions.

The expression of *spo0A* was not affected in the Cap37::int mutant. To activate Spo0A via phosphorylation, histidine kinases were employed. Many histidine kinases have been found in the *C. acetobutylicum* genome, including Cac0323, Cac0903, Cac3319 and Cac0437 (with exceptional dephosphorylation activity) [119]. In the wild-type strain, CA_C3319

exhibits the highest expression among the four. In the acidogenic Cap37::int mutant, the *CA_C3319* histidine kinase-encoding gene (as well as the other three) and *sigF* were significantly down-regulated. In the solventogenic Cap37::int mutant, the phosphorylation of Spo0A was not affected as the histidine kinase-encoding gene expression remained unchanged. In addition, we observed a significant up-regulation of all of the sporulation sigma factors: *sigE*, *sigF*, *sigG*, *sigK* and *spoIIAA* (Figure 6-16). The high expression of *sigF* along with the high expression of *sigE* and *sigG* was in agreement with previous findings that SigF is necessary for the expression of other sigma factors [119].

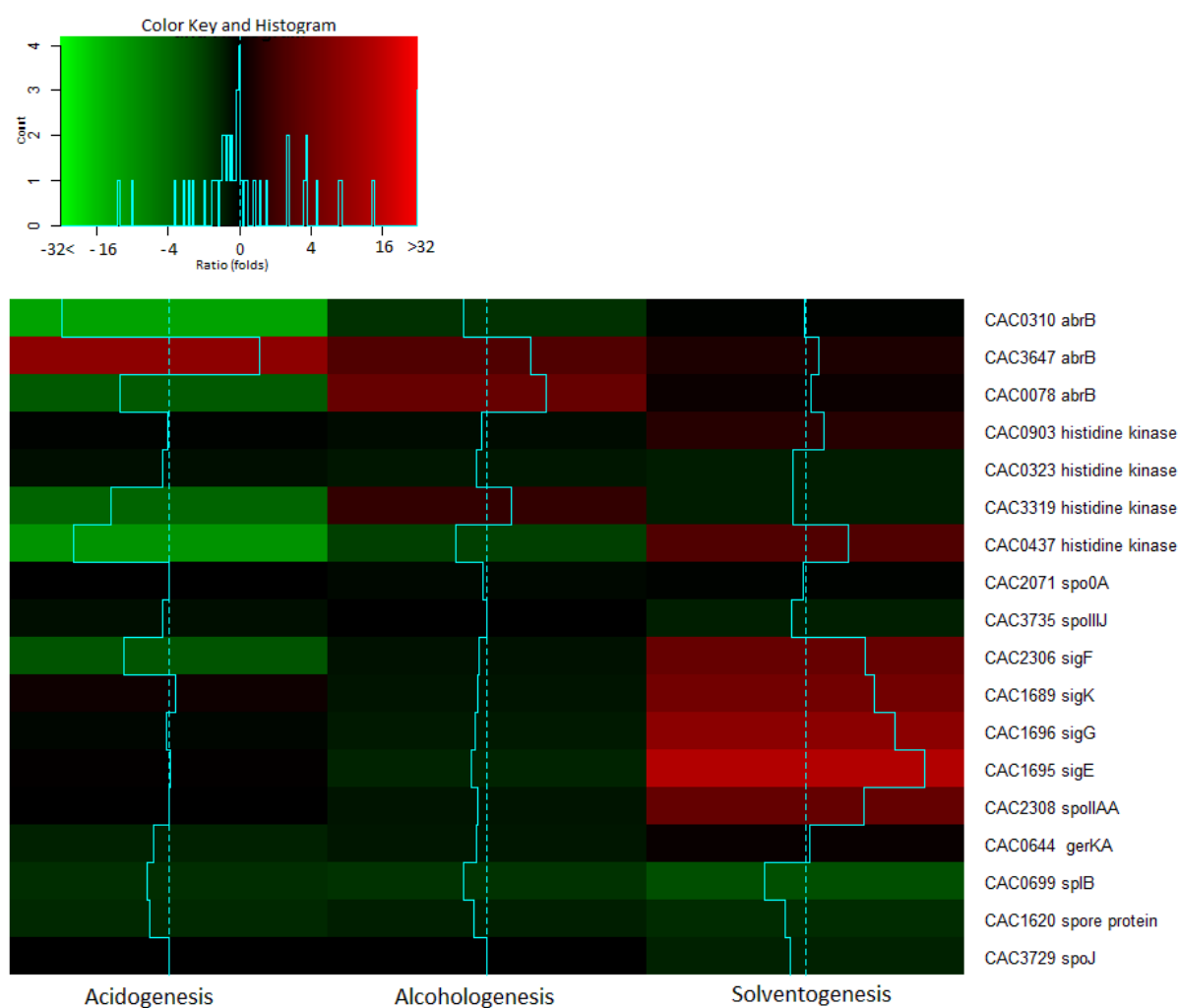


Figure 6-16 Expression profiles of genes involved in sporulation in acidogenesis, alcohologenesis and solventogenesis state

6.3.6.8 CA_P0037 inactivation affects the expression of the Rex, Fur and PerR regulons

6.3.6.8.1 In the Cap37::int mutant the Rex regulon transcription is affected.

The redox sensing transcriptional repressor (Rex) has been found to modulate its DNA-binding activity in response to the NADH/NAD⁺ ratio and plays a role in the solventogenic shift of *C. acetobutylicum*. Rex binding sites were determined in the promoters of different operons: *adhE2*, *ldhA*, *crt*, *thlA*, *asrT*, *ptb*, and *nadA* [110,134]. Those operons were up-regulated in the Rex mutant. The gene *ldhA* was the most induced gene followed by *adhE2* [134]. In the Cap37::int mutant, Rex expression was only 1.4-fold higher in acidogenesis and was unchanged in alcohologenesis and solventogenesis. However, except for the *thlA* gene, all of the other genes of the Rex regulon, such as *adhE2*, *ldhA*, *crt-bcd-etfAB-hbd*, *nadABC*, and *ptb-buk*, were up-regulated to varying degrees (Table 6-2). As there are no putative Cap0037 binding sites in the promoter region of all the genes or operons shown to be controlled by Rex, it is tempting to speculate that the effect of Cap0037 inactivation might be indirect and occur through an increase in the NADH/NAD⁺ ratio. The fact that *thlA* was unchanged or slightly down-regulated in the Cap37::int mutant but highly up-regulated in the Rex mutant suggests a more complex regulatory mechanism of *thlA* expression.

Table 6-2 Expression profiles of Rex regulon in the three metabolic states: acidogenesis (AC), alcohologenesis (AL) and solventogenesis (SO)

ORF	Function	Relative expression MT/WT (fold)			
		AC	AL	SO	RexMT/WT*
CAC0267	L-lactate dehydrogenase	83.37	31.58	3.54	13.75
CAC1023	Nicotinate-nucleotide pyrophosphorylase	2.62	0.84	0.34	1.79
CAC1024	Aspartate oxidase	2.25	0.86	0.63	1.79
CAC1025	Quinolate synthase	2.33	0.93	0.54	2.53
CAC1512	Formate/nitrite family of transporter	n.d.	n.d.	n.d.	2.04
CAC1513	Anaerobic sulfite reductase (Fe-S subunit)	n.d.	n.d.	n.d.	2.41
CAC1514	Anaerobic sulfite reductase, B subunit	0.61	0.72	0.77	2.67
CAC1515	Anaerobic sulfite reduction protein C, reductase	n.d.	n.d.	n.d.	1.69
CAC2708	β -hydroxybutyryl-CoA dehydrogenase	1.86	0.32	0.53	2.79
CAC2709	Electron transfer flavoprotein alpha-subunit	2.60	0.51	0.75	1.55
CAC2710	Electron transfer flavoprotein beta-subunit	2.02	0.38	0.68	2.17
CAC2711	Butyryl-CoA dehydrogenase	1.46	0.96	0.83	2.81
CAC2712	Crotonase (3-hydroxybutyryl-COA dehydratase)	1.85	0.94	1.04	2.43
CAC2873	Acetyl-CoA acetyltransferase	0.98	0.25	0.43	12.46
CAC3075	Butyrate kinase, BUK	0.89	0.65	0.71	1.97
CAC3076	Phosphate butyryltransferase	1.41	0.79	0.86	6.07
CAP0035	Aldehyde-alcohol dehydrogenase, ADHE2	49.81	0.11	4.47	164.75
CAC2713	AT-rich DNA-binding protein	1.44	0.79	0.92	n.a.

* Data from Zhang et al. 2014, n.a.: not available; n.d.: not detected

6.3.6.8.2 In the Cap37::int mutant the PerR regulon transcription is affected.

In a *perR* mutant or after the wild-type strain is exposed to oxygen, the genes encoding for proteins involved in oxygen and ROS detoxification, including the reverse rubrerythins (*rbr3A-rbr3B*), desulfoferrodoxin (*dfx*), glutaredoxin (*CA_C2777*), rubredoxin (*rd*, *CA_C2778*), NADH-dependent rubredoxin oxidoreductase (*nror*), the oxygen-reducing flavoproteins (*fprA1* and *fprA2*) and a flavodoxin (*CA_C2452*), were up-regulated [129]. The Cap0037::int mutant behaved quite similarly to the *perR* mutant under acidogenic and alcohologenic conditions (Table 6-3). As shown above, a putative Cap0037 binding site was located 7 bp upstream of the -35 promoter region of *CA_C2777-CA_C2778* encoding a glutaredoxin and a rubredoxin, respectively. However, no putative Cap0037 binding sites were identified in the promoter region of all of the other genes or operons shown to be controlled by PerR, and it is probable that the effect of Cap0037 inactivation is indirect.

6.3.6.8.3 In the Cap37::int mutant the Fur regulon transcription is affected.

Interestingly, we also found a common regulatory effect of Cap0037 with the *fur* mutant. Table 6-3 compares the Cap0037::int mutant with the *fur* mutant and the iron-limiting wild-type. One of the most critical similarities is the over-expression of genes involved in iron transport and metabolism. As discussed above, the ferrous transport system FeoA (encoded by *CA_C1029-CA_C1032*), which contains a putative Cap0037 box and a Fur box in the promoter, was strikingly over-expressed in the Cap37::int mutant, the Fur mutant and in the iron-limited WT. The second putative ferrous iron-uptake system FeoB (encoded by *CA_C0447-CA_C0448*) for Fe^{2+} and the ferrichrome-dependent system (encoded by *CA_C0788-CA_C0791*) were also highly up-regulated in all cases, although less significantly in the iron-limited wild-type strain. A flavodoxin (*fld1*)-encoding gene (*CA_C0587*) was strikingly up-regulated in the iron-limited WT, a *fur* mutant and the Cap37::int mutant (Table 6-3). A Fur box was found in the promoter of *fld1* (*CAC0587*), but no putative Cap0037 box was identified. It has been well-documented that ferredoxin is reduced dramatically in Clostridia under low iron conditions, *fld* is induced and flavodoxin replaces ferredoxin as a single electron carrier. However, in the Cap37::int mutant, the expression level of *fdx* (*CA_C0303*), encoding the main ferredoxin, was not reduced. In iron deficiency, or in a *fur* mutant or Cap0037 inactivation, the levels of genes belonging to the *ribGBAH* operon, which is involved in riboflavin synthesis, are strongly induced. Riboflavin biosynthesis genes were

previously shown to be up-regulated only by butyrate stress and not by acetate or butanol stress [197]. Riboflavin is a precursor of the FMN and FAD coenzymes. The expression of the *rib* operon was shown to be regulated by a FMN-sensing riboswitch element [217] in *B. subtilis*, but the regulatory mechanism in *C. acetobutylicum* has yet to be characterized.

To summarize, the Cap37::int mutant cells behave as if they are Fe²⁺-limited, although they exist under iron excess. However, the up-regulated Fur regulon in the mutant is an indirect regulatory effect rather than reflecting a real shortage in the intracellular Fe²⁺ concentration, and except for the FeoA-encoding operon (*CA_C1029-1032*), no Cap0037 box was identified in the promoter regions of the up-regulated genes.

Table 6-3 Transcriptional profiles of selected genes in Cap0037::int mutant in the three metabolic states compared with those of oxygen exposed WT, ΔPerR mutant, iron limited WT and Fur::int mutant

ORF	Gene	Protein function	Cap0037 ::int vs WT				ΔPerR vs WT ^a			Fur ::int vs WT ^b		
			AC	AL	SO	Cap37 box	+O ₂ /-O ₂	ΔPerR	Per box	-Fe/+Fe	Fur :int	Fur box
CAC1027	<i>fprA1</i>	Flavoprotein	0.97	1.06	1.01	–	4.4	16.1	+	n.a.	n.a.	–
CAC2449	<i>fprA2</i>	Flavoprotein	7.01	3.28	1.21	–	7.3	22.5	+	n.a.	n.a.	–
CAC2448	<i>nror</i>	NROR	6.30	2.54	1.12	–	7.8	16.8	+	n.a.	n.a.	–
CAC2450	<i>dfx</i>	desulfoferrodoxin, superoxide reductase	2.39	1.11	0.39	–	18.4	29.6	+	n.a.	n.a.	–
CAC2777		Glutaredoxin	7.20	3.45	1.62	+	3.5	6.0	+	n.a.	n.a.	–
CAC2778	<i>rd</i>	Rubredoxin	8.47	4.30	2.00	+	5.2	14.4	+	n.a.	n.a.	–
CAC3306	<i>tpx</i>	Thiolperoxidase	2.16	0.77	0.63	–	3.0	3.4	+	n.a.	n.a.	–
CAC3597	<i>rbr3B</i>	Reverse rubrerythrin	1.69	0.78	1.45	–	5.6	26.9	+	n.a.	n.a.	–
CAC0447	<i>feoB</i>	Feo protein, involved in Fe ²⁺ transport	18.67	70.71	5.65	–	2.3	1.22	–	n.a.	n.a.	–
CAC0448	<i>feoB</i>	Fe ²⁺ transport protein B	6.58	17.92	2.71	–	2.8	1.34	–	3.76	n.a.	–

CAC0788		Ferrichrome transport permease	76.51	32.37	33.6	–	3.3	n.a.	–	207	33.6	+
CAC0789	<i>fhuB</i>	Permease	125.0	50.73	59.5	–	3.4	n.a.	–	199	37.4	+
CAC0790	<i>fhuD</i>	Ferrichrome-binding periplasmic protein	195.4	128.3	50.7	–	2.6	n.a.	–	86.7	54.4	+
CAC0791	<i>fhuC</i>	Ferrichrome ABC transporter, ATP-binding protein	∞	90.69	∞	–	2.3	n.a.	–	122.2	67.0	+
CAC1029	<i>feoA</i>	FeoA-like protein, involved in iron transport	195.6	243.3	76.2	+	5.8	0.99	–	100.5	140.5	+
CAC1030	<i>feoA</i>	FeoA-like protein, involved in iron transport	252.1	283.6	83.1	+	8.1	1.06	–	113.8	154.7	+
CAC1031		FeoB-like GTPase, responsible for iron uptake	220.9	337.1	68.6	+	15.9	1.13	–	246.7	235.9	+
CAC1032		Predicted transcriptional regulator	318.1	452.7	82.4	+	4.3	0.94	–	218.1	159.9	+
CAC0587	<i>fldI</i>	Flavodoxin	461.9	1197	119	–	11.4	n.a.	–	275.7	312.4	+
CAC2452	<i>fld</i>	Flavodoxin	6.15	3.90	1.58	–	3.6	11.6	+	n.a.	n.a.	–
CAC0590	<i>ribD</i>	Pyrimidine deaminase and	8.13	1.78	3.36	–	2.08	1.22	–	65.4	78.9	+

		pyrimidine reductase										
CAC0591	<i>ribB</i>	Riboflavin synthase subunit α	9.07	1.31	3.68	–	1.72	1.1	–	55.3	57.2	+
CAC0592	<i>ribA</i>	GTP cyclohydrolase/3,4-dihydroxy-2-butanone 4-phosphate synthase	10.83	1.30	3.56	–	1.5	1	–	44.2	61.9	+
CAC0593	<i>ribH</i>	6,7-dimethyl-8-ribityllumazine synthase	5.88	1.02	1.98	–	1.71	1.05	–	41.0	49.0	+

a, b Data from Hillmann et al. 2009 and Vasileva et al. 2012, respectively; n.a.: not available

6.4 Conclusion

Collectively, these results demonstrate that in the strict anaerobe *C. acetobutylicum*, Cap0037 directly or indirectly controls a large set of genes belonging to the Rex, PerR and Fur regulons. This control drastically affects metabolism under both acidogenic and alcohologenic conditions. More experiments will be needed to understand the role of Cap0037 in the regulatory network of *C. acetobutylicum*.

6.5 Experimental Procedure

6.5.1 Bacterial strains, plasmids and culture media

The bacterial strains, plasmids and primers used in this study are listed in Table S- . *C. acetobutylicum* $\Delta cac1502\Delta upp$ and its derivative mutants were grown in Clostridial growth medium (CGM) for routine manipulation and in synthetic medium (SM) for sporulation and the production of solvents. A Reinforced Clostridial Agar (RCA) plate with antibiotic at a pH of 5.8 was used after electroporation and a pH of 6.5 was used for the replication of transformants. The *E. coli* strain was grown in Luria-Bertani (LB) medium at 37°C. Antibiotics were used at the following concentrations: erythromycin and clarithromycin at 40 µg/ml and thiamphenicol at 10 µg/ml for *C. acetobutylicum* and ampicillin at 50 µg/ml for *E. coli*.

6.5.2 Plasmid construction

Plasmid pMTL007C-E2::Cap0037 (277a) was generated from pMTL007C-E2 [84]. The retargeting region was designed via a Perutka algorithm and was constructed using the relative primers in Table S- . The pMTL007C-E2::Cap0037 (277a) plasmid was then transformed into *Clostridium acetobutylicum* $\Delta cac1502\Delta upp$ by electroporation without prior methylation [167], and the transformants were selected on an RCA plate at a pH of 6.5 containing thiamphenicol at 40 µg/ml. After that, integrants were selected on an RCA plate at a pH of 6.5 containing erythromycin at 40 µg/ml and were analyzed by PCR using the CAP0037_F and CAP0037_R primers that flank the 277/278 position. We were unable to detect the intron inserted at the target position in the integrants. As the integrants were Erm resistant, the intron should have been integrated elsewhere in the genome. Another construction used was pCUI-Cap0037 (189s). This plasmid is derived from the pCUI plasmid, which was developed by our lab [182]. The intron in this vector is smaller than in the pMTL007C-E2 system because it does not contain a marker gene. Therefore, in order to

select for mutants, screening by PCR with 2 external primers must be performed. In the plasmid pCUI-Cap0037 (189s), the retargeting region was designed via the Perutka algorithm for the position 189/190 sense strand. In the presence of the *ltrA* gene, the integrants could lose the intron due to mRNA maturation. As a result, after screening and selecting for insertion, the plasmid pCUI-Cap0037 (189s) in the mutant was cured by spreading the cells on RCA at a pH of 6.5 containing 5'fluorouracil at 1 mM [167]. The colonies were then checked for plasmid loss by testing their sensitivity to Tm, Erm and PCR using LtrA 5D and LtrA 3R primers that are specific for the vector backbone. The mutant Cap0037::int was successfully generated by this second strategy.

For the complementation experiment, *CAP0037* CDS and its ribosomal binding site (RBS) were amplified by Cap0037-F-RBS *Bam*HI and Cap0037-R-*Sfo*I and cloned into a pSOS95 shuttle vector [154] by two restriction sites, *Bam*HI and *Sfo*I, to obtain the plasmid pSOS95-Cap0037. This pSOS95 vector allows the expression of *CAP0037* under the constitutive *thlA* promoter. The plasmid pSOS95-Cap0037 was then electroporated into mutant Cap0037::int and selected on an RCA plate with Erm at 40 µg/ml. The complementary strain carrying the plasmid was maintained in liquid MS culture under pressure from clarithromycin at 40 µg/ml.

Plasmid pBS2 was used for heterologous production of Cap0037. For construction of plasmid pBS2, the 630 bp *CA_P0037* open reading frame was amplified by PCR using primers ygaS1hinCI and yghSrückCI with *C. acetobutylicum* WT genomic DNA as template. The PCR product was digested with *Nde*I and *Hind*III and ligated into respectively digested vector pET29a(+). Vector pET29a(+) contains a His-tag sequence that is added to the C-terminus of the produced protein.

Plasmids pDrive_86kurz, pDrive_144, pDrive_*adc*Start, and pDrive_*sol* were used for sequencing reaction of DNaseI protection assay (DNA Footprinting). For construction of pDrive_86kurz, pDrive_144, pDrive_*adc*Start, and pDrive_*sol*, the probes 85 (256 bp), 144 (249 bp), *adc*Start (263 bp), and *sol* (245 bp) were amplified by PCR using genomic DNA from *C. acetobutylicum* WT as template and primers 86 fwd and 85 rev, 144 rev and 145 fwd, *adc*start FP EcoRI fwd and *adc*start FP EcoRV rev, and *sol* FP EcoRI fwd and *sol* FP EcoRV rev, respectively. PCR products were ligated into pDrive cloning vector. The generated PCR probes 85 (256 bp), 144 (249 bp), *adc*Start (263 bp), and *sol* (245 bp) which contain parts of

respective promoter regions were used as probes for DNaseI protection assay (DNA Footprinting).

6.5.3 Continuous culture

Continuous fermentation was performed as previously described [42]. The cultures were maintained under acidogenesis, solventogenesis and alcohologenesis conditions, fed with 995 mM total carbon, and maintained at a dilution rate of 0.05 h^{-1} . Samples for the production profile analysis and mRNA extraction were prepared under steady states. Results from the mutant *Clostridium acetobutylicum* $\Delta\text{cac1502}\Delta\text{uppCap0037::int}$ were compared with the control strain *Clostridium acetobutylicum* $\Delta\text{cac1502}\Delta\text{upp}$ [42].

For the complementary strain, continuous culture under acidogenic condition was performed in the presence of clarithromycin at $40\text{ }\mu\text{g/ml}$ for a short period to avoid losing the plasmid. The production profile was analyzed and compared with the mutant Cap0037::int.

6.5.4 Isolation of total mRNA and microarray

Total RNA was isolated from continuous cultures using the RNEasy kit protocol (Qiagen, Hilden, Germany), except for the cell breaking step that we adapted for *C. acetobutylicum*: 3 ml of culture were taken from the chemostat, transferred into a 15-ml Falcon tube, and immediately frozen in liquid nitrogen. The frozen culture was mechanically crushed in a mortar containing liquid nitrogen. After that, the extraction was processed using the Qiagen protocol. The quality and quantity of the mRNA samples were controlled before proceeding to labeling, hybridization, one color microarray-based gene expression, and scanning using a Tecan MS200 scanner, followed by analysis using Feature Extraction V.11.5.1.1. The protocol and microarray slides used in this study were the same as described elsewhere [42] to obtain comparable results.

6.5.5 Analytical methods

The concentrations of the substrate and fermentation products were measured by high-pressure liquid chromatography (HPLC, Agilent 1200 series, Massy, France). The separation was obtained with an Aminex HPX-87H column (300 x 7.8 mm, Biorad), at 14°C with $0.5\text{ mM H}_2\text{SO}_4$ at a flow rate of 0.5 ml/min . Detection was achieved by reading the refractive index and UV absorbance (210 nm).

6.5.6 Southern blot analysis

Genomic DNA of the mutant Cap0037::int was digested with *EcoRI* restriction enzyme. The probe for hybridization was prepared by PCR using Intron probe-F and Intron probe-R primers and was then labeled using the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche). The protocols for hybridization and detection were performed according to the Roche instructions.

6.5.7 Expression and purification of His-tagged protein

E. coli BL21 (DE3) cells (Novagen R & D Systems) containing pBS2 was used for overproduction of protein Cap0037. Cells were grown aerobically at 37 °C in LB medium supplemented with either kanamycin (50 µg/ml) or ampicillin (100 µg/ml) to OD₆₀₀ 0.7 and induced by addition of 1 mM isopropyl β-D-1 thiogalactopyranoside (IPTG) at 37 °C, until stationary phase was reached. Cells were then harvested by centrifugation (4300 x g, 10 min, 4 °C) and washed twice with buffer containing imidazole (10 mM). Wet weight of the cell pellet was determined and solubilized in three volumes of the buffer mentioned before. Cells were disrupted using a French[®] pressure cell (4 MPa) (FRENCH[®] Pressure Cell/Press, G. HEINEMANN Ultraschall- und Labortechnik, Schwäbisch Gmünd, Germany). The fusion protein was purified by nickel-nitrilotriacetic acid (NTA) affinity chromatography using Ni-NTA Agarose (Qiagen GmbH, Hilden, Germany). For elution, imidazole buffers with 75, 100, 200, and 250 mM imidazole were used. Purity of the protein was analyzed by SDS polyacrylamide gel electrophoresis, and protein concentration was quantified by Pierce BCA protein assay kit (Pierce, Rockford, USA) according to the instructions provided by manufacturer.

6.5.8 Electromobility shift assays (EMSAs)

For electromobility shift assays of the *adc*, *CA_P0037/CA_P0036*, and *sol* operon, respective DNA probes were amplified by PCR using genomic DNA from *C. acetobutylicum* ATCC 824 as template. For amplifying the 5'untranslated region of the *adc* gene, the *adc* probe (318 bp), primers *adc_forward* and *adc_reverse* were used. For EMSAs to the 5'untranslated region of the *CA_P0037/CA_P0036* operon, the CAP0037 probe (539 bp), primers *94_fwd* and *PEX1_kurz* were used. For EMSAs to the 5'untranslated region of the *sol* operon, *sol* probe (219 bp), primers *sol_frag_F* and *sol_frag_R* were used. The *adcint* probe (233 bp) was used as negative control and was amplified out of the internal region of the *adc*

gene, using primers *adcm1* and *adcm2*. In the binding assays, DNA probes were mixed with increasing amounts of the respective fusion proteins in 20- μ l reactions with 0.5 μ l poly[d(I-C)] (2 μ g/ μ l) and 2 μ l band shift buffer (200 mM Tris, 500 mM Na-glutamate, 100 mM MgCl₂ x 6 H₂O, 63.6 mM EDTA, 0.5 % Nonidet P-40 (v/v), 50 % glycerol (v/v)). After incubation for 20 min at room temperature, samples were loaded on a 2 % agarose gel in 1x TAE buffer (200 mM Tris, 100 mM acetate, 5 mM EDTA, pH 7.5), and separated by electrophoresis (70 V, 50 min). The gel was stained using ethidium bromide.

For electromobility shift assays of *ldhA* and *fld1*, PCR DNA fragments of *ldhA* and *fld1* promoters were generated by two pairs of primers (*ldh-prom-F* and *ldh-prom-R*; *fld1-prom-F* and *fld1-prom-R*) and DIG-labeled (with DIG High Prime DNA Labeling and Detection Starter Kit II (Roche)). Binding assays were done as described above. The detection of DNA shifts was performed according to Roche protocol.

6.5.9 DNase I protection assays (DNA Footprinting)

To determine the exact binding sites of protein Cap0037 at the 5' untranslated region of the *CA_P0037/CA_P0036* operon, *adc* operon, and *sol* operon, different DNA fragments (probes) were amplified by PCR (6.5.2). The primers generated restriction sites at both ends of the probes (6.5.2). Probes were radioactively 5' labeled using [γ -³²P] ATP. Labeling was performed in a 25 μ l reaction with 20 pmol DNA, 2 μ l polynucleotide kinase (10 U/ μ l) (Thermo Scientific, St. Leon-Rot, Germany), 2.5 μ l Reaction Buffer (10x Binding Buffer (see below) and 2 μ l BSA 100 mg/ml (w/v)), and 4 pmol [γ -³²P] ATP (20 μ Ci). The reaction was incubated at 37 °C for 1 h. One end of the DNA was digested using *EcoRV*, removing a 5' labeled phosphate. Purification of the labeled DNA was carried out by Microspin™ G-25 columns (GE Healthcare, Buckinghamshire, UK) according to the instructions of the manufacturer. For protein binding, respective radio labeled DNA probe was mixed with increasing amounts of Cap0037 in a 20 μ l reaction together with 0.5 μ l poly[d(I-C)] (2 μ g/ μ l) and 2 μ l 10x Binding Buffer (500 mM Tris, 500 mM KCl, 50 mM MgCl₂ x 6 H₂O, 10 mM EDTA). After incubation for 30 min at 37 °C, 1.1 μ l Buffer R1 (120 mM MgCl₂ x 6 H₂O, 120 mM CaCl₂ x 2 H₂O) and 0.1 U DNase (Thermo Scientific, St. Leon-Rot, Germany) were added. Reaction was stopped after just 1 min by adding 1 μ l EDTA (0.5 M, pH 8) and transfer on ice. The sample was purified by phenol chloroform extraction, the DNA was ethanol precipitated, and suspended in 3 μ l of sequencing gel loading buffer (30 mg bromphenol blue,

30 mg xylene cyanole, 200 mEDTA (pH 8), ad 10 ml formamide). After an incubation at 80 °C for 2 min the DNase I digestion product was loaded on a 7 M urea 6 % polyacrylamide gel and separated at 1500 V, 46 mA, and 60 W for 2 h. Dideoxy sequencing reaction was performed using respective plasmid (6.5.2) and T7 Sequencing Kit (Affymetrix, Santa Clara, USA), and run in the same gel as the footprint sample. Gel was vacuum dried, exposed on Hyperfilm™ MP (Amersham Biosciences Europe GmbH; Freiburg) for 24 h, and developed with VCURIX 60V (AGFA Graphics GmbH & Co. KG, Düsseldorf, Germany).

6.5.10 Bioinformatic tools

6.5.10.1 Blast searches, amino acid sequence alignments and phylogenetic analysis

Identification of orthologs was performed using the BLAST tool provided by NCBI [218]. Amino acid sequence alignment and the phylogenetic tree were processed using the phylogeny program [219].

6.5.10.2 Transmembrane protein prediction

Cap0036 amino acid sequence was used as an input to find possible transmembrane regions using two bioinformatics web-based tools: TMPred [200] and OCTOPUS [201].

6.5.10.3 DNA binding motif search and promoter' elements prediction

Amino acid sequences of Cap0037 were subject to a search for the helix-turn-helix DNA binding motif using the NPS@ web-based bio-tool (<https://npsa-prabi.ibcp.fr/>) [202].

The prediction the promoter of different genes, BPROM web-based tool was used [203].

6.5.10.4 Putative Cap0037 regulon

Construction of putative Cap0037 regulon was performed using an established comparative genomics method [220] implemented in the RegPredict webserver (<http://regpredict.jbl.gov>) [212] and Artemis Java Webstart [221]. Genes with potential upstream binding sites that had high scores and/or were conserved were rebuilt to improve search accuracy. Genome sequences and annotations of *C. acetobutylicum* were obtained from GenBank (<http://www.ncbi.nlm.nih.gov/GenBank/>).

6.5.11 Microarray data accession number

The microarray data can be accessed at GEO through accession number GSE81273 and GSE69973 for the Cap0037::int mutant and the control strain [42], respectively.

6.6 Acknowledgement

We thank Sophie Lamarre and Lidwine Trouilh for help with the data analysis.

This work was financially supported by the European Community's Seventh Framework Program "CLOSTNET" (PEOPLE-ITN-2008-237942) to Ngoc-Phuong-Thao NGUYEN

6.7 Supplementary

Table S- 6-1 Bacterial strains, plasmids and primers used in this study

Table S- 6-2 Relative transcript levels of selected genes of Cap0037::int *C. acetobutylicum* mutant in the three metabolic states: acidogenesis (AC), alcohologenesis (AL) and solventogenesis (SO)

Table S- 6-1 Bacterial strains, plasmids and primers used in this study

Strain/plasmid	Relevant characteristics	Source
<i>Clostridium acetobutylicum</i> and <i>Escherichia coli</i> strains		
Control strain	<i>Clostridium acetobutylicum</i> Δ cac1502 Δ upp	[42]
Mutant Cap37::int	<i>Clostridium acetobutylicum</i> Δ cac1502 Δ uppcap0037::int(189s)	This study
Complementary strain	<i>Clostridium acetobutylicum</i> Δ cac1502 Δ uppcap0037::int(189s) pSOS95-cap0037	This study
<i>Escherichia coli</i> TOP10	<i>F</i> - mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80lacZ Δ M15 Δ lacX74 nupG recA1 araD139 Δ (ara-leu)7697 galE15 galK16 rpsL(Str ^R) endA1 λ ⁻	Invitrogen
Plasmids		
pCR4-TOPO-Blunt	Ap ^r Km ^r	Invitrogen
pMTL007C-E2	Cm ^R , Clostridial expression vector for the expression of ClosTron containing ErmRAM	[84]
pMTL007C-E2::cap37(277a)	Cm ^R , Clostridial expression vector for the expression of ClosTron containing ErmRAM, retargeting CA_P0037_277 278a	This study
pCUI	catP, erm, ltrA, repL, upp, ColE1 origin, intron	[182]
pCUI-cap37(189s)	catP, erm, ltrA, repL, upp, ColE1 origin, intron retargeting CA_P0037 189 190s	This study
pSOS95	Ap ^r MLS ^r , repL gene, ColE1 origin	[154]
pSOS95-Cap0037	Ap ^r MLS ^r , repL gene, ColE1 origin, Cap0037 operon	This study
pBS2	pET29a(+), adcRx6His-Fusion	This study
pDrive	Ap ^R , Km ^R , pMB1 origin, lacPOZ'	This study
pDrive_adcR	pDrive vector with the promoter region of CA_P0037/CA_P0036 (+487 to -108)	This study
pDrive_86kurz	pDrive vector with the promoter region of CA_P0037/CA_P0036 (-219 to +21)	This study

pDrive_144	pDrive vector with the promoter region of <i>CA_P0037/CA_P0036</i> (-233 to +6)	This study
pDrive_adcStart	pDrive vector with the promoter region of the <i>adc</i> operon (-105 to +49)	This study
pDrive_sol	pDrive vector with the promoter region of the <i>sol</i> operon (-233 to -15)	This study

Primer	Sequence
CA_P0037n-189 190s -IBS	aaaaaagcttataattatccttagaaaacattaaggtgcgccagatagggtg
CA_P0037n-189 190s -EBS1d	cagattgtacaaatgtggtgataacagataagtcattaagagtaacttacctttcttgt
CA_P0037n-189 190s -EBS2	tgaacgcaagtttctaatttcgatttttctcgatagaggaaaagtgtct
EBS universal	cgaaattagaaacttgcgttcagtaaac
CAP0037-277 278a-IBS	aaaaaagcttataattatccttatgttctgtccagtgcgccagatagggtg
CAP0037-277 278a-EBS1d	cagattgtacaaatgtggtgataacagataagtcgtccagctaacttacctttcttgt
CAP0037-277 278a-EBS2	tgaacgcaagtttctaatttcggttaacatcgatagaggaaaagtgtct
CAP0037-R-b	ttctctagcgtttacttttagaacaagag
CAP0037-F-b	atggatgaaactaatttaatttctaaaa
CAP0037_F	attttcccaaacttttcagagctttc
CAP0037_R	taagttcattatagtttgcgaaacaatg
cap37-F-RBS-BamH1	aaaaaaggatccaggatctcaagcaaatcgagaggacacagggtatggatgaa actaatttaatttctaaaaagaacttc
cap37-R-SfoI	tttttttggcgccttatagtgataattttgattttaattcttctaaacattctc
Intron probe-F	accctcttcgggaaccgtacg
Intron probe-R	gacttaacactaactggggataccc
REP-upp-F	aaaacagctgggaggaatgaaataatgagtaaagttacac
REP-upp-R	aaaacagctgttattttgtaccgaataatctatctccagc
ygas1hincI	gaatagccatatggatgaaac
ygas2rückcl	tcctttaagctttagtgataattttgattttaattcttctaaac
PEX1_kurz	ggcaagctcatctaataatgatacatccc
PEX2_lang	cccggatttttaggaacctatctaataatagtc
adcRpDrive	cactccttatctatagttcg
adcRpDrive_rev	ctacgtcacaaatcgggtcatc
86 fwd	<u>gaattc</u> ccctccaatattatc
85 rev	<u>gatatc</u> ccctccagttattctc
144 rev	agaga <u>attccc</u> attttaccttctattctaac
145 fwd	agaga <u>tatcc</u> attataacacatattg

adestart FP EcoRV fwd	acagatatcggcgaagttaatggcgtg
adestart FP EcoRI rev	acagaattcgcagaatttttaggaagtg
sol FP EcoRI fwd	acagaattcctcaattctatttatgctcc
sol FP EcoRV rev	acagatatccaataaaaagcataggtc
adc_forward	gccaacattgcctataag
adc_reverse	gttaatggcgtgctaattg
94_fwd	ctttatcaacttcctcaacggg
ademit1	gatgagcccttagtcagg
ademit2	cctactaaagtatctgaatcc
sol_frag_F	actcaattctatttatgctcc
Sol_frag_R	gcataggtcaaaaatataacag
ldh-prom-F	ctgcataatataaaatgcagattttttaaac
ldh-prom-R	acaaatcctgctccgattactg
fld1-prom-F	gttatgtagaatttatatacataattagagc
fld1-prom-R	caataaattatgtt tattttcacc

Table S- 6-2 Relative transcript levels of selected genes of Cap0037::int *C. acetobutylicum* mutant in the three metabolic states: acidogenesis (AC), alcohologenesis (AL) and solventogenesis (SO)

Genes	Fold regulation		
	AC	AL	SO
Arginine transport and metabolism			
CAC0316 argF/I	1.548	0.456	1.2632
CAC0366 Predicted permease	nd	nd	0.792
CAC0367 rocB	1.061	0.462	1.261
CAC0368 gabT	1.023	0.645	1.121
CAC0837 potD	0.589	0.951	0.844
CAC0838 potC	0.587	0.780	0.632
CAC0839 potB	0.424	0.604	0.481
CAC0840 potA	0.579	0.822	0.721
CAC0841 Predicted TR	0.960	0.996	0.944
CAC0973 argG	2.919	1.145	1.550
CAC0974 argH	2.733	0.984	1.660
CAC1054 arginase	0.530	1.159	0.746
CAC2388 argD	5.208	0.737	1.148
CAC2389 argB	3.312	0.818	1.223
CAC2390 argC	2.274	0.825	1.091
CAC2391 argJ	4.518	1.139	1.866
CAC2645 carA	6.963	4.182	3.305
CAC2770 ABC transporter	0.677	0.418	0.393

CAC2771 ABC transporter	0.644	0.327	0.350
CAC3618 ABC transporter	1.984	1.131	2.993
CAC3619 ABC transporter	0.944	0.597	1.944
CAC3620 ABC transporter	1.722	0.938	2.567
CAC2074 ArgR	0.963	1.176	0.722
Histidine and Tryptophan biosynthesis			
CAC0935 hisZ	5.395	0.950	1.827
CAC0936 hisG	5.853	1.139	2.055
CAC0937 hisD	3.958	1.040	1.583
CAC0938 hisB	4.852	1.036	1.803
CAC0939 hisH	4.781	0.953	1.578
CAC0940 hisA	5.298	1.322	2.023
CAC0941 hisF	3.473	0.775	1.146
CAC0942 hisI-1	6.371	1.549	2.077
CAC0943 hisI-2	3.827	0.958	0.991
CAC0944 tkt	2.498	1.299	1.169
CAC3031 hisC	0.714	0.557	0.515
CAC3157 trpA	1.742	3.949	0.527
CAC3158 trpB	2.810	4.231	1.069
CAC3159 trpF	3.354	5.361	0.873
CAC3160 trpC	2.053	4.431	0.642
CAC3161 trpD	3.661	8.180	0.932
CAC3162 pabA	2.838	6.782	0.586
CAC3163 pabB	2.603	6.534	0.529

The putative CymR and IscR regulons			
CAC0093 LysR family TR	1.236	0.969	0.734
CAC0094 FNR	1.163	2.140	1.263
CAC0102 O-ahs sulfhydrylase	29.638	0.176	21.117
CAC0103 cysC	33.316	0.100	43.961
CAC0104 adenylylsulfate red	43.769	0.086	50.771
CAC0105 ferredoxin	25.441	0.083	27.181
CAC0106 sulfate transporter	16.736	0.107	10.029
CAC0107 sulfate transporter	18.367	0.127	31.006
CAC0108 sulfate transporter	32.540	0.111	41.051
CAC0109 cysD	34.231	0.087	48.408
CAC0110 cysN	58.430	0.109	29.921
CAC0389 hyp. protein	1.493	0.746	1.110
CAC0390 metB	2.168	0.837	1.861
CAC0391 metC	1.404	0.683	1.165
CAC0878 ABC transporter	2.950	1.019	1.175
CAC0879 ABC transporter	7.680	1.462	2.092
CAC0880 ABC transporter	9.142	1.156	1.734
CAC0929 methyltransferase	1.858	0.638	1.825
CAC0930 metB	3.005	0.374	4.302
CAC0931 cysK	3.096	0.424	3.582
CAC0984 ABC transporter	1.098	0.559	2.064
CAC0985 ABC transporter	0.803	0.452	2.208
CAC0986 lipoprotein	0.976	0.714	2.682

CAC1675 iscR	0.642	0.326	0.408
CAC2235 cysK	8.393	0.524	6.179
CAC2236 cymR	3.808	0.505	3.769
CAC3325 ABC transporter	12.794	0.337	5.551
CAC3326 ABC transporter	7.706	0.192	4.219
CAC3327 ABC transporter	12.580	0.328	5.630
Nucleic acid transport and metabolism			
CAC2644 carA	13.827	9.332	10.624
CAC2645 carB	6.963	4.181	3.305
CAC2650 pyrD	68.714	26.067	34.395
CAC2651 pyrZ	95.625	35.050	38.377
CAC2652 pyrF	101.855	31.113	39.511
CAC2653 pyrI	53.008	17.150	18.604
CAC2654 pyrB	80.451	25.042	33.751
CAC1390 purE	1.569	0.753	1.224
CAC1391 purC	3.023	1.018	0.799
CAC1392 purF	6.053	0.830	0.865
CAC1393 purM	2.866	0.505	0.456
CAC1394 purN	5.316	0.830	0.747
CAC1395 purH	3.984	0.812	0.745
CAC1396 purD	3.690	0.999	0.700
CAC1655 purQ/purL	4.219	0.756	0.746
CAC2112 UraA/PyrP	6.361	1.778	3.019
Cation transporters			

CAC3551 Na ⁺ ABC transporter	0.774	1.710	1.077
CAC3550 Na ⁺ ABC transporter	0.301	0.670	0.377
CAC3679 kdpX	4.335	1.070	1.297
CAC3681 kdpC	3.287	0.000	0.863
CAC3680 kdpD	6.938	1.004	1.028
Glycerol lipid metabolism			
CAC1319 glpF	0.432	0.271	0.393
CAC1321 glpK	0.461	0.207	0.333
CAC1322 glpA	3.428	0.297	2.549
CAC0427 Glycerol3P ABC transporter	0.117	10000.000	1.451
CAC0428 sugar permease	0.189	0.382	1.834
CAC0429 Glycerol3P ABC transporter	0.177	0.000	1.292
CAC0430 Glycerophosphodiester phosphodiesterase	0.168	0.753	1.254
Mobility			
CAC0120 mcp	2.324	3.572	0.468
CAC0250 mcp	0.967	1.670	0.620
CAC0805 mcp	1.101	2.062	0.758
CAC0815 mcp	0.053	0.680	0.487
CAC0909 mcp	0.576	0.775	0.704
CAC1387 mcp	3.790	4.590	0.959
CAC2617 mcp	0.830	1.355	0.770
CAC3352 mcp	0.369	1.401	0.213
CAC3476 mcp	1.005	1.148	0.612
CAC3688 mcp	1.009	2.150	0.762

Chemotaxis			
CAC0304 motA	1.113	3.168	0.277
CAC0305 motB	0.550	1.496	0.181
CAC1845 motB	1.543	1.399	0.747
CAC1846 motA	0.952	0.903	0.562
CAC2139 flgG	0.686	1.638	1.223
CAC2140 flgG	0.719	1.031	0.661
CAC2147 flhA	0.554	1.158	0.330
CAC2148 fliR/flhB	0.337	0.621	0.186
CAC2149 fliQ	0.215	0.408	0.193
CAC2150 fliP	0.228	0.459	0.183
CAC2151 fliZ	0.532	1.093	0.371
CAC2154 flgE	0.473	0.788	0.258
CAC2155 hyp. protein	0.674	1.114	0.238
CAC2156 flgD	0.734	1.352	0.335
CAC2157 fliK	0.937	1.793	0.398
CAC2158 fliJ	0.433	0.874	0.351
CAC2159 flg ATPase	0.875	1.670	0.920
CAC2160 fliH	0.574	1.057	0.652
CAC2161 fliG	0.548	1.094	0.729
CAC2162 fliF	0.632	1.217	0.809
CAC2163 fliE	0.673	1.246	1.003
CAC2164 flgC	0.689	1.189	1.038
CAC2165 flgB	0.745	1.139	0.918

CAC2205 fliD	0.696	2.464	0.395
CAC2206 fliS	1.092	4.284	0.358
CAC2211 flagellin	0.550	2.410	0.295
CAC2212 flgK	0.476	1.728	0.212
CAC2214 flgM	0.591	1.826	0.486
CAC2215 fliY	0.487	2.733	0.529
CAC2216 fliM	0.359	2.091	0.487
Sporulation			
CAC0310 abrB	0.097	0.600	0.971
CAC3647 abrB	7.097	2.539	1.316
CAC0078 abrB	0.343	3.632	1.133
CAC0903 histidine kinase	0.970	0.875	1.498
CAC0323 histidine kinase	0.861	0.793	0.748
CAC3319 histidine kinase	0.282	1.699	0.755
CAC0437 histidine kinase	0.126	0.503	2.552
CAC2071 spo0A	0.994	0.912	0.957
CAC3735 spoIIJ	0.863	0.984	0.736
CAC2306 sigF	0.373	0.841	3.638
CAC1689 sigK	1.149	0.821	4.438
CAC1696 sigG	0.947	0.776	5.000
CAC1695 sigE	1.035	0.698	5.000
CAC2308 spoIIAA	0.000	0.815	3.529
CAC0644 gerKA	0.719	0.796	1.103
CAC0699 splB	0.625	0.594	0.406

CAC1620 spore protein	0.656	0.737	0.632
CAC3729 spoJ	1.017	0.994	0.712

Chapter 7 Conclusion and future perspectives

The metabolic engineering part of this thesis describes two recombination system to engineer *C. acetobutylicum*. Both of them uses *upp* as a counter-selectable marker. One system uses replicative plasmid and the other non-replicative plasmid. The non-replicative homologous recombination system is time-saving because it doesn't need the flippase involving step to remove the selective marker as the replicative recombination system does to obtain a marker-less strain. Furthermore, the non-replicative homologous recombination system i) does not leave an FRT scar in the genome, ii) offers the possibility to introduce point mutations and iii) allows the simultaneous deletion of two genes. Those two methods have been successfully used to engineer *C. acetobutylicum* strains. The replicative homologous system was employed to delete *cac1502*, *upp*, *cac3535*, *ctfAB*, *ldhA*, *ptb*, *buk*, *rex*, and to directly replace *hbd* with *hbd1* gene. The non-replicative plasmids were successfully used to directly replace *thlA* with synthetic *atoB* gene and also to simultaneously delete *hydA* and *thlA* genes.

The deletion of *cac1502* and *upp* (with or without the deletion of *cac3535*) produced a convenient background strain that is useful for further metabolic engineering and for functional genomic studies of *C. acetobutylicum*.

The deletion of *ctfAB*, *ldhA*, *ptb*, and *buk* resulted in *C. acetobutylicum* CAB1057 strain unable to produce lactate, acetone, and butyrate. *Rex* gene was deleted to eliminate the inhibition effect on the main metabolism fluxes. The replacement of *hbd* by *hbd1* and of *thlA* by *atoB* has successfully pulled carbon flux to n-butanol production. In the final strain CAB1060, n-butanol was produced at high yield and selectivity. Along with the decrease in acetate production, the hydrogen production was also significantly low and most of the electrons from reduced ferredoxin were used to produced NAD(P)H. This is the best strain ever reported that produces n-butanol at high yield and selectivity. This strain was subjected to a new continuous fermentation process using i) in situ extraction of alcohols by distillation under low pressure and ii) high cell density cultures. This process proved to increase the titer, yield and productivity of n-butanol production to levels that have never been previously achieved in any organism. This process provides a means to selectively produce n-butanol

from different feedstocks at performance levels that are now compatible with a commercial process. However, this mutant could be further optimized to eliminate totally the acetate production. To do so, all the electrons from reduced ferredoxin should be used to produced NAD(P)H, thus, the ferredoxin-NAD(P)H- reductase has to be identified and overexpressed.

By using the Genome Scale Model developed in our group, we demonstrated that *C. acetobutylicum* can potentially continuously produce several bio-chemicals if we manage to block hydrogen production. By simultaneously inactivating *hydA* and *thlA*, we managed to obtain a platform strain that no longer produces hydrogen. This platform strain was further engineered to continuously produce L-lactate and 1,3-propanediol at high yield, titer and productivities. Further work is currently carried out to construct strains producing either, isopropanol, 1,3-butanediol or isobutene.

Lastly, in the strict anaerobe *C. acetobutylicum*, Cap0037 was found to directly or indirectly control a large set of genes belonging to the Rex, PerR and Fur regulons. This control drastically affects metabolism under both acidogenic and alcohologenic conditions. More study is needed to understand the regulatory mechanism of Cap0037, including clarification of the particular location of Cap0037 in the cell structure, finding the effectors and probably a second protein that may interact or compete with Cap0037 (such as Cap0036 or Spo0A) in response to cell signals. This would be a valuable information to understand the function of this conserved protein in other firmicutes.

Abbreviations

5'-FC	5-fluorocytosine
5'-FOA	5-fluoroorotic acid
5-FU	5-fluorouracil
AB	acetone, butanol
ABE	acetone, butanol, ethanol
AC	acidogenesis
ACE	Allele couple exchange
ACK	acetate kinase
ADC	acetoacetate decarboxylase
ADHE	aldehyde alcohol dehydrogenase
ADP	adenosine-5' diphosphate
AL	alcohologenesis
ALSD	acetolactate decarboxylase
ALSS	acetolactate synthase
Apr	ampicillin resistance
asRNA	antisense RNA
ATP	adenosine-5' triphosphate
BCD	butyryl coA dehydrogenase
BDH	butanol dehydrogenase
BUK	butyrate kinase
CcpA	catabolite control protein
CGM	Clostridium growth medium

CGMMG	CGM with 50g/L Glucose and 0.1M MES, pH 6.3
Cm^r	chloramphenicol resistance
Cm	chloramphenicol
CoA	Coenzyme A
COGs	Cluster Orthologous Groups
CRT	crotonase
CsrA	carbon storage regulator
CTFA/B	coA-transferase subunit A/B
Ctrl	Control
DMSO	Dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothretol
EMP	Embden-Meyerhof glycolytic pathway
ETFA/B	electron transfer flavoprotein subunit A/B
Ery^R	Erythromycin resistant
FAD	flavin adenine nucleotide
Fd-ox	oxidized ferredoxin
FDP	flavodiiron protein
Fd-red	reduced ferredoxin
FLP	flippase
FMN	flavin mononucleotide
FNOR	NAD(P) ferredoxin oxidoreductase
FRT	flippase recognition target
Fur	ferric uptake regulator

gDNA	genomic DNA
GSM	Genome scale metabolic
HBD	hydroxybutyryl-coA dehydrogenase
HYD	hydrogenase
j3TI	j3TI methyltransferase
LDH	lactate dehydrogenase
McrBC⁻	lacking methylcytosine-specific restriction system
MLS^r	macrolide lincosamide and streptogramin B resistance
mRNA	messenger RNA
MT	mutant
NAD⁺	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NADP⁺	nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate
ORF	open reading frame
PCR	polymerase chain reaction
PerR	peroxide repressor
PFOR	pyruvate ferredoxin oxidoreductase
PTA	phosphotransacetylase
PTB	phosphotransbutyrylase
PTS	phosphotransferase systems
RBS	ribosome binding site
RCM	reinforced clostridial medium
RecA⁻	homologous recombination abolished

repL	Gram-positive origin of replication from pIM13
Rex	redox sensing transcriptional repressor
SM	synthetic medium
SO	solventogenesis
sRNA	small RNA
THL	thiolase
Th^R	thiamphenicol resistant
Tm	thiamphenicol
UPRTase	uracil phosphoribosyl-transferase
WT	wild-type
Zeo^R	zeomycin resistance

References

1. C.L. G. Butanol Fermentation Process. Ind. Eng. Chem. 1928;20:1063.
2. Jones DT, Woods DR. Acetone-butanol fermentation revisited. Microbiol. Rev. American Society for Microbiology (ASM); 1986;50:484–524.
3. Yadav S, Rawat G, Tripathi P, Saxena RK. A novel approach for biobutanol production by *Clostridium acetobutylicum* using glycerol: A low cost substrate. Renew. Energy. 2014;71:37–42.
4. Ndaba B, Chiyanzu I, Marx S. Direct fermentation of sweet sorghum juice by *Clostridium acetobutylicum* and *Clostridium tetanomorphum* to produce bio-butanol and organic acids. Biofuel Res. J. 2015;2:248–52.
5. Dong H, Tao W, Zhang Y, Li Y. Development of an anhydrotetracycline-inducible gene expression system for solvent-producing *Clostridium acetobutylicum*: A useful tool for strain engineering. Metab. Eng. Elsevier; 2012;14:59–67.
6. Wu P, Wang G, Wang G, Børresen BT, Liu H, Zhang J. Butanol production under microaerobic conditions with a symbiotic system of *Clostridium acetobutylicum* and *Bacillus cereus*. Microb. Cell Fact. 2016;15:8.
7. Wang Z, Cao G, Zheng J, Fu D, Song J, Zhang J, et al. Developing a mesophilic co-culture for direct conversion of cellulose to butanol in consolidated bioprocess. Biotechnol. Biofuels. BioMed Central; 2015;8:84.
8. Pang Z-W, Lu W, Zhang H, Liang Z-W, Liang J-J, Du L-W, et al. Butanol production employing fed-batch fermentation by *Clostridium acetobutylicum* GX01 using alkali-pretreated sugarcane bagasse hydrolysed by enzymes from *Thermoascus aurantiacus* QS 7-2-4. Bioresour. Technol. 2016;212:82–91.
9. Cai D, Wang Y, Chen C, Qin P, Miao Q, Zhang C, et al. Acetone–butanol–ethanol from sweet sorghum juice by an immobilized fermentation-gas stripping integration process. Bioresour. Technol. 2016;211:704–10.
10. Nanda S, Dalai AK, Kozinski JA. Butanol and ethanol production from lignocellulosic feedstock: biomass pretreatment and bioconversion. Energy Sci. Eng. 2014;2:138–48.
11. Gaida SM, Liedtke A, Jentges AHW, Engels B, Jennewein S. Metabolic engineering of *Clostridium cellulolyticum* for the production of n-butanol from crystalline cellulose. Microb. Cell Fact. BioMed Central; 2016;15:6.
12. Gottumukkala LD, Parameswaran B, Valappil SK, Mathiyazhakan K, Pandey A, Sukumaran RK. Biobutanol production from rice straw by a non acetone producing

Clostridium sporogenes BE01. Bioresour. Technol. 2013;145:182–7.

13. Boonsombuti A, Tangmanasakul K, Nantapipat J, Komolpis K, Luengnaruemitchai A, Wongkasemjit S. Production of biobutanol from acid-pretreated corncob using *Clostridium beijerinckii* TISTR 1461: Process optimization studies. Prep. Biochem. Biotechnol. 2016;46:141–9.

14. Minton NP, Clarke DJ, editors. Clostridia. Boston, MA: Springer US; 1989.

15. Tracy BP, Jones SW, Fast AG, Indurthi DC, Papoutsakis ET. Clostridia: The importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. Curr. Opin. Biotechnol. 2012;23:364–81.

16. Schwarz WH. The cellulosome and cellulose degradation by anaerobic bacteria. Appl. Microbiol. Biotechnol. 2001;56:634–49.

17. Hein S, Söhling B, Gottschalk G, Steinbüchel A. Biosynthesis of poly(4-hydroxybutyric acid) by recombinant strains of *Escherichia coli*. FEMS Microbiol. Lett. 1997;153:411–8.

18. Weyer ER, Rettger LF. A comparative study of six different strains of the organism commonly concerned in large scale production of butyl alcohol and acetone by the biological process. J. Bacteriol. 1927;14:399–424.

19. Nölling J, Breton G, Omelchenko M V, Kira S, Zeng Q, Gibson R, et al. Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum* Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum*. J. Bacteriol. 2001;183:a823–4838.

20. Girbal L, Soucaille P. Regulation of solvent production in *Clostridium acetobutylicum*. Trends Biotechnol. 1998;16:11–6.

21. Girbal L, Vasconcelos I, Saint-Amans S, Soucaille P. How neutral red modified carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH. FEMS Microbiol. Rev. 1995;16:151–62.

22. Vasconcelos I, Girbal L, Soucaille P. Regulation of carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol. J. Bacteriol. 1994;176:1443–50.

23. Grupe H, Gottschalk G. Physiological Events in *Clostridium acetobutylicum* during the Shift from Acidogenesis to Solventogenesis in Continuous Culture and Presentation of a Model for Shift Induction. Appl. Environ. Microbiol. 1992;58:3896–902.

24. Husemann MW, Papoutsakis ET, Husemann MHW, Papoutsakis ET. Comparison between in vivo and in vitro enzyme activities in continuous and batch fermentations of *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. Springer-Verlag; 1989;30:585–95.

25. Tangney M, Winters GT, Mitchell WJ. Characterization of a maltose transport system in *Clostridium acetobutylicum* ATCC 824. J. Ind. Microbiol. Biotechnol. 2001;27:298–306.
26. Meinecke B, Bertram J, Gottschalk G. Purification and characterization of the pyruvate-ferredoxin oxidoreductase from *Clostridium acetobutylicum*. Arch. Microbiol. Springer-Verlag; 1989;152:244–50.
27. Demuez M, Cournac L, Guerrini O, Soucaille P, Girbal L. Complete activity profile of *Clostridium acetobutylicum* [FeFe]-hydrogenase and kinetic parameters for endogenous redox partners. FEMS Microbiol. Lett. 2007;275:113–21.
28. Girbal L, von Abendroth G, Winkler M, Benton PMC, Meynial-Salles I, Croux C, et al. Homologous and heterologous overexpression in *Clostridium acetobutylicum* and characterization of purified clostridial and algal Fe-only hydrogenases with high specific activities. Appl. Environ. Microbiol. 2005;71:2777–81.
29. King PW, Posewitz MC, Ghirardi ML, Seibert M. Functional studies of [FeFe] hydrogenase maturation in an *Escherichia coli* biosynthetic system. J. Bacteriol. 2006;188:2163–72.
30. Girbal L, Croux C, Vasconcelos I, Soucaille P, Bioingdnierie C De, Durand G, et al. Regulation of metabolic shifts in *Clostridium acetobutylicum*. FEMS Microbiol. Rev. 1995;17:287–97.
31. Vasconcelos I, Soucaille P. Regulation of Carbon and Electron Flow in *Clostridium butyricum* VPI 3266 grown on glucose-glycerol mixtures. Society. 2001;183:1748–54.
32. Winzer K, Lorenz K, Zickner B, Dürre P. Differential regulation of two thiolase genes from *Clostridium acetobutylicum* DSM 792. J. Mol. Microbiol. Biotechnol. 2000;2:531–41.
33. Stim-Herndon KP, Petersen DJ, Bennett GN. Characterization of an acetyl-CoA C-acetyltransferase (thiolase) gene from *Clostridium acetobutylicum* ATCC 824. Gene. 1995;154:81–5.
34. Wiesenborn DP, Rudolph FB, Papoutsakis ET. Thiolase from *Clostridium acetobutylicum* ATCC 824 and Its Role in the Synthesis of Acids and Solvents. Appl. Environ. Microbiol. American Society for Microbiology (ASM); 1988;54:2717–22.
35. Mann MS, Lütke-Eversloh T. Thiolase engineering for enhanced butanol production in *Clostridium acetobutylicum*. Biotechnol. Bioeng. 2013;110:887–97.
36. Kim S, Jang Y-S, Ha S-C, Ahn J-W, Kim E-J, Hong Lim J, et al. Redox-switch regulatory mechanism of thiolase from *Clostridium acetobutylicum*. Nat. Commun. Nature Publishing Group; 2015;6:8410.

37. Sommer B, Garbe D, Schrepfer P, Brück T. Characterization of a highly thermostable β -hydroxybutyryl CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824. J. Mol. Catal. B Enzym. Elsevier B.V.; 2013;98:138–44.
38. Lehmann DD, Lütke-Eversloh T. Switching *Clostridium acetobutylicum* to an ethanol producer by disruption of the butyrate/butanol fermentative pathway. Metab. Eng. 2011;13:464–73.
39. Kim EJ, Kim J, Ahn JW, Kim YJ, Chang JH, Kim KJ. Crystal structure of (s)-3-hydroxybutyryl-CoA Dehydrogenase from *Clostridium butyricum* and its mutations that enhance reaction kinetics. J. Microbiol. Biotechnol. 2014;24:1636–43.
40. Madan VK, Hillmer P, Gottschalk G. Purification and properties of NADP-dependent L(+)-3-hydroxybutyryl-CoA dehydrogenase from *Clostridium kluyveri*. Eur. J. Biochem. 1973;32:51–6.
41. Kim E-J, Kim Y-J, Kim K-J. Structural insights into substrate specificity of crotonase from the n-butanol producing bacterium *Clostridium acetobutylicum*. Biochem. Biophys. Res. Commun. 2014;451:431–5.
42. Yoo M, Bestel-Corre G, Croux C, Riviere A, Meynial-Salles I, Soucaille P. A Quantitative System-Scale Characterization of the Metabolism of *Clostridium acetobutylicum*. MBio. American Society for Microbiology (ASM); 2015;6:e01808–15.
43. Li F, Hinderberger J, Seedorf H, Zhang J, Buckel W, Thauer RK. Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri*. J. Bacteriol. American Society for Microbiology; 2008;190:843–50.
44. Boynton ZL, Bennet GN, Rudolph FB. Cloning, sequencing, and expression of clustered genes encoding beta-hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and butyryl-CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824. J. Bacteriol. 1996;178:3015–24.
45. Zdravko D, MANN M. Improvement of clostridial butanol production by gene overexpression. Google Patents; 2014.
46. Garza E, Zhao J, Wang Y, Wang J, Iverson A, Manow R, et al. Engineering a homobutanol fermentation pathway in *Escherichia coli* EG03. J. Ind. Microbiol. Biotechnol. 2012;39:1101–7.
47. Winzer K, Lorenz K, Dürre P. Acetate kinase from *Clostridium acetobutylicum*: a highly specific enzyme that is actively transcribed during acidogenesis and solventogenesis. Microbiology. 1997;143:3279–86.
48. Boynton ZL, Bennett GN, Rudolph FB. Cloning, sequencing, and expression of

genes encoding phosphotransacetylase and acetate kinase from *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. 1996;62:2758–66.

49. Walter KA, Nair R V, Cary JW, Bennett GN, Papoutsakis ET. Sequence and arrangement of two genes of the butyrate-synthesis pathway of *Clostridium acetobutylicum* ATCC 824. Gene. 1993;134:107–11.

50. Cary JW, Petersen DJ, Papoutsakis ET, Bennett GN. Cloning and expression of *Clostridium acetobutylicum* phosphotransbutyrylase and butyrate kinase genes in *Escherichia coli*. J. Bacteriol. 1988;170:4613–8.

51. Wiesenborn DP, Rudolph FB, Papoutsakis ET. Phosphotransbutyrylase from *Clostridium acetobutylicum* ATCC 824 and its role in acidogenesis. Appl. Environ. Microbiol. American Society for Microbiology (ASM); 1989;55:317–22.

52. Huang KX, Huang S, Rudolph FB, Bennett GN. Identification and characterization of a second butyrate kinase from *Clostridium acetobutylicum* ATCC 824. J. Mol. Microbiol. Biotechnol. 2000;2:33–8.

53. Hartmanis MG. Butyrate kinase from *Clostridium acetobutylicum*. J. Biol. Chem. 1987;262:617–21.

54. Sullivan L, Cates MS, Bennett GN. Structural correlations of activity of *Clostridium acetobutylicum* ATCC 824 butyrate kinase isozymes. Enzyme Microb. Technol. 2010;46:118–24.

55. Green EM, Boynton ZL, Harris LM, Rudolph FB, Papoutsakis ET, Bennett GN. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology. Microbiology Society; 1996;142 (Pt 8:2079–86.

56. Bochner BR, Huang HC, Schieven GL, Ames BN. Positive selection for loss of tetracycline resistance. J. Bacteriol. American Society for Microbiology (ASM); 1980;143:926–33.

57. Freier D, Gottschalk G. L(+)-lactate dehydrogenase of *Clostridium acetobutylicum* is activated by fructose-1,6-bisphosphate. FEMS Microbiol. Lett. 1987;43:229–33.

58. Contag PR, Williams MG, Rogers P. Cloning of a lactate dehydrogenase gene from *Clostridium acetobutylicum* B643 and expression in *Escherichia coli*. Appl. Environ. Microbiol. 1990;56:3760–5.

59. Liu D, Chen Y, Ding F, Guo T, Xie J, Zhuang W, et al. Simultaneous production of butanol and acetoin by metabolically engineered *Clostridium acetobutylicum*. Metab. Eng. Elsevier; 2015;27:107–14.

60. Wardwell SA. Metabolism of acetoin in *Clostridium acetobutylicum* ATCC 824. Rice University; 1999.

61. Frädriich C, March A, Fiege K, Hartmann A, Jahn D, Härtig E. The transcription factor AlsR binds and regulates the promoter of the alsSD operon responsible for acetoin formation in *Bacillus subtilis*. J. Bacteriol. 2012;194:1100–12.
62. Petersen DJ, Cary JW, Vanderleyden J, Bennett GN. Sequence and arrangement of genes encoding enzymes of the acetone-production pathway of *Clostridium acetobutylicum* ATCC 824. Gene. 1993;123:93–7.
63. Fischer RJ, Helms J, Dürre P. Cloning, sequencing, and molecular analysis of the sol operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. J. Bacteriol. 1993;175:6959–69.
64. Sauer U. Differential induction of genes related to solvent formation during the shift from acidogenesis to solventogenesis in continuous culture of *Clostridium acetobutylicum*. FEMS Microbiol. Lett. 1995;125:115–20.
65. Wiesenborn DP, Rudolph FB, Papoutsakis ET. Coenzyme A transferase from *Clostridium acetobutylicum* ATCC 824 and its role in the uptake of acids. Appl. Environ. Microbiol. American Society for Microbiology (ASM); 1989;55:323–9.
66. Lehmann D, Hönicke D, Ehrenreich A, Schmidt M, Weuster-Botz D, Bahl H, et al. Modifying the product pattern of *Clostridium acetobutylicum*: Physiological effects of disrupting the acetate and acetone formation pathways. Appl. Microbiol. Biotechnol. 2012;94:743–54.
67. Ravagnani A, Jennert KCB, Steiner E, Grunberg R, Jefferies JR, Wilkinson SR, et al. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Mol. Microbiol. Blackwell Science Ltd; 2000;37:1172–85.
68. Cary JW, Petersen DJ, Bennett GN, Papoutsakis ET. Methods for cloning key primary metabolic enzymes and ancillary proteins associated with the acetone-butanol fermentation of *Clostridium acetobutylicum*. Ann. N. Y. Acad. Sci. 1990;589:67–81.
69. Cornillot E, Nair R V, Papoutsakis ET, Soucaille P. The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. J. Bacteriol. 1997;179:5442–7.
70. Fontaine L, Meynial-Salles I, Girbal L, Yang X, Croux C, Soucaille P. 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. American Society for Microbiology; 2002;184:821–30.
71. Petersen DJ, Welch RW, Rudolph FB, Bennett GN. Molecular cloning of an alcohol (butanol) dehydrogenase gene cluster from *Clostridium acetobutylicum* ATCC 824. J. Bacteriol. American Society for Microbiology (ASM); 1991;173:1831–4.

72. Walter KA, Bennett GN, Papoutsakis ET. Molecular characterization of two *Clostridium acetobutylicum* ATCC 824 butanol dehydrogenase isozyme genes. J. Bacteriol. 1992;174:7149–58.
73. Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET. Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. Biotechnology. (N. Y). 1992;10:190–5.
74. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The Clostron: a universal gene knock-out system for the genus *Clostridium*. J. Microbiol. Methods. 2007;70:452–64.
75. Lee SY, Mermelstein LD, Bennett GN, Papoutsakis ET. Vector construction, transformation, and gene amplification in *Clostridium acetobutylicum* ATCC 824. Ann. N. Y. Acad. Sci. 1992;665:39–51.
76. Tummala SB, Welker NE, Papoutsakis ET. Development and characterization of a gene expression reporter system for *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. American Society for Microbiology (ASM); 1999;65:3793–9.
77. Heap JT, Pennington OJ, Cartman ST, Minton NP. A modular system for *Clostridium* shuttle plasmids. J. Microbiol. Methods. 2009;78:79–85.
78. Lee S-H, Kwon M-A, Choi S, Kim S, Kim J, Shin Y-A, et al. A New Shuttle Plasmid That Stably Replicates in *Clostridium acetobutylicum*. J. Microbiol. Biotechnol. 2015;25:1702–8.
79. Girbal L, Mortier-Barriere I, Raynaud F, Rouanet C, Croux C, Soucaille P. Development of a Sensitive Gene Expression Reporter System and an Inducible Promoter-Repressor System for *Clostridium acetobutylicum*. Appl. Environ. Microbiol. American Society for Microbiology (ASM); 2003;69:4985–8.
80. Zhang J, Liu Y-J, Cui G-Z, Cui Q. A novel arabinose-inducible genetic operation system developed for *Clostridium cellulolyticum*. Biotechnol. Biofuels. BioMed Central; 2015;8:36.
81. Feustel L, Nakotte S, Dürre P. Characterization and development of two reporter gene systems for *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 2004;70:798–803.
82. Desai RP, Papoutsakis ET. Antisense RNA Strategies for Metabolic Engineering of *Clostridium acetobutylicum*. Appl. Environ. Microbiol. American Society for Microbiology; 1999;65:936–45.
83. Tummala SB, Junne SG, Paredes CJ, Papoutsakis ET. Transcriptional analysis of product-concentration driven changes in cellular programs of recombinant *Clostridium acetobutylicum* strains. Biotechnol. Bioeng. 2003;84:842–54.

84. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, et al. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *J. Microbiol. Methods*. 2010;80:49–55.
85. Kuehne SA, Minton NP. ClosTron-mediated engineering of *Clostridium*. *Bioengineered*. 2012;3:247–54.
86. Lütke-Eversloh T. Application of new metabolic engineering tools for *Clostridium acetobutylicum*. *Appl. Microbiol. Biotechnol*. 2014;98:5823–37.
87. Ehsaan M, Kuit W, Zhang Y, Cartman ST, Heap JT, Winzer K, et al. Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. *Biotechnol. Biofuels*. 2016;9:4.
88. Shao L, Hu S, Yang YY, Gu Y, Chen J, Yang YY, et al. Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. *Cell Res*. 2007;17:963–5.
89. Jang Y-S, Lee JYJJ, Lee JYJJ, Park JH, Im JA, Eom M, et al. Enhanced Butanol Production Obtained by Reinforcing the Direct Butanol-Forming Route in *Clostridium acetobutylicum*. *MBio*. 2012;3:1–9.
90. Belfort M, Wood DW, Stoddard BL, Derbyshire V, editors. *Homing Endonucleases and Inteins*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2005.
91. Cooksley CM, Zhang Y, Wang H, Redl S, Winzer K, Minton NP. Targeted mutagenesis of the *Clostridium acetobutylicum* acetone-butanol-ethanol fermentation pathway. *Metab. Eng*. 2012;14:630–41.
92. Harris LM, Welker NE, Papoutsakis ET. Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *J. Bacteriol*. 2002;184:3586–97.
93. Heap JT, Ehsaan M, Cooksley CM, Ng Y-KK, Cartman ST, Winzer K, et al. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic Acids Res*. 2012;40:1–10.
94. Al-Hinai MA, Fast AG, Papoutsakis ET. Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. *Appl. Environ. Microbiol*. 2012;78:8112–21.
95. Zhang N, Shao L, Jiang Y, Gu Y, Li Q, Liu J, et al. I-SceI-mediated scarless gene modification via allelic exchange in *Clostridium*. *J. Microbiol. Methods*. 2015;108:49–60.
96. Green EM, Bennett GN. Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. *Appl. Biochem. Biotechnol*. Humana Press; 1996;57-58:213–21.

97. Lee S-HJ, Kim HJ, Shin Y-A, Kim KH, Lee S-HJ. Single crossover-mediated markerless genome engineering in *Clostridium acetobutylicum*. J. Microbiol. Biotechnol. 2016;
98. Dong H, Tao W, Gong F, Li Y, Zhang Y. A functional recT gene for recombining of *Clostridium*. J. Biotechnol. Elsevier B.V.; 2014;173:65–7.
99. Zhang Y, Xu S, Chai C, Yang S, Jiang W, Minton NP, et al. Development of an inducible transposon system for efficient random mutagenesis in *Clostridium acetobutylicum*. FEMS Microbiol. Lett. 2016;
100. Sillers R, Chow A, Tracy B, Papoutsakis ET. Metabolic engineering of the non-sporulating, non-solventogenic *Clostridium acetobutylicum* strain M5 to produce butanol without acetone demonstrate the robustness of the acid-formation pathways and the importance of the electron balance. Metab. Eng. 2008;10:321–32.
101. Xu M, Zhao J, Yu L, Tang I-C, Xue C, Yang S-T. Engineering *Clostridium acetobutylicum* with a histidine kinase knockout for enhanced n-butanol tolerance and production. Appl. Microbiol. Biotechnol. 2015;99:1011–22.
102. Ventura J-RS, Hu H, Jahng D. Enhanced butanol production in *Clostridium acetobutylicum* ATCC 824 by double overexpression of 6-phosphofructokinase and pyruvate kinase genes. Appl. Microbiol. Biotechnol. 2013
103. Yang Y, Lang N, Yang G, Yang S, Jiang W, Gu Y. Improving the performance of solventogenic clostridia by reinforcing the biotin synthetic pathway. Metab. Eng. 2016;35:121–8.
104. Hou X, Peng W, Xiong L, Huang C, Chen X, Chen X, et al. Engineering *Clostridium acetobutylicum* for alcohol production. J. Biotechnol. 2013;166:25–33.
105. Kuit W, Minton NP, López-Contreras AM, Eggink G. Disruption of the acetate kinase (ack) gene of *Clostridium acetobutylicum* results in delayed acetate production. Appl. Microbiol. Biotechnol. 2012;94:729–41.
106. Lehmann D, Radomski N, Lütke-Eversloh T. New insights into the butyric acid metabolism of *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. 2012;96:1325–39.
107. Jiang Y, Xu C, Dong F, Yang Y, Jiang W, Yang S. Disruption of the acetoacetate decarboxylase gene in solvent-producing *Clostridium acetobutylicum* increases the butanol ratio. Metab. Eng. 2009;11:284–91.
108. Jang Y-S, Im JA, Choi SY, Lee JI, Lee SY. Metabolic engineering of *Clostridium acetobutylicum* for butyric acid production with high butyric acid selectivity. Metab. Eng. 2014;23:165–74.
109. Harris LM, Blank L, Desai RP, Welker NE, Papoutsakis ET. Fermentation

characterization and flux analysis of recombinant strains of *Clostridium acetobutylicum* with an inactivated solR gene. J. Ind. Microbiol. Biotechnol. 2001;27:322–8.

110. Wietzke M, Bahl H. The redox-sensing protein Rex, a transcriptional regulator of solventogenesis in *Clostridium acetobutylicum*. Appl. Microbiol. Biotechnol. 2012;96:749–61.

111. Honicke D, Lutke-Eversloh T, Liu Z, Lehmann DD, Liebl W, Ehrenreich A, et al. Chemostat cultivation and transcriptional analyses of *Clostridium acetobutylicum* mutants with defects in the acid and acetone biosynthetic pathways. Appl. Microbiol. Biotechnol. 2014;98:9777–94.

112. Lee J, Yun H, Feist AM, Palsson BØ, Lee SY. Genome-scale reconstruction and in silico analysis of the *Clostridium acetobutylicum* ATCC 824 metabolic network. Appl. Microbiol. Biotechnol. 2008;80:849–62.

113. Senger RS, Papoutsakis ET. Genome-scale model for *Clostridium acetobutylicum* : Part I. Metabolic network resolution and analysis. Biotechnol. Bioeng. 2008;101:1036–52.

114. McAnulty MJ, Yen JY, Freedman BG, Senger RS. Genome-scale modeling using flux ratio constraints to enable metabolic engineering of clostridial metabolism in silico. BMC Syst. Biol. 2012;6:42.

115. Dash S, Mueller TJ, Venkataramanan KP, Papoutsakis ET, Maranas CD. Capturing the response of *Clostridium acetobutylicum* to chemical stressors using a regulated genome-scale metabolic model. Biotechnol. Biofuels. BioMed Central; 2014;7:144.

116. Tracy BP, Jones SW, Papoutsakis ET. Inactivation of σ E and σ G in *Clostridium acetobutylicum* illuminates their roles in clostridial-cell-form biogenesis, granulose synthesis, solventogenesis, and spore morphogenesis. J. Bacteriol. 2011;193:1414–26.

117. Jones MAAH, Papoutsakis ET. σ K of *Clostridium acetobutylicum* is the first known sporulation: Specific sigma factor with two developmentally separated roles, one early and one late in sporulation. J. Bacteriol. 2014;196:287–99.

118. Jones SW, Tracy BP, Gaida SM, Papoutsakis ET. Inactivation of σ F in *Clostridium acetobutylicum* ATCC 824 blocks sporulation prior to asymmetric division and abolishes σ E and σ G protein expression but does not block solvent formation. J. Bacteriol. American Society for Microbiology (ASM); 2011;193:2429–40.

119. Al-Hinai MA, Jones SW, Papoutsakis ET. The *Clostridium* sporulation programs: diversity and preservation of endospore differentiation. Microbiol. Mol. Biol. Rev. 2015;79:19–37.

120. Nair R V., Green EM, Watson DE, Bennett GN, Papoutsakis ET. Regulation of the sol locus genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824

by a putative transcriptional repressor. J. Bacteriol. American Society for Microbiology (ASM); 1999;181:319–30.

121. Thormann K, Feustel L, Lorenz K, Nakotte S, Dürre P. Control of butanol formation in *Clostridium acetobutylicum* by transcriptional activation. J. Bacteriol. 2002;184:1966–73.

122. Steiner E, Scott J, Minton NP, Winzer K. An agr quorum sensing system that regulates granulose formation and sporulation in *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 2012;78:1113–22.

123. Bahl H, Müller H, Behrens S, Joseph H, Narberhaus F. Expression of heat shock genes in *Clostridium acetobutylicum*. FEMS Microbiol. Rev. 1995;17:341–8.

124. Tomas CA, Welker NE, Papoutsakis ET. 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. American Society for Microbiology (ASM); 2003;69:4951–65.

125. Hillmann F, Fischer R-J, Bahl H. The rubrerythrin-like protein Hsp21 of *Clostridium acetobutylicum* is a general stress protein. Arch. Microbiol. 2006;185:270–6.

126. Fiedler T, Mix M, Meyer U, Mikkat S, Glocker MO, Bahl H, et al. The two-component system PhoPR of *Clostridium acetobutylicum* is involved in phosphate-dependent gene regulation. J. Bacteriol. 2008;190:6559–67.

127. Scotcher MC, Rudolph FB, Bennett GN. Expression of abrB310 and SinR, and effects of decreased abrB310 expression on the transition from acidogenesis to solventogenesis, in *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. 2005;71:1987–95.

128. Hillmann F, Fischer R-JJ, Saint-Prix F, Girbal L, Bahl H. PerR acts as a switch for oxygen tolerance in the strict anaerobe *Clostridium acetobutylicum*. Mol. Microbiol. 2008;68:848–60.

129. Hillmann F, Döring C, Riebe O, Ehrenreich A, Fischer R-J, Bahl H. The role of PerR in O₂-affected gene expression of *Clostridium acetobutylicum*. J. Bacteriol. 2009;191:6082–93.

130. Vasileva D, Janssen H, Hönicke D, Ehrenreich A, Bahl H. Effect of iron limitation and fur gene inactivation on the transcriptional profile of the strict anaerobe *Clostridium acetobutylicum*. Microbiol. (United Kingdom). 2012;158:1918–29.

131. Tan Y, Liu ZZ-Y, Liu ZZ-Y, Zheng H-J, Li F-L. Comparative transcriptome analysis between csrA-disruption *Clostridium acetobutylicum* and its parent strain. Mol. Biosyst. 2015;11:1434–42.

132. Wu Y, Yang Y, Ren C, Yang C, Yang S, Gu Y, et al. Molecular modulation of

pleiotropic regulator CcpA for glucose and xylose coutilization by solvent-producing *Clostridium acetobutylicum*. Metab. Eng. 2015;28:169–79.

133. Davison SP, Santangelo JD, Reid SJ, Woods DR. A *Clostridium acetobutylicum* regulator gene (regA) affecting amylase production in *Bacillus subtilis*. Microbiology. 1995;141 (Pt 4:989–96.

134. Zhang L, Nie X, Ravcheev DA, Rodionov DA, Sheng J, Gu Y, et al. Redox-responsive repressor Rex modulates alcohol production and oxidative stress tolerance in *Clostridium acetobutylicum*. J. Bacteriol. 2014;196:3949–63.

135. Mota LJ, Sarmiento LM, de Sá-Nogueira I. Control of the arabinose regulon in *Bacillus subtilis* by AraR in vivo: crucial roles of operators, cooperativity, and DNA looping. J. Bacteriol. American Society for Microbiology (ASM); 2001;183:4190–201.

136. Zhang L, Leyn SA, Gu Y, Jiang W, Rodionov DA, Yang C. Ribulokinase and transcriptional regulation of arabinose metabolism in *Clostridium acetobutylicum*. J. Bacteriol. American Society for Microbiology; 2012;194:1055–64.

137. Ternan N. Small regulatory RNA molecules in bacteria.OA Microbiology. OA Microbiol. 2013;1:1.

138. Chen Y, Indurthi DC, Jones SW, Papoutsakis ET. Small RNAs in the genus *Clostridium*. MBio. 2011;2:e00340–10.

139. Venkataramanan KP, Jones SW, McCormick KP, Kunjeti SG, Ralston MT, Meyers BC, et al. The *Clostridium* small RNome that responds to stress: the paradigm and importance of toxic metabolite stress in *C. acetobutylicum*. BMC Genomics. 2013;14:849.

140. Liu CC, Qi L, Yanofsky C, Arkin AP. Regulation of transcription by unnatural amino acids. Nat. Biotechnol. Nature Publishing Group; 2011;29:164–8.

141. Soucaille P, Figge R, Croux C. Process for Chromosomal Integration and DNA Sequence Replacement in Clostridia. US: International Patent Application; 2008.

142. Mermelstein LD, Papoutsakis ET. In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. American Society for Microbiology (ASM); 1993;59:1077–81.

143. Fabret C, Ehrlich SD, Noirot P. A new mutation delivery system for genome-scale approaches in *Bacillus subtilis*. Mol. Microbiol. 2002;46:25–36.

144. Samuelson J, Xu S, O’Loane D. Method for cloning and expression of AcuI restriction endonuclease and AcuI methylase in *E. coli*. US patent No.7,011,966; 2006.

145. Janulaitis A, Vaisvila R, Timinskas A, Klimasauskas S, Butkus V. Cloning and

sequence analysis of the genes coding for Eco57I type IV restriction-modification enzymes. *Nucleic Acids Res.* 1992;20:6051–6.

146. Noirot P, Petit MA, Ehrlich SD. Plasmid replication stimulates DNA recombination in *Bacillus subtilis*. *J. Mol. Biol.* 1987;196:39–48.

147. Petit MA, Mesas JM, Noirot P, Morel-Deville F, Ehrlich SD. Induction of DNA amplification in the *Bacillus subtilis* chromosome. *EMBO J.* 1992;11:1317–26.

148. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 2000;97:6640–5.

149. Stephan J, Stemmer V, Niederweis M. Consecutive gene deletions in *Mycobacterium smegmatis* using the yeast FLP recombinase. *Gene.* 2004;343:181–90.

150. Millat T, Voigt C, Janssen H, Cooksley CM, Winzer K, Minton NP, et al. Coenzyme A-transferase-independent butyrate re-assimilation in *Clostridium acetobutylicum*—evidence from a mathematical model. *Appl. Microbiol. Biotechnol.* Springer Berlin Heidelberg; 2014;98:9059–72.

151. Soucaille P. Metabolic engineering of *Clostridium acetobutylicum* for enhanced production of n-butanol. International patent application WO2008052973; 2013.

152. Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET. Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. *Biotechnology.* (N. Y). Nature Publishing Group; 1992;10:190–5.

153. Cherepanov PP, Wackernagel W. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of FLP-catalyzed excision of the antibiotic-resistance determinant. *Gene.* 1995;158:9–14.

154. Raynaud C, Sarçabal P, Meynial-Salles I, Croux C, Soucaille P. Molecular characterization of the 1,3-propanediol (1,3-PD) operon of *Clostridium butyricum*. *Proc. Natl. Acad. Sci. U. S. A.* 2003;100:5010–5.

155. Peguin S, Goma G, Delorme P, Soucaille P. Metabolic flexibility of *Clostridium acetobutylicum* in response to methyl viologen addition. *Appl. Microbiol. Biotechnol.* Springer-Verlag; 1994;42:611–6.

156. Sambrook J, Fritsch EF, Maniatis T. *Molecular cloning: A laboratory manual*. Second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; 1989.

157. Payne MS, K. SP, Hsu AK-H, Nair R V., Valle F, Soucaille P, et al. Promoter and plasmid system for genetic engineering. US20090156430 A1; 2009.

158. Meynial-salles I, Cervin M a, Soucaille P. New Tool for Metabolic Pathway Engineering. *Appl. Environ. Microbiol.* 2005;71:2140–4.

159. Weizmann C. U.S. Patent 1,315,585. 1919.
160. Weizmann C. UK Patent 4845. 1915.
161. Ndaba B, Chiyanzu I, Marx S. N-Butanol derived from biochemical and chemical routes: A review. *Biotechnol. Reports*. Elsevier B.V.; 2015;8:1–9.
162. Cornillot E, Soucaille P. Solvent-forming genes in clostridia. *Nature*. 1996;380:489–489.
163. Anbarasan P, Baer ZC, Sreekumar S, Gross E, Binder JB, Blanch HW, et al. Integration of chemical catalysis with extractive fermentation to produce fuels. *Nature*. Nature Publishing Group; 2012;491:235–9.
164. Bormann S, Baer ZC, Sreekumar S, Kuchenreuther JM, Toste FD, Blanch HWH, et al. Engineering *Clostridium acetobutylicum* for production of kerosene and diesel blendstock precursors. *Metab. Eng.* 2014;25:124–30.
165. Green EM. Fermentative production of butanol—the industrial perspective. *Curr. Opin. Biotechnol.* 2011;22:337–43.
166. Papoutsakis ET. Reassessing the Progress in the Production of Advanced Biofuels in the Current Competitive Environment and Beyond: What Are the Successes and Where Progress Eludes Us and Why. *Ind. Eng. Chem. Res.* 2015;54:10170–82.
167. Croux C, Nguyen N-P-T, Lee J, Raynaud C, Saint-Prix F, Gonzalez-Pajuelo M, et al. Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. *Biotechnol. Biofuels*. 2016;9:23.
168. Duncombe GR, Frerman FE. Molecular and catalytic properties of the acetoacetyl-coenzyme A thiolase of *E. coli*. *Arch. Biochem. Biophys.* 1976;176:159–70.
169. Atsumi S, Cann AF, Connor MR, Shen CR, Smith KM, Brynildsen MP, et al. Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab. Eng.* 2008;10:305–11.
170. Angov E, Legler P, Mease R. Adjustment of codon usage frequencies by codon harmonization improves protein expression and folding. *Methods Mol. Biol.* 2011;705:1–13.
171. Soucaille P, Nguyen N-P-T, Percheron B, Croux C, Meynial-Salles I. *Clostridium acetobutylicum* unable to produce hydrogen and useful for the continuous production of chemicals and fuels, Patent WO/2016/042160. 2016.
172. Amador-Noguez D, Brasg IA, Feng XJ, Roquet N, Rabinowitz JD. Metabolome remodeling during the acidogenic-solventogenic transition in *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* 2011;77:7984–97.

173. Dugar D, Stephanopoulos G. Relative potential of biosynthetic pathways for biofuels and bio-based products. *Nat. Biotechnol.* 2011;29:1074–8.
174. Qureshi N, David H, Vertes A. *Biorefineries: Integrated biochemical processes for liquid biofuels*. 1st ed. Elsevier; 2014.
175. Zigha A. Process for butanol production, US20150004664 A1. 2012.
176. Soni BK, Soucaille P, Goma G. Continuous acetone-butanol fermentation: a global approach for the improvement in the solvent productivity in synthetic medium. *Appl. Microbiol. Biotechnol.* 1987;25:317–21.
177. Minier M, Grateloup R, Goma G. Extractive Acetonobutylic Fermentation by Coupling Ultrafiltration and Distillation. *Biotechnol. Bioeng.* 1990;35:861–9.
178. Dusséaux S, Croux C, Soucaille P, Meynial-Salles I. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for the high-yield production of a biofuel composed of an isopropanol/butanol/ethanol mixture. *Metab. Eng.* 2013;18:1–8.
179. Hartmanis MG, Gatenbeck S. Intermediary Metabolism in *Clostridium acetobutylicum*: Levels of Enzymes Involved in the Formation of Acetate and Butyrate. *Appl. Environ. Microbiol.* 1984;47:1277–83.
180. Inui M, Suda M, Kimura S, Yasuda K, Suzuki H, Toda H, et al. Expression of *Clostridium acetobutylicum* butanol synthetic genes in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 2008;77:1305–16.
181. Girbal L, Soucaille P. Regulation of *Clostridium acetobutylicum* metabolism as revealed by mixed-substrate steady-state continuous cultures: role of NADH/NAD ratio and ATP pool. *J. Bacteriol.* 1994;176:6433–8.
182. Soucaille P. Process for the stable gene interruption in clostridia, U.S. 61/220,606. 2010.
183. Sabathé F, Croux C, Cornillot E, Soucaille P. amyP, a reporter gene to study strain degeneration in *Clostridium acetobutylicum* ATCC 824. *FEMS Microbiol. Lett.* 2002;210:93–8.
184. Zhong J, Karberg M, Lambowitz AM. Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Res.* 2003;31:1656–64.
185. Lutke-Eversloh T, Bahl H. Metabolic engineering of *Clostridium acetobutylicum*: Recent advances to improve butanol production. *Curr. Opin. Biotechnol.* 2011;22:634–47.
186. Nair R V, Bennett GN, Papoutsakis ET. Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. *J.*

Bacteriol. 1994;176:871–85.

187. Atsumi S, Liao JC. Metabolic engineering for advanced biofuels production from *Escherichia coli*. Curr. Opin. Biotechnol. 2008;19:414–9.

188. Bahl H, Andersch W, Gottschalk G. Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a two-stage phosphate limited chemostat. Eur. J. Appl. Microbiol. Biotechnol. Springer-Verlag; 1982;15:201–5.

189. Grimmmler C, Janssen H, Krauß D, Fischer RJ, Bahl H, Dürre P, et al. Genome-wide gene expression analysis of the switch between acidogenesis and solventogenesis in continuous cultures of *Clostridium acetobutylicum*. J. Mol. Microbiol. Biotechnol. 2011;20:1–15.

190. Peguin S, Soucaille P, Peguin S, Soucaille P. Modulation of Carbon and Electron Flow in *Clostridium acetobutylicum* by Iron Limitation and Methyl Viologen Addition . These include : Modulation of Carbon and Electron Flow in *Clostridium acetobutylicum* by Iron Limitation and Methyl Viologen Addition. 1995;61:403–5.

191. Millat T, Janssen H, Thorn GJ, King JR, Bahl H, Fischer RJ, et al. A shift in the dominant phenotype governs the pH-induced metabolic switch of *Clostridium acetobutylicum* in phosphate-limited continuous cultures. Appl. Microbiol. Biotechnol. 2013;97:6451–66.

192. Hönicke D, Janssen H, Grimmmler C, Ehrenreich A, Lütke-Eversloh T. Global transcriptional changes of *Clostridium acetobutylicum* cultures with increased butanol: Acetone ratios. N. Biotechnol. 2012;29:485–93.

193. Wang Q, Venkataramanan KP, Huang H, Papoutsakis ET, Wu CH. Transcription factors and genetic circuits orchestrating the complex, multilayered response of *Clostridium acetobutylicum* to butanol and butyrate stress. BMC Syst. Biol. BMC Systems Biology; 2013;7:120.

194. Janssen H, Grimmmler C, Ehrenreich A, Bahl H, Fischer RJ. A transcriptional study of acidogenic chemostat cells of *Clostridium acetobutylicum*-Solvent stress caused by a transient n-butanol pulse. J. Biotechnol. Elsevier B.V.; 2012;161:354–65.

195. Alsaker K V, Spitzer TR, Eleftherios T, Papoutsakis ET. Transcriptional Analysis of spo0A Overexpression in *Clostridium acetobutylicum* and Its Effect on the Cell ' s Response to Butanol Stress Transcriptional Analysis of spo0A Overexpression in *Clostridium acetobutylicum* and Its Effect on the Cell ' s Response. J. Bacteriol. 2004;186:1959–71.

196. Schwarz KM, Kuit W, Grimmmler C, Ehrenreich A, Kengen SWM. A transcriptional study of acidogenic chemostat cells of *Clostridium acetobutylicum*--cellular behavior in adaptation to n-butanol. J. Biotechnol. 2012;161:366–77.

197. Alsaker K V, Paredes C, Papoutsakis ET. 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. 2010;105:1131–47.

198. Davison SP, Santangelo JD, Reid SJ, Woods DR. A *Clostridium acetobutylicum* regulator gene (regA) affecting amylase production in *Bacillus subtilis*. DNA Seq. 1995;141 (Pt 4:989–96.

199. Janssen H, Döring C, Ehrenreich A, Voigt B, Hecker M, Bahl H, et al. A proteomic and transcriptional view of acidogenic and solventogenic steady-state cells of *Clostridium acetobutylicum* in a chemostat culture. Appl. Microbiol. Biotechnol. 2010;87:2209–26.

200. Hofmann K. and Stoffel W. TMbase-a database of membrane spanning proteins segments. Biol. Chem. Hoppe Seyler. 1993;347.

201. Viklund H, Elofsson A. OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. Bioinformatics. 2008;24:1662–8.

202. Dodd IB, Egan JB. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Res. 1990;18:5019–26.

203. Solovyev V, Salamov A. Metagenomics and its Applications in Agriculture; Biomedicine and Environmental Studies. Autom. Annot. Microb. Genomes Metagenomic Seq. Nova Science; 2011.

204. Thormann K, Feustel L, Lorenz K, Dürre P, Nakotte S. Control of Butanol Formation in *Clostridium acetobutylicum* by Transcriptional Activation. J. Bacteriol. 2002;184:1966–73.

205. Ravagnani A, Jennert KCB, Steiner E, Grunberg R, Jefferies JR, Wilkinson SR, et al. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Mol. Microbiol. Blackwell Science Ltd; 2000;37:1172–85.

206. Gerischer U, Dürre P. mRNA analysis of the adc gene region of *Clostridium acetobutylicum* during the shift to solventogenesis. J. Bacteriol. American Society for Microbiology (ASM); 1992;174:426–33.

207. Sheldon JR, Heinrichs DE. Recent developments in understanding the iron acquisition strategies of gram positive pathogens. FEMS Microbiol. Rev. The Oxford University Press; 2015;39:592–630.

208. Jiménez-López C, Collette JR, Brothers KM, Shepardson KM, Cramer RA, Wheeler RT, et al. *Candida albicans* induces arginine biosynthetic genes in response to host-derived reactive oxygen species. Eukaryot. Cell. 2013;12:91–100.

209. Brenner M, Ames B. The histidine operon and its regulation-metabolic pathways.

Academic Press; 1971.

210. Winkler M. Biosynthesis of histidine-*Escheria coli* and *Samonella*: cellular and molecular biology. ASM Press; 1996.

211. Kulis-Horn RK, Persicke M, Kalinowski J. Histidine biosynthesis, its regulation and biotechnological application in *Corynebacterium glutamicum*. Microb. Biotechnol. 2014;7:5–25.

212. Novichkov PS, Laikova ON, Novichkova ES, Gelfand MS, Arkin AP, Dubchak I, et al. RegPrecise: a database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes. Nucleic Acids Res. 2010;38:D111–8.

213. Broadbent JR, Larsen RL, Deibel V, Steele JL. Physiological and transcriptional response of *Lactobacillus casei* ATCC 334 to acid stress. J. Bacteriol. American Society for Microbiology (ASM); 2010;192:2445–58.

214. Tanous C, Soutourina O, Raynal B, Hullo M-F, Mervelet P, Gilles A-M, et al. The CymR regulator in complex with the enzyme CysK controls cysteine metabolism in *Bacillus subtilis*. J. Biol. Chem. 2008;283:35551–60.

215. Paredes CJ, Alsaker K V, Papoutsakis ET. A comparative genomic view of clostridial sporulation and physiology. Nat. Rev. Microbiol. 2005;3:969–78.

216. Zhao Y, Tomas CA, Rudolph FB, Papoutsakis ET, Bennett GN. Intracellular butyryl phosphate and acetyl phosphate concentrations in *Clostridium acetobutylicum* and their implications for solvent formation. Appl. Environ. Microbiol. 2005;71:530–7.

217. Mironov AS, Gusarov I, Rafikov R, Lopez LE, Shatalin K, Kreneva RA, et al. Sensing Small Molecules by Nascent RNA. Cell. 2002;111:747–56.

218. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.

219. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008;36:W465–9.

220. Rodionov DA. Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. Chem. Rev. 2007;107:3467–97.

221. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics. 2000;16:944–5.

Publications

Croux C*, **Nguyen NPT***, Lee J, Raynaud C, Saint-Prix F, Gonzalez-Pajuelo M, Meynial-Salles I, Soucaille P. *Biotechnol Biofuels*. **2016**. Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum* (* **co-first authors**)

The upcoming ones

Nguyen NPT, Soucaille P. et al. **2016**. CAP0037, a Novel Global Regulator of *Clostridium acetobutylicum* Metabolism: A system scale analysis. *Under review by mBio journal*

Nguyen NPT, Soucaille P. et al. **2016**. The Weizmann process revisited for the continuous production of n-butanol. *In preparation*